visited every 18 days by USDA/ASCS field personnel, except for the Finney County, Kansas, and Hand County, South Dakota, "supersites," which are visited every 9 days. The 11 ITS's in Canada are monitored every 18 days by personnel from the Canadian Agriculture Department.

The crop calendar model used by LACIE is a modification of the biometeorological time scale (BMTS) developed by Robertson. The Robertson model predicts the rate of progression of wheat through its biological development. The principal output of the model is a daily increment of development through six physiological stages of growth (table 5-6). Daily maximum and minimum temperatures and day length are variables used to implement this model, which is often referred to as the adjustable crop calendar (ACC).

The crop calendar model was developed using phenological data from Canada for spring wheat. Terms and coefficients are the same for all locations. In 1976, Feyerherm developed a scalar multiplier which was applied to the initial ACC equations between emergence and heading and which reflected the effect of dormancy on winter wheat (ref. 1). During Phase III, the ACC winter wheat estimates were modified by the use of a multiplier at each crop calendar station to improve the accuracy of the model. In addition to the multipliers, another control was introduced to the model to prohibit crop calendar advancement beyond stage 2.85 before January 1 to prevent the model from predicting jointing prior to spring greening.

All the growth stages defined by Robertson in the BMTS model development are not easily observable by field personnel; for example, BMTS stage 3.0 (jointing) can be observed only by plant dissection. A different set of stages has therefore been developed for ground observations. The ground-observed growth stage of each ITS must be developed by relating the ITS growth-stage observations to the related BMTS stage. After emergence, the earliest stage at which there is no ambiguity in this relationship is at heading. The BMTS stage 3.0 (jointing) is known to occur after tillering and before booting, which are observed by ground personnel; jointing is estimated

TABLE 5-6.— ROBERTSON BMTS AND OBSERVED ITS WHEAT PHENOLOGICAL STAGES

Stage	Robertson BMTS	ITS growth stage code	Description
Planted	1.0	01	Planted
		02	Planted, no emergence
Emergence	2.0	03	Emergence
Jointing	3.0	04	Tillering, prebooting, prebudding
	3.5	05	Booted or budded
Heading	4.0	06	Beginning to head or flower
	4.5	07	Fully headed or flowered
Soft dough	5.0	08	Beginning to ripen
Ripening	6.0	09	Ripe to mature
Harvest	7.0	10	Harvest

by extrapolating between these observations. An error of a few days is probable in relating ground observations to BMTS stages.

The ACC is prepared biweekly in a meteorological summary for all regions being examined by LACIE. The BMTS stages of wheat are based on inputs from each reporting meteorological station, and these estimates are used to develop BMTS contours as shown in figure 5-6. The ITS BMTS estimate is then determined from its location on this contour map and compared to that determined by ground observations (figs. 5-7 and 5-8). The SD, $\pm 1\sigma$, of these ground-observed estimates on a field-to-field basis is also shown in these figures. Note in the Oldham County, Texas, example (fig. 5-7) that the ground-computed stage contains the ACC-estimated stage within one SD in the periods from midjointing (3.5) to soft dough (5.0). Before 3.5 and after 5.0, the ACC was ahead of the ground truth by a few days and more than one SD. However, in most cases, the ACC BMTS estimate was somewhat more accurate than if a normal or average growth stage was assumed.

Tables 5-7, 5-8, and 5-9 show the differences in days between the historical, the computed, and the observed development curves at each of the BMTS stages. From these data, the biases, the SD's, and the RMSE's were calculated over all sites at each development stage for both the model estimates versus ITS observed data and historical versus ITS.

Apparently increased accuracy was obtained in Phase III by using the scalar multipliers to generate the winter wheat estimates. The magnitude of the mean biases computed for the model versus ITS estimates at the development stages of jointing through ripening varied from 1.5 to 2.9 days. The corresponding SD's and RMSE's varied from 6.1 to 9.7 days. Comparisons of the historical versus the ground-observed development curves at these same stages produced mean biases whose magnitudes ranged from 1.2 to 8.4 days. The calculated SD's and RMSE's over all sites ranged from 8.3 to 12.0 days.

The difference between the ACC estimates and the gound-observed values over all ITS's at all development stages was 7 days or less 65 percent of the

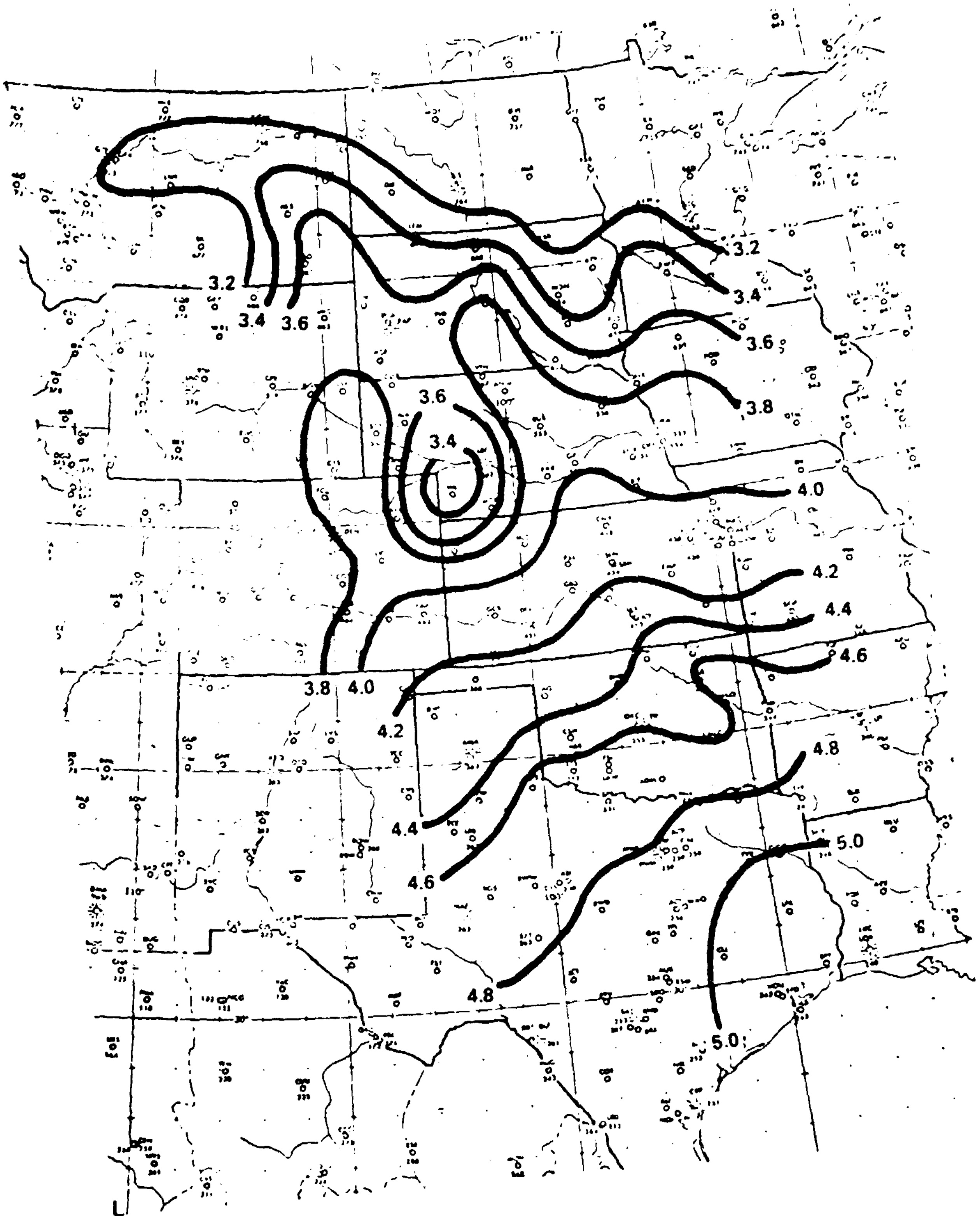


Figure 5-6.— Winter wheat BMTS isolines as predicted by the LACIE ACC meteorological data through May 1, 1977.

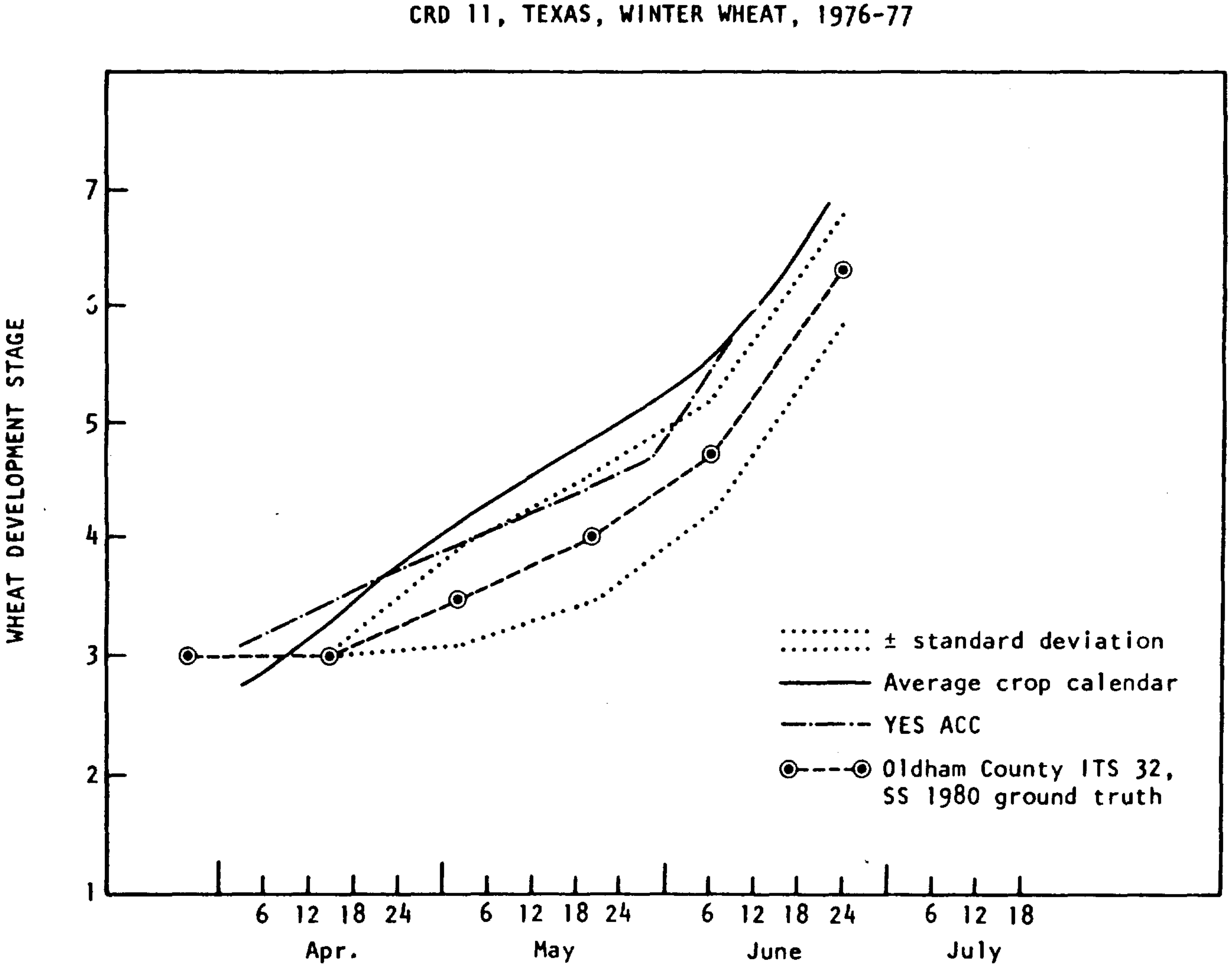
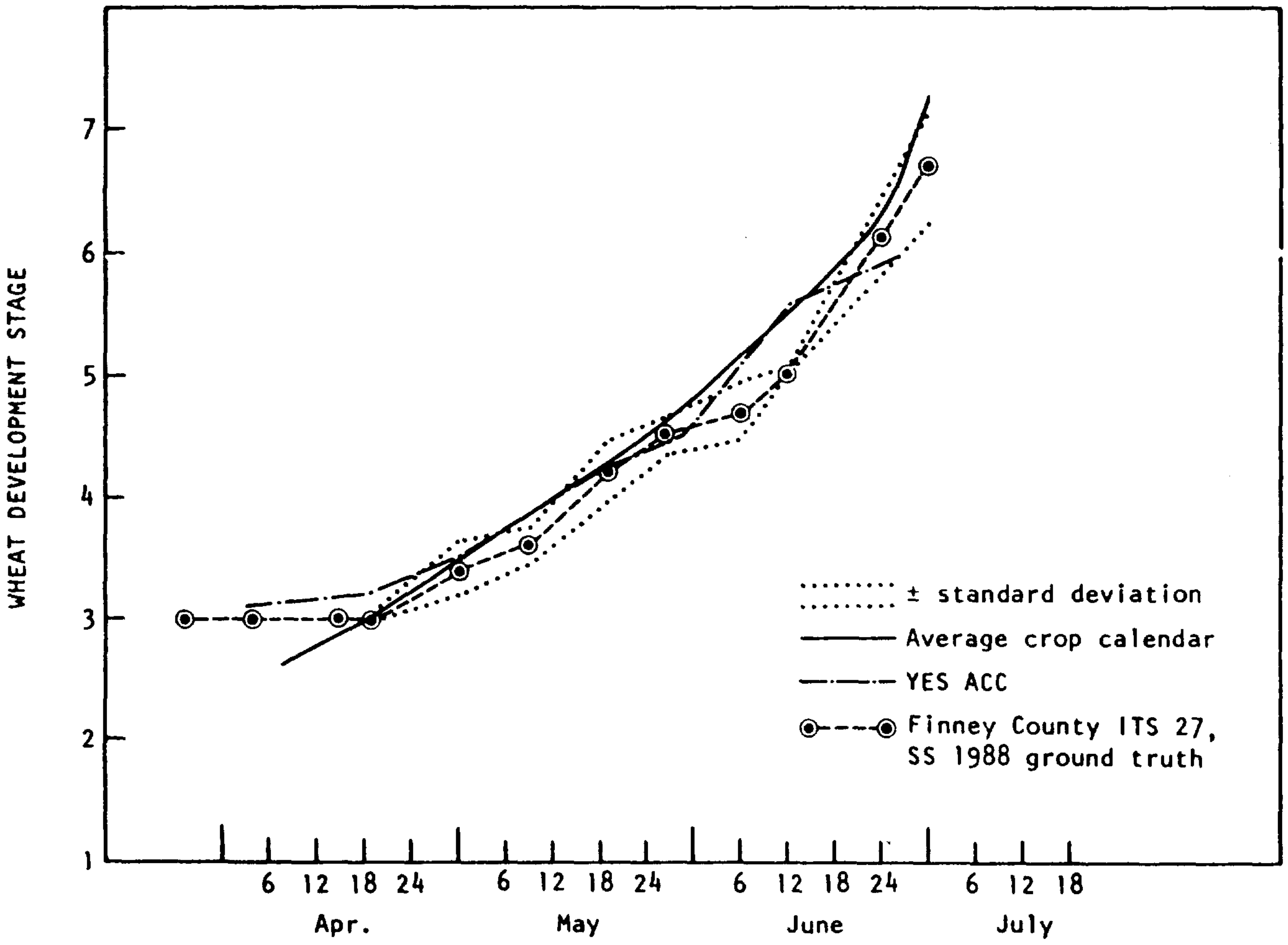


Figure 5-7.– Comparison of observed and predicted crop calendar stages for Oldham County, Texas.



CRD 30, KANSAS, WINTER WHEAT, 1976-77

Figure 5-8.– Comparison of observed and predicted crop calendar stages for Finney County, Kansas.

TABLE 5-7.– LACIE PHASE III ACC AND HISTORICAL CRD CALENDARS WITH OBSERVED DEVELOPMENT STAGES IN THE 1976-77 WINTER WHEAT ITS'S[a]

[Comparison by days]

ITS		ITS minus ACC						ITS minus historical					
County	State	Jointing		Heading		Soft dough	Ripe	Jointing		Heading		Soft dough	Ripe
		3.0	3.5	4.0	4.5	5.0	6.0	3.0	3.5	4.0	4.5	5.0	6.0
Oldham	Texas	-4	17	17	9	9	8	-13	19	21	20	16	8
Randall	Texas	3	7	5	4	8	8	-9	-2	3	11	13	5
Finney	Kansas	4	5	-3	3	8	-5	-23	4	3	2	9	3
Rice	Kansas	-12	0	-5	-14	0	7	-10	-4	-8	-10	-2	-3
Ellis	Kansas	-11	-3	-8	-15	1	-11	-27	-3	-6	-10	0	-4
Saline	Kansas	4	0	-3	-3	6	11	-10	-3	-4	-2	3	4
Morton	Kansas	2	0	1	0	5	8	-19	3	1	3	5	0
Boone	Indiana	10	9	2	0	2	5	-5	1	-4	-10	-13	5
Madison	Indiana	10	6	1	0	8	5	-5	0	-5	-10	-5	-8
Shelby	Indiana	10	-1	-3	-1	-4	2	-6	-8	-8	-9	-16	-11
Bannock	Idaho	15	3	0	-1	8	13	4	7	0	-11	-12	-19
Franklin	Idaho	NA	NA	NA	11	14	NA	NA	NA	NA	8	-1	NA
Oneida	Idaho	-11	-7	-7	-7	-5	-3	3	10	6	-2	-12	-23
Whitman (2)	Washington	-5	10	-3	-9	2	7	0	14	4	-8	-10	-16
Hill	Montana	3	-8	-9	-10	5	10	-13	-13	-23	-23	-9	1
Toole	Montana	-4	-8	-6	-9	-8	-8	-10	-11	-11	-12	-8	-5
Hand (1)	S. Dakota	17	5	-5	0	-6	-5	0	-7	-13	-3	-6	-13
Hand (2)	S. Dakota	17	7	0	-4	-3	-3	0	-1	-5	-12	-3	-10
Bias		2.8	2.5	-1.5	-2.6	2.8	2.9	-8.4	1.2	-2.9	-4.3	-2.8	-5.1
SD		9.6	6.8	6.1	7.2	6.1	7.5	8.8	8.5	9.5	10.2	9.2	9.2
RMSE		9.7	7.1	6.1	7.4	6.6	7.6	12.0	8.3	9.6	10.8	9.4	10.3

---

[a] Historical crop calendars are over the following periods: Kansas, 1963-1973; Montana, 1969-1975; North Dakota, 1952-1964; South Dakota, 1960-1969; and Texas, 1964-1966 and 1971-1973. The period of years (obtained from USDA Bulletin 283) is not specified for Indiana, Idaho, Minnesota, and Washington.

TABLE 5-8.– LACIE PHASE III ACC AND HISTORICAL CRD CALENDARS WITH OBSERVED DEVELOPMENT STAGES IN THE 1977 SPRING WHEAT ITS'S

[Comparison by days]

ITS		ITS minus ACC						ITS minus historical					
		Jointing		Heading		Soft dough	Ripe	Jointing		Heading		Soft dough	Ripe
County	State	3.0	3.5	4.0	4.5	5.0	6.0	3.0	3.5	4.0	4.5	5.0	6.0
Hand (1)	S. Dakota	-10	-5	-2	-8	1	5	-11	-17	-20	-15	-8	-3
Hand (2)	S. Dakota	-10	-8	-2	-3	-3	-3	-10	-19	-19	-14	-11	-10
Burke	N. Dakota	NA	NA	NA	NA	22	21	NA	NA	NA	NA	8	2
Williams	N. Dakota	0	5	2	4	12	10	-6	-4	-10	-9	-3	-8
Hill	Montana	10	12	6	6	15	14	-10	-14	-22	-20	-9	-5
Liberty	Montana	19	22	19	11	27	34	3	0	-7	-8	11	20
Toole	Montana	2	0	-1	6	12	17	-6	-11	-17	-8	4	15
Polk	Minnesota	-7	-5	-2	6	8	5	-19	-21	-20	-5	0	-3
Bias		0.6	3.0	2.9	3.1	11.8	12.9	-8.4	-12.3	-16.4	-11.3	-1.0	1.0
SD		10.9	10.8	7.7	6.4	10.0	11.4	6.7	7.8	5.7	5.2	8.2	10.9
RMSE		10.1	10.5	7.7	6.7	15.0	16.7	10.4	14.3	17.3	12.3	7.7	10.2

TABLE 5-9.— LACIE PHASE III ACC AND HISTORICAL CRD CALENDARS WITH OBSERVED DEVELOPMENT STAGES IN THE 1977 SPRING WHEAT CANADIAN ITS

[Comparison by days]

ITS		ITS minus ACC						ITS minus historical					
Town	Province	Jointing		Heading		Soft dough	Ripe	Jointing		Heading		Soft dough	Ripe
		3.0	3.5	4.0	4.5	5.0	6.0	3.0	3.5	4.0	4.5	5.0	6.0
Fort Saskatchewan	Alberta	-1	0	-7	-11	4	-19	-9	-6	-6	-6	4	-4
Olds	Alberta	10	7	4	3	14	NA	0	2	2	1	23	NA
Lethbridge	Alberta	12	13	10	9	7	0	-2	0	-3	0	2	-10
Dawson Creek	British Columbia	-3	2	-3	-5	-6	4	0	4	5	7	8	5
Stony Mountain	Manitoba	6	3	1	2	3	4	-3	-7	-9	-5	-2	-5
Starbuck	Manitoba	4	0	-3	-3	0	5	-4	-8	-12	-10	-6	-8
Altona	Manitoba	3	-1	-8	-9	-6	-6	0	-8	-14	-12	-11	-17
Delisle	Saskatchewan	11	5	0	10	8	10	7	5	2	14	13	14
Swift Current	Saskatchewan	9	5	-4	7	4	0	2	-4	-10	4	3	0
Torquay	Saskatchewan	7	3	-2	-2	1	6	0	-4	-8	-6	-3	-7
Melfort	Saskatchewan	9	9	7	6	12	-7	0	0	0	0	6	-3
Bias		6.1	4.2	-0.5	0.6	3.7	-0.3	-0.8	-2.4	-4.8	-1.2	3.4	-3.5
SD		4.9	4.2	5.6	7.2	6.4	8.4	4.0	4.8	6.4	7.7	9.3	8.5
RMSE		7.7	5.8	5.4	6.9	7.2	8.0	3.9	5.1	7.8	7.4	9.5	8.8

time. Correspondingly, the difference between historical and ITS estimates was 7 days or less 46 percent of the time.

In the U.S. spring wheat region, the average ACC estimates were ahead of the ground-truth estimates for the entire development of the plant. The difference between the two estimates was smallest at jointing (0.6 day) and generally increased as the crop progressed toward maturity to a value of almost 13 days at the ripening stage. The average historical values were approximately 8 days later than the ITS at jointing, regressed further at heading to 16 days behind, and then approached the ground-observed values at soft dough and ripening. At heading, the historical estimates were significantly different from zero at the 1-percent level, whereas at soft dough the model's adjusted calendars were significantly different from zero at the 5-percent level.

During Phase III, the effects of the extended drought in the northern intermountain and western regions were still being felt in Montana and to a lesser degree in North Dakota. Planting was delayed at numerous fields in these two states, especially at the ITS in Liberty, Montana. The spring wheat starter model did not account for these deferred plantings and thus generated planting dates for these states that were early. An abundance of rain fell in July, after the wheat had headed, and tended to slow the crop's actual development. The model did not respond to this slower development rate and thus in Montana and North Dakota advanced still further ahead of the ITS values as the crop proceeded toward ripening. At the same time, the combined historical calendars, which had averaged some 16 days behind ground truth at heading, were within a day of the observed value at ripening.

In Canada's spring wheat areas, the average model estimates were ahead of ground observations by 4 to 6 days at jointing and by 4 days at soft dough. There was little difference, on the average, between the observed and the predicted values at heading and ripening. SD's varied between 4 days at midjointing to 8 days at ripening. The historical calendars also proved to be close to the ground-observed values, with average differences at the various stages ranging from 5 days behind (at heading) to 3 days ahead (at soft dough). Corresponding SD's varied from 4 days at jointing to 9 days at soft dough. At the 5-percent level, the normal calendar at the development stage of heading was significantly different from zero.

While these results indicate that overall the model estimates provided more accurate information than was available from the historical averages, several issues should be addressed before the ACC technology can be considered adequate. For CAMS, the analyst must know, early in the season, the expected spectral appearance of the wheat canopy. This signature, however, is related not only to the wheat's growth stage but also to other factors; for example, whether the field is irrigated and whether it was fallow the previous year and the soil color. Thus, a signature model incorporating the ACC parameter as input would be a more desirable product from the analyst's point of view. Another major issue to be addressed is understanding just how crop calendar errors affect labeling accuracy. At present, these effects are only qualitatively understood.

Whatever the ACC model requirements, the model can be improved for winter wheat by developing an additional model to predict the actual planting date. Currently, the LACIE ACC is "started" (i.e., the clock is set to 1.0, and meteorological data are used as variables in the model) on a date determined to be the historical average planting date for the CRD in which the segment is situated. Since this average planting date can vary considerably from one year to the next, a sizable error can be introduced into growth-stage estimation before dormancy for winter wheat. In tests where the ACC has been "started" on the basis of the ground-observed planting date, the ACC BMTS estimates have been more accurate prior to dormancy.

6. AA SPECIAL STUDIES

The results of special studies by AA during Phase III are presented in this section. These studies were undertaken to evaluate methods employed by CAMS, YES, CAS, and AA in the computation and analysis of LACIE products.

Included among these special studies are the evaluation of the dot-labeling errors, the effect of area and yield biases on the production bias, the effect of ratio errors, the comparison of dot-count and digitized ground-truth wheat proportions, the evaluation of bias correction, and the evaluation of the registration of ground-truth images, etc.

6.1 CONTRIBUTIONS OF AREA AND YIELD ERROR TO PRODUCTION ESTIMATION

This section contains a discussion of the contributions of area and yield errors to the LACIE production estimates for winter, spring, and total wheat. Actually, a study of this kind should be concerned with pseudozone level estimates, but USDA/SRS area and yield estimates are available only at the state level. Therefore, this analysis consists of a cursory examination of plots which depict LACIE winter, spring, and total wheat production estimates with either area or yield errors removed. The respective USDA/SRS production estimates are included in the plots.

The following calculations were made for each crop type and month:

a. For LACIE production estimates without area errors, the LACIE state-level yield estimates were multiplied by the USDA/SRS state-level area estimates and the products were summed for each wheat crop type (winter, spring, and total).

b. For LACIE production estimates without yield errors, the LACIE state-level area estimates were multiplied by the USDA/SRS state-level yield estimates and the products were summed for each wheat crop type.

The area and yield error contributions to production estimates are shown in figure 6-1.

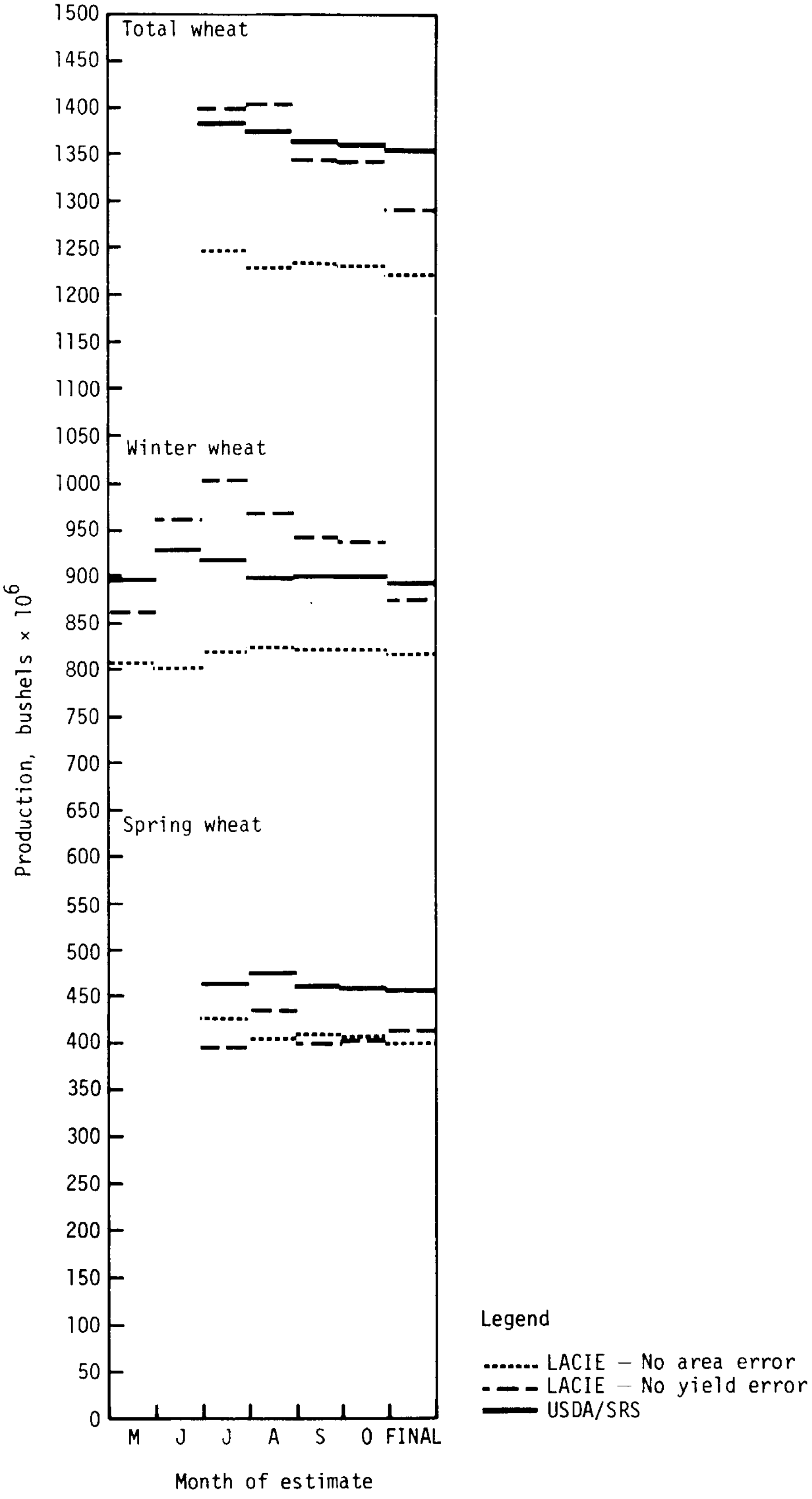


Figure 6-1.– Area and yield error contributions to production estimates.

For winter wheat, the LACIE production estimates calculated without yield errors were invariably closer to the corresponding USDA/SRS monthly figures than were the LACIE figures calculated without area errors, which indicates that the yield error introduced more bias than the area error. For spring wheat, the indications were not as consistent and the results were inconclusive.

6.2 CONTRIBUTION OF THE CLASSIFICATION AND RATIO ERRORS TO WHEAT PROPORTION ESTIMATION

The wheat proportion estimate for a segment is determined by multiplying two terms, a wheat-to-small-grains ratio and the CAMS small-grains proportion estimate. The values of the first term are determined by two methods, (1) ratios derived for the five USSGP states from an analysis of county historical data and (2) ratios derived for the USNGP states using an econometric model for predicting confusion crop ratios. The values of the second term are computed through LACIE classification procedures. Thus, an error in either term affects the wheat proportion estimate. The objective of the following analysis is to determine which term contributes more to the total error.

The true ratio of wheat to small grains and the true proportion of small grains are determined from a digital analysis of aerial photographs delineated to describe each agricultural and nonagricultural feature according to the code specified in appendix G. The digital analysis sums the areas and produces area proportions of each feature delineated. The true proportions for each crop in a segment are thus known, and the accuracy of the classification and the historical ratio values can be assessed.

It is necessary to differentiate between winter small grains and spring small grains in order to compute the true ratio values. In the winter wheat (USSGP) states, all small grains (winter wheat, barley, rye, flax, oats) are considered to be winter small grains. However, in the USNGP states, the econometric model defines winter small grains as consisting of winter wheat and rye and the spring small grains as consisting of spring wheat, barley, oats, and flax. These definitions apply to the winter segments in the mixed winter-spring wheat states of South Dakota and Montana and in the spring wheat states of Minnesota and North Dakota.

The bias (B) and the mean-square error (MSE) of the wheat proportion estimate for a segment may be estimated [denoted by a circumflex (^)]

$$\hat{B} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{r}_i\hat{X}_i - r_iX_i\right)$$

and

$$\widehat{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{r}_i\hat{X}_i - r_iX_i\right)^2$$

where

n is the number of blind sites,

r_i is the true ratio of wheat to small grains ($i = 1, 2, \cdots, n$),

$\hat{r}_i$ is the estimate of r_i (Phase III CAS ratio),

X_i is the true proportion of small grains, and

$\hat{X}_i$ is the estimate of X_i (Phase III CAMS final estimate).

It is clear that these errors are caused by two factors: the CAMS classification of small grains and the estimated ratio of wheat to small grains. The contribution of a particular error factor may be measured by the reduction in the bias or MSE which would be achieved if that error factor were omitted. Specifically, the following formulas are used in this study.

a. Proportion bias estimate without ratio error:

$$\hat{B}' = \frac{1}{n}\sum_{i=1}^{n}\left(r_i\hat{X}_i - r_iX_i\right)$$

b. Proportion bias estimate without classification error:

$$\hat{B}'' = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{r}_i X_i - r_i X_i\right)$$

c. Proportion MSE without ratio error:

$$\hat{MSE}' = \frac{1}{n}\sum_{i=1}^{n}\left(r_i \hat{X}_i - r_i X_i\right)^2$$

d. Proportion MSE without classification error:

$$\hat{MSE}'' = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{r}_i X_i - r_i X_i\right)^2$$

e. The 90-percent confidence limits for the biases:

$$\hat{B} \pm t_{(0.05,n-1)}\frac{S}{\sqrt{n - 1}}$$

where $t_{(0.05,n-1)}$ is the value of the 95-percent point from the Student's t-distribution with n - 1 degrees of freedom and S is the SD of the classification error.

Data from 123 blind sites in the nine states of the USGP were used in the analysis. Total small-grain proportion estimates from all 123 sites were used by the CAS aggregation (appendix D).

The average biases and MSE's of the wheat proportion estimates (Phase III final) and the predicted biases and MSE's for LACIE proportion estimates without ratioing error and without classification error are given in table 6-1. Also shown are standard error of the biases, the percentage of reduction in bias and MSE for the estimates without ratioing and classification errors, and 90-percent confidence limits for the true biases.

TABLE 6-1.— BIASES AND MSE's

(a) Pure spring wheat states

	Bias of estimate, %	SD of bias	Reduction in absolute bias, %	90% confidence limits for bias	MSE of estimate	Reduction in MSE, %
North Dakota (19 sites)						
Phase III final	-4.4	7.3	—	(-7.4, -1.4)	70	—
No ratioing error	-3.8	5.6	13.27	(-6.1, -1.5)	44	37.61
No classification error	-0.9	3.9	78.95	(-2.5, 0.7)	15	78.20
Minnesota (11 sites)						
Phase III final	-5.4	8.6	—	(-10.4, -0.5)	97	—
No ratioing error	-2.2	2.9	59.59	(-3.9, -0.5)	13	87.01
No classification error	-1.0	6.7	80.81	(-4.9, 2.8)	42	57.01

TABLE 6-1.— Continued.

(b) Mixed wheat states

	Bias of estimate, %	SD of bias	Reduction in absolute bias, %	90% confidence limits for bias	MSE of estimate	Reduction of MSE, %
South Dakota — spring wheat sites (9 sites)						
Phase III final	-3.7	6.4	—	(-7.9, 0.5)	50	—
No ratioing error	-4.4	4.3	-16.94	(-7.1, -1.6)	35	29.89
No classification error	0.6	4.4	116.22	(-2.3, 3.6)	18	64.16
Montana — winter wheat sites (13 sites)						
Phase III final	-0.7	4.9	—	(-3.2, 1.9)	23	—
No ratioing error	-0.7	4.9	0.0	(-3.2, 1.9)	23	0.0
No classification error	0	0	100.0	(0, 0)	0	100.0
Montana — spring wheat sites (7 sites)						
Phase III final	-0.8	4.9	—	(-4.8, 3.2)	21	—
No ratioing error	-0.7	1.8	12.50	(-2.1, 0.8)	3	85.71
No classification error	-0.2	5.3	75.00	(-4.5, 4.2)	24	-14.29

TABLE 6-1.— Continued.

(c) Pure winter wheat states

	Bias of estimate, %	SD of bias	Reduction in absolute bias, %	90% confidence limits for bias	MSE of estimate	Reduction in MSE, %
Colorado (11 sites)						
Phase III final	-2.4	5.9	—	(-5.8, 0.9)	37	—
No ratioing error	-3.1	5.6	-26.03	(-6.3, 0.2)	38	-2.29
No classification error	0.8	1.4	131.40	(0, 1.6)	2	93.81
Nebraska (16 sites)						
Phase III final	-3.0	6.2	—	(-5.8, -0.2)	45	—
No ratioing error	-3.6	5.6	-22.63	(-6.2, -1.1)	43	4.49
No classification error	0.8	1.3	127.70	(0.3, 1.39)	2	95.08
Kansas (13 sites)						
Phase III final	-3.9	4.7	—	(-6.4, -1.5)	36	—
No ratioing error	-5.1	4.5	-30.46	(-7.4, -2.9)	45	-24.88
No classification error	1.8	3.9	144.42	(-0.3, 3.8)	17	52.58

TABLE 6-1.— Continued.

(c) Concluded

	Bias of estimate, %	SD of bias	Reduction in absolute bias, %	90% confidence limits for bias	MSE of estimate	Reduction in MSE, %
Oklahoma (15 sites)						
Phase III final	-6.1	14.5	—	(-12.9, 0.7)	232	—
No ratioing error	-6.0	14.2	1.64	(-12.7, 0.7)	224	3.66
No classification error	-0.3	1.1	95.58	(-0.8, 0.3)	1.	99.45
Texas (9 sites)						
Phase III final	-2.2	9.9	—	(-8.7, 4.3)	91	—
No ratioing error	-1.0	9.4	53.88	(-7.2, 5.2)	80	12.54
No classification error	-1.4	2.2	37.90	(-2.8, 0.1)	6	93.44

TABLE 6-1.— Continued.

(d) Subgroup states

	Bias of estimate, %	SD of bias	Reduction in absolute bias, %	90% confidence limits for bias	MSE of estimate	Reduction in MSE, %
Pure spring wheat states (30 sites)						
Phase III final	-4.8	7.7	—	(-7.2, -2.3)	80	—
No ratioing error	-3.2	4.8	32.77	(-4.7, -1.7)	32	59.56
No classification error	-1.0	5.0	79.83	(-2.5, 0.6)	25	68.79
Mixed wheat states (29 sites)						
Phase III final	-1.6	5.4	—	(-3.4, -0.1)	31	—
No ratioing error	-1.8	4.4	-26.86	(-3.2, -0.4)	22	28.98
No classification error	0.2	3.4	93.14	(-1.0, 1.3)	11	63.12
USSGP — winter wheat (64 sites)						
Phase III final	-3.7	8.9	—	(-5.6, -1.9)	92	—
No ratioing error	-3.8	8.8	-2.16	(-5.6, -2.0)	91	1.49
No classification error	0.4	2.3	111.89	(0, 0.9)	6	93.97

TABLE 6-1.— Concluded.

(e) USGP totals

	Bias of estimate, %	SD of bias	Reduction in absolute bias, %	90% confidence limits for bias	MSE of estimate	Reduction in MSE, %
USGP — 9 states (123 sites)						
Phase III final	-3.5	7.9	—	(-4.7, -2.3)	75	—
No ratioing error	-3.3	7.1	6.57	(-4.3, -2.2)	60	19.08
No classification error	0.0	3.4	99.43	(-0.5, 0.5)	12	84.41

Tables 6-1(a) through (d) show the results of the analysis by state. The states are also subgrouped in table 6-1(d) as spring wheat states, mixed wheat states, and winter wheat states. The USGP totals are shown in table 6-1(e).

A review of tables 6-1(a) through (c) shows that in all cases except Texas, the reduction in the bias was greater when the classification error was eliminated (i.e., all errors were due to ratioing error) than when the ratioing error was eliminated. In four of the nine states, the positive ratioing errors compensated for the negative classification errors, resulting in slight improvement of the final error.

In the summations of subgroups of states [table 6-1(d)], the reduction in bias was greater when the classification error was eliminated (i.e., all errors were caused by the ratioing error) than when the ratioing error was eliminated.

When data from all 123 blind sites in all nine states of the USGP [table 6-1(e)] are included in the sample, the bias resulting from the classification error is negative and almost as large as the bias of the final estimates. The bias resulting from ratioing error was not significant.

Examination of the MSE's again indicates that errors resulting from classification are a bigger problem than errors resulting from ratioing.

6.3 DETAILED ANALYSIS OF CAMS PROCEDURES FOR PHASE III USING GROUND-TRUTH INVENTORIES

The digitized ground-truth inventories developed during Phase III may be used in detailed analyses of the CAMS classification procedure (P1 only) to determine the proportion of small grains in a segment. All of the processings used in this study were passed to the aggregation system as good, although some were not used in the aggregations.

To conduct this pixel level study, it is necessary to use only those CAMS classifications which do not have any pixels designated other (DO) or designated unidentifiable (DU). Examples of DO pixels are mountain ranges and other nonagricultural features which should be excluded from a segment; examples of DU pixels are clouds which obscure the ground surface and cloud shadows which effectively change the spectral values on the ground. The number of CAMS classifications used in this study is consequently considerably less than those classifications used in other spectral studies (see section 6.7) in this report.

The CAMS classification procedure (P1) follows these steps:

a. The analysts label two sets of dots as wheat or nonwheat.

b. One set of analyst-labeled dots (type 1 dots) is used as seed pixels to group all the pixels in the segment into clusters on the basis of their spectral values.

c. Each of the clusters is labeled as wheat or nonwheat by the type 1 analyst-labeled dot closest to the mean of the cluster.

d. On the basis of the means and variances for each cluster, every pixel in the segment is classified as either wheat or nonwheat.

e. The second set of analyst-labeled dots (type 2 dots) is used as a random sample of the segment to correct the machine classification proportion for any bias introduced by the classification process.

The proportion of wheat in a segment can be estimated at four steps in the procedure:

a. The type 2 dots can be used as a random sample of the segments to determine a proportion.

b. At the machine clustering stage, a proportion can be determined using the analyst label for each cluster.

c. The machine classification proportion is calculated using CAMS procedures.

d. Bias-corrected machine proportion is calculated using CAMS procedures.

If the procedure is effective, the proportion estimate should improve at each step. The CAMS procedures will be evaluated by calculating the proportion of small grains at each of these four steps: type 2 dots as a random sample, machine clusters, machine classification, and bias-corrected machine classification.

The results of these studies will be given for three groups: winter wheat segments, spring wheat segments, and mixed wheat segments. The winter wheat segments were those located in Colorado, Kansas, Nebraska, Oklahoma, and Texas; the spring wheat segments were those in Minnesota and North Dakota. All of the segments in Montana and South Dakota were grouped as mixed wheat although some of these segments were processed as winter or spring wheat.

When necessary to aggregate the pixels in a segment into small grains and nonsmall grains, winter wheat, spring wheat, barley, rye, flax, and oats were aggregated as small grains and all other crops were aggregated as nonsmall grains.

6.3.1 CAMS CLASSIFICATION RESULTS

Figures 6-2 through 6-5 show the errors in the estimates at each of the four stages in the CAMS procedure, using the last processing for each segment. The errors are plotted as a function of the true proportion of small grains for each segment. The general trend with each of the four estimates is an underestimation of the small grains', proportion, with the worst errors occurring for large proportions of small grains.

The mean error and the SD of the mean error were calculated to quantify the errors. The mean error gives a measure of the bias of the estimator, and the SD is a measure of the variability. The MSE, a measure of the overall performance because it considers both bias and variability, was also calculated. The mean error, the SD, and the MSE are shown in table 6-2; these results indicate that the estimate of small-grain proportion did not improve significantly from one step to the next. In all cases, the bias was approximately 6 percent with an SD of approximately 10 percent.

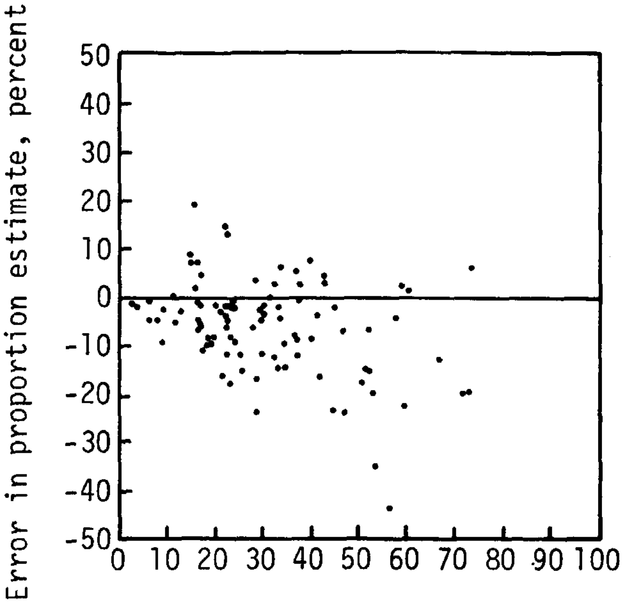


Figure 6-2.— Analyst-labeled type 2 dots as random sample.

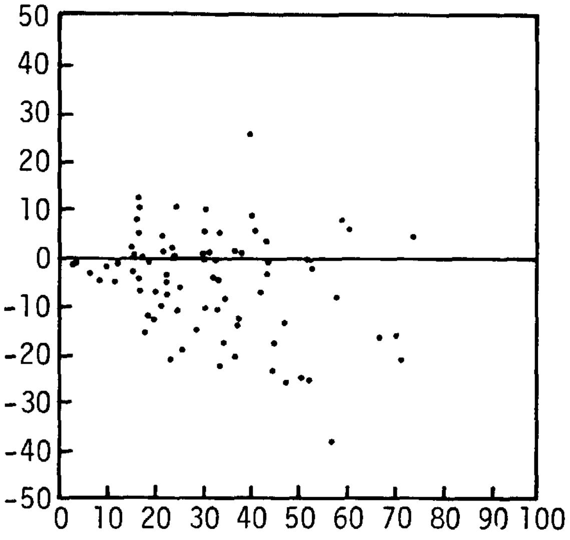


Figure 6-3.— Machine clusters with analyst labels.

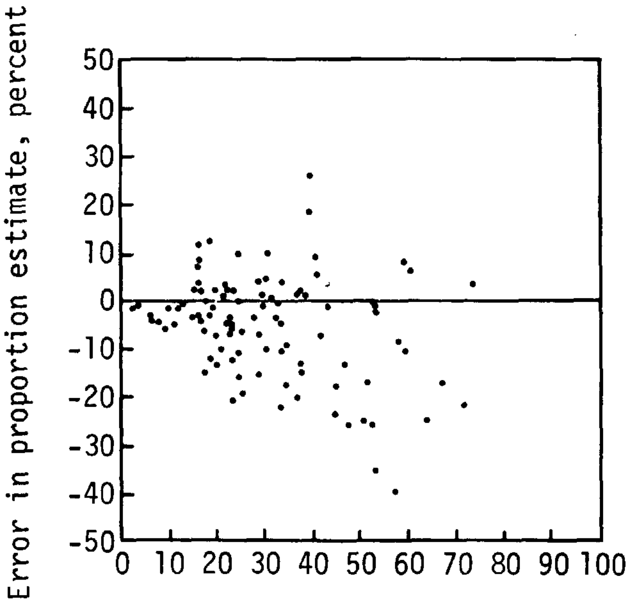


Figure 6-4.— Machine classification.

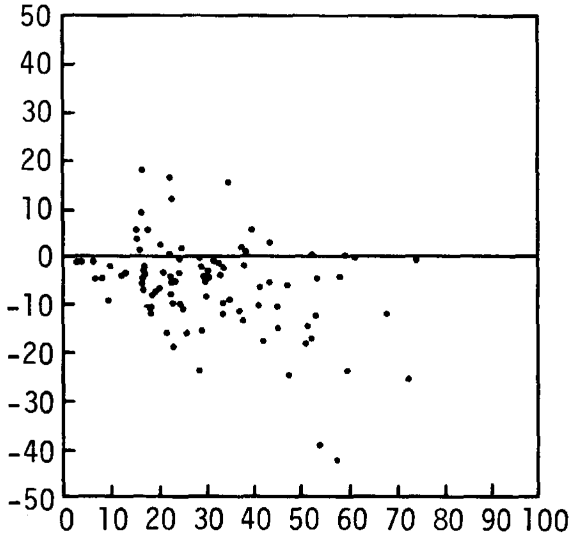


Figure 6-5.— Bias-corrected machine classification.

TABLE 6-2.– CAMS CLASSIFICATION ERRORS

[Using final processing for each segment]

Source of classification result	Winter wheat				Spring wheat				Mixed wheat				All categories			
	No. of processings	Mean error	SD	MSE	No. of processings	Mean error	SD	MSE	No. of processings	Mean error	SD	MSE	No. of processings	Mean error	SD	MSE
Type 2 dots as random sample	43	-6.4	10.3	146	26	-7.7	9.5	146	27	-5.4	9.2	110	96	-6.5	9.8	136
Machine clusters with type 1 labels	33	-5.6	11.7	163	21	-4.7	11.2	142	24	-5.8	9.3	118	78	-5.4	10.7	143
Machine classification	43	-5.7	10.3	135	26	-6.0	12.9	197	27	-5.8	9.1	111	96	-5.8	10.6	145
Bias-corrected machine classification	43	-6.6	9.9	140	26	-7.7	9.6	147	27	-5.4	8.4	97	96	-6.6	9.4	130

Another way of analyzing the procedure is to calculate the improvement (the difference in absolute value) in the error between any two steps. A positive improvement indicates that the error was less in the latter step than in the earlier step. The percentage of processings in which there was an improvement can also be calculated. If the step is effective, the percentage of processings improved should be greater than 50 percent, and the mean improvement should be greater than zero. These calculations for the CAMS results are shown in table 6-3. All of the comparisons indicate very little improvement in the error in any of the steps; overall, about half the processings improved, and half the processings became worse. The mean improvement was less than 0.5 percent.

Analysis of the differences between machine classification estimates and machine clustering estimates showed the mean improvement to be 0.04 percent with an SD of the mean improvement of 0.46 percent. In performing a linear regression of the machine classification error against the machine clustering error, the analyst found the slope to be 1.003 with an intercept of -0.185. The coefficient of determination for the regression was 0.9985. This result indicates that the classification results are essentially the same as the clustering results. A plot of the classification error as a function of clustering error is shown in figure 6-6. A pixel-level comparison was made between the classification results and the clustering results to investigate this relationship further. This comparison indicates that an average of 96 percent of the pixels do not change their label from the clustering to the classification stage and that the average net change in pixel counts was only 0.3 percent, indicating that the classification is unnecessary.

6.3.2 CAMS CLASSIFICATION RESULTS USING GROUND-TRUTH DOT LABELING

The bias and the variability in the estimates produced by the CAMS procedure are caused by the procedure itself and by bad input data in the form of mislabeled type 1 and type 2 dots. If one could reprocess the segments using the true labels for the type 1 and type 2 dots, any bias or variability in the results would be due to the procedure itself and not to bad input data.

TABLE 6-3.— CAMS CLASSIFICATION IMPROVEMENT

[Using final processing for each segment]

Classification sources compared	Winter wheat		Spring wheat		Mixed wheat		All categories	
	Processings improved, %	Mean improvement	Processings improved, %	Mean improvement	Processings improved, %	Mean improvement	Processings improved, %	Mean improvement
Clusters vs. type 2 dots	55	0.4	43	-1.4	42	-0.9	47	-0.5
Machine classification vs. clusters	52	0.1	24	-0.2	38	0.0	40	0.0
Bias-corrected machine classification vs. machine classification	51	-0.3	58	1.9	52	0.5	53	0.5
Machine classification vs. type 2 dots	58	0.7	38	-1.3	48	0.1	50	0.0
Bias-corrected machine classification vs. type 2 dots	53	0.4	50	0.6	44	0.6	50	0.5

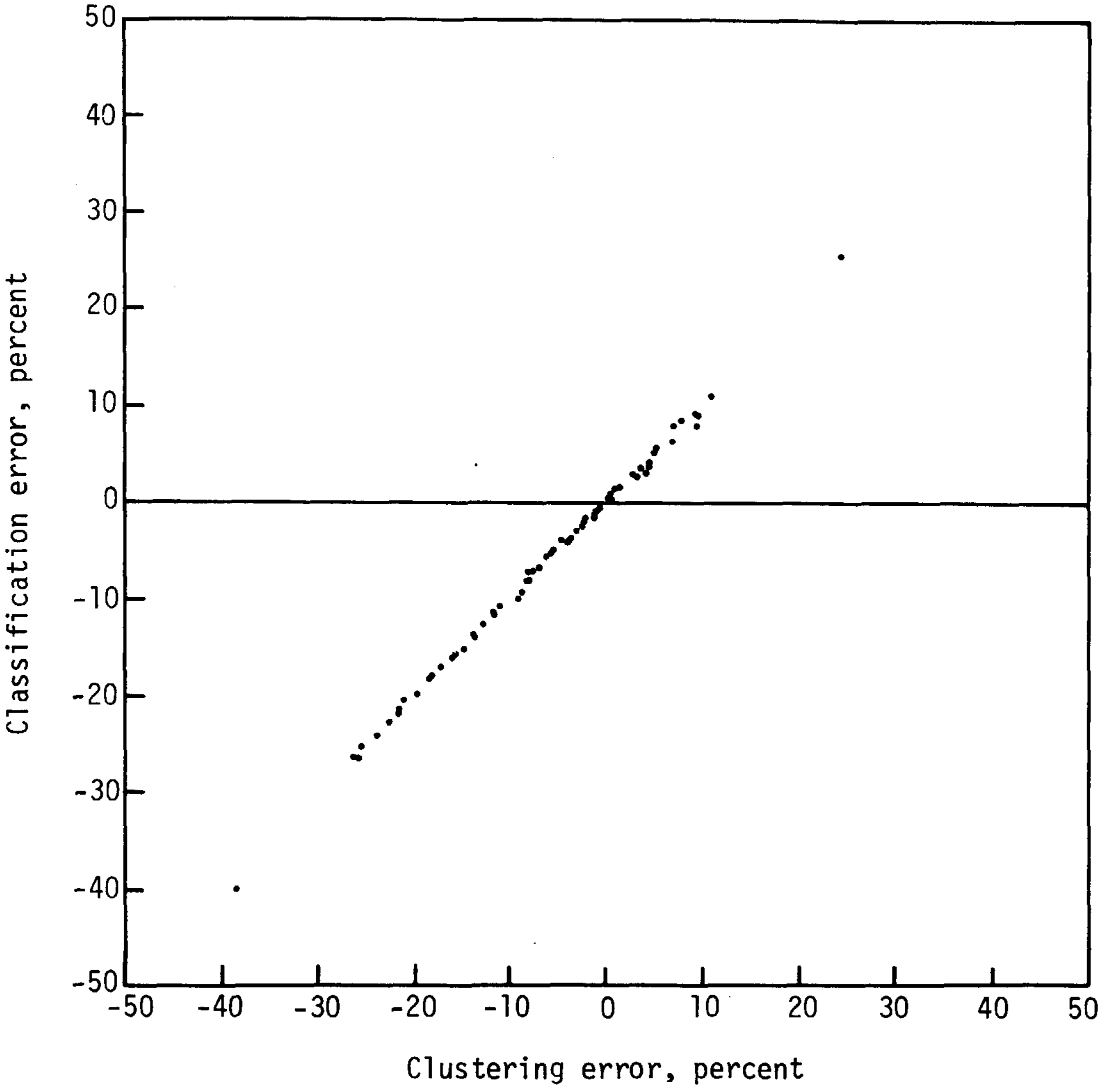


Figure 6-6.— Classification versus clusters.

Reprocessing all of the segments would be a big project; an easier way is to modify the CAMS results to reflect true dot labels instead of analyst labels. For the random-sample estimate using type 2 dots, it is a simple matter to replace the analyst labels with the true labels and recalculate the proportion. The clustering proportion is determined by aggregating the clusters on the basis of the analyst label for the dot closest to the mean of each cluster. The ground-truth clustering proportion can be determined by aggregating the clusters on the basis of the true dot label instead of the analyst dot label. It is not possible to reproduce the machine classification results using true labels because means and variances of the clusters are used to classify the pixels. One does not have this information based on true labels. However, comparison of the classification results with the clustering results using analyst labels indicates that the results are identical. It can be assumed, therefore, that the classification results would be identical to the clustering result if true labels were used. The bias correction can be performed by comparing the ground-truth labels for type 2 dots with the label for the cluster in which the dot lies. The CAMS results can thus be reproduced by using ground-truth labels without reprocessing the segments.

The CAMS results using ground-truth labels for type 1 and type 2 dots are shown in figures 6-7, 6-8, and 6-9, which can be compared with the actual CAMS results in figures 6-2 through 6-5. The scatter in the error is much less using ground-truth labels, and there is no underestimation for large proportions of small grains. The clustering estimates have more variability than the random sample and the bias-corrected estimates. The mean error, the SD, and the MSE for the CAMS results using ground-truth labels are shown in table 6-4. As could be expected, the clustering estimates have a great deal more variability than the random sample or bias-corrected estimates. The bias of the clustering estimate was less than 0.5 percent, indicating that the classification is essentially unbiased. The clustering does increase the variability significantly. The bias correction reduces the errors to about the same level as for the random sample.

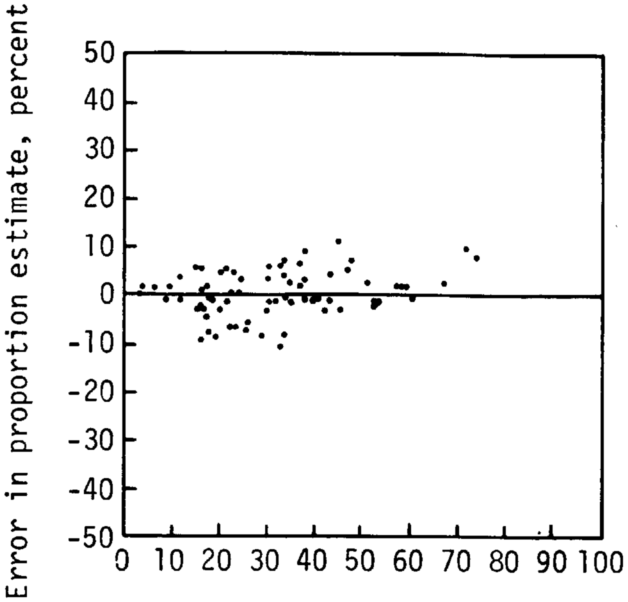


Figure 6-7.— Ground-truth labeled type 2 dots (limited).

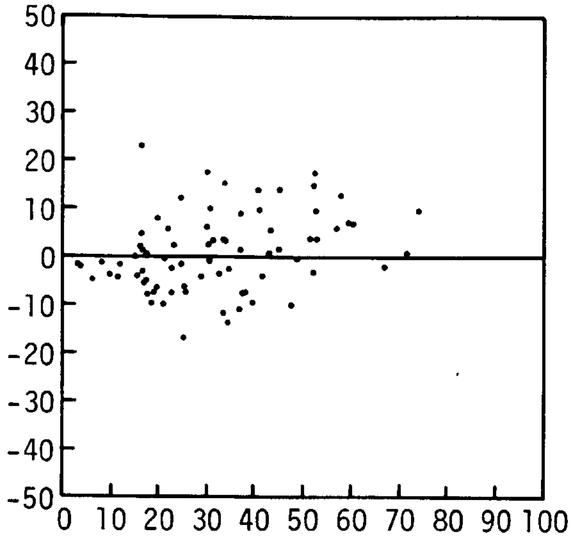


Figure 6-8.— Machine clusters with ground-truth labels.

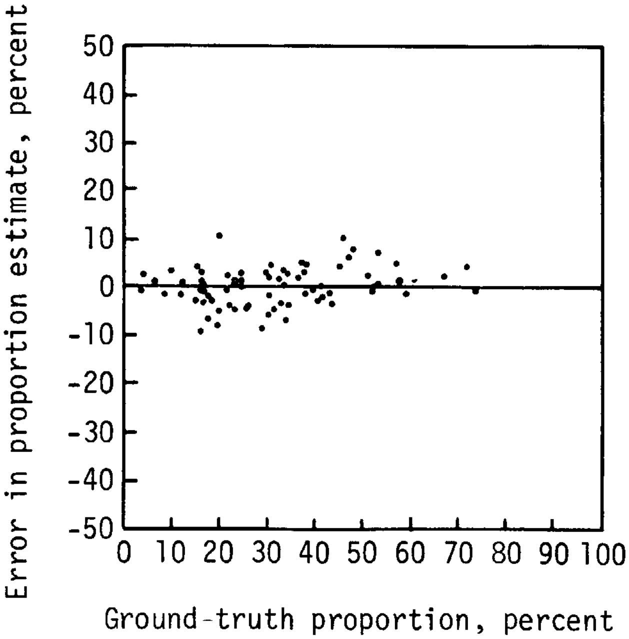


Figure 6-9.— Bias-corrected machine clusters.

TABLE 6-4.— CAMS CLASSIFICATION ERRORS FOR GROUND-TRUTH DOT LABELS

[Using final processing for each segment]

Source of classification result	Winter wheat				Spring wheat				Mixed wheat				All categories			
	No. of processings	Mean error	SD	MSE	No. of processings	Mean error	SD	MSE	No. of processings	Mean error	SD	MSE	No. of processings	Mean error	SD	MSE
Type 2 dots as random sample	31	-1.1	5.0	26	20	1.0	4.2	18	24	0.4	4.7	21	75	-0.1	4.7	22
Machine clusters (machine classification)	31	0.0	8.2	64	20	4.1	8.0	78	24	-2.4	6.5	46	75	0.4	7.9	62
Bias-corrected machine classification	31	-1.2	3.8	16	20	0.9	3.8	15	24	0.9	4.3	19	75	0.1	4.1	16

Table 6-5 shows the relative improvement between the three estimates. Clustering made the estimate worse for 71 percent of the segments. The bias-corrected estimate was better than the random sample for 57 percent of the processings, but the mean improvement was only 0.5 percent.

The results using ground-truth dot labels indicate that the 6-percent negative bias and about half of the variability are due to analyst dot-labeling errors. The procedure is capable of producing an unbiased estimate with an SD of about 4 percent.

6.3.3 ANALYST DOT-LABELING ACCURACY

Because analyst dot-labeling errors are so important, the analyst labeling accuracy was studied in detail. The labeling accuracy was determined for over 7000 type 1 dots and 12 037 type 2 dots. The dots used in this study were from all processings for each segment; classification results presented in section 6.3.1 were for only the last processing for each segment.

Tables 6-6 and 6-7 show the analyst dot-labeling accuracy for type 1 and type 2 dots. The analysts labeled small-grains dots correctly about 61 percent of the time; the labeling accuracy for nonsmall grains was about 93 percent. In the strip-fallow categories, the dots were labeled as small grains about 42 percent of the time. Because strip-fallow categories are half small grains and half nonsmall grains, the strip-fallow dots should be labeled as small grains 50 percent of the time. Therefore, the labeling accuracy for strip-fallow categories is really 85 percent, which is better than the 61-percent labeling accuracy for small grains.

These results are consistent with the underestimation of the small-grain proportion by the CAMS procedure. The analyst does a good job of labeling nonsmall-grain pixels but mislabels many of the small-grain pixels.

The accuracy for labeling type 1 dots is slightly better than for type 2 dots, probably because type 1 dots are not labeled if they fall on field boundaries, whereas type 2 dots are labeled regardless of where they fall.

TABLE 6-5.— IMPROVEMENT IN CAMS CLASSIFICATION FOR GROUND-TRUTH DOT LABELS

[Using final processing for each segment]

Classification sources compared	Winter wheat		Spring wheat		Mixed wheat		All categories	
	Processings improved, %	Mean improvement	Processings improved, %	Mean improvement	Processings improved, %	Mean improvement	Processings improved, %	Mean improvement
Clusters (classification) vs. type 2 dots	35	-2.2	5	-4.6	42	-1.4	29	-2.6
Bias-corrected clusters vs. clusters	68	2.9	75	4.8	58	1.9	67	3.1
Bias-corrected clusters vs. type 2 dots	55	0.8	65	0.2	54	0.5	57	0.5

TABLE 6-6.— ANALYST DOT-LABELING ACCURACY FOR PHASE III PROCESSING — TYPE 1 DOTS

Classification	Winter wheat		Spring wheat		Mixed wheat		All categories	
	No. of dots	Correctly labeled, %	No. of dots	Correctly labeled, %	No. of dots	Correctly labeled, %	No. of dots	Correctly labeled, %
Small grains								
Winter wheat	483	61			75	57	558	61
Spring wheat			432	73	140	63	572	70
Barley			187	75	139	38	326	60
Flax			21	24	17	6	38	16
Oats	25	28	152	45	227	71	404	58
Total small grains	508	60	792	67	598	58	1898	62
Strip-fallow small grains[a]								
Winter wheat	48	35			107	46	155	43
Spring wheat			51	37	45	47	96	42
Barley					21	21	21	24
Total strip-fallow small grains	48	35	51	37	173	43	272	41
Nonsmall grains								
Alfalfa	49	90	106	90	151	79	306	85
Beans	19	95					19	95
Corn	159	98	193	95	225	92	577	94
Sunflower			104	98			104	98
Sudan grass	10	90			12	100	22	95
Sorghum	178	92			26	100	204	93
Soybeans and guar	40	100	36	100	11	82	137	99
Sugar beets			27	93	14	100	41	95
Grass	47	98	67	94	125	90	239	93
Hay	25	88	63	89	116	83	204	85
Pasture	933	97	354	92	1218	96	2505	96
Trees	27	85	42	88	41	100	110	92
Cotton	32	97					32	97
Water	27	100	80	100	86	100	193	100
Nonagricultural	37	100	40	98	39	97	116	98
Homestead	51	98	22	91	45	69	118	86
Idle cropland — stubble	13	85			12	92	25	88
Idle cropland — cover crop	10	90					10	90
Idle cropland — residue	33	94			16	100	49	96
Idle cropland — fallow	190	95	139	94	167	93	496	94
Total nonsmall grains	1880	96	1323	94	2304	93	5507	94

[a] The percent correctly labeled for strip-fallow assumes that small grains is the correct label.

TABLE 6-7.— ANALYST DOT-LABELING ACCURACY FOR PHASE III PROCESSING — TYPE 2 DOTS

Classification	Winter wheat		Spring wheat		Mixed wheat		All categories	
	No. of dots	Correctly labeled, %	No. of dots	Correctly labeled, %	No. of dots	Correctly labeled, %	No. of dots	Correctly labeled, %
Small grains								
Winter wheat	712	61			149	55	861	60
Spring wheat			738	68	217	56	955	66
Barley			282	70	210	40	492	57
Rye					16	38	16	38
Flax			27	11	23	30	50	20
Oats	32	19	281	59	440	59	753	58
Total small grains	744	59	1328	66	1055	53	3127	60
Strip-fallow small grains[a]								
Winter wheat	86	36			179	54	277	47
Spring wheat			75	32	107	41	182	37
Barley					69	38	69	38
Total strip-fallow small grains	86	36	75	32	355	47	528	43
Nonsmall grains								
Alfalfa	53	81	159	89	264	78	476	82
Beans			11	91			11	91
Corn	220	97	228	93	366	92	814	94
Sunflower			170	94	29	93	199	94
Sudan grass	14	86	10	100	11	100	35	94
Sorghum	291	95			55	95	346	95
Soybeans and guar	51	82	105	94			156	90
Sugar beets			41	93			41	93
Grass	65	86	120	88	217	89	402	88
Hay	53	98	76	89	188	90	317	91
Pasture	1271	96	478	95	1993	93	3742	95
Trees	46	96	77	81	95	96	218	90
Cotton	57	81					57	81
Millet					14	79	14	79
Water	36	100	95	100	86	100	217	100
Nonagricultural	58	97	55	93	69	100	182	97
Homestead	68	96	48	60	84	85	200	83
Idle cropland — stubble	22	91			11	91	33	91
Idle cropland — cover crop	10	90	12	100			22	95
Idle cropland — residue	25	100			44	95	69	97
Idle cropland — fallow	343	91	244	81	244	94	831	89
Total nonsmall grains	2683	94	1929	90	3770	92	8382	92

[a] The percent correctly labeled for strip-fallow assumes that small grains is the correct label.

The CAMS procedure allows the analyst to change the labels of type 2 dots after the machine classification has been performed. Table 6-8 shows a comparison of the proportion errors for those segments in which type 2 dot labels were changed. There was an overall improvement in the errors when the relabeled dots were used, but in the mixed wheat segments, the errors became worse. To investigate this problem further, the analyst calculated the improvement in dot labeling accuracy for those processings whose dot labels were changed; the results of these calculations are shown in table 6-9. The overall improvement in labeling small-grains dots was 4 percent. In the strip-fallow and nonsmall-grain categories, the improvement was 1 percent; in the mixed wheat segments, the accuracy of labeling small grains went down by 2 percent and that of nonsmall grains went up by 3 percent. The less accurate labeling of small grains coupled with the more accurate labeling of nonsmall grains caused the increased proportion errors observed in the mixed wheat segment.

6.3.4 ANALYSIS OF CLUSTERING EFFECTIVENESS

In the CAMS results using ground-truth dot labels, clustering increased the variability of the estimate from 4 to 7 percent. To investigate this problem, the analyst calculated the cluster purity for all clusters of all processings. A histogram of cluster purity is given in figure 6-10. The number of clusters with a given proportion of small grains is plotted as a function of the small grains' proportion within the cluster. Ideally, this histogram would show a maximum value near zero purity to reflect clustering of nonsmall grains, a second maximum near 100-percent purity to reflect clustering of small grains, and a minimum near 50 percent. Examination of figure 6-10 reveals the nonsmall grains peaking near 5 percent, a minimum near 57 percent, and a maximum barely discernible between 80 and 90 percent. These results show that the clustering does not separate the small grains from the nonsmall grains.

Each cluster is labeled by the dot closest to the cluster mean. If a cluster of small grains is defined as a cluster with more than 50 percent small grains,

TABLE 6-8.— IMPROVEMENT IN CAMS CLASSIFICATION RESULTS

[Relabeled type 2 dots]

Source of classification result	Winter wheat				Spring wheat				Mixed wheat				All categories			
	Mean error	SD	Processing improved, %	Mean improvement	Mean error	SD	Processing improved, %	Mean improvement	Mean error	SD	Processing improved, %	Mean improvement	Mean error	SD	Processing improved, %	Mean improvement
Type 2 dots																
Original	-5.9	9.1			-10.6	9.2			-3.1	9.7			-6.7	9.6		
Relabeled	-4.7	9.8	50	0.0	-7.4	8.2	80	3.2	-6.8	10.1	43	-1.7	-6.3	9.2	58	0.6
Bias correction																
Original	-6.9	8.8			-10.2	7.8			-2.9	8.7			-6.7	8.8		
Relabeled	-5.8	8.8	57	1.1	-7.0	7.7	73	3.1	-6.5	8.2	43	-0.6	-6.5	8.1	58	1.3

TABLE 6-9.— IMPROVEMENT IN ANALYST DOT-LABELING ACCURACY FOR PHASE III PROCESSING

[Relabeled type 2 dots]

Classification	Winter wheat			Spring wheat			Mixed wheat			All categories		
	No. of dots	Original correct, %	Improvement, %	No. of dots	Original correct, %	Improvement, %	No. of dots	Original correct, %	Improvement, %	No. of dots	Original correct, %	Improvement, %
Small grains												
Winter wheat	218	58	+10				27	59	-15	245	58	+8
Spring wheat				218	69	+5	48	50	-2	266	65	+4
Barley				82	54	+3	51	53	-8	133	53	0
Flax				8	0	0	6	67	-17	14	29	-8
Oats	4	25	-25	67	66	+7	118	68	+2	189	66	+4
Total small grains	222	58	+10	375	63	+5	250	60	-2	847	61	+4
Strip-fallow small grains												
Winter wheat	30	50	+3				51	45	-8	81	43	0
Spring wheat				23	22	+17	69	45	0	92	43	0
Barley				1	0	0	32	38	+3	33	39	0
Total strip-fallow small grains	30	50	+3	24	22	+17	152	43	-2	206	42	+1
Nonsmall grains												
Alfalfa	10	70	+10	34	94	0	72	79	+10	116	83	+7
Corn	67	99	-3	39	90	0	93	96	+2	199	95	0
Sunflower				47	96	+2	6	83	0	53	94	+2
Sudan grass	6	100	-17	4	100	0	4	100	0	14	100	-7
Sorghum	77	94	-1				7	100	0	84	94	0
Soybeans and guar	30	87	-4	16	94	-7	1	100	0	1	100	0
Sugar beets	2	100	0	10	100	0				12	100	0
Grass	22	91	0	38	87	-5	40	88	+2	100	88	-1
Hay	13	92	0	16	88	-7	48	88	+2	77	88	0
Pasture	263	97	0	136	97	-1	583	94	+2	982	95	+2
Trees	5	80	0	20	95	-5	23	91	+9	48	92	+2
Cotton	48	81	+7							48	81	+7
Water	3	100	0	30	100	0	5	100	0	38	100	0
Nonagricultural	4	75	+25	20	95	0	5	100	0	29	93	+4
Homestead	24	88	+4	9	56	-23	14	93	-7	47	83	-4
Idle cropland – stubble	7	100	0							7	100	0
Idle cropland – cover				5	100	0	1	100	0	6	100	0
Idle cropland – residue	8	100	0				3	100	0	11	100	0
Idle cropland – fallow	108	92	+1	66	79	+3	60	95	0	234	89	+1
Total nonsmall grains	697	93	+1	490	92	-1	965	92	+3	2106	93	+1

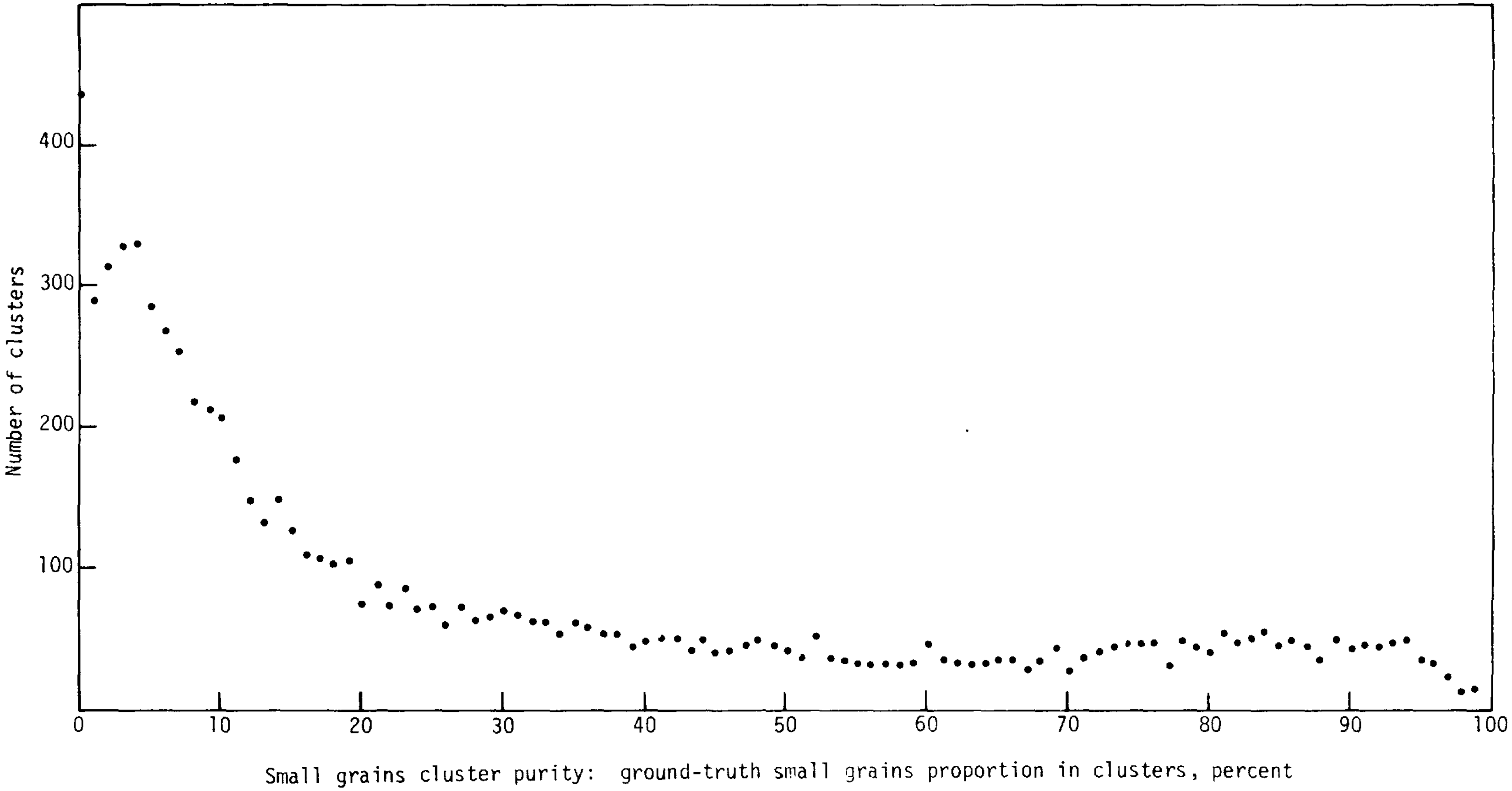


Figure 6-10.— Histogram of cluster purity.

the labeling logic correctly labels the small-grain cluster 70 percent of the time, based on the analyst dot labels. The nonsmall-grain clusters are labeled correctly 91 percent of the time. If ground-truth labels were used instead of the analyst labels, the small-grain clusters were labeled correctly 80 percent of the time and the nonsmall-grains clusters were correctly labeled 83 percent of the time. This indicates that the labeling logic is nearly as effective on small-grain clusters as on nonsmall-grains clusters.

6.3.5 CONCLUSIONS

Based on these studies, the following conclusions are reached:

a. The CAMS proportion estimates have a bias of -6 percent with an SD of 10 percent.

b. The -6 percent bias and half of the SD are caused by the analyst's dot-labeling errors.

c. If the dot labeling were completely accurate, the proportion estimates would be unbiased with an SD of 4 percent.

d. The proportions based on the type 2 dots as a random sample produce as good an estimate as the final bias-corrected result.

e. The proportion estimate produced by the machine classification is identical to the estimate produced by clustering; therefore, machine classification is nonproductive.

f. The -6 percent bias is due to the analysts' labeling nonsmall-grain dots quite well while mislabeling a large portion of the small-grain dots.

g. Relabeling the type 2 dots improved the proportion estimates overall but produced worse estimates in mixed wheat states.

h. Using the clusters in all the P1 processings in the data set, machine clustering does not effectively separate small grains from the nonsmall grains (corn, soybeans, grasses, trees, etc.).

i. The greatest improvement in results would be produced by improving the analyst dot-labeling accuracy.

j. A significant improvement in results would be produced with better clustering.

6.4 ANALYST DOT-LABELING ERROR SOURCES

The results presented in this section are from a comparison of ground-observed and analyst-designated labels of dots from 51 blind sites located in North Dakota, Minnesota, Montana, Colorado, and Oklahoma. These type 2 dots are used to perform the stratified area estimation part of P1. The accuracy of the segment-level proportion estimate is critically dependent upon the labeling accuracy of these type 2 pixels.

Table 6-10 shows at-harvest total omission (O) and commission (C) error rates (as a percentage of total pixels labeled in a state) for each state and the omission and commission error rates for the three major error sources identified in Phase III. An omission error is the result of mislabeling small-grain pixels as nonsmall grains, whereas a commission error is the result of mislabeling nonsmall-grain pixels as small grains. The omission error is shown to be consistently larger than the commission error by state and by error source (table 6-10), typically leading to underestimation of the small-grain proportion in a segment, as was found in the blind site analyses of proportion estimation error described in section 6.2.

TABLE 6-10.– PHASE III LABELING ERROR CAUSES

State	N. Dakota		Minnesota		Montana[a]		Colorado		Oklahoma	
Number of blind sites	18		6		10		6		11	
Cause of error	O	C	O	C	O	C	O	C	O	C
Abnormal signatures	4.4	0.5	2.6	0.3	1.4	0.9	2.8	-	3.3	1.4
Boundaries	3.2	0.7	4.0	1.1	1.0	0.6	2.3	0.8	2.2	0.8
Inadequate acquisitions	1.5	1.0	-	-	0.5	-	-	-	3.0	-
Other	2.1	0.8	2.5	1.2	1.9	0.6	0.9	-	1.4	3.3
Total errors	11.2	3.0	9.1	2.6	4.8	2.1	6.0	0.8	9.9	5.5

[a] Nonresolvable small-grain strip-fallow pixels are excluded.

The three major sources of labeling error were found to be abnormal signatures, boundaries, and inadequate acquisitions, as detailed in the following.

a. Abnormal signatures are those in which the crop signature does not follow the expected temporal sequence under the conditions believed by the analyst to be occurring in the segment.

b. Boundaries consist of border and edge pixels; the signature of a border pixel is spectrally mixed, representing both small-grain and nonsmall-grain areas. The signature of an edge pixel, on the other hand, is spatially mixed; and on the acquisitions used by the analyst for proportion estimation, the edge pixel moves at least once from a small-grain field to a nonsmall-grain field because of misregistration.

c. Inadequate acquisitions are those labeling errors that occur when the analyst attempts to label a segment for which key acquisitions are missing, usually guessing for many of the pixels when an estimate should not be made. This particular error occurred in only one or two blind sites per state; but when it did occur, both the labeling error and the proportion estimation error were large. For example, the 3-percent omission error caused by inadequate acquisitions for the 11 blind sites in Oklahoma came from one segment; this particular segment accounts for one of the two extreme underestimates in Oklahoma (section 4.2.4). The other outlier in Oklahoma, not included in this study, had the same acquisition history.

Miscellaneous labeling errors in the "other" category of table 6-10 include clerical and inconsistent labeling errors. The latter occurs when an analyst has labeled several pixels correctly and then incorrectly labels one or two pixels following the same temporal sequence in the same segment.

Note in table 6-10 that the nonresolvable small-grain strip-fallow pixels were excluded from the study of Montana. For these pixels, the multispectral scanner's resolution is not fine enough to show the strips in the imagery acquired by the Land Satellite (Landsat). The signature is integrated for the whole field and therefore cannot be called a boundary-type signature. However, because either a small-grain or a nonsmall-grain label could be considered correct for these areas, they were omitted. In Montana, 10.3 percent of the labeled pixels fell in the nonresolvable strip-fallow areas. Of these, the analyst labeled 54 percent nonsmall grains and 46 percent small grains, in good agreement with the expected 50 percent of small grains in these areas.

About 11 percent of the pixels labeled in Montana fell in the resolvable strip-fallow areas. The relatively low error rate for boundaries (1.0 percent omission, 0.6 percent commission) indicates that the analyst labeled quite accurately in these areas. Overall, the Montana small-grain signatures were found to be quite good, with very few abnormal signatures and good separation of the small-grain and nonsmall-grain signatures. In the proportion estimation error study, neither the winter wheat nor the spring wheat blind site analyses indicated a bias for the Montana proportion estimates.

Excluding the outlier for Oklahoma, the largest total labeling errors in the study were for Minnesota and North Dakota. These errors were due primarily to omission errors for abnormal signatures and boundaries. In the spring-wheat proportion estimation error study, Minnesota and North Dakota were the only two states for which a negative bias was indicated; the large errors of omission apparently caused this low proportion estimate.

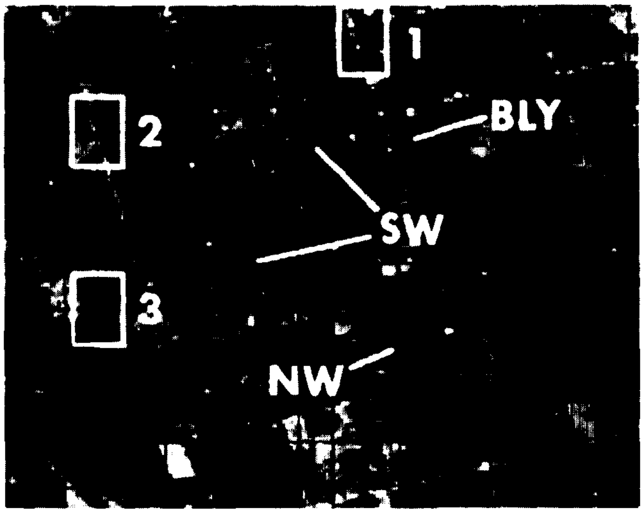
Figure 6-11 depicts the two largest sources of omission errors in Minnesota and North Dakota. The blind site depicted is located in Grant County, Minnesota. The pixels identified as 1, 2, and 3 are examples of a border pixel, an edge pixel, and an abnormal signature, respectively. (The upper left corner of the grid intersection is designated as the exact location of the pixel.)

Pixel 1 lies on the border between a spring-wheat field and a sunflower field. From the ground-truth map, it was determined that the pixel contained more spring wheat than sunflowers; however, the analyst labeled the pixel as nonsmall grains. The more accurate ground-truth determination is possible because the ground observations are made at a subpixel level one-sixth the size of a pixel. The evaluator thought that the analyst should have labeled the pixel as small grains since close inspection of the imagery revealed that the pixel was more red than green in the heading acquisition and more green than red in the turning acquisition.

Pixel 2 is a classic example of an edge pixel. In the heading acquisition, the pixel is on a road; in the turning acquisition, the pixel is in a spring wheat field. The analyst chose the turning acquisition as the base acquisition for this segment and registered the grid intersections of the other acquisitions to the base acquisition for labeling. In this example, pixel 2 was labeled as nonsmall grains, but it should have been labeled as small grains in agreement with the base acquisition. This may have been a clerical error.

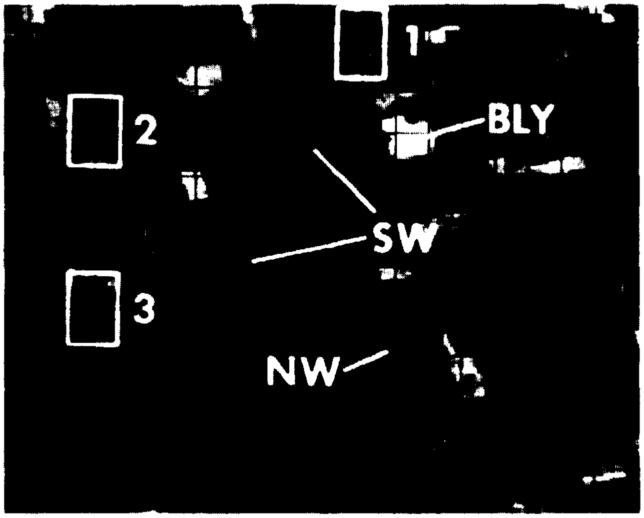
Pixel 3 is an example of an abnormal signature. The pixel is green in the heading acquisition and red in the turning acquisition; however, this pixel lies on the edge of a small body of water, and ground truth indicated that

NASA-S-78-16398



Heading - June 23, 1977

SW - red
NW - green



Turning - July 29, 1977

SW - green
NW - red

SW = Spring wheat
NW = Nonwheat
BLY = Barley

1 = Border pixel — spectral confusion of spring wheat and sunflowers

2 = Edge pixel — shifts from road (during the heading of spring wheat) to spring wheat (turning)

3 = Abnormal signature — excess water retarded spring wheat development

Figure 6-11.— Phase III omission labeling error examples the wheat field extended to the edge of the water. The analyst labeled the pixel as nonsmall grains. The evaluator, who thought that the analyst believed the pixel to be grass growing on the edge of the water, determined that the pixel was actually spring wheat, as indicated by the ground truth. The development of the spring wheat in this pixel was delayed because of excess moisture and was still in the heading stage, although most of the wheat in the segment was in the turning stage.

6.5 COMPARISON OF 400 DOT-COUNT GROUND TRUTH (400 TO 500 DOTS) AND DIGITIZED GROUND-TRUTH WHEAT PROPORTIONS

This section contains a discussion of the dot-count ground truth and digitized ground-truth wheat proportions for the spring and winter wheat segments in the USGP, which are compared to their respective digitized ground-truth wheat proportions at the state and regional levels.

The dot-count ground truth is obtained from the labeling of 400 to 500 dots corresponding to grid intersections overlaying aerial photographs. Digitized ground truth is a result of the machine processing of complete ground-truth information for each blind site. The dot-count ground-truth proportion is available several months earlier than the digitized ground-truth proportion, thus providing a more timely use for the blind site investigation of proportion estimation error. The comparison of 400 dot-count ground truth and digitized ground-truth wheat proportions provides an assessment of the reliability of the 400 dot-count ground truth.

Table 6-11 contains the statistical results for comparing the average dot-count ground-truth wheat proportions to the average digitized ground-truth wheat proportions for the nine states in the USGP. The following factors are noted:

a. The average dot-count ground-truth wheat proportion, $\overline{X}'$.

b. The average digitized ground-truth wheat proportion, $\overline{X}$.

c. The average difference, $\overline{D}' = \overline{X}' - \overline{X}$.

TABLE 6-11.— COMPARISON OF DOT-COUNT AND DIGITIZED GROUND-TRUTH WHEAT PROPORTIONS[a]

Region	n/M	$\bar{X}$	$\bar{X}'$	$\bar{D}'$	$S_{\bar{D}'}$	90% confidence limit for population $\bar{D}$
Winter wheat						
Colorado	13/31	19.52	19.52	0.0	0.27	(-0.49, 0.49) N
Kansas	15/121	28.85	27.65	-1.19	0.73	(-2.48, 0.10) N
Montana	18/58	12.96	12.38	-0.58	0.26	(-1.03, -0.13) S
Nebraska	18/56	17.68	16.31	-1.37	0.92	(-2.97, 0.23) N
Oklahoma	17/46	37.00	36.52	-0.48	0.70	(-1.69, 0.73) N
South Dakota	8/21	2.38	3.35	0.97	0.43	(0.15, 1.79) S
Texas	12/35	20.31	20.67	0.35	0.72	(-0.94, 1.64) N
USGP-7	101/368	21.09	20.60	-0.48	0.24	(-0.97, 0.19) N
Spring wheat						
Minnesota	13/47	19.55	21.25	1.69	0.60	(0.62, 2.76) S
Montana	13/48	10.88	11.31	0.43	0.43	(-3.24, 1.18) N
North Dakota	22/103	25.66	25.31	-0.35	0.60	(-1.38, 0.68) N
South Dakota	14/37	8.14	8.58	0.44	0.57	(-0.57, 1.45) N
USNGP	62/268	17.32	17.74	0.42	0.31	(-0.19, 1.23) N
USGP	163/557	19.65	19.51	-0.14	0.19	(-0.51, 0.35) N

[a] Symbols:

S = Significantly different from zero at the 10-percent level.

N = Not significantly different from zero at the 10-percent level.

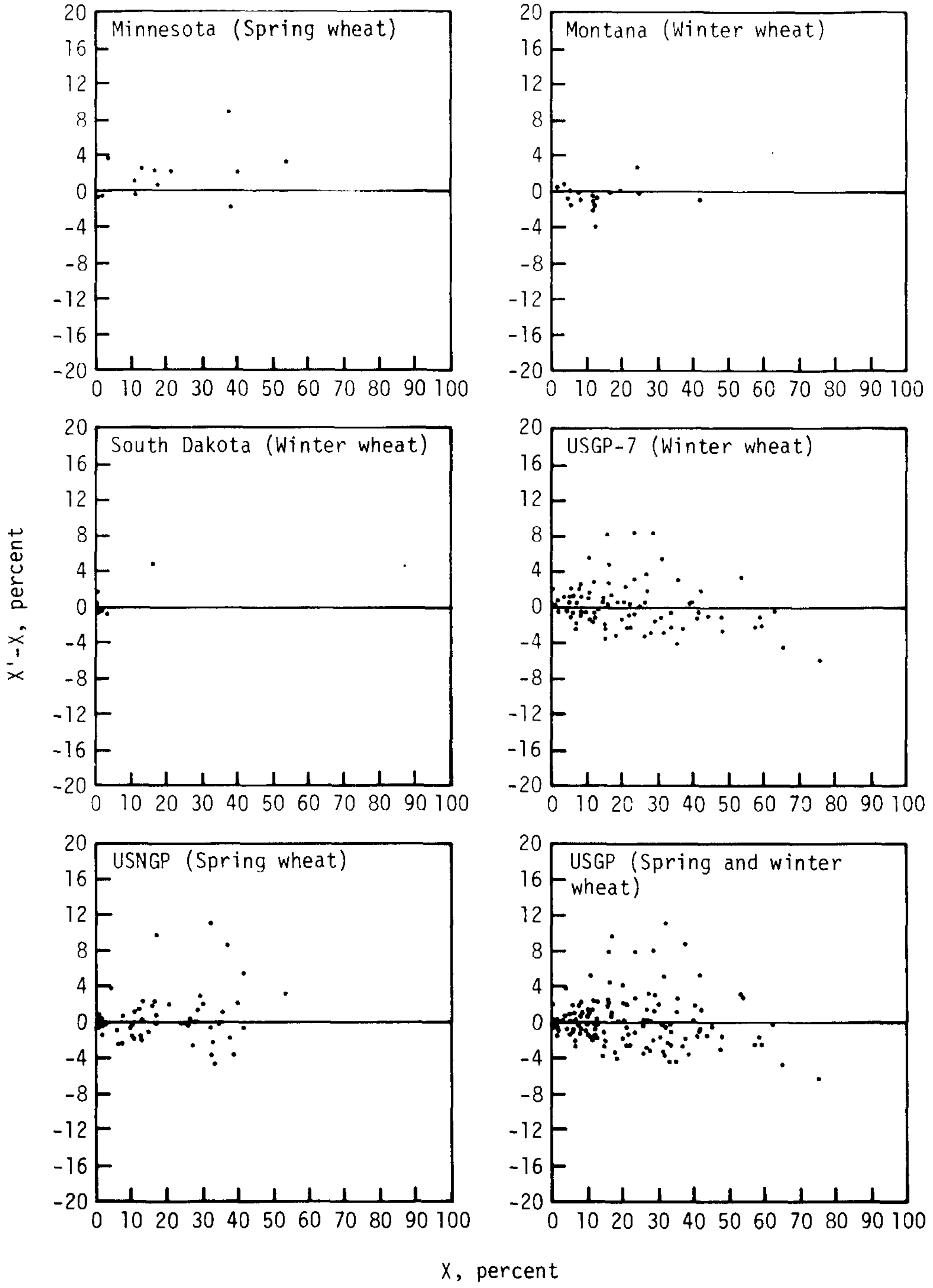
d. The standard error the average difference, $S_{\bar{D}'}$.

e. The 90-percent confidence limits for the population average difference, $\bar{D}$.

In the winter wheat region, the difference between the average dot-count ground-truth and digitized ground-truth wheat proportion was not significantly different from zero at the 10-percent level for Colorado, Kansas, Nebraska, Oklahoma, and Texas. Similarly, in the spring wheat region, the difference was not significant for Montana, North Dakota, and South Dakota. However, there was a significant difference between the average wheat proportions for Montana and South Dakota in the winter wheat region and for Minnesota in the spring wheat region (table 6-11).

Figure 6-12 shows plots of the difference between the dot-count ground-truth and digitized ground-truth wheat proportions $(X' - X)$ versus X, where X' is the dot-count ground-truth wheat proportion and X is the digitized ground-truth wheat proportion. Plots for Minnesota spring wheat, Montana and South Dakota winter wheat, combined spring wheat states, and combined winter wheat states are included. A point lying below the horizontal line $X' - X = 0$ corresponds to a segment for which the dot-count ground-truth wheat proportion is less than the digitized ground-truth wheat proportion. From the plots in figure 6-12, it is generally apparent that the differences between the dot-count ground-truth and the digitized ground-truth wheat proportions are negligible except for those caused by a few segments in Minnesota, Montana, and South Dakota. Furthermore, the t-test at the 10-percent level indicated that the average difference for the USGP is not significantly different from zero.

Figure 6-13 shows the scatter plot of the 400 dot-count ground-truth wheat proportions versus the corresponding digitized ground-truth proportions for 163 blind sites in the USGP. There was a high positive correlation ($r = 0.98292$), which was significantly different from zero at the 1-percent level, between the 400 dot-count ground-truth wheat proportions and the digitized ground-truth proportion.



LEGEND:

X' = Dot-count ground-truth wheat proportion.

X = Digitized ground-truth wheat proportion.

Figure 6-12.— Comparison of dot-count ground-truth and digitized ground-truth wheat proportions.

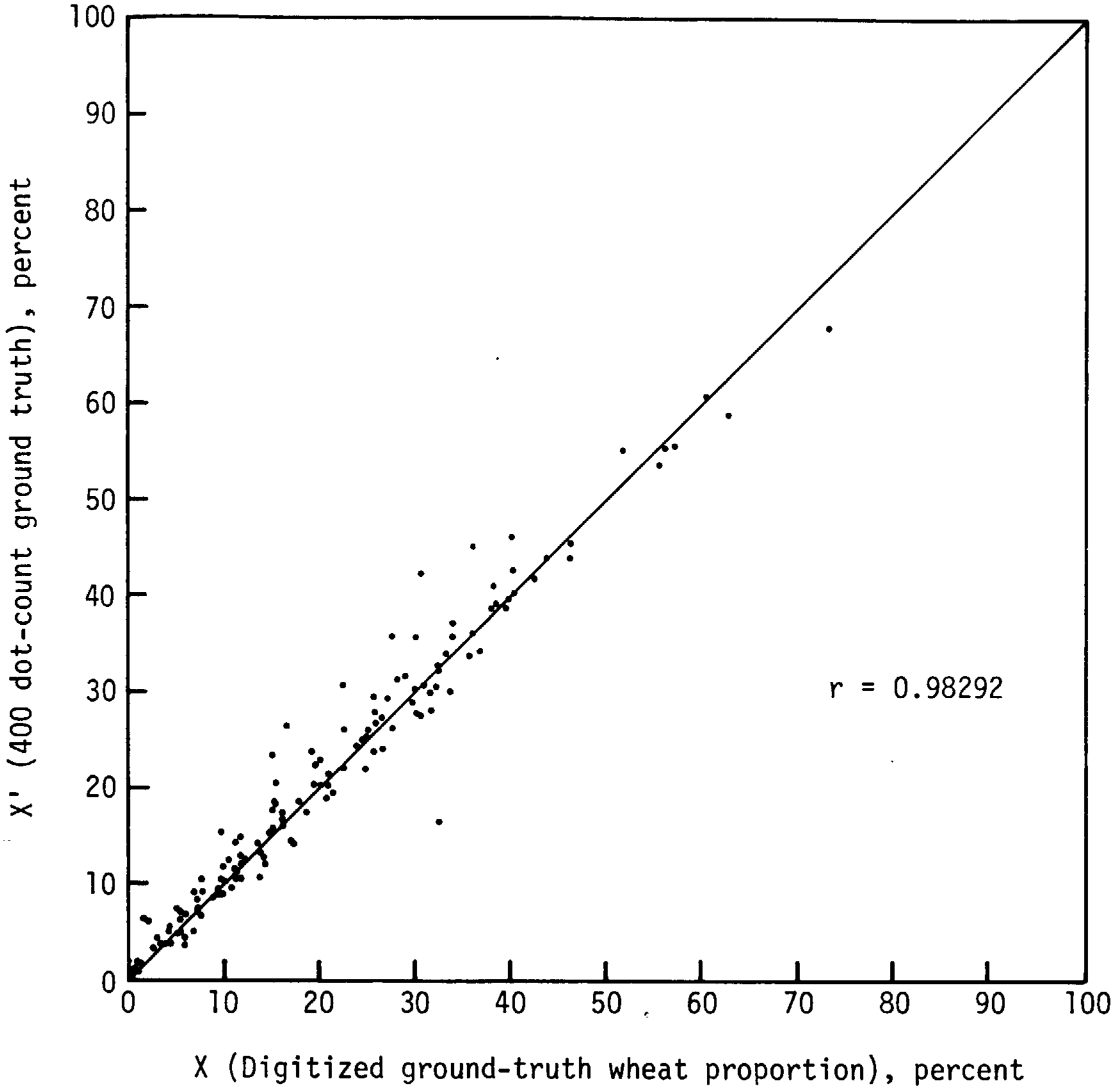


Figure 6-13.— Scatter plot of the 400 dot-count ground-truth wheat proportions versus the corresponding digitized ground truth in the USGP.

6.6 EFFECT OF VARIOUS VARIABLES ON SMALL-GRAIN PROPORTION ERRORS

This section reports the relationship between four variables and the small-grain proportion estimation error for segments in the USGP. The small-grain proportion estimation error ($\hat{X} - X$) is the difference between the LACIE small-grain proportion estimate ($\hat{X}$) and the digitized ground-truth small-grain proportion (X). The variables considered in this section as possibly being related to the small-grain proportion estimation error for the segments are total small-grain proportion, pasture proportion, corn proportion, and pure pixel proportion, where a pure pixel is either entirely grain or nongrain. There are 122 observations for each of the variables, except for corn proportion which has 77 data points corresponding to those blind sites with fields planted to corn.

The plots in figure 6-14 show the small-grain proportion estimation errors plotted as a function of each variable for segments in the USGP. The small-grain proportions used in this study were obtained from the final CAS data base, which generates the final LACIE estimates. A point lying below the horizontal line $\hat{X} - X = 0$ corresponds to a segment for which the digitized ground-truth small-grain proportions are less than the LACIE small-grain proportion estimate. A linear regression of each variable on the small-grain proportion error and the corresponding t-test of significance for the slope of the line describe the effect of the variable on small-grain proportion estimation error.

The plot of $\hat{X} - X$ versus X shows that $\hat{X}$ tends to underestimate larger values of X by a large margin. This compares directly with the blind site studies using the dot-count ground-truth proportions.

The correlation coefficients on the small-grain proportion errors with the pasture proportion ($r = 0.12547$) and with the corn proportion ($r = 0.18414$) are not significantly different from zero. No relationship is indicated between pasture or corn proportion and small-grain proportion estimation error. The proportion of pasture or of corn does not affect small-grain proportion estimation error, as indicated by nonsignificant regression

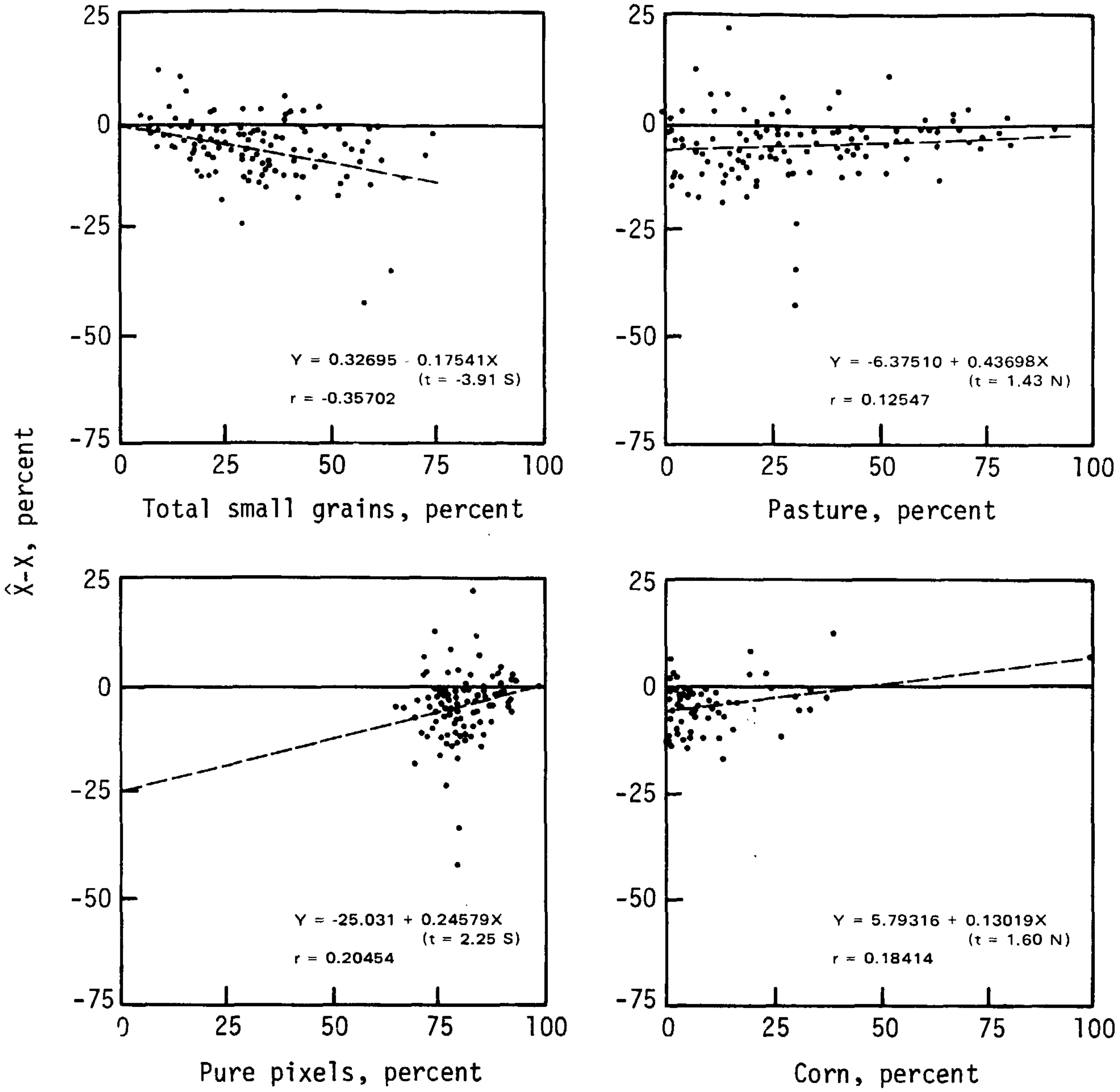


Figure 6-14.– Various variables versus total small-grain proportion estimation errors in all states in the USGP combined.

results. The line of best fit for each variable indicated consistent underestimation of the LACIE estimate for both corn and pasture proportions.

The correlation coefficients of the small-grain proportion errors with the pure pixel proportion (r = 0.20454) and with the total small-grain proportion (r = -0.35702) are significantly different from zero at the 5-percent level. Although there was underestimation of the LACIE estimate, there was a significant positive regression (slope = 0.24579) of the pure pixel proportion on the small-grain proportion estimation error indicating that the purer the pixel, the closer the regression line is to the zero line (i.e., the higher the pure pixel proportion, the lesser is the small-grain proportion estimation error). However, there was a significant negative regression (slope = -0.17541) of the total small-grain proportion on the small-grain proportion estimation error indicating that the higher the total small-grain proportion, the higher is the tendency to underestimate the small-grain proportion.

6.7 COMPARISON OF UNCORRECTED, BIAS-CORRECTED, REWORKED, AND RANDOM SAMPLING PROPORTION ESTIMATES

This section presents an evaluation of the proportion estimation procedure using 122 digitized and 8 dot-count ground-truth proportions in the USGP.

Two types of small-grain proportion estimates are generated by CAMS, uncorrected (machine-classified) and bias-corrected estimates, both calculated on the same computer run. The results of this computer run are examined by the analyst who occasionally sees where borderline decisions in the type 2 dot labeling could be reworked to improve the bias correction factor. This latter step is usually a simple hand calculation and does not require another computer run. The bias-corrected proportion is passed to CAS as the official CAMS estimate unless it is superseded by the rework calculation. The random sampling proportion estimates for small grains are calculated by AA personnel as the ratio of the number of type 2 dots labeled as small grains to the total number of the type 2 dots. For each blind site, the four proportion estimates are acquired from the data used in the last acceptable classification run.

Let $\hat{X}_i$, $\hat{X}'_i$, $\hat{X}''_i$, and $\hat{X}'''_i$ $(i = 1, 2, \cdots n)$ be estimates of X_i for the *ith* blind site, where

X_i is the true small-grain proportions;

$\hat{X}_i$ is the estimate derived from the type 1 dots, machine-classified proportion;

$\hat{X}'_i$ is the estimate derived from type 2 dots, bias-corrected proportion;

$\hat{X}''_i$ is the estimate derived from rework of type 2 dots;

$\hat{X}'''_i$ is the random sampling estimate from type 2 dots; and

n is the number of available blind sites.

From these basic data, the proportion estimation errors are

$\hat{X}_i - X_i$ for machine-classified error,

$\hat{X}'_i - X_i$ for bias-corrected error,

$\hat{X}''_i - X_i$ for reworked error, and

$\hat{X}'''_i - X_i$ for random sampling error.

The 130 segments were distributed as 34 spring wheat states, 33 mixed wheat states, and 63 winter wheat states. The MSE and the SD of the errors for the segments in each state were computed as shown in table 6-12. The succession of estimates for the blind sites in the USGP being studied is given in table 6-12. The effect of the correction factor applied at each step is seen in the reduction in the mean error and the MSE. Except for the order of presentation in the tables, the results in table 6-12 can be compared with data presented in table 6-2 (section 6.3). Major differences between the tables can be due to different sample sizes.

In Montana and South Dakota mixed wheat states and Colorado and Texas winter wheat states, the mean error of the bias-corrected estimate is smaller in magnitude than that of the machine estimate. On the other hand, in North Dakota and Minnesota (spring wheat states) and Nebraska, Kansas, and Oklahoma (winter wheat states), the mean error is increased in magnitude with the bias correction to the machine estimate.

TABLE 6-12.– ERROR ANALYSIS OF MACHINE, BIAS-CORRECTED, AND REWORKED SMALL-GRAIN ESTIMATES

Small-grain estimate	Mean error, %	SD	MSE	Reduction	
				Mean error, %	MSE, %
Colorado (11 winter wheat segments)					
Machine	-5.65	6.95	75.89	–	–
Bias corrected	-4.75	6.83	64.93	15.93	14.44
Final CAMS	-3.78	5.80	44.89	33.10	40.85
Random sampling	-5.06	6.70	66.43	10.44	12.46
Nebraska (16 winter wheat segments)					
Machine	-3.99	9.32	97.44	–	–
Bias corrected	-5.62	6.11	66.58	-40.85	31.67
Final CAMS	-4.18	5.76	48.56	-4.76	50.16
Random sampling	-5.22	7.61	81.58	-30.83	16.28
Kansas (15 winter wheat segments)					
Machine	-4.03	8.33	81.05	–	–
Bias corrected	-5.91	9.86	125.68	-46.65	-55.06
Final CAMS	-4.36	7.56	72.43	-8.19	10.64
Random sampling	-7.64	10.47	160.68	-89.58	-98.25
Oklahoma (12 winter wheat segments)					
Machine	-4.30	12.48	161.13	–	–
Bias corrected	-4.44	14.01	199.50	-3.26	-23.81
Final CAMS	-3.36	13.34	174.43	21.86	-8.25
Random sampling	-3.95	14.75	215.08	8.14	-33.48
Texas (9 winter wheat segments)					
Machine	-1.56	9.05	75.17	–	–
Bias corrected	-1.11	11.69	122.66	28.85	-63.18
Final CAMS	-1.17	9.44	80.61	25.00	-7.24
Random sampling	0.76	9.37	78.56	148.72	-4.51
Winter wheat states (63 segments)					
Machine	-4.00	9.17	98.72	–	–
Bias corrected	-4.66	9.66	113.69	-16.50	-15.16
Final CAMS	-3.57	8.40	82.16	10.75	16.77
Random sampling	-4.59	10.16	122.76	-14.75	-24.35

TABLE 6-12.– Concluded.

Small-grain estimate	Mean error, %	SD	MSE	Reduction	
				Mean error, %	MSE, %
North Dakota (22 Spring wheat segments)					
Machine	-9.32	12.41	243.46	–	–
Bias corrected	-10.24	8.83	179.27	-9.87	26.36
Final CAMS	-8.09	8.81	139.40	13.20	42.74
Random sampling	-9.76	9.29	177.61	-4.72	27.05
Minnesota (12 spring wheat segments)					
Machine	-0.39	4.32	17.28	–	–
Bias corrected	-6.61	6.15	78.35	-1594.87	-353.41
Final CAMS	-4.29	4.74	38.97	-1000.00	-125.52
Random sampling	-4.81	4.92	89.90	-1133.33	-420.25
Spring wheat states (34 segments)					
Machine	-6.17	11.37	163.63	–	–
Bias corrected	-8.96	8.08	143.59	-45.22	12.25
Final CAMS	-6.75	7.76	103.45	-9.40	36.78
Random sampling	-8.01	9.22	146.65	-29.82	10.38
South Dakota (16 spring and winter wheat segments)					
Machine	-5.42	8.82	102.28	–	–
Bias corrected	-3.47	9.32	93.92	35.98	8.17
Final CAMS	-5.40	8.47	96.48	0.37	5.67
Random sampling	-3.20	10.11	106.04	40.96	-3.68
Montana (17 spring and winter wheat segments)					
Machine	-4.60	7.87	79.38	–	–
Bias corrected	-3.34	6.27	48.11	27.39	39.39
Final CAMS	-3.45	5.62	41.56	25.00	47.64
Random sampling	-1.45	6.07	36.78	68.48	53.66
Mixed spring-winter wheat states (33 segments)					
Machine	-5.00	8.22	90.49	–	–
Bias corrected	-3.40	7.77	70.08	32.00	22.55
Final CAMS	-4.40	7.10	68.19	12.00	24.64
Random sampling	-2.30	8.19	70.36	54.00	22.24
USGP (130 segments)					
Machine	-4.82	9.54	113.61	–	–
Bias corrected	-5.47	9.01	110.44	-13.49	5.56
Final CAMS	-4.62	7.97	84.31	4.15	16.46
Random sampling	-4.90	9.61	115.71	-1.66	-0.63

An examination of the succession of the MSE in each state shows that the sequence improves the result in North Dakota, Montana, Colorado, and Nebraska. However, there was no improvement of results in Minnesota, South Dakota, Oklahoma, Kansas, and Texas, as indicated by smaller changes in values of the MSE's. Here, the errors of the estimates are maintained through the bias correction and rework.

It is important to look for an improvement trend in each step of succession, ending with the final CAMS estimate. When all the segments in the USGP are considered, the final CAMS estimates are the best of the four estimates, as seen in the mean error, the SD, and the MSE. This is not always true for the subgroups of states (spring wheat states, winter wheat states, and mixed wheat states) that comprise the USGP. This tends to support the contention that the sample sizes of the subgroups are not large enough to reflect improvement trends at each step of succession. The results tend to confirm that an analyst review of the bias-corrected results improves the final proportion estimate.

6.8 COMPARISON OF RATIOED AND DIRECT SPRING WHEAT ESTIMATES

For the CMR's of August, September, and October, CAMS personnel made two types of proportion estimates for the segments in North Dakota. First, as usual, CAMS estimated the spring small-grain proportion for each segment; these estimates were passed to CAS and ratioed down to spring-wheat proportions before aggregation. Second, CAMS estimated spring-wheat proportions directly for these same segments; these estimates were also aggregated by CAS.

The results of the two aggregations and the corresponding USDA/SRS estimates are shown in table 6-13. The CV's for the direct wheat estimates are slightly larger than those for the ratioed wheat estimates in all 3 months. However, the RD's for August and September are larger in absolute value for the ratioed wheat estimates. In October, the RD for the direct wheat estimate was larger. In August, both estimates were significantly different from the USDA/SRS estimate. In September, the direct wheat estimate was not significantly different from the USDA/SRS estimate; in October, the ratioed wheat estimate was not significantly different from the USDA/SRS estimate.

TABLE 6-13.— COMPARISON OF RATIOED AND DIRECT SPRING WHEAT (AGGREGATION) AREA ESTIMATES FOR NORTH DAKOTA[a]

Month of estimate	USDA/SRS area estimate, ac × 10^3	LACIE				RD, %		Value of test statistic	
		Estimate, ac × 10^3		CV, %					
		Ratioed wheat	Direct wheat	Ratioed wheat	Direct wheat	Ratioed wheat	Direct wheat	Ratioed wheat	Direct wheat
August	9530	6761	7 525	8.6	9.6	-41.0	-26.6	-4.8 S	-2.8 S
September	9530	8678	9 828	4.6	5.2	-9.8	3.0	-2.1 S	0.6 N
October	9530	9173	10 604	4.4	4.8	-3.9	10.1	-0.9 N	2.1 S

[a] Symbol definitions;

S = LACIE estimate is significantly different from USDA/SRS estimate at the 10-percent level.
N = LACIE estimate is not significantly different from USDA/SRS estimate at the 10-percent level.

A blind site study was performed using the ratioed and direct estimates for North Dakota from the October 11 CMR. Figure 6-15 shows plots of the proportion error ($\hat{X}$ - X) versus X, where $\hat{X}$ is the LACIE proportion estimate and X is the dot-count ground-truth proportion. Table 6-14 shows that the average proportion error was smaller for the direct estimates than for the ratioed estimates. Like the aggregation study, the blind site study indicated a higher degree of variability in the direct wheat estimates, as evidenced by the plots shown in figure 6-15.

In both studies, the October results reveal underestimation in the ratioed wheat estimate and overestimation in the direct wheat estimate.

TABLE 6-14.— COMPARISON OF RATIOED AND DIRECT SPRING WHEAT BLIND SITE PROPORTION ESTIMATES FOR NORTH DAKOTA[a]

[October 11 CMR estimates, percentages]

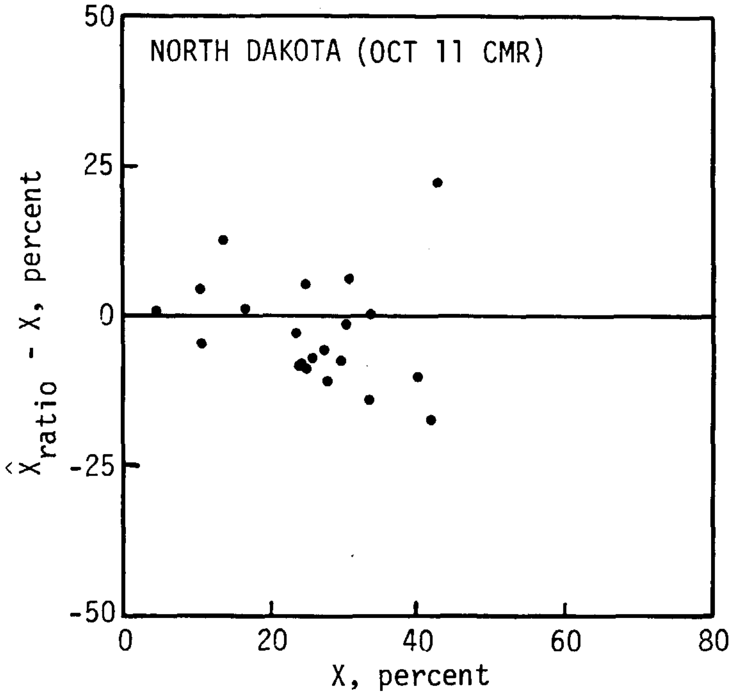
Estimate	n/M	$\bar{\hat{X}}$	$\bar{X}$	$\bar{D}$	$S_{\bar{D}}$	90% confidence interval for $\bar{D}$
Ratioed	20/103	21.0	25.1	-4.1	1.5	(-6.7, -1.5) S
Direct	20/103	25.6	25.1	0.5	1.8	(-2.6, 3.6) N

[a] Symbols:

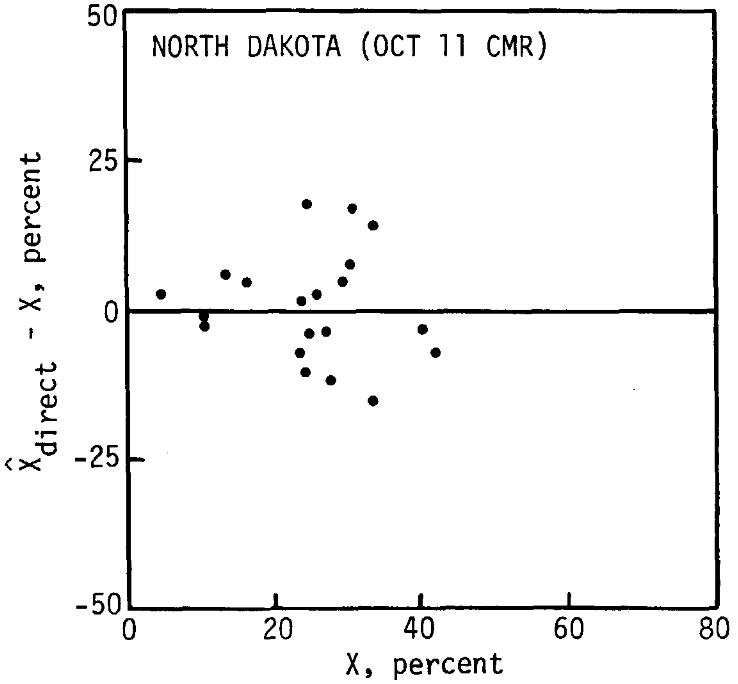
S = Significantly different from 0 at the 10-percent level
N = Not significantly different from 0 at the 10-percent level

6.9 REQUIREMENTS FOR REGISTRATION OF GROUND-TRUTH IMAGES

In order for the ground-truth inventory data to be usefully transformed into digital form, the ground-truth field boundaries of the Bendix-100 (field) vertices) representation must be registered to the Landsat image representation for each segment. The registration is accomplished by a straightforward application of a mapping transformation to a set of common control points measured from each inventory and Landsat segment. The final registration must be as accurate as possible, particularly if the 209 dots are to be correctly identified in the ground-truth image.



(a) October ratioed spring wheat estimates



(b) October direct spring wheat estimates

Figure 6-15.— Plots of proportion estimation errors versus dot-count ground-truth proportion estimates for blind sites in North Dakota.

The importance of good registration in either the small-grain class or "other" was demonstrated by obtaining seven different registrations for two Phase III blind sites and comparing the 209 dot labels obtained. The two segments, 1048 and 1602, are described in table 6-15. Using the Bendix-100 system, each segment was first processed using the nominal transformation procedure. Then, six precise offsets of 1/2, 1, and 1-1/2 pixels along the X (line) and y (pixel) axes were applied to the data to produce six additional Bendix-100 products. The program SPATL was used to label the 209 dots for all 14 products.

The labeled dots were classified as either small grains or nonsmall grains. Comparing the labels for each offset product to those for the reference product, the analyst computed the percentage change in the dot labels and conditional probability matrices for mislabeling for both segments. The results are given in table 6-15.

Note that a shift of only 1/2 pixel yields a change in 4 to 8 percent of the labels (8 to 17 mislabeled dots). A few USNGP segments have many more fields (as many as 675) than the 428 delineated for segment 1602. For these segments, a 1/2-pixel shift would be expected to yield as many as 29 mislabeled dots (14 percent based on a linear relationship). Table 6-15 contains the off-diagonal elements of the probability matrix,

$$P(l,l/l) = \frac{N(l,k)}{N_0(l)} \quad ; \quad l \neq k$$

where $N(l,k)$ is the number of dots labeled as class l on the reference product that are labeled as class k on the offset product. The diagonal pair of matrix elements

$$P(l,l/l) = \frac{N(l,l)}{N_0(l)}$$

may be obtained from the relation

$$P(l,l/l) + P(l,k/l) = 1$$

The implications of this study are clear. The inventories must be registered to Landsat better than ±1/2 pixel root mean square to accurately assess

TABLE 6-15.— CHANGE STATISTICS FOR SIX CASES OF MISREGISTRATION IN 209 DOT NAMES

Segment	Change in 209 dot names, % (a)						Conditional probability matrix (a)						Element P(1,k/1), % (b)					
							1/2x		x		1-1/2x		1/2y		y		1-1/2y	
	1/2x	x	1-1/2x	1/2y	y	1-1/2y	ϕ/SG	SG/ϕ	ϕ/SG	SG/ϕ	ϕ/SG	SG/ϕ	ϕ/SG	SG/ϕ	ϕ/SG	SG/ϕ	ϕ/SG	SG/ϕ
1048	4.3	7.2	9.7	3.8	8.6	14.4	3.6	4.8	9.6	5.6	10.8	8.7	4.8	3.2	12.0	6.3	18.1	11.9
1602	7.2	13.9	17.7	8.1	13.4	19.6	8.6	6.2	16.0	12.5	23.5	14.1	9.9	7.0	14.8	12.5	24.7	16.4

[a] The following legend applies to this column:
1/2x = 1/2 pixel misregistration.
x = line axis.
1/2y = 1/2 line misregistration.
y = pixel axis.
ϕ = crops other than small grains.
SG = small grains.

[b] The element P(1,k/1) equals the number of pixels labeled as class k on the offset scene which were labeled as class 1 on the reference scene divided by the total number of pixels labeled as class 1 on the reference scene; i.e., the proportion of mislabeled 1-pixels.

analyst dot-labeling error for the small grains versus nonsmall-grain representation. At present, a LACIE production film converter (PFC) is used as the reference scene to register the ground-truth inventories. This product is the primary limitation on the accuracy of the registration process. A careful study of the registration of segment 1602 and an analysis of the registration procedure indicates that the current registration capability is ±1/2 pixel.

6.10 REGISTRATION OF GROUND-TRUTH IMAGES

Until recently, registration was difficult to verify experimentally. Two techniques have now been developed to verify that registration of the ground-truth maps meets the requirements indicated by section 6.9.

In the first method developed, the ground-truth map and three channels (1, 2, and 4) of the LACIE image are read into the refresh memory of the General Electric Interactive Multispectral Image Analysis System (I-100). Using existing programs and procedures, the analyst transfers selected fields from the ground-truth map to the I-100 themes, where the field patterns can be conveniently superimposed upon the LACIE segment to permit precise verification of the registration of the ground-truth maps to the LACIE segment. This method is more precise but more time consuming.

The second method used to verify the registration of the ground-truth maps was developed concurrently with the verification of the registration of the Phase III blind sites with the objectives of reducing the time required for verification and producing a permanent record of the registration. In this method, a registration plot which contains all the field boundaries in the ground-truth map is generated on the Gerber plotter, which is registered to the photographic image of the LACIE segment. The registration plot can then be superimposed upon the photographic image and the registration verified.

Early in the growing season, Goddard Space Flight Center (GSFC) specifies for all LACIE segments a LACIE acquisition of a Landsat image, which will be the acquisition designated as the reference image for registration of the other

LACIE acquisitions. However, in the 1977 AA procedures for developing ground-truth maps, the LACIE acquisition which was used for the "last" successful classification was the acquisition designated by AA personnel for registration of the ground-truth map. Because some of the blind sites had already been registered to other acquisitions before the designation of the "last" successful classification and because registration by GSFC between acquisitions was not considered to be a problem, the blind sites were not reregistered. The registration of the ground-truth map was always checked against the AA designated acquisition. Nine of the blind sites which were registered to these other acquisitions rather than the AA designated acquisition were not satisfactorily registered when compared to the AA designated acquisition because of misregistration between the two LACIE images. These nine blind sites are marked with an asterisk in table 6-16. For pixel-level studies, use of these nine sites should acknowledge the alternate acquisition, not the AA designated one.

Satisfactory digitized ground-truth maps were prepared for 146 Phase III blind sites. The I-100 was used to check the registration of all of these sites with the requirement that the registration be accurate to within one pixel throughout the LACIE segment. Of these 146 ground-truth maps, 124 met the one-pixel criterion. These 124 blind sites are considered to be registered to the LACIE segments and therefore to the LACIE classification maps with sufficient accuracy to permit the analysis of analyst dot labeling errors and classification errors. The 22 blind sites which should not be used for pixel level studies are indicated by a 1 in the first column of appendix C.

6.11 COMPARISON OF DELTA AND P1 ESTIMATORS

The purpose of this section is to display the results of an analysis of the delta ($\hat{X}_\Delta$) estimator as an at-harvest winter small-grain proportion estimator for the Phase III blind sites in the USSGP five-state region. Results are also given for the at-harvest LACIE segment proportion estimate, $\hat{X}$, of winter small grains obtained using P1 for the same set of blind sites.

TABLE 6-16.— PHASE III BLIND SITES WITH REGISTERED GROUND TRUTH

Segment	YYDDD†	State	Segment	YYDDD	State
1009	77178	COLORADO	1378	77178	NEBRASKA
1005	77177	COLORADO	1398	77209	NEBRASKA
100[illegible]	77195	COLORADO	1450	77191	NEBRASKA
10[illegible]2	77194	KANSAS	1479	77157	NEBRASKA
1033	76325*	KANSAS	1489	77176	SOUTH DAKOTA
1059	77157	TEXAS	1498	77210	SOUTH DAKOTA
1060	77158	TEXAS	1502	77197	COLORADO
1079	77157	TEXAS	1506	77178	COLORADO
1094	77181	COLORADO	1507	77196	COLORADO
1099	77122*	COLORADO	1512	77193	MINNESOTA
1103	77200	MONTANA	1513	77175	MINNESOTA
1104	77200	MONTANA	1515	77193	MINNESOTA
1158	77155	KANSAS	1520	77174	MINNESOTA
1166	77190	KANSAS	1521	77210	MINNESOTA
1175	77191*	KANSAS	1522	77210	MINNESOTA
1180	77153	KANSAS	1523	77175*	MINNESOTA
1183	77154	KANSAS	1529	77220	MONTANA
1219	77156	OKLAHOMA	1531	77220	MONTANA
1220	77156	OKLAHOMA	1532	77236	MONTANA
1222	77083	OKLAHOMA	1537	77127	MONTANA
1223	77066	OKLAHOMA	1544	77198	MONTANA
1228	77155	OKLAHOMA	1560	77197	NEBRASKA
1231	77156	OKLAHOMA	1564	77213	NEBRASKA
1233	77066	OKLAHOMA	1566	77214	NEBRASKA
1236	77101	OKLAHOMA	1568	77178	NEBRASKA
1239	77155	OKLAHOMA	1576	77155	NEBRASKA
1263	77156	TEXAS	1577	77120	NEBRASKA
1266	77156	TEXAS	1586	77194	NEBRASKA
1272	77154	TEXAS	1588	77193	NEBRASKA
1275	77171	TEXAS	1592	77156	NEBRASKA
1279	77194	KANSAS	1594	77155	NEBRASKA
1290	77067*	KANSAS	1595	77192	NEBRASKA
1293	77067*	KANSAS	1604	77143	NORTH DAKOTA
1325	77155	TEXAS	1606	77197	NORTH DAKOTA
1355	77156	OKLAHOMA	1616	77159	NORTH DAKOTA
1362	77155	OKLAHOMA	1619	77175	NORTH DAKOTA
1365	77155	OKLAHOMA	1625	77233	NORTH DAKOTA
1367	77155	OKLAHOMA	1635	77159	NORTH DAKOTA
1371	77050	TEXAS	1637	77194	NORTH DAKOTA

†YY = last two digits of year; DDD = day of the year.

*Referenced images different from AA referenced image.

TABLE 6-16.– Concluded.

Segment	YYDDD[†]	State	Segment	YYDDD	State
1640	77211	NORTH DAKOTA	1803	77178	SOUTH DAKOTA
1644	77140	NORTH DAKOTA	1805	77211	SOUTH DAKOTA
1648	77179	NORTH DAKOTA	1807	77211	SOUTH DAKOTA
1652	77197	NORTH DAKOTA	1811	77211	SOUTH DAKOTA
1661	77159	NORTH DAKOTA	1830	77211	MINNESOTA
1663	77175	NORTH DAKOTA	1835	77174	MINNESOTA
1669	77179	SOUTH DAKOTA	1839	77174	MINNESOTA
1675	77230	SOUTH DAKOTA	1849	77136	MINNESOTA
1681	77192	SOUTH DAKOTA	1850	76327	COLORADO
1686	77194	SOUTH DAKOTA	1853	77193	KANSAS
1694	77213	SOUTH DAKOTA	1859	76326	KANSAS
1699	77194	SOUTH DAKOTA	1864	76326*	KANSAS
1725	77170	MONTANA	1873	77192	MINNESOTA
1730	76281*	MONTANA	1894	77153	MINNESOTA
1739	77222	MONTANA	1897	77214	NORTH DAKOTA
1741	77203	MONTANA	1899	77193	NORTH DAKOTA
1742	77113	MONTANA	1903	77179*	NORTH DAKOTA
1747	77184	MONTANA	1913	77215	NORTH DAKOTA
1750	77221	MONTANA	1924	77176	NORTH DAKOTA
1752	77203	MONTANA	1927	77230	NORTH DAKOTA
1753	77184	MONTANA	1937	77203	MONTANA
1800	77210	SOUTH DAKOTA	1944	77199	MONTANA
1802	77211	SOUTH DAKOTA	1948	77184	MONTANA

[†] YY = last two digits of year; DDD = day of the year.

*Referenced images different from AA referenced image.

The winter small-grain proportion estimates obtained using the Δ-estimator were provided by the Mission Planning and Analysis Division (MPAD) of the Johnson Space Center (JSC). The Δ-estimator is a product of the Δ-classifier program, a basic tool used by MPAD for CAMS quality assessment (ref. 2). The MPAD indicated that of the nine proportion estimates provided, the one given by the "fixed point 41" estimator was preferred. The two estimators, $\hat{X}_\Delta$ and $\hat{X}$, were compared only for the blind sites that were common to both those listed in the CAR which were worked using the P1 estimator (LACIE) and the list of Δ-workable segments provided by the MPAD. This gave a total of 27 segments over which to compare the two estimates. The following table gives the total number of segments worked by LACIE and MPAD in the USGP in Phase III and the distribution of the 27 blind sites by state.

State	Total no. of sites allocated	Total sites worked by LACIE	Total sites worked by MPAD	No. of sites common to LACIE and MPAD
Colorado	31	24	21	8
Kansas	121	108	43	8
Nebraska	56	39	29	8
Oklahoma	46	41	0	0
Texas	35	29	8	3
Total	289	241	101	27

These results are for those segments considered to have the best acquisition pattern for obtaining a good Δ-estimator proportion estimate.

The ground-truth winter small-grain proportions, X_i, used in this analysis were obtained using the 400 dot-count procedure. The AA effort has demonstrated that the differences between the 400 dot-count proportions and the proportions obtained over a subset of the blind sites using the wall-to-wall inventory results are random sampling errors. These random errors reduce the power of any statistical tests performed but do not bias the results in favor of one estimator over the other.

Table 6-17 gives, by state, the blind site segment number, the ground-truth winter small-grain proportion, the P1 proportion estimate, the Δ-proportion estimate, and the corresponding differences between estimates and ground-truth used in this analysis. The following display gives the average difference and the estimated SD of the differences between each estimator and the corresponding ground truth. The t-statistic is also given for testing whether or not the true classification bias of each estimator is significantly different from zero.

Parameter	P1 estimator, %	Δ-estimator, %
Average difference	-3.43	4.90
SD	5.39	19.28
t-statistic	-3.31*	1.32

*This indicates that the P1 estimator has a bias significantly different from 0 at the 5-percent level. The large SD observed for the Δ-estimator makes it uncertain, for this sample size 27, whether the Δ-estimator is biased.

The 90-percent confidence limits for the true bias of P1 is (-5.14, -1.72) and (-1.20, 11.00) for the Δ-estimator. Obviously, the width of the confidence interval for the true bias of the P1 estimator (3.42 percent) is much smaller than that of the Δ-estimator (12.20 percent). This indicates the lack of accuracy in the Δ-estimator for a fixed confidence level. To obtain the 90-percent confidence limits for the true bias of the Δ-estimator with a width of that observed for the P1 estimator over 27 blind sites, the analyst would have to obtain ground truth and Δ-estimates for 344 sample segments.

Plots of the error versus the ground-truth proportion by segment for each estimator are given in figures 6-16 and 6-17. Comparison of the SD's observed in table 6-17, using the standard F-test, indicates that the standard error of the average difference for the Δ-estimator is significantly larger than that of the P1 estimator.

Rather than compare the average differences from ground-truth for the two estimators, the analyst employed a nonparametric test comparing the absolute

TABLE 6-17.— DATA USED IN DELTA-P1 COMPARISON

State	Segment number	Ground truth winter small-grain proportion (X)	P1 proportion estimate ($\hat{X}$)	Delta proportion estimate ($\hat{X}_\Delta$)	$\hat{X}-X$	$\hat{X}_\Delta-X$
Colorado	1501	11.0	14.0	14.2	3.0	3.2
	1000	41.9	46.0	53.2	4.1	11.3
	1502	15.6	13.0	29.0	-2.6	13.4
	1506	22.2	22.0	42.5	-0.2	20.3
	1507	10.2	3.3	12.1	-6.9	1.9
	1005	37.2	19.9	47.8	-17.3	10.6
	1091	10.5	15.0	63.1	4.5	52.6
	1099	25.9	19.4	37.7	-6.5	11.8
Kansas	1021	29.9	24.7	51.2	-5.2	21.3
	1032	38.7	40.5	65.8	1.8	27.1
	1170	63.0	63.0	69.3	0.0	6.3
	1175	43.9	40.5	23.1	-3.4	-20.8
	1340	56.8	45.8	35.2	-11.0	-21.6
	1279	30.4	29.0	44.4	-1.4	14.0
	1864	35.8	30.8	37.9	-5.0	2.1
	1885	54.3	45.0	25.2	-9.3	-29.1
Nebraska	1560	38.7	35.1	78.9	-3.6	40.2
	1566	32.6	24.0	57.1	-8.6	24.5
	1571	10.9	8.3	10.0	-2.6	-0.9
	1579	9.0	6.6	7.6	-2.4	-1.4
	1582	17.8	18.0	21.2	0.2	3.4
	1586	17.3	21.5	25.0	4.2	7.7
	1588	21.9	20.7	14.6	-1.2	-7.3
	1595	34.0	26.0	7.2	-8.0	-26.8
Texas	1263	43.1	32.3	33.2	-10.8	-9.9
	1272	21.0	14.0	10.4	-7.0	-10.6
	1275	28.4	31.0	17.5	2.6	-10.9

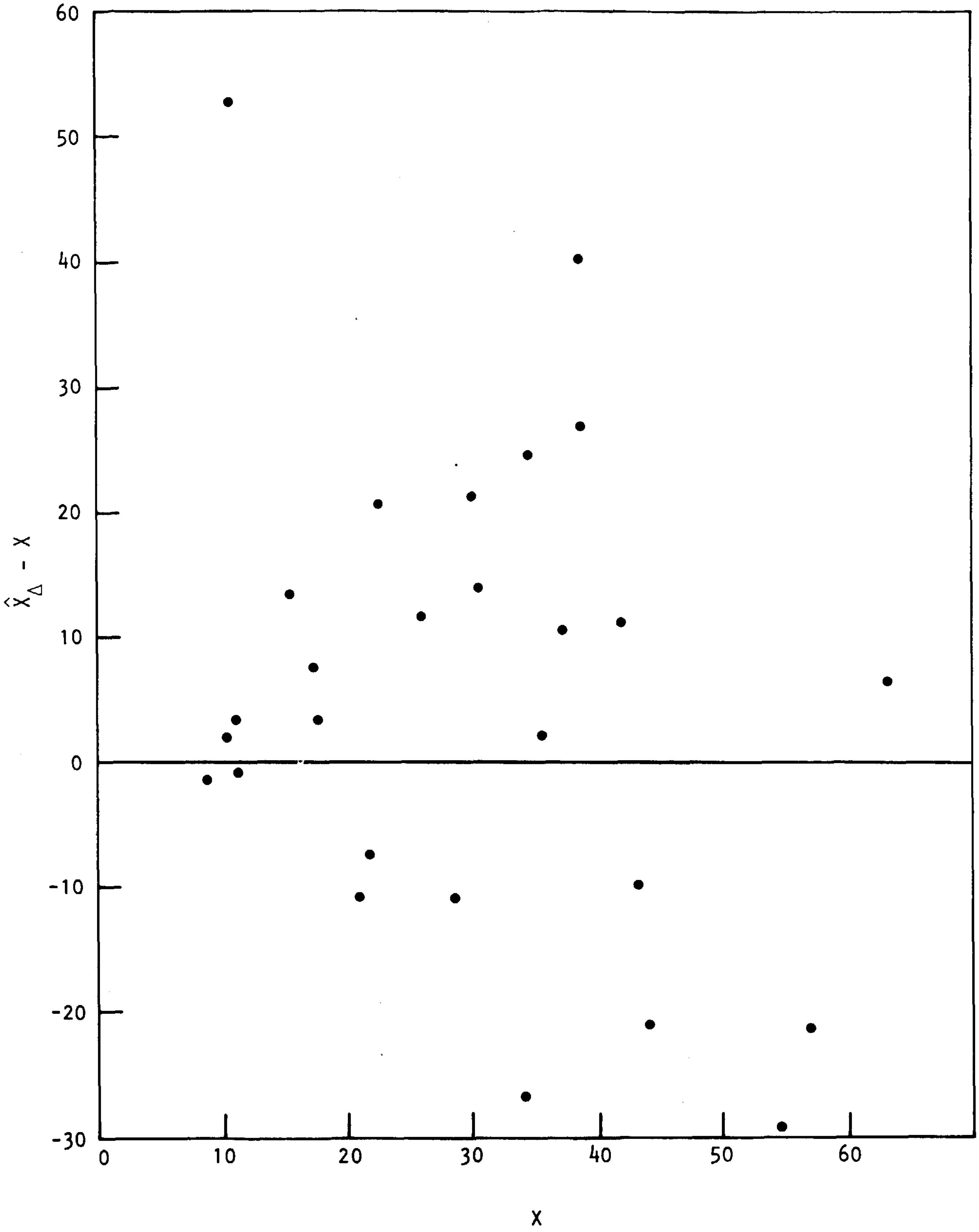


Figure 6-16.— Plot of the Δ-estimator winter small-grain proportion errors versus 400 dot-count ground-truth winter small-grain proportions.

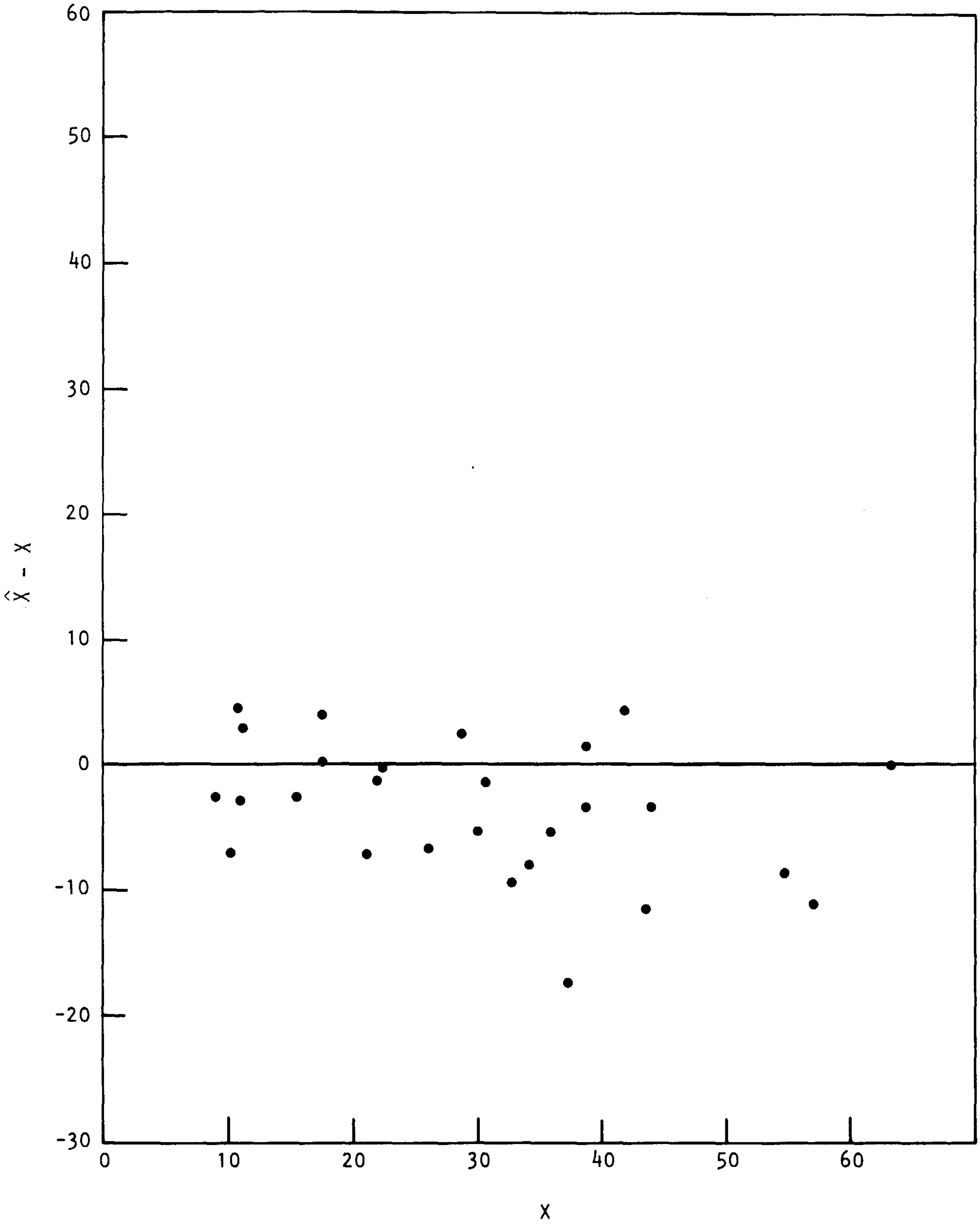


Figure 6-17.— Plot of the P1 estimator winter small-grain proportion errors versus 400 dot-count ground-truth winter small-grain proportions.

errors of the estimates. Comparison of absolute errors will indicate which, if either, is a better estimator of segment-level winter small-grain proportions.

Let $D_i = |\hat{X}_{\Delta i} - X_i| - |\hat{X}_i - X_i|$ for each blind site (i = 1, ···, 27). The D_i's, ranked in order of absolute value but with the sign of the difference indicated, are given below.

$$D = |\hat{X}_\Delta - X| - |\hat{X} - X|$$

Rank	D_i	Rank	D_i	Rank	D_i
1	0.2	10	5.3	19	15.9
2	-0.9	11	6.1	20	16.1
3	-1.0	12	6.3	21	17.4
4	-1.7	13	-6.7	22	18.8
5	-2.9	14	7.2	23	19.8
6	3.2	15	8.3	24	20.1
7	3.5	16	10.6	25	25.3
8	3.6	17	10.8	26	36.6
9	-5.0	18	12.6	27	48.1

Note that the P1 estimate had a larger absolute error than the Δ estimate in only six cases. The Wilcoxon signed-rank test statistic for paired samples is used to test the null hypothesis that the population of differences is symmetric with median zero. The test statistic, in this case, is the sum of the ranks with negative sign, which is 36. The lower critical point from tables for this test statistic is 84 for an α-level of 0.01. Since 36 is less than 84, the null hypothesis is rejected and the conclusion is reached that the absolute error at the segment level using the Δ-estimator is significantly larger than the absolute error using P1.

To determine the correlation between the estimates and ground-truth proportions and to determine whether bias could be eliminated using regression techniques, the ground-truth proportions were regressed on the corresponding estimates. Fitting the model $X = a + b\hat{X}_{\Delta} + \delta$ over the 27 data points yielded $r^2 = 0.21$, $\hat{a} = 18.28$, and $\hat{b} = 0.33$. A plot of X versus $\hat{X}_{\Delta}$ is given in figure 6-18 along with the fitted regression line and the one-to-one line. Fitting a similar model, using the P1 estimator, $X = c + d(\hat{X}) + \varepsilon$ over the 27 data points yielded $r^2 = 0.87$, $\hat{c} = 3.50$, and $\hat{d} = 1.00$. A plot of X versus $\hat{X}$ is given in figure 6-19. Obviously, the P1 estimator yields estimates of winter small grains that are much better correlated with ground-truth proportions of winter small grains than the Δ-estimator. Further, the high correlation between the P1 estimates and ground-truth and the estimated slope being 1.00 indicates that the fitted regression equation could be used to correct for the negative bias of the P1 estimator for the USSGP segment proportion estimates whose acquisition patterns are similar to those in the sample. Since the Δ-estimator has such large variability, linear regression is unsafe and inappropriate for attempting to remove the bias.

In conclusion, the previous analyses over 27 USSGP blind sites indicate that the Δ-estimator is not reliable as an estimator of winter small-grain proportions. On the other hand, for those segments having the acquisition pattern that was thought to be required for a successful Δ-classification, the P1 estimator provides impressive, though biased, results as an estimator of winter small-grain proportions. The regression analysis indicates that the bias could possibly be eliminated using the linear regression of ground-truth winter small-grain proportions on the P1 winter small-grain proportion estimates.

As stated previously in this section, the use of the 400 dot-count proportion estimates as ground truth lowers the power of the Wilcoxon signed-rank test. However, because the null hypothesis was rejected, the conclusion that P1 estimator had a significantly smaller absolute error than did the Δ-estimator is still valid because the inclusion of the random-error component does not bias the results in favor of one estimator over the other.

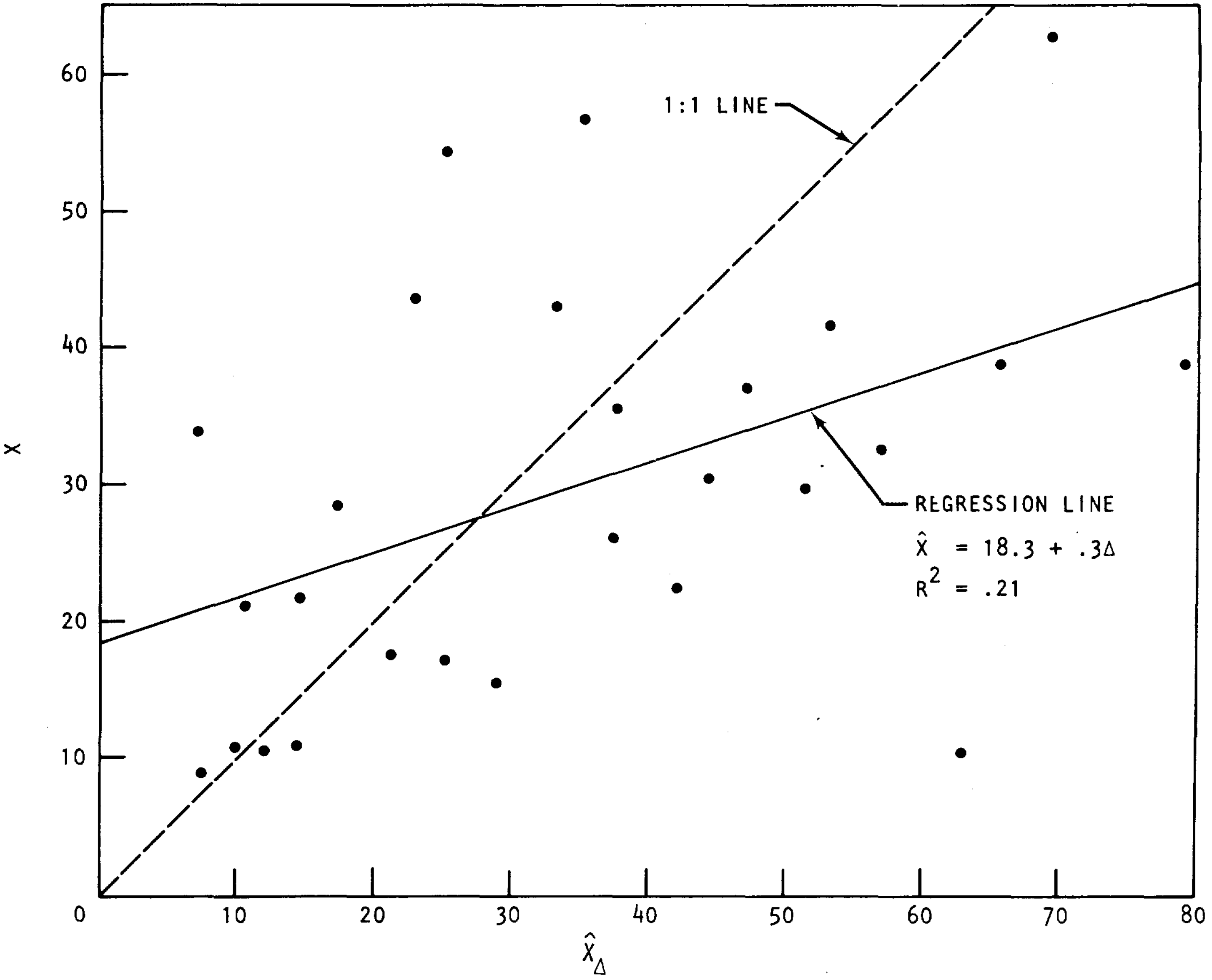


Figure 6-18.— Plot of 400 dot-count ground-truth winter small-grain proportions versus the Δ-estimator winter small-grain proportions.

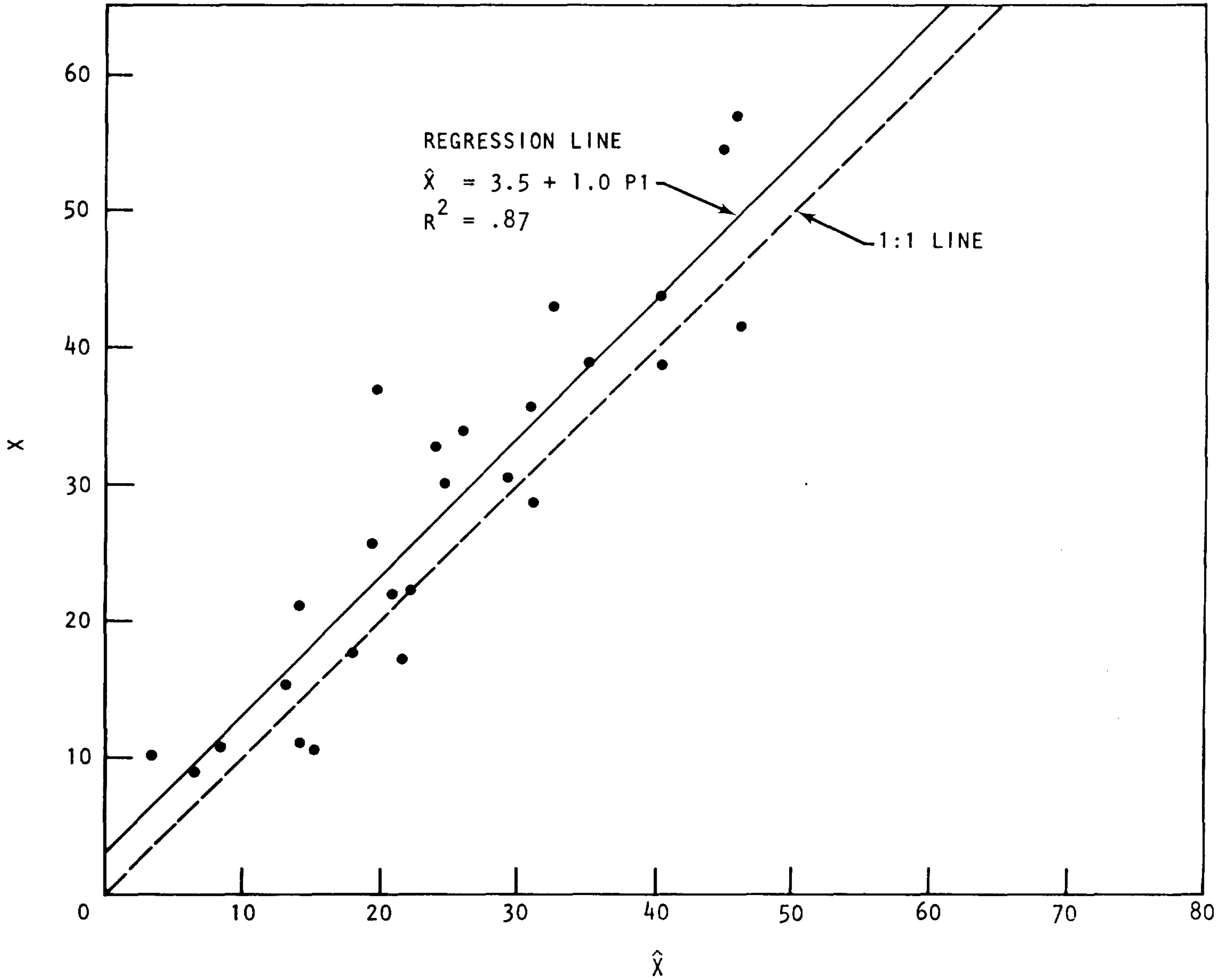


Figure 6-19.– Plot of 400 dot-count ground-truth winter small-grain proportions versus the P1 estimator winter small-grain proportions.

6.12 EFFECTS OF ANALYST, ACQUISITION HISTORY, AND BIAS CORRECTION ON PROPORTION ESTIMATION ERROR

The I-100 processor and data from eight U.S. blind sites were used in an experiment wherein each site was analyzed by three analysts to give a P1 "raw" and a bias-corrected estimate of the proportion of small grains in each segment. The segments were of two types, those having acquisitions in all four biophases and those having only early-season acquisitions. The segments were chosen at random from the blind sites for which detailed ground truth was available.

The objectives of the experiment were as follows:

a. To evaluate the performance of P1 in terms of absolute proportion estimation error and its repeatability over analyst-interpreters' products.

b. To make comparisons between bias-corrected and raw P1 estimates.

c. To determine whether the performance was better when acquisitions from all biostages were used than when only the early-season acquisition was used.

The third objective could not be achieved properly because only four of each type of segment was used. It was later estimated that to make effective comparisons of this type in a fully nested design, one would need about 10 times as many segments. The efficiency of the test could be improved if the same segments were analyzed first using only early-season acquisitions and then using all acquisitions; however, there would be potential biasing problems in such replication if the same analyst analyzed the segment under both the early-season and full-season conditions. If different analysts performed the analysis, the resulting large variability would reduce the power of the test.

Table 6-18 shows the absolute proportion estimation error $|\hat{X} - X|$, where X is the ground-truth small-grain proportion and $\hat{X}$ is the analyst's estimate of X for the various treatment combinations. Averages are blocked off from the basic data; for example, the average absolute error for analyst B on early-season segments was 11.6 for the raw estimate and 11.8 for the bias-corrected

TABLE 6-18.— I-100 P1 DATA

[$|\hat{X} - X|$ (small grains)]

Acquisition history	Segment	Raw				Bias correction				Overall average
		Analyst			Average	Analyst			Average	
		A	B	C		A	B	C		
Early season only	1642	16.5	10.8	2.0		18.9	8.7	16.8		
	1651	11.4	18.5	21.3		5.6	18.3	19.7		
	1660	9.7	14.6	30.3		8.0	11.9	19.5		
	1662	8.4	2.5	7.0		1.6	8.2	1.5		
Average		11.5	11.6	15.2	12.8	8.5	11.8	14.4	11.6	12.2
Full season	1603	0.8	1.4	0.9		1.4	1.4	2.0		
	1614	5.2	10.6	31.7		9.7	32.9	32.6		
	1637	1.3	0.3	15.1		7.2	5.0	14.0		
	1656	1.7	4.7	2.4		2.7	2.5	2.5		
Average		2.2	4.3	12.5	6.3	5.3	10.5	12.8	9.5	7.9
Overall average		6.9	7.9	13.8	9.5	6.9	11.1	13.6	10.5	10.0

estimate. The average absolute error on all segments was 7.9 for raw estimates and 11.1 for bias-corrected estimates. The average absolute error for all three analysts was 12.8 for raw early-season estimates, 6.3 for raw full-season estimates, and 9.5 for all eight segments with the raw estimate. The overall absolute proportion estimation error was 10.0.

The most obvious feature of table 6-18 is the large variability between analysts and between segments. If this variation is taken to be typical, future experiments should be designed to account for the variability of segments, analysts, and treatments.

Analysis of variance was used to test for the effects of analysts, time (i.e., early-season versus all acquisitions), method (raw versus bias correction), and their interactions. The results lead to the following conclusions:

a. The large disparity between data from various analysts was not consistent over segments; i.e., the analyst would do better on one segment than on another one.

b. There was no significant difference between methods; i.e., the use of bias correction exchanged one random error for another one of comparable magnitude.

c. Any test involving "times" was not significant; these tests had extremely low power because of an insufficient number of segments.

6.13 DIGITIZATION OF CANADIAN PHASE III TEST SITES

During Phase III, CAMS provided 22 acceptable LACIE classifications on 20 Canadian test sites in Saskatchewan (table 6-19). A program to acquire digitized ground truth for these 20 sites was established using aircraft imagery and annotated overlays designating in the image all nonagricultural and agricultural features.

The digitization of ground truth was a means to produce crop proportion estimates to evaluate the classification proportion accuracy and to produce ground-truth labels for evaluation of analyst dot labeling accuracy.

TABLE 6-19.— PHASE III CANADIAN TEST SITES

Segment number	Location		Strip fallow, %
	Lat., N.	Long., W.	
3064	49°29'	104°38'	0
3075	49°03'	105°30'	25.4
3080	49°19'	107°14'	10.8
3083	49°32'	106°20'	70.8
3093	49°06'	108°16'	22.1
3099	49°51'	108°22'	81.4
3112	50°19'	104°50'	10.1
3132	51°03'	105°48'	7.7
3143	50°53'	108°01'	11.3
3147	51°09'	109°10'	18.3
3159	51°25'	104°04'	0
3163	51°59'	104°55'	0
3165	51°16'	106°47'	5.7
3166	51°28'	105°40'	0
3169	51°30'	106°58'	.5
3185	52°05'	109°09'	3.9
3197	53°21'	103°19'	0
3201	52°24'	105°35'	0
3207	52°47'	106°44'	0
3214	53°08'	109°39'	0

The aerial photographs for each site were taken in 1970 and were scaled to 1:24 000 in order to define field boundaries. The inventories of all fields in the test site were taken near harvest time (late summer 1977).

After two sites had been digitized, the products were reviewed and two problems became apparent. The first problem was that sizable portions of the test sites were not inventoried; i.e., designated neither nonagricultural nor agricultural, as is evident in the ground-truth map of test site 3075, in which 51.4 percent of the segment was noninventoried agriculture (fig. 6-20). The second problem was that field boundaries on Landsat imagery (1977) did not always agree with field boundaries on aerial photographs (1970) as seen in the crop inventory (1977) using the 1970 photographs. Some, but not all, field boundaries were redrawn on the overlays to the 1970 photographs to reflect the changed field boundaries. Figure 6-21 shows the field boundaries from the ground-truth map superimposed on a false-color image of test site 3075. The ground-truth field boundaries in the blue-shaded areas do not agree with the field boundaries on the Landsat image.

The percentage of strip fallow (crop strips alternating with fallowed soil strips) is listed in table 6-19. The variability of strip fallow in the fields in the test sites is large, from 0 percent to 81 percent. Strip fallow noted on aerial photographs occasionally is noted to disagree significantly with the inventory.

Since analyst dot labels are prepared for the entire segment, the accuracy of the pixel labeling and classification can only be done for a small number of analyst labels and portions of the scene. This is shown by noting that in test site 3075, only 12 of the 32 analyst dot labels were confirmed with ground truth; in test site 3080, only 8 of the 25 analyst labels were confirmed with ground truth.

Because of the problems incurred, digitization of the remaining Canadian test sites was discontinued. The problems encountered in this program have prompted review of aerial photographs to assure that the image reasonably

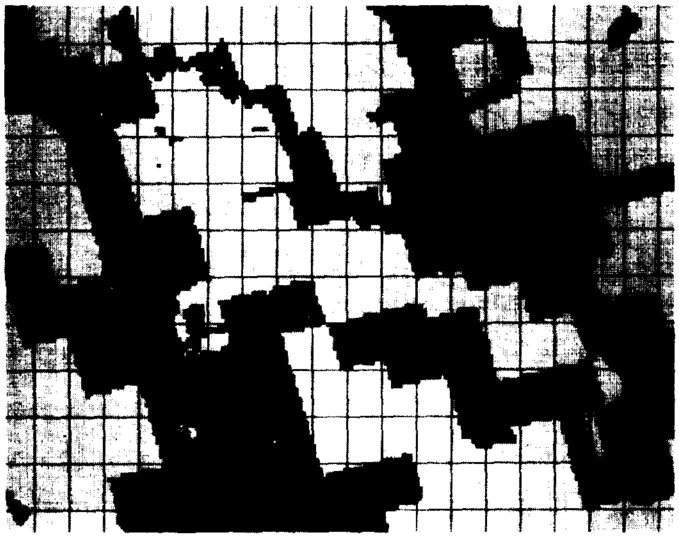


Figure 6-20.– Digitized ground truth for Canadian test site Hart Butte, Saskatchewan, registered to Landsat imagery of LACIE segment 3075, 1977 crop inventory.

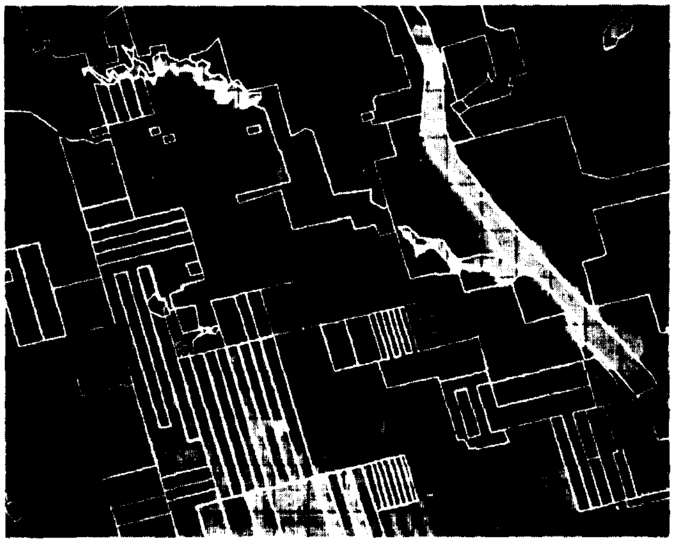


Figure 6-21.– Ground-truth boundaries registered to Landsat imagery, showing area of apparent boundary error.

reflects current agricultural crop boundaries. In addition, instructions to the agricultural agents who perform the inventories will emphasize the need for annotation of the entire segment.

6.14 ANALYSIS OF THE CLASSIFICATION OF U.S. AND CANADIAN ITS's

During Phase III of LACIE, 24 U.S. and 10 Canadian ITS's were scheduled for processing on the I-100, using the P1 hybrid classification system. This section describes an evaluation of the results obtained.

In this study, the only segments analyzed were those with (1) satisfactory classifications and (2) complete classification and ground-truth data. This eliminated 22 of the 34 segments: 4 segments were not classified; 6 segments were eliminated because of unsatisfactory classifications; and 12 segments were eliminated because of inadequate classification or ground-truth data, leaving 12 segments with 13 estimates. (Segment 1968 is a mixed-wheat segment; so estimates for this segment are available for both spring and winter wheat.)

It was originally intended to investigate accuracy in both labeling and proportion estimation; but this part of the investigation was abandoned because insufficient dot-labeling data were available.

6.14.1 METHOD

In the procedures used on the I-100, the analyst could give proportion estimates for winter wheat, winter grains, spring wheat, or spring grains; but estimates could not be compared directly with corresponding ground-truth proportions because the ground truth did not cover the whole segment. Therefore, an analyst's estimate of the proportion in the ground-truth area was obtained and compared with the ground-truth value in the following manner.

The image, classification map, and a map of the ground-truth area were read into the I-100. Using existing hardware and software, the analyst counted the number of pixels within the ground-truth area and the number of these pixels classified as wheat or grains. After deleting DO pixels, the analyst used the pixel counts to calculate the "uncorrected" proportion estimate for the ground-truth area. The bias correction alpha table, as stored in the I-100, was then used to obtain a bias-corrected estimate for the ground-truth area.

6.14.2 RESULTS

The results of the evaluation are shown in table 6-20. The table identifies each segment, the acquisition used for the estimate, the type of estimate (i.e., spring wheat, winter wheat, spring grains, and winter grains), and the Robertson biostage as determined from the ACC for the latest imagery used. In each case, the uncorrected proportion estimate, $\hat{X}$, and the bias-corrected proportion estimate $\hat{X}_C$, which are expressed as percentages, are given for the entire segment. The corresponding quantities $\hat{Y}$ and $\hat{Y}_C$ for the ground-truth area within the segment are given in the next two columns. Estimates passed to CAS are identified in table 6-20 by an asterisk. The ground-truth proportion, Y, which was determined from field reports, and the difference, $D = \hat{Y}_C - Y$, which is the proportion error for the ground-truth area, are also given.

LACIE segment 1973 (Whitman County, Washington) had the largest difference value, D = -33.8. Labeling data were available for this classification; they showed that all labeling errors were for wheat labeled nonwheat. Of 98 labeled type 1 and type 2 dots, approximately 25 percent were wheat dots labeled nonwheat. Note that 90 percent of the cultivated crops within the ground-truth area (excluding pasture and summer fallow) were winter wheat.[3]

The largest overestimate (17.2 percent) occurred for LACIE segment 1992. This segment had only 2.3 percent spring wheat within the ground-truth area, but the large amount of barley raised the total spring grains to 40 percent. The high wheat estimate was caused primarily by the inclusion of a large

---

[3] Ground truth was available for approximately one-third of this segment. This area contained 53.5 percent winter wheat, 3.9 percent spring barley, 2.4 percent dry peas, and 40.1 percent fallow and pasture.

TABLE 6-20.— LACIE PHASE III CLASSIFICATION OF U.S. AND CANADIAN ITS's USING THE I-100 HYBRID SYSTEM

Number	Location	Segment				Segment estimate, %		Ground-truth area estimate, %		Ground-truth (Y), %	Bias-corrected estimate minus ground-truth ($\hat{Y}_C$ - Y), %
		Number of acquisitions used	Latest acquisition used	Type of estimate	Robertson biostage	Uncorrected proportion estimate ($\hat{X}$)	Bias corrected proportion estimate ($\hat{X}_C$)	Uncorrected ($\hat{Y}$)	Bias corrected ($\hat{Y}_C$)		
					United States						
1963	Kansas	1	77101	Winter wheat	3.4	35.2*	41.1	30.8	37.0	34.3	2.7
1964	Kansas	3	77193	Winter wheat	6.5	29.8	33.6*	37.4	37.8	48.6	-10.8
1965	North Dakota	1	77143	Spring wheat	2.7	27.4*	21.9	26.7	22.1	43.1	-21.0
1973	Washington	1	77118	Winter grains	2.8	18.5	18.3*	22.2	19.7	53.5	-33.8
1975	Idaho	1	77112	Winter grains	2.2	9.3*	13.9	7.3	13.8	7.0	6.8
1976	Idaho	1	77166	Winter wheat	4.0	12.8	8.0*	20.4	12.6	6.8	5.8
1983	Indiana	2	77130	Winter wheat	3.8	10.5	14.5*	13.2	16.2	2.8	13.5
1986	South Dakota	3	77194	Winter wheat	5.9	2.5*	5.5	2.3	4.6	.5	4.1
				Spring wheat	4.9	5.6*	16.4	5.3	13.5	3.8	9.7
1988	Kansas	3	77211	Winter wheat	7.0	46.8	42.8*	46.1	42.3	41.5	.8
					Canada						
1958	Saskatchewan	2	77182	Spring wheat	4.0	47.8	34.8*	40.6	29.6	15.3	14.3
1991	Manitoba	3	77230	Spring grains	6.9	58.0	49.5*	50.0	44.0	61.3	-17.3
1992	Alberta	1	77170	Spring wheat	3.4	32.4	18.5*	34.2	19.5	2.3	17.2

portion of barley in the wheat estimate. It is particularly difficult to separate spring wheat and spring barley without an acquisition showing a difference in crop development such as that which occurs near the ripening stage of barley.

The estimate for LACIE segment 1975 is probably better than is indicated by the data in table 6-20. Examination of the imagery and the classification map indicates that the wheat proportion was probably much higher in the area outside rather than inside the ground-truth area. The bias correction factors which were calculated for the entire segment are probably valid for the segment; but because the wheat proportions are significantly different in the ground-truth area, these bias correction factors are inappropriate. The bias correction was nevertheless applied to the ground-truth area for consistency. (See table 6-20.)

Figure 6-22 is a plot of the proportion errors as a function of the ground-truth proportions for all 13 estimates. The plot shows that the errors are overestimates for segments with low wheat proportions and that the errors are underestimates for segments with high wheat proportions. This behavior is similar to that observed for blind sites during LACIE Phases I and II (ref. 3).

With the exception of LACIE segment 1973, figure 6-22 shows that the largest proportion errors occurred in spring wheat and spring grains. Of the five spring-wheat and spring-grain estimates, two segments (1958 and 1992) had very large overestimates of spring wheat in segments with very large proportions of barley and oats (46.0 and 38.6 percent, respectively).

6.14.3 CONCLUSIONS

Insufficient data are available to provide a clear picture of the accuracy of the I-100 classifications, but some specific observations are valid.

a. Labeling of wheat rather than total grains, particularly with only one acquisition, led to significant overestimates in some segments.

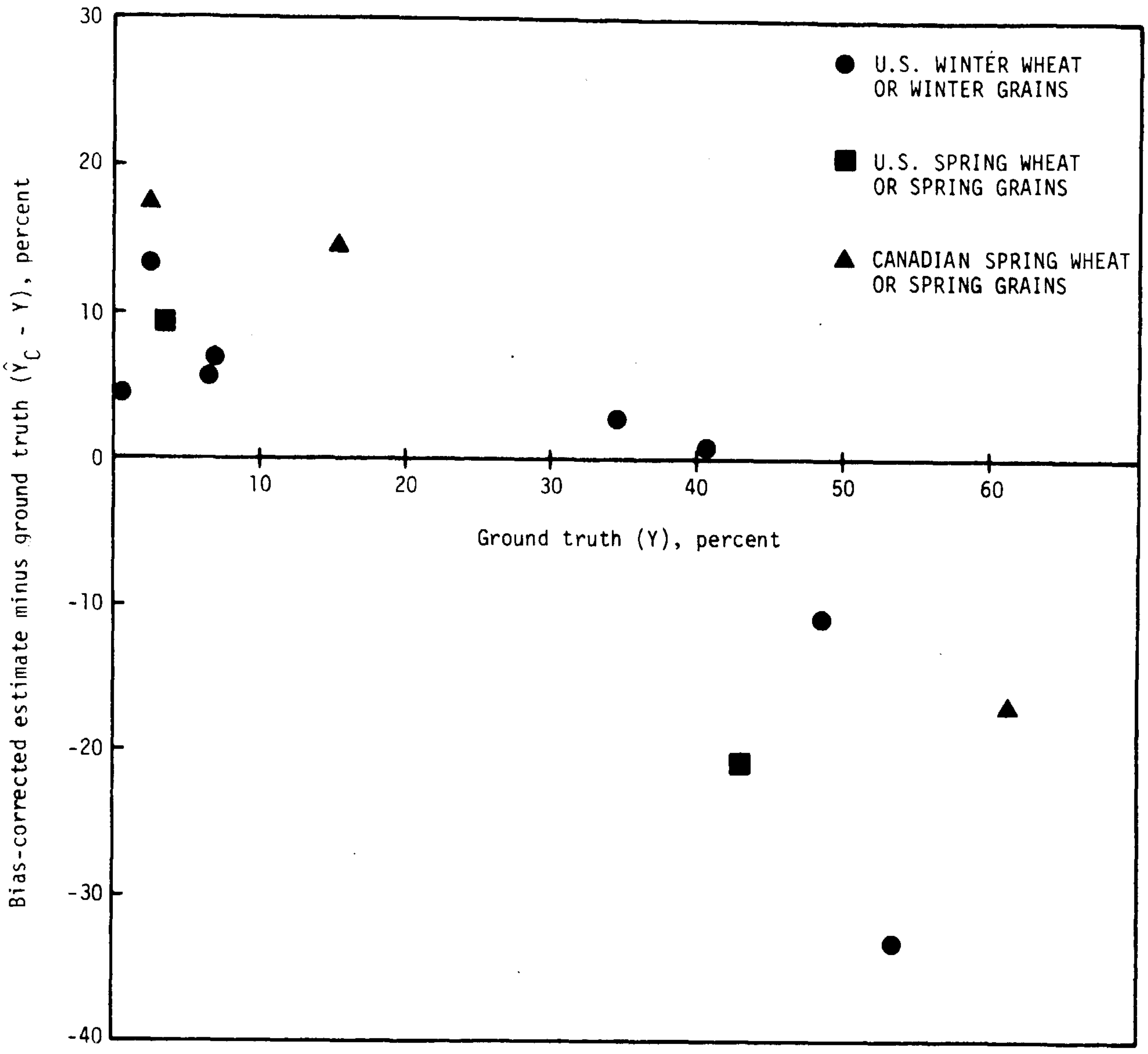


Figure 6-22.— Plot of proportion errors as a function of ground-truth proportions.

b. The I-100 software and the procedures were written to facilitate classification of the LACIE segments but were not designed to record data for later AA. A much better evaluation would have been possible if AA data had been collected after each satisfactory classification.

c. A more thorough quality assurance check on the I-100 analysis would probably have prevented reporting to CAS some of the segments (such as 1973 and 1992) which had large estimation errors.

6.15 A SIMULATION STUDY OF LACIE TECHNOLOGY

An assessment of the performance of the LACIE system requires several years of LACIE results. Two types of problems are thus presented: (1) It will be several years before these data are available; and (2) the LACIE system is evolving from year to year, so that the results obtained over several years are actually representative of several different LACIE systems.

The LACIE performance predictor (LPP) is a set of computer programs which simluate the performance of a given LACIE system (i.e., the system used in a given year or phase of LACIE). The LPP can be used to evaluate the system by simulating the input and thereby simulating the results that would be obtained in several years of operation of that system and to study the effect of various error sources on the final LACIE estimates.

This study describes several runs that were made with the LPP, each of which simulated 15 years of LACIE Phase II operations. The runs correspond to different sets of assumptions about the basic error sources in the LACIE system.

6.15.1 THE LPP

The LPP simulates these major elements of the LACIE system: segment acquisition, estimation of wheat proportion within the segment, yield estimation, and area and production aggregation.

The procedures used to perform these simulation tasks are described in following sections.

6.15.1.1 Segment Acquisition

The first major task the LPP performs is to simulate the acquisition of sample segment data acquired by Landsat. The segments are located at positions determined by the LACIE Phase II allocation. The LPP calculates the orbit of the Landsat and prepares a file containing the dates on which Landsat acquired data for each segment.

Subsequently, an allowance is made for cloud cover because acquisitions with cloud cover above a given threshold are not used by the LACIE system. Historical cloud cover data from weather observations are used to simulate the cloud cover on each acquisition of each segment. This is done by randomly choosing a cloud cover whose probability of a given cloud cover percentage being selected is equal to the frequency it was observed in the past. If the percentage of the simulated cloud cover is greater than the threshold value, the acquisition is rejected.

6.15.1.2 Simulation of County Proportions

It is assumed that the county proportion P_i for the *ith* county is distributed according to a beta distribution; i.e.,

$$P_i \sim \beta(\mu_i, \xi_i)$$

where μ_i is the mean and ξ_i is the SD of the distribution. The means are taken to be the 1975 proportions as determined by the USDA/SRS. For the *ith* county, the proportion is denoted by $P_{75,i}$. The SD's are calculated for each county by taking the SD of the historical proportions for that county for the years 1965 through 1974. However, the LPP does not accept $P_{75,i}$ and ξ_i directly but instead requires the following inputs:

$$CV_{1,i} = \xi_i / \overline{P}_{H,i}$$

$$\delta PW_i = (\mu_i - \overline{P}_{H,i}) / \overline{P}_{H,i}$$

where $\overline{P}_{H,i}$ equals the average proportion for the *ith* county for the years 1965 through 1974 and μ_i equals $P_{75,i}$.

Both $CV_{1,i}$ and δPW_i are manually calculated from ξ_i, $\overline{P}_{H,i}$, and $P_{75,i}$ and used as inputs to the LPP, which calculates μ_i and ξ_i for each county. A simulated "true" county proportion P_i is then calculated by the LPP for each county i by choosing a random number generated for the distribution $\beta(\mu_i, \xi_i)$.

6.15.1.3 Simulation of True Proportions for Segments

For segments in county i, it is assumed that the segment proportions X_i are distributed according to a beta distribution $\beta(P_i, \Theta_i)$, where P_i is the true county proportion described above and Θ_i^2 is the within-county variance of segment wheat proportions. To determine Θ_i^2, analysts use previous studies that provide an estimate of the within-county variance of small-grain proportions. It is assumed that this estimate is equal to Θ_i^2. The studies are based on LACIE analysts' interpretation of Landsat imagery of all the counties in the USGP. Each county was partitioned into segments, and estimates were made of the total agriculture proportion of each segment in each county. These estimates were used to calculate the average agriculture proportion $\overline{X}_{ag}$ and the agriculture variance $\Theta_{ag,i}^2$ for all the counties in the USGP.

Another task consisted of doing the same type of analysis to produce estimates of average small-grain proportion $\overline{X}_{sg,i}$ and small-grain variance $\Theta_{sg,i}^2$. However, the results for small grains are limited to a subset of approximately 45 counties. A simple regression model based upon $\overline{X}_{ag,i}$, $\Theta_{ag,i}^2$, $\overline{X}_{sg,i}$, and $\Theta_{sg,i}^2$ in the subset is used to obtain values of $\Theta_{sg,i}^2$ for all the counties. As stated previously, it is assumed that $\Theta_i = \Theta_{sg,i}$.

The LPP does not accept as input the variance Θ_i^2 itself but the CV, which is given by $CV_{2,i} = \Theta_i/P_i$. These could be calculated if the P_i were known; but unfortunately the P_i, which are calculated by the LPP, are not available in advance to compute $CV_{2,i}$. Therefore, the following procedure is used to determine the $CV_{2,i}$. First, the following quantities are calculated:

$$\overline{P} = \frac{1}{n} \sum_{i=1}^{n} P_{75,i}$$

$$\overline{\Theta^2} = \frac{1}{n} \sum_{i=1}^{n} \Theta_i^2$$

$$CV_2 = \frac{\sqrt{\overline{\Theta^2}}}{\overline{P}}$$

where n is the number of counties in a given state. The value obtained for CV_2 is then input to the LPP for each of the nine states. For the counties in each state, the LPP calculates the quantities:

$$\Theta_i' = (CV_2)P_i$$

These are taken as the estimates of the within-county variance of wheat proportion to be used in the model.

A "true" wheat proportion X_{ij} is then simulated for each segment j in the *ith* county by choosing a random number generated for the distribution $\beta(P_i, \Theta_i')$.

6.15.1.4 Simulation of the CAMS Estimate

It is assumed that the CAMS estimates of wheat proportions $\hat{X}_{ij}$ for the *jth* segment in the *ith* county are distributed according to the beta distribution $\beta(X_{ij} + B_{ij}, \sigma_{ij})$, where B_{ij} is the CAMS bias for the *jth* segment in the *ith* county and σ_{ij} is the variance in the CAMS errors (i.e., in $\hat{X}_{ij} - X_{ij}$) for these segments. The values B_{ij} and σ_{ij} were estimated using blind site data and CAMS estimates as follows:

$$B_{ij} = \left(\frac{B}{\overline{X}}\right) X_{ij}$$

and

$$\sigma_{ij} = \left(\frac{\sigma}{\overline{X}}\right) X_{ij}$$

$$B = \frac{1}{N_B} \sum_{k=1}^{N_B} \left(\hat{X}_k - X_k\right)$$

$$\sigma^2 = \frac{1}{N_B - 1} \sum_{k=1}^{N_B} \left(\hat{X}_k - X_k - B\right)^2$$

$$\overline{X} = \frac{1}{N} \sum_{i=1}^{n} \sum_{j=1}^{n_i} X_{ij}$$

where

- N_B is the number of blind sites in the state.
- N is the number of segments in the state.
- n is the number of counties in the state.
- n_i is the number of segments in the *ith* county.
- X_k is the measured ground-truth wheat proportion (the true wheat proportion).
- $\hat{X}_k$ is the CAMS estimate of the wheat proportion for the *kth* blind site in the state.

The quantities $B/\overline{X}$ and $\sigma/\overline{X}$ are input to the LPP, which performs a multiplication by X_{ij} to obtain B_{ij} and σ_{ij}. Subsequently, it generates the $\hat{X}_{ij}$ by choosing a random number generated for the distribution $\beta(X_{ij} + B_{ij}, \sigma_{ij})$.

Different values of B and σ^2 are computed for each of the four biowindows, and the appropriate values are used to simulate a value of $\hat{X}_{ij}$ corresponding to an acquisition in a given biowindow. In principle, one could also calculate values of B and σ^2 corresponding to various combinations of biowindows. However, this was not done for the runs described in this paper because not enough blind sites had the required combinations of acquisitions. All of the estimates of $\hat{X}_{ij}$ described here correspond to a single acquisition.

6.15.1.5 Simulation of Yield Estimates

Yield estimates are simulated by the LPP for each CRD in the USGP. The true yield Y_i for the *ith* CRD is taken to be the 1975 yield estimates by the USDA/SRS. The final yield estimates corresponding to biowindow 4 for the *ith* CRD, Y_i, are assumed to be distributed normally; i.e.,

$$\hat{Y}_i \sim n(Y_i, \phi_{4i})$$

where the SD ϕ_{4i} is determined from the results of the 10-year test made by the CCEA of the yield model used in LACIE.

SD's of yield estimates for earlier biowindows have to reflect the increasingly unreliable nature of CCEA yield estimates made at earlier dates in the growing season. To do this, each SD input for a CRD for a particular biowindow is assumed to be 4 percent larger than the SD input for the biowindow that followed in the season. Working backwards from harvest, one derives the following.

$$\phi_{3,i} = 1.04\phi_{4,i}$$

$$\phi_{2,i} = 1.04\phi_{3,i}$$

$$\phi_{1,i} = 1.04\phi_{2,i}$$

6.15.1.6 Simulation of the LACIE Aggregation Procedure

The LPP simulates the LACIE aggregation procedure to produce estimates for each year of the harvested wheat area, the wheat yield, and the wheat production by CRD, state, region, and country. Estimates of the CV's (SD divided by the true value) are also produced at these levels identically to how they are produced by the actual LACIE aggregation procedures.

Any LPP aggregation corresponds to a particular date, and the CAMS estimate $\hat{X}_{ij}$ is based on the latest acquisition prior to that date. The time taken by the actual LACIE system to process an acquisition to the point where it is ready for aggregation is not considered in the LPP.

Two kinds of aggregations corresponding to the kinds of error included in the aggregation estimates are performed:

- *Sampling error only,* performed by aggregating the simulated true segment proportions X_{ij} with yield $\hat{Y}_i$ set equal to Y_i.
- *Sampling, classification, and yield errors,* performed by aggregating the CAMS estimates $\hat{X}_{ij}$ to the CRD and multiplying by the yield estimates, $\hat{Y}_i$, for the CRD to obtain a production estimate for the CRD. The acreage and the production estimates are then summed to obtain estimates for higher levels.

6.15.2 DESCRIPTION OF RUNS

In the evaluations described here, the outer loop shown in figure 6-23 is run four times, once corresponding to no clouds and three times using the regular cloud cover data, thereby producing four different sets of acquisition dates. By design, each produces the same set of values for the true county proportion P_i by using the same random number seed for the generation of the P_i in all four runs. Each run of the outer loop produces a data tape containing the results of that run. The run is an input to the inner loop (fig. 6-23), which is a separate set of programs. In all, two separate runs can be made with the inner loop for each of the four runs made with the outer loop, as listed in table 6-21. Each of the eight runs could be made to simulate any desired number of "years" of the system.

The runs corresponding to S0, S1, S2, and S3 in table 6-21 were each made to simulate 15 separate "years" of LACIE operations. After each area estimate $\hat{A}_i$ for the nine-state USGP region was calculated, the CV for that and all the previous years was calculated, as shown in figure 6-24. It appeared that at 15 years the CV had converged sufficiently well to a constant value to stop the processing.

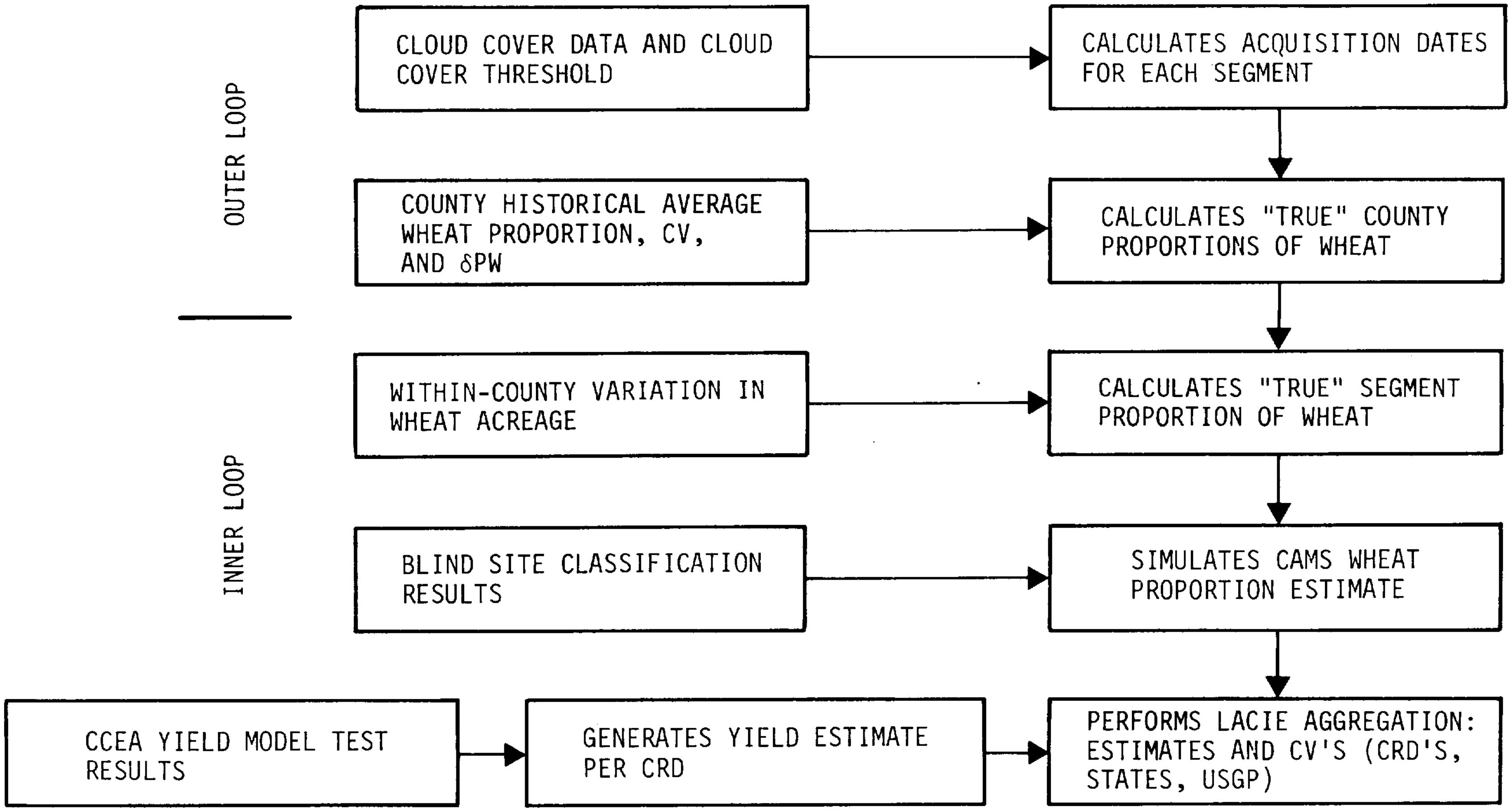


Figure 6-23.— LPP data flow.

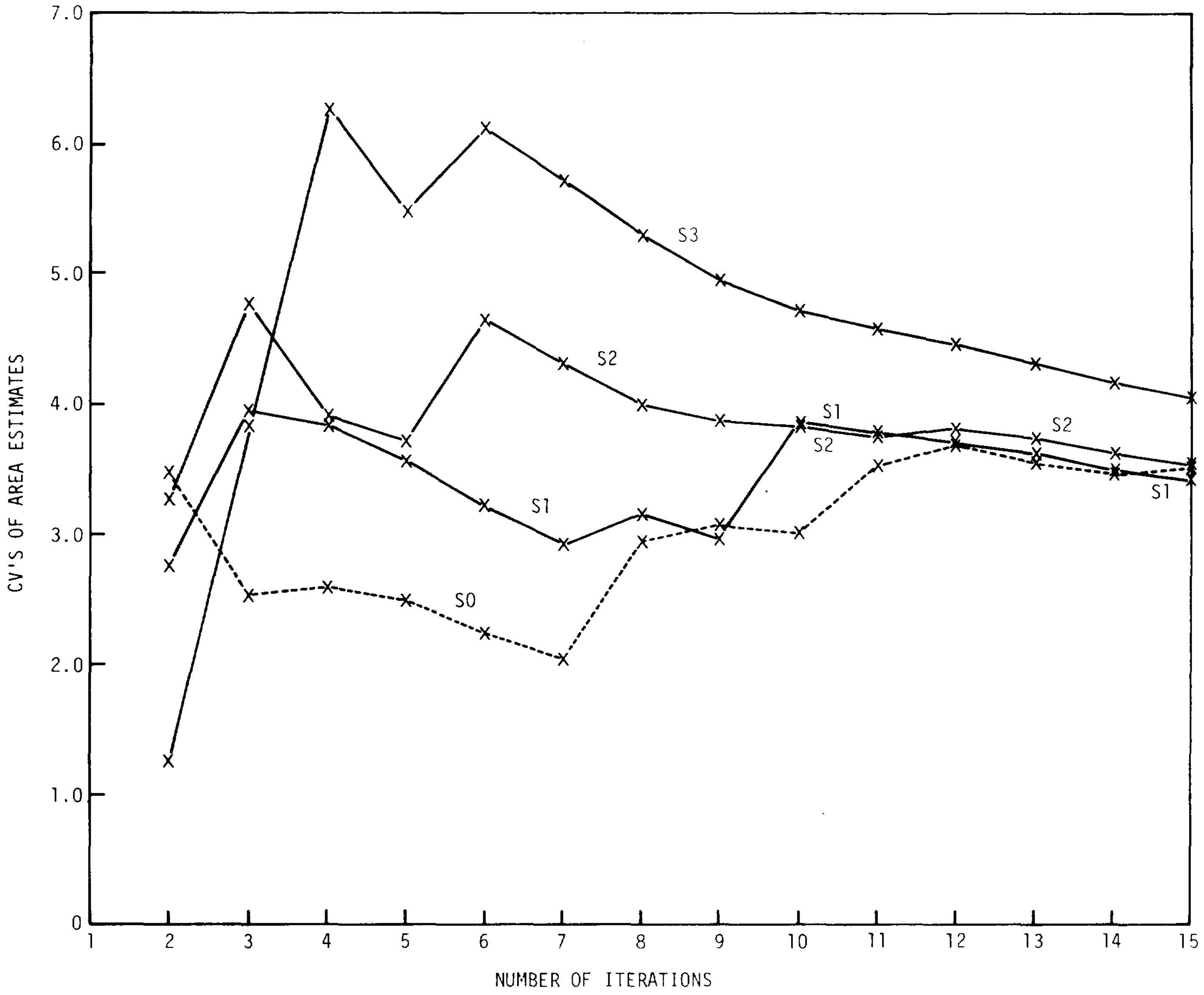


Figure 6-24.— CV's for area estimates for the USGP as a function of the number of iterations.

TABLE 6-21.— RUNS MADE WITH THE LPP

Outer-loop run	Inner-loop runs	
	Sampling error only	Sampling classification and yield errors
1	S0	SCY0
2	S1	SCY1
3	S2	SCY2
4	S3	SCY3

6.15.3 RESULTS

6.15.3.1 Segment Acquisition

The results of these runs were used to make a study of the acquisition simulation part of the LPP. The fraction of the sample segments having at least one acquisition as determined by the LPP was plotted as a function of time and compared with the number actually obtained in LACIE. The results are shown in figure 6-25. The curve labeled A is the LPP results for which zero cloud coverage was assumed; i.e., the cloud cover simulator was programmed to always produce a cloud cover of zero. By December 1, all the winter wheat segments had been acquired and the curve was flat until April 1, when the acquisition of spring wheat sites began. All sites had been acquired at least once by July 1.

The three curves labeled B correspond to simulations of three different years of LACIE operations in which the only factor which varied was the cloud cover (i.e., S1, S2, and S3). A threshold of 50 percent was used and was chosen to obtain approximately the same total number of acquisitions over the year as was obtained in LACIE Phase II.[4] The three curves are quite close together, indicating only a small effect of a cloud cover on acquisition

---

[4] Actually the threshold was too high; the three curves labeled B correspond to about 15 percent more acquisitions than the 2249 acquisitions in LACIE Phase II.

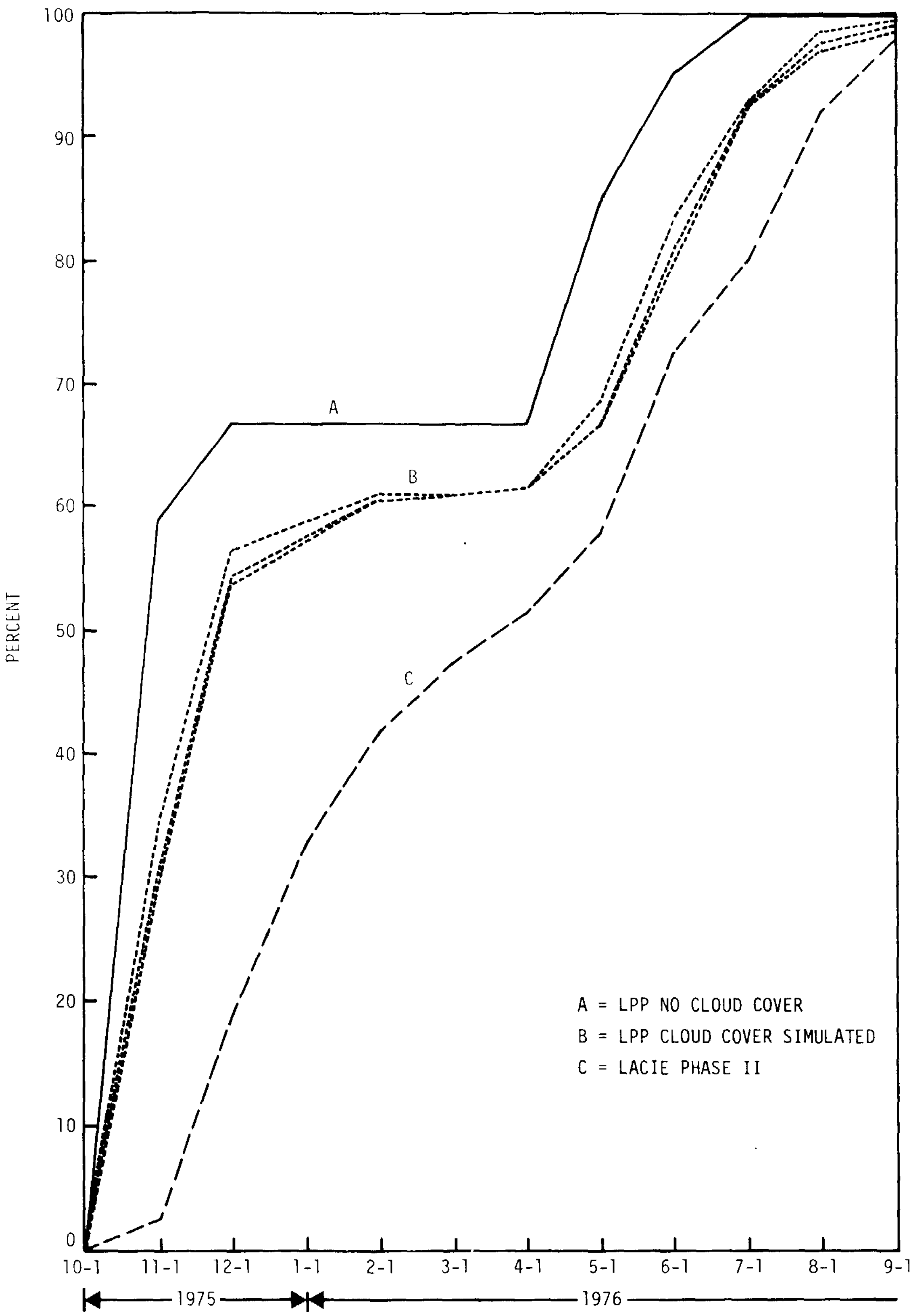


Figure 6-25.– Percentage of sites acquired as a function of date.

history. This is probably due to the fact that in the LPP it is assumed that cloud cover at each segment is independent of the cloud cover at all other segments, whereas in fact there is probably a high degree of correlation between the amounts of cloud cover over segments that are reasonably close together.

Curve C is the actual LACIE Phase II acquisition history. It is lower than the curves produced by the LPP for all dates, partly because the cloud cover threshold of 50 percent was too high and partly because the discrepancy is quite large early in the year. The reason for this is not known.

6.15.3.2 True Proportions for Segments

The true wheat proportions for the blind site segments generated by the LPP in run number S1 are compared with the actual blind site proportions in figure 6-26. The LPP produced more segments with 0 to 4 percent wheat and more segments with a high proportion of wheat (greater than 55 percent). A Kolmogorov-Smirnov test was performed; it showed that there was no significant difference between the two distributions.

6.15.3.3 CAMS Proportion Estimates

Originally, it was planned to also make all the runs corresponding to column 2 in table 6-21. However, because each run took about 3 hours of computer time, it was decided to drop the unrealistic case SCY0. Because there was little difference between the results of the runs of S1, S2, and S3, it was decided to drop SCY2 and SCY3. Thus, the only run made which included more than sampling error alone was SCY1, which used the same true county proportions as S1.

Figure 6-27 shows a histogram of the LACIE errors ($\hat{X}_k - X_k$) for all the blind sites in the USGP region. Figure 6-28 shows a histogram of the errors simulated by the LPP in run SCY1; i.e., $\hat{X}_{ij} - X_{ij}$ for all blind sites in the USGP. These histograms should be similar if the LPP is correctly simulating the results of the CAMS classification procedures.