EVALUATION OF «STRAW MAN» MODEL 1, THE SIMPLE LINEAR MODEL, FOR SOYBEAN YIELDS IN IOWA, ILLINOIS AND INDIANA

Jeanne L. Sebaugh

ESS Staff Report No. AGESS810304

Statistical Research Division
Economics and Statistics Service
U.S. Department of Agriculture
Columbia, Missouri 65201

March 1981

EVALUATION OF "STRAW MAN" MODEL 1, THE SIMPLE LINEAR MODEL, FOR SOYBEAN YIELDS IN IOWA, ILLINOIS, AND INDIANA. By Jeanne L. Sebaugh; Statistical Research Division, Economics and Statistics Service, U.S. Department of Agriculture, Columbia, Missouri 65201; February 1981.
ESS Staff Report No. AGESS 8102

ABSTRACT

Straw man model 1 is one of the simplest regression models which can be used to predict crop yields. A one-line regression of yield over time, it represents the increases in yield which have occurred through the adoption of improved varieties, and the increased use of fertilizer and other cultural practices. The performance of this model in predicting soybean yields in Iowa, Illinois, and Indiana is evaluated. Eight model characteristics are discussed. Indicators of yield reliability obtained from bootstrap testing show that the bias is generally small for this model. However, the model is unable to predict the low and high yields accurately. The model is objective, adequate, timely, simple, and not costly. However it does not consider known scientific relationships and does not provide a good current measure of modeled yield reliability.

Key words: Model Evaluation, yield modeling, linear regression.

ACKNOWLEDGMENTS

The author wishes to thank Wendell Wilson, other members of the Yield Evaluation Staff and various AgRISTARS Yield Development Project personnel for their comments and assistance, and Joan Wendt for typing this report.

EVALUATION OF "STRAW MAN" MODEL 1,
THE SIMPLE LINEAR MODEL,
FOR SOYBEAN YIELDS IN
IOWA, ILLINOIS AND INDIANA

by
Jeanne L. Sebaugh
Mathematical Statistician

This research was conducted as part of the AgRISTARS Yield Model Development Project. It is part of task 3 (subtask 2) in major project element number 1, as identified in the 1980 Yield Model Development Project Implementation Plan. As an internal project document, this report is identifed as shown below.

AgRISTARS
Yield Model Development
Project

YMD-1-3-2(81-03.1)

FOREWORD

Development and application of techniques for crop yield model test and evaluation are important parts of the Yield Model Development Project in AgRISTARS. 1/ Promising yield models available in the literature or from various researchers will be subjected to performance test and evaluation. In order that there may be a common reference for describing the capabilities and limitations of these models, criteria for doing so have been developed and described in a document entitled Crop Yield Model Test and Evaluation Criteria (Wilson, et al., 1980). These criteria are used both in the evaluation of a single model and in comparisons between models.

The purpose of preparing this document is to gain some experience in the application of the criteria for evaluative purposes. A follow-up document will use the same criteria to compare models. It is anticipated that the evaluation and comparison of other models will be done in a similar manner.

The models to be evaluated and compared were chosen to be quite simple since the focus of attention is on the "pilot test" of the procedures. The models involved are the "straw man" crop yield models developed and discussed by Kestle (1981). This document evaluates the simplest of the straw man models, the one line model, regressing yield on year.

Jeanne L. Sebaugh
Mathematical Statistician
Yield Evaluation Staff
Yield Research Branch
Statistical Research Division

1/ Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing (AgRISTARS) is a multi-agency research program to meet some current and new information needs of the U.S. Department of Agriculture.

Table of Contents

Page

Summary. 1

Description of the Model 1
Straw Man Models Describe Technological Trends 1
Straw Man Model 1 - Uniform Trend Over Time. 1

Evaluation Methodology 2
Eight Model Characteristics to be Discussed. 2
Bootstrap Technique Used to Generate Indicators of Yield Reliability. 2
Review of Indicators of Yield Reliability. 3
Indicators Based on Differences between Y and $\hat{Y}$ (d) Demonstrate Accuracy, Precision and Bias 3
Indicators Based on Relative Differences between Y and $\hat{Y}$ (rd) Demonstrate Worst and Best Performance. 3
Indicators Based on Y and $\hat{Y}$ Demonstrate Correspondence Between Actual and Predicted Yields. 6
Current Measure of Modeled Yield Reliability Defined by a Correlation Coefficient 6
Model Evaluation... 7
Indicators of Yield Reliability Based on Differences between Y and $\hat{Y}$ (d) Show Small Bias and a Standard Deviation between 1½-3¼ Quintals/Hectare. 7
Indicators of Yield Reliability Based on Relative Differences between Y and $\hat{Y}$ (rd) Show 1974 as Worst Year and 20-40 Percent of the Years Have (rd) Greater than 10 Percent. 7
Indicators of Yield Reliability Based on Y and $\hat{Y}$ Show Low Correspondence Between the Direction of Change in Predicted as Compared to Actual Yields 14
Base Period Indicates More Precision Than Independent Tests Can Confirm. 14
Model is Objective 23
Model Does Not Consider Known Scientific Relationships 23
Model is Adequate. 25
Model is Timely. 25
Model is Not Costly. 25
Model is Simple. 25
Model Has Poor Current Measure of Modeled Yield Reliability. . . . 25

Conclusions. 28

References 29

List of Figures and Tables

Page

Table 1: Average Production and Yield for Test Years 1970-1979. . 4

Figure 1: Production of soybeans by CRD (1970-79 average), as a percent of the regional total. 5

Table 2: Indicators of Yield Reliability Based on D = Predicted - Actual Yield 8

Figure 2: Root mean square error (RMSE) for soybeans in quintals per hectare based on test years 1970-1979. 9

Table 3: Indicators of Yield Reliability Based on RD = 100 * ((Predicted-Actual Yield)/Actual Yield). 10

Figure 3: Percent of test years (1970-1979) the absolute value of the relative difference is greater than ten percent for soybeans 11

Figure 4: Largest absolute value of the relative difference for soybeans during the test years 1970-1979 12

Figure 5: Next largest absolute value of the relative difference for soybeans during the test years 1970-1979 13

Figure 6: Iowa State Model. Actual and Predicted Yields for the Test Years 1970-1979 15

Figure 7: Illinois State Model. Actual and Predicted Yields for the Test Years 1970-1979 16

Figure 8: Indiana State Model. Actual and Predicted Yields for the Test Years 1970-1979 17

Table 4: Indicators of Yield Reliability Based on Actual and Predicted Yields 18

Figure 9: Percent of test years (1970-1979) the direction of change in predicted yield from the previous year agrees with the direction of change in actual soybean yield . . 19

Figure 10: Percent of test years (1970-1979) the direction of change in predicted yield from the previous three year average agrees with the direction of change in actual soybean yield. 20

Figure 11: Pearson correlation coefficient between actual and predicted soybean yields in the test years (1970-1979) . 21

Table 5: Residual Mean Square as an Indicator of the Fit of the Model Based on the Model Development Base Period 22

Page

Table 6: Correlation Between Observed and Predicted Yields as an Indicator of the Fit of the Model Based on the Model Development Base Period. 24

Table 7: Current Indication of Modeled Yield Reliability. 26

Figure 12: Spearman correlation coefficient between the estimate of the standard error of a predicted value from the base period model and the absolute value of the difference between the predicted and actual soybean yield in the test years (1970-1979) 27

Evaluation of "Straw Man" Model 1, The Simple Linear Model, for Soybean Yields in Iowa, Illinois, and Indiana

Jeanne L. Sebaugh

SUMMARY

Straw man model 1, simple linear regression of yield over time, describes a uniform increase in soybean yields over time. Indicators of yield reliability obtained from bootstrap testing are used as a basis of comparison between competing models and the results for straw man model 1 do not appear very promising. The bias is generally small, however, the model is unable to predict the low and high yields accurately. The model is objective, adequate, timely, simple, and not costly. However, it does not consider known scientific relationships and does not provide a good current measure of modeled yield reliability.

DESCRIPTION OF THE MODEL

Straw Man Models Describe Technological Trends

All of the straw man models attempt to explain differences in crop yields over time by simply fitting trend lines to the yield data. Improvements in technology, including varieties, hybrids, fertilizers, insecticides, herbicides, farming practices, equipment, etc., have resulted in steady improvements in yields. There are occasional set-backs, primarily due to weather, but the overall trend has been towards increasing yields.

The straw man models demonstrate how much of the year-to-year difference in yield can be explained simply by this technological trend. These models are not expected to be particularly accurate in predicting the yield for any future year since that year's particular weather conditions are not used by the model. However, as pointed out by Kestle (1981), these models may be treated as "below base" models. Any candidate model which cannot substantially outperform a straw man model is of questionable value.

Straw Man Model 1 - Uniform Trend Over Time

Straw man model 1 is a simple linear regression over time. The statistical model is $E(Y) = \beta_o + \beta_1 X$, where Y is the soybean yield in quintals per hectare and X is the corresponding year number (1950=0).

The inherent assumption in a simple linear regression model is that the rate of change in the Y variable is constant over the entire range of the X values. In our case, this means that the year-to-year increases in yield are assumed to be the same later on in the time period as they were earlier. Under that assumption, β_o is the yield in 1950 and β_1 is the increase in yield between any two adjacent years in the time period being modeled.

EVALUATION METHODOLOGY

Eight Model Characteristics to be Discussed

The document, Crop Yield Model Test and Evaluation Criteria, (Wilson, *et al.*, 1980), states:

> "The model characteristics to be emphasized in the evaluation process are: yield indication reliability, objectivity, consistency with scientific knowledge, adequacy, timeliness, minimum costs, simplicity, and accurate current measures of modeled yield reliability."

Each of these characteristics will be discussed with respect to straw man model 1.

Bootstrap Technique Used to Generate Indicators of Yield Reliability

Indicators of yield reliability (reviewed below) require that the parameters of the regression model be computed for a set of data and that a yield prediction be made based on that data for a given "test" year. The values required to generate indicators of yield reliability include the predicted yield, $\hat{Y}$, the actual (reported) yield, Y, and the difference between them, $d = \hat{Y}-Y$, for each test year. It is desirable that the data used to generate the parameters for the model not include data from the test year.

In order to accomplish this, the "bootstrap" technique is used. Years from an earlier base period are used to fit the model and obtain a prediction equation. The values of the independent variables for the test year following the base period are inserted into the equation and a predicted yield is generated. Then, the base period is shifted one year forward and the process is repeated. Continuing in this way, ten (1970-1979) predictions of yield are obtained, each independent of the data used to fit the model.

The $\hat{Y}$ and d values for the ten-year test period are obtained from models derived at the crop reporting district (CRD) level, state level, and region level. Another set of $\hat{Y}$ values are obtained at the state level by using a weighted average of the predicted yields from the CRD models. Predicted yields for the region are also obtained using a weighted average of the predicted yields from the CRD models and from the state models. The weighting factor used is harvested acreage for the year the prediction is made.

For Illinois and Indiana, data for 1947-1969 (23 years) are used to fit prediction models for 1970, data for 1948-1970 (23 years) are used to fit prediction models for 1971, etc. For Iowa, data for 1950-1969 are used to fit prediction models for 1970 (20 years), data for 1950-1970 are used to fit prediction models for 1971 (21 years), etc. When shifting the base period forward, the earliest year is dropped if it would result in more than 23 years of data. A base period of consistent size is desired because of the type of trend models with which straw man 1 will be compared and is not necessarily a standard bootstrap procedure.

The average and percent production and the yield over the ten year test period are listed in Table 1 for each geographic region. The percentage of regional production contributed by each CRD is shown graphically in Figure 1. Darker shades indicate higher production.

Review of Indicators of Yield Reliability

The Y, $\hat{Y}$ and d values for the ten-year test period at each geographic area may be summarized into various indicators of yield reliability.

Indicators Based on the Differences between Y and $\hat{Y}$ (d) Demonstrate Accuracy, Precision and Bias

From the d value, the mean square error (root and relative root mean square error), the variance (standard deviation and relative standard deviation), and the bias (its square and the relative bias) are obtained.

The root mean square error (RMSE) and the standard deviation (SD) indicate the accuracy and precision of the model and are expressed in the original units of measure (quintals/hectare). It is about 68% probable that the absolute value of d for a future year will be less than one RMSE and 95% probable that it will be less than twice the RMSE. So, accurate prediction capability is indicated by a small RMSE.

A non-zero bias means the model is, on the average, overestimating the yield (positive bias) or underestimating the yield (negative bias). The SD is smaller than the RMSE when there is non-zero bias and indicates what the RMSE would be if there were no bias. If the bias is near zero, the SD and the RMSE will be close in value. We prefer an unbiased model, i.e. bias close to zero.

Indicators Based on Relative Differences between Y and $\hat{Y}$ (rd) Demonstrate Worst and Best Performance

The relative difference, rd=(100d/Y), is an especially useful indicator in years where a low actual yield is not predicted accurately. This is because years with small observed actual yields and large differences have the largest rd values.

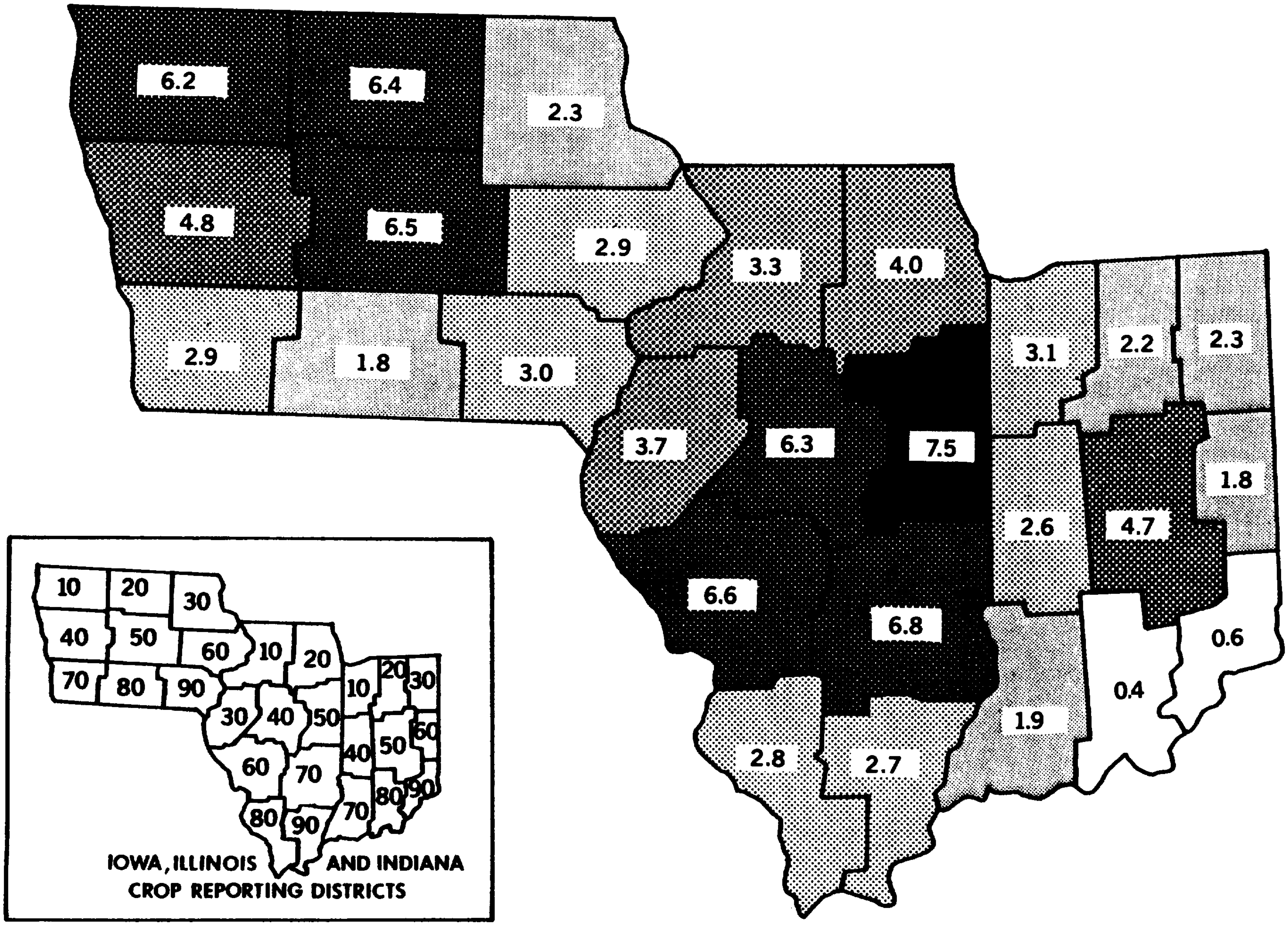
Several indicators are derived using relative differences. In order to calculate the proportion of years beyond a critical error limit, we count the number of years in which the absolute value of the relative difference exceeds the critical limit of 10 percent. Values between 5 and 25 percent were investigated and a critical limit of 10 percent was found most useful in describing model performance. The worst and next to worst performance during the test period are defined as the largest and next to largest absolute value of the relative difference. The range of yield indication accuracy is defined by the largest and smallest absolute values of the relative difference.

TABLE 1
AVERAGE PRODUCTION AND YIELD
FOR TEST YEARS 1970-79

SOYBEANS
IOWA, ILLINOIS, INDIANA

STATE	CRD	PRODUCTION (1,000) QUINTALS	BUSHELS	PERCENT OF STATE	REGION	YIELD QNTL/HA	BU/ACRE
IOWA	10	10,734	39,439	16.9	6.2	23.4	34.8
	20	10,992	40,389	17.3	6.4	22.7	33.8
	30	3,929	14,435	6.2	2.3	21.7	32.3
	40	8,189	30,090	12.9	4.8	22.3	33.1
	50	11,207	41,177	17.7	6.5	23.7	35.3
	60	4,996	18,358	7.9	2.9	24.5	36.4
	70	5,016	18,430	7.9	2.9	22.1	32.9
	80	3,107	11,415	4.9	1.8	20.4	30.4
	90	5,187	19,060	8.2	3.0	23.1	34.3
STATE		63,357	232,793		36.8	22.9	34.0
ILLINOIS	10	5,670	20,834	7.5	3.3	24.0	35.6
	20	6,960	25,575	9.2	4.0	22.2	33.0
	30	6,331	23,263	8.4	3.7	23.5	35.0
	40	10,855	39,885	14.4	6.3	25.0	37.2
	50	12,870	47,288	17.1	7.5	24.2	36.0
	60	11,412	41,931	15.1	6.6	23.2	34.6
	70	11,739	43,133	15.6	6.8	20.8	30.9
	80	4,800	17,637	6.4	2.8	19.2	28.6
	90	4,694	17,248	6.2	2.7	17.4	25.8
STATE		75,333	276,795		43.7	22.4	33.3
INDIANA	10	5,258	19,320	15.6	3.1	22.2	33.0
	20	3,717	13,659	11.1	2.2	21.5	32.0
	30	3,897	14,319	11.6	2.3	20.8	31.0
	40	4,443	16,326	13.2	2.6	22.5	33.5
	50	8,100	29,761	24.1	4.7	23.6	35.1
	60	3,142	11,544	9.3	1.8	21.0	31.2
	70	3,304	12,139	9.8	1.9	21.0	31.3
	80	709	2,604	2.1	0.4	18.3	27.3
	90	1,042	3,827	3.1	0.6	18.8	27.9
STATE		33,612	123,500		19.5	21.9	32.5
REGION		172,301	633,088			22.5	33.4

Figure 1. Production of soybeans by CRD (1970-79 average), as a percent of the regional total. Darker shades indicate CRDs with higher production.



Indicators Based on Y and $\hat{Y}$ Demonstrate Correspondence Between Actual and Predicted Yields

Another set of indicators demonstrates the correspondence between actual and predicted yields. It would be desirable for increases in actual yield to be accompanied by increases in predicted yields. It would also be desirable for large (small) actual yields to correspond to large (small) predicted yields.

Two indicators relate the change in direction of actual yields to the corresponding change in predicted yields. One looks at change from the previous year (nine observations) and the other at change from the average of the previous three years (seven observations). A base period of three years is used since a longer base period would further decrease the number of observations, while a shorter period would not be very different from the comparison to a single previous year.

Finally, the Pearson correlation coefficient, r, between the set of actual and predicted values for the test years is computed. It is desirable that $r(-1 \leq r \leq +1)$ be large and positive. A negative r indicates smaller predicted yields occurring with larger observed yields (and vice versa).

Current Measure of Modeled Yield Reliability Defined by a Correlation Coefficient

One of the model characteristics to be evaluated is its ability to provide an accurate, current measure of modeled yield reliability. Although a specific statistic was not discussed in the paper, Crop Yield Model Test and Evaluation Criteria, (Wilson, et al., 1980), it was stated that:

> "This 'reliability of the reliability' characteristic can be evaluated by comparing model generated reliability measures with subsequently determined deviation between modeled and 'true' yield."

For regression models, this suggests the use of a correlation coefficient between two variables generated for each test year. One variable is an indicator of the precision with which a prediction for the next year can be made, based on the model development base period. The other variable (obtained retrospectively) is an indicator of how close the predicted value for the next year actually is to the "true" value. The estimate of the standard error of a predicted value from the base period model is used for the first value, $s_{\hat{Y}}$, and the absolute value of the difference between the predicted and actual yield in the test year is used as the second variable, $|d|$.

A non-parametric (Spearman) correlation coefficient, r, is employed since the assumption of bivariate normality cannot be made. A positive value of $r(-1 \leq r \leq +1)$ indicates agreement between $s_{\hat{Y}}$ and $|d|$, i.e., a smaller (larger) value of $s_{\hat{Y}}$ is associated with a smaller (larger) value of $|d|$. An r value close to +1 is desirable since it indicates that a small standard error of prediction (and therefore a narrow confidence interval about the true predicted value) is associated with small discrepancies between predicted and actual yields. If this were the case, one would have confidence in $s_{\hat{Y}}$ as an indicator of the accuracy of $\hat{Y}$.

MODEL EVALUATION

Indicators of Yield Reliability Based on Differences between Y and $\hat{Y}$ (d) Show Small Bias and a Standard Deviation Between $1\frac{1}{2}$-$3\frac{1}{4}$ Quintals/Hectare

The CRD, state and region values of indicators of yield reliability based on d for this simple linear model are given in Table 2. The bias for CRDs is generally less than half a quintal in Iowa and Illinois and less than a quintal in Indiana. The CRDs in Iowa and Illinois have a relative bias of less than five percent. In Indiana, three CRDs have a relative bias between five and ten percent, while the rest are less than five percent.

The root and relative root mean square error values (RMSE and RRMSE) are somewhat lower in Iowa and higher in Illinois, as can be seen for the RMSE values in Figure 2. CRD values for RMSE range from 1.49 to 3.21 quintals/hectare and values for RRMSE range from 6.6 percent to 15.0 percent.

Generally, as the level of aggregation increases in size, the bias becomes closer to zero and the RMSE becomes smaller. This demonstrates the greater accuracy obtained with the data which has been stabilized through the aggregation process. The results are very similar regardless of whether the aggregation is done prior to fitting the model (state and region models) or after the models are fit (CRDs aggregated and states aggregated).

Indicators of Yield Reliability Based on Relative Differences Between Y and $\hat{Y}$ (rd) Show 1974 as Worst Year and 20-40 Percent of the Years Have rd Greater than 10 Percent

The CRD, state, and region values for indicators of yield reliability based on rd are given in Table 3. CRD values are also shown in Figure 3-5. Two to four of the ten test years have absolute relative differences greater than 10 percent in most (21 out of 27) of the CRDs (Figure 3). The very low yield in 1974 caused the largest absolute relative difference in most CRDs, ranging from 15.8 percent to 57.4 percent (Figure 4). The range in values for the next largest absolute relative difference is 7.4 percent to 30.4 percent (Figure 5). The smallest absolute relative difference is sometimes zero (four CRDs) and ranged up to 3.3 percent. These small absolute relative differences result in the range being very much like the largest absolute relative difference varying over CRDs from 14.9 percent to 55.4 percent.

As compared to the CRD results, the state and regional aggregate values for the largest and smallest absolute relative difference are somewhat lower. There are fewer years with absolute relative differences greater than 10 percent. The method of aggregation makes little difference.

TABLE 2
INDICATORS OF YIELD RELIABILITY
BASED ON D = PREDICTED - ACTUAL YIELD

STRAW MAN MODEL 1 - SOYBEANS
IOWA, ILLINOIS, INDIANA

MSE, VAR, B-SQR (QUINTALS/HECTARE SQUARED)
RMSE, SD, BIAS (QUINTALS/HECTARE)
RRMSE, RSD, RB (PERCENT OF AVERAGE YIELD)

STATE	CRD	MSE	RMSE	RRMSE	VAR	SD	RSD	B-SQR	BIAS	RB
IOWA	10	7.29	2.70	11.5	6.66	2.58	11.4	0.62	-0.79	-3.4
	20	2.21	1.49	6.6	2.18	1.48	6.5	0.03	0.18	0.8
	30	4.52	2.13	9.8	4.05	2.01	9.6	0.48	-0.69	-3.2
	40	6.77	2.60	11.7	6.77	2.60	11.6	0.00	0.06	0.3
	50	6.76	2.60	11.0	6.62	2.57	10.7	0.14	0.37	1.6
	60	4.80	2.19	9.0	4.54	2.13	8.5	0.26	0.51	2.1
	70	4.77	2.18	9.9	3.65	1.91	8.2	1.12	1.06	4.8
	80	8.22	2.87	14.0	8.20	2.86	13.9	0.02	0.14	0.7
	90	7.08	2.66	11.5	6.96	2.64	11.6	0.12	-0.35	-1.5
STATE MODEL		3.79	1.95	8.5	3.79	1.95	8.5	0.00	0.00	0.0
CRDS AGGR.		3.80	1.95	8.5	3.80	1.95	8.5	0.00	0.01	0.0
ILLINOIS	10	8.56	2.93	12.2	8.48	2.91	12.3	0.08	-0.29	-1.2
	20	7.35	2.71	12.2	6.81	2.61	12.2	0.53	-0.73	-3.3
	30	7.86	2.80	11.9	7.80	2.79	12.0	0.06	-0.24	-1.0
	40	10.32	3.21	12.8	10.31	3.21	12.8	0.00	0.04	0.2
	50	7.52	2.74	11.3	7.39	2.72	11.4	0.13	-0.36	-1.5
	60	6.21	2.49	10.7	6.07	2.46	10.4	0.14	0.38	1.6
	70	6.11	2.47	11.9	6.09	2.47	11.9	0.02	-0.14	-0.7
	80	5.71	2.39	12.4	5.70	2.39	12.4	0.00	0.05	0.3
	90	5.94	2.44	14.0	5.94	2.44	14.0	0.00	-0.00	-0.0
STATE MODEL		6.32	2.51	11.2	6.32	2.51	11.2	0.00	-0.05	-0.2
CRDS AGGR.		6.38	2.53	11.3	6.37	2.52	11.3	0.01	-0.12	-0.5
INDIANA	10	4.98	2.23	10.0	4.63	2.15	10.0	0.35	-0.59	-2.7
	20	7.01	2.65	12.3	6.20	2.49	12.1	0.81	-0.90	-4.2
	30	5.48	2.34	11.2	4.24	2.06	10.4	1.23	-1.11	-5.3
	40	7.13	2.67	11.8	7.11	2.67	11.9	0.02	-0.13	-0.6
	50	6.30	2.51	10.6	5.47	2.34	10.3	0.83	-0.91	-3.9
	60	3.21	1.79	8.5	2.74	1.66	8.2	0.46	-0.68	-3.2
	70	3.73	1.93	9.2	3.33	1.83	8.4	0.40	0.63	3.0
	80	7.53	2.74	15.0	4.50	2.12	10.6	3.03	1.74	9.5
	90	7.50	2.74	14.6	4.88	2.21	10.8	2.62	1.62	8.6
STATE MODEL		4.19	2.05	9.4	4.02	2.01	9.4	0.17	-0.41	-1.9
CRDS AGGR.		4.14	2.03	9.3	3.96	1.99	9.3	0.18	-0.42	-1.9
REGION MODEL		3.85	1.96	8.7	3.85	1.96	8.8	0.00	-0.07	-0.3
CRDS AGGR.		3.96	1.99	8.9	3.95	1.99	8.9	0.01	-0.11	-0.5
STATES AGGR.		4.01	2.00	8.9	4.00	2.00	8.9	0.00	-0.07	-0.3

Figure 2. Root mean square error (RMSE) for soybeans in quintals per hectare based on test years 1970-1979. Darker shades indicate CRDs with higher production.

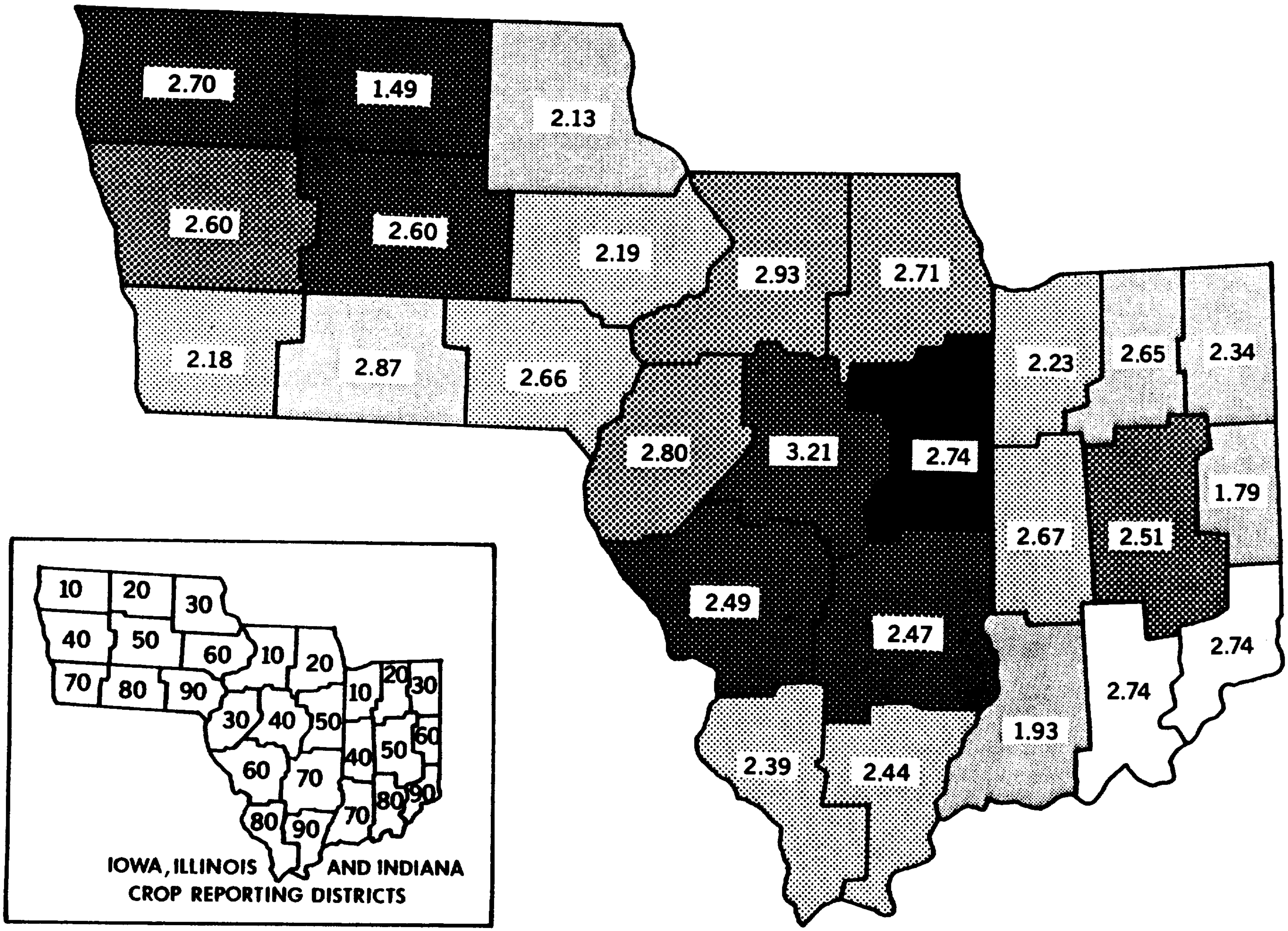


TABLE 3
INDICATORS OF YIELD RELIABILITY
BASED ON RD = 100 * ((PREDICTED-ACTUAL YIELD)/ACTUAL YIELD)

STRAW MAN MODEL 1 - SOYBEANS
IOWA, ILLINOIS, INDIANA

STATE	CRD	PERCENT OF YEARS |RD|>10%	LARGEST |RD| RD (YEAR)	NEXT LARGEST	SMALLEST |RD|	RANGE |RD|
IOWA	10	60	-17.4 (1972)	17.0	1.2	16.2
	20	10	19.6 (1974)	7.4	-0.4	19.1
	30	40	15.8 (1974)	-14.4	-1.0	14.9
	40	40	25.4 (1976)	16.5	0.0	25.4
	50	40	28.0 (1974)	16.2	0.4	27.5
	60	20	27.1 (1974)	11.9	0.4	26.7
	70	30	26.5 (1974)	13.1	-1.3	25.2
	80	20	57.4 (1974)	-12.7	2.0	55.4
	90	20	38.1 (1974)	-12.4	-0.4	37.7
STATE MODEL		20	23.9 (1974)	12.5	-0.4	23.5
CRDS AGGR.		20	23.9 (1974)	12.5	0.0	23.9
ILLINOIS	10	40	41.2 (1974)	-12.8	-0.4	40.8
	20	40	26.5 (1974)	-16.6	0.0	26.5
	30	20	42.8 (1974)	-10.3	2.1	40.6
	40	10	52.7 (1974)	-9.6	0.8	51.9
	50	40	36.9 (1974)	-12.5	1.2	35.7
	60	20	34.3 (1974)	-10.3	-3.3	31.0
	70	40	36.8 (1974)	-12.4	0.0	36.8
	80	50	26.6 (1974)	-13.9	-1.1	25.5
	90	50	26.6 (1974)	19.9	-1.7	25.0
STATE MODEL		20	37.6 (1974)	-10.9	1.9	35.7
CRDS AGGR.		30	37.0 (1974)	-11.7	1.9	35.0
INDIANA	10	40	25.4 (1974)	-13.4	1.3	24.1
	20	40	31.0 (1974)	-15.5	2.0	29.0
	30	30	24.5 (1974)	-18.9	-0.5	24.1
	40	20	48.1 (1974)	-12.9	-0.4	47.6
	50	30	28.2 (1974)	-13.0	-2.7	25.5
	60	20	18.6 (1974)	-10.8	-1.0	17.7
	70	20	22.6 (1974)	16.7	0.5	22.1
	80	40	33.8 (1975)	27.6	0.0	33.8
	90	40	33.1 (1973)	30.4	3.0	30.1
STATE MODEL		20	28.0 (1974)	-12.0	0.0	28.0
CRDS AGGR.		20	28.0 (1974)	-11.6	-0.5	27.5
REGION MODEL		10	29.9 (1974)	-8.1	-0.9	28.9
CRDS AGGR.		10	29.9 (1974)	-8.5	-0.9	28.9
STATES AGGR.		10	30.5 (1974)	-8.1	-0.9	29.5

Figure 3. Percent of test years (1970-1979) the absolute value of the relative difference is greater than ten percent for soybeans. Darker shades indicate CRDs with higher production.

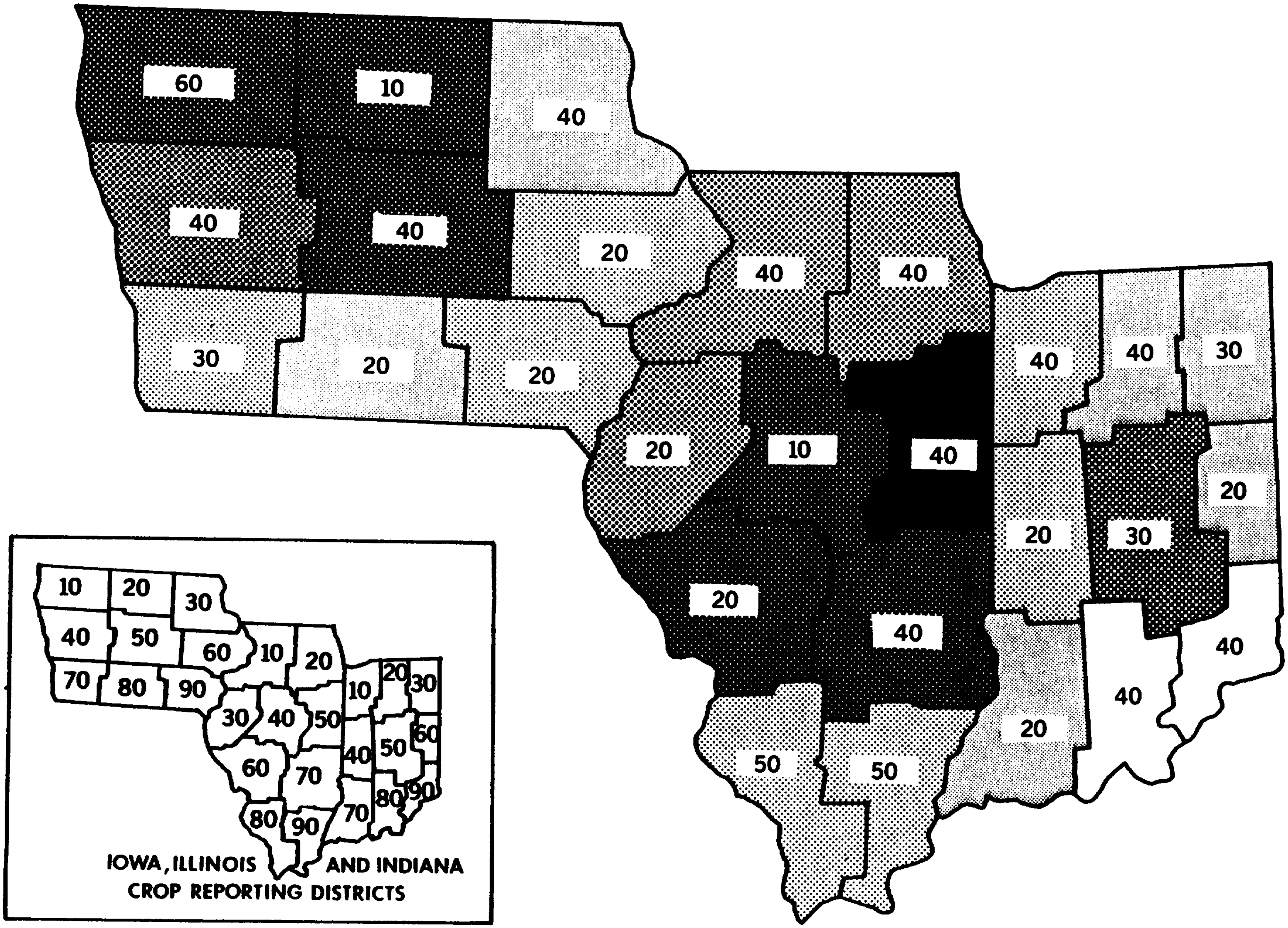


Figure 4. Largest absolute value of the relative difference for soybeans during the test years 1970-1979. Darker shades indicate CRDs with higher production.

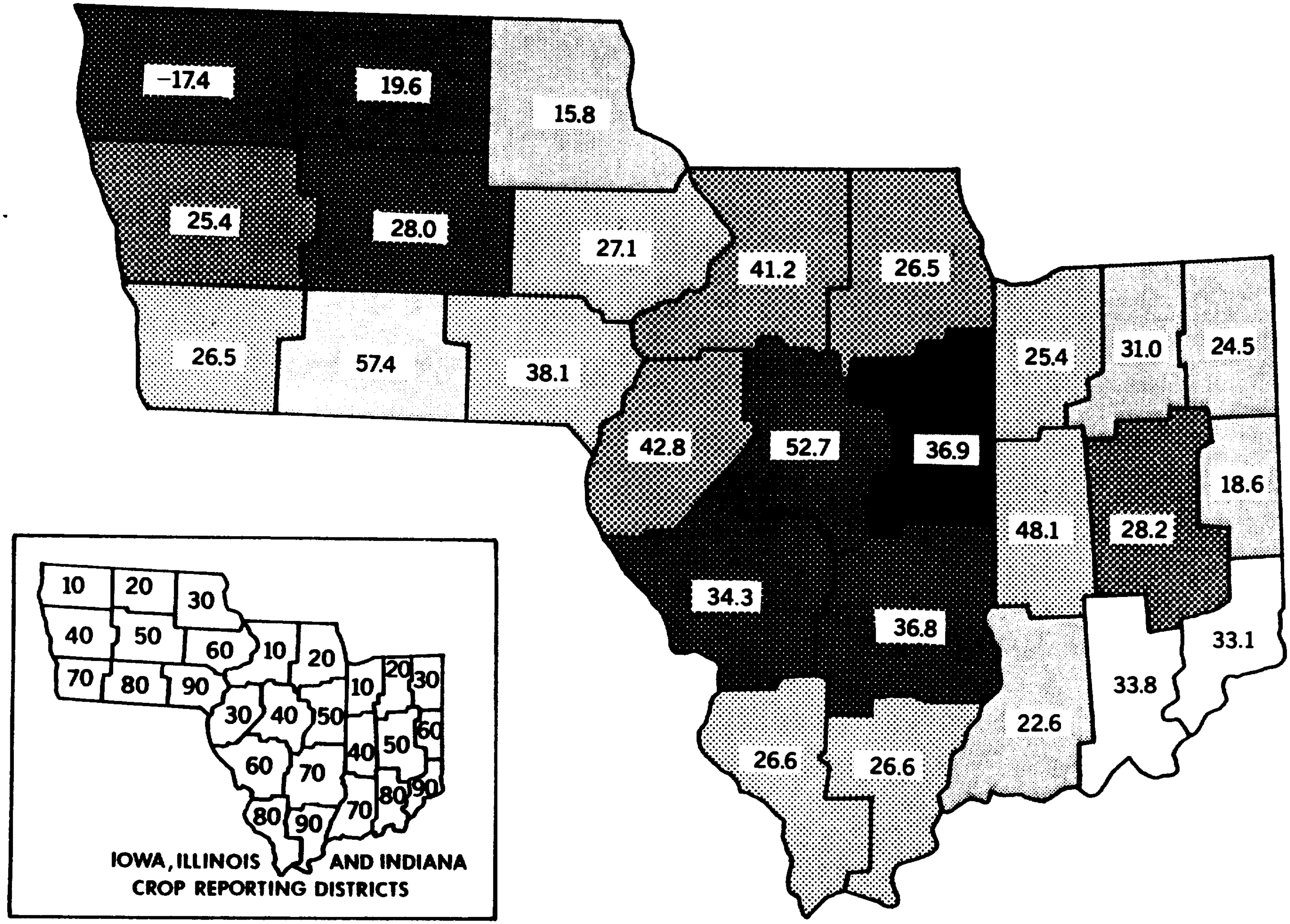
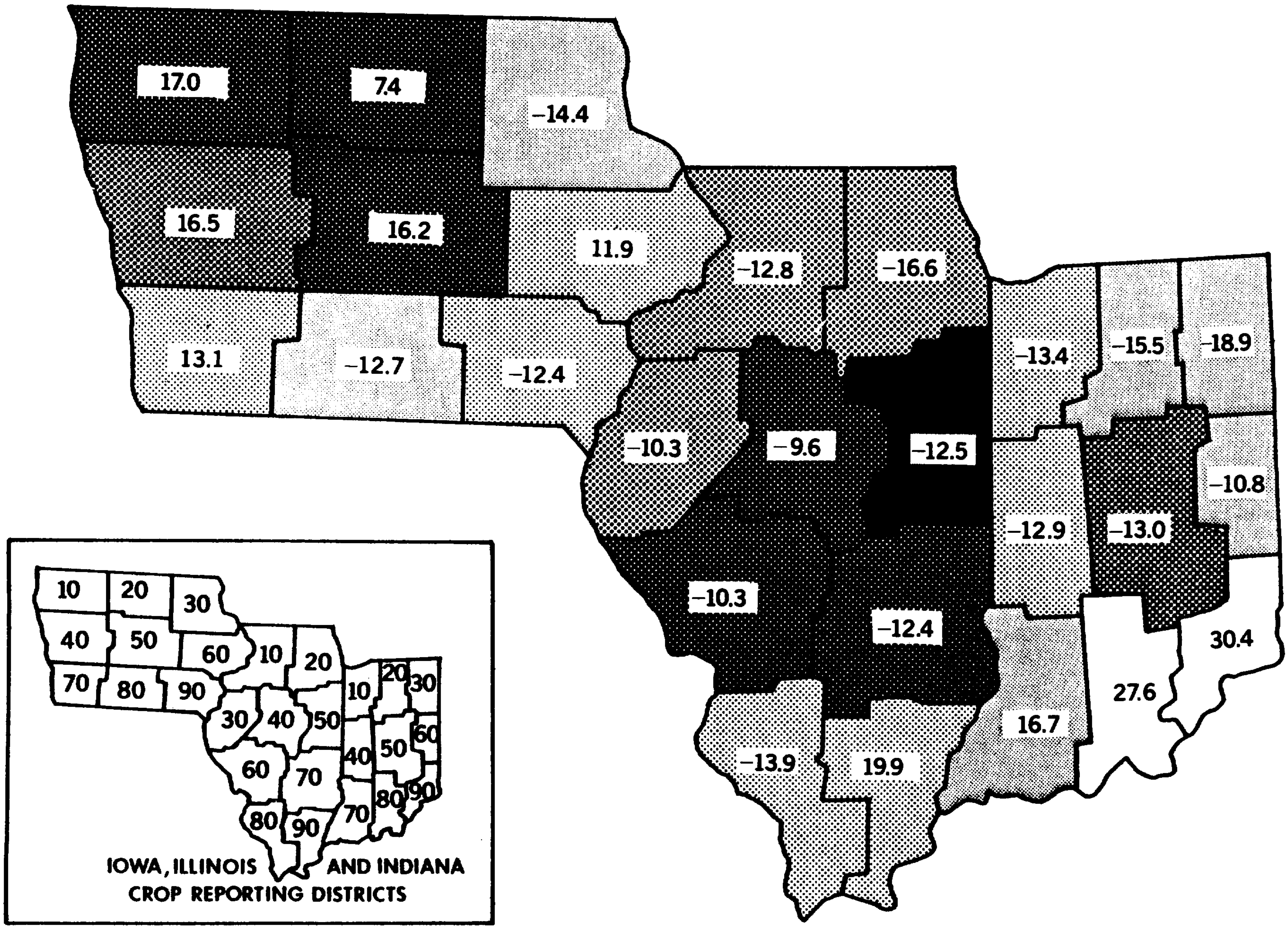


Figure 5. Next largest absolute value of the relative difference for soybeans during the test years 1970-1979. Darker shades indicate CRDs with higher production.



Indicators of Yield Reliability Based On Y and $\hat{Y}$ Show Low Correspondence Between the Direction of Change in Predicted as Compared to Actual Yields

Plots of the actual and predicted yields over the ten-year test period using state level models are displayed in Figures 6, 7 and 8. The CRD, state and region values for indicators of yield reliability based directly on actual and predicted yields are given in Table 4. CRD values are also shown in Figures 9-11.

The results for this model are poor. In only 3 out of 27 CRDs does the change in direction of predicted yields agree with the change in direction of actual yields from the previous year in over half of the test years (Figure 9). When the direction of change is based on an average of the three previous years, the direction of change is in agreement over half the time in only 10 of the 27 CRDs (Figure 10). Results are not much better at the state or region level. The Pearson r is negative for five of the CRDs (Figure 11). The largest positive r is 0.53. State and region results are not much better. This indicates that the model does a poor job of predicting high and low yields.

Change of predicted yield from previous forecasts within the current year cannot be investigated with a straw man model since the prediction for the current year only requires the addition of the actual yield for the previous year. No additional forecasts are made during the growing season unless more accurate figures for yield in previous years become available.

Base Period Indicates More Precision Than Independent Tests Can Confirm

Certain statistics generated from the regression analysis of the base period data are often used to provide some indication of expected yield reliability. However, these statistics only reflect how well the model describes the data used to generate the model, i.e., fit of the model, rather than how well the model can predict given new data. Therefore, it is important to compare these indicators of fit of the model to the independent indicators of yield reliability discussed in the preceding sections. In this way, one can see how these base period indicators of fit of the model do or do not correspond to independent test indicators of yield reliability.

One indicator of yield reliability, the mean square error (MSE), is the sum of squared d values ($d = \hat{Y} - Y$) for the independent test years divided by the number of test years (Table 2). The direct analogue for the model development base period is the residual mean square. The residual mean square is obtained by first generating the usual least squares prediction equation using the base period years. Then instead of predicting the yield for the following test year, yields are predicted for each of the base period years. The residual mean square is the sum of squared d values for these base period years divided by the appropriate degrees of freedom (number of years minus number of parameters estimated in fitting the model). Whereas one value of MSE is generated for each geographic area over the entire test period, a value of the residual mean square is generated for each base period corresponding to a test year in that area. The low, high, and average of the base period values for each area are given in Table 5.

FIGURE 6

IOWA
State Model

Actual and Predicted Yields for
the Test Years 1970-1979

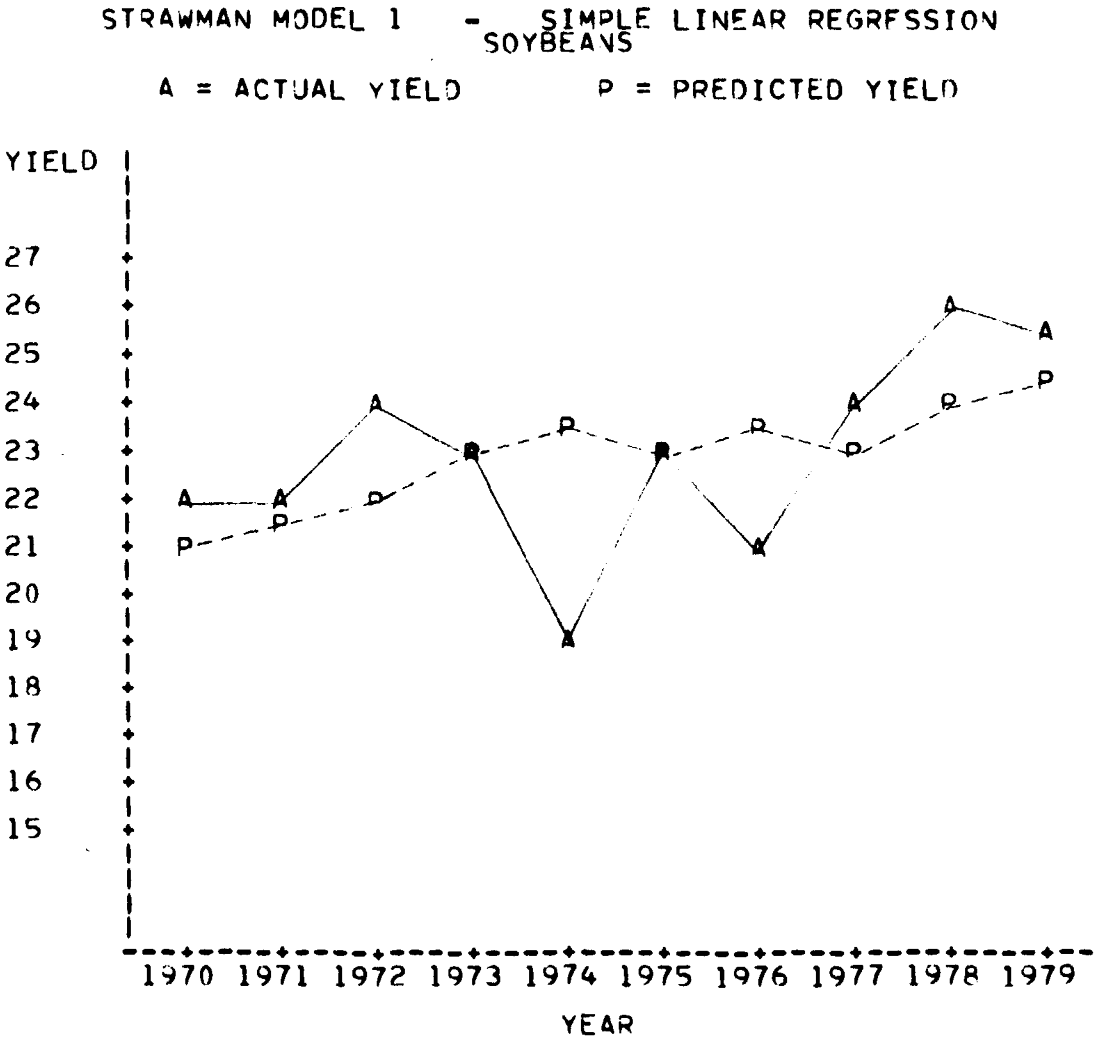


FIGURE 7

ILLINOIS
State Model

Actual and Predicted Yields for the Test Years 1970-1979

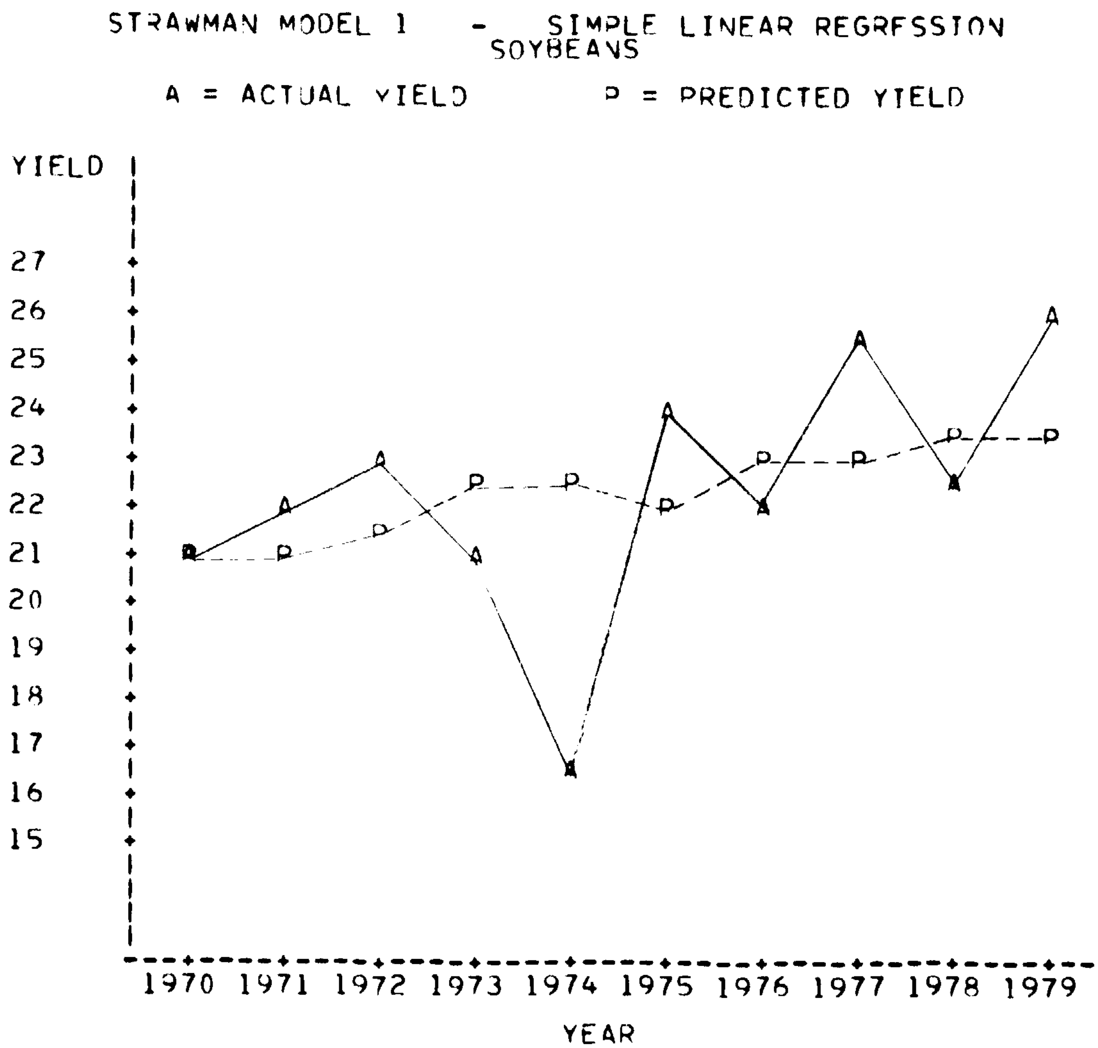


FIGURE 8

INDIANA
State Model

Actual and Predicted Yields for the Test Years 1970-1979

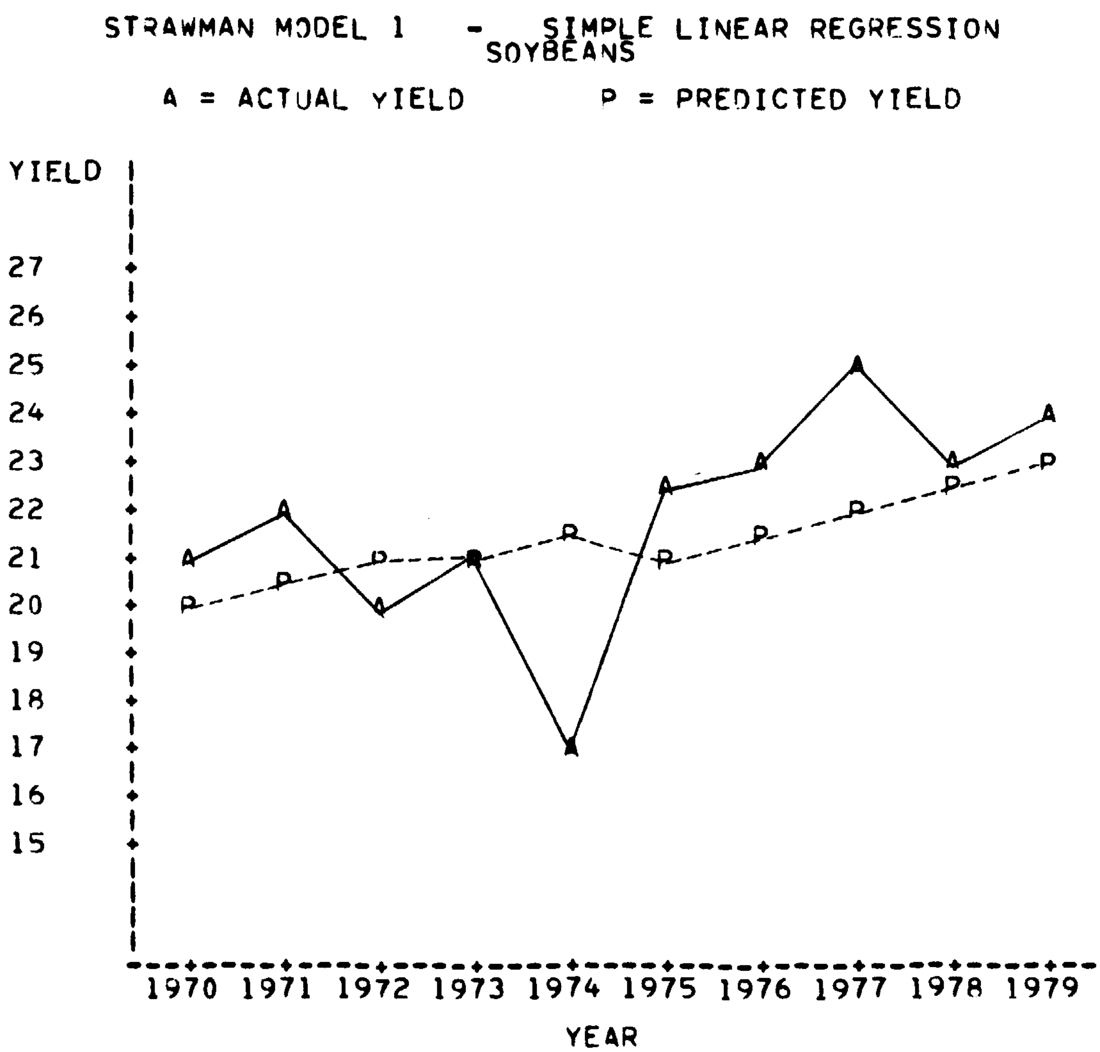


TABLE 4
INDICATORS OF YIELD RELIABILITY
BASED ON ACTUAL AND PREDICTED YIELDS

STRAW MAN MODEL 1 - SOYBEANS
IOWA, ILLINOIS, INDIANA

STATE	CRD	PERCENT OF YEARS DIRECTION OF CHANGE IS CORRECT FROM PREVIOUS YEAR	FROM BASE PERIOD	PEARSON CORR. COEF.
IOWA	10	22	71	0.46
	20	22	71	0.37
	30	33	57	0.51
	40	22	57	0.38
	50	22	43	-0.12
	60	33	43	0.25
	70	22	43	-0.43
	80	56	29	-0.49
	90	44	57	0.13
STATE MODEL		22	71	0.34
CRDS AGGR.		22	71	0.33
ILLINOIS	10	22	29	0.14
	20	11	43	0.28
	30	33	43	0.04
	40	33	43	0.05
	50	33	43	0.08
	60	11	43	0.23
	70	44	43	0.48
	80	11	29	0.30
	90	22	29	0.07
STATE MODEL		11	43	0.20
CRDS AGGR.		11	43	0.19
INDIANA	10	44	43	0.36
	20	33	57	0.30
	30	56	57	0.48
	40	56	57	0.18
	50	44	57	0.47
	60	33	71	0.44
	70	33	29	0.53
	80	22	29	-0.74
	90	33	0	-0.77
STATE MODEL		56	57	0.44
CRDS AGGR.		56	57	0.45
REGION MODEL		44	57	0.38
CRDS AGGR.		33	57	0.35
STATES AGGR.		33	57	0.33

Figure 9. Percent of test years (1970-1979) the direction of change in predicted yield from the previous year agrees with the direction of change in actual soybean yield. Darker shades indicate CRDs with higher production.

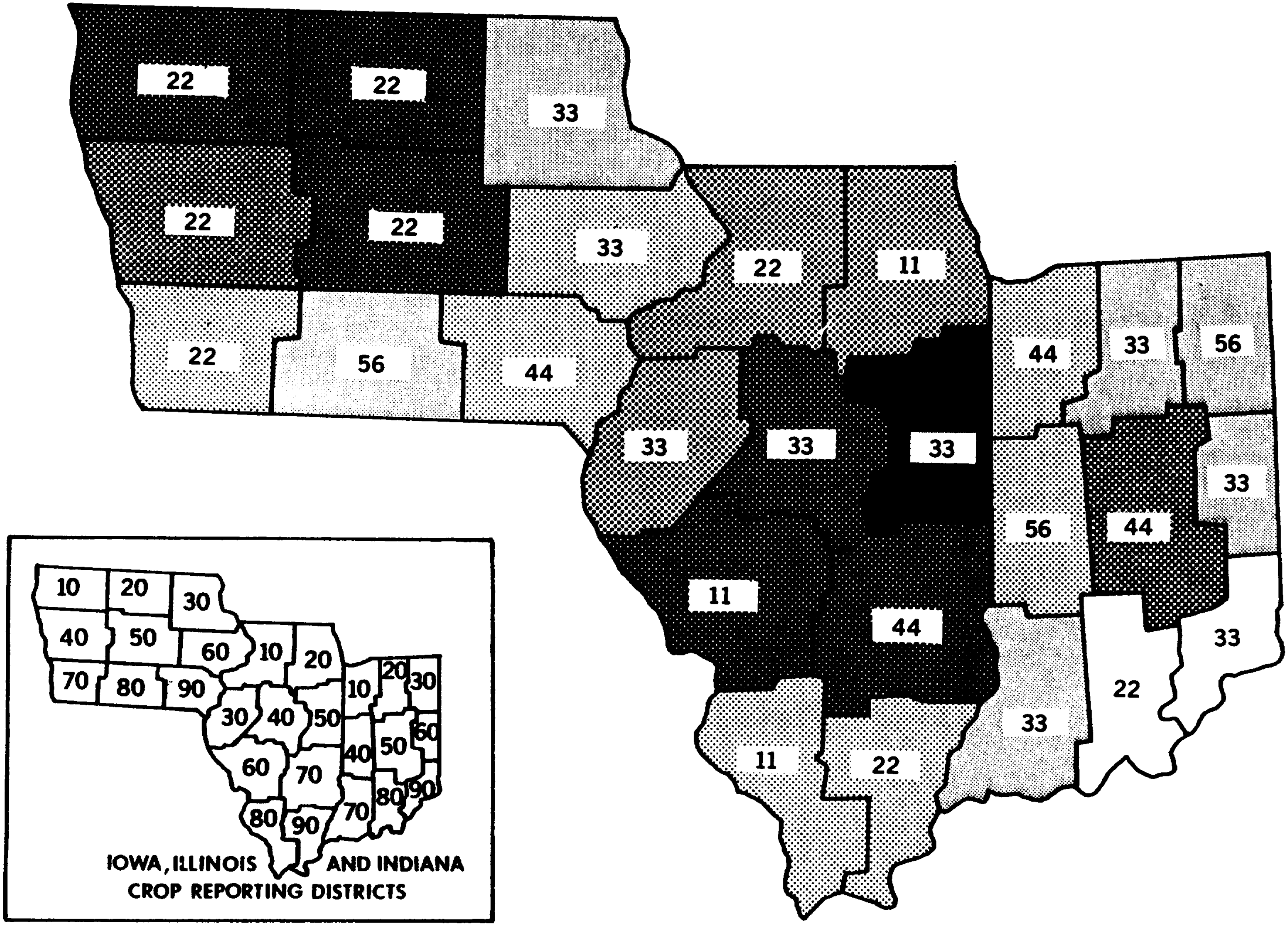


Figure 10. Percent of test years (1970-1979) the direction of change in predicted yield from the previous three year average agrees with the direction of change in actual soybean yield. Darker shades indicate CRDs with higher production.

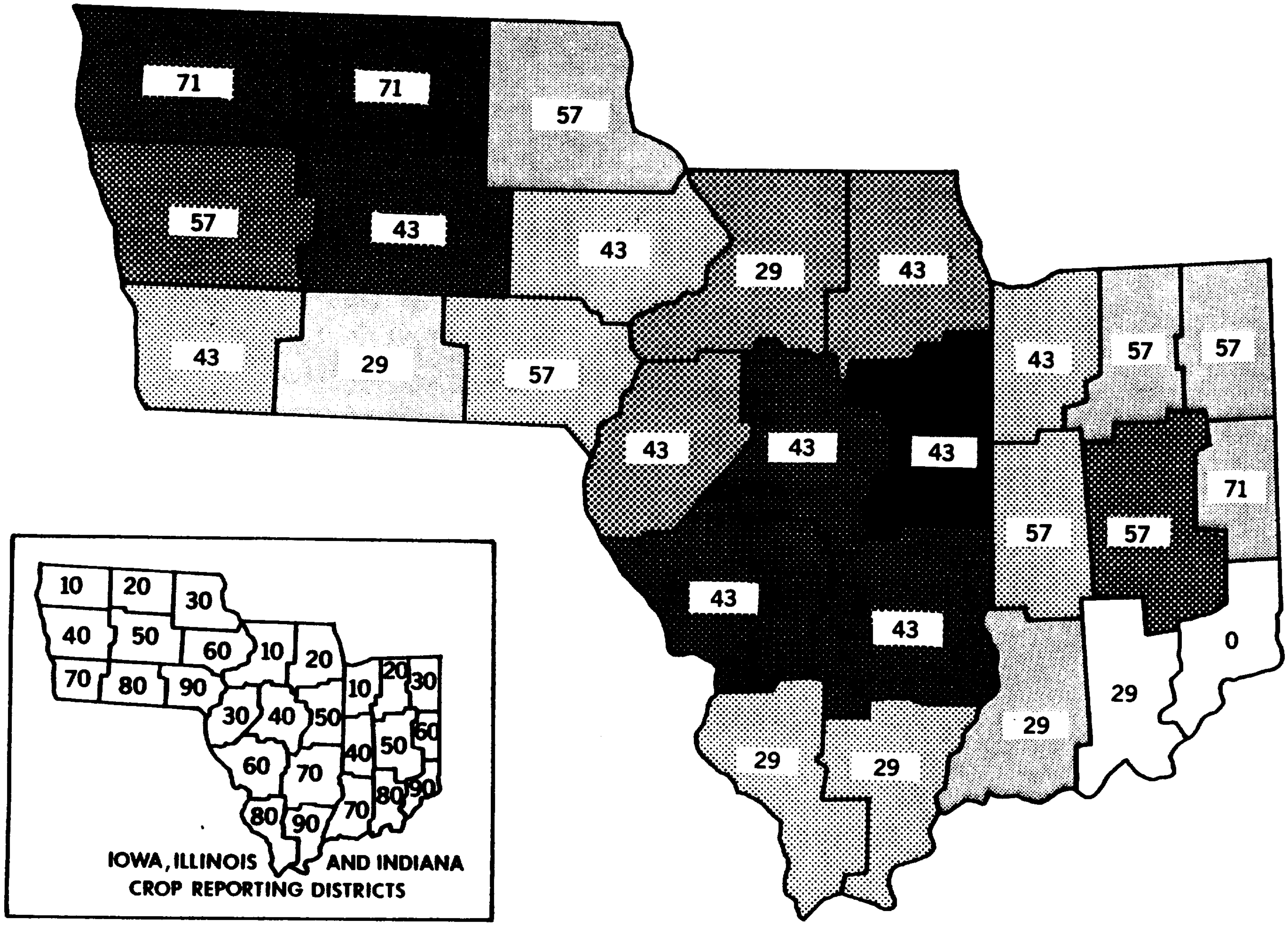


Figure 11. Pearson correlation coefficient between actual and predicted soybean yields in the test years (1970-1979). Darker shades indicate CRDs with higher production.

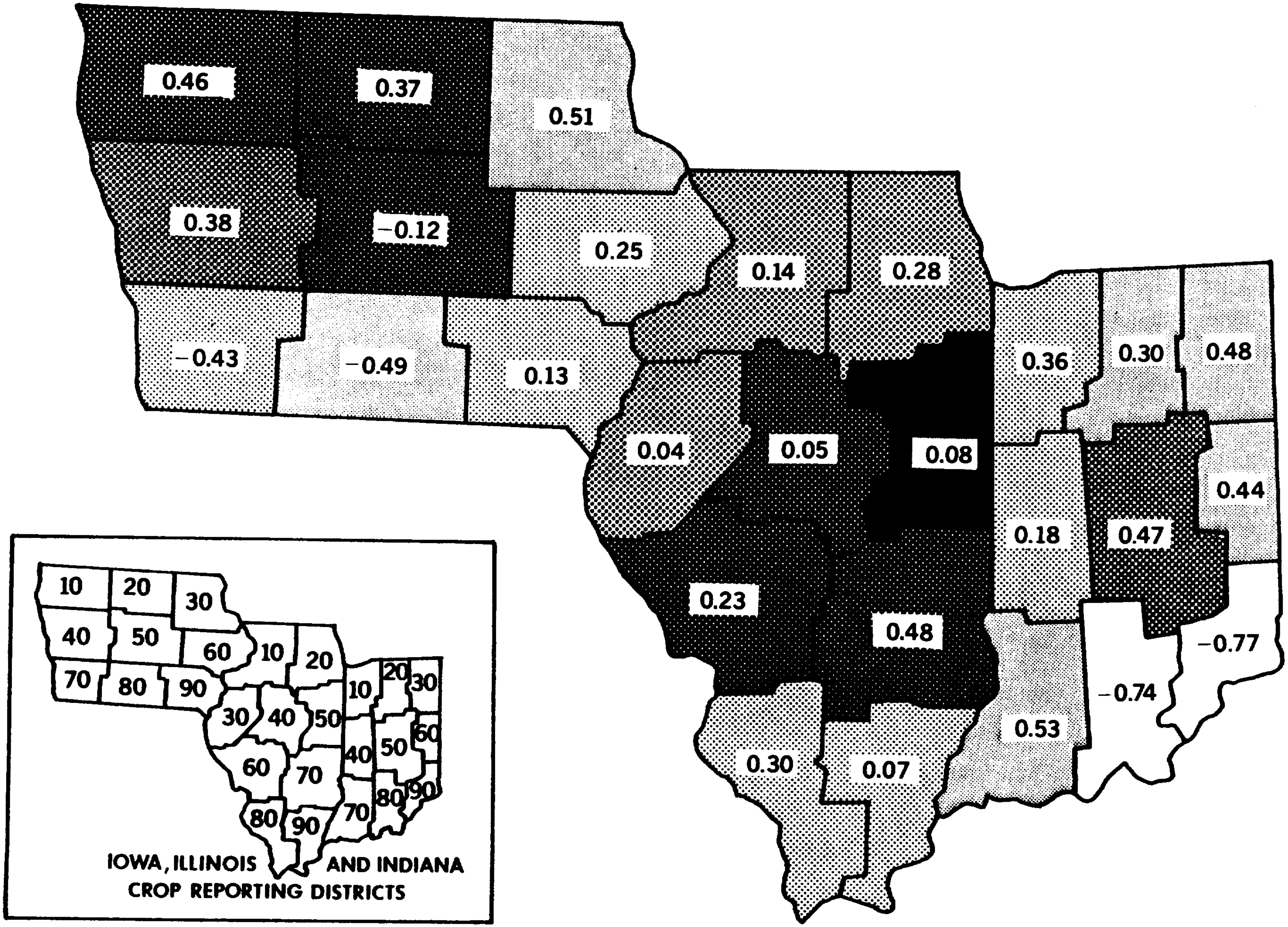


TABLE 5
RESIDUAL MEAN SQUARE AS AN
INDICATOR OF THE FIT OF THE MODEL
BASED ON THE MODEL DEVELOPMENT BASE PERIOD

STRAW MAN MODEL 1 - SOYBEANS
IOWA, ILLINOIS, INDIANA

STATE	CRD	BASE PERIOD RESIDUAL MEAN SQUARE LOW	HIGH	AVERAGE	INDEPENDENT TEST MSE
IOWA	10	4.49	5.91	5.24	7.29
	20	1.78	2.18	2.02	2.21
	30	1.75	2.56	2.12	4.52
	40	4.50	5.39	4.87	6.77
	50	2.75	4.20	3.40	6.76
	60	1.50	2.77	2.10	4.80
	70	3.05	4.36	3.94	4.77
	80	3.60	5.94	4.61	8.22
	90	2.61	4.38	3.33	7.08
STATE MODEL		2.07	2.89	2.45	3.79
ILLINOIS	10	1.31	3.84	2.57	8.56
	20	1.56	3.57	2.44	7.35
	30	1.79	3.71	2.91	7.86
	40	1.90	5.23	3.56	10.32
	50	1.67	4.08	2.87	7.52
	60	3.36	4.65	3.91	6.21
	70	3.59	4.96	4.24	6.11
	80	3.57	6.83	5.60	5.71
	90	3.33	4.77	4.09	5.94
STATE MODEL		1.83	3.42	2.65	6.32
INDIANA	10	1.93	3.09	2.51	4.98
	20	2.12	4.25	3.04	7.01
	30	2.40	3.75	2.94	5.48
	40	2.17	4.53	3.35	7.13
	50	2.86	5.14	3.85	6.30
	60	3.32	3.87	3.65	3.21
	70	2.11	3.25	2.63	3.73
	80	1.81	4.53	2.95	7.53
	90	2.07	4.19	3.26	7.50
STATE MODEL		1.55	2.92	2.16	4.19
REGION MODEL		1.00	2.10	1.51	3.85

The MSE values in Table 2 are also given in Table 5. The average residual mean square is almost always less than the MSE, many times being less than half as large. In fact, the largest residual mean square is almost always less than the MSE. Therefore, the results from the independent test indicate less reliability than one might have expected from base period model development results.

Another indicator of yield reliability is the correlation coefficient, r, between the observed and predicted yields for the independent test years (Table 4). It is desirable for r to be close to +1, even though it can be negative. The analogue for the model development base period is the square root of R^2, the coefficient of multiple determination. The square root of R^2 expressed as a proportion, R ($0 \leq R \leq 1$), may be interpreted as the correlation between observed and predicted values for the base period years. The low, high, and average values of R for each geographic area are given in Table 6.

The Pearson correlation coefficient values in Table 4 are also given in Table 6. The highest positive value of r is 0.53 and some r values for negative. Average CRD values of R range from 0.57 to 0.88. The values of r from the independent tests are certainly much lower than the values of R from the base period. It is obvious that levels of R (or alternatively R^2) for a model development base period are of no value in indicating independent performance of this model. In fact, the base period R or R^2 can be very misleading.

Model is Objective

Since the single independent variable is objectively defined (year minus 1950), no subjective inputs are required to run the model.

Results might differ if the set of years used to generate the models were changed. In this evaluation, the post World War II period was used, resulting in a maximum of twenty-three years on which to base the model (1947-1969, 1948-1970, etc.). Iowa had some slightly shorter periods since comparably defined yields were not available until 1950.

Once the decisions on the time period to use for model development and on the regression method to use (least squares) are fixed, the operation becomes completely objective.

Model Does Not Consider Known Scientific Relationships

The straw man models do not consider factors which have a recognized causal relationship with crop yields. For example, it is well known that year-to-year variations in weather have an important effect on yield. Therefore, if weather data were available, it would be consistent with scientific knowledge to include weather variables in a model predicting crop yields. Weather variables are excluded from the straw man models yet nothing is done to account for the fact that the yields have been influenced by weather. The yields may also have been influenced by other non-technology factors. However, since no adjustment is made to the yields for these non-technology factors and since these factors are not included as independent variables in the model, the straw man model results will be affected by non-technology influences. As was anticipated, the calculated slope of the regression line for straw man model 1 is positive for each base period at each geographic location.

TABLE 6
CORRELATION BETWEEN OBSERVED AND PREDICTED YIELDS AS AN INDICATOR OF THE FIT OF THE MODEL BASED ON THE MODEL DEVELOPMENT BASE PERIOD

STRAW MAN MODEL 1 - SOYBEANS
IOWA, ILLINOIS, INDIANA

TEST STATE	CRD	BASE PERIOD CORRELATION COEF. LOW	HIGH	AVERAGE	INDEPENDENT CORR. COEF.
IOWA	10	0.66	0.79	0.73	0.46
	20	0.85	0.91	0.88	0.37
	30	0.85	0.90	0.87	0.51
	40	0.67	0.77	0.71	0.38
	50	0.71	0.85	0.78	-0.12
	60	0.83	0.93	0.88	0.25
	70	0.67	0.84	0.76	-0.43
	80	0.53	0.76	0.63	-0.49
	90	0.68	0.82	0.75	0.13
STATE MODEL		0.79	0.88	0.84	0.34
ILLINOIS	10	0.64	0.87	0.75	0.14
	20	0.51	0.72	0.62	0.28
	30	0.68	0.85	0.75	0.04
	40	0.64	0.86	0.74	0.05
	50	0.69	0.84	0.76	0.08
	60	0.69	0.80	0.74	0.23
	70	0.66	0.77	0.71	0.48
	80	0.60	0.74	0.68	0.30
	90	0.47	0.71	0.63	0.07
STATE MODEL		0.71	0.86	0.78	0.20
INDIANA	10	0.71	0.82	0.77	0.36
	20	0.55	0.71	0.64	0.30
	30	0.46	0.69	0.57	0.48
	40	0.71	0.86	0.79	0.18
	50	0.57	0.70	0.65	0.47
	60	0.56	0.74	0.65	0.44
	70	0.82	0.90	0.85	0.53
	80	0.60	0.92	0.78	-0.74
	90	0.66	0.92	0.80	-0.77
STATE MODEL		0.73	0.85	0.80	0.44
REGION MODEL		0.82	0.92	0.86	0.38

The assumption of straw man model 1--simple linear regression--is also open to question. That is, it may not be reasonable to expect that the rate of change in yields has stayed constant over the model development base period. Although technology has resulted in increasing yields, the rate of increase may be greater or less over different portions of the time period. Also, it is known that contributions to yield from technology may deviate widely from the trend in any given year. For example, temporary fuel shortages may decrease potential benefits from fertilizer/herbicide/pesticide applications.

Model is Adequate

The model can provide estimates for any geographic area having historic yield information. This basis information would be required for any modeling effort. Therefore, straw man model 1 is at least as adequate as any other model would be.

Model is Timely

As soon as reliable figures are available for this year's yield, the model can be developed for estimating next year's yield.

Model is Not Costly

The only data required are the year and actual yield. These data are readily available at no additional cost. The least squares, simple linear regression model can be fit using any standard statistical packaged program or statistical calculator.

Model is Simple

The model is simple. Users can clearly understand the basis for predicted yields. The model is easy to use. The X values in the model are simply the year minus 1950. Thus to estimate the yield for 1980, multiply the slope by 30 and add the intercept from a model developed using years 1957-1979.

Model Has Poor Current Measure of Modeled Yield Reliability

The CRD, state, and region values for the Spearman correlation coefficient between the estimate of the standard error of a predicted yield value and the absolute value of the difference between the predicted and actual yield are computed. They are given in Table 7. The CRD correlation coefficient values are displayed in Figure 12. In 22 of 27 CRDs, the correlation is negative. The largest positive value is 0.31. Thus, the model does not provide a good measure as to how close the predicted values will be to the actual values. Instances of years with smaller confidence intervals about the true predicted value are all too often associated with larger observed discrepancies between the actual and predicted values. The accuracy of a predicted yield cannot be reliably judged using information provided by the model.

The value of the standard error of a predicted yield is a function of the residual mean square and the distance of the independent variable values in the prediction year from their average during the base period. Since the distance value is constant over the independent test years for the straw man model, the variability in the standard error is simply a function of the size of the residual mean square. As can be seen from Figures 6, 7 and 8, years with larger differences between predicted and actual yields, which increase the value of the residual mean square, alternate with years having smaller differences. Therefore, the above results are not surprising.

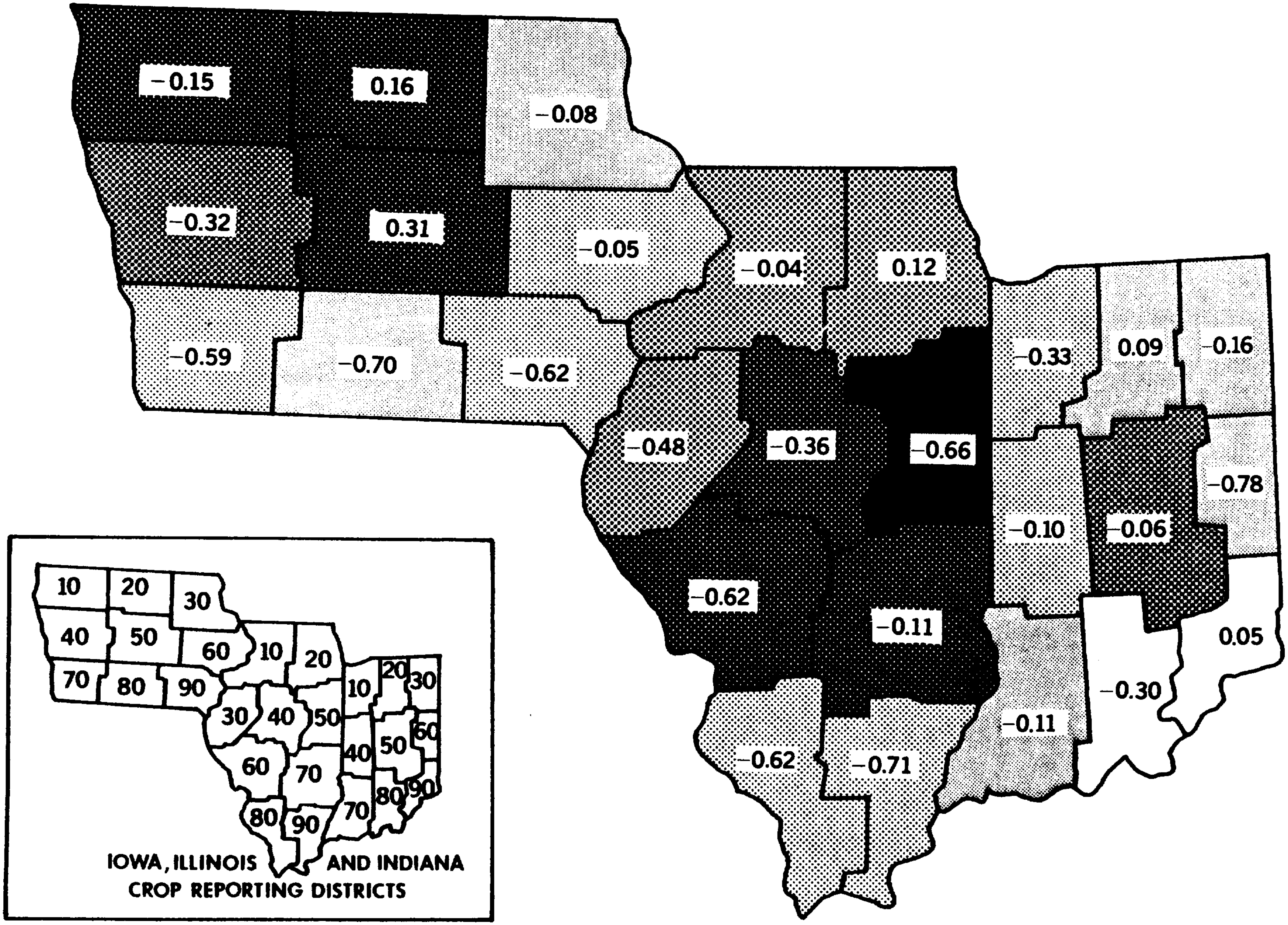
TABLE 7
CURRENT INDICATION OF
MODELED YIELD RELIABILITY

AGREEMENT BETWEEN BASE PERIOD PREDICTED
AND TEST YEAR ACTUAL ACCURACY

STRAW MAN MODEL 1 - SOYBEANS
IOWA, ILLINOIS, INDIANA

STATE	CRD	SPEARMAN CORRELATION COEF.
IOWA	10	-0.15
	20	0.16
	30	-0.08
	40	-0.32
	50	0.31
	60	-0.05
	70	-0.59
	80	-0.70
	90	-0.62
STATE MODEL		-0.28
CRDS AGGR.		-0.21
ILLINOIS	10	-0.04
	20	0.12
	30	-0.48
	40	-0.36
	50	-0.66
	60	-0.62
	70	-0.11
	80	-0.62
	90	-0.71
STATE MODEL		-0.30
CRDS AGGR.		-0.15
INDIANA	10	-0.33
	20	0.09
	30	-0.16
	40	-0.10
	50	-0.06
	60	-0.78
	70	-0.11
	80	-0.30
	90	0.05
STATE MODEL		-0.21
CRDS AGGR.		0.02
REGION MODEL		-0.18
CRDS AGGR.		0.17
STATES AGGR.		0.02

Figure 12. Spearman correlation coefficient between the estimate of the standard error of a predicted value from the base period model and the absolute value of the difference between the predicted and actual soybean yield in the test years (1970-79). Darker shades indicate CRDs with higher production.



CONCLUSIONS

Straw man model 1, simple linear regression of yield over time, describes a uniform increase in soybean yields over time. Indicators of yield reliability obtained from bootstrap testing are used as a basis of comparison between competing models and the results for straw man model 1 do not appear very promising. The bias is generally small, however, the model is unable to predict the low and high yields accurately. The model is objective, adequate, timely, simple, and not costly. However, it does not consider known scientific relationships and does not provide a good current measure of modeled yield reliability.

In conclusion, as expected, straw man model 1 is truly a "below base" model. Competing models, requiring additional inputs, will certainly be less simple and more costly, will probably be less timely, and will possibly be less adequate and objective. However, it is hoped that these models will provide more accurate indicators of yield reliability and current measures of modeled yield reliability.

REFERENCES

KESTLE, RICHARD A., 1981. Analysis of Crop Yield Trends and Development of Simple Corn and Soybean "Straw Man" Models for Indiana, Illinois, and Iowa. AgRISTARS Yield Model Development Project, Document YMD-2-11-1 (80-11.1). ESS Staff Report AGESS810114.

WILSON, WENDELL W., BARNETT, THOMAS L., LeDUC, SHARON K., WARREN, FRED B., 1980. Crop Yield Model Test and Evaluation Criteria. AgRISTARS Yield Model Development Project, Document YMD-1-1-2 (80-2.1).