United States
Department of
Agriculture

National
Agricultural
Statistics
Service

Research and
Applications
Division

NASS Staff Report
Number SRB-87-05

August 1987

The Feyerherm Winter Wheat Model: A Performance Report for Eight Objective Yield States

Fatu G. Bigsby

THE FEYERHERM WINTER WHEAT MODEL: A PERFORMANCE REPORT FOR EIGHT OBJECTIVE YIELD STATES by Fatu G. Bigsby, Research and Applications Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C. 20250, August, 1987. Staff Report No. SRB-87-05.

ABSTRACT

This study examines the contribution from the Feyerherm winter wheat model in predicting May 1 to July 1 winter wheat yield in eight Objective Yield Survey states from 1977 to 1983. The Feyerherm winter wheat model is composed of a trend term to account for improved technological effects and weather variables derived from daily weather data. Forecast errors from three different composite forecasts are compared: NASS yield indications only, NASS indications plus a long-term trend forecast, and NASS indications plus the Feyerherm forecast for the month. Results show no significant differences among the forecast errors for the three types of composite forecasts. Therefore it is recommended that the Feyerherm model not be used in the operational program of NASS.

ACKNOWLEDGMENTS

The author would like to thank Benjamin Klugh and Paul Cook for their helpful input and comments on this paper and to all NASS State Statistical Offices that contributed to the data used in the analyses. Prior work on the Feyerherm model by staff of the Yield Evaluation Section made it possible to run the model on a real-time basis in Kansas, Nebraska and Oklahoma and to simulate the performance of the model for these states and the other states included in this report. Typing of the final copy of this report by Jennifer Kotch is appreciated.

Washington, D. C. August, 1987

TABLE OF CONTENTS

SUMMARY

Based on comparative analyses in eight winter wheat Objective Yield Survey states from 1977 to 1983, the Feyerherm wheat model was found to provide no significant improvement in forecasting ability over current yield indications of the National Agricultural Statistics Service(NASS) and a long-term trend model. The analyses were based on composite estimation techniques and nonparametric methods. Composite forecasts were developed using NASS monthly farm report indications, objective yield indications, the Feyerherm model forecasts, and a long-term trend.

No significant differences were detected between composite forecasts when forecasts from the Feyerherm winter wheat model were excluded from or included with yield indications from NASS. No significant differences were detected between errors for the forecasts from a long-term trend model which uses no weather data and forecasts from the Feyerherm model based on daily weather data. It is therefore recommended that the Feyerherm model not be used by NASS.

THE FEYERHERM WINTER WHEAT MODEL: A PERFORMANCE REPORT FOR EIGHT OBJECTIVE YIELD STATES

Fatu G. Bigsby[1]

INTRODUCTION

This study assesses the usefulness of the Feyerherm winter wheat model in improving the May 1 to July 1 winter wheat yield forecasts in eight Objective Yield Survey states. It compares the accuracy of three different composite forecasts in predicting winter wheat yields for the period 1977 to 1983. The first type of composite forecast consists of two to four NASS (National Agricultural Statistics Service of the USDA) indications[2]. A second composite forecast is computed using NASS indications and the forecast derived from long-term trend which excludes weather information for the growing season. The last composite forecast includes NASS indications and the monthly Feyerherm forecast made during the growing season, using daily weather information. States included in the analyses are Colorado, Illinois, Indiana, Kansas, Nebraska, Montana, Ohio and Oklahoma.

The sections below present the procedures and the results of the analyses. First, a background of the study is given, which is followed by the methodologies used to obtain the Feyerherm forecasts and the composite forecasts. After methods for comparison of the forecasts are presented, conclusions and recommendations are made.

BACKGROUND

The '84 Feyerherm winter wheat model was developed by Dr. Arlin Feyerherm at Kansas State University for the improvement of winter wheat forecasts during the growing season, under contract with NASS . This model was preceded by the Feyerherm '82 winter wheat model [2]. Recommendations from analysis of the '82 model led to the development of the '84 model [4,5]. The '84 winter wheat model was used to obtain forecasts for Kansas, Nebraska and Oklahoma during the 1985 and 1986 crop years in the NASS operational program.

----

[1] The author is a mathematical statistician with the National Agricultural Statistics Service, U.S. Department of Agriculture, Washington D.C.

[2] An indication is a forecast of crop yield derived from a specific survey or model. One or more indications are used to forecast a monthly crop yield.

METHODOLOGIES

This section includes an outline of the procedures for generating Feyerherm model forecasts and composite forecasts. Methodology for examining the change in forecast errors when the Feyerherm forecast is combined with NASS indications to form composite forecasts is also included.

Feyerherm Forecasts

The Feyerherm model estimated for each Crop Reporting District (CRD) in a state is

$$\text{Yield} = \text{long term trend} + \text{weather influence} \qquad (1)$$

where

$$\text{long term trend} = a_0 + b_1(T) \qquad (1a)$$

and

$$\text{weather influence} = b_2(\text{DWYF}) \qquad (1b)$$

a_0, b_1, and b_2 are regression parameters;

T equals current year minus 1951 and is used to account for the influence of technology; DWYF equals the difference between the weather influence in the current year and a long-term average computed for the previous 30 years.

The base period used to estimate equation (1) is 1951 up to the year preceding the forecast year. The long-term trend, $a_0 + b_1(T)$, can be used to forecast yield at anytime during the year and is used to forecast yield at planting by Feyerherm. After the trend yield is computed, the second term in the model is used to determine the influence of weather on yield at the current stage of crop development. A regression model called the "weather-yield function" was estimated by Feyerherm using experiment station data to reflect the relationship between yield and daily weather data during eight wheat growth stages. DWYF is obtained by subtracting the current year's weather-yield function (WYF) value from the mean of the base period. The value of the weather-yield function for the current year is calculated by including all WYF regression model terms associated with the current and all previous crop development stages. Appendix 1 presents a description of the model. The method used to obtain a yield forecast during a specific growth stage is given below.

To Forecast Yield at:	Model is Used as Follows:
Wintering	Yield=Trend + b_2(DWYF for planting to winter)
Spring green-up	Yield=Trend + b_2(DWYF for planting to spring)
Jointing	Yield=Trend + b_2(DWYF for planting to jointing)
Flag	Yield=Trend + b_2(DWYF for planting to flag leaf)
Heading	Yield=Trend + b_2(DWYF for planting to heading)
Milk	Yield=Trend + b_2(DWYF for planting to milk)
Dough	Yield=Trend + b_2(DWYF for planting to dough)
Ripe	Yield=Trend + b_2(DWYF for planting to ripe)

For example, to forecast yield at heading, portions of the weather-yield function pertaining to stages that occur after heading (heading to milk, milk to dough, and dough to ripe) are not used. Because of weather conditions unique to each state, the time or degree of occurrence of crop stages are not uniform across states. For example, winter wheat development in Kansas and Nebraska is likely to occur later than winter wheat development in Oklahoma. For this reason, the stage used to forecast for a state in a given month is the stage whose long-term average is closest to the forecast date. The same stage may be used for more than one forecast date. The ripe stage, for example, is used for both July 1 and August 1 forecasts in Kansas. The difference between the two forecasts is the availability of weather data for a longer period for the August 1 forecast, which should make it more accurate. Forecasts are obtained for the state by weighting the CRD forecasts to state level, using harvested acres as weights.

Composite Forecasts

Composite forecasts for 1977 to 1983 are used to determine if addition of the Feyerherm forecast to the indications used by NASS increases the accuracy of the May 1, June 1, and July 1 forecasts. The accuracy of each type of forecast for a given month is determined by computing its root mean-square error for the period, 1977 to 1983. The root mean-square error(in bushels per acre) is defined as

$$\text{RMSE}=\text{square root of } [1/n \sum_{i=1}^{n} (C_{jh} - \text{Final Board Yield}_h)^2] \qquad (2)$$

Where C_{jh} is the composite forecast for month j in year h and n is the number of years from 1977 to 1983=7.

NASS's final yield is considered the true yield for computing the root mean-square error. This assumption implies that the final yield published by the Agricultural Statistics Board is the best estimate using all data available. However the best estimate is not necessarily the "true universal value". Therefore this measure of performance is not a pure measure of mean-square error but provides one method of approximation. Both the forecast obtained from long-term trend (which uses no weather information) and the Feyerherm monthly forecast are analyzed. This is done to determine if the weather influence information supplied by DWYF improves the accuracy of the yield forecast.

For any given month from May through July, there are a maximum of four NASS indications which can be combined to obtain composite forecasts for winter wheat [8]. The NASS indications are condition-yield, mean locality yield, multiple regression yield, and objective yield. Respondents are asked to report the condition of their crops as compared to normal in a nonprobability survey. The value of condition reported in this survey and trend are regressed on final yield to derive the condition-yield indication. The mean locality yield indication is also derived from the same nonprobability survey in which respondents report the probable average yield in their localities. The physical characteristics of the crop, such as number of heads and head-weight are obtained from a probability survey to derive the objective-yield indication. The multiple regression indication is the forecast from a regression model in which yield is the dependent variable with monthly precipitation, trend, and condition of crop (early in the season) or probable yield (later in the season), as the independent variables.

The method used to obtain a composite forecast is given below:

Let X_{ijk} = the historic value of the ith indication (or model forecast) for the jth month in year k, k=1969,1970,...,year preceding forecast year

M_{ij} = the mean-square error of the ith indication for the jth month and for the period 1969 to year preceding forecast year

$$= 1/n \sum_{i=1}^{n} (X_{ijk} - \text{Final Board Yield}_k)^2$$

where n= number of years from 1969 to year preceding forecast year

W_{ij} = $1/M_{ij}$, the weight given to the ith indication for the jth month and the period 1969 to year preceding forecast year

Then the composite forecast which includes m indications for month j in forecast year h is:

$$C_{jh} = \sum_{i=1}^{m}(X_{ijh})(W_{ij}) \div \sum_{i=1}^{m} W_{ij} \qquad (3)$$

W_{ij} is defined above, and X_{ijh} is the current value of the ith indication for the jth month in forecast year h, h=1977,1978,...,1983

As shown above, a composite forecast for any year is obtained by using data for prior years to compute the mean-square errors of the indications which are used as weights. For example, since data for the NASS indications are available from 1969, composite forecasts for 1979 are obtained by using data for 1969 to 1978 to compute mean-square errors for weighting the indications.

Analyses of Forecasts

The root mean-square errors (RMSEs) of the composite forecasts are compared to determine their relative accuracies. As stated above in (2), the root mean-square error for the jth month is computed for the period 1977 to 1983 and is defined as:

$$\text{RMSE} = \text{square root of } [1/n \sum_{i=1}^{n} (C_{jh} - \text{Final Board Yield}_h)^2]$$

In making the comparisons, the primary question of interest is does the Feyerherm model provide information to improve early season yield forecasts by reducing forecast error. Examining the difference between the errors of the Feyerherm monthly forecast and the long-term trend forecast is one means of determining the value of the Feyerherm monthly forecast. Equation (1a) shows that the long-term trend forecast is $\{a_0 + b_1(\text{forecast year} -1951)\}$. The difference between the Feyerherm monthly forecast and the long-term trend forecast is the inclusion of weather information in the Feyerherm monthly forecast. Another means of determing the value of the Feyerherm forecast is to examine the difference in errors between composite forecasts that includes or excludes it. Both methods are used, individually or in combination, as shown below.

Three types of comparisons are made. One comparison examines the differences in the RMSEs when all NASS indications available for the month are combined with or without the long-term trend forecast or the Feyerherm forecast. The terminology "all NASS indications" used here and in the remainder of this report means the use of all four indications described earlier(condition-yield, mean locality yield, multiple regression yield, and objective yield).

The second comparison examines the RMSEs of the best composite forecasts that include a) only NASS indications, b) NASS indications and the long-term trend forecast, and c) NASS indications and the Feyerherm forecast for the month. The best forecast for a month is selected for each of the categories a, b, and c given above. For each category, the root mean-square errors of all possible forecasts for the month are compared and the forecast with the minimum root mean-square error is chosen.

In the third comparison, the magnitudes of the errors for the Feyerherm monthly forecasts and the long-term trend forecast are compared, independent of NASS indications.

In the first and second comparisons in which composite forecasts that include NASS indications are compared with those that exclude NASS indications, the hypothesis tested is:

$$H_0: r1=r2=r3 \quad\quad (4)$$
$$\text{versus} \quad H_a: \text{not all of the } r_i\text{'s are equal, } i=1,2,3$$

where r_1 is the RMSE of the forecast that contains only NASS indications, r_2 is the RMSE of the forecast that contains NASS indications and long-term trend, and r_3 is the RMSE of the forecast that contains NASS indications and the Feyerherm monthly forecast. The hypothesis given in (4) can be restated as follows,

$$H_0: a_i = 0 \text{ , } i=1,2,3 \quad\quad (5)$$
$$H_a: \text{not all of the } a_i\text{'s}=0$$

where a_i is defined in the following model which is used to test the null hypotheses, H_0:

$$Y_{ij} = u.. + a_i + b_j + e_{ij} \quad\quad (6)$$

where

- Y_{ij} is the value of the RMSE for the ith forecast and the jth state
- u.. is the overall mean
- a_i is the effect of the ith type of forecast
- b_j is the effect of the jth state
- e_{ij} is the deviation of Y_{ij} from its mean

Friedman's two-way nonparametric rank sums test [6] is performed to determine if there exists a significant difference among r_1, r_2 and r_3. A nonparametric test is performed instead of an analysis of variance test because the former does not require the assumption that the e_{ij}'s come from a Normal distribution, which may not hold for this analysis. The factors involved in the analysis are type of forecast and state. The forecast effect is of primary interest . The state effect is included in the analysis for efficiency, since it is unlikely that the performance of the forecasts is the same across states. The two-way model given in (6) has one observation per cell and it assumes no interaction between the forecast and state effects.

The Friedman rank sums test is

$$\text{Reject } H_0 \text{ if } S >= s(\alpha,k,n) \qquad (7)$$

where

$$S = (12n/k(k+1))*[\sum_{i=1}^{k}(R_{.i} - R_{..})^2]$$

k = number of types forecast
n = number of states
α = the probability of rejecting the null hypothesis when it is true (.05,in this study)
$R_{.i}$ = average of the ranks of the RMSEs for the ith forecast
$R_{..}$ = (average of ranks over all forecasts)/n

$s(\alpha,k,n)$ is the critical value of the test. It is the value of the test statistic S, for which the probability of rejecting the null hypothesis is equal to α. If there exists one or more ties among the ranks of the RMSEs of the forecasts for a state, the test statistic S is modified and is denoted by S':

$$S' = \frac{12\sum_{i=1}^{k}(R_i - nR_{..})^2}{nk(k+1) - [1/(k-1)]\sum_{i=1}^{n}\left\{\sum_{i=1}^{gj} t^3_{i,j}) - k\right\}}$$

In S', g_j = the number of tie groups in the jth state

t_{ij} = size of the ith tie group in the jth state

If the null hypothesis is true and the RMSEs are not significantly different then we would expect the differences between $R_{.i}$ and $R_{..}$ to be small, which would lead to a small value for S. Therefore the null hypothesis is rejected for large values of S. The null hypothesis is rejected if the probability of obtaining a larger value of S (p value) for a specified n and k is greater than α, given that the RMSEs are equal. In the case in which the rank sums test concludes that the RMSEs are not equal, a simultaneous multiple comparison procedure is used to determine which pairs are significantly different.

As stated earlier, the third comparison examines the differences between the Feyerherm monthly forecasts and the long-term trend forecast, independent of NASS indications. Of primary interest is the benefit from the weather-yield function in reducing the size of the forecast error. This is determined by using the trend forecast as a 'control' and comparing its root mean-square error against the root mean-square errors of the May 1, June 1 and July 1 Feyerherm forecasts simultaneously. The hypothesis tested is:

$$H_o: r_m = r_t \;, \; m=1,\ 2,\ 3 \qquad (8)$$

$$H_a: r_m \neq r_t \text{ for at least one } m$$

In the null hypothesis, r_m is the root mean-square error of each of the monthly forecasts and r_t is the root mean-square error of the forecast from the long-term trend; r_t is constant for any given year and therefore does not vary from month to month. A two-way nonparametric multiple comparison procedure which compares a control against treatments is conducted [6]. The test is based on a large sample approximation in which H_o is rejected for

$$|R_t - R_m| >= [|m|(\alpha,k-1,.5)]*[nk(k+1)/6]1/2 = 12.1 \qquad (9)$$

In (9), R_t and R_m are the sums of the ranks of the RMSEs for trend and for each of the monthly forecasts, respectively. $|m|(\alpha,k-1,.5)$ is the largest absolute value of k-1 standard normal random variables with a common correlation of .5 [3].

RESULTS

Results are given in this section for the three types of comparisons made. Results are first given for comparison of the differences in the RMSEs when all NASS indications available for a month are combined with or without the long-term trend or the Feyerherm forecast. The long-term trend forecast is equal to $\{a_0 + b_1(\text{forecast year} - 1951)\}$.

Next, results are given for the comparison of the RMSEs of the best composite forecasts that include a) only NASS indications, b) NASS indications and the long-term trend forecast, and c) NASS indications and the Feyerherm forecast for the month.

Finally, results of comparing the errors for the Feyerherm forecasts with that for the long-term trend forecast, independent of NASS indications, are given.

I. Composite Forecasts Including All Indications

Table 1 gives the RMSEs of the three types of composite forecasts that include all of NASS indications available for the month. The indications included in the composite forecasts are given in Appendix 2 by month and state. They are combined as shown in (3). The row in Table 1 entitled "NASS Ind" refer to composite forecasts that include only NASS indications. In the table, "NASS and Trend" refer to the composite forecasts that include NASS indications and the forecast from long-term trend. The row "NASS and Feyerherm" refer to composite forecasts that include NASS indications and the Feyerherm monthly forecast. The Board forecasts are included to provide a point of reference. The table shows that the forecast errors across states generally decrease as the season progresses but that for a given month there is not a marked difference between the three types of forecast errors for a state. This observation is reinforced in the results for the rank sums test reported in Table 2. It shows that the forecast effect is insignificant, which means that addition of the Feyerherm forecast to all NASS indications does not improve the accuracy of the composite forecast. Plots of the composite forecasts for each state in Appendix 4 show how close the three types of forecasts are.

Table 1--Root mean-square errors(bushels per acre) of composite forecasts including all indications, for the period 1977-1983

DATE		STATE							
MAY 1		CO	IL	IN	KS	MT	NE	OH	OK
	NASS Ind.	3.52	4.09	4.14	3.68	3.71	4.97	3.24	3.69
	NASS Ind. and Trend	3.88	3.36	3.67	3.85	3.91	4.15	2.92	3.79
	NASS Ind. and Feyerherm	3.87	3.29	3.89	3.92	3.99	4.08	2.99	3.55
	BOARD	3.93	3.76	4.63	4.43	4.23	3.44	4.29	3.44
JUNE 1		CO	IL	IN	KS	MT	NE	OH	OK
	NASS Ind.	3.05	1.58	2.77	3.09	3.82	4.51	4.58	2.62
	NASS Ind. and Trend	3.39	1.42	2.59	3.17	3.30	4.28	4.20	2.74
	NASS Ind. and Feyerherm	3.29	1.38	2.73	3.32	3.50	4.24	4.22	2.66
	BOARD	2.98	2.04	3.44	3.49	3.33	3.44	4.09	2.33
JULY 1		CO	IL	IN	KS	MT	NE	OH	OK
	NASS Ind.	1.82	1.56	2.84	1.08	2.25	1.98	3.12	0.80
	NASS Ind. and Trend	2.14	1.50	2.80	1.19	2.29	1.86	2.98	0.80
	NASS Ind. and Feyerherm	2.01	1.52	2.95	1.28	2.32	1.83	2.99	0.75
	BOARD	2.18	1.51	2.85	1.49	2.55	2.14	3.00	1.00

Table 2--Friedman's Rank Sums Test: Root mean-square errors for composite forecasts including all indications

MONTH	S	p value
May	.25	.967
June	.75	.794
July	.25	.967

II. <u>Composite or Individual Forecasts with Best Monthly Performance</u>

In the second comparison all possible forecasts for the month are examined and the one with the minimum root mean-square error is selected. This is carried out by type of forecast to allow each type to be judged by its best performance. The indications included in the best forecast are given in Appendix 3 by month and by state; equation (3) shows how the indications are combined. The best forecast with only NASS indications include a minimum of one indication while the best composite forecast with NASS indication(s) and trend or the Feyerherm monthly forecast include a minimum of two components. Table 3 gives the root mean-square errors of the forecasts. The titles of the rows in Table 3 have the same meaning as those in Table 1. For example, the row entitled "NASS Ind. and Feyerherm" refer to RMSEs of composite forecasts that include NASS indications and the Feyerherm monthly forecast.

The results given in Table 3 are similar to the ones found in Table 1. That is, the errors across states become smaller as the year progresses but do not show marked differences within a state for a given month. Introducing a long-term trend improves May 1 forecasts in four states and adds a degree of stability to that early forecast. The Friedman's rank sums tests results for the best forecasts are given in Table 4. S' is used instead of S for July because one tie group exists among the RMSEs for one state. The forecast effects for all three months are insignificant.

Table 3--Root mean-square errors(bushels per acre) of forecasts with best performance for the period, 1977-1983

DATE		STATE							
		CO	IL	IN	KS	MT	NE	OH	OK
MAY 1	NASS Ind.	3.40	4.09	3.73	3.42	2.88	4.54	3.24	3.58
	NASS Ind. and Trend	3.88	2.82	3.33	3.66	3.76	3.51	2.92	3.70
	NASS Ind. and Feyerherm	3.87	2.09	3.56	3.78	3.82	3.47	2.99	3.23
	BOARD	3.93	3.76	4.63	4.43	4.23	3.44	4.29	3.44
JUNE 1	NASS Ind.	2.21	1.58	2.17	2.74	3.37	2.68	3.53	2.50
	NASS Ind. and Trend	2.99	1.42	2.28	2.83	2.81	2.42	3.21	2.63
	NASS Ind. and Feyerherm	2.73	1.31	2.46	2.88	3.06	2.40	3.18	2.53
	BOARD	2.98	2.04	3.44	3.49	3.33	3.44	4.09	2.33
JULY 1	NASS Ind.	1.37	1.50	2.70	1.08	2.02	1.63	3.06	0.80
	NASS Ind. and Trend	1.81	1.45	2.63	1.19	1.95	1.50	2.96	0.80
	NASS Ind. and Feyerherm	1.66	1.45	2.78	1.28	2.03	1.55	2.91	0.75
	BOARD	2.18	1.51	2.85	1.49	2.55	2.14	3.00	1.00

Table 4--Friedman's Rank Sums Test: Root mean-square error for best monthly forecast

MONTH	S	S'	p value
May	.25	.	.967
June	.25	.	.967
July		.56	>.794 <.967

III. Comparison of Feyerherm Forecasts

In the last comparison, Feyerherm forecasts are examined independent of NASS indications. As stated earlier ,the difference between the Feyerherm monthly forecast and the long-term trend forecast is the inclusion of weather information in the monthly forecast. Unlike the Feyerherm monthly forecast which varies from month to month, the long-term trend forecast is constant for any given year. Table 5 gives the root mean-square errors of the forecasts. In this table RMSEs of the long-term trend forecast are given in the row entitled "Trend Yield" and the RMSEs of the Feyerherm monthly forecast for May, June and July, are given in the rows entitled "Feyerherm May Yield", "Feyerherm June Yield" and "Feyerherm July Yield", respectively. Another means of comparing the errors is given in Table 6. In this table a plus sign indicates that the RMSE of the Feyerherm monthly forecast is smaller than the the RMSE of the long-term trend forecast, a minus sign indicates that the forecast for the month has a larger RMSE than the long-term trend forecast, and O indicates that the two RMSE's are equal.

Tables 5 and 6 show that generally across states, the results are mixed; that is, there is little indication that the monthly forecast which is derived using daily weather data is more accurate than the forecast obtained using trend. These findings are also reinforced by the absolute values of (R_m - R_t) given in Table 7. In the table D is the critical value for the test. To reject the null hypothesis that the RMSE of a monthly forecast is the same as the RMSE of the trend forecast the absolute value of (R_m - R_t) has to be greater than or equal to D. This is not the case and the null hypothesis is not rejected. It is interesting that the smallest absolute values of (R_m - R_t) are for the June 1 and July 1 forecasts, which one would expect to be the most different from the trend forecast.

Table 5--Root mean-square errors(bushels per acre) of Trend and May 1 to July 1 Feyerherm forecasts for the period, 1977-1983

FORECASTS	STATE							
	CO	IL	IN	KS	MT	NE	OH	OK
Trend Yield	5.18	3.36	3.63	4.55	4.60	3.16	3.59	4.56
Feyerherm May Yield	4.83	3.92	4.45	4.84	4.78	3.10	3.59	4.65
Feyerherm June Yield	4.10	3.79	4.37	4.58	4.09	3.10	3.63	4.29
Feyerherm July Yield	3.41	3.77	4.28	4.25	3.81	3.38	3.67	3.82

Table 6--Comparison of Root mean-square errors for Trend forecast and the Feyerherm monthly forecasts for the period, 1977-1983

	STATE							
MONTH	CO	IL	IN	KS	MT	NE	OH	OK
MAY	+	-	-	-	-	+	O	-
JUNE	+	-	-	-	+	+	-	+
JULY	+	-	-	+	+	-	-	+

\+ Means RMSE for Feyerherm monthly forecast smaller than RMSE for long-term trend
\- Means RMSE for Feyerherm monthly forecast larger than RMSE for long-term trend
O Means RMSEs for both forecasts are equal

Table 7. Multiple Comparison Results for Feyerherm monthly forecasts with Trend forecast as control, for the period 1977-1983

MONTH	D	$\|R_m - R_t\|$
May	12.1	8
June	12.1	2
July	12.1	2

CONCLUSIONS

Results of comparative analyses performed in eight winter wheat Objective Yield Survey states from 1977 to 1983 show that the Feyerherm winter wheat model provides no significant improvement in forecasting ability over current Agency yield indications and a long-term trend forecast. That is, the change in forecast errors when the Feyerherm forecasts are added to NASS monthly yield indications to form composite forecasts is insignificant. The indications used were the mean locality yield, the condition-yield, multiple regression yield and objective yield. The results also show that differences between long-term trend forecast and the Feyerherm forecast for the month are insignificant. The forecast made from long-term trend utilizes no weather information while the monthly forecast is based on a complex function which is estimated by using values of daily weather variables.

RECOMMENDATIONS

According to the conclusions above the Feyerherm model forecasts do not make a significant contribution to the ability of NASS to predict May 1 to July 1 winter wheat yields. A straight forward long-term trend is also as accurate as the Feyerherm monthly forecast which uses daily weather data. Therefore it is recommended that the Feyerherm monthly forecast not be used in NASS's operational program.

It is recommended that a study focusing on NASS indications be conducted for wheat, corn, soybeans and cotton to

a) examine the possibility of finding a " best fixed " composite indication for all states, for a given month and crop.

b) compare the performance of the composite indication found in (a) with the performance of those used by the Crops Branch.

Some preliminary analysis related to recommendation (a) has been carried out for wheat[7].

REFERENCES

1. Baier, W., D. Z. Chaput, D. A. Russels, and W.R.Sharp,1972. Soil Moisture Estimator Program System. Tech. Bull. 78, Agrometeorology Section, Agric. Canada, Ottawa

2. Cotter, James J., Evaluation of the Feyerherm Winter Wheat Model for Estimating Yields in Indiana, Kansas, Montana and Ohio, U.S. Dept. AGR., Statistical Reporting Service, October 1984.

3. Dunnett, C.W.,1964. New Tables for Multiple Comparisons with a Control, Biometrics , 20: 482-91.

4. Feyerherm, A.M. and Paulsen, G. M., 1986. Development of a Wheat-Yield Function for Winter Wheat. To be published in the Agronomy Journal.

5. Feyerherm, A.M. and Paulsen, G. M., 1981a. Development of a Wheat Yield Prediction Model. Agronomy Journal, 73:277-282.

6. Hollander, Myles, and Douglas A. Wolfe. Nonparametric Statistical Methods. New York: John Wiley & Sons, 1973

7. Klugh, B. F. and Bigsby, F. G., 1987. Unpublished memo to John Witzig.

8. U.S. Department of Agriculture, Statistical Reporting Service, Scope and Methods of the Statistical Reporting Service, September 1983.

Appendix 1

Description of Feyerherm '84 Winter Wheat Model

The Feyerherm winter wheat model uses the relationship between experiment station yields and daily weather variables derived from precipitation, maximum temperature, and minimum temperature to forecast yields [4,5]. This relationship, which is the core of the model, is estimated using regression analyses and is called the weather-yield function. To estimate the weather-yield function, weather data are accumulated based on crop growth stages. The crop growth stages are 1) planting to wintering, 2) wintering to spring green-up, 3) spring green-up to jointing, 4) jointing to flag, 5) flag to heading, 6) heading to milk, 7) milk to dough and 8) dough to ripe. The weather-yield function is abbreviated WYF. The weather variables included in the weather-yield function are called weather-related variables and are abbreviated WRVs. The model also uses a trend term (T) to attempt to account for influences on yield that are not directly weather-related. Examples of such influences are increases in plant population, fertilizer applications, etc. The Feyerherm model for a region is:

$$\text{Yield} = a_0 + b_1(T) + b_2(\text{DWYF}) \qquad (1)$$

The trend term T, is calculated by subtracting the value of each year in the data series from the first year in the series. The coefficient b_1 represents the amount by which yield increases from year to year due to changes in applications of technology to agriculture (assuming the effect of weather is constant). DWYF represents the amount by which the value of a given year's weather-yield function differs from the long-term mean(30 years). DWYF multiplied by b_2 represents the effect of weather on the level of yield (assuming that the effect of technology is constant).

Development of Model

The Feyerherm winter wheat model was estimated in two stages. In the first stage the regression relationship between experiment station yields and weather variables was estimated and called WYF. In the second stage output from this regression relationship was used to develop models for forecasting regional yields.

Stage I. Historical Development of Weather-Yield Function

The data base included 876 observations (yields) for Montana, South Dakota, Nebraska, Kansas, Colorado, Oklahoma, Texas, Missouri, Illinois, Indiana and Ohio experiment stations. Winter wheat yields for each experiment station for the same year were adjusted to one standard yield. The adjustment was based on the 'differential yielding ability' of each variety using the formula:

$$\text{standard yield} = \sum_{k=1}^{n} (Q_K * DYA_K) + \sum_{k=1}^{n} Q_k$$

where

Q = percent of area planted to variety k
DYA = differential yielding ability of variety K (average of the differences between the yields for variety K and the yield for a standard variety)
n = number of varieties

Crop calendar dates which started at planting and ended at ripe were used for accumulating the daily weather data according to wheat growth stages. Planting and heading dates were derived by using regression models developed for this purpose. Other stages occured a fixed number of days from planting or heading. Duration of a stage was the same for all locations, except for the winter to spring green-up period. The following formulas were used to compute the crop calendar:

Stage	50% Date of Occurance
Planting	154.3 + 2.16*(mean of average monthly temperatures for September October and November)
Wintering	Planting Date+60 days, if long-term average temperature in January exceeds 27°F, or is the day of year when long-term temperature reaches 35°F(if long-term average in January does not exceed 28°F).
Spring green-up	Heading date-80 days
Jointing	Heading date-40 days
Flag	Heading date-20 days
Heading	253.8 - 2.06*(mean of average monthly temperatures for March, April and May)
Milk	Heading date+10 days
Dough	Heading date+20 days
Ripe	Heading date+30 days

The Baier-Robertson soil moisture estimator system [1] was used to estimate moisture demand and supply for the crop. The system used daily precipitation and temperature extremes to calculate the daily contents of six soil moisture zones from which the crop can extract moisture. The soil capacity for a location (total inches in all six soil moisture zones) was set at 10 inches. Daily precipitation, maximum and minimum temperatures and simulated evapotranspiration were accumulated over crop growth stages. Averages were computed for temperature variables. Adjusted yields were regressed on weather-related variables to estimate the weather-yield function. Coefficients given below were estimated for the weather-related variables included in the function:

WYF= 80.8

+2.91(sum of daily simulated evapotranspiration from planting to winter)

-0.338(mean of daily maximum temperature from planting to winter)

-0.089(moisture amounts in excess of 10-inch soil capacity, from winter to spring)

-0.00213(mean of daily maximum temperature from winter to spring)

-0.181(mean of daily minimum temperature from spring to jointing)

+12.36(sum of daily simulated evapotranspiration from jointing to flag)

-2.655(Square root of sum of daily simulated evapotranspiration from jointing to flag)

-0.544(sum of increments of minimum temperature below a treshold during jointing to flag. Treshold starts at 20°F and increases by .5 degrees each day during the 20 day period between jointing and flag. At the end of the flag leaf stage the treshold is 30°F.)

-0.71(moisture amounts in excess of 10-inch soil capacity, from jointing to flag)

-0.181(mean of daily minimum temperatures from jointing to flag)

+1.43(sum of daily simulated evapotranspiration from flag to heading)

-0.71(moisture amounts in excess of 10-inch soil capacity in model, from flag to heading)

- 0.181(mean of daily minimum temperatures from flag to heading)

-0.544(sum of increments of minimum temperatures below a treshold during flag to heading. Treshold starts at 30°F at heading and increases each day by the same amount during the 20 period from flag to heading. At heading the treshold is 32°F)

-0.544(sum of increments of minimum temperatures below 32°F during heading to milk)

-0.1525(mean of daily maximum temperatures from heading to milk)

+1.08(sum of daily simulated evapotranspiration from heading to milk)

-0.1535(mean of daily maximum temperatures from milk to dough)

+1.08(sum of daily simulated evapotranspiration from milk to dough)

+1.08(sum of daily simulated evapotranspiration from dough to ripe)

-0.81(amount of precipitation which exceeds 5 inches from dough to ripe)

WYF for planting to ripe=WYF for planting to wintering +
WYF for wintering to spring greenup +
WYF for spring green-up to jointing +
WYf for jointing to flag +
WYF for flag to heading +
WYF for heading to milk +
WYF for milk to dough +
WYF for dough to ripe

STAGE II. Development of Regional Model

The procedures described below used daily weather data obtained from the Oklahoma Climatological Survey in Norman, Oklahoma and the National Weather Service in Asheville, North Carolina. Weather stations were selected based on the ability to obtain daily long-term historical data for them. The number of stations used in each state ranged fron 9 (Colorado) to 16 (Kansas). Each weather station was assigned to one region or two adjoining regions within a state. Regional weather-related variables were obtained by computing a simple average of the weather-related variables for stations within a region. Previous analyses of the Feyerherm winter wheat model found that a denser network of weather stations did not improve the accuracy of the model [2].

Values of WYF were obtained for each region within a state by multiplying the regional weather-related variables by their respective experiment station level coefficients given above .
DWYF was derived by computing the long-term average of WYF and subtracting it from the WYF value for each year. NASS final yields were regressed on DWYF and T to estimate regional models for each state. When the Feyerherm model is used for a state, the WYF coefficients are not re-estimated. Feyerherm's purpose in using experiment stations from a number of wheat states to develop the WYF was to create a model that is "universal for similar growing regions".

Appendix 2

Indications included in Analyses, by month and State

CT = Condition Yield
MR = Multiple Regression Yield
ML = Mean Locality Yield
OY = Objective Yield

STATE	MONTH	INDICATION
Colorado	May	CT,MR
Illinois		CT,MR
Indiana		CT,MR
Kansas		OY,CT,MR
Montana		CT,MR
Nebraska		CT,MR
Ohio		CT,MR
Oklahoma		OY,CT,MR
Colorado	JUNE	ML,OY,CT,MR
Illinois		ML,OY,CT,MR
Indiana		ML,OY,CT,MR
Kansas		ML,OY,CT,MR
Montana		ML,OY,CT,MR
Nebraska		ML,OY,CT,MR
Ohio		ML,OY,CT,MR
Oklahoma		ML,OY,CT,MR
Colorado	JULY	ML,OY,CT,MR
Illinois		ML,OY,MR
Indiana		ML,OY,MR
Kansas		ML,OY,MR
Montana		ML,OY,CT,MR
Nebraska		ML,OY,MR
Ohio		ML,OY,MR
Oklahoma		ML,OY,MR

Appendix 3

Indications included in forecasts with best monthly performance, by month and state

CT = Condition Yield
MR = Multiple Regression Yield
ML = Mean Locality Yield
OY = Ojective Yield

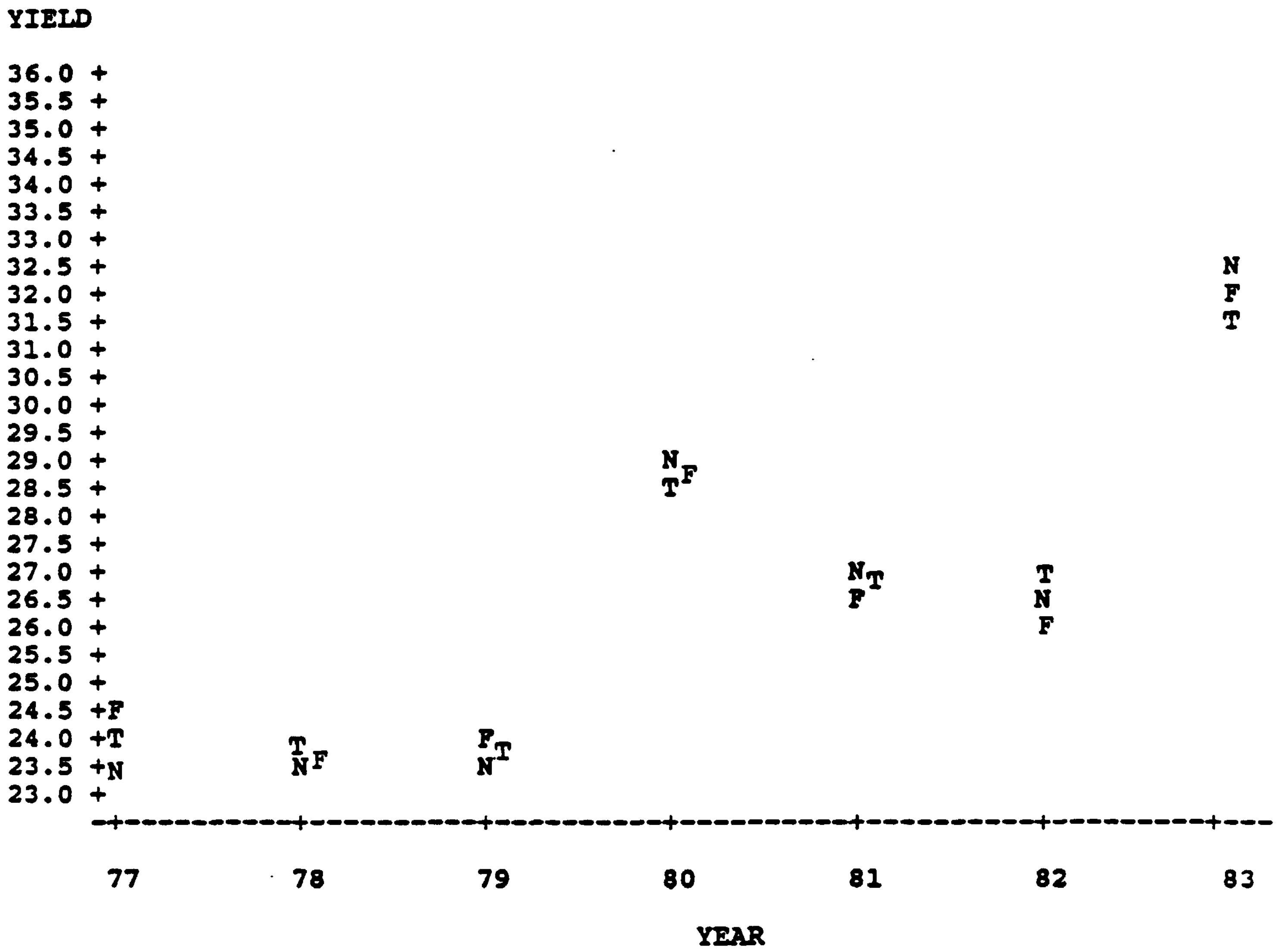
STATE	MONTH	NASS	NASS,Trend	NASS,FY
Colorado	MAY	MR	MR,CT	MR,CT
Illinois		MR,CT	CT	CT
Indiana		CT	CT	CT
Kansas		OY,CT	OY,CT	OY,CT
Montana		CT	CT	CT
Nebraska		CT	CT	CT
Ohio		MR,CT	MR,CT	MR,CT
Oklahoma		OY,MR	OY	OY
Colorado	JUNE	ML	OY,ML	OY,ML,CT
Illinois		ML,MR,OY,CT	ML,MR,OY,CT	CT
Indiana		CT	CT	CT
Kansas		OY,ML	OY,ML	OY,ML,CT
Montana		OY,ML	ML	OY,ML
Nebraska		OY	OY	OY,ML
Ohio		OY,ML,CT	OY,CT	OY,ML
Oklahoma		OY,ML,MR	OY,ML	OY,ML
Colorado	JULY	OY,ML,MR	OY,ML,MR	OY,ML,MR
Illinois		OY,ML	OY,ML	OY,ML
Indiana		ML,MR	ML,MR	ML,MR
Kansas		OY,ML,MR	OY,ML,MR	OY,ML,MR
Montana		OY,ML	ML	ML
Nebraska		OY	OY	OY,ML
Ohio		OY,ML	ML,MR	OY,ML
Oklahoma		OY,ML,MR	OY,ML,MR	OY,ML,MR

Appendix 4

June Composite Forecasts With all NASS Indications(N),
NASS Indications Plus Monthly Feyerherm Forecast(F),
NASS Indications Plus Trend(T)

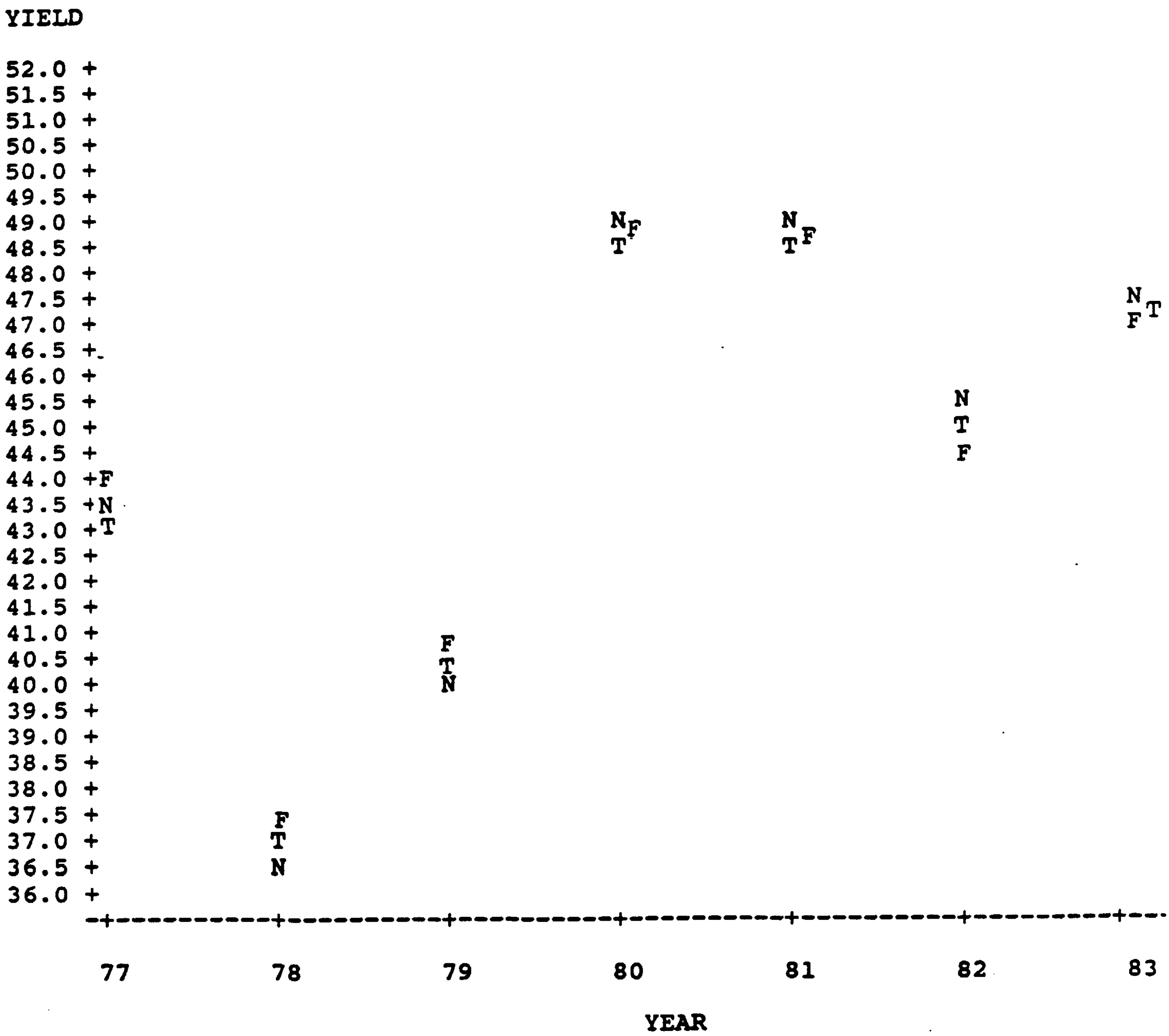
COLORADO

June Composite Forecasts With all NASS Indications(N),
NASS Indications Plus Monthly Feyerherm Forecast(F),
NASS Indications Plus Trend(T)



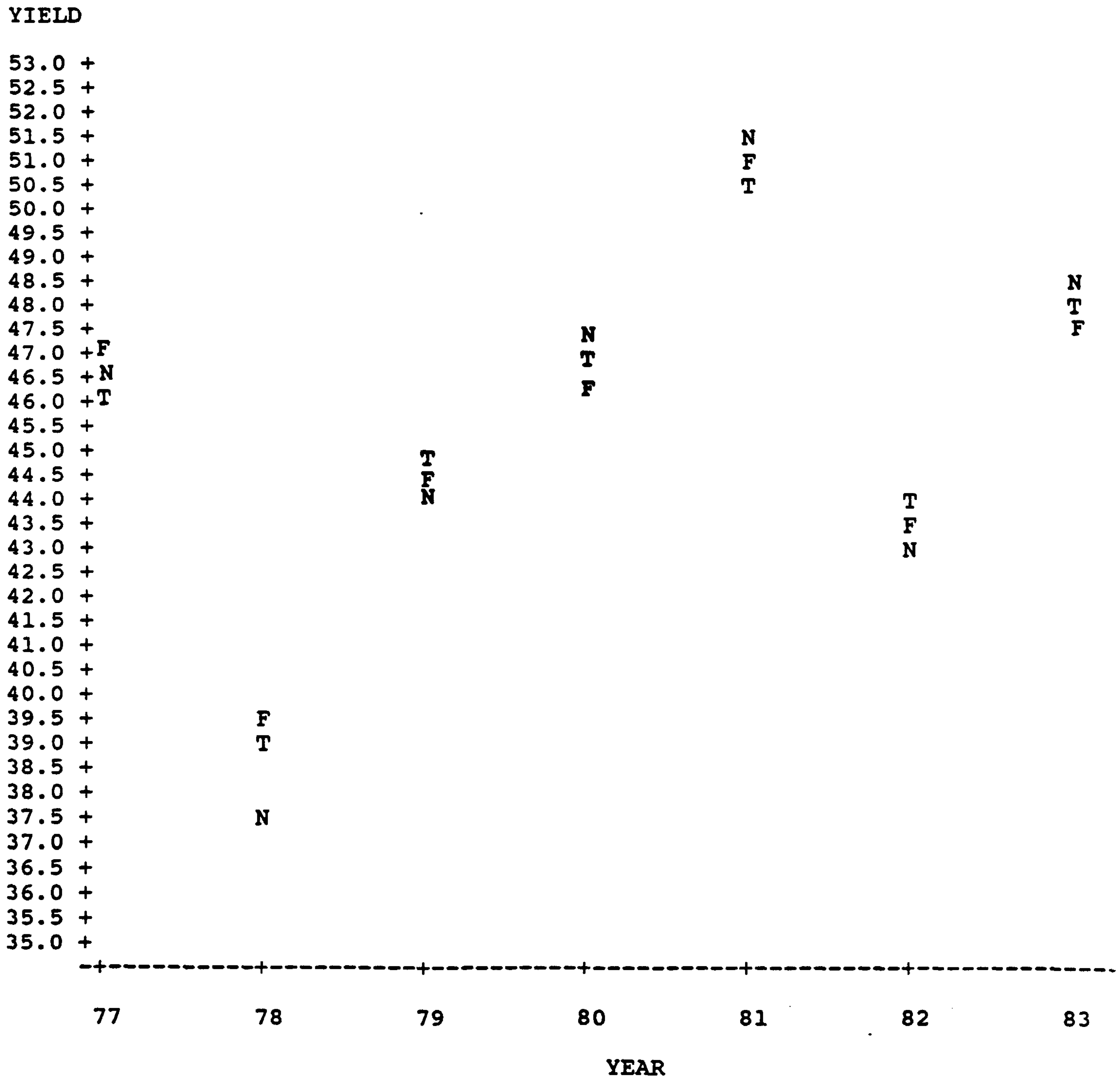
ILLINOIS

June Composite Forecasts With all NASS Indications(N),
NASS Indications Plus Monthly Feyerherm Forecast(F),
NASS Indications Plus Trend(T)



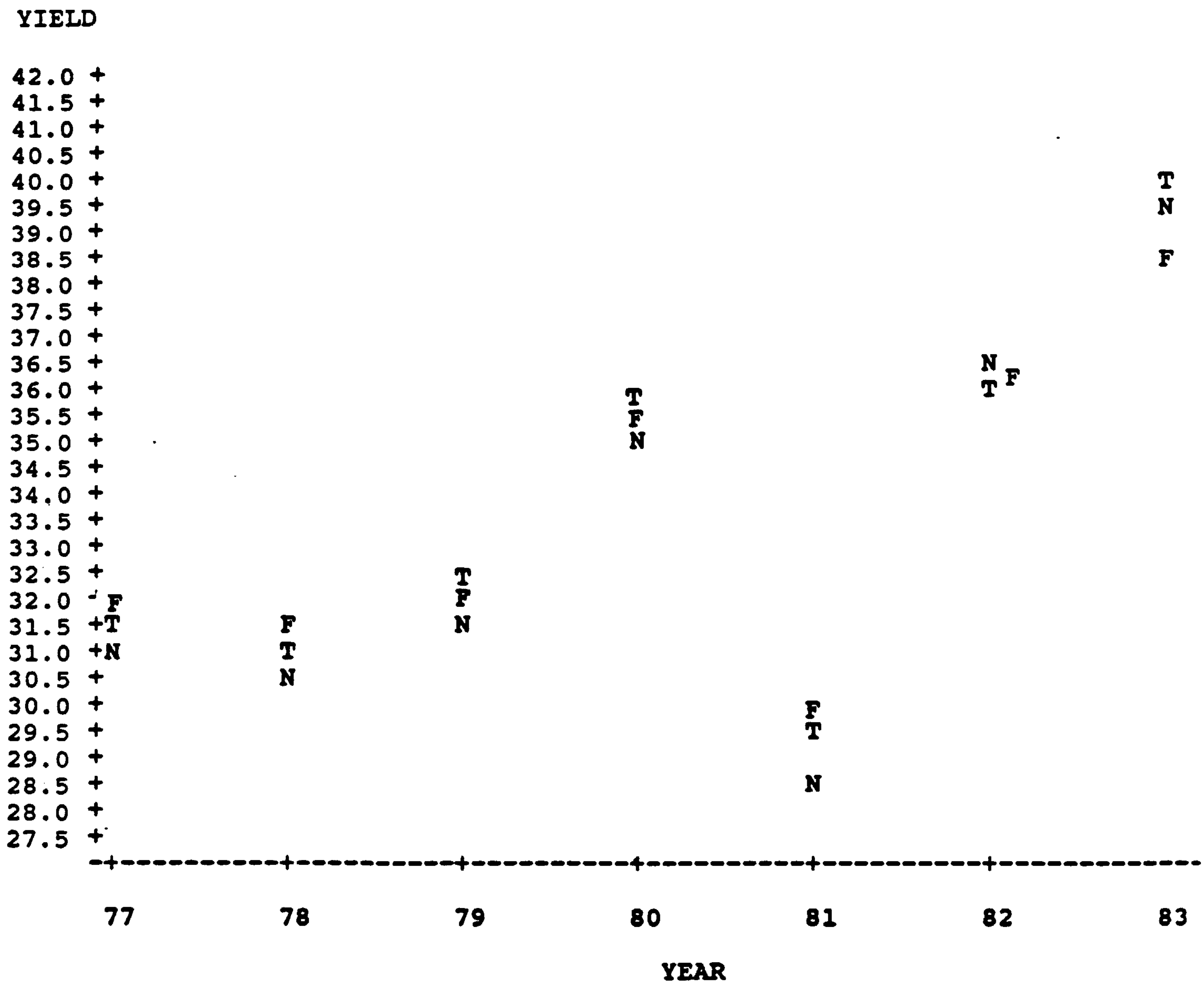
INDIANA

June Composite Forecasts With all NASS Indications(N),
NASS Indications Plus Monthly Feyerherm Forecast(F),
NASS Indications Plus Trend(T)



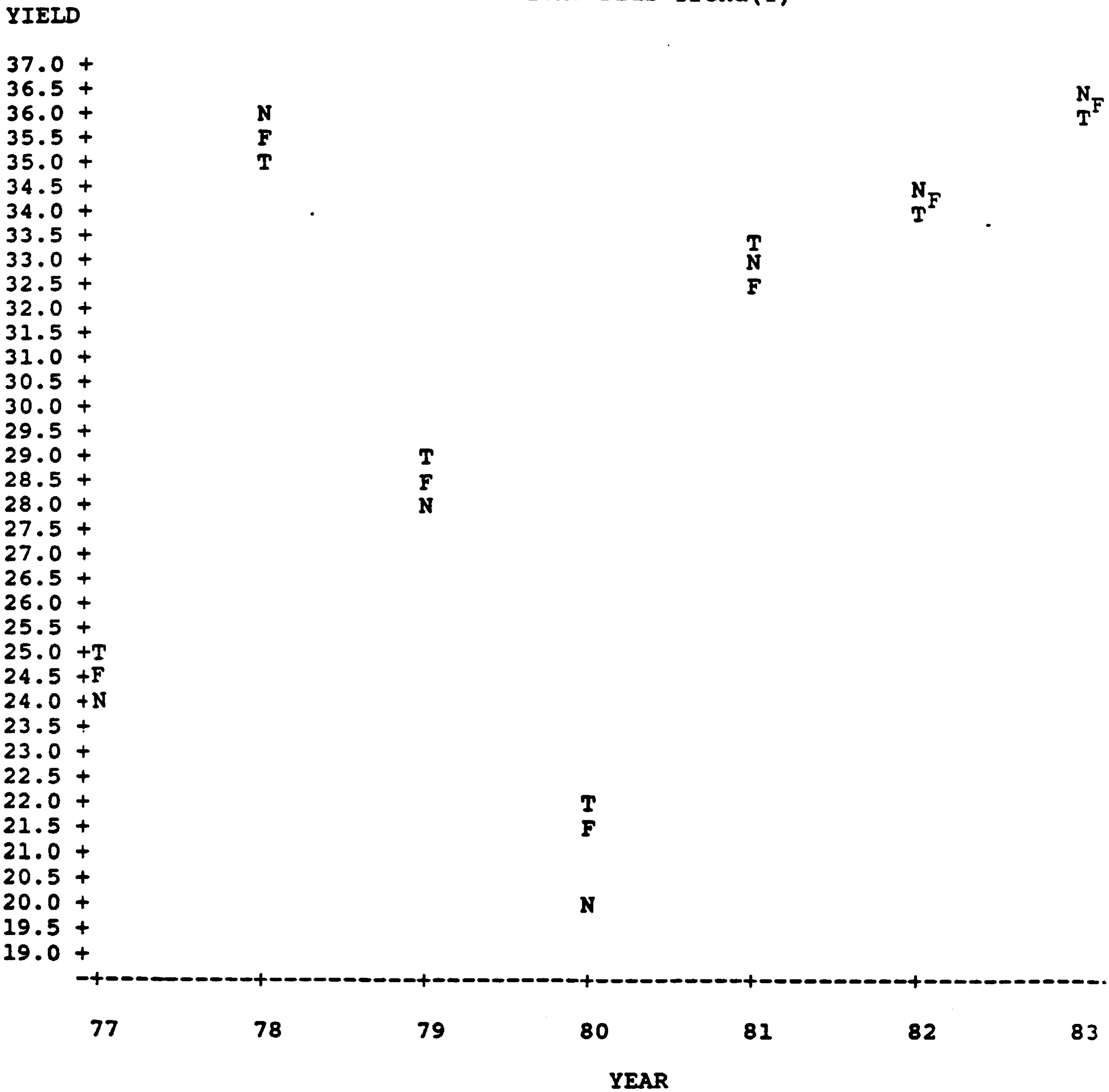
KANSAS

June Composite Forecasts With all NASS Indications(N),
NASS Indications Plus Monthly Feyerherm Forecast(F),
NASS Indications Plus Trend(T)

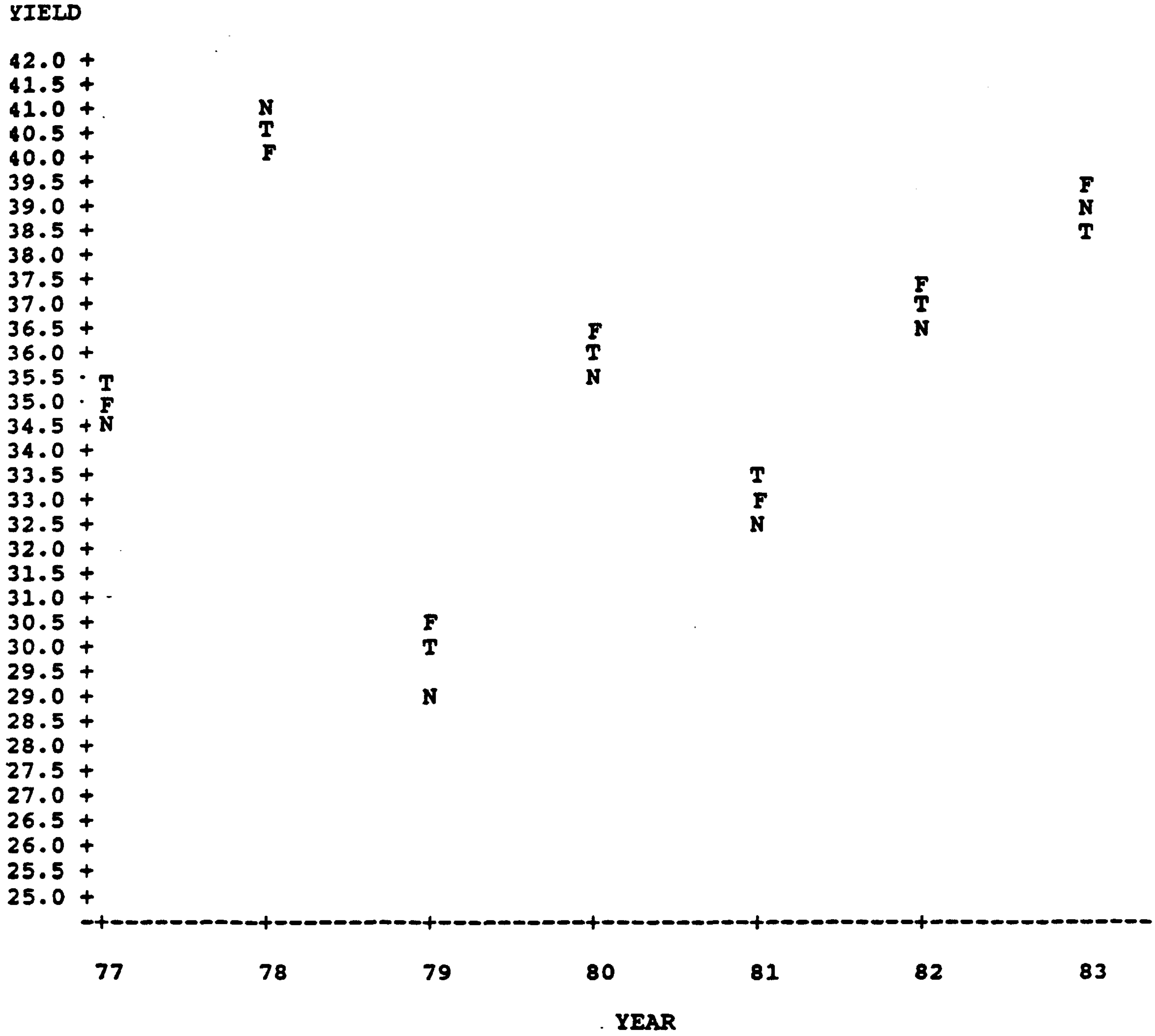


MONTANA

June Composite Forecasts With all NASS Indications(N),
NASS Indications Plus Monthly Feyerherm Forecast(F),
NASS Indications Plus Trend(T)

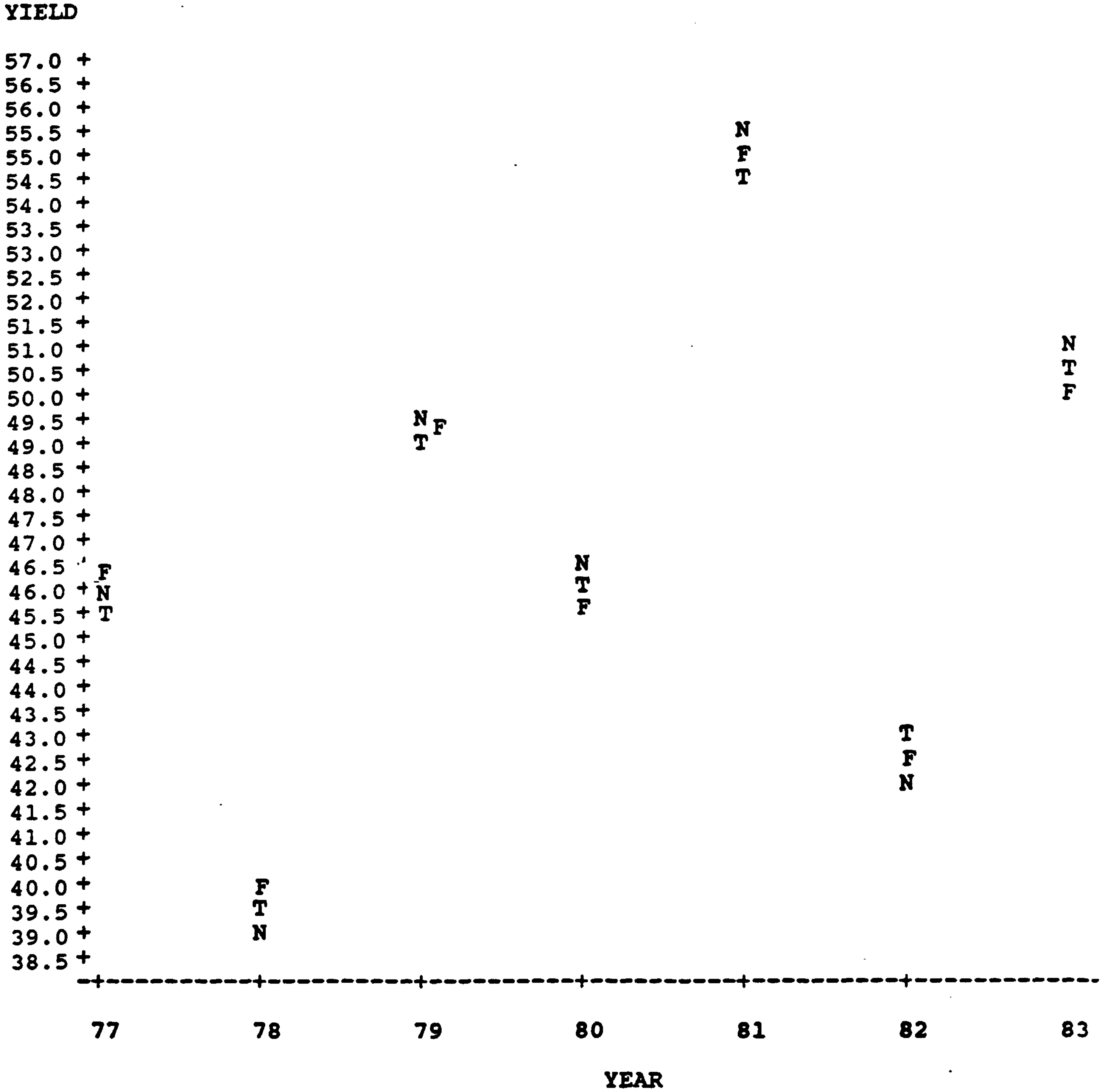


NEBRASKA
June Composite Forecasts With all NASS Indications(N),
NASS Indications Plus Monthly Feyerherm Forecast(F),
NASS Indications Plus Trend(T)



OHIO

June Composite Forecasts With all NASS Indications(N),
NASS Indications Plus Monthly Feyerherm Forecast(F),
NASS Indications Plus Trend(T)



OKLAHOMA

June Composite Forecasts With all NASS Indications(N),
NASS Indications Plus Monthly Feyerherm Forecast(F),
NASS Indications Plus Trend(T)