Forecasting 1977 Kansas Wheat Growth

Economics,
Statistics, and
Cooperatives Service

U. S. Department
of Agriculture

Washington, D. C.
20250

FORECASTING

1977

KANSAS WHEAT GROWTH

By

Greg A. Larsen

Mathematical Statistician
Research and Development Branch
Statistical Research Division
Economics, Statistics and Cooperatives Service
United States Department of Agriculture

Washington, D.C. August 1978

Contents

Index of Diagrams, Tables, and Figures

Page

Acknowledgments

I would like to thank the staff of the Kansas State Statistical Office for supervising the data collection phase of the 1977 Within-Year Wheat Growth Study. A group of approximately 30 enumerators were employed and the office staff did a good job in coordinating their efforts. The keypunching, also done in the state office, was nearly perfect.

I would also like to recognize the fine job that Jack Nealon did in setting up the sampling plan and preparing all the manuals, forms, and supplies necessary for the data collection.

Introduction

Research has been conducted for the past several years in modeling the weight of grain for wheat. The purpose of this effort has been to develop a model capable of making early season forecasts of yield per acre based entirely on current season data as a supplement to the regular forecasting program. The regular objective yield forecast puts current year plant data into a model developed from previous years. While performing satisfactorily in typical years, models which have been developed from historic sources often falter in atypical years. Within-year growth modeling research has been conducted with the idea of providing supplemental information or sensitivity in unusual years.

Research efforts in past years have been done on a comparatively small scale. By contrast, the 1977 study was on a state level with the intent of collecting sufficient data to answer several of the questions remaining from the previous work. The main question is whether the within-year growth model can provide a reliable yield forecast by June 1. This requirement is necessary for the growth model to be of help to the regular objective yield forecasting program. Another question is whether there are differences between early, normal and late developing fields which would undermine efforts to fit a single model for the entire state.

Sample Design

The 1977 study was conducted in Kansas. The sample was a simple random subsample of one-fourth of the regular objective yield fields. This provided an initial sample of 80 fields distributed throughout the state. Refusals and other problems reduced the final sample size to 67. The regular objective yield sample was selected with probability proportional to acreage from wheat fields identified in a December area frame land use survey.

Within each field, there were two randomly and independently located plots. Each plot consisted of one row and its associated row middle and was approximately five feet in length. If rows were not discernable, the plots were six inches wide. Stalk counts were made in the five-foot plot after most of the stalks had at least one leaf off of the main stem. Every tenth plant was tagged until a total of 30 was reached. The 30 tagged plants could lie within the five-foot section or go beyond it depending on the plant density. The stalk counts in 1977 were such that the sample generally extended somewhat past the five-foot area.

Weekly visits were made by trained enumerators to observe when the tagged plants had fully emerged heads and when flowering occurred. Once at least 80% of the tagged plants in a particular field had flowered, clipping began. A predetermined random sample of four heads per plot per weekly visit was clipped, placed in air tight plastic tubes and mailed to the state lab.

Once in the lab, the wet and dry weights were determined for each head. The heads were dried for 46 hours at 150 degrees Fahrenheit. Head emergence and flowering continued to be observed until two weeks after clipping began or until all tagged plants had flowered, whichever came first. Heads continued to be clipped on a weekly basis until harvest. Time since flowering and time since full head emergence were calculated for each clipped head.

Aggregation

The sampling design is hierarchial with plots nested within fields and stalks nested within plots. This implies that individual head weights are not independent and the data should be aggregated to a level in which the resultant means are independent. Aggregation took place on a time interval basis so that only heads with similar time values were combined. The flowering date was determined by taking the average of the date of the visit when flowering was first observed and the date of the previous visit. Since visits were generally made on a weekly basis, the flowering date has a maximum error of 3.5 days in either direction. For this reason, time values computed to be within a few days of each other can be considered equivalent. A plot of the data appears in Figure 1. Since visits were made on about the same day each week, the data readily divides into seven time intervals. Observations within a time interval were assumed to have equivalent time values. Individual head weights within a time interval were aggregated to the plot level. Since plots within a field are independent, the two plot means per time interval were averaged together equally.

The Logistic Growth Model

Previous research in the area of within-year growth models for wheat[1/] and corn[2/] has demonstrated that a growth model could be successfully used to describe the time-growth relationship for these two crops. The basic growth model that has been used is as follows:

$$(1) \qquad y_i = \frac{\alpha}{1 + \beta\rho^{t}} + \varepsilon_i \text{ where } i = 1, 2, \ldots, n$$

$\alpha>0$, $\beta>0$, $0<\rho<1$

y_i = dependent growth variable

t_i = independent time variable

ε_i = error term

[1/] Nealon, Jack, 1976, The Development of Within-Year Forecasting Models for Winter Wheat. Research and Development Branch, Statistical Research Division, ESCS, USDA.

[2/] House, Carol C., 1977, A Within-Year Growth Model Approach to Forecasting Corn Yields. Research and Development Branch, Statistical Research Division, ESCS, USDA.

Logistic Growth Model

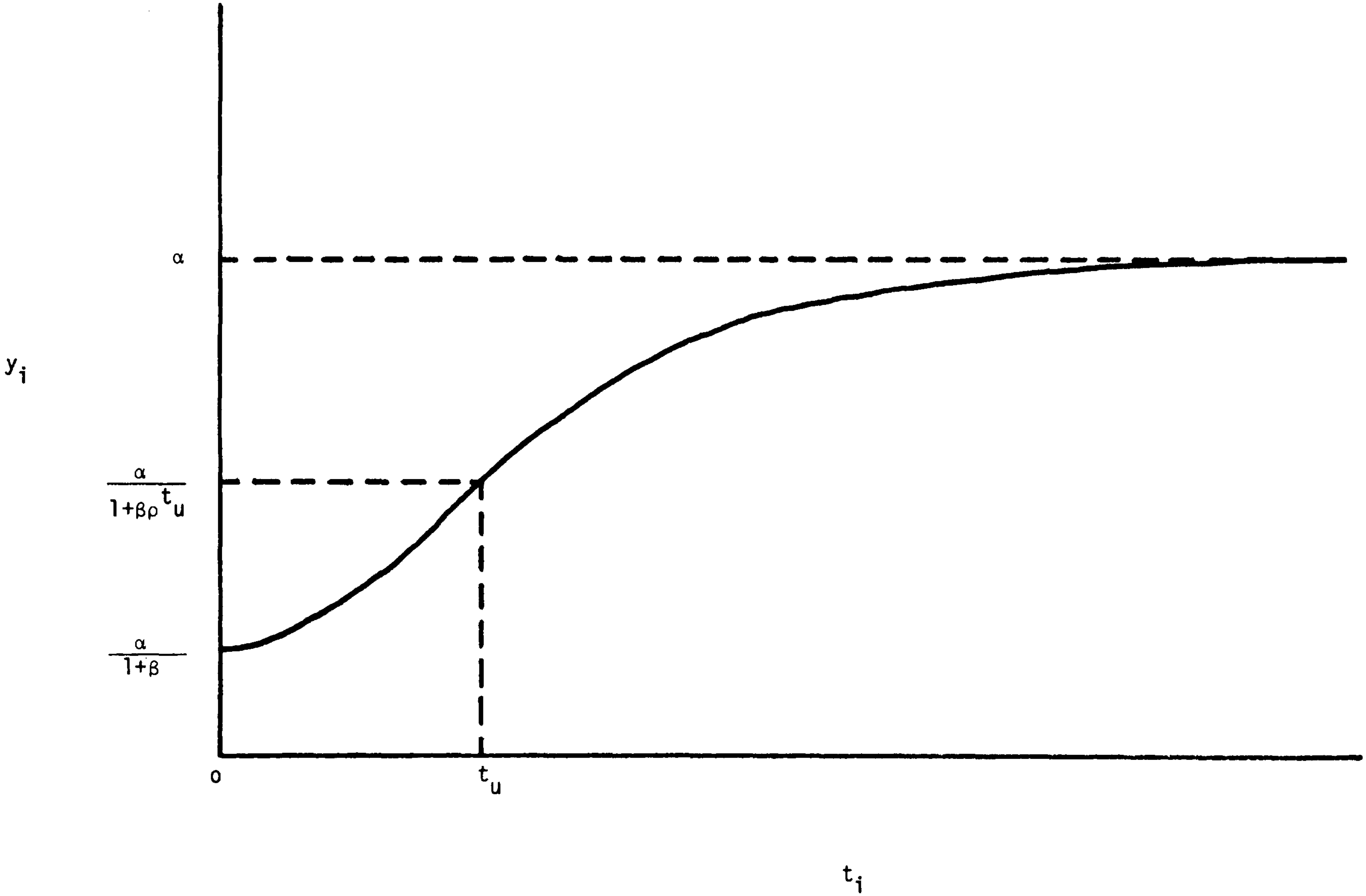


Diagram 1

Least squares theory was used to estimate the parameters α, β, and ρ. This requires the following assumptions about the nature of the model.

(2) $E(\varepsilon_i) = 0$ for all i

(3) $Var(\varepsilon_i) = E(\varepsilon_i^2) = \sigma^2$ for all i

(4) $Cov(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i \varepsilon_j) = 0$ for all $i \neq j$

The parameter which we are most interested in estimating is the asymptote, α (see Diagram 1). The estimate of α is the average amount of dry matter which has been accumulated at maturity. The dependent growth variable used in this report is dried head weight. Dried kernel weight has also been used in the past but the dried head weight was shown to be preferable.[1/] The independent time variable is time since flowering. Previous research has demonstrated that time since flowering is preferable to time since full head emergence.[2/] Examination of the 1977 data has verified this result.

Biological Yield

The asymptotic result (α) of the within-year growth model provides a forecast of dry weight per head for heads surviving until maturity. The dry weight can then be adjusted for threshing and moisture to the standard moisture grain weight. The stalk counts that are made in the five-foot plots can be adjusted by the proportion which have heads at maturity and expanded to an average number of heads per acre at maturity. Multiplying the standard moisture grain weight per head times the average number of heads per acre at maturity produces a forecast of biological yield per acre. This biological yield could then be adjusted to reflect harvest loss using information obtained in the regular objective yield forecasting program.

To adjust the dry head weight for threshing and moisture, four extra heads per plot were clipped on the last weekly visit before harvest. These extra heads were paired with the regular sample by clipping the untagged stalk immediately following the tagged stalk to be clipped. By extracting the kernels, the extra heads were used to obtain wet and dry grain weights per head. The drying procedure for the grain was consistent with the standard methods for grain moisture determination. It consisted of drying the grain for 16 hours

---

1/ Nealon, p. 9.

2/ Nealon, p. 7.

at 266 degrees Fahrenheit. The dry grain weight is adjusted to the standard 12% moisture by dividing by 0.88. The dry grain weight at the standard moisture is then related to the dry head weight from the corresponding stalks. Since the grain weight from the extra heads is obtained at maturity, the data must come from a previous year to be of use in a forecasting mode. The relationship between the historic standard moisture grain weight and the head weight forecasted by the growth model gives a method of producing forecasts of standard moisture grain weight per head during the current year.

To obtain the average number of heads per acre at maturity, a survival ratio was computed and applied to the stalk counts made in the five-foot plots. As mentioned earlier, after 80% of the tagged stalks had flowered, weekly visits were made to clip the heads from four randomly selected stalks per plot. The sampled stalks received a code based on whether or not a head was present. Partially emerged heads and heads still in the boot were classified as having a head. The survival ratio was the number of stalks with heads divided by the total number of sampled stalks. The survival ratio was very consistent from week to week as the season progressed so it was not necessary to forecast what the ratio would be at maturity. After the stalk counts in the five-foot plots are adjusted to stalks with heads at maturity, they are expanded to an average number of heads per acre at maturity. The biological yield per acre can then be computed by taking the product of the standard moisture grain weight per head and the average number of heads per acre.

Weighting

As described in the previous section, the stalk population and survival ratio are used after an average weight per head is obtained from the growth model. If either is strongly correlated with dry head weight then a change needs to be considered in the model form. The average number of stalks per square foot was compared to the dry head weight on a plot basis for time since flowering between 28 and 33 days. A single time interval was used so that head weights would not be varying over time. This time interval was selected because it has the largest range of head weights and because it has nearly all the fields represented. There was a highly significant negative correlation between stalk population and head weight. This means that the denser the stalk population the less the weight per head. The survival ratio on a plot basis was compared with both the head weight averaged over time and the stalk population. The survival ratio did not appear to be correlated with either.

This suggests that the stalk population should be incorporated into the model to take advantage of its relationship with head weight. This can be accomplished with a weighted regression. The weight to be used is stalks per square foot. This was done to avoid the large numbers encountered with larger units. (1) now becomes:

$$(5) \quad w_i y_i = w_i \left[\frac{\alpha}{1 + \beta\rho^t} + \epsilon_i \right] \quad i = 1, 2, \ldots, n$$

where w_i = some function of the average number of stalks per square foot.

The subscript i has been used in a general sense and simply refers to the observations of the dependent variable. The method of aggregation already discussed indicates that y_i is actually a field level mean dry head weight in a particular time interval. A more specific notation which will let us indicate the method of aggregation is as follows:

y_{jkm} = mean dry head weight in the j^{th} field, the k^{th} time interval and the m^{th} plot.

To aggregate to the field level within time intervals we would average the plot means. Thus,

$$y_{jk.} = \frac{y_{jk1} + y_{jk2}}{2}$$

So, $y_{jk.}$ is the same as y_i in the general notation. Now, from (5) we'd like to find some function of the average number of stalks per square foot, w_i. Since there is an estimate of the stalk population from each of the two plots within a field, an obvious choice for w_i would be an average of the two estimates for each field. Define the following variable.

sp_{jkm} = average number of stalks per square foot in the j^{th} field, the k^{th} time interval, and the m^{th} plot.

Since sp is the same for all time intervals in a plot, the subscript k could be omitted. The average of the two stalk population estimates in each field is then:

$$sp_{jk.} = \frac{sp_{jk1} + sp_{jk2}}{2}$$

So, $sp_{jk.}$ is a candidate for w_i in (5). However, the data revealed that there was quite often a large difference between stalk counts from plots within the same field. This suggests that it would be more effective to weight on a plot basis than a field basis. Therefore, we would like to find some w_i which when multiplied times $y_{jk.}$ gives the equivalent to weighting on a plot basis. This implies that:

$$(6) \quad y_{jk.}w_i = \frac{y_{jk1}sp_{jk1} + y_{jk2}sp_{jk2}}{2}$$

Define a new variable $z_{jk.}$ where:

$$z_{jk.} = y_{jk.}w_i$$

Solving for w_i gives:

$$(7) \quad w_i = \frac{z_{jk.}}{y_{jk.}}$$

Using (7) as the weight in (5) provides the equivalent of plot level weighting even though the observations going into the model are at the field level.

Data Analysis

General

The nonlinear procedure (NLIN) in the Statistical Analysis System (SAS)[1/] can be used to fit either (1) or (5) to the data. There are two ways to fit (5) which provides equivalent sums of squares and parameter estimates. The NLIN procedure can be applied to (5) or a weight statement can be used with NLIN applied to (1). The value of the weight in the weight statement is multiplied times the diagonal elements of the variance-covariance matrix of the dependent variable. From (7), the value of the weight would be $(z_{jk.}/y_{jk.})^2$. The difference between the two ways of fitting (5) is that without the weight statement the residuals are in terms of $y_i w_i$ while the residuals with the weight statement are in terms of the original dependent variable y_i. The weight statement will be used for purposes of weighting by stalks per square foot. A heteroscedasticity adjustment, to be discussed next, will not use the weight statement since it is necessary to have the residuals in terms of an adjusted variable.

Heteroscedasticity

The NLIN procedure was used to fit (1) to the data. As discussed earlier, the y_i's are field level averages of dried head weight at t days since flowering. Figure 1 shows a plot of the data and the estimated growth curve. Figure 2 is a plot of the residuals versus time. The tendency of the residuals to spread out as time increases is evidence of heteroscedasticity. The correlation between the absolute value of the residuals and time is .34 which is highly significant. The assumption in (3) necessary to make least squares parameter estimates is violated since the variance of the dry head weights increases with time rather than being constant. The effects of heteroscedasticity on our ability to make reliable forecasts and methods of adjusting have been discussed in greater detail in an earlier paper.[2/] Two different methods of adjustment will be used in this report. The first will be referred to as the YHAT adjustment.[3/]

1/ Barr, Anthony J., Goodnight, James H., et. al. A User's Guide to SAS '76. SAS Institute Inc., Raleigh, North Carolina, pp. 193-199.

2/ Larsen, Greg A., 1978, Alternative Methods of Adjusting for Heteroscedasticity in Wheat Growth Data, Research and Development Branch, Statistical Research Division, ESCS, USDA.

3/ Larsen, pp. 14-21.

$$(8)\quad \frac{y_i}{\sqrt{k_i}} = \frac{\alpha}{\sqrt{k_i}\ (1 + \beta\rho^t)} + \frac{\varepsilon_i}{\sqrt{k_i}} \qquad \text{where } i = 1, 2, \ldots, n$$

$$(9)\quad \sqrt{k_i} = \frac{\sigma_t}{\sigma}$$

where σ_t = true standard deviation of y at a point in time

σ = true population standard deviation

$$(10)\quad \hat{\sigma}_t = a + b\,\hat{y}$$

The estimate of σ_t used in (10) is made by calculating the standard deviation of the mean dry head weights in each of the seven time intervals mentioned earlier. This standard deviation is weighted by the number of observations in each mean. The $\hat{y}$ values come from a fit of either (1) or (5) depending on whether or not the stalk population is considered. An estimate of σ^2 is obtained by calculating a weighted MSE (weighted by the number of observations in each residual). (10) is obtained by performing a linear regression of $\hat{\sigma}_t$ on $\hat{y}$. An estimate of $\sqrt{k_i}$ is made by dividing both sides of (10) by $\hat{\sigma}$. This estimate of $\sqrt{k_i}$ is called the YHAT adjustment and will provide uncorrelated residuals if the original residuals are approximately normally distributed about the fitted growth curve. Some examples using simulated data are discussed in Appendix A.

The second method of adjustment to be used in this report is the LOG adjustment.[1/]

(11) $$\ln (Ay_i + B) = \ln \left[\frac{A\alpha}{1 + \beta\rho^t} + B \right] + \varepsilon_i \qquad \text{where } i = 1, 2, \ldots, n$$

Values for A and B are chosen so that the range of the transformed dependent variable is the same as that of the untransformed dependent variable. This can be done very easily by finding the minimum and maximum values of y_i and solving constraints which make the minimum and maximum in the transformed set identical to those in the original data set. Keeping the transformed data in the same range also maintains roughly the same overall variation. The log transformation has the affect of spreading smaller values of y_i farther apart and making larger values closer together. A look at Figure 1 reveals that this is what needs to happen if there is to be a constant variance over the range of time.

In some cases, (1) may not reduce the heteroscedasticity to the extent that the residuals become uncorrelated. It is then possible to use a double log transformation.[2/]

(12) $$\ln (A \ln (Ay_i + B) + B) = \ln \left(A \ln \left[\frac{A\alpha}{1 + \beta\rho^t} + B \right] + B\right) + \varepsilon_i$$

where $i = 1, 2, \ldots, n$

The same transformation is made a second time. The values of A and B remain the same since the minimum and maximum did not change. The double log may over adjust in some cases and cause the residuals to become significantly negatively correlated which may not be any more desirable than the original problem.

Comparisons Among Early, Normal, and Late Developing Fields

The ability of a single growth model to forecast an average dry head weight on a state level rests upon the assumption that early developing wheat yields no differently than the later developing fields. The stage of development depends upon variety in combination with a host of environmental and cultural factors. Development may correspond to certain geographic areas from year to

---

1/ Larsen, pp. 8-10.

2/ Larsen, pp. 12-14.

year particularly if weather factors are the primary determinants of stage of development. If early developing areas yield differently than those which are later, it may be necessary to fit growth models for different geographic areas. The influence of varietal differences on fitting a single model for the state is discussed later.

To investigate this question, all the fields were classified as being early, normal or late based on the date that flowering was first observed and the date when clipping began. About half of the 67 fields were classified as normal with the remainder split fairly equally between early and late. Timing of flowering showed some correspondence to geographic areas. The South Central Crop Reporting District (CRD) was predominately early while the western third of the state tended to be normal to late. The southeastern CRD had very few samples but appeared to be early.

The NLIN procedure in SAS was used to fit the growth model to early, normal, late and to all fields. The unadjusted model (1), the weighted model (5) and the YHAT adjustment on (5) were run. Table 1 summarizes the results. Estimates are shown for α in grams and the relative standard error of α in a percent. R_p is the Pearson correlation between the absolute value of the residuals and time. A significance probability is given for each correlation coefficient. This is the probability that an equal or greater (in absolute terms) correlation than the one calculated would have arisen from another random sample given that the residuals and time are truly uncorrelated. For a significance level of .05, we would accept the hypothesis of no correlation if the significance probability is greater than .05. We would fail to accept for probabilities less than or equal to .05 and conclude that the residuals and time are not uncorrelated. The lower portion of Table 1 shows the average number of stalks per square foot, the number of observations and the mean dry head weight over time for each group of fields.

The main result in Table 1 is that there is very little difference between the $\hat{\alpha}$'s or between the mean dry head weights for the early and normal fields. However, $\hat{\alpha}$ for the late fields is about 14% above the normal fields. The mean dry head weight over time is approximately 7% higher for the late fields. The fact that early and normal fields are very similar in dry matter production is good from a forecasting point of view. Late fields could be modeled separately but since they are late, there most likely would'nt be sufficient data for an early season forecast. The $\hat{\alpha}$ for all fields is very close to the early and normal $\hat{\alpha}$'s so the fact that late fields behave differently would not be sufficient reason in and of itself to use separate models.

Another result from Table 1 is that weighting by stalk population does have an affect. The decreases in the $\hat{\alpha}$ levels indicate that the lighter heads are receiving more weight. The decrease in the average number of stalks per square foot for fields which develop later is of interest. All stalk counts were made

TABLE 1

Comparisons Among Early, Normal, and Late Fields

		Models		
		Unadj.	Weighted	YHAT Adj.
Early	$\hat{\alpha}$	1.038	.972	.972
	$\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$	10.3	9.1	9.5
	R_p	.313	.334	.075
	Prob>$\lvert R_p \rvert$	.0014	.0006	.4513
Normal	$\hat{\alpha}$	1.045	.980	1.010
	$\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$	4.2	3.4	4.2
	R_p	.339	.339	.005
	Prob>$\lvert R_p \rvert$	.0001	.0001	.9485
Late	$\hat{\alpha}$	1.189	1.103	1.131
	$\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$	9.1	9.8	11.3
	R_p	.279	.319	.033
	Prob>$\lvert R_p \rvert$	.0162	.0057	.7777
All	$\hat{\alpha}$	1.058	.982	1.001
	$\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$	3.6	3.1	3.7
	R_p	.338	.354	.054
	Prob>$\lvert R_p \rvert$	.0001	.0001	.3023

	Obs.	Average Stalks Per Ft^2	Mean Dry Head Wt. (g)
Early	102	56.4	.696
Normal	190	47.6	.705
Late	74	36.1	.749
All	366	46.6	.711

at approximately the same time so the decrease is not caused by a time differential. The nearly constant reduction in $\hat{\alpha}$ from the unadjusted to the weighted models points out that each category has some relatively high stalk populations with relatively low dry head weights. Also, when all observations are used, the stalk population weights the early fields more heavily than the late. This is as it should be since we are ultimately interested in forecasting a yield which is a function of the head weight times the stalk population.

In conclusion, a single state model appears to be a reasonable alternative at least for the 1977 crop. The YHAT adjustment for heteroscedasticity was successful in creating uncorrelated residuals and the relationships among the categories were preserved.

Forecasting at Cut Off Dates

As discussed earlier, the main purpose of a within-year growth model is to provide early season forecasts which can then be used to supplement those obtained from the regular objective yield program. If successful, the within-year growth model would be of particular value in atypical years when a historically based model might perform poorly.

The regular objective yield program makes forecasts on May 1, June 1, and July 1. Normally, very little heading occurs in Kansas prior to May 1 so a within-year forecast using only the growth model is impossible on this date. The June 1 forecast then becomes the critical date since by July 1 there is usually enough yield data to make a reliable forecast in the regular operating program. From a research standpoint, a mid-June update of expected yield might be contemplated. However, we concentrated on the possibility of supplementing the June 1 forecast.

In 1977, there were three weeks of data available for a June 1 forecast. Unadjusted, weighted, YHAT adjusted, LOG adjusted, and double LOG adjusted models were run for each weekly cut off date. The unadjusted model is unweighted while the weighted model has not been adjusted for heteroscedasticity. The YHAT, LOG, and double LOG adjusted models have all been weighted by the stalk population. The information is summarized in Table 2. A general observation can be made. The estimate of α does not start to settle down until the fourth week and it doesn't approach the final level until the fifth week in each of the models. From the standpoint of a June 1 forecast, this result is rather discouraging.

There are several ways to compare the models in Table 2 to determine which would be preferable for a particular cut off date. The MSE's are not directly comparable between models other than to notice that, with the exception of the unadjusted model, they are at about the same level within a particular cut off date. This is particularly true as the number of weeks of data approaches the total. This means that the MSE can not be used as a determining factor in selecting one model over another. The R^2 values are over .9 in all cases except when only one week of data was used. Therefore, comparisons between R^2 values are not particularly meaningful. The criteria that will be used to select the most appropriate model in a given situation are the significance of the correlation between the residuals and time and the stability of the forecast of mean dry head weight as successive weeks of data are included.

As mentioned earlier, the presence of heteroscedasticity influences our ability to make reliable parameter estimates. Since $\hat{\alpha}$ is the parameter estimate which corresponds to the mean dry head weight at maturity, it is important that we do not have our ability to estimate α impaired by heteroscedasticity. Therefore, the significance of the correlation between the absolute value of the residuals and time becomes an important factor. The Pearson correlation coefficient and significance probability is given in Table 2 for each model and weekly cut off date. Clearly, the unadjusted and weighted models have a problem with heteroscedasticity. This is also evidenced by the cone-shaped appearance of the residuals in Figure 2. The YHAT adjusted model has residuals which are uncorrelated at the 5% level when the data set has 2 weeks of data and 5 weeks or more. The LOG adjusted model has significantly correlated residuals at the 5% level for all cut off dates although the correlations are considerably smaller than when no adjustment is made. The double LOG adjusted model successfully reduces the correlation when there is one to five weeks of data but over adjusts with more than 5 weeks of data to the extent that the correlation becomes significant in the negative direction.

The second determinant in selecting a model is the stability of $\hat{\alpha}$. There are several things that need to be considered. The relative standard error is an indicator of how much variability is associated with the parameter estimate. The relative standard errors associated with $\hat{\alpha}$ in terms of a percent are given in Table 2. With three weeks of data the errors are in excess of 20% for all models indicating that there is a considerable amount of uncertainty associated with the forecast of mean dry head weight. With five weeks of data, however, the errors have dropped to around 5% which tends to indicate that a fairly good forecast could be made by mid-June.

While the relative standard error can be used to indicate the stability of $\hat{\alpha}$ as the season progresses and more data becomes available, it has been pointed out[1/] that it is of questionable value as a determining factor in choosing a preferred model. In other words, the model with the smaller error on $\hat{\alpha}$ in a given situation may not necessarily provide a more reliable forecast. The reason for this is that heteroscedasticity influences the estimation of the standard error. In a generalized linear regression setting, least squares parameter estimation with the presence of a heteroscedastic disturbance tends

---

1/ Larsen, p. 22

TABLE 2

Cut Off Dates

		Week of Dates								
		1	2	3	4	5	6	7	8	All
	OBS	21	67	137	208	291	324	348	362	366
Unadjusted	MSE	.0273	.0230	.0308	.0262	.0354	.0417	.0430	.0402	.0394
	R^2	.862	.923	.901	.931	.930	.925	.929	.934	.936
	$\hat{\alpha}$	.5304	2.2308	1.2559	1.1180	1.0317	1.0446	1.0963	1.0588	1.0582
	$\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$	28.7	66.6	27.1	10.0	4.9	4.7	4.7	3.8	3.6
	R_p	.59	.40	.53	.42	.39	.42	.38	.37	.34
	Prob>$\lvert R_p \rvert$	.0049	.0007	.0001	.0001	.0001	.0001	.0001	.0001	.0001
Weighted	MSE	-	57.47	77.06	65.05	81.19	93.26	91.71	88.46	87.28
	R^2	-	.926	.915	.932	.939	.932	.935	.941	.941
	$\hat{\alpha}$	-	2.048	1.1699	1.0961	.9684	.9944	1.0239	.9803	.9822
	$\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$	-	63.7	24.2	12.1	4.5	4.7	4.6	3.2	3.1
	R_p	-	.41	.51	.43	.39	.41	.38	.38	.35
	Prob>$\lvert R_p \rvert$	-	.0006	.0001	.0001	.0001	.0001	.0001	.0001	.0001
YHAT Adjustment	MSE	-	63.23	70.42	70.64	82.75	90.47	91.23	89.70	89.03
	R^2	-	.919	.921	.927	.934	.930	.933	.937	.937
	$\hat{\alpha}$	-	2.173	1.2339	1.0899	.9980	1.0239	1.0459	1.0000	1.0013
	$\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$	-	105.6	32.2	14.1	5.5	5.6	5.2	3.8	3.7
	R_p	-	.22	.19	.14	.06	.08	.08	.07	.05
	Prob>$\lvert R_p \rvert$	-	.0793	.0281	.0490	.2778	.1340	.1225	.1563	.3023
LOG Adjustment	MSE	89.49	78.66	82.92	67.78	87.12	119.32	114.84	95.28	94.21
	R^2	.878	.935	.941	.952	.961	.961	.963	.964	.965
	$\hat{\alpha}$	.5630	2.1223	1.1113	1.0865	.9636	.9776	1.0065	.9716	.9740
	$\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$	45.2	101.3	26.2	13.9	5.2	5.3	5.0	3.7	3.7
	R_p	.55	.26	.36	.22	.17	.17	.14	.15	.13
	Prob>$\lvert R_p \rvert$	.0094	.0311	.0001	.0011	.0028	.0025	.0093	.0044	.0140
Double LOG	MSE	89.62	99.79	82.66	67.11	83.04	121.32	114.57	87.14	86.19
	R^2	.901	.945	.960	.966	.975	.977	.979	.979	.979
	$\hat{\alpha}$	.5341	2.1596	1.0981	1.0737	.9716	.9784	1.0054	.9728	.9750
	$\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$	49.4	158.4	31.5	16.4	6.9	7.3	6.8	4.9	4.9
	R_p	.41	.05	.14	-.05	-.11	-.19	-.22	-.17	-.18
	Prob>$\lvert R_p \rvert$	.0650	.6985	.1026	.5021	.0516	.0014	.0001	.0015	.0005

to understand the standard error.[1/] This means that the standard error estimates will tend to increase as the heteroscedasticity is removed. The models that we are using are, of course, nonlinear but this does make the comparison of relative standard errors between models with varying degrees of heteroscedasticity questionable. It can be seen in Table 2 that the adjusted models had higher relative standard errors than their unadjusted counterparts in every case.

There are other ways to assess the stability of $\hat{\alpha}$ which allow comparisons between models. Several of these have been discussed in an earlier report.[2/] The general idea is that a good forecasting model tends to limit fluctuations in $\hat{\alpha}$ as additional data becomes available. The earlier in the growing season that a model stabilizes, the better. Of course, some of the fluctuation exhibited by the model is due simply to legitimate changes in yield prospects as the season progresses.

Diagram 2 illustrates how each of the models compare during the growing season to the final mean dry head weight that was obtained from the YHAT adjusted model. The double LOG model is not presented because it is very similar to the LOG model. The first two cut off dates are not included because they fluctuated wildly and would have required a much smaller scale. With one week of data, the models that converged produced $\hat{\alpha}$'s which were about 55% of the final level. By contrast, with two weeks of data, the $\hat{\alpha}$'s jumped up to around 215% of the final level. The main thing to notice in Diagram 2 is that the models all behave about the same. Stability is reached near the fifth week with an increase up to week seven and a general decline to the final level. It was seen earlier that the comparatively late fields had a higher mean dry head weight which could explain the rise between weeks 5 and 7. The cause of the decline following week seven is unclear especially since it can be seen in Table 2 that only 22 observations were added to the data set after week seven.

The similarity among the models is to be expected since they are really all the same basic model with differences in weighting and degree of heteroscedastic disturbance. Diagram 3 shows a comparison of the percent of the $\hat{\alpha}$ level during the growing season to the final $\hat{\alpha}$ level for each of the models. Diagram 3 again shows the similarity among the models. In terms of stability, it can be seen that a slight overall edge would go to the LOG adjusted model.

1/ Goldberger, Arthur S., 1964, Econometric Theory. John Wiley & Sons, Inc., New York, London, Sidney, pp. 238-241.

2/ House, pp. 14-16

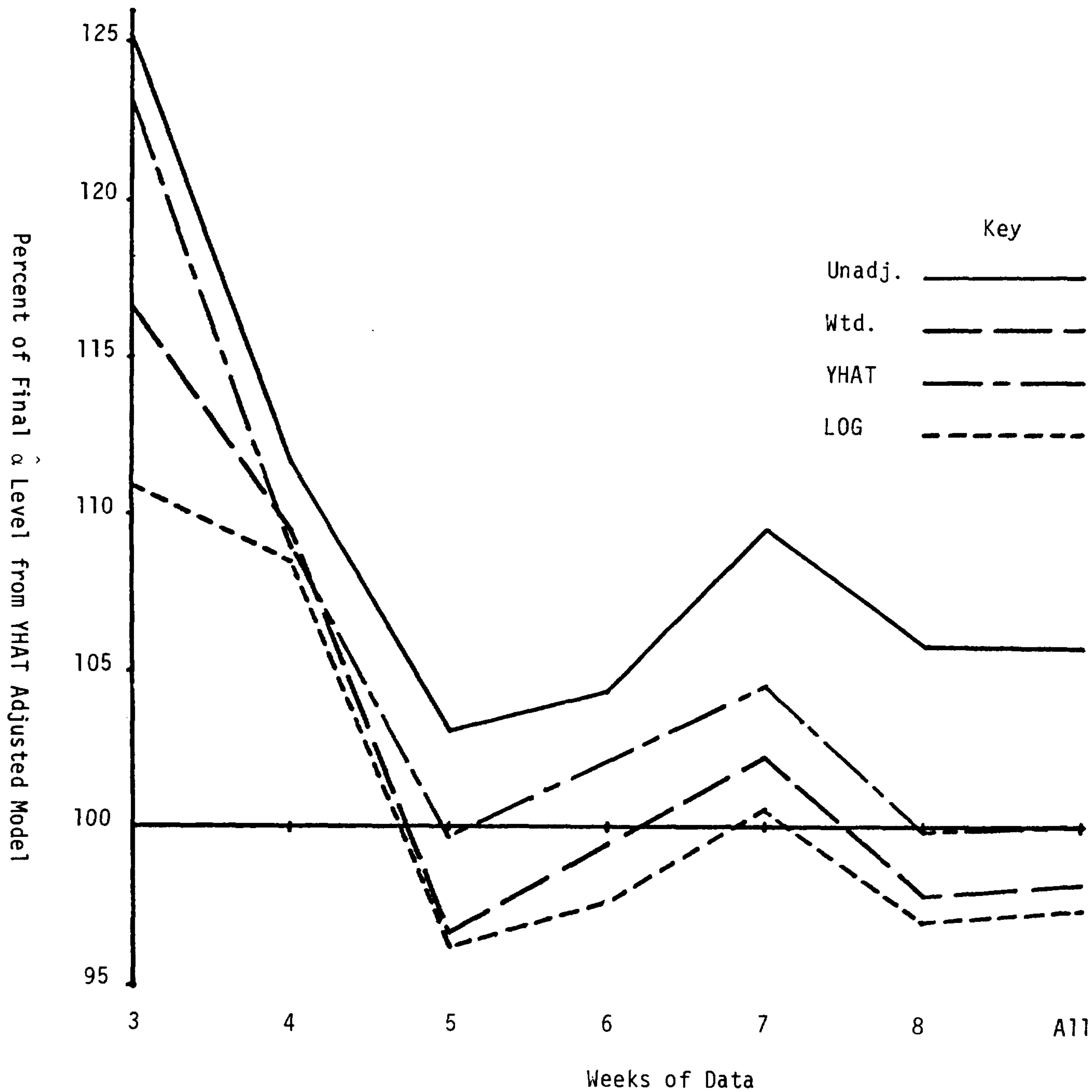
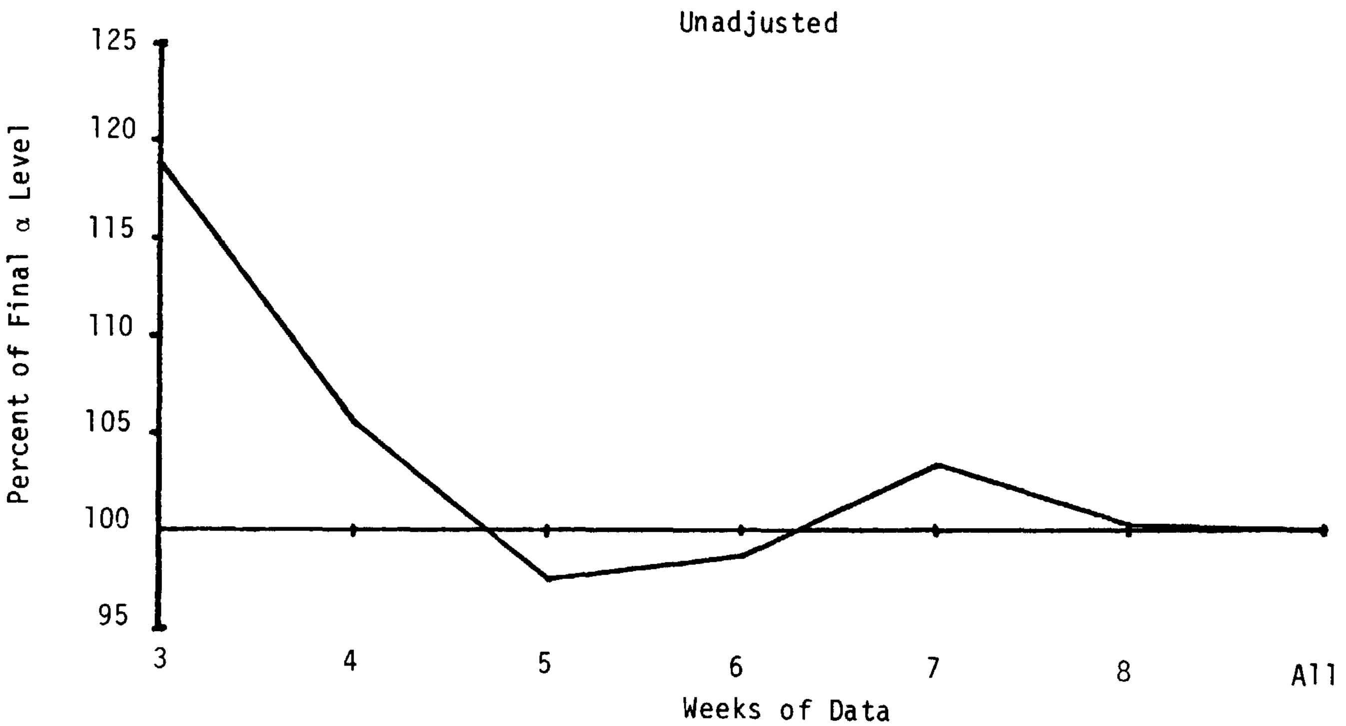

Percent of Final α̂ Level from YHAT Adjusted Model
125
120
115
110
105
100
95
3
4
5
6
7
8
All
Weeks of Data
Key
Unadj.
Wtd.
YHAT
LOG

Diagram 2

Diagram 3


Unadjusted
125
120
115
110
105
100
95
Percent of Final $\hat{\alpha}$ Level
3
4
5
6
7
8
All
Weeks of Data

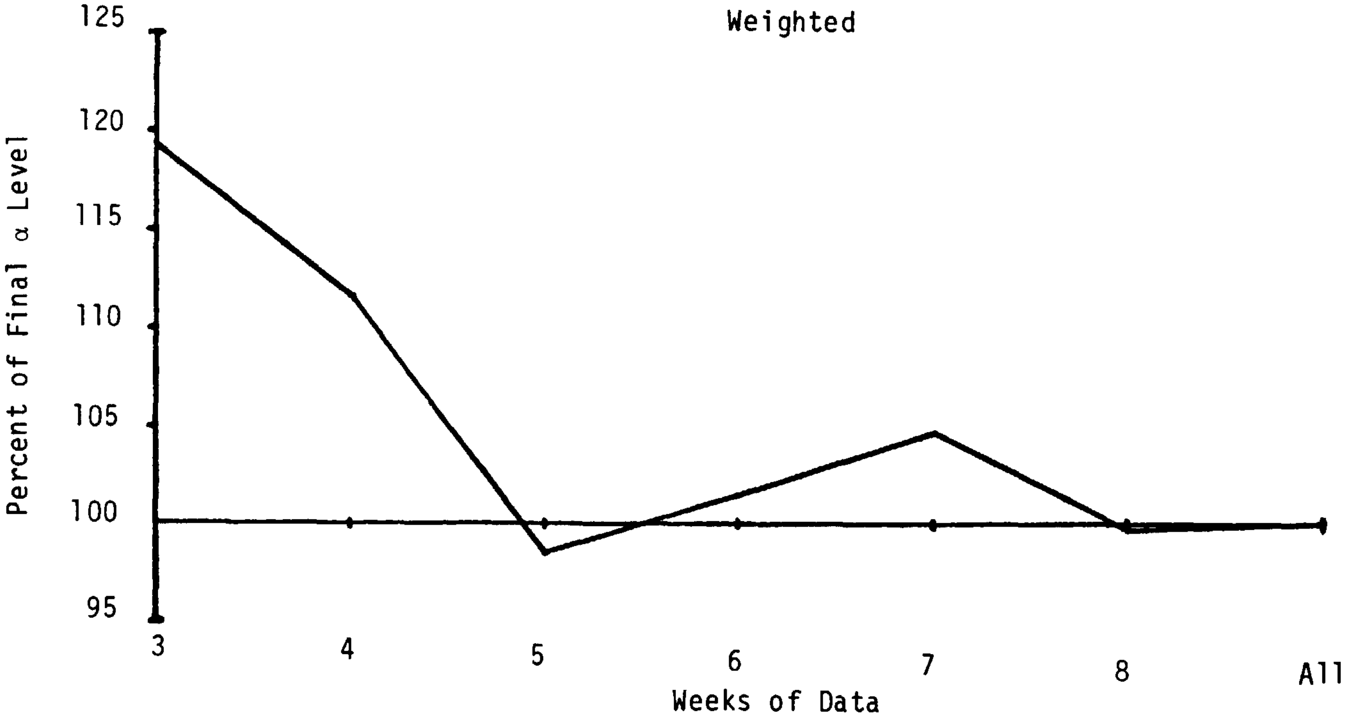
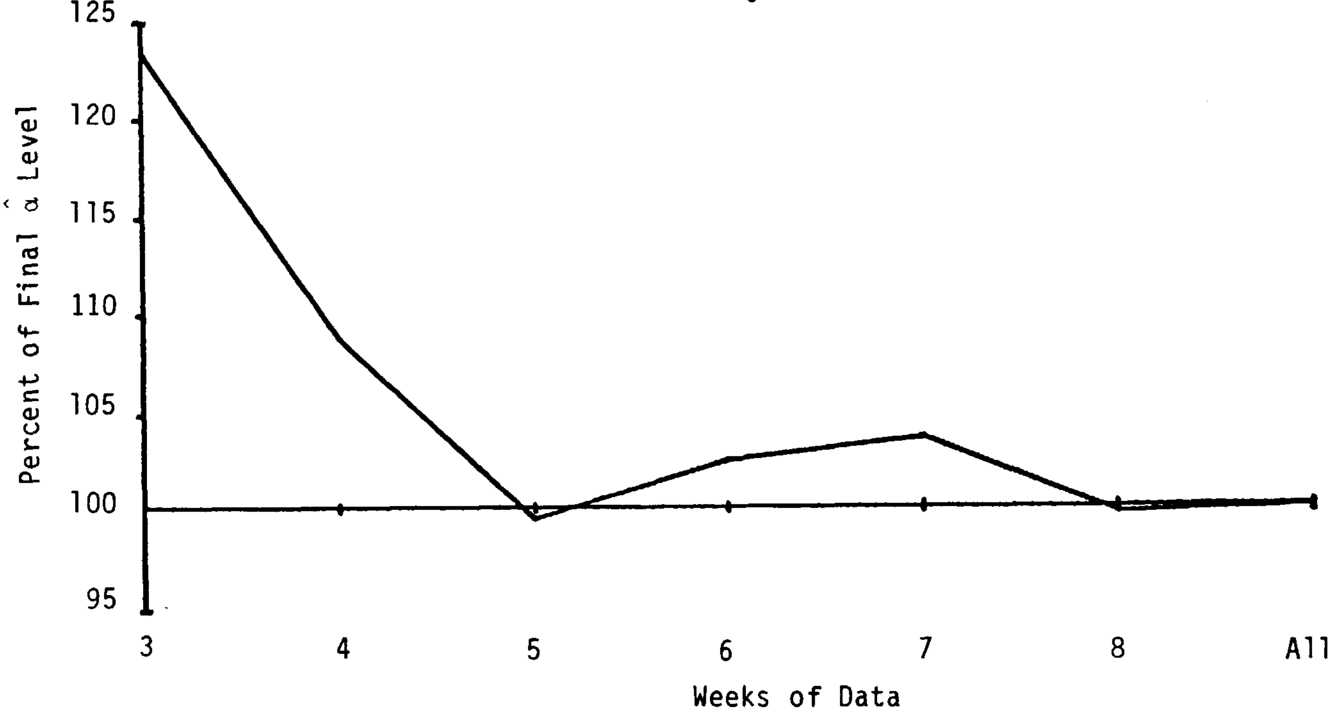
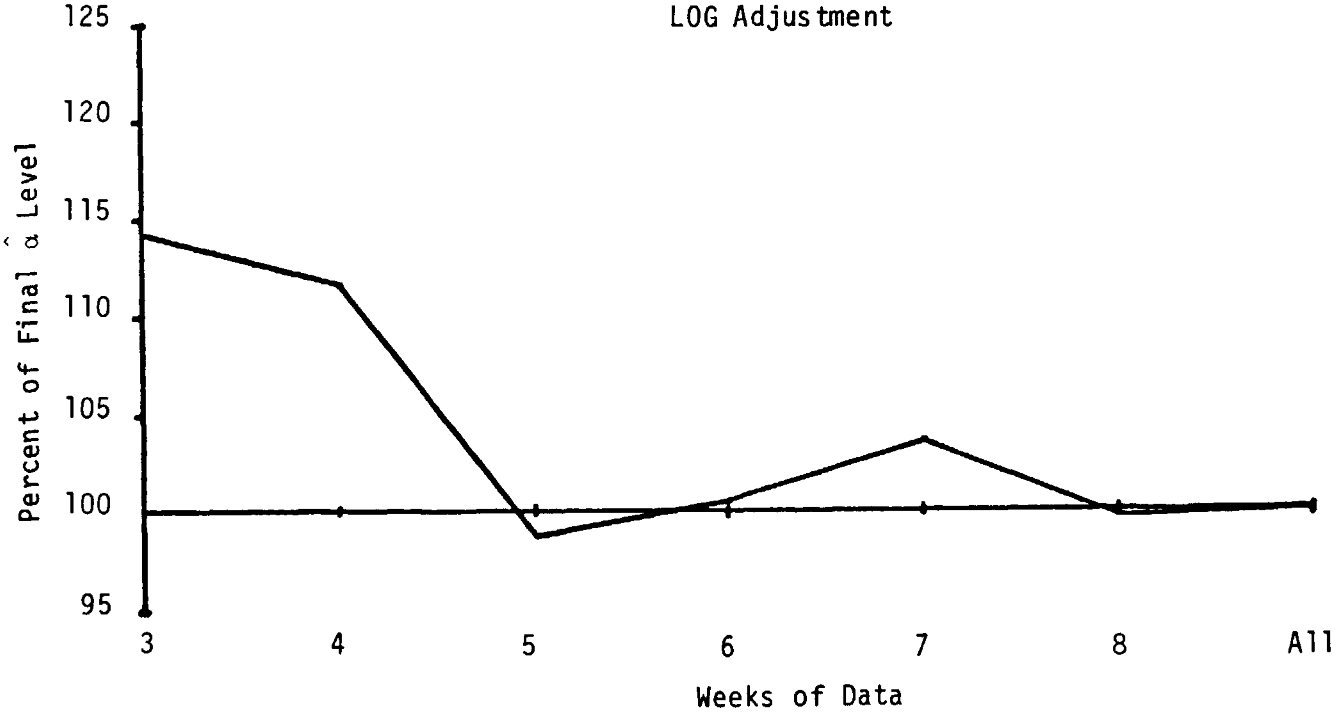

Weighted
125
120
115
110
105
100
95
Percent of Final $\hat{\alpha}$ Level
3
4
5
6
7
8
All
Weeks of Data

Diagram 3 (cont'd)

YHAT Adjustment



LOG Adjustment



Since the stability of the models is about the same in all cases, the main determinant in selecting a preferred model in a given situation is the correlation coefficients. The double LOG model is preferred for weeks one through four while the YHAT adjusted model is preferred for the fifth week through the end of the season. Notice that this choice of models provides correlations which are not significant at the α = .05 level and the resultant forecasts show more stability than those from any of the individual models.

Influence of Varietal Differences

It has been previously discussed that, as far as the 1977 season is concerned, it would not improve the forecast of mean dry head weight to fit separate models for early, normal, and late developing fields. The reason for this is that early and normal fields showed very little difference in the forecasted mean dry head weight and, even though the late fields were at a substantially higher $\hat{\alpha}$ level, it would not be of help to model the late fields separately from a forecasting point of view. It has also been previously discussed that, with three weeks of data available, a reliable June 1 forecast of mean dry head weight could not be made. While the principal reason for this failure is an insufficient amount of data, a contributing factor is the between field variation that is going into the model. The fact that different fields have different yields and, presumably, different mean dry head weights does not in itself inhibit the ability of the growth model to forecast a mean dry head weight over all fields. However, fields that have different rates of growth, and hence, differently shaped growth curves, would make fitting a single model with only early season data more difficult. The reason for this is that in the early season many of the fields would not have enough data to get much beyond the inflection point of the growth curve. The inflection point is the point on the curve where the slope of the tangent is at a maximum. In Diagram 1, t_u would be the time where the growth is the most rapid and therefore corresponds to the point of inflection. If different fields have different rates of growth and there is not sufficient data to get much beyond the inflection point in many fields, the estimate of α will have additional unreliability.

While weather factors and other growing conditions affect the rate of growth in a field, it is also recognized that varieties of wheat are developed with different maturity characteristics. Early maturing varieties are more likely to escape damage from hot winds, drought, and rust but are more susceptible to late spring freezes. The degree of winter hardiness that a variety exhibits tends to be inversely related to the maturity characteristics in that the early maturing varieties are less winter hardy than the later varieties. For this reason, early varieties are recommended for the southern and eastern portions of Kansas while the later varieties are generally planted in the northern and western parts of the state. In addition to maturity and winter hardiness, other criteria used to select a variety in a particular environment include quality, hereditary yield potential and resistance to insects and disease. While varieties have some correspondence to geographic location, several varieties generally dominate a particular area rather than just one and some varieties are versatile enough to be planted in any part of Kansas.

A limited investigation was made to see if growth curves could be fitted on a variety basis, and if so, examine the differences between the curves. Fifty of the 67 fields in the study fell into five major varieties. The remaining 17 fields could be combined with the major varieties by combining varieties with similar characteristics. For the purpose of this discussion, however, this was not done.

The five varieties that occurred most frequently in the order of importance are Scout, Eagle, Sage, Triumph, and Centurk. Eagle is a selection from Scout. Sage is a cross between Scout and Agent. Scout, Eagle, and Sage are quite similar and adapt well to any area of the state. Triumph is an early maturing variety which is suitable for the southern and eastern parts of Kansas while Centurk is the latest maturing of the five varieties and is recommended for northern and western protions of the state. Table 3 contains information on each of the five varieties and was taken from a Cooperative Extension Service publication.[1/]

Table 3

Varietal Agronomic Characteristics

Variety	Yield Potential	Maturity[1]	Winter Hardiness[2]
Scout	Very good	3	3
Eagle	Equal or superior to Scout	3	3
Sage	Potentially one of highest yielding varieties. Wide adaptation for yield	3	3
Triumph	Very good in its area of adaptation	1	7
Centurk	Excellent	4	2

1. Rating scale of 0 to 9. 0 is earliest, 9 is latest.

2. Rating scale of 0 to 9:

0 = excellent
1 to 3 = good
4 to 6 = average
7 to 9 = poor

[1/] Wheat Production Handbook, 1975. Cooperative Extension Service, Kansas State University, Manhattan, Kansas.

The weighted model (5) with no heteroscedasticity adjustment was fit to the data for each of the five varieties. Figures 7 through 11 show a plot of the data for each variety and the corresponding fitted growth curve. Table 4 summarizes the pertinent information for each variety. Several of the entries in the table need explanation. The estimated values of β and ρ have been provided for each variety. If α and ρ are held constant, increasing values of β would cause the resulting growth curves to have steeper slopes. This would imply that the wheat head is growing more rapidly. If α and β are held constant, increasing values of ρ would make the resulting growth curves expand horizontally. This implies that the time it takes for a wheat head to reach its full weight would increase. Therefore, a relatively high value of β and a low value of ρ would correspond to an early maturing variety. The opposite would reflect a late maturing variety.

The mean dry head weight over the entire range of time and the standard deviation of the mean dry head weights are provided for each variety and are denoted by $\overline{Y}$ and $\hat{\sigma}_Y$ respectively.

The magnitude of $\overline{Y}$ should correspond to the magnitude of $\hat{\alpha}$ if the growth curve is doing a good job of representing the data. Similarly, the same should be true for $\hat{\sigma}_Y$ and the MSE. The reason the MSE's are of higher magnitude than the standard deviations is because the weighted model was used. It can be seen in Table 4 that $\overline{Y}$ and $\hat{\sigma}_Y$ correspond well to $\hat{\alpha}$ and the MSE for each of the varieties.

In addition to the estimates of β and ρ, another indication of the maturity characteristic of each variety is the time it takes for the fitted growth curve to reach the asympotote. Since an infinite number of days are needed to reach the exact asymptote, the number of days it took for each growth curve to reach 98% of $\hat{\alpha}$ is presented in Table 4.

The point of inflection for the growth curve is found by setting the second derivative with respect to t equal to zero and solving for t. If this is done, t is found to be $-(\ln \beta)/(\ln \rho)$. The point of inflection for each of the fitted growth curves is in the neighborhood of 10 to 15 days since flowering.

Another entry in Table 4 is the average number of stalks per square foot. This is a mean in which the number of stalks in a plot was weighted by the number of observations from the plot. The average number of stalks per square foot is from early in the growing season prior to heading and does not take into account mortalities and stalks that do not produce flowered heads. To get an indication of the yield potential of each variety, the optimum biological yield was calculated by taking the product of the average number of stalks per square foot and $\hat{\alpha}$.

Table 4

Comparisons Among the Five Major Varieties

	Scout	Eagle	Sage	Triumph	Centurk
No. of fields	18	12	8	7	5
Observations	91	62	45	42	31
MSE	61.08	72.80	123.58	91.99	69.70
R^2	.951	.928	.909	.957	.972
$\hat{\hat{\alpha}}$	1.0440	.8234	1.0807	1.0850	.8655
$\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$	7.5	8.7	22.9	6.9	4.9
$\hat{\beta}$	4.7467	4.0443	3.6066	4.6135	6.7363
$\hat{\rho}$	.8973	.8750	.9208	.8881	.8581
$\overline{Y}$	.696	.655	.724	.850	.625
$\hat{\sigma}_Y$	.292	.320	.351	.321	.274
Time to 98% of $\hat{\alpha}$ (days)	50	40	63	41	38
Point of Inflection (days)	14.4	10.5	15.6	12.9	12.5
Avg. Stalks per Ft^2	42.65	40.64	48.78	50.18	67.76
Opt. Bio. Yield	44.53	33.46	52.72	54.45	58.65

In comparing Table 4 and Table 3, it can be seen that Triumph was one of the earlier maturing varieties and showed the second highest yield potential. Centurk did not mature late in 1977 as would be indicated in Table 3. It was the earliest maturing variety. It did, however, show the highest yield potential. Sage turned out to be the latest maturing variety and showed a comparatively high yield potential. In general, the data tends to support many of the general characteristics outlined in Table 3.

Since Figures 7-11 have different scales, Diagram 4 shows each of the five fitted growth curves. The dotted lines show the point where each curve reached 98% of the $\hat{\alpha}$ value. From 10 to 20 days which is around the points of inflection and a few days beyond, the fitted curves for Scout, Eagle, and Sage are very similar. Centurk and Triumph are increasing at a faster rate in that interval. With the three weeks of data that were available for a June 1 forecast, most of the data was less than 20 days and none of it went past 35 days. If a single model is fit to the data, the varietal differences in rate of growth would cause a poorer fit and add unreliability to the estimate of α. If the variety curves were identical in shape so that β and ρ were the same for each, a single curve would estimate a value for α which would be the weighted average of the $\hat{\alpha}$'s from the individual varieties. Therefore, small differences in rate of growth for varieties would not be of much harm when fitting a single model. Since the growth is quite different, this suggests that modeling by variety would improve early season forecasts.

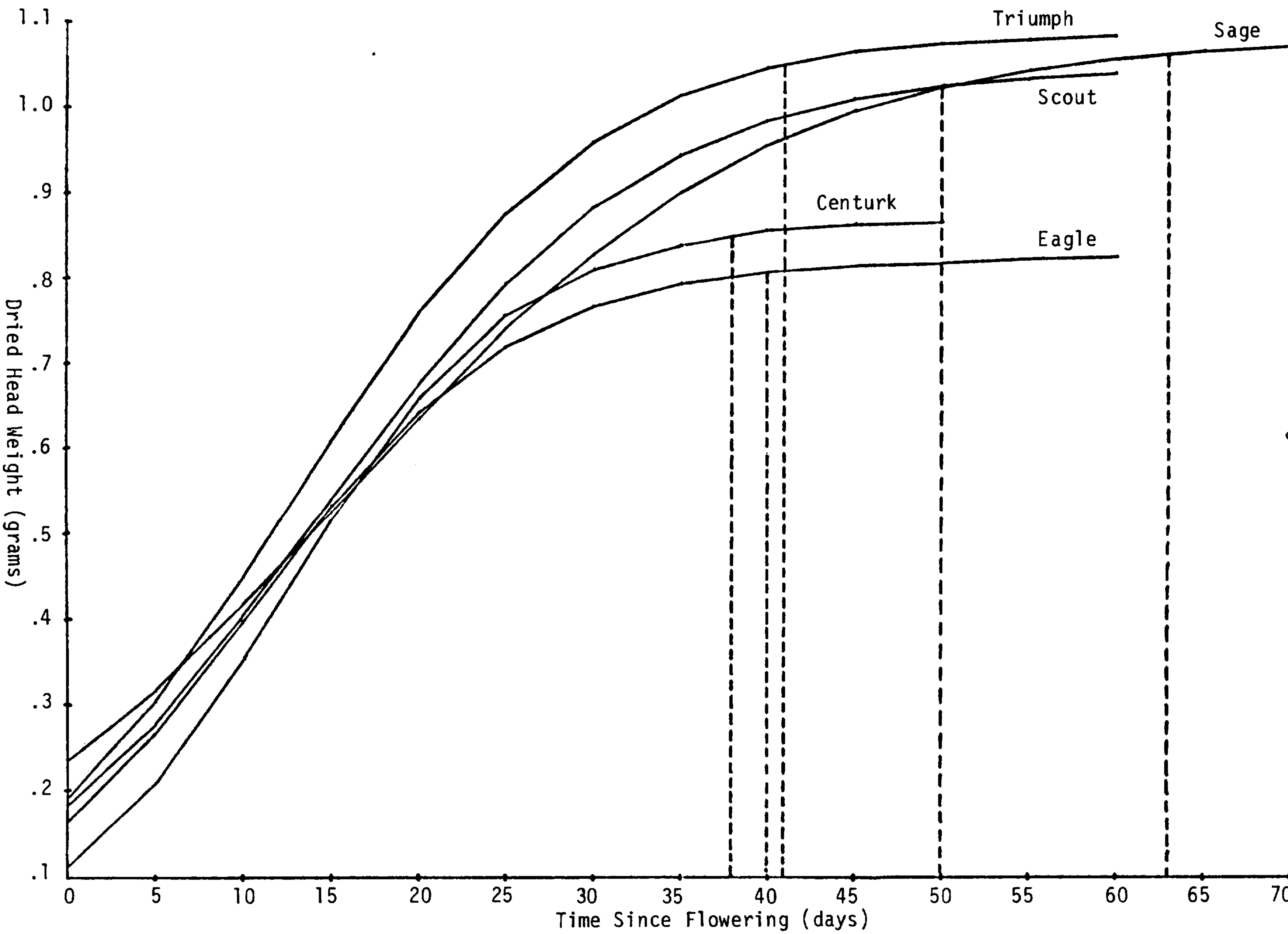

Dried Head Weight (grams)
1.1
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
Time Since Flowering (days)
Triumph
Sage
Scout
Centurk
Eagle

Diagram 4

Conclusions and Additional Research

The main result from the 1977 wheat research study was that a reliable forecast of mean dry head weight at maturity could not be made on June 1. However, the within-year growth model was able to make a forecast with an error of approximately 5% in mid-June.

The possibility of fitting separate models for different geographic areas based on the relative development of the wheat did not appear to be of use. It was found that "early" developing fields were very similar to "normal" developing fields in terms of mean dry head weight. The "late" developing fields demonstrated a higher mean dry head weight but separate modeling would not be of much help in a forecasting mode. Fitting separate models by variety did, however, appear to be valuable. The inability of a single model to provide a reliable forecast by June 1 can be at least partially explained by the between field variation caused by varietal growth rate differences. It was possible to fit a separate model for major varieties, and presumably, minor varieties could be combined with the major classes on the basis of similar agronomic and maturity characteristics. Forecasts from separate variety models could be weighted together to form a state forecast which might prove to be more reliable around June 1. Data in the 1977 study was not sufficient to fully investigate this possibility.

Adjustments to the model to correct for heteroscedastic disturbance were able to create residuals which were not significantly correlated with time. The adjustments worked extremely well when the data was simulated.

Several problems which came to light during the analysis will be discussed at this time. The first problem relates to the method of aggregation. As discussed earlier, individual head weights were aggregated to the field level based on intervals of time since flowering. This was done so that the observations entering the model would be independent. Another generally excepted method of aggregation is to combine all observations made on a particular visit within a field. Because of the two-level nested design, plots within a field are independent so means are calculated for each plot and plot means are averaged. This provides one mean dry head weight per visit in each field. The difference between this method of aggregation and the one which was used is that individual head weights are combined on a calendar date basis rather than with regard to the time since flowering. The 1977 data revealed that for a fixed calendar date visit, the number of days since flowering for clipped heads within a field differed by as much as two weeks. Aggregation requires that averages be taken over both time and weight. Since the success of the growth model is dependent upon the relationship between time since flowering and head weight, there is concern that aggregating by visit could obscure any real time-growth relationship. Aggregation based on intervals of time since flowering was used so that any real time-growth relationship could be better preserved.

A second problem encountered in the analysis concerns the weight which each field receives in the model. The regular wheat objective yield fields are selected with probability proportional to acreage so that field level yield observations can be weighted together equally to form an average state yield. This yield multiplied by acreage will provide a statistically correct production figure. Since the research fields are a random subset of the objective yield sample, they should also receive equal weight. A problem which seems to be inherent in the sample design which has been used is that the number of observations per field varies. This is due to the fact that data is collected weekly until harvest. Regardless of whether aggregation is done by time intervals or visits, the variability in the data collection period is likely to produce differing numbers of observations within fields. This means that fields in which harvest occurs closer to flowering would have less weight in the model than fields which have longer periods between flowering and harvest. Clearly the fields do not receive equal weight in the model as the sampling method intends. A possible solution to this problem is to weight each mean by the inverse of the total number of time intervals for the field from which it came. It is recommended that this approach be investigated in future research.

A third problem is that the sample design did not provide an estimate of flowered stalks per acre. The reason that it didn't is because of the 80% rule mentioned earlier which signals the start of clipping. This rule was employed so that field enumerators could have a simple way to know when to begin clipping. It also eliminated the close monitoring of crop development by office staff which would have been necessary to make decisions as to when clipping should begin in each field. As a result of using the general 80% rule, some heads were clipped before they had a chance to flower, and hence, a total count of flowered tagged stalks per plot is not available. As discussed earlier, it was possible to make an estimate of heads per acre at maturity. Experience has shown that flowering is a better indicator of grain production than is just the presence of a head. This is because stalks that do not flower do not produce grain and not all stalks with heads flower.

An additional problem encountered in the data analysis is that the relationship between the dry head weight and the standard grain weight from paired heads which was described earlier showed considerable variation. The pairing of adjacent heads apparently does not exhibit the similarity that was hoped for. The use of this relationship would add to the variability of the biological yield forecast.

Additional within-year growth model research is being conducted for the 1978 season. Results from the 1977 study have brought about a different emphasis for the 1978 study. Because of the unreliability of the June 1 forecast on a state level, the 1978 study will avoid the problem of between field differences in growth rate by operating on a field level. A net yield is to be forecasted for each of four commercial fields located in Ellsworth County, Kansas. A sufficient number of plots and tagged stalks will be used to account for the within field variability of dry head weights. To at least partially alleviate the aggregation problems in the 1977 study, only flowered heads will be sampled and it will be done on the basis of time since flowering rather than calendar dates. This complicates the data collection and requires that the field enumerators be thoroughly trained and that their work be closely monitored.

An estimate of the number of flowered heads at maturity will be available. Shortly before harvest, dry kernel weights will be determined on paired heads as well as the regularly clipped heads. This will allow the estimation of a dry head weight to standard kernel weight adjustment from the same head along with a comparison to the adjustment obtained from paired heads. Post-harvest plots similar to those used in the regular wheat objective yield program will be used to obtain harvest loss estimates for each field. The data collected for the 1978 study will provide forecasts of net yield per acre for each field.

The size of each research field will be determined and the actual production will be obtained at the elevator. This will provide an actual yield per acre for each field which can then be compared with the yield forecasts coming from the within-year growth models. Modeling at the field level greatly reduces the problem of variety related maturity differences. If wheat yield can be successfully modeled on a field level, this should provide insight into problems encountered in large area modeling. It also is conceivable that if costs for field level forecasts can be made low enough, large area forecasts might be obtained from aggregation of individual field results.

Appendix A

A. Simulation

Simulated data is of value in determining the effectiveness of heteroscedasticity adjustments when the data is generated in such a way that it closely resembles the actual data. Dry head weights tend to be normally distributed for a particular time value in that the heaviest concentration of points is around the mean. While normality is not a necessary assumption for obtaining least squares parameter estimates, it is assumed when various hypotheses are statistically tested. A data set was simulated with a random number generator which produced dry head weights appearing to be normally distributed with a specified mean and variance. The mean was the expected value of the dependent variable when the unadjusted model was fit to the data. The variance was the square of the standard deviation of the mean dry head weights in each of the seven time intervals. A plot of the simulated data and the fitted regression curve appears in Figure 3. The corresponding residual plot is shown in Figure 4. Although the scales are different, a comparison of Figure 3 with Figure 1 shows that the simulated data looks much like the actual data.

Both of the heteroscedasticity adjustments mentioned earlier were applied to the simulated data. Some of the results are summarized in Tables 5.1 and 5.2. The plot of the YHAT adjusted simulated data and the fitted adjusted regression curve is shown in Figure 5. The corresponding residual plot appears in Figure 6. From Table 5.1 it can be seen that the YHAT adjustment did an excellent job of producing uncorrelated residuals evidencing a relatively constant variance over the range of time. In this case, the estimate of α decreased very slightly when the adjustment was used. To make sure that the apparent success of the YHAT adjustment was not dependent upon this particular simulation, independent data sets were simulated. Results for these simulations were very similar to those for the initial data set.

Table 5.2 shows that the LOG adjustment performed very well also. The simulated data used in Table 5.2 came from a separate run and hence, the results from the unadjusted model are different from those in Table 5.1. In this particular simulation, the LOG adjustment caused about a 2% decrease in $\hat{\alpha}$. Other simulation runs showed similar reductions in the correlation.

Table 5.1

Simulated Dry Head Weights

	Models	
	Unadjusted	YHAT Adjusted
MSE	.0331	.0331
R^2	.944	.940
$\hat{\alpha}$	1.0549	1.0513
$\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$	3.5	3.8
R_p	.301	.001
Prob>$\lvert R_p \rvert$	.0001	.9846

Table 5.2

Simulated Dry Head Weights

	Models	
	Unadjusted	LOG Adjusted
MSE	.0374	.0331
R^2	.939	.965
$\hat{\alpha}$	1.0103	.9859
$\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$	3.0	3.4
R_p	.268	.016
Prob>$\lvert R_p \rvert$	.0001	.7623

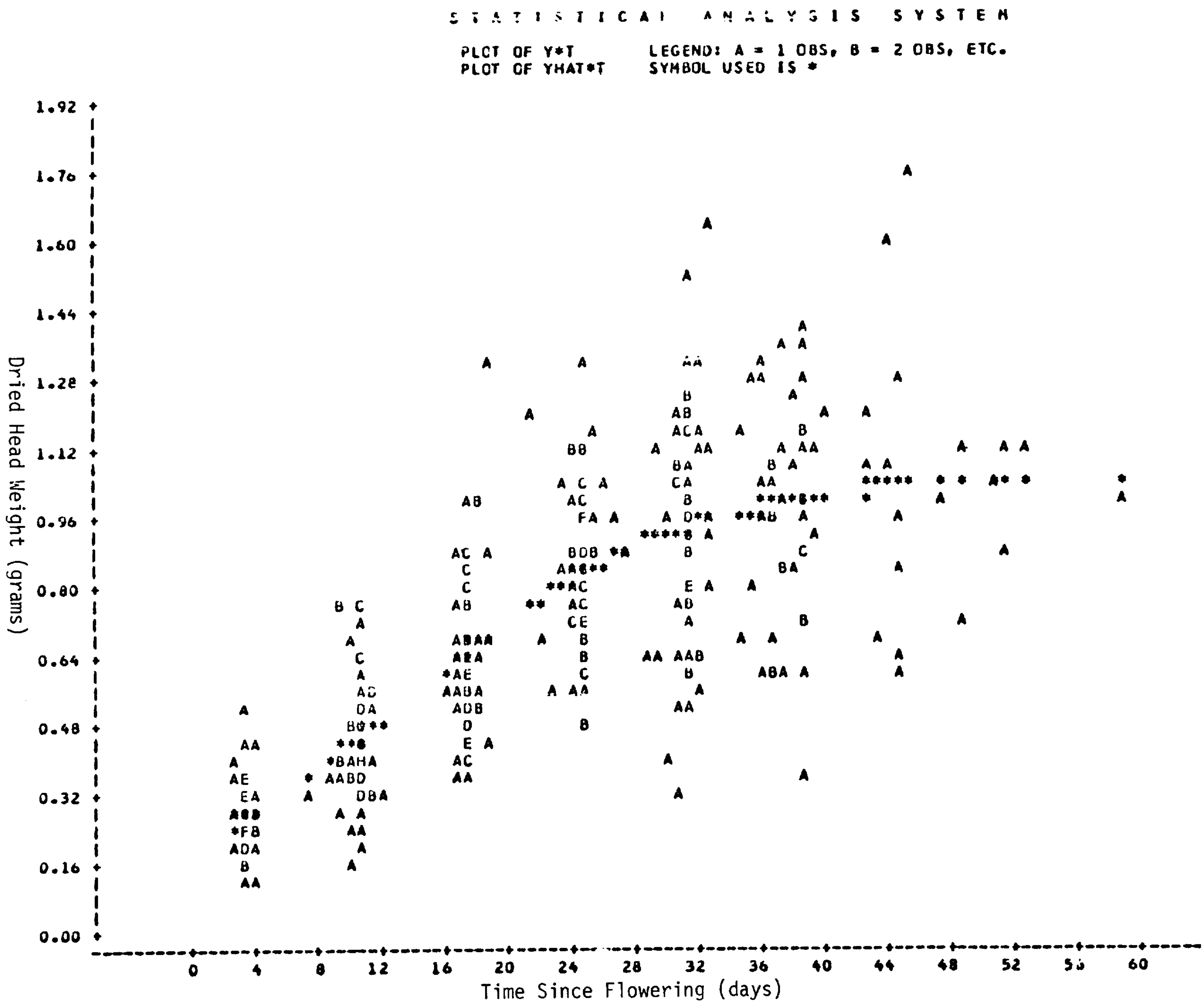

STATISTICAL ANALYSIS SYSTEM
PLOT OF Y*T LEGEND: A = 1 OBS, B = 2 OBS, ETC.
PLOT OF YHAT*T SYMBOL USED IS *
Dried Head Weight (grams)
1.92
1.76
1.60
1.44
1.28
1.12
0.96
0.80
0.64
0.48
0.32
0.16
0.00
0
4
8
12
16
20
24
28
32
36
40
44
48
52
60
Time Since Flowering (days)

Figure 1

STATISTICAL ANALYSIS SYSTEM

PLOT OF RES*T LEGEND: A = 1 OBS, B = 2 OBS, ETC.

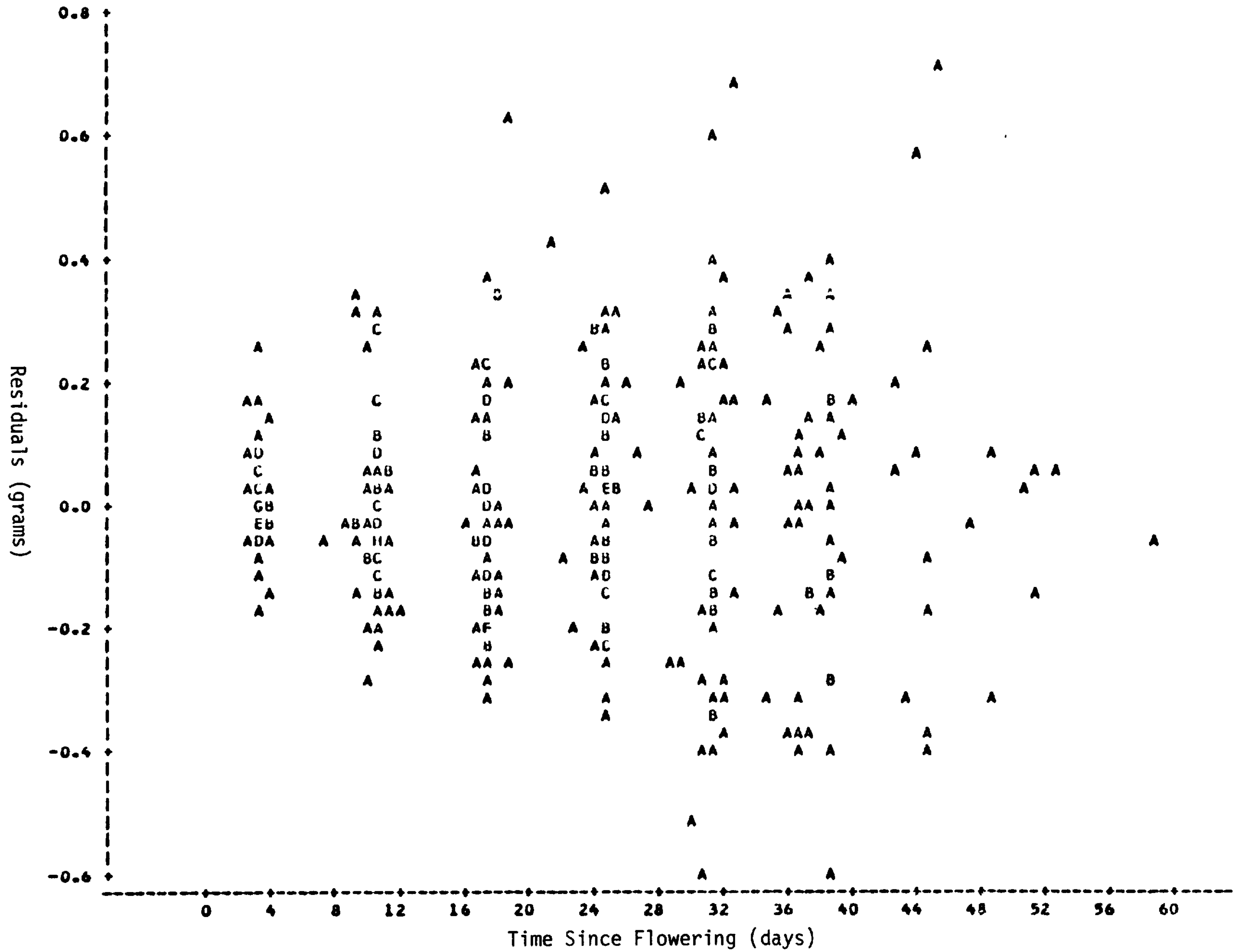


Figure 2

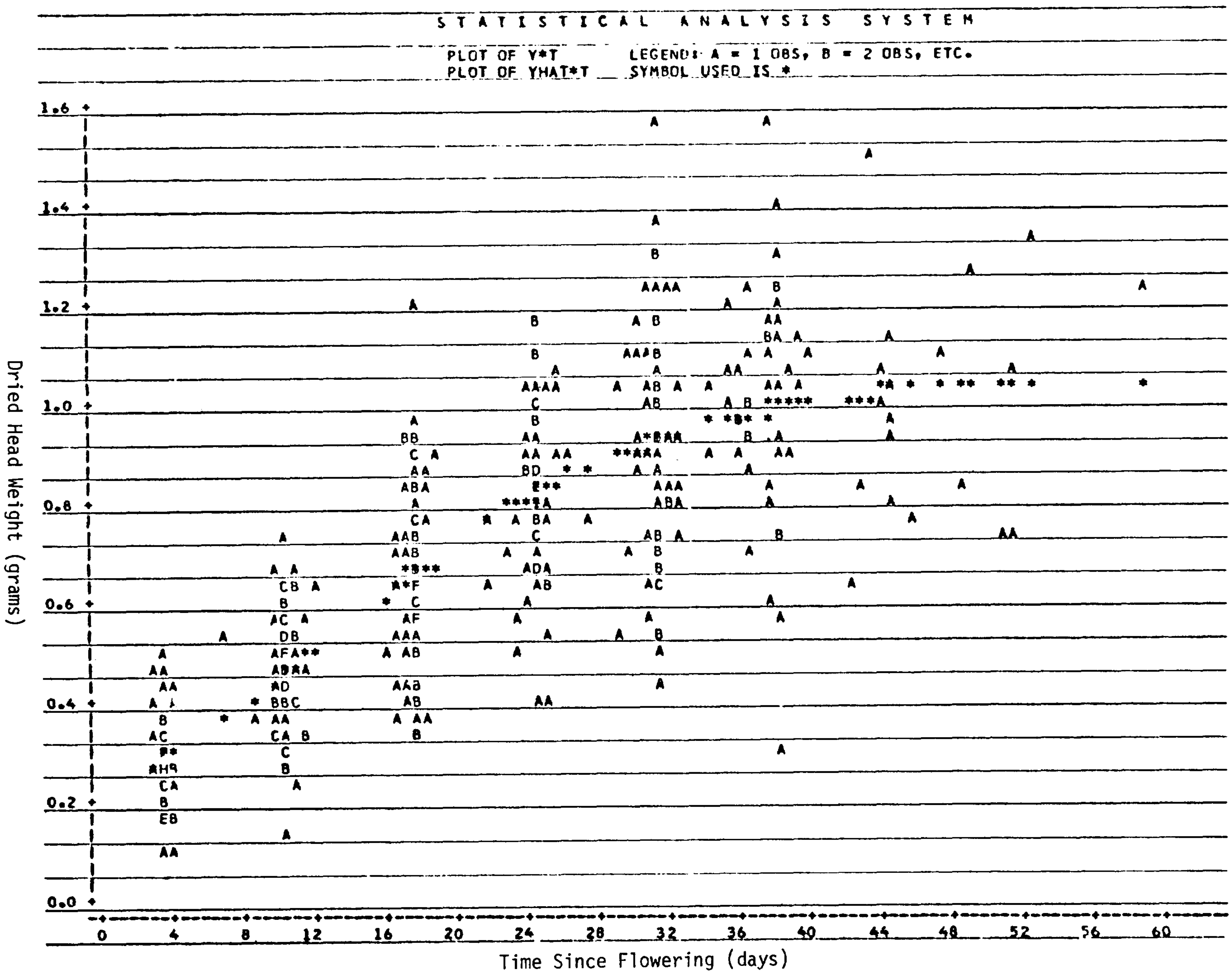
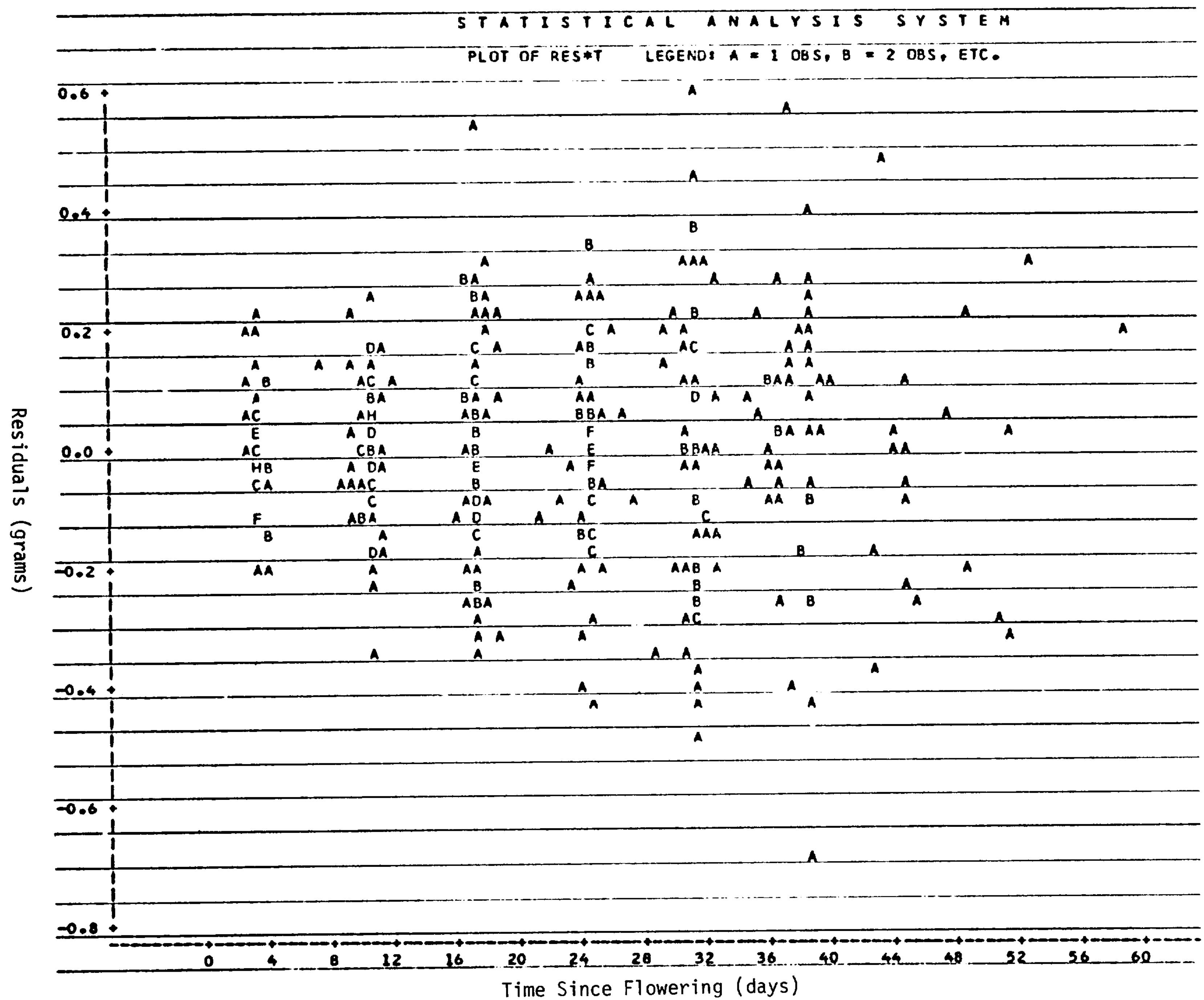

STATISTICAL ANALYSIS SYSTEM
PLOT OF Y*T LEGEND: A = 1 OBS, B = 2 OBS, ETC.
PLOT OF YHAT*T SYMBOL USED IS *
Dried Head Weight (grams)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
0
4
8
12
16
20
24
28
32
36
40
44
48
52
56
60
Time Since Flowering (days)

Figure 3



STATISTICAL ANALYSIS SYSTEM
PLOT OF RES*T LEGEND: A = 1 OBS, B = 2 OBS, ETC.
Residuals (grams)
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
0 4 8 12 16 20 24 28 32 36 40 44 48 52 56 60
Time Since Flowering (days)

Figure 4

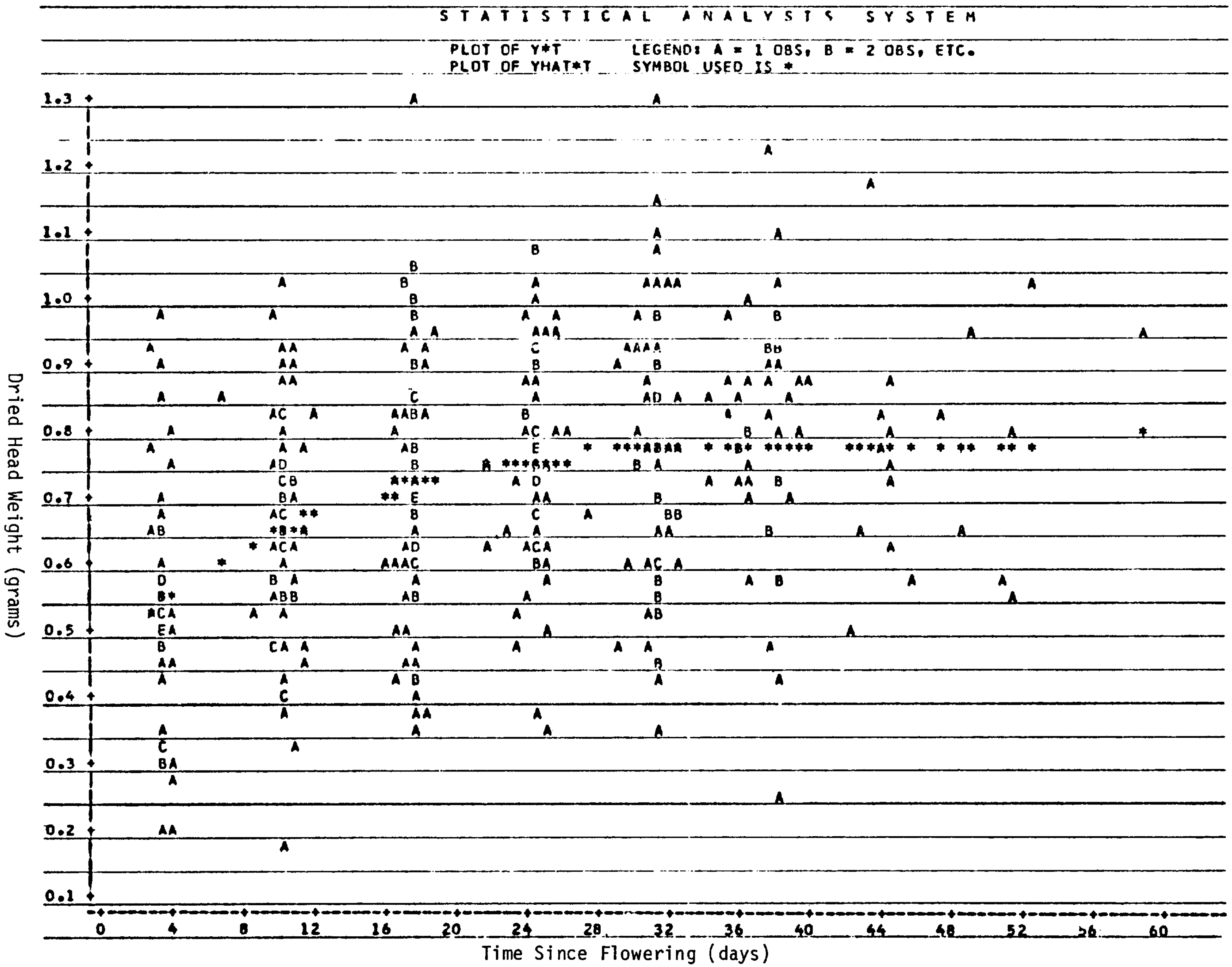

STATISTICAL ANALYSIS SYSTEM
PLOT OF Y*T LEGEND: A = 1 OBS, B = 2 OBS, ETC.
PLOT OF YHAT*T SYMBOL USED IS *
Dried Head Weight (grams)
Time Since Flowering (days)

Figure 5

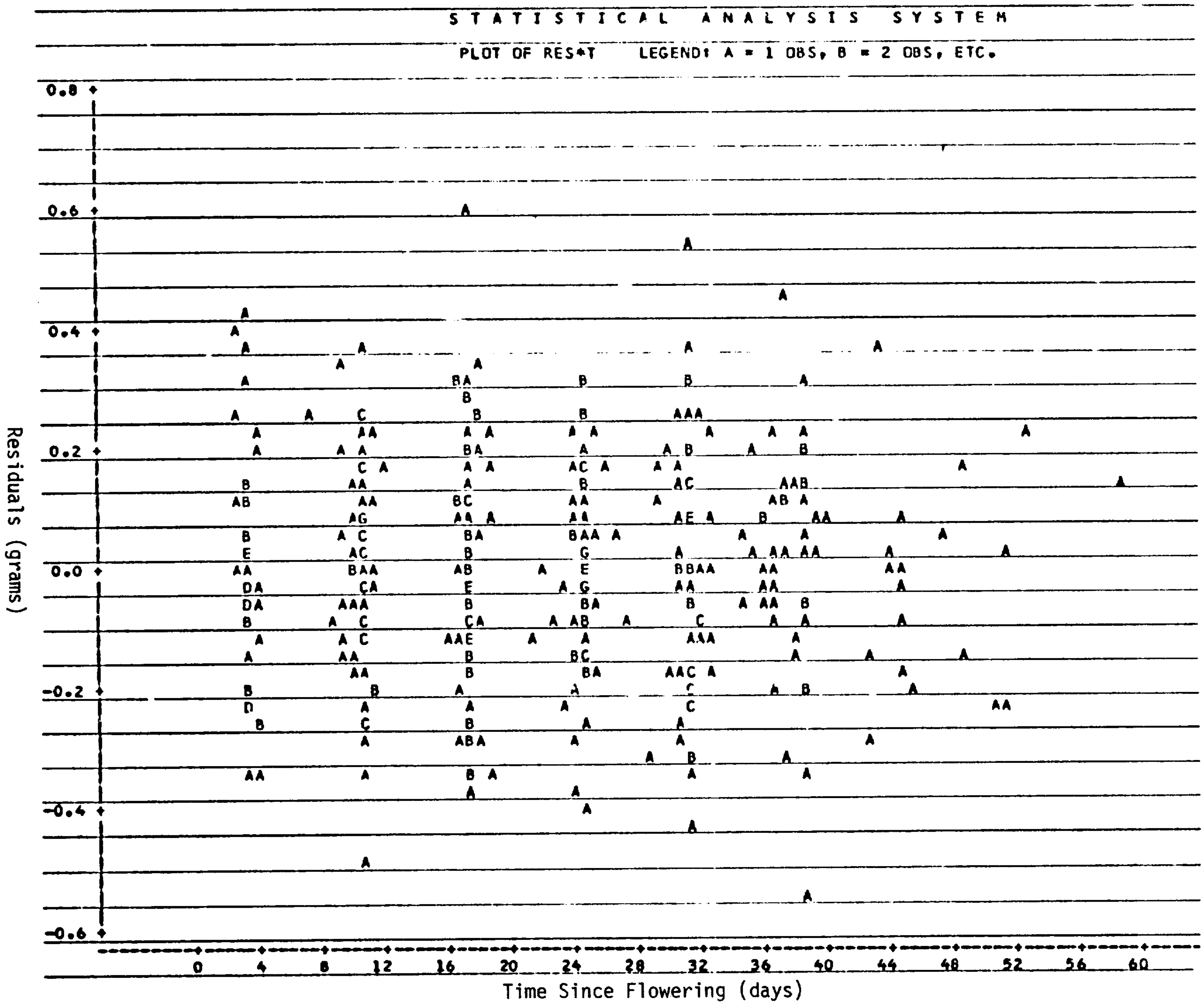

S T A T I S T I C A L   A N A L Y S I S   S Y S T E M
PLOT OF RES*T   LEGEND: A = 1 OBS, B = 2 OBS, ETC.
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
Residuals (grams)
0 4 8 12 16 20 24 28 32 36 40 44 48 52 56 60
Time Since Flowering (days)

Figure 6

STATISTICAL ANALYSIS SYSTEM

PLOT OF Y*T   LEGEND: A = 1 OBS, B = 2 OBS, ETC.
PLOT OF YHAT*T   SYMBOL USED IS *

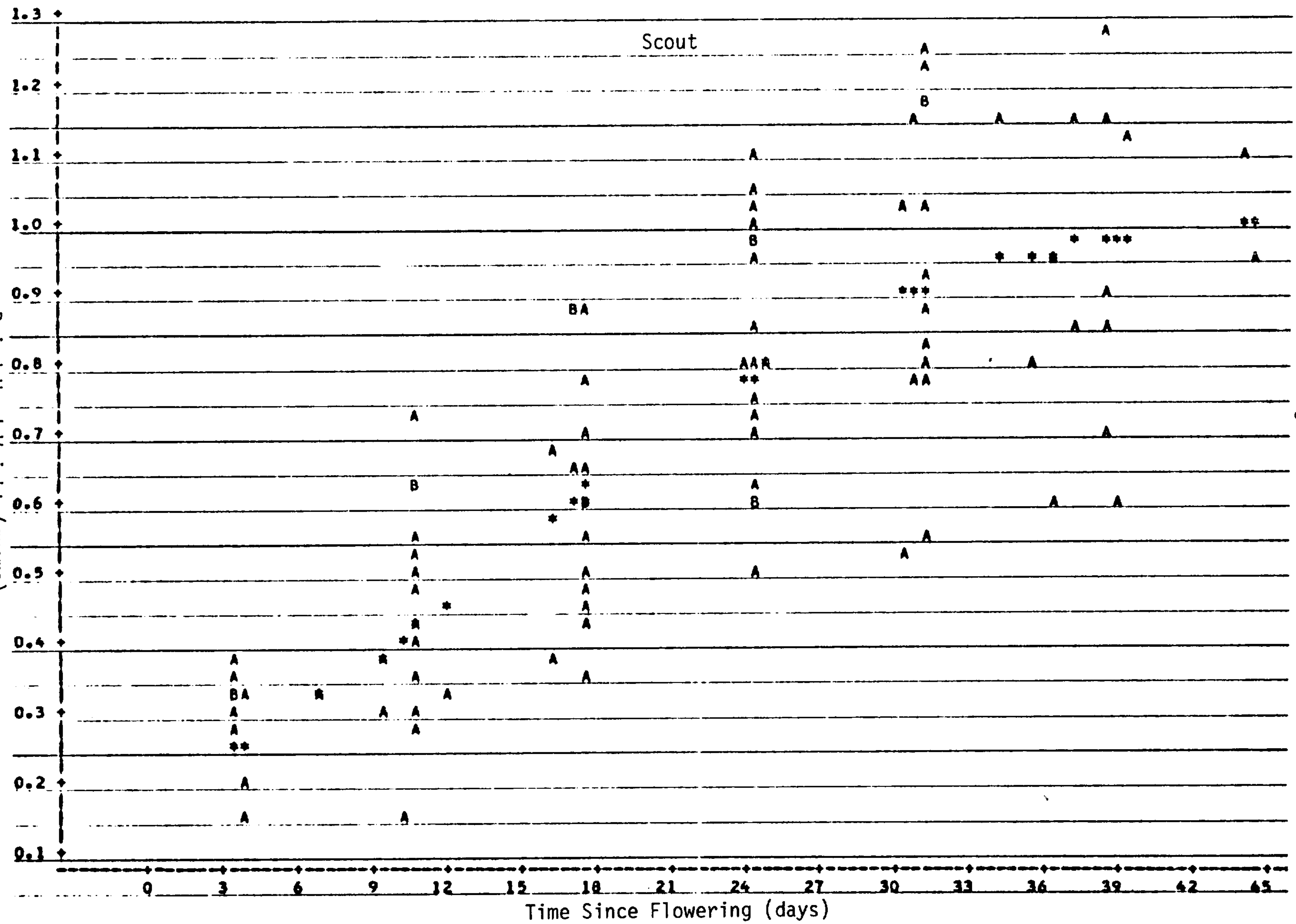


Figure 7

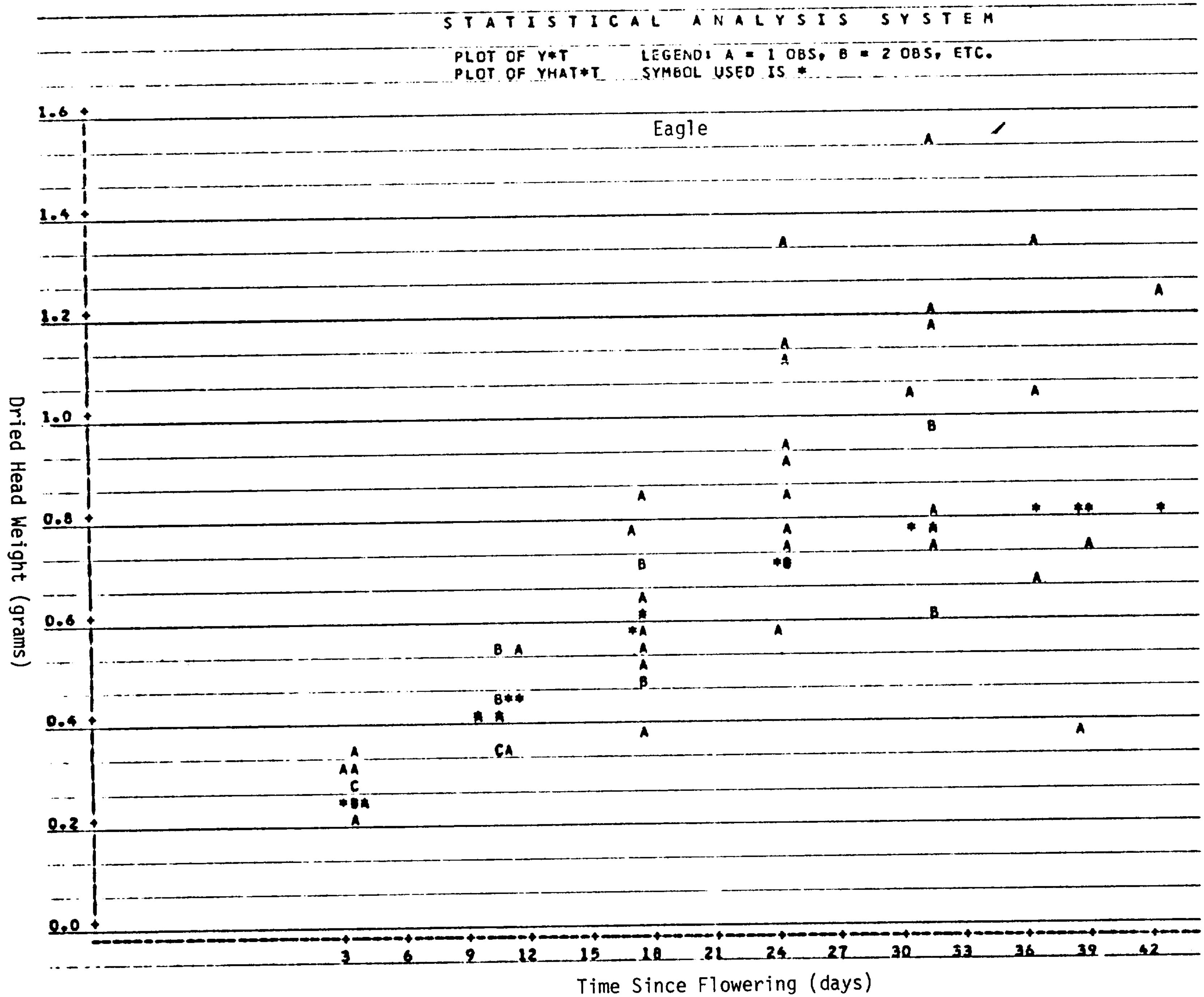

STATISTICAL ANALYSIS SYSTEM
PLOT OF Y*T LEGEND: A = 1 OBS, B = 2 OBS, ETC.
PLOT OF YHAT*T SYMBOL USED IS *
Eagle
Dried Head Weight (grams)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
3
6
9
12
15
18
21
24
27
30
33
36
39
42
Time Since Flowering (days)

Figure 8

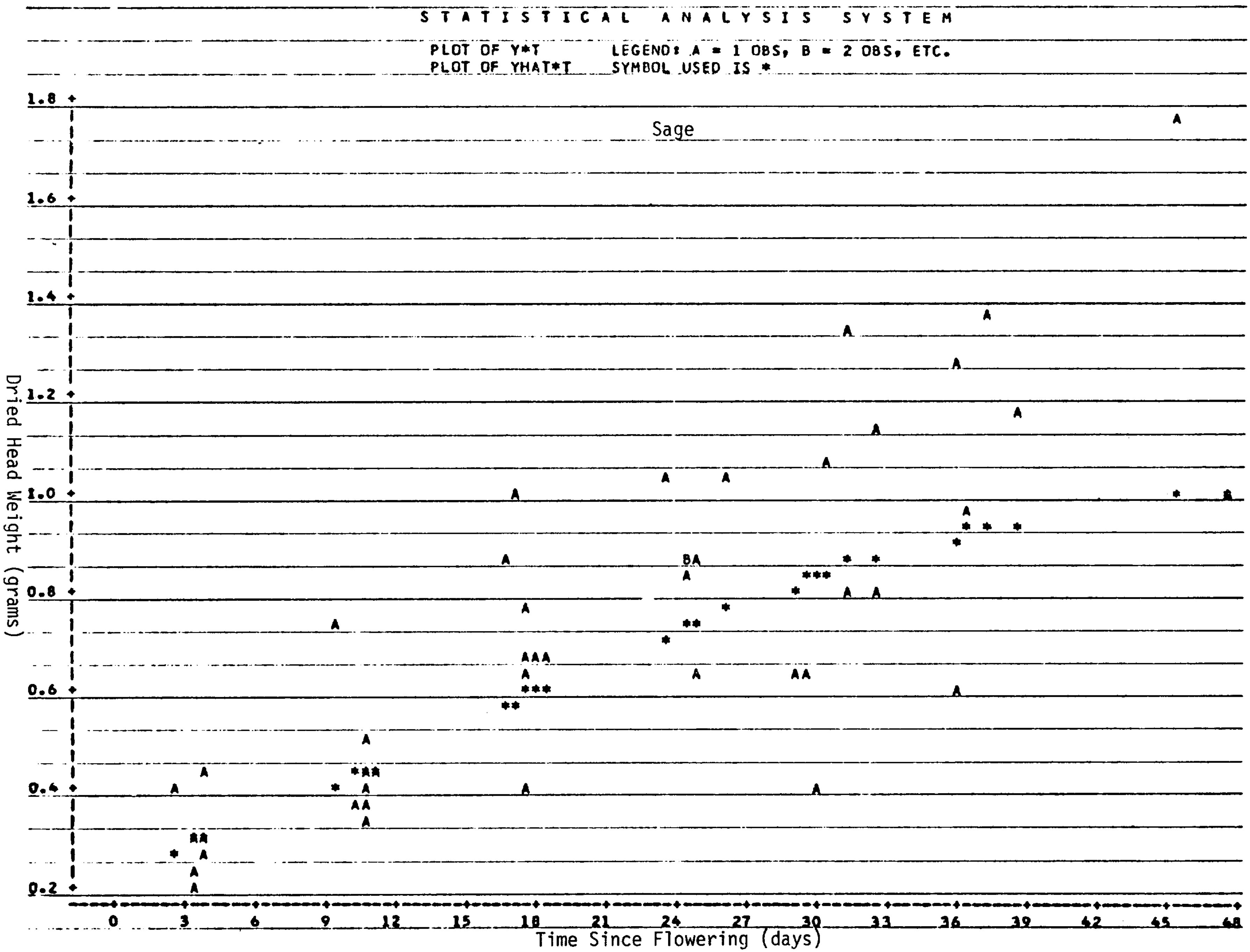

STATISTICAL ANALYSIS SYSTEM
PLOT OF Y*T LEGEND: A = 1 OBS, B = 2 OBS, ETC.
PLOT OF YHAT*T SYMBOL USED IS *
Sage
Dried Head Weight (grams)
Time Since Flowering (days)

Figure 9

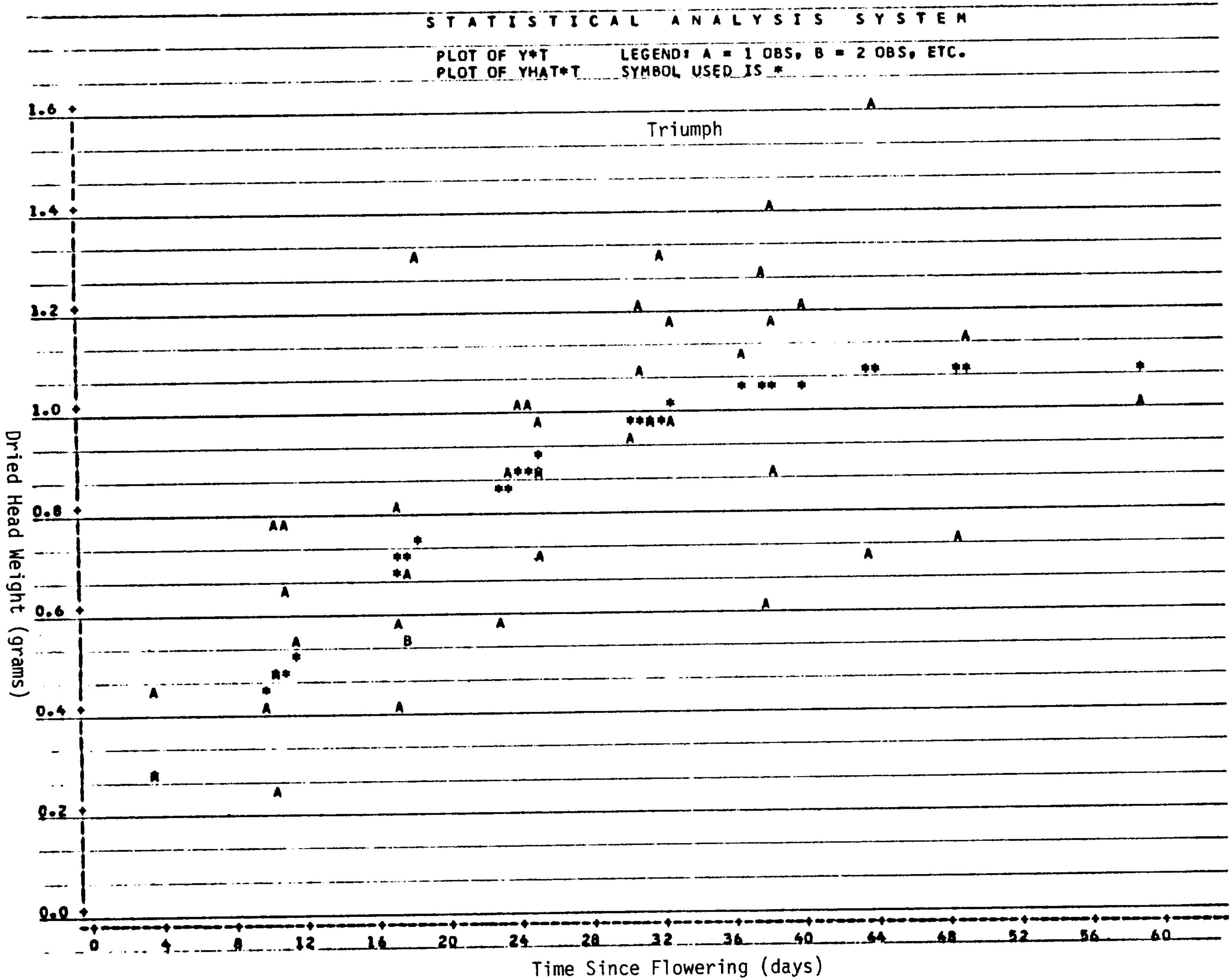

STATISTICAL ANALYSIS SYSTEM
PLOT OF Y*T LEGEND: A = 1 OBS, B = 2 OBS, ETC.
PLOT OF YHAT*T SYMBOL USED IS *
Triumph
Dried Head Weight (grams)
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0 4 8 12 16 20 24 28 32 36 40 44 48 52 56 60
Time Since Flowering (days)

Figure 10

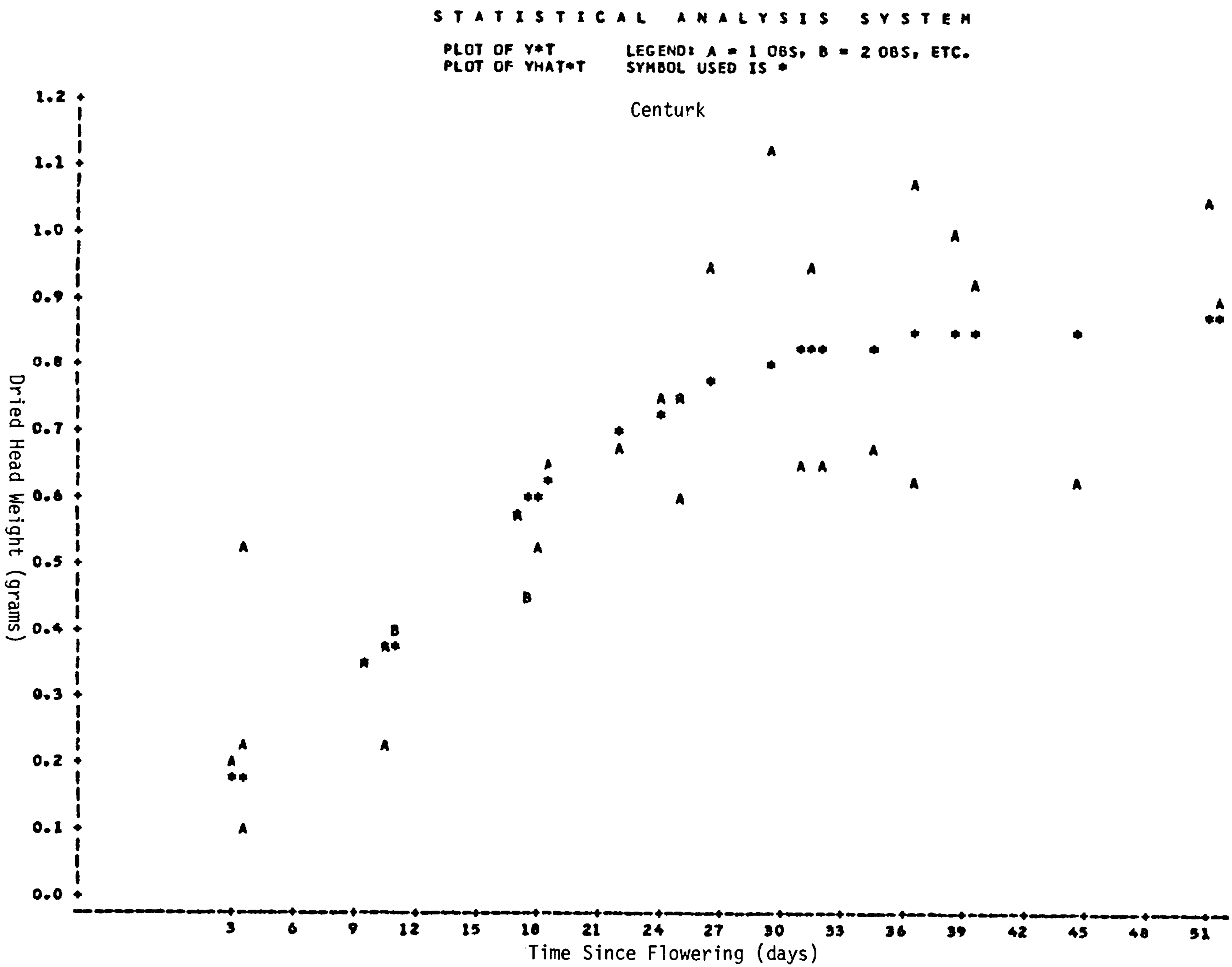

STATISTICAL ANALYSIS SYSTEM
PLOT OF Y*T LEGEND: A = 1 OBS, B = 2 OBS, ETC.
PLOT OF YHAT*T SYMBOL USED IS *
Centurk
Dried Head Weight (grams)
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51
Time Since Flowering (days)

Figure 11