# WITHIN-YEAR SPRING WHEAT GROWTH MODELS

Research and Development Branch
Research Division
Statistical Reporting Service
U. S. Department of Agriculture

January 1976

# WITHIN-YEAR SPRING WHEAT GROWTH MODELS

by

Jack Nealon

## Table of Contents

## ACKNOWLEDGEMENTS

Gary Lochow and Ron Dittus from the North Dakota State Statistical Office performed an excellent job collecting all data. Dwight Rockwell's assistance designing the data collection forms and data analysis was very much appreciated.

## INTRODUCTION

Research endeavors since 1973 by the Yield Assessment Section of the Research Division of the Statistical Reporting Service have emphasized development of crop yield models for which the parameters are derived solely from current year data for use in forecasting pertinent components of crop yield. These models are referred to as within-year models.

Between-year crop yield models presently utilized by the Statistical Reporting Service depend upon data from a base period (usually three years) to estimate the parameters in the models. These models assume the current year is part of the composite population of the base period years, which provide the parameter values. Within-year crop yield models would have the advantage of providing crop yield forecasts without the dependence on a base period to estimate the parameters. Within-year models would reflect unique characteristics of the year for which the forecast is desired.

Within-year models could be a valuable supplement to between-year models. Supplemental information from within-year models may be beneficial in improving crop yield forecasts for atypical years when growing conditions differ greatly from the base period years that generated the parameters in the between-year models.

In addition to providing supplemental information to the present crop yield forecasting system, within-year models could be very useful in developing forecasting models for crops not in the present crop yield forecasting system. Usually, three to five years of data must be collected before a reliable between-year model can be implemented. A within-year model could be developed in a shorter time period since data from a base period are not required.

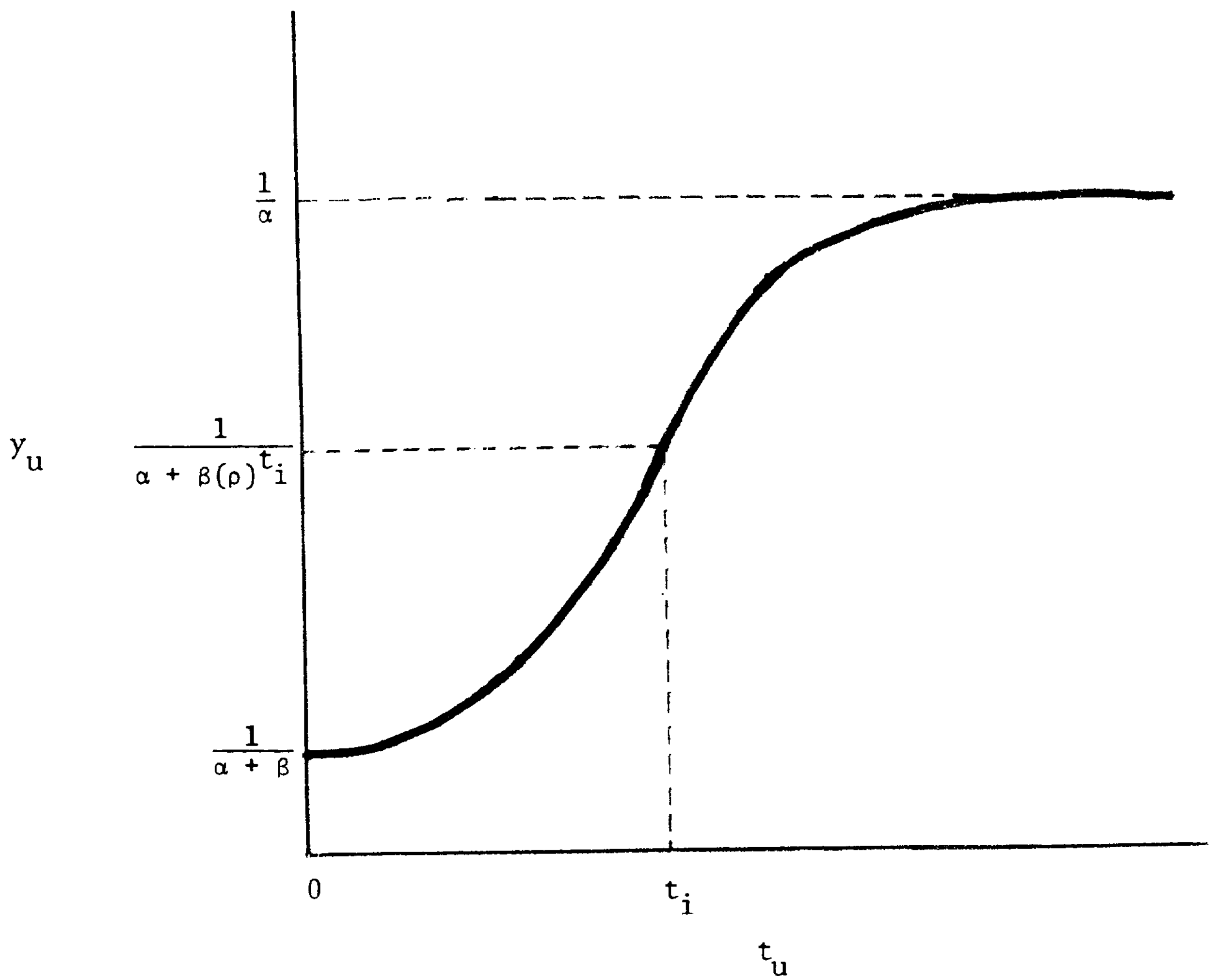
The Yield Assessment Section has investigated the applicability of within-year models to corn. Various within-year models have been examined to determine if biological laws relate corn growth, in terms of dry kernel weight, to time scales closely associated with initial kernel formation. Results demonstrate that a within-year model known as the logistic growth model accurately describes the process by which dry kernel weight accumulates in corn. Therefore, this model was adopted to investigate its potentiality as a forecasting methodology for spring wheat.

The logistic growth model is a non-linear model that uses incipient data for the independent variable and dependent variable to estimate values for the parameters in the model and provide a forecast of the dependent variable for a particular value of the independent variable. The form of the logistic growth model is:

$$y_u = \frac{1}{\alpha + \beta(\rho)^{t_u}} + e_u \quad ; \quad u = 1, 2, \ldots, n.$$

For this study, data collected during the growing season for the independent and dependent variables provided a forecast for the dependent variable at maturity. The independent variable, $t_u$ , is a time variable associated with a phenological event such as time since flowering in days. The dependent variable,

$y_u$ , is the mean dry kernel weight for stalks corresponding to a particular value of the time variable. The parameters, which are estimated, are $\alpha$, $\beta$ and $\rho$. These parameters must each be greater than zero. In addition, $\rho$ must be less than one. The disturbance term, $e_u$, is assumed to be normally distributed with zero mean and constant variance. The logistic growth model is shown graphically below.




When $t_u = 0$, the dependent variable, $y_u$, is $\frac{1}{\alpha + \beta}$ . The value of the dependent variable then increases at an increasing rate until $t_u = t_i$. For this value of time, the value of $y_u$ is $\frac{1}{\alpha + \beta(\rho)^{t_i}}$ . The dependent variable then increases at a decreasing rate until the asymptotic value, $\frac{1}{\alpha}$, is attained for $y_u$. That is, for large values of $t_u$, $\beta(\rho)^{t_u}$ approaches zero since $\rho$ is between zero and one. Therefore, the asymptotic value is the forecast of dry kernel weight per stalk at maturity. Since $\frac{1}{\alpha}$ provides the forecasting component, $\alpha$ is known as the primary parameter, and $\beta$ and $\rho$ are referred to as the secondary parameters.

## DATA COLLECTION

### SAMPLE

A nonprobabilistic sample of three spring wheat fields was selected in Cass County, North Dakota. Three varieties (Bounty, Era and Olaf) were represented in the fields. Within each field, three units 50 feet in length were randomly selected. Within each unit, the stalk at each foot mark was identified uniquely with a numerically labeled tag.

### PHENOLOGICAL EVENT OBSERVATIONS

The purpose of the first phase of data collection, phenological event observations, was to provide the date of occurrence of each phenological event for the 50 stalks in each unit. The occurrence date assists in determining the value of the independent variable, time since a phenological event. Observations commenced when the sample fields had passed the jointing stage.

Phenological events observed were head swelling, head emergence and flowering. Head swelling was identified by observing the protuberance in the sheath caused by a partially developed wheat head. Head emergence occurred when at least one spikelet on the wheat head was visible. When anthers protruded from the florets on the wheat head, the stalk was in the flowering stage.

Stalks were observed at two to three day intervals so that the occurrence date for each phenological event could be accurately identified. Table 1 shows the cumulative percentage of stalks in each phenological event category by visit. The rapid changes in stalk development indicate the necessity of frequent visits to observe the stalks. For example, the time period for observing flowering is only about nine days. Phenological event observations terminated when all surviving stalks had flowered. On visit 9 shown in Table 1, stalks not in the flowering category had perished. That is, 5.3% of the stalks did not survive.

Table 1

| Visit | Date | Head Swelling (%) | Head Emergence (%) | Flowering (%) |
|---|---|---|---|---|
| 1 | July 12 | .2 | | |
| 2 | July 14 | 16.0 | .7 | |
| 3 | July 16 | 46.4 | 17.3 | |
| 4 | July 18 | 78.0 | 56.7 | |
| 5 | July 21 | 94.2 | 86.4 | 28.4 |
| 6 | July 23 | 94.2 | 90.7 | 50.2 |
| 7 | July 25 | 95.6 | 94.0 | 72.4 |
| 8 | July 28 | 96.2 | 95.1 | 90.7 |
| 9 | July 30 | 96.2 | 96.2 | 94.7 |

## STALK CHARACTERISTIC OBSERVATIONS AND PREPARATION OF HEADS FOR LABORATORY DETERMINATIONS

The second phase of data collection commenced when flowering had occurred for all surviving stalks. Stalk characteristics were observed and heads prepared for laboratory determinations every seven days.

On each weekly visit a random sample of 10 stalks from the 50 stalks in each unit was observed for certain stalk characteristics. For each stalk, characteristics examined were:

(1) fertile spikelet count
(2) sterile spikelet count
(3) head length measurement

The counts and measurement for each stalk were performed without disturbing the stalk.

Five stalks were randomly selected from these 10 stalks and prepared for laboratory determinations. This preparation involved clipping the head from each stalk and placing the head in a separate bag with a card uniquely identifying the head. These bags were then transported to the laboratory.

Stalk characteristic observations on the 10 stalks were used as auxiliary information to refine the logistic growth model generated from the sample of five stalks so that the model would represent a greater proportion of stalks in the unit.

Time since a phenological event was defined for a stalk as the sampling date for laboratory determinations minus the occurrence date for the phenological event. For example, if a stalk was estimated to have flowered on July 23 and was sampled for laboratory determinations July 28, the time since flowering for the stalk would be five days.

## LABORATORY DETERMINATIONS

The primary purpose of the third and final phase of data collection was to obtain the dry kernel weight for each stalk sampled for laboratory determinations. The secondary purpose was to determine if each stalk characteristic observation from the field was accurately obtained.

Laboratory determinations commenced the day following the preparation of sampled heads in the field. In the laboratory, fertile and sterile spikelet counts and the head length measurement were obtained for sampled heads. These laboratory observations should be more accurate than field observations since wheat heads can be observed with less difficulty in the laboratory. Comparison of field and laboratory observations will provide an indication of the accuracy of counts and measurements in the field. The kernels were then extracted from each head and placed in a uniquely labeled dish. The dishes, which contained wheat kernels, were oven-dried at 140°F for 42 hours. Results from previous testing showed that an oven temperature of 140°F did not burn immature kernels and a substantial portion of the moisture was removed from the kernels in less than 42 hours of oven-drying. Upon completion of oven-drying, the dry kernels were weighed to the nearest milligram using an analytical scale.

## DATA ANALYSIS

### LOGISTIC GROWTH MODELS

#### Without Double Sampling Refinement

Logistic growth models were generated for each of the three phenological events with the unit as days for the independent variable, time since a phenological event. The dependent variable, $y_u$, was defined as the mean dry kernel weight in grams for stalks from the same sample field with the same value of the time variable. Therefore, for a sample field each distinct value of time since a phenological event constituted an observation. Since flowering occurred during a shorter time period than head emergence or head swelling, the number of observations was the least for time since flowering.

In Table 2, the number of observations, the estimated value and relative standard error for each parameter and the estimated value of $y_u$ at maturity are given for each phenological event. No phenological event appears to be superior with respect to the relative standard errors of the parameters. The fits of the data to the logistic growth model for each phenological event using the estimated parameter values are shown in Figures 1, 2 and 3. All figures are displayed in the APPENDIX.

Table 2

| Phenological Event | n | $\hat{\alpha}$ | $\hat{\beta}$ | $\hat{\rho}$ | $\sigma_{\hat{\alpha}}/\hat{\alpha}$ % | $\sigma_{\hat{\beta}}/\hat{\beta}$ % | $\sigma_{\hat{\rho}}/\hat{\rho}$ % | $\lim_{t_u \to \infty} y_u$ |
|---|---|---|---|---|---|---|---|---|
| Flowering | 40 | 1.7016 | 35.125 | .80761 | 7.14 | 57.68 | 4.52 | .588 |
| Head Emergence | 63 | 1.7253 | 184.66 | .78174 | 6.01 | 80.75 | 4.56 | .580 |
| Head Swelling | 66 | 1.6593 | 165.25 | .8051 | 6.88 | 79.36 | 4.09 | .603 |

When the time variable is defined in units of days, the assumption is that each day has an equivalent effect on the growth behavior of the dependent variable. Since weather conditions influence stalk development, the unit of the independent variable was defined as heat summation days to reflect daily temperature conditions.

The mean daily temperature was obtained from the weather station closest to the sample fields. Since the minimum temperature for wheat growth is 40°F, a heat day was defined as the mean daily temperature minus 40°F. For a particular day, the heat summation days would be the sum of previous heat days divided by the mean heat day for the growing season. For example, if the mean heat day for the growing season was 40°F and the mean daily temperatures for the first, second and third days were 90°F, 70°F and 60°F, respectively, then the heat summation days for the third day would be:

$$\frac{((90°F + 70°F + 60°F) - 3(40°F))}{40°F} = 2.5 .$$

With the independent variable defined in units of heat summation days rather than days, data for each phenological event were fitted to the logistic growth model. Pertinent information concerning the parameters is given in Table 3. Comparison of the relative standard errors of the parameters in Tables 2 and 3 does not provide conclusive evidence concerning the more desirable unit for the time variable. The associated data fits are shown in Figures 4, 5 and 6.

Table 3

| Phenological Event | n | $\hat{\alpha}$ | $\hat{\beta}$ | $\hat{\rho}$ | $\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$ % | $\hat{\sigma}_{\hat{\beta}}/\hat{\beta}$ % | $\hat{\sigma}_{\hat{\rho}}/\hat{\rho}$ % | $\lim_{t_u \to \infty} y_u$ |
|---|---|---|---|---|---|---|---|---|
| Flowering | 49 | 1.5376 | 38.003 | .83019 | 9.76 | 57.77 | 3.87 | .650 |
| Head Emergence | 67 | 1.6249 | 141.66 | .81626 | 7.51 | 64.70 | 3.32 | .615 |
| Head Swelling | 66 | 1.4913 | 146.61 | .82763 | 8.11 | 63.42 | 3.02 | .671 |

Alternative definitions for the dependent and independent variables were then tested. The dependent variable was defined as the mean dry kernel weight in grams for stalks from the same sample field for a sampling visit, and the independent variable was the mean time in days since a phenological event for stalks from the same sample field for a sampling visit. Therefore, on each weekly sampling visit for laboratory determinations one observation was generated for each sample field for the logistic growth model. Since two fields were visited on four occasions, and the third field had five weekly visits, the total observations were 13. Again, data were fitted to the logistic growth model for each phenological event (Figures 7, 8 and 9). Parameter estimates, their estimated relative standard errors and the forecast for the dependent variable at maturity are displayed in Table 4. Using the alternative definitions, the relative standard error of each parameter for each phenological event increased with respect to the relative standard errors in Table 2. However, this increase is to be expected since the number of observations was greatly reduced.

Table 4

| Phenological Event | n | $\hat{\alpha}$ | $\hat{\beta}$ | $\hat{\rho}$ | $\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$ % | $\hat{\sigma}_{\hat{\beta}}/\hat{\beta}$ % | $\hat{\sigma}_{\hat{\rho}}/\hat{\rho}$ % | $\lim_{t_u \to \infty} y_u$ |
|---|---|---|---|---|---|---|---|---|
| Flowering | 13 | 1.7732 | 38.646 | .79283 | 9.26 | 83.29 | 6.89 | .564 |
| Head Emergence | 13 | 1.8373 | 129.33 | .78473 | 7.69 | 102.09 | 6.20 | .544 |
| Head Swelling | 13 | 1.8377 | 219.37 | .78233 | 7.29 | 110.18 | 6.02 | .544 |

No phenological event appeared to be superior with respect to producing the smallest relative standard errors for the parameters. Therefore, since the time variable should correspond to an event closely related to initial kernel formation, flowering was chosen for further analysis.

## With Double Sampling Refinement

Figure 10 contains for each value of time since flowering in days the square of the correlation coefficient ($R^2$) between the dry kernel weight and the following variables:

(1) head weight in the laboratory
(2) fertile spikelet count in the field
(3) head length measurement in the field

Each bracketed number in Figure 10 corresponds to the number of observations determining the $R^2$ for that value of time since flowering.

Head weight demonstrated the best relationship with dry kernel weight. The $R^2$ were very similar for the fertile spikelet count and head length measurement. The large variations in the $R^2$ for different values of time since flowering are unreasonable. These variations may be due to:

(1) the limited amount of data
(2) different relationships between the dry kernel weight and the variables for the different varieties of wheat

Figures 11, 12 and 13 demonstrate that the head weight, fertile spikelet count and head length measurement vary among the wheat varieties. For example, the head weights were heaviest for Era, and the fertile spikelet counts and head length measurements were smallest for Bounty and Olaf. However, due to the limited amount of data, the $R^2$ variations could not feasibly be analyzed.

Since head weight was not observed in the field for the larger sample of 10 stalks, this variable could not be used as a double sampling refinement for the dry kernel weight from the smaller sample of five stalks sent to the laboratory.

The fertile spikelet count and head length measurement for the larger sample in the field were each used as an auxiliary variable to refine the dry kernel weight obtained in the laboratory for the smaller sample by establishing a double sampling regression estimate of the dry kernel weight for the larger sample.

The form of the double sampling regression estimate for the dry kernel weight from the larger sample is:

$$\overline{Y}_L = \overline{Y}_S + \hat{\beta}_1(\overline{X}_L - \overline{X}_S).$$

$\overline{X}_L$ represents the mean value of the auxiliary variable in the larger sample for each time interval, and $\overline{Y}_S$ and $\overline{X}_S$ are the mean values for the dry kernel weight and auxiliary variable from the smaller sample for the same time interval, respectively. The regression coefficient, $\hat{\beta}_1$, is obtained by regression data from the smaller sample for the auxiliary variable and dry kernel weight for the same time interval. If the regression coefficient was significantly different from zero for a time interval, the dry kernel weight from the smaller sample was adjusted to reflect the dry kernel weight in the larger sample.

Logistic growth models were generated with the refinement on the dry kernel weight for each auxiliary variable. The dependent variable was the refined mean dry kernel weight in grams for stalks from the same sample field with the same value of the independent variable, time since flowering in days. Table 5 exhibits the refined estimate for each parameter and remaining relevant parameter information for each auxiliary variable. The logistic growth models produced for each auxiliary variable are shown in Figures 14 and 15. Comparison of the relative standard errors in Tables 2 and 5 for the phenological event, flowering, indicates that the use of the auxiliary variables to refine the dry kernel weight did not improve the model. The large $R^2$ variations (Figure 10) seem to have affected the refinements.

Table 5

| Auxiliary Variable | n | $\hat{\alpha}$ | $\hat{\beta}$ | $\hat{\rho}$ | $\sigma_{\hat{\alpha}}/\hat{\alpha}$ % | $\sigma_{\hat{\beta}}/\hat{\beta}$ % | $\sigma_{\hat{\rho}}/\hat{\rho}$ % | $\lim_{t_u \to \infty} y_u$ |
|---|---|---|---|---|---|---|---|---|
| Fertile Spikelet Count | 40 | 1.0988 | 31.07 | .81607 | 7.73 | 56.17 | 4.42 | .589 |
| Head Length | 40 | 1.7317 | 34.46 | .80722 | 7.17 | 58.12 | 4.61 | .577 |

## Heteroscedastic-Error Adjustment

A relevant result from the data analysis of the 1974 Corn Growth Research was that an assumption for the logistic growth model, namely, residuals possess constant variance, was violated. That is, as the time variable increases, the difference between the observed dry kernel weight and the functional value for the dry kernel weight produced by the logistic growth model increases. This violation is known as heteroscedasticity, and if present, less efficient estimates for the parameters are generated.

Plots of the residuals versus time since a phenological event were examined for this study, and demonstrated that the residuals (observed value minus functional value) increased as time since a phenological event increased. Figure 16 illustrates that as time since flowering in heat summation days increases, the residuals increase.

The assumption of constant variance for the residuals can be attained by restructuring the logistic growth model to reflect the dependence of residuals on time. The form of the logistic growth model with the heteroscedastic-error adjustment is:

$$\frac{y_u}{f(t_u)} = \frac{1}{f(t_u)} \frac{1}{\alpha + \beta(\rho)^{t_u}} + \frac{e_u}{f(t_u)} \quad ; \quad u = 1, 2, \ldots, n.$$

The function, $f(t_u)$, reflects the relationship between the absolute value of the residuals and time since a phenological event. The revised disturbance term, $e_u/f(t_u)$, has a constant variance.

The heteroscedastic-error refinement was tested for the independent variables, time since flowering in units of days and heat summation days, and the dependent variable, mean dry kernel weight in grams for stalks from the same sample field with the same value of the independent variable. An exponential function described the relationship between time since flowering in days or heat summation days and the absolute value of the residuals better than a linear or quadratic function. However, the correlation coefficients were only .408 and .431 when the unit for the independent variable was days and heat summation days, respectively. The form of the logistic growth model with the heteroscedastic-error adjustment was:

$$\frac{y_u}{\exp[\hat{\beta}_0 + \hat{\beta}_1 t_u]} = \frac{1}{\exp[\hat{\beta}_0 + \hat{\beta}_1 t_u]} \frac{1}{\alpha + \beta(\rho)^{t_u}} + \frac{e_u}{\exp[\hat{\beta}_0 + \hat{\beta}_1 t_u]} \; ; \; u = 1, 2, \ldots, n.$$

Refined estimates of the parameters, associated relative standard errors and the forecast of the dependent variable at maturity are given in Table 6. Comparison of Tables 2 and 3 to Table 6 demonstrates that the relative standard error for $\beta$ and $\rho$ decreased substantially using the heteroscedastic-error adjustment. However, the relative standard error for $\alpha$ increased slightly with the adjustment. Figures 17 and 18 demonstrate the refined fits of the data to the logistic growth model for time since flowering in days and heat summation days, respectively.

Table 6

| Unit | n | $\hat{\alpha}$ | $\hat{\beta}$ | $\hat{\rho}$ | $\hat{\sigma}_{\hat{\alpha}}/\hat{\alpha}$ % | $\hat{\sigma}_{\hat{\beta}}/\hat{\beta}$ % | $\hat{\sigma}_{\hat{\rho}}/\hat{\rho}$ % | $\lim_{t_u \to \infty} y_u$ |
|---|---|---|---|---|---|---|---|---|
| Days | 40 | 1.7711 | 44.093 | .78948 | 7.40 | 36.62 | 3.45 | .565 |
| Heat Summation Days | 49 | 1.5753 | 41.726 | .8236 | 10.55 | 27.64 | 2.44 | .635 |

## Maturity Category Approach

Since phenological event observations required frequent field visits to accurately define the occurrence date for each phenological event for each stalk, these observations were a costly method for determining the time variable. Therefore, an attempt to define the independent variable during the weekly sampling visits was pursued.

Each stalk observed for stalk characteristics was classified into a maturity category, which described a distinct stage of stalk development. The maturity categories, which were identical to maturity categories for the Wheat Objective Yield Survey, were:

(1) Flag or Early Boot
(2) Late Boot or Flower includes watery kernels
(3) Milk
(4) Soft Dough
(5) Hard Dough
(6) Ripe

Defining the time variable by maturity categories will ensure reduced data collection costs. However, the accuracy of the time variable may be affected for two reasons: (1) Stalks in the same maturity category were assigned the same time value regardless of their distinct stages of development within the maturity category. (2) Stalks classified into maturity categories were observed non-destructively, and classifications were therefore very subjective. An advantage of classification into maturity category is that the time variable would be determined by the stalk's stage of development at the time of sampling, not by the stage of development the stalk should be at because a phenological event was observed for the stalk a certain number of days previously.

For each maturity category, mean time since flowering in days was computed using data collected during phenological event observations. If the relationship between each maturity category and time since a phenological event is constant from year to year, classification into maturity categories would provide an appealing time variable since survey costs would be drastically reduced.

With the independent variable, mean time in days since flowering for a maturity category, and the dependent variable, mean dry kernel weight in grams for stalks from the same sample field with the same maturity category, a logistic growth model was generated and is shown in Figure 19. Pertinent parameter data are displayed in Table 7. The relative standard error of each parameter has not been significantly affected by the adjustment of the time variable.

Table 7

| n | $\hat{\alpha}$ | $\hat{\beta}$ | $\hat{\rho}$ | $\sigma_{\hat{\alpha}}/\hat{\alpha}$ % | $\sigma_{\hat{\beta}}/\hat{\beta}$ % | $\sigma_{\hat{\rho}}/\hat{\rho}$ % | $\lim_{t_u \to \infty} y_u$ |
|---|---|---|---|---|---|---|---|
| 14 | 1.8612 | 44.225 | .78512 | 8.33 | 79.54 | 6.70 | .537 |

LOGISTIC SURVIVAL MODEL

The forecast generated by the logistic growth model for mean dry kernel weight for stalks at maturity provides the component, yield per stalk. Another component, stalks per acre at harvest, is required to forecast yield per acre. Therefore, a method was designed to forecast stalks per acre at maturity.

During the first weekly sampling visit for laboratory determinations, stalks with potential kernel formation per acre were estimated. On each subsequent weekly field visit, the ratio of stalks surviving to total stalks was obtained from the stalks observed for stalk characteristics. This provided a survival ratio utilized to forecast the survival ratio at maturity of the stalks with potential kernel formation per acre on the first sampling visit.

The logistic survival model is derived from the logistic growth model. For the logistic growth model, at $t_u = 0$, the dependent variable is $\frac{1}{\alpha + \beta}$.

For the survival model, on the first sampling visit, time equals zero and the survival ratio is one. Therefore, since $\frac{1}{\alpha + \beta} = 1$, $\beta = 1 - \alpha$. From the logistic growth model the form of the logistic survival model is:

$$y_u = \frac{1}{\alpha + \beta(\rho)^{t_u}} = \frac{1}{\alpha + (1 - \alpha)(\rho)^{t_u}} = \frac{1}{\alpha(1 - \rho^{t_u}) + \rho^{t_u}}$$

with $\alpha \geq 1$, $0 < \rho < 1$, and the forecast of the dependent variable, survival ratio, is $\frac{1}{\alpha}$.

Observations for the dependent variable, survival ratio for stalks from the same sample field with the same value of the independent variable, time in days since the first sampling visit, were fitted to the logistic survival model (Figure 20). The forecasted survival ratio and parameter information is given in Table 8.

Table 8

| $n$ | $\hat{\alpha}$ | $\hat{\rho}$ | $\sigma_{\hat{\alpha}}/\hat{\alpha}$ % | $\sigma_{\hat{\rho}}/\hat{\rho}$ % | $\lim_{t_u \to \infty} y_u$ |
|---|---|---|---|---|---|
| 10 | 1.0307 | .000010133 | 1.11 | 0.0 | .970 |

COMPARISON OF STALK CHARACTERISTIC OBSERVATIONS IN THE FIELD AND LABORATORY

As stated previously, the purpose of the comparison between field and laboratory data for stalk characteristic observations was to determine the accuracy of field observations. A parametric univariate paired t-test and a non-parametric Wilcoxon paired signed rank test were performed for each stalk characteristic observation to determine if a significant difference existed between the field and laboratory data.

For the univariate paired t-test, the difference between the stalk characteristic observation in the field and laboratory was computed for each characteristic. The statistic, $t_{(n-1)} = \frac{(\sqrt{n})(\bar{d})}{S_d}$, was used to test the hypotheses:

$H_0$: $\delta = 0$

$H_A$: $\delta \neq 0$ where n is the number of paired observations; $\bar{d}$ and $S_d$ are the mean difference and standard deviation of the differences, respectively. The term, $\delta$, is the mean difference in the infinite population. The calculated value of t is squared and compared to the critical value $t^2_{(n-1)}$ at the 95% probability level. If the calculated $t^2$ is greater than the critical value, the null hypothesis is rejected. Table 9 demonstrates that at the 95% probability level no significant difference existed between the fertile spikelet count in the field and laboratory. Conversely, the sterile spikelet count and head length measurement each differed significantly in the field and laboratory. Therefore, according to this study, the accuracy of the fertile spikelet count is not signficantly hindered by counting in the field rather than in the laboratory.

Table 9

| n | Stalk Characteristic | $t^2$ calculated | Critical Value 95% level |
|---|---|---|---|
| 161 | Fertile Spikelet Count | .549 | 3.902 |
| 161 | Sterile Spikelet Count | 4.344 | 3.902 |
| 161 | Head Length Measurement | 16.555 | 3.902 |

The non-parametric Wilcoxon paired signed rank test produced the same conclusions concerning the accuracy of the stalk characteristic observations. If the absolute value of the calculated Z statistic is greater than the critical value for Z, the null hypothesis that the mean difference in the infinite population is zero is rejected. Results are shown in Table 10.

Table 10

| n | Stalk Characteristic | $\lvert Z \rvert$ calculated | Critical Value 95% level |
|---|---|---|---|
| 61 | Fertile Spikelet Count | .535 | 1.96 |
| 53 | Sterile Spikelet Count | 2.386 | 1.96 |
| 115 | Head Length Measurement | 5.745 | 1.96 |

## CONCLUSION

Results based on a nonprobabilistic sample of three spring wheat fields in Cass County, North Dakota provide evidence that within-year growth and survival models may be useful for forecasting relevant components of spring wheat yield. However, statistical inferences concerning the applicability of within-year growth and survival models to spring wheat cannot be stated due to the nonprobabilistic selection of fields. A probability sample of field is recommended for 1976.

The initial values for dry kernel weight should correspond to a time variable very closely associated with the beginning of kernel formation. Since kernel formation commences at flowering, the independent variable, time since flowering, is the most desirable time variable. Head swelling and head emergence, as defined for this study, are not closely related to initial kernel formation. To strengthen this relationship for head emergence, this phenological event should be redefined as head fully emerged. Because of discrepancy in reliably identifying head swelling, this phenological event should not be observed in future studies.

No selection of the dependent and independent variables appeared to perform best. However, to maintain independence in the observations on a sampling visit for the dependent variable, one observation per sample field should be generated for the model. This will be the approach in future studies.

The use of field observations to refine the dependent variable has the appealing attribute of inexpensively representing a larger percentage of the population. Therefore, if a stalk characteristic is highly correlated with dry kernel weight, a larger sample of stalks should be observed in the field and fewer stalks should be sent to the laboratory. Surely, this would reduce survey expenditures.

Because head weight is highly correlated with dry kernel weight, this characteristic should be observed in future studies to refine the dry kernel weight. Fertile spikelet counts were accurately obtained in the field. Therefore, this stalk characteristic should be observed in future research efforts. Conversely, head length measurements were significantly different in the field and laboratory. Unless these measurements are performed in the laboratory, they should be discontinued.

The double sampling refinements did not improve the performance of the model for the time variable, time since flowering in days. A larger sample will be necessary to determine if the relationship between each stalk characteristic and dry kernel weight is good enough to improve the model. Data from numerous varieties should be compared to determine if these relationships are similar among varieties. In addition, each variety will be fitted to the logistic growth model to decide whether a distinct model is necessary for each variety or group of similar varieties.

Time since each phenological event is available for the stalks observed during the sampling visits in the field. Therefore, this information should be used to refine the time variable from the smaller sample. Also, if the time variable and dry kernel weight are correlated for the smaller sample, the dry kernel weight can be refined utilizing the time data from the larger sample.

Because the model assumption of constant variance for the residuals was violated, the heteroscedastic-error adjustment was necessary to correct this violation. Additional research is needed to determine the function or functions that best describe the dependence of residuals on time.

The validity of the maturity category approach to define the independent variable cannot be commented upon until more research is conducted.

# REFERENCES

[1] Rockwell, Dwight A., Nonlinear Estimation, Research and Development Branch, Research Division, Statistical Reporting Service, U. S. Department of Agriculture, Washington, D. C., 1975.

[2] Wilson, Wendell W., Preliminary Report on the Use of Time Related Growth Models in Forecasting Components of Corn Yield, Research and Development Branch, Research Division, Statistical Reporting Service, U. S. Department of Agriculture, Washington, D. C., 1974.

APPENDIX

Figures 1 - 20

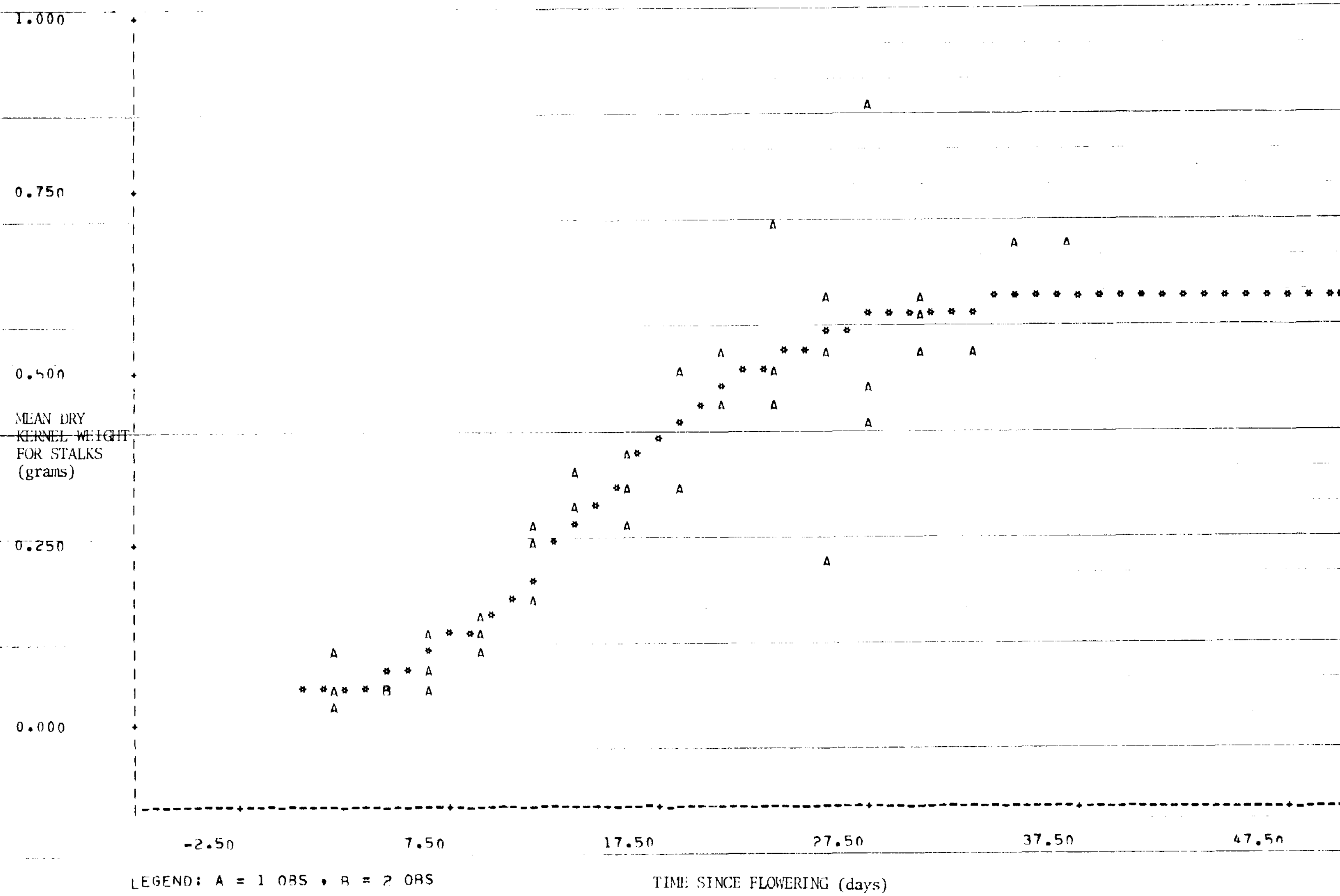

1.000
0.750
0.500
0.250
0.000
MEAN DRY KERNEL WEIGHT FOR STALKS (grams)
-2.50
7.50
17.50
27.50
37.50
47.50
TIME SINCE FLOWERING (days)
LEGEND: A = 1 OBS , B = 2 OBS

Figure 1

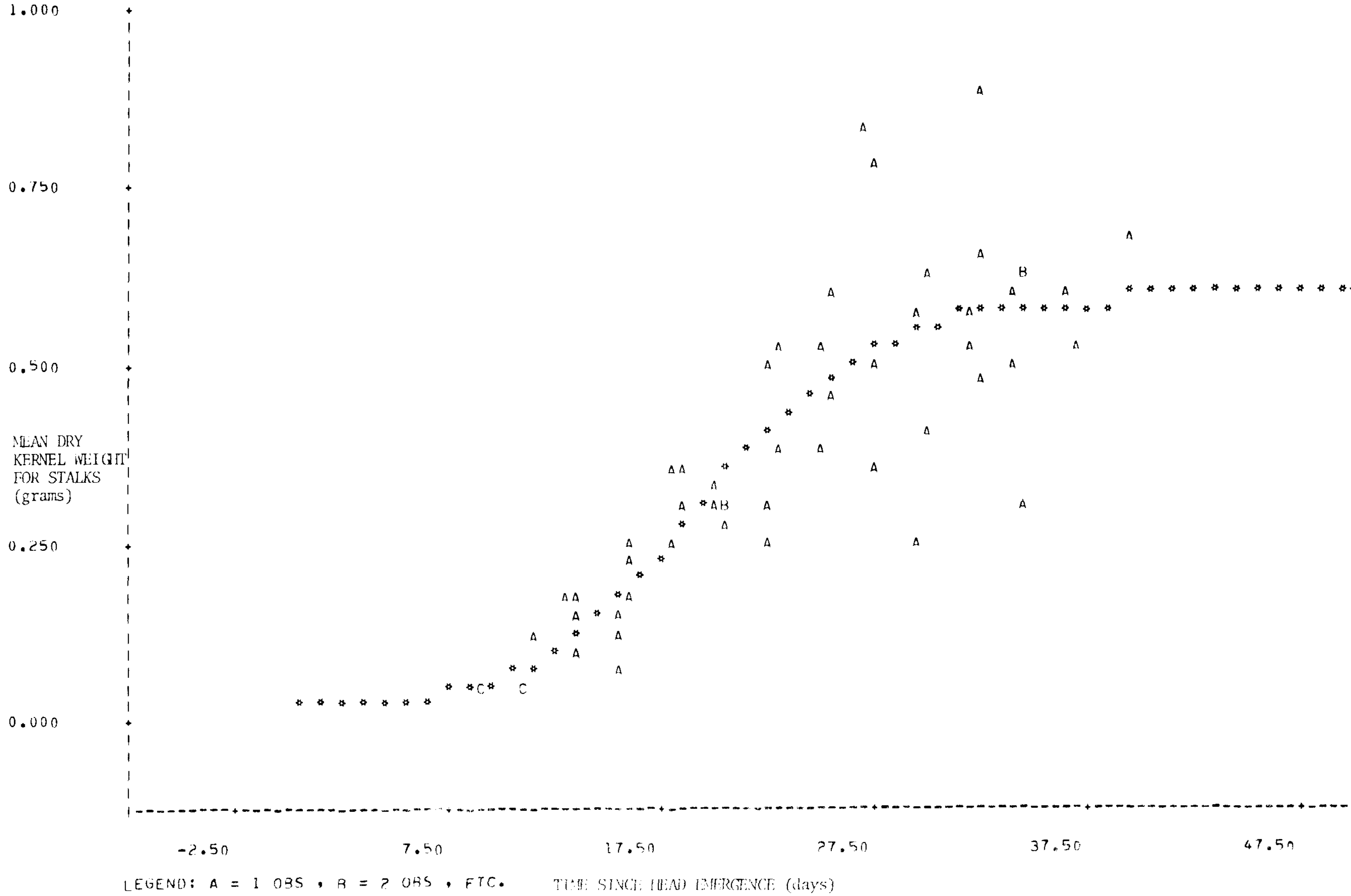

1.000
0.750
0.500
0.250
0.000
MEAN DRY KERNEL WEIGHT FOR STALKS (grams)
-2.50
7.50
17.50
27.50
37.50
47.50
LEGEND: A = 1 OBS , B = 2 OBS , ETC.
TIME SINCE HEAD EMERGENCE (days)

Figure 2

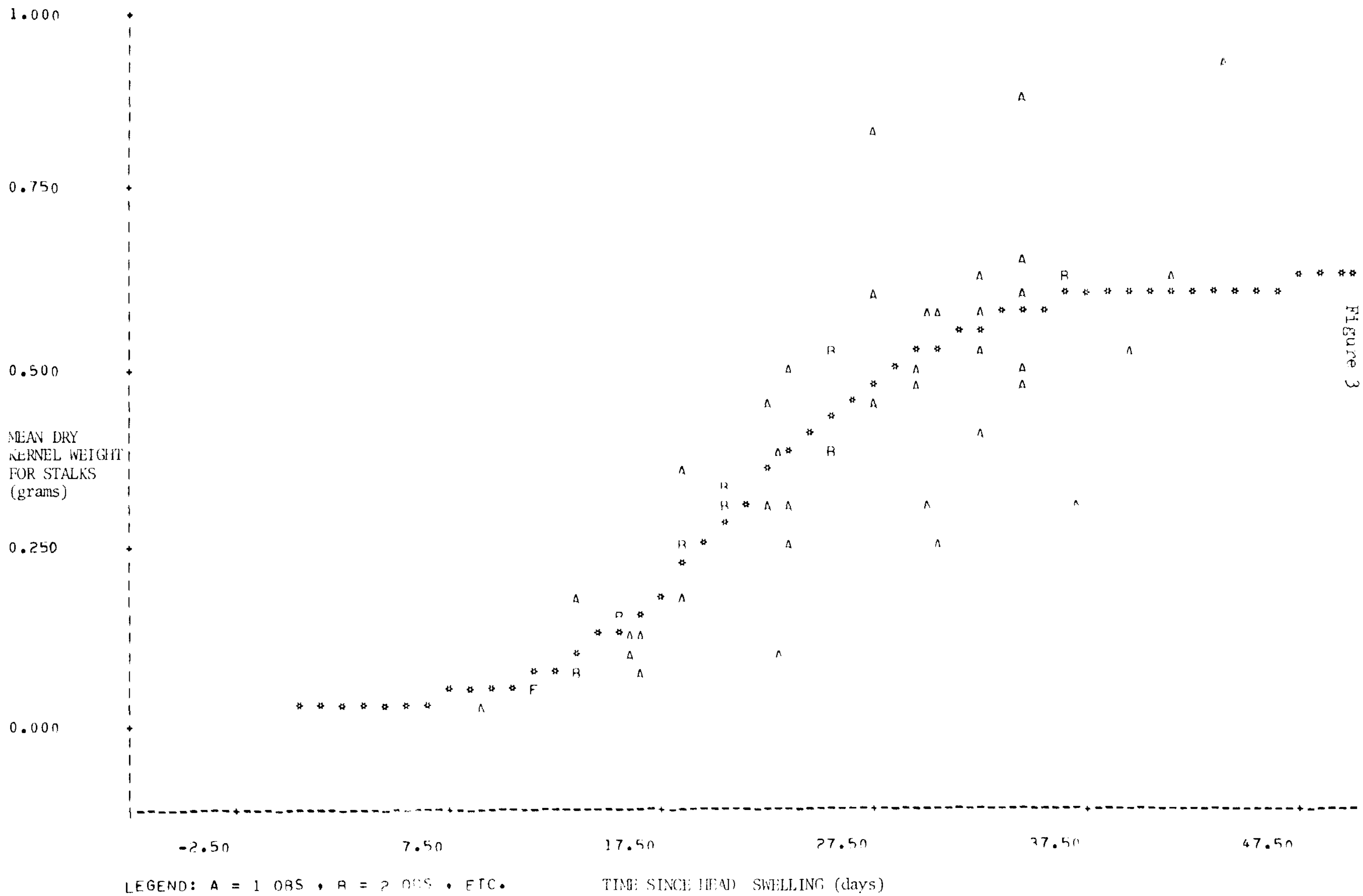

1.000
0.750
0.500
0.250
0.000
MEAN DRY
KERNEL WEIGHT
FOR STALKS
(grams)
-2.50
7.50
17.50
27.50
37.50
47.50
LEGEND: A = 1 OBS , B = 2 OBS , ETC.
TIME SINCE HEAD SWELLING (days)

Figure 3

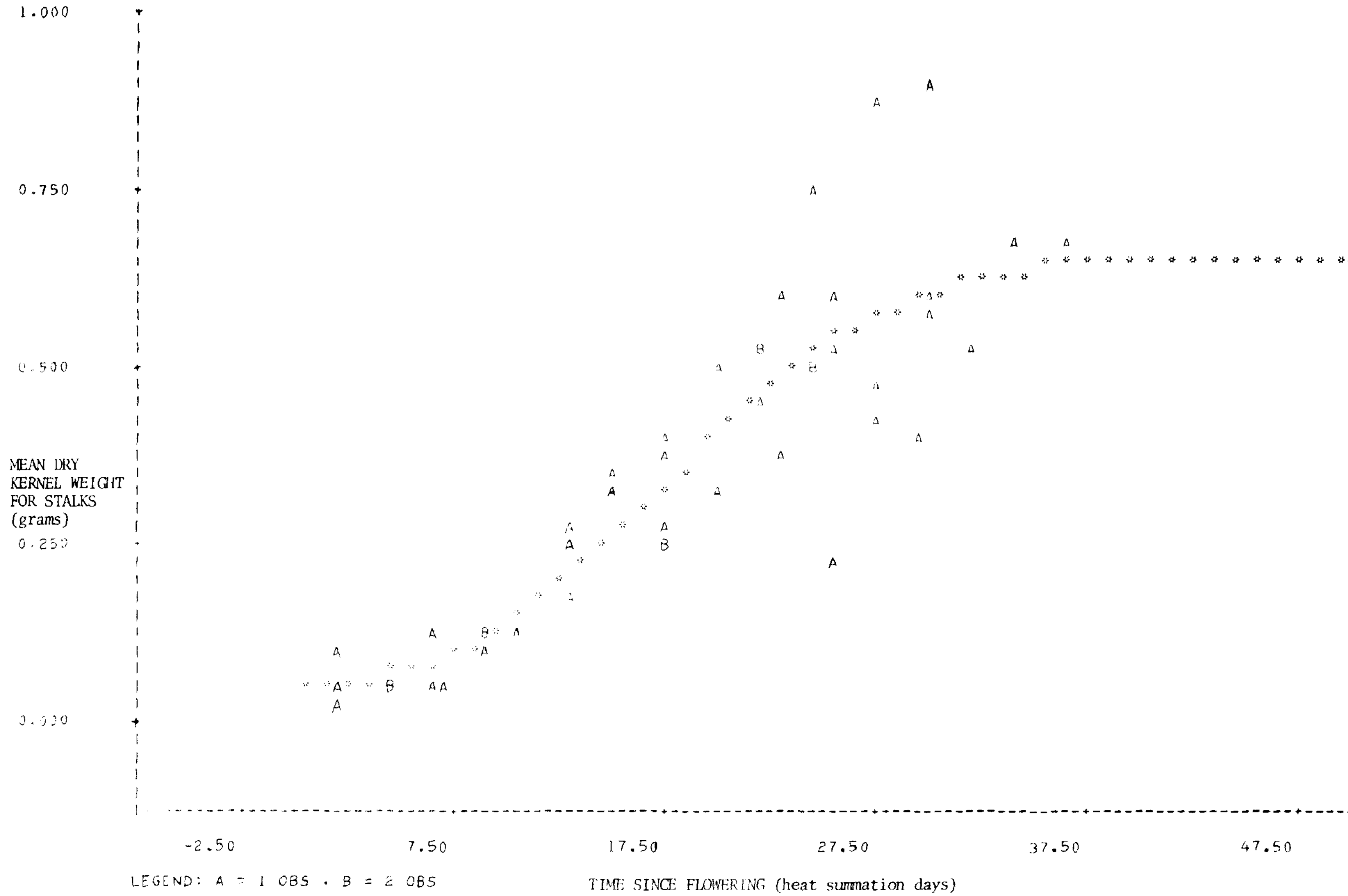

1.000
0.750
0.500
0.250
0.000
MEAN DRY
KERNEL WEIGHT
FOR STALKS
(grams)
-2.50
7.50
17.50
27.50
37.50
47.50
LEGEND: A = 1 OBS , B = 2 OBS
TIME SINCE FLOWERING (heat summation days)

Figure 4

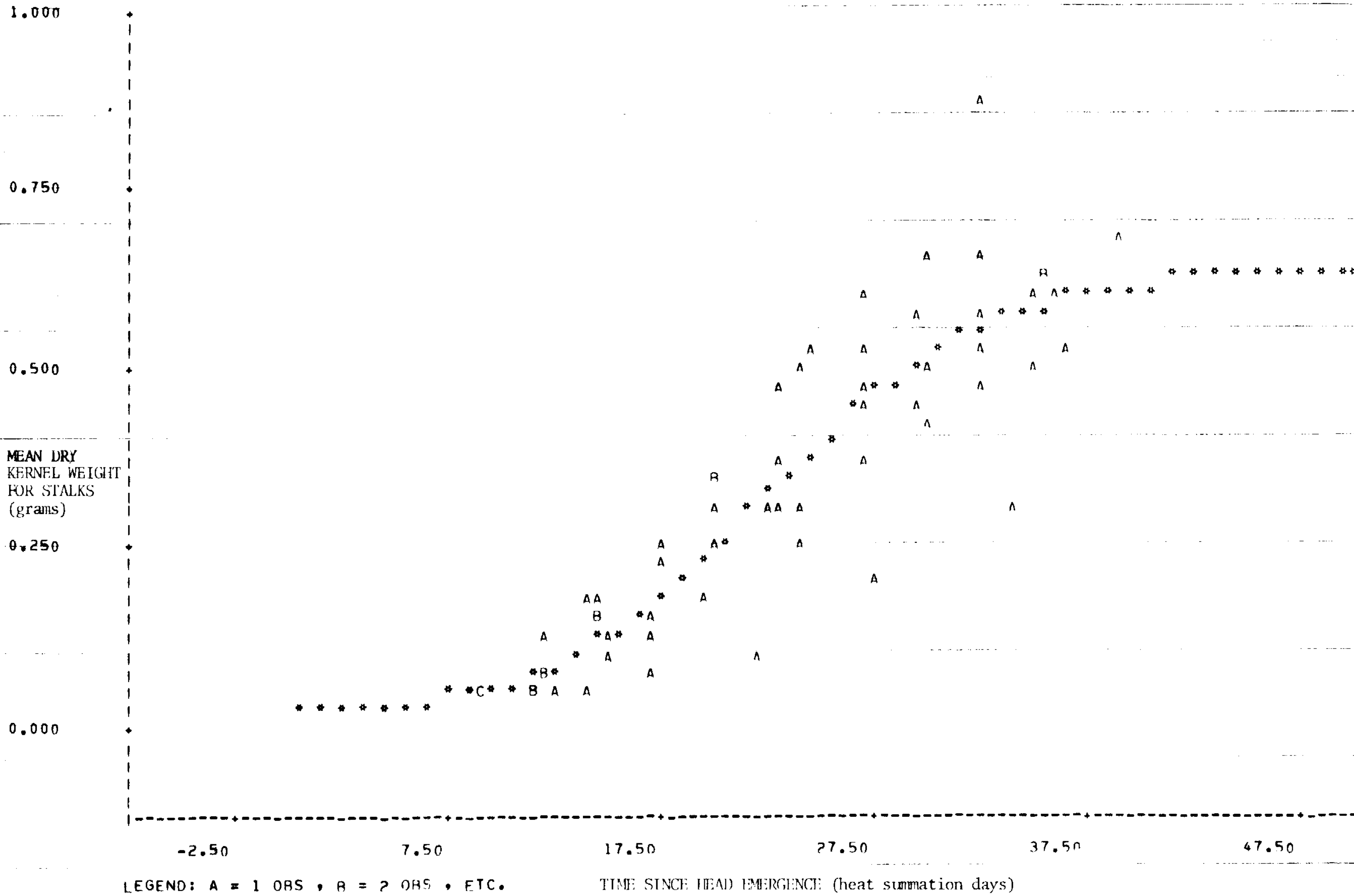

1.000
0.750
0.500
MEAN DRY
KERNEL WEIGHT
FOR STALKS
(grams)
0.250
0.000
-2.50
7.50
17.50
27.50
37.50
47.50
LEGEND: A = 1 OBS , B = 2 OBS , ETC.
TIME SINCE HEAD EMERGENCE (heat summation days)

Figure 5

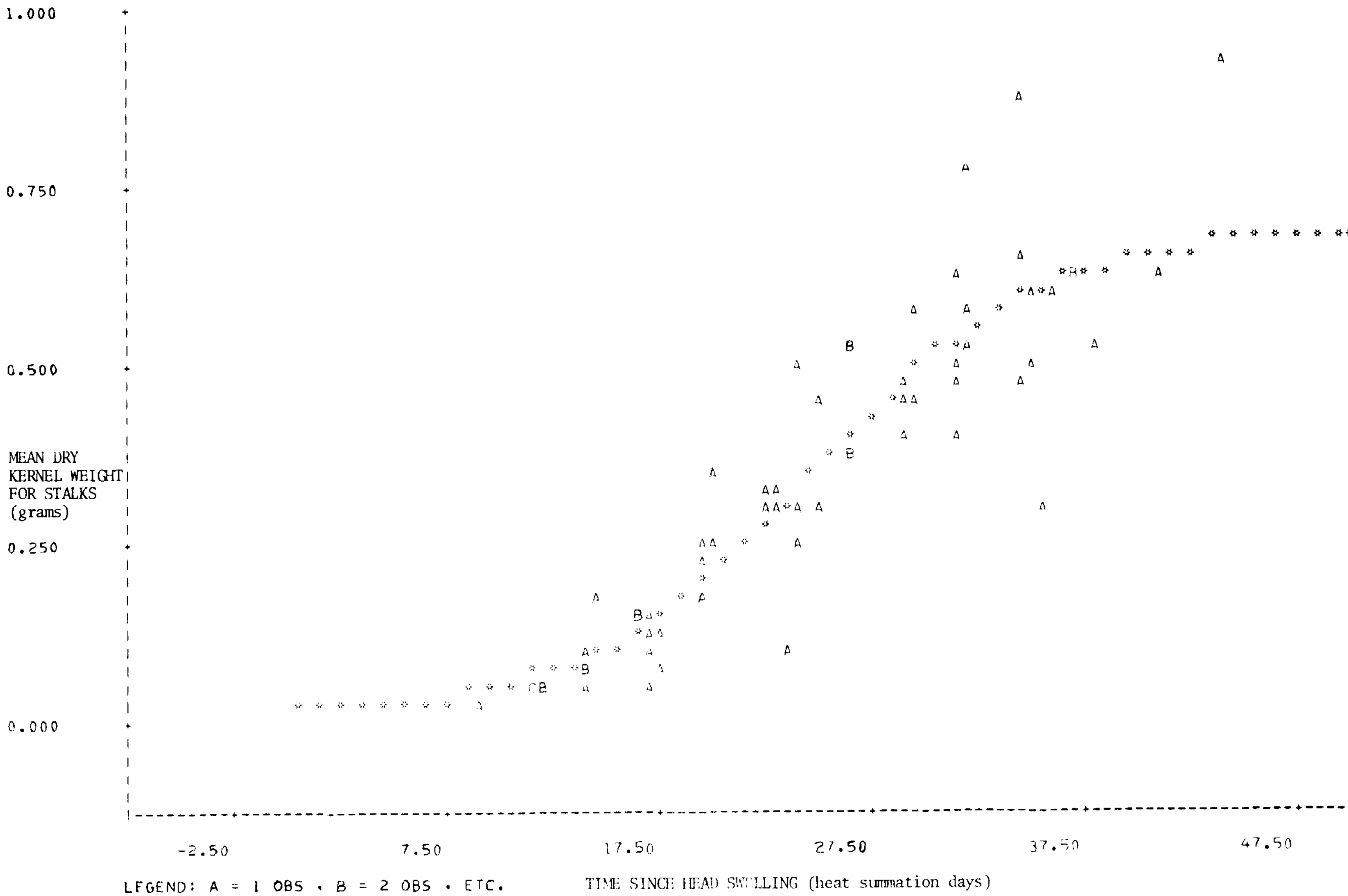

1.000
0.750
0.500
0.250
0.000
MEAN DRY KERNEL WEIGHT FOR STALKS (grams)
-2.50
7.50
17.50
27.50
37.50
47.50
TIME SINCE HEAD SWELLING (heat summation days)
LEGEND: A = 1 OBS , B = 2 OBS , ETC.

Figure 6

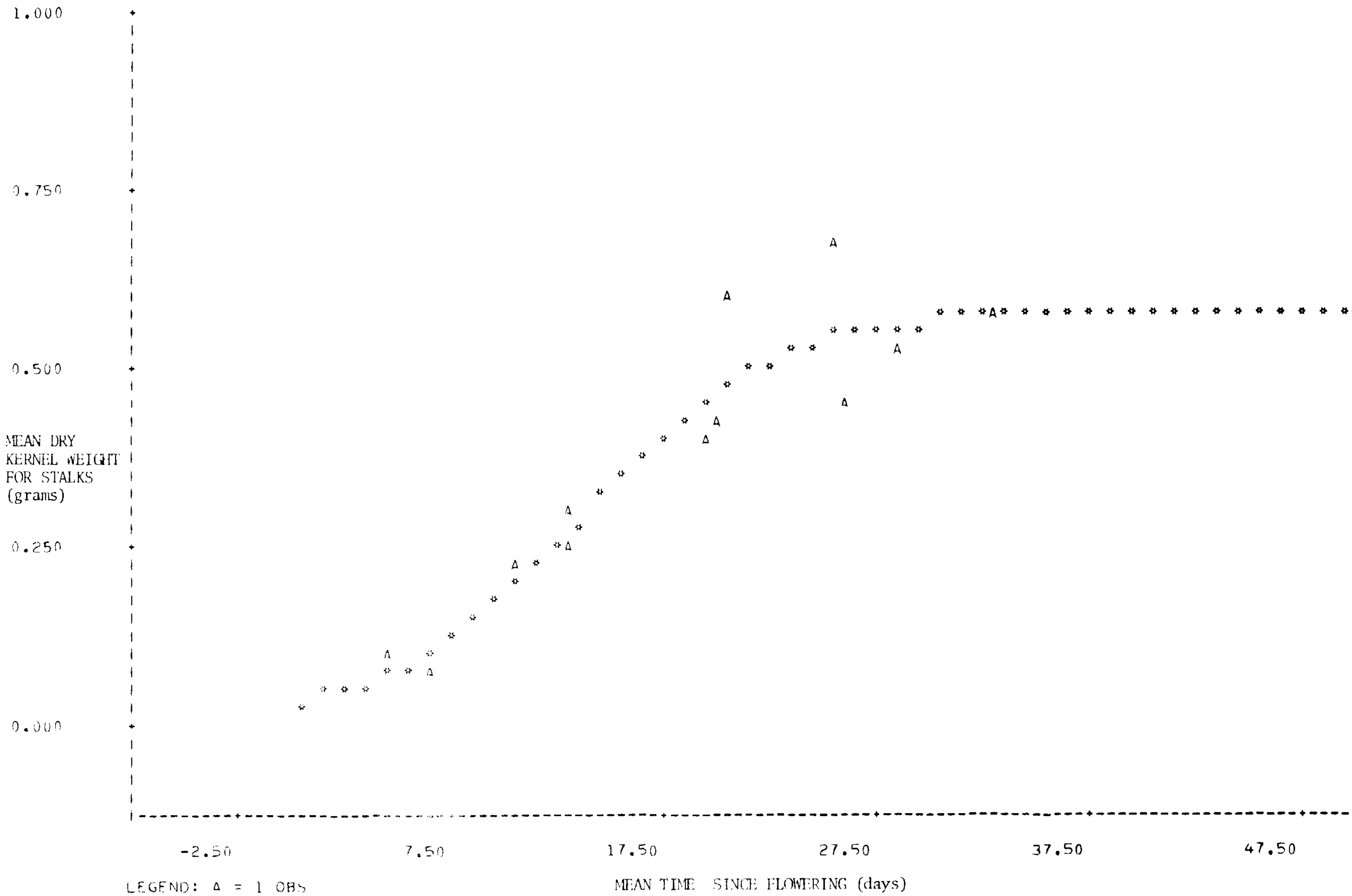

1.000
0.750
0.500
0.250
0.000
MEAN DRY
KERNEL WEIGHT
FOR STALKS
(grams)
-2.50
7.50
17.50
27.50
37.50
47.50
LEGEND: A = 1 OBS
MEAN TIME SINCE FLOWERING (days)

Figure 7

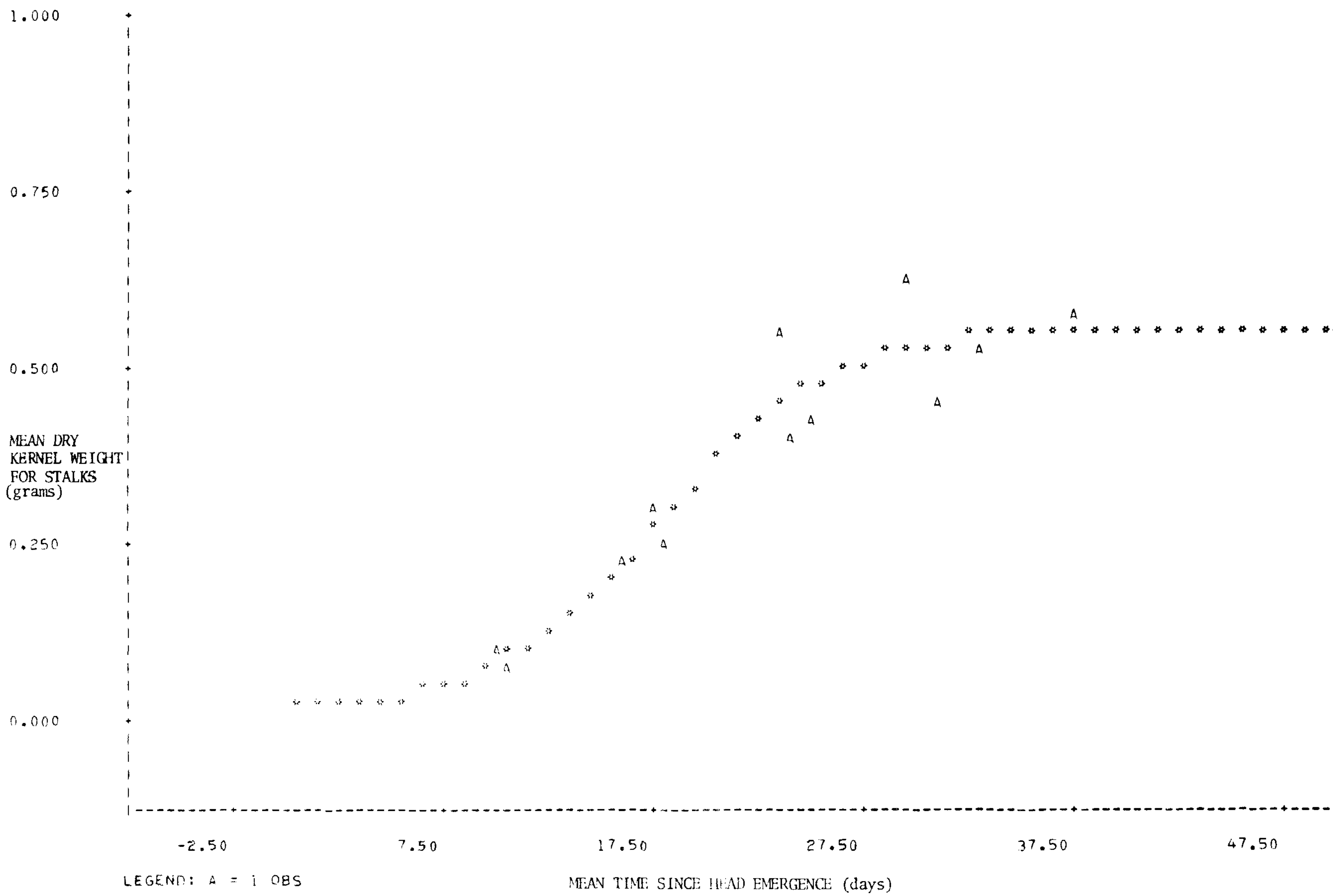

1.000
0.750
0.500
0.250
0.000
MEAN DRY
KERNEL WEIGHT
FOR STALKS
(grams)
-2.50
7.50
17.50
27.50
37.50
47.50
LEGEND: A = 1 OBS
MEAN TIME SINCE HEAD EMERGENCE (days)

Figure 8

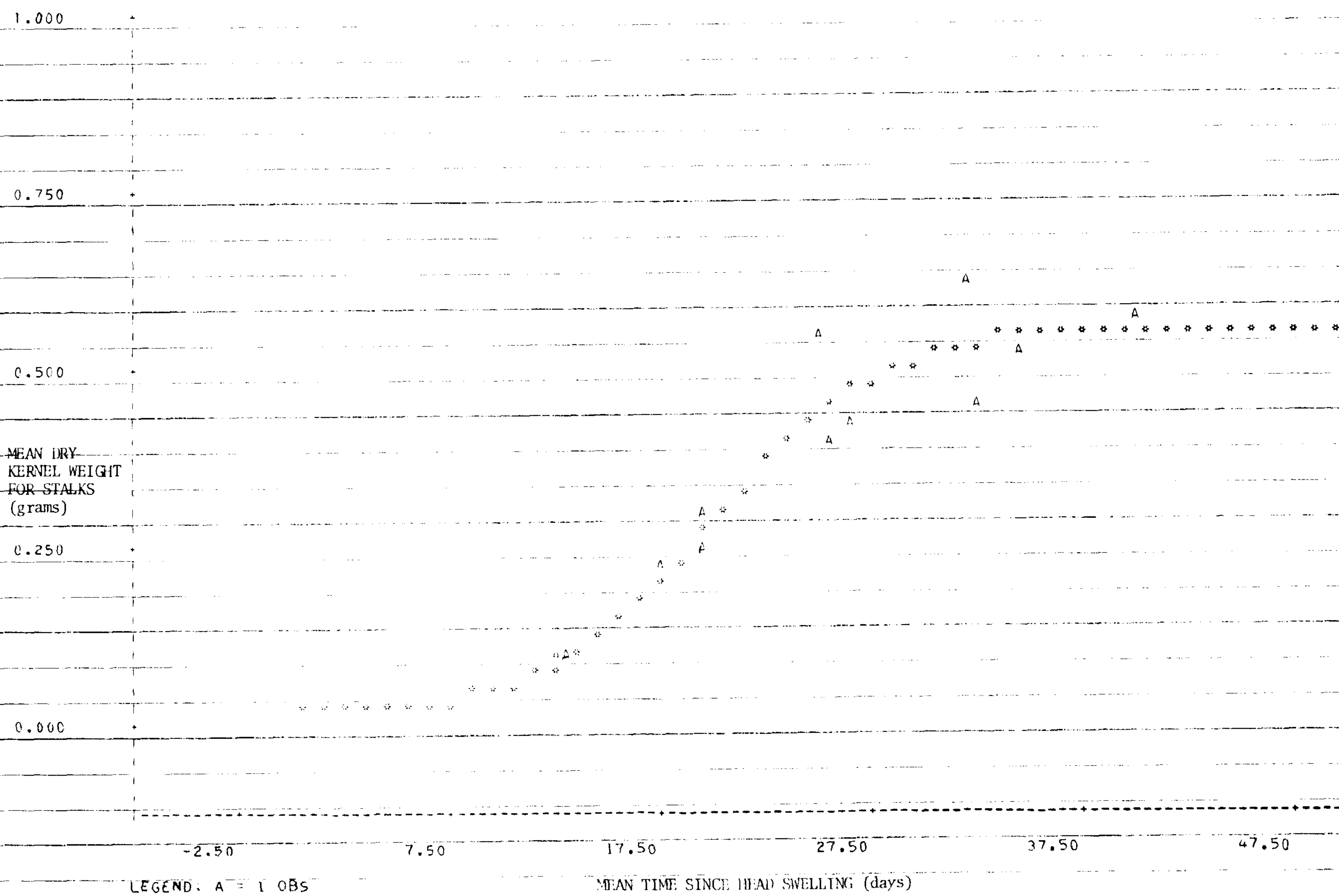

1.000
0.750
0.500
0.250
0.000
MEAN DRY KERNEL WEIGHT FOR STALKS (grams)
-2.50
7.50
17.50
27.50
37.50
47.50
MEAN TIME SINCE HEAD SWELLING (days)
LEGEND: A = 1 OBS

Figure 9

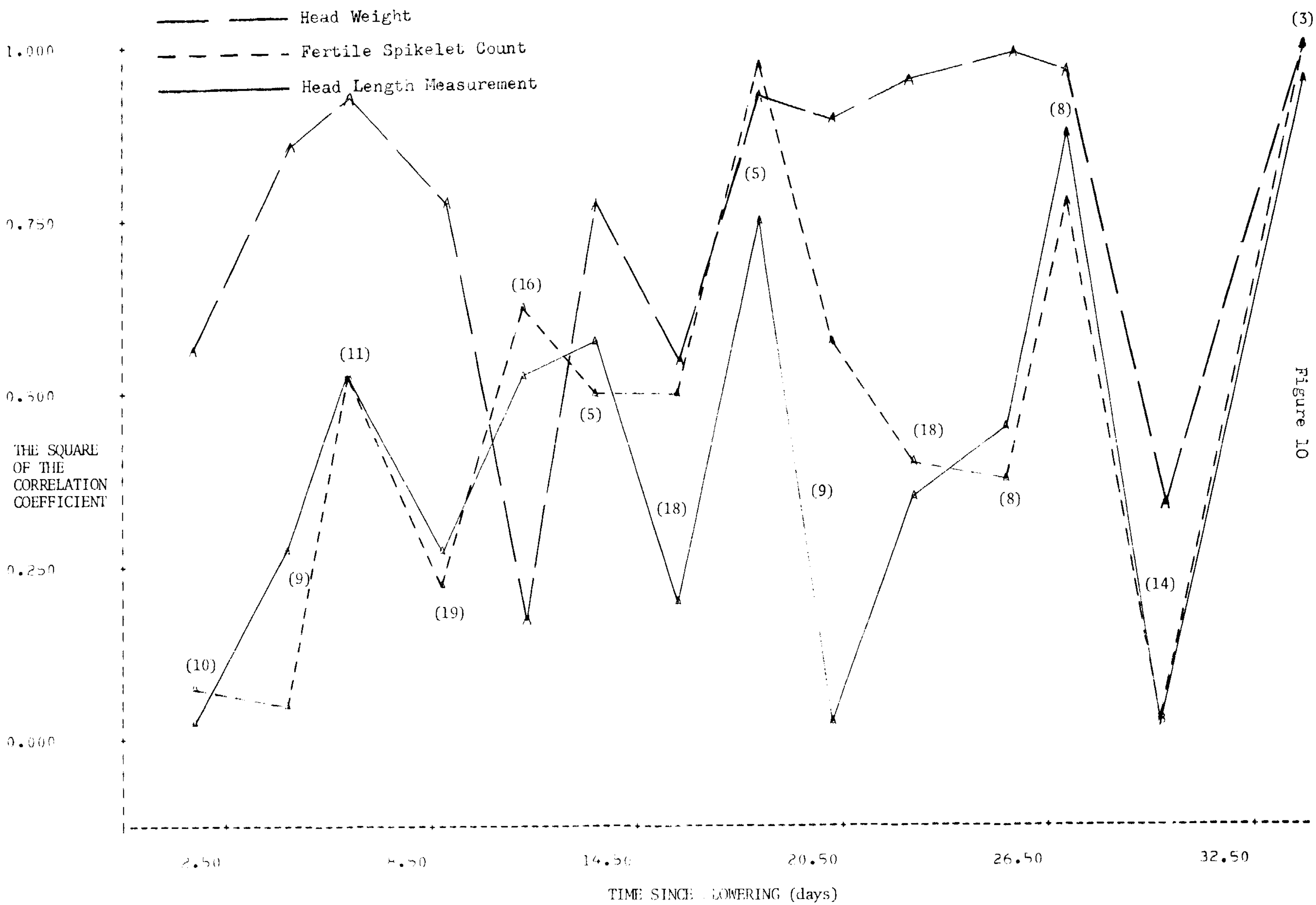

Head Weight
Fertile Spikelet Count
Head Length Measurement
1.000
0.750
0.500
0.250
0.000
THE SQUARE OF THE CORRELATION COEFFICIENT
(3)
(8)
(5)
(16)
(11)
(5)
(18)
(9)
(18)
(8)
(9)
(19)
(10)
(14)
32.50
26.50
20.50
TIME SINCE FLOWERING (days)

Figure 10

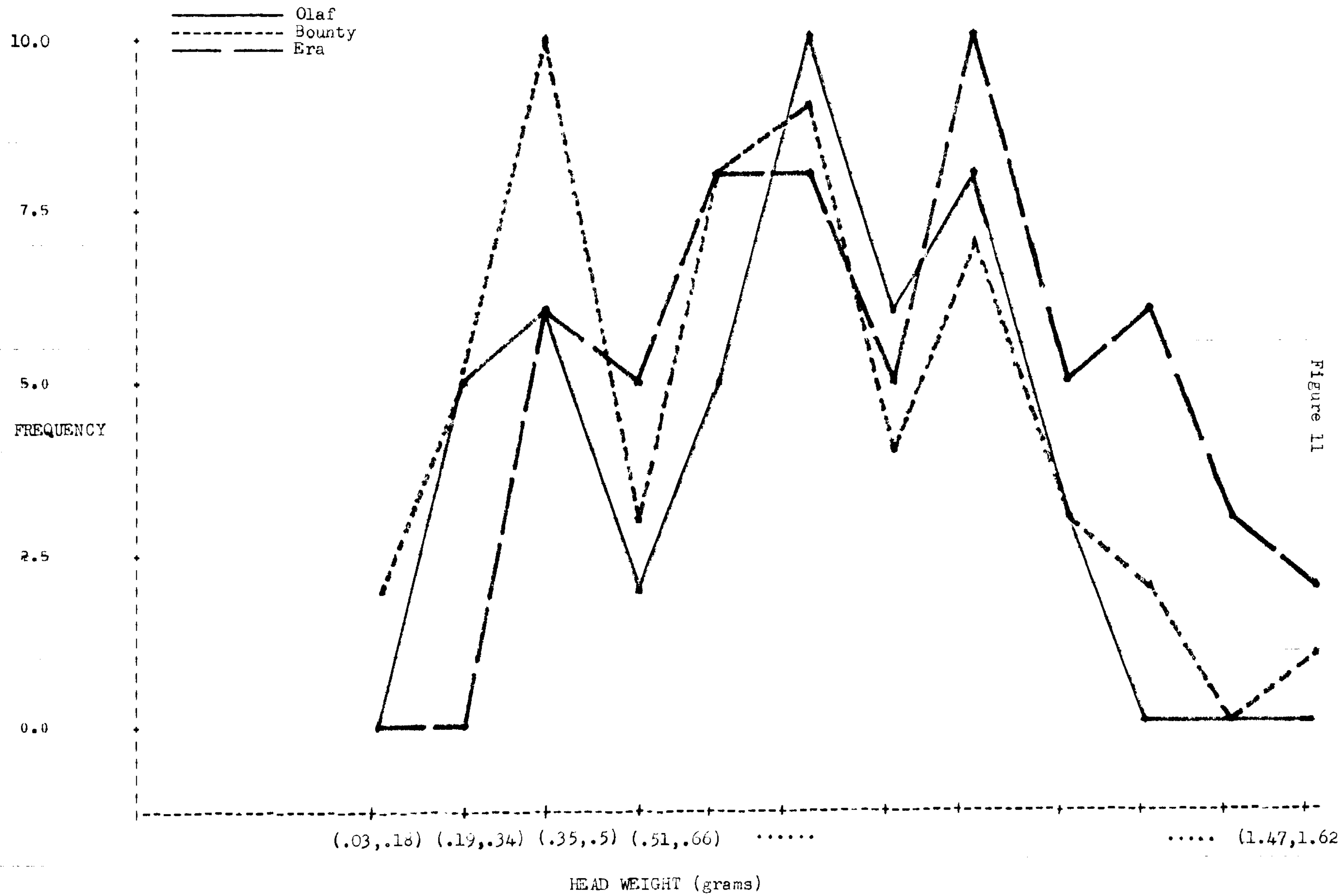

Olaf
Bounty
Era
10.0
7.5
5.0
2.5
0.0
FREQUENCY
(.03,.18) (.19,.34) (.35,.5) (.51,.66) ......
..... (1.47,1.62)
HEAD WEIGHT (grams)

Figure 11

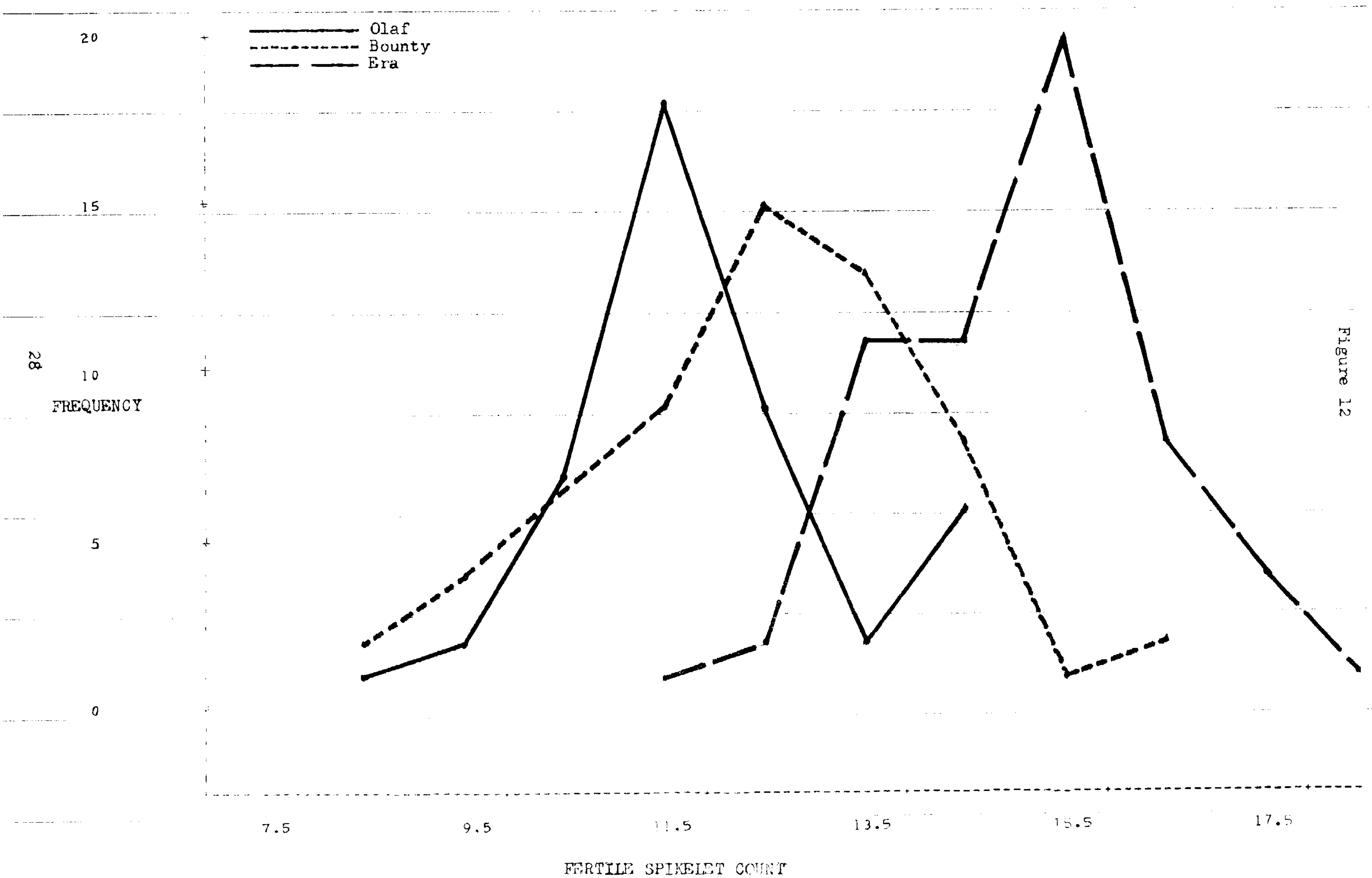

Olaf
Bounty
Era
20
15
10
5
0
FREQUENCY
7.5
9.5
11.5
13.5
15.5
17.5
FERTILE SPIKELET COUNT

Figure 12

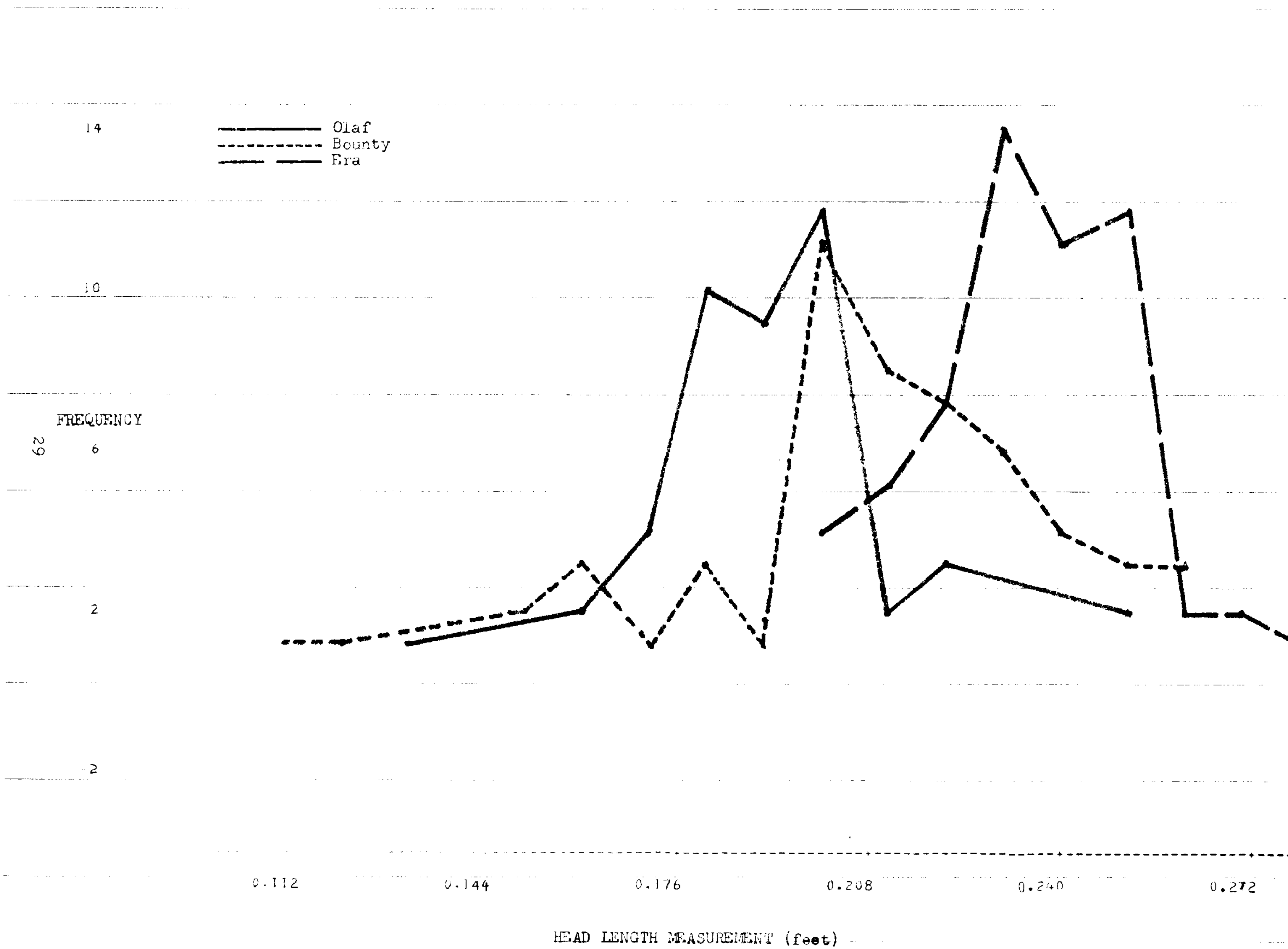

Olaf
Bounty
Era
14
10
FREQUENCY
6
2
2
0.112
0.144
0.176
0.208
0.240
0.272
HEAD LENGTH MEASUREMENT (feet)

Figure 13

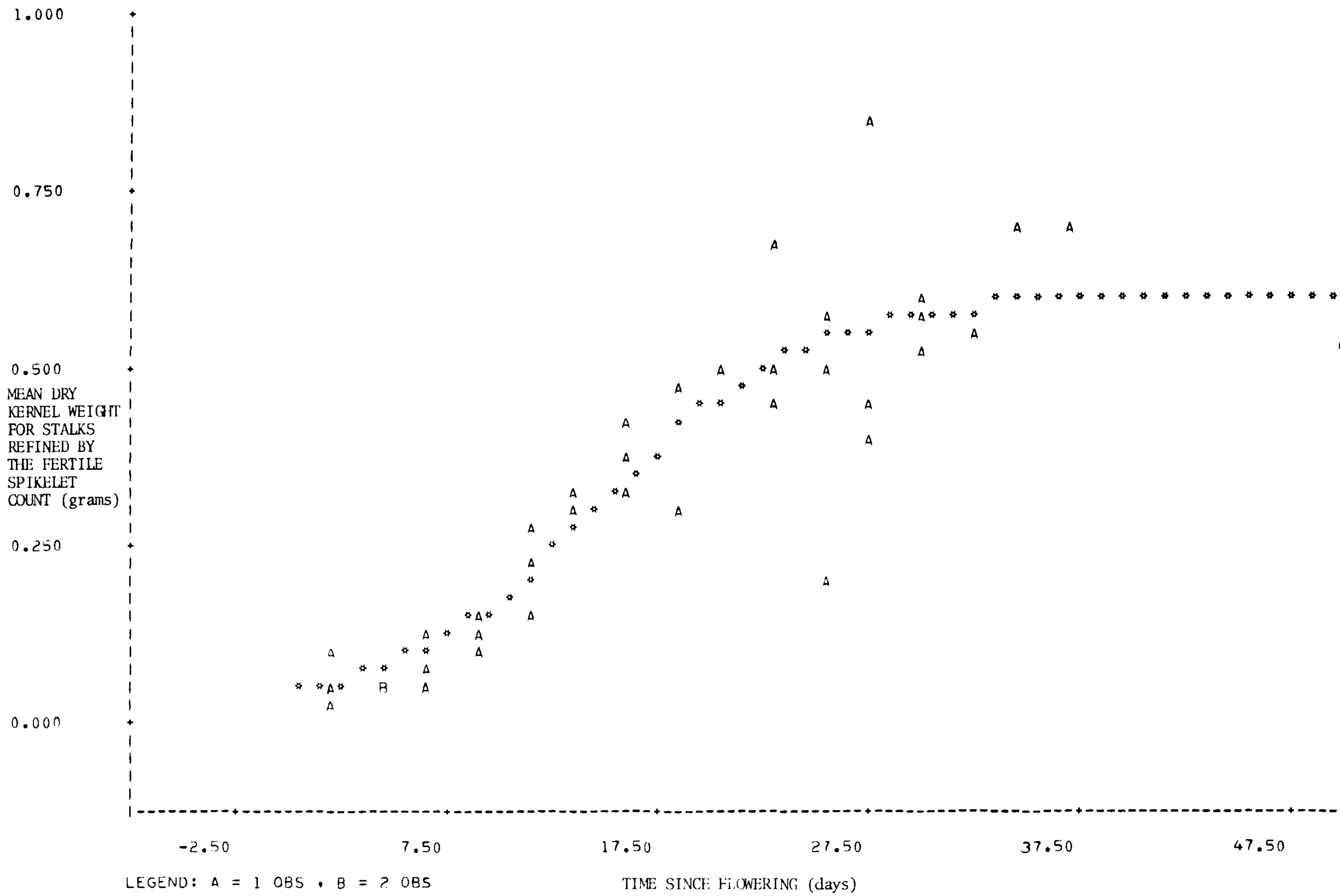
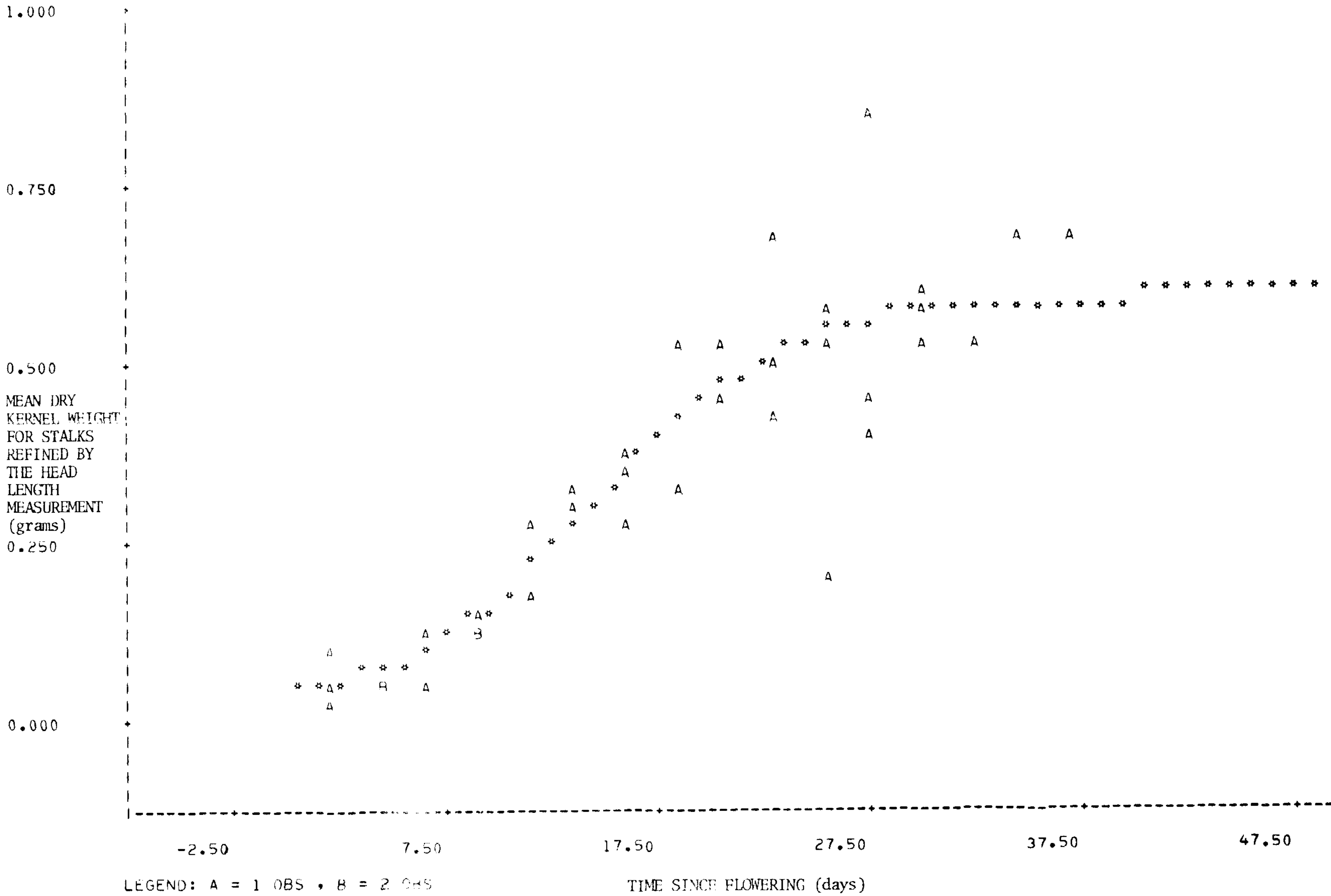

1.000
0.750
0.500
0.250
0.000
MEAN DRY KERNEL WEIGHT FOR STALKS REFINED BY THE FERTILE SPIKELET COUNT (grams)
-2.50
7.50
17.50
27.50
37.50
47.50
LEGEND: A = 1 OBS , B = 2 OBS
TIME SINCE FLOWERING (days)

Figure 14



1.000
0.750
0.500
0.250
0.000
MEAN DRY KERNEL WEIGHT FOR STALKS REFINED BY THE HEAD LENGTH MEASUREMENT (grams)
-2.50
7.50
17.50
27.50
37.50
47.50
LEGEND: A = 1 OBS , B = 2 OBS
TIME SINCE FLOWERING (days)


Figure 15

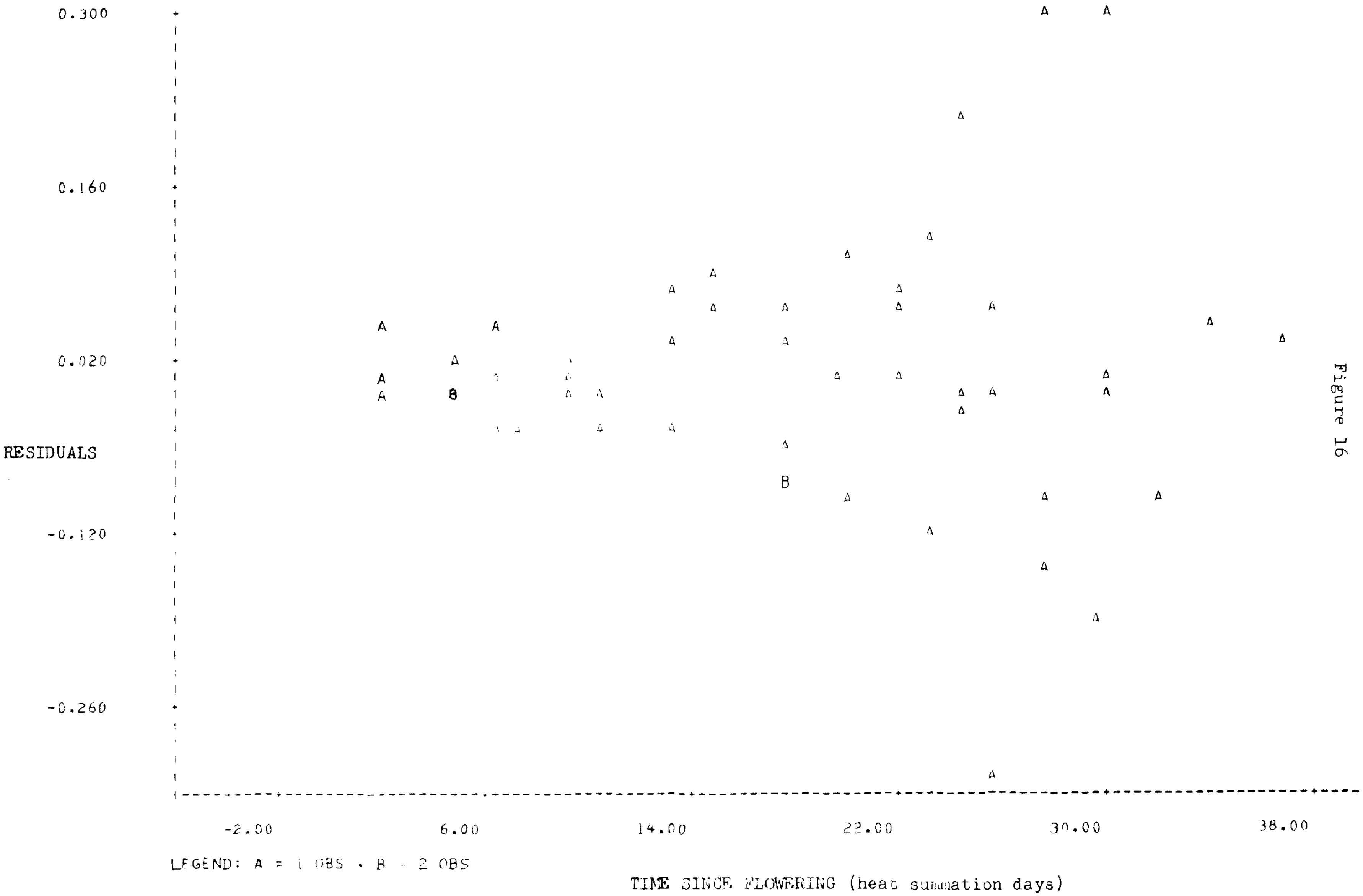

0.300
0.160
0.020
-0.120
-0.260
RESIDUALS
A
B
-2.00
6.00
14.00
22.00
30.00
38.00
LEGEND: A = 1 OBS , B = 2 OBS
TIME SINCE FLOWERING (heat summation days)

Figure 16

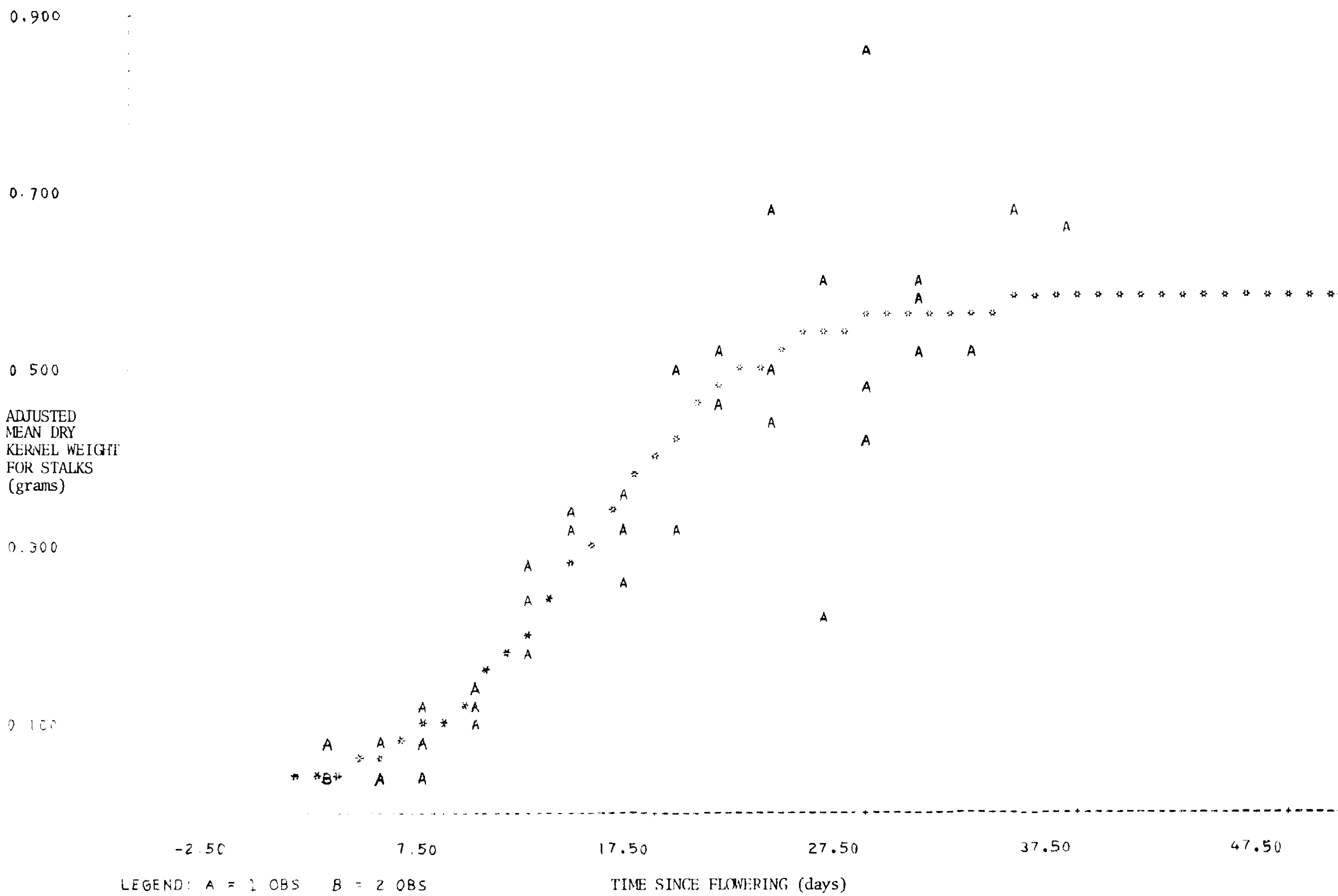

0.900
0.700
0.500
ADJUSTED MEAN DRY KERNEL WEIGHT FOR STALKS (grams)
0.300
0.100
-2.50
7.50
17.50
27.50
37.50
47.50
LEGEND: A = 1 OBS B = 2 OBS
TIME SINCE FLOWERING (days)

Figure 17

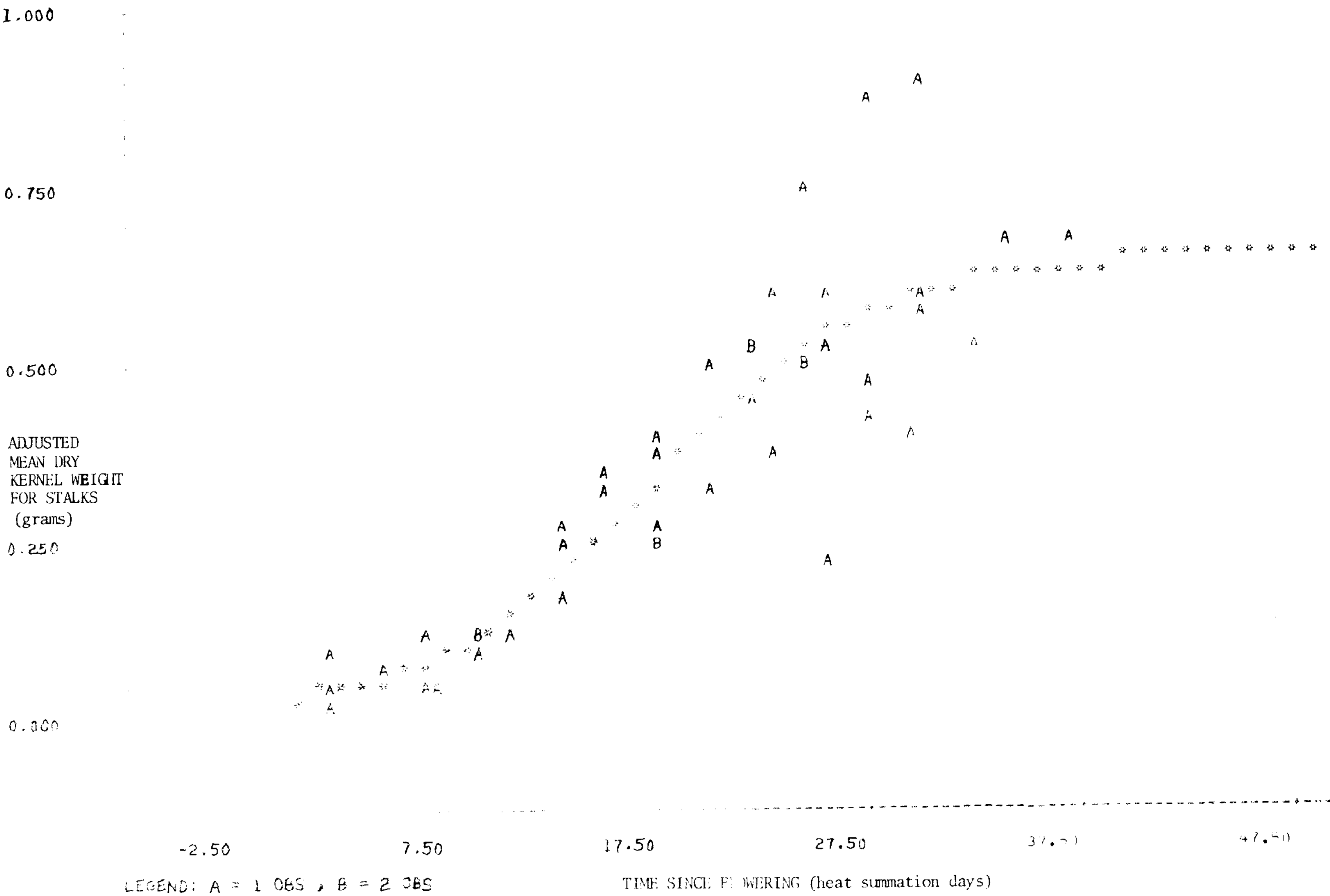

1.000
0.750
0.500
0.250
0.000
ADJUSTED MEAN DRY KERNEL WEIGHT FOR STALKS (grams)
-2.50
7.50
17.50
27.50
TIME SINCE FLOWERING (heat summation days)
LEGEND: A = 1 OBS , B = 2 OBS

Figure 18

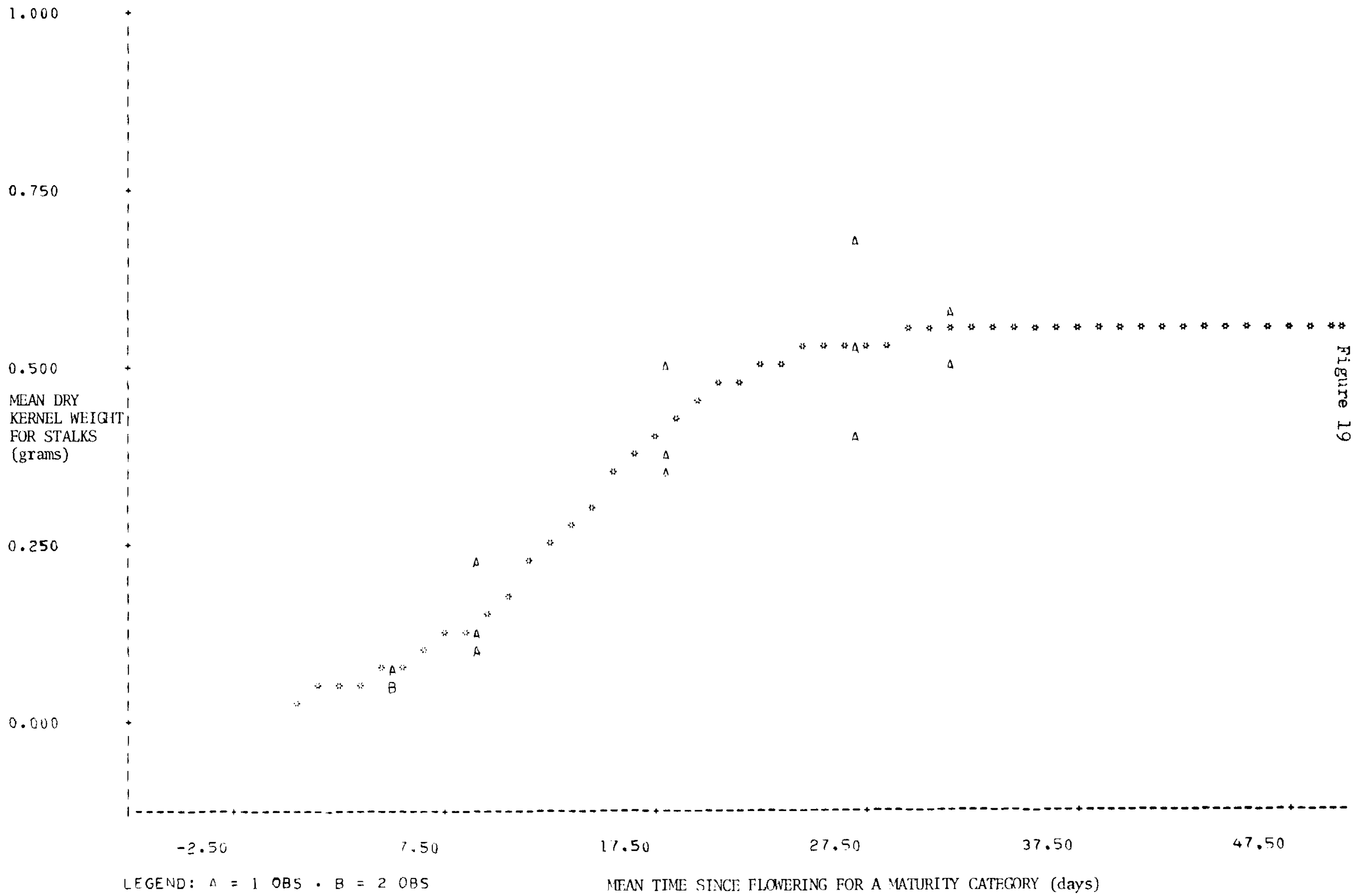

1.000
0.750
0.500
0.250
0.000
MEAN DRY KERNEL WEIGHT FOR STALKS (grams)
-2.50
7.50
17.50
27.50
37.50
47.50
LEGEND: A = 1 OBS · B = 2 OBS
MEAN TIME SINCE FLOWERING FOR A MATURITY CATEGORY (days)

Figure 19

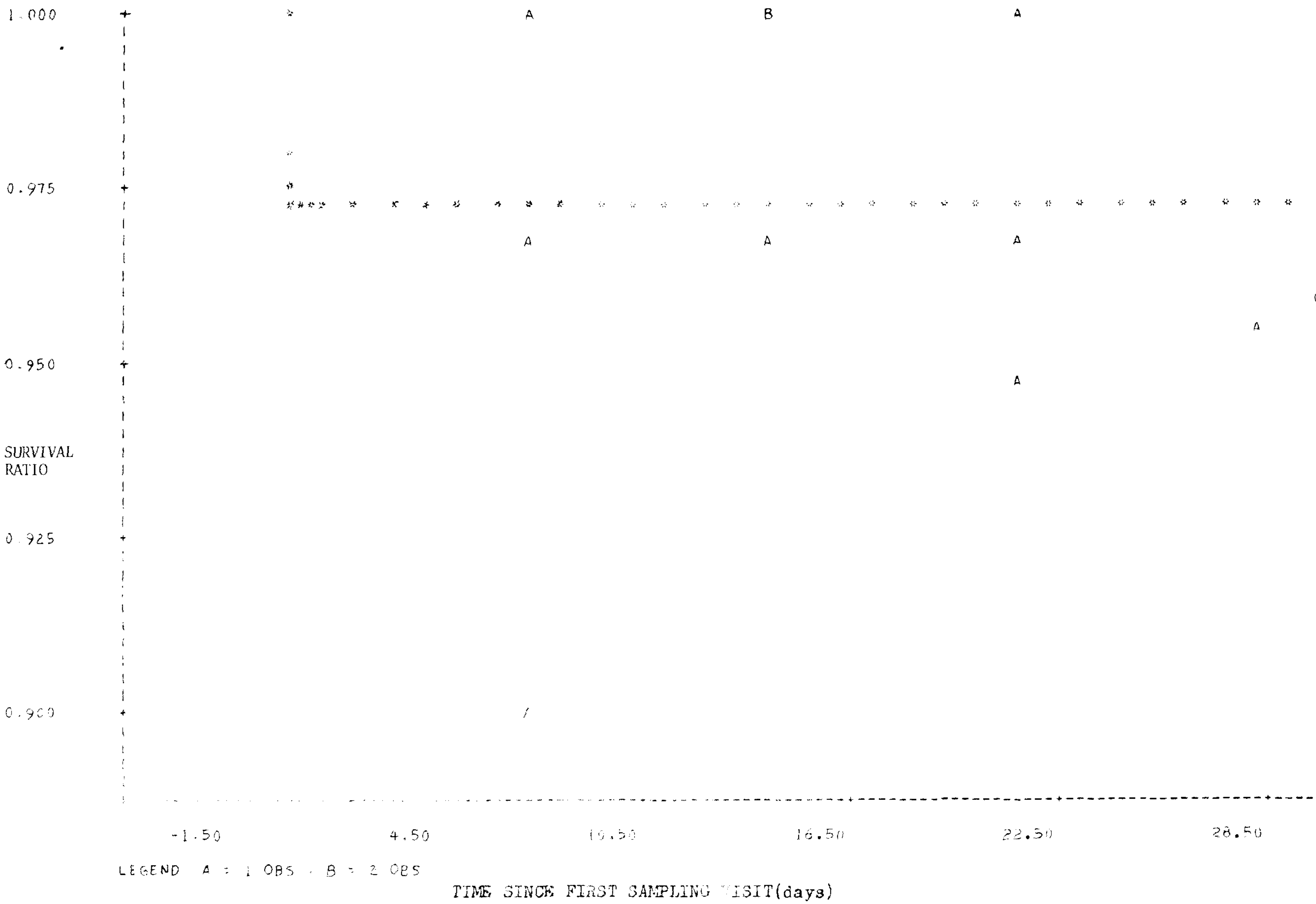

1.000
0.975
0.950
0.925
0.900
SURVIVAL
RATIO
-1.50
4.50
10.50
16.50
22.50
28.50
LEGEND A = 1 OBS , B = 2 OBS
TIME SINCE FIRST SAMPLING VISIT(days)

Figure 20