A STUDY OF WEATHER FACTORS AFFECTING COTTON YIELDS

by

W. D. Foster

Institute of Statistics
Mimeograph Series #37
For Limited Distribution

TABLE OF CONTENTS

Appendix of Tables and Figures

FINAL REPORT TO WEATHER BUREAU

A STUDY OF WEATHER FACTORS AFFECTING COTTON YIELDS

Walter D. Foster[1]

INTRODUCTION

The relationship of crop and weather would be studied far more effectively if data taken from sources which vary in both time and space could be analyzed by statistical techniques designed for most efficient combination of such material. Unfortunately, this statistical problem is a most complex one. While it is not the purpose of this paper to search for such an analysis, the latest approaches to this problem will be used in analyzing cotton-weather data which were obtained from five experiment stations throughout the South and Southwest over a period of three years.

The advantages of using multiple sources are at once apparent. A greater range of weather variables is afforded. A more representative sample of field conditions is obtained. The amount of data desired can be acquired over a much shorter length of time. With careful measurements taken from a uniform experimental design, practically all variation in yield except that portion due to weather and soil can be removed and the residual variation, the far greatest proportion of which is ascribable to weather, is available for statistical treatment. This method is inherently more accurate because it is a truer picture of plant response to climatic variation under actual conditions.

---

[1] Research Assistant, Department of Statistics, N. C. State College, August 1948

OBJECTIVES

1. To apply the latest statistical methods of combining data taken from diverse sources so that extraneous variations involved in general pooling are excluded.

2. To use amended regression technique and to search for new techniques which will use information already known about the cotton plant in order to refine and make more accurate the mathematical expressions of the relation between cotton and weather.

3. To predict cotton yields from weather data.
   - A. To compare predictors.
     - i. By regression.
     - ii. By empirical methods.
   - B. To compare the various periods in growth of plant.
     - i. For independent weather variables.
     - ii. For all weather variables jointly.

From this portion of the investigation, an attempt is made to devise an accurate forecast which will be available before the end of the growing season.

4. To investigate the distribution of weather factors used in this study.

5. To devise a punch card code for the most efficient method of recording the weather data to be collected from the Soil Weather project.

DESCRIPTION OF COTTON EXPERIMENT[1]

In a complex design repeated identically at five stations throughout the South and Southwest, four varieties of cotton were grown at two levels of fertilizer treatment. At each station, there were two planting dates per year for a series of three consecutive years.

---

[1] This experiment is known as the Cotton-Weather project, Sponsored by the ~~U.S.D.A.~~ Agricultural Marketing Service.

As will be shown by the preliminary tests, there are sufficiently large yield differences between planting dates and between stations to consider them as influenced by different weather conditions. In this way, independent observations are made available as though the experiment had been continued over a 30 year period at one location, except for soil differences.

The five stations cooperating in this experiment are as follows: Florence, South Carolina; Griffin, Georgia; Stoneville, Mississippi; Marianna, Arkansas; and Lawton, Oklahoma. From geography it is obvious that a large range of climatic factors were available during the growing season. At these stations, the four varieties of cotton grown were Oklahoma Triumph, Dixie Triumph, Shafter Acala, and Stoneville, all of which are varieties of American upland short staple cotton.

Thus far we have controlled five major effects influencing cotton yields; variety, date of planting, fertilizer treatment, station and year. To these is added a sixth, blocks. The field plan, Table 1, was replicated twice at every station to give an estimate of experimental error. Furthermore, to prevent carry-over effects from year to year such as soil depletion from affecting yield, fresh, completely different plots of ground were used each year. In this way, a continuity correlation was avoided, though possibly at the expense of a slightly larger soil variation. However, it was believed by the experimenters that this source of additional variation was negligible.

Of the many measurements taken during the experiment, only those pertaining to weather will be enumerated. Temperatures recorded daily include the maximum, minimum, and mean. Precipitation was measured both by the Weather Bureau standard eight inch rain gauge and by a recording gauge. Evaporation rates were obtained from a standard white atmometer. Soil moisture at four levels, 6", 12", 24", and 36" were measured by tensiometers, duplicated at each station for greater stability of index. Other weather phenomena were indicated in the form of remarks by the observers, such as "windy", "heavy rain", "partly cloudy", etc.

The administration of the field work was centralized in order to make the procedure as uniform as possible with every possible care taken to introduce no variation other than that which is actually encountered in field conditions. The advantage of working with data from an experimental arrangement such as this can be seen at once in the homogeneity of conditions in each sub-class in which the measurements were taken. Although the soils at the various stations are known to be dissimilar, the fertility of each was raised to a maximum by the application of a dressing optimum to each location. By this means, the effect of the soil differences is minimized except for the water content, an index of which is available. Thus the contribution of the soil factor is believed to be small compared to the weather effects and it is under this assumption that the relation of weather and cotton yield will be estimated in this study.

No mention has been made thus far of the insect factor which principally is the boll weevil. The ordinary precautions of dusting the crop early in the season against boll weevil was followed throughout the three years. For the first two years, 1939 and 1940, there was no noticeable effect due to insects. However, in the third year, the infestation was particularly heavy at Florence with varying degrees of attack at other stations. An estimate of the amount of cotton lost due to boll weevil was obtained by sampling, hence some adjustment for these figures can be made.

PRELIMINARY STATISTICAL TESTS ON DATA

As we have seen, the data consists of means of yield of seed cotton in pounds per acre classified by variety, treatment, date, year, station and block. A table of these means is shown in Table 2 where it can be seen how widely these figures vary. In order to investigate the magnitude of these differences, the first preliminary test was the analysis of variance on yield for variation within a station for a given year. For example, at Stoneville in 1939, the main effects tested were blocks, varieties, date and treatment with the first order interactions that are of primary interest, i.e., those showing variation in time. Table 3 summarizes these tests. The customary significance levels of 5% and 1% are indicated in deference to the precedence in the literature, although it is felt that the significance levels in weather and crop weather data should be more on the order of 1/2% and 1/10% due to the skewed distributions often encountered in this type of data.

The assumption that different weather conditions were encountered in the two dates of planting at a single station was well substantiated according to Table 3. Furthermore, looking at the F value for blocks, it is seen that for the most part there were no differences attributable to soil variations within stations. Nor can the variety by date interaction be considered significant, that is, to say, the varieties reacted similarly among themselves to a given planting date.

Further information obtained from Table 3 shows that in all but two instances a respectably small coefficient of variation existed, indicating the homogeneity of response per classification.

The most important information obtained from Table 3 is the general magnitude of the error mean square, which is a quantity arising from random variation between replications, such as block by treatment, block by variety, block by variety by date, etc. It is obvious from the range of these error terms that pooling among stations without further investigation would be hazardous. As an index, Bartlett's [1/] test for homogeneity of variances was applied to these errors by years. Results of this test are shown in Table 4 where it is seen that the probabilities of

---

[1/] M.S. Bartlett, Supplement to the Journal of the Royal Statistical Society, 4:137 (1937).

obtaining such widely varying figures in random sampling from the same population are very, very small. Hence we are warned against casual pooling of errors among stations within years for any exact test.

The next preliminary test is concerned with investigating the magnitude of the differences between stations and between years. Again, our hypothesis that different weather conditions existed at the five stations was substantiated by the significant between stations mean square, (Table 5).

As expected, the main effects account for the great bulk of variation with station, date, and years which carry the climatic factor, appearing considerably larger than the other contributors. As a matter of fact, the items in Table 5 which are due purely to weather and soil comprise 85% of the total variation, i.e. ignoring treatment and variety.

SELECTION OF PREDICTORS

Of the various weather elements measured during the experiment, three seem most comprehensive and indicative of the total influences exerted by weather during the growing season. These three are soil moisture, mean temperature, and evaporation. Precipitation was not used since it is obvious that soil moisture readings are a much better indication of the moisture available to the plant. Relative humidity, a function of temperature and amount of moisture in the air, was measured at only three points during the day and appears to offer nothing that is not included in an index of evaporation. Evaporation results from the action and interaction of solar radiation, air temperature, wind velocity, and humidity and thus presents in a single figure the influence of a number of atmospheric elements which if taken as individual predictors would make regression extremely cumbersome and empirical techniques too bulky. Further, evaporation has been found [1/] to be the

---

[1/] Miller, E. C., 1931. Plant Physiology. McGraw-Hill Book Company best single measure of transpiration; a first approximation to the effect of transpiration on yield is hereby afforded.

Since soil moisture was measured at four levels of depth, it was felt that some way of combining these readings into one would expedite handling of the data without serious loss of information. Brown[1/] found that in general the tap root of the cotton plant extended to 2 1/2 feet while the laterals reached that depth and sometimes went deeper. Although the maximum depth of the roots is about four feet, the bulk of the roots seemed to lie between one and two feet. From this it follows that the moisture supply at 12" and 24" is the most important while that of 6" and 36" might be classed together. Hence a weighted average of these depths was computed as

$$SM_{av} = \frac{SM_6 + 2SM_{12} + 2SM_{24} + SM_{36}}{6}$$

The figure, SM_{av}, was used throughout this study as the soil moisture index.

It is worthwhile mentioning the character of the instrument which was used to measure soil moisture. The tensiometer consists of a porous cup located at a desired depth and connected to the surface by an air-tight tube. The cup is filled with water and the resulting tension caused by a thirsty soil endeavoring to "drink" the water from the cup is measured by mercury column. Thus, the drier the soil, the greater the tension while a saturated soil exerts no tension on the cup. Although this instrument has been very popular, it has two drawbacks.[2/] First, its range extends over one atmosphere of tension, while the tension range from saturated soil to an average wilting point is several times as great. On the other hand, plant growth beyond the range of the tensiometer approaches a limit rather

---

[1/] Brown, H. B., 1938. Cotton. McGraw-Hill Book Company

[2/] Baver, L. D., 1940. Soil Physics. John Wiley and Sons Book Co. p. 207.

quickly. The second disadvantage is an inherent hysteresis. For a given amount of soil moisture, the instrument does not give the same reading when moisture is increasing as it does when moisture is decreasing, although this difference is probably much less than the sampling variation encountered in field conditions. Despite these minor limitations, the tensiometer is believed to offer a reliable index of available soil water.

The original weather records were replete with temperature readings, offering daily maximum, minimum, and mean. While it was a temptation to extend the study to include investigation of all three, all indications from the literature and the cotton experts in the N. C. Experiment Station advised narrowing the field to the mean temperature, a factor which combines the values of both maximum and minimum readings into one, though with questionable efficacy.

STATISTICAL ADJUSTMENT OF DATA

Conforming for a while to precedence, the unit of time was chosen as the week and the daily weather records were transformed into weekly means, measured from the date of planting. This maneuver erased some of the difficulty of data missing due to instrument failure, etc.

In comparing the problem at hand, that of explaining the variation in cotton yields by means of weather factors, to the problem of estimation of a true yield from data taken from sources which vary in either time or space, it is seen that each is a segment of a parent problem. When the data are collected from different centers in the same year, Cochran 1/ assumed a constant response with heterogeneous variances. (A constant response with homogeneous variances is handled by the familiar analysis of variance without adjustment.) His solution shows the relative

---

1/ Cochran, W. G., 1937. Problem Arising in the Analysis of a Series of Similar Experiments. Sup. to Jour. of Roy. Stat. Soc. Vol. IV 31.

efficiency of weighting each mean inversely as the observed variance compared to the true variance for varying degrees of freedom at each center, provided the number of centers, k, is infinitely large. Attacking the same problem, Porter[1] computed the relative efficiencies for steps of k from 2 to 20 and for varying degrees of freedom at each center.

A further complication to this problem is added when the responses at the several centers are different as well as the variances. This is the nature of the investigation headed by Cochran at present in a Navy sponsored research program.[2] Thus far, it has been learned that the most efficient estimate of the true yield is the unweighted mean in the case when the variance of the different responses remains the same from center to center. In the alternative case, when the variance of the different responses varies from center to center a system of partial or semi-weights is the most effective. It is planned that the next phase of the investigation will cover the estimation of true yield when data are available not only from different locations but in different years.

It is here that the two problems begin to overlap. While the crop-weather question attempts to explain the variation as mainly due to weather, the estimation of the true yield is accomplished after all variation except that inherently due to the crop itself has been removed. The figures in this cotton study are best unweighted according to the foregoing for use in regression. To reduce to a minimum all variation except that due weather, only the yields from the Stoneville 2B variety of cotton were selected for regression, Stoneville being a popular, high yielding variety of cotton grown widely throughout the cotton belt. Further, only yields from plants given the high level of fertilizer were considered. By this means, the soil differences between stations are minimized since each station

---

[1] Porter, S. H., 1948. The Relative Accuracy of Weighting Inversely as the Estimated Variance. Master's Thesis. N. C. State College.

[2] Progress Report, 1948, to Office of Naval Research.

applied as the high level of fertilizer that dressing considered optimum by the agronomists for the local soil.

REGRESSION

There are many possible ways regression could be applied to this problem. In order to handle as many predictors as possible, it is first necessary to assume a linear relationship between the climatic elements and final yield. This assumption is by no means a new one in crop-weather research and seems justified only when a slight improvement in the analysis by use of higher degrees is compared to the enormous work required in obtaining it. For instance, in this study there are not a sufficient number of observations to allow a least squares fit for second or higher degree. A second assumption is more critical. It is necessary to suppose that the effect of a weather factor at any given period is independent of the effect at another period during the growing season. Perhaps one of the best ways to investigate the validity of this hypothesis would be to compute various serial correlations and if they appear to be more than trivial, examine them in an effort to find a form of regression or statistical technique which would be more effective in explaining the crop-weather relationship. However, it is believed that this assumption involves no great error provided the weather variables do not deviate from their expected values too greatly.

A second great saving in the number of constants to be fitted in regression was effected by R. A. Fisher 1/ when he introduced the use of orthogonal polynomials to multiple regression. This method has been used considerably and descriptions of the process of fitting are abundant. 2/ By the use of Fisher's method, the

---

1/ Fisher, R. A. 1924. The Influence of Rainfall on the Yield of Wheat at Rothamsted. Roy. Soc. Phil. Trans. Ser. B, 213, 89-142.

2/ Davis, F. E. and Pallesen, J. E. 1940. Effects of the amount and distribution of rainfall and evaporation during the growing season on yields of corn and spring wheat. Jour. Agr. Res. 60:1-33.

Hopkins, J. W. 1935. Weather and wheat yield in western Canada. Canada Jour. Res. 12:306-334.

Davis, F. E. and Harrell, G. D., 1942. Relation of Weather and its Distribution to Corn Yields. U.S.D.A. Tech. Bull. 806.

change in yield due to a unit of the weather factor at a given week can be computed, thus showing the critical stages with regard to a given weather factor during the growing season.

While Fisher's model represents a great improvement in the technique for studying the crop-weather relationship, it still assumes that only one weather factor is to be studied at one time, or if several are used, that each is independent of the effect of the other. Hendricks and Scholl[1] have amended Fisher's approach by setting up a model which studies the effect of weather elements simultaneously as well as separately. For example, in their paper temperature, precipitation, and the joint effect of temperature and precipitation are used to account for variation in corn yield.

In this study, regressions modeled both from Fisher and Hendricks and Scholl are computed. In a gross way, a comparison of the two is provided, although the main purpose is to study the effectiveness of various predictors. For instance, the first regression patterned after Fisher has as independent predictors soil moisture, evaporation and mean temperature. Let us call this function regression # 1. The second, regression # 2, follows Hendricks and Scholl and includes evaporation, soil moisture, and the joint effect of evaporation and soil moisture. Regression # 3 expresses yield as a function of mean temperature, soil moisture, and the joint effect of temperature and soil moisture.

The models of these regressions are shown in Table 6. It is not proposed in this paper to offer a complete description of the development of these techniques since full details can be obtained from Hendricks' and Scholl's article if desired.

The data for these regressions, as explained before, consist of yield of seed cotton in pounds per acre, while the weather elements have been averaged over

---

[1] Hendricks, W. A. and Scholl, J. C. 1943. Techniques in Measuring Joint Relationships. Tech. Bull. 74. Agr. Exp. Sta., N. C. State College.

weekly periods. Using the date of planting, as the zero or reference point, the weeks were numbered 1, 2, 3 etc. Since there was considerable variation in the length of records kept by each station cooperating in this experiment, it was necessary to select those weeks affording the longest consecutive series when all five stations were considered at once. For instance, if station A shows weather records from weeks 1 to 6 and station B shows weather records from weeks 2 to 7, then weeks 2 to 6 are available for study for both stations. In this manner, nine consecutive weeks starting with the eighth week after planting and ending with the 16th week were available. In terms of the cotton plant, this overall period includes half of the squaring period, blooming and the setting of young bolls. The periods in this study were based on the actual phonological dates and have been defined in terms of weeks after planting as follows: squaring, 6-9 weeks; blooming, 10-12 weeks, setting young bolls, 13-16 weeks, opening of bolls, 17-19 weeks. These weeks correspond almost exactly to those described in the literature. C. B. Doyle[1] and Brown[2] have indicated that the period of blooming is the most critical in the life of the plant, while the water requirements for squaring and setting bolls are high.

A further step in handling the data was to compensate for the boll weevil damage. Though the experiment as a whole was relatively free from this pest, certain of the yield figures are low because of it. The rate of infestation was estimated by taking random samples of the cotton and counting the number of locks per boll destroyed by the weevil. This rate was expressed as a percentage and the adjusted yield would be that figure which when decreased by the rate of infestation becomes the actual recorded yield. For instance, at Station A the

---

[1] Doyle, C. B. 1941. Yearbook of Agriculture: Climate and Man. pp. 348-363.

[2] Brown, H. B., Loc cit.

rate of infestation is 25% and its recorded yield is 1200 lbs./acre. Then the adjusted is 1200/.75 or 1600 lbs./acre. Despite the apparent logic of this method, in some cases it fails to adjust the yield to a credible figure. Nevertheless, for the lack of a better method, the above procedure was followed. Later in this paper, a second method of adjustment will be described though it is not applicable to regression.

A comparison of the three regressions is best seen in tabular and graphic form. In Table 6 in the appendix where the models, regression coefficients, tests of significance, and the multiple correlation coefficients are given, it is seen that the joint relationship of soil moisture and evaporation was easily the most effective of the three models with a correlation of .73, (R^2 = .537). Further inspection shows that regression # 1 is fitted with nine constants instead of ten, while regression # 3 has only five. In regression # 1, a_o was deleted from the predictors because it was so closely correlated to A_o. When such a condition is present, solution for the regression coefficients is possible by the abbreviated Doolittle method of matrix inversion only when an extremely large number of digits is carried throughout the computation. Since eight digits were carried in these calculations, it was deemed more feasible to drop the predictor than to attempt the operation with 12 or more digits. This problem was compounded in regression # 3 where there were so many dependencies in the matrix of normal equations that its virtual rank was reduced to five.

From these observations, it immediately appears that mean temperature as a predictor is the least important of the three despite the slight evidence to the contrary offered by the t-test for the regression coefficients in regression # 1 in which the three temperature coefficients present a questionably best showing.

The quadratic terms in regression # 2 without exception showed the smallest probabilities of being obtained by chance. The fact that the coefficients for evaporation appear to be better estimated than those for either soil moisture or the joint effect of evaporation or soil moisture is presented without comment.

The first partial derivative of yield in regression # 2, with respect to evaporation when evaluated for t and arbitrary values of soil moisture offers a family of curves which shows the effect of a unit of evaporation for different water tension values on the final yield of Stoneville cotton at different weeks in the growing season. These are shown in Figure # 2. Likewise, the partial with respect to soil moisture leads to another family of curves, also shown in Figure # 2. The labels and orientation of the curves may be confusing unless one remembers that soil water is measured as tension with the result that a high tension indicates a low amount of soil water. The other units are direct; evaporation, for instance, is low for a low index. There is another precaution to observe. Since only a quadratic was fitted to the data, it is best not to extrapolate these curves. As a matter of fact, even near the borders of the chart, some of the curves tend toward the incredible.

Similar charts appear for the other two regressions, although it is immediately apparent that with such large confidence limits as indicated by the t-tests for the coefficients, the true location and orientation of the curves could easily be entirely different.

SCORE METHOD

Thus far, it has been seen that regression has been hampered by restrictive assumptions and more severely by the limited number of constants that could be fitted. These penalities are the price paid for the advantages of tests of significance, the availability of confidence limits, estimation of the contribution of each weather factor for a given week on the final yield and following the well-known path of previous investigation. As was indicated in the preliminary report, a different approach, not necessarily new, was intended whose results were to be judged by comparison with the regressions rather than by statistical tests of significance.

The basic idea in this empirical investigation is to put on a quantitative basis the lore and experience of those who have worked many years with cotton. By assembling the many observations of cotton yields as affected by various types of weather, a little experimentation in assigning values to weather factors according as they are desirable or undesirable leads to a score method which has several patent advantages over regression. The primary advantage is the complete freedom in the simultaneous use of as many weather factors as desired. In our regression studies, only two variables could be handled by the joint relationship technique. The flexibility innate in this approach offers a choice of many comparisons and investigations with a minimum of calculation. The problem of adjusting for boll weevil infestation is solved much more effectively. Computation is considerably simplified, both in construction and in use as a prediction method. The improvement and expansion of this or similar methods lends considerable promise towards a more accurate definition of the cotton-weather relationship which C. B. Doyle[1] has concluded is not likely through the usual investigations of precipitation and temperature.

From Doyle,[1] Brown,[2] McNamara[3] and associates, Hawkins[4] and associates, Kramer[5] and others, generalities about the believed optimum climatic requirements of the

---

[1] Doyle, C. B. Loc cit.

[2] Brown, H. B. Loc cit.

[3] McNamara, H. C., Hooton, D. R., and Porter, D. D. 1940. Differential Growth Rates in Cotton Varieties and their Responses to Seasonal Conditions at Greenville, Texas. Tech. Bull. 710 USDA.

[4] Hawkins, R. S., Matlock, R. L., and Hobart, C. 1933. Physiological Factors Affecting the Fruiting of Cotton with Special Reference to Boll Shedding. Tech. Bull. 46 Univ. of Arizona Agriculture Exp. Sta.

[5] Kramer, P. S. 1944. Soil Moisture in Relation to Plant Growth. Botanical Review 10: 505-556.

cotton plant for each of the periods were obtained. For instance, during the period of blooming, the water requirement is believed to be at a maximum, i.e., a very moist soil is best. Evaporation or transpiration is believed to have little effect except when the soil is dry. Temperature is thought to have little effect. On the other hand, a dry soil causes excessive shedding of blossoms.

The optimum conditions were set up in a qualitative sense for this study purposely without regard to the weevil and the weather conditions for its spread. Although it is known that the weevil thrives and spreads rapidly when soil water is abundant, relative humidity high, and temperature warm, these facts were omitted from consideration, since the presence of these conditions by no means guarantees the presence of the weevil. A correction for the weevil is shown later.

The construction of a chart to assign quantitative values to certain combinations of weather factors was achieved first by dividing the range of each weather variable into suitable blocks. The range of soil moisture tension was found to be 0 to 60. It was convenient to define 12 blocks, running 0-5, 5-10, 10-15, etc., with the convention that a reading of 10 would fall in the 10-15 block. After a great deal of experimentation, the most effective size of block for temperature was found to be ten degrees and for evaporation ten unit blocks up to 50, then 15 units to 65, and over 65 for the last block. These coordinates can be seen in Figure # 4.

The second step in the construction was to define five basic classifications into which the various weather combinations would fall. These classifications, excellent, good, fair, poor, and very poor are used to characterize the effect of weather on final yield according to the writers on this subject. The classification of excellent confers a high score while very poor on the other extreme confers a low score; a plant with a high total score is expected to have a high yield. Thus the assignment of values to the classifications is seen to become a problem in interpolation.

Once the number of blocks was decided, the assignment of values to these blocks was undertaken. In Figure # 4 where there are 216 blocks, the range of values was taken as 1 to 216. Then the five classifications were allotted a number of blocks to correspond approximately with a normal distribution (although no use is made of this arbitrary decision.) Of the 216 blocks, "excellent" and "very poor" each received 22, "good" and "poor" each received 52, and "fair" received 68. Because of the impossibility of ranking every block, that is, to say of two adjacent blocks near optimum that one is better than the other, each classification was further graded into three approximately equal sets, denoted by different colors and each block of the same color in the same major classification assigned the middle value of its own range. For instance, the "excellent" classification was allotted 22 blocks. These were further graded into eight blocks colored orange, seven blocks green and seven blocks brown. Counting from 216, 22 blocks would end with 195. If the orange blocks are to be assigned the highest values because they represent the best growing conditions for a given period, then they would occupy the values 216 down through 209. Using the middle value of this range, all the orange blocks in the "excellent" classification receive the value 212. (When a choice between two middle values occurs, the even number was taken.) Similarly, the green blocks were evaluated at 205, the brown at 198. The rest of the scheme can be seen on the margin of Figure # 4. This arrangement was followed for each of the four periods, squaring, blooming, setting young bolls, and opening of bolls. Since there was considerable leeway in arranging the order of the weather conditions from best to worst, considerable experimentation was involved before the most successful plan was selected. The relative effectiveness of the last three plans can be seen in Table 8.

Adjustment for boll weevil damage becomes extremely simple. An estimate of the rate of infestation is required and can be obtained by sampling the crop and reporting, as was done in the Cotton-Weather project, the ratio of damaged locks to total locks as a percentage. After the scores for the weeks have been added, the total score is adjusted by multiplying it by the percentage of good locks. Thus, if the total score is 1500 and the rate of infestation is 12%, then the adjusted score becomes 1500 x 88% or 1320. Should other measurable conditions arise, further adjustment is made in a similar manner.

As an example, the steps in obtaining a score will be described in detail. At Florence, South Carolina for the normal planting in 1939, the soil moisture reading for the 10th week was 11.1 while the mean temperature was 84.3 and the evaporation reading was 45.7. Referring to Figure 4 and using the chart for blooming, since the 19th week is defined to belong to blooming, the correct square is found in the bottom section in "good" and is colored green. The value for this square is 167 according to the code on the margin. Likewise, the score for the 11th week was found to be 198. Repetition of this process through the growing season gives a series of scores whose total is 1440. Since there was no weevil damage reported, no adjustment of this score is necessary. Plotting this score against the recorded yield of 2366 lbs./acre for the Stoneville variety, we obtain a point, labelled for reference as 1, (see Legend, Table 7). The plot of the rest of the points comprises a scatter diagram, Figure 5A. Curiosity immediately required the application of the scoring method to the yields of the other three varieties with weeks 10-19 taken as the scoring interval. These results are also shown in Figure 5, where it is apparent that the two high yielding varieties, Oklahoma Triumph and Stoneville, have almost identical scatters while the two lower yielding varieties, Acala and Dixie Triumph are very similar. A cubic was fitted by least squares to the averaged yields of Oklahoma and Stoneville and modified near the end points where the characteristics of a polynomial fit are not true to actual trend. The lower mean yields are satisfied with a linear fit. Thus, Figure 6 becomes a prediction chart based on a scoring interval of 10-19 weeks for varieties of cotton similar to types studied here.

However, it is desired to offer a forecasting tool which is available earlier in the season. By taking advantage of the flexibility of the scoring method, it is easy to obtain a direct comparison of any combination of periods of growth and to learn if a scoring method which terminates before the time of harvest is any less accurate and if so, by how much. By totaling the scores over any three of the four periods and assuming a linear fit, the corresponding correlations were computed and compared in Table 8. Then, if the omission of given period reduces the correlation by more than any other omission, it is concluded that the period omitted has the greatest effect on yield and is the most critical from the standpoint of climatic factors. By a simple ranking test, Table 8, the blooming period was easily the most important to final yield, followed by the periods of setting bolls, opening and squaring. According to scoring method # 1 which is the one adopted in this study, there is very little difference whether the scoring interval is taken as 6-16 weeks or 10-19 weeks. Then by observing the scatter diagrams in Figure # 7 and 8, where the same scheme has been followed as described above for Figure 5 and 6, a forecasting diagram is present which makes a forecast available three weeks before the first of the cotton is ready to be ginned. Apparently it is a characteristic of the scoring diagram which consistently suggests a cubic fit for the 10-19 week scoring interval and a linear fit for the 6-16 week interval.

In order to obtain sufficient yield data to evaluate all of these periods when nine weeks of data were available for regression, a small amount of extrapolation of certain weather values was necessary. In nearly all of these extrapolations, concomitant measurements existed in sufficient detail to afford a very intelligent guess. As a further justification, the range of values in the individual blocks is large enough to absorb considerable variation from the true value. Even a gross error would probably result in no more than the substitution of an adjacent block for the correct one, with an undetectable result in the final correlation.

While the scoring method presented is far from elegant, it has certainly held its own in comparison with the other approaches here. A systematic experimentation in the construction of these or similar charts by agronomic experts would undoubtedly lead to more accurate predictions than offered here.

FREQUENCY DISTRIBUTION OF WEATHER FACTORS

In the search for a stable pattern of frequencies of occurrence, the three weather factors, evaporation, soil moisture and mean temperature were each divided into small classes of five, three and two units respectively such that there were at least 20 classes in the range of values for each factor. Whether the values enumerated were daily or weekly, the histograms were obviously inconsistent both for individual stations and for pooled observations, although it was noted that mean temperature tended to show the least skewness of the three. Since these frequency patterns are best investigated when extensive data are available; the effort in this direction was purposely curtailed.

CODE FOR PUNCH CARDS

A preliminary code is given in Table 9 for recording on IBM punch cards the large quantities of weather data which are being accumulated by the Southern Cooperative Group. At the present time, this code is being revised by its author, Jay T. Wakeley, to improve the index of temperature and sunlight while other changes are probable.

SUMMARY

Measurements of soil moisture, evaporation, and mean temperature have been employed to study the relationship of weather to final yield of cotton in an identical experiment conducted in 1939, 1940 and 1941 at five experimental stations widely scattered throughout the cotton belt.

The preliminary tests have shown that differences in yields between stations and between years were by no means attributable to chance variation. These tests also proved that the differences between a normal and a late planting date at the same location were statistically significant. Hence, from the control of other variables in the experiment, it is argued that the differences in yield are due to weather and soil, other miscellaneous factors contributing practically nothing. Further, the variation caused by soil differences is minimized by the application of fertilizer optimum for each location, leaving weather as the only major influence affecting final yield. The damage from boll weevil was measured and the yields were adjusted accordingly.

Linear regression in the form developed by R. A. Fisher and amended by W. A. Hendricks was applied to these data in preference to any other type of regression, since it is felt that these models offer the latest and most comprehensive approach. Three regressions, each with ten constants to be fitted, were computed, using various combinations of weather variables. The greatest amount of variation was explained by using evaporation, soil moisture and the joint effect of evaporation and soil moisture. A second combination, evaporation, soil moisture and mean temperature taken independently, was less effective, while the third, soil moisture, mean temperature and the joint effect of soil moisture and mean temperature was found to contain five internal dependencies which reduced the number of parameters to five. It was concluded that mean temperature was the least effective of the three weather factors, either singly or in combination.

Applied to cotton data, the regressions used in this study were felt to be insufficient, since the data permitted a maximum of only ten predictors. Cotton is known to be extremely sensitive to weather, a fact which requires a more complete description of the climatic factors than apparently afforded by the regression studies presented here. In an effort to manufacture a device which approximated the joint effect of all three weather variables simultaneously, it was necessary to abandon regression and endure the loss of the associated tests of significance.

Drawing upon the experience of cotton experts, a score method was devised which allotted a high score for the combination of weather factors considered good and a low score for undesirable weather. In terms of final results, the score method was considerably more accurate in predicting cotton yields than any of the regressions. Due to the flexibility of the score method, a forecast of final yield could be made available at the end of the 16th week from planting without loss of accuracy. Purely as a prediction tool, the score method or similar schemes appear to be considerably more effective than regression and certainly far easier to construct. The penalties for these advantages are the inability to place confidence limits on estimates and the loss of tests of significance.

An attempt was made to find a recurring pattern in distribution of the weather elements measured. Nothing definite was discovered in several methods of constructing frequency diagrams. Further, should any distribution have been offered, its validity would be open to serious question due to the relative paucity of observations.

Table 1: Sample Planting Plan From Cotton Weather Experiment

Plots 103 feet, 8 1/2 inches - 1/30 acre, 4-row plots, width of row - 3 feet, 6 inches

Subblock:	a		b	Border Cotton	c		d	
BORDER ENDS	C109 $V_4D_1T_2$ Acala (Shafter) Normal Planting Date Low Fertilizer	ALLEY	C209 $V_2D_2T_1$ Dixie Triumph 12 Late Planting Date High Fertilizer	ALLEY	C309 $V_4D_1T_2$ Acala (Shafter) Normal Planting Date Low Fertilizer	ALLEY	C409 $V_4D_2T_2$ Acala (Shafter) Late Planting Date Low Fertilizer	BORDER ENDS
	C110 $V_3D_1T_2$ Stoneville 2B Normal Planting Date Low Fertilizer		C210 $V_1D_1T_2$ Okla. Triumph 29 - 44 Normal Planting Date Low Fertilizer		C310 $V_3D_2T_1$ Stoneville 2B Late Planting Date High Fertilizer		C410 $V_4D_1T_1$ Acala (Shafter) Normal Planting Date High Fertilizer	
	C111 $V_4D_2T_1$ Acala (Shafter) Late Planting Date High Fertilizer		C211 $V_3D_2T_2$ Stoneville 2B Late Planting Date Low Fertilizer		C311 $V_1D_2T_2$ Okla. Triumph 29 - 44 Late Planting Date Low Fertilizer		C411 $V_1D_1T_2$ Okla. Triumph 29 - 44 Normal Planting Date Low Fertilizer	
	C112 $V_2D_2T_2$ Dixie Triumph 12 Late Planting Date Low Fertilizer		C212 $V_4D_2T_2$ Acala (Shafter) Late Planting Date Low Fertilizer		C312 $V_3D_1T_2$ Stoneville 2B Normal Planting Date Low Fertilizer		C412 $V_3D_2T_2$ Stoneville 2B Late Planting Date Low Fertilizer	
	C113 $V_1D_2T_2$ Okla. Triumph 29 - 44 Late Planting Date Low Fertilizer		C213 $V_1D_2T_1$ Okla. Triumph 29 - 44 Late Planting Date High Fertilizer		C313 $V_2D_2T_2$ Dixie Triumph 12 Late Planting Date Low Fertilizer		C413 $V_2D_1T_2$ Dixie Triumph 12 Normal Planting Date Low Fertilizer	
	C114 $V_2D_1T_1$ Dixie Triumph 12 Normal Planting Date High Fertilizer		C214 $V_4D_1T_1$ Acala (Shafter) Normal Planting Date High Fertilizer		C314 $V_1D_1T_1$ Okla. Triumph 29 - 44 Normal Planting Date High Fertilizer		C414 $V_3D_1T_1$ Stoneville 2B Normal Planting Date High Fertilizer	
	C115 $V_3D_2T_1$ Stoneville 2B Late Planting Date High Fertilizer		C215 $V_3D_1T_1$ Stoneville 2B Normal Planting Date High Fertilizer		C315 $V_4D_2T_1$ Acala (Shafter) Late Planting Date High Fertilizer		C415 $V_1D_2T_1$ Okla. Triumph 29 - 44 Late Planting Date High Fertilizer	
	C116 $V_1D_1T_1$ Okla. Triumph 29 - 44 Normal Planting Date High Fertilizer		C216 $V_2D_1T_2$ Dixie Triumph 12 Normal Planting Date Low Fertilizer		C316 $V_2D_1T_1$ Dixie Triumph 12 Normal Planting Date High Fertilizer		C416 $V_2D_2T_1$ Dixie Triumph 12 Late Planting Date High Fertilizer	
	MAIN BLOCK I		BORDER		COTTON		MAIN BLOCK II	

Table 2: Table of Mean Yields of Seed Cotton (lbs./acre) for Planting Dates, Varieties and Treatments[1] for all Stations and Years.

Station & Year	Treatment	Normal Planting Date V_1	V_2	V_3	V_4	Late planting Date V_1	V_2	V_3	V_4	Mean
S.C. '39	1	2052	2374	2366	2054	2250	2192	2170	1952	2176
	2	1572	1641	1638	1540	1524	1570	1726	1519	1591
	Mean	1812	2008	2002	1797	1887	1881	1948	1736	1884
'40	1	1560	1722	1716	1548	1158	1438	1611	1458	1526
	2	964	1410	1575	1455	1120	1116	1533	1347	1315
	Mean	1262	1566	1646	1502	1139	1277	1572	1402	1420
'41	1	1048	579	812	914	255	118	172	194	512
	2	681	586	585	783	267	100	105	62	396
	Mean	864	582	698	848	261	109	138	128	454
Ga. '39	1	1137	1110	1230	1146	765	663	849	624	940
	2	1002	1134	1233	1110	897	843	792	558	946
	Mean	1070	1122	1232	1128	831	753	820	591	943
'40	1	1239	1203	1227	1338	1122	1041	882	984	1130
	2	1038	975	1572	1008	1056	1083	858	699	1036
	Mean	1138	1089	1400	1173	1089	1062	870	842	1083
'41	1	1180	914	745	962	922	476	642	514	794
	2	897	728	572	738	786	458	459	428	633
	Mean	1038	821	658	850	854	467	550	471	714
Miss. '39	1	2206	2227	2515	1716	1658	1424	1877	1082	1838
	2	1898	2144	2327	1514	1754	1648	1502	1175	1745
	Mean	2052	2186	2421	1615	1706	1536	1690	1128	1793
'40	1	2130	1820	2344	1410	1381	1127	1492	616	1538
	2	1974	1944	2302	1548	1248	1177	1423	720	1542
	Mean	2052	1882	2323	1479	1314	1152	1458	668	1540
'41	1	2476	2172	2378	2294	1701	1299	884	1630	1854
	2	1981	2137	2161	2180	1603	1103	1232	1391	1724
	Mean	2228	2154	2270	2237	1652	1201	1058	1510	1789

---

[1] Treatment 1 is "high fertilization" for the station in question. Treatment 2 is "low fertilization" for the station in question. These levels are consistent except for Lawton, Oklahoma, where (10" between rows) and (20" between rows) were the treatments.

Table 2: (Continued)

Station & Year	Treatment	Normal Planting Date V$_1$	V$_2$	V$_3$	V$_4$	Late Planting Date V$_1$	V$_2$	V$_3$	V$_4$	Mean
Ark. '39	1	1614	1524	1735	1486	1478	1296	1634	1451	1527
	2	1370	1381	1871	1522	1364	1436	1753	1280	1497
	Mean	1492	1452	1803	1504	1421	1366	1694	1366	1512
'40	1	2433	1848	2415	1863	2070	1404	2082	1566	1960
	2	1983	1764	2469	1668	1974	1578	1617	1116	1771
	Mean	2208	1806	2442	1766	2022	1491	1850	1341	1866
'41	1	2097	1557	1992	1980	1896	1488	2013	1752	1847
	2	1683	1542	1545	1824	1713	1302	1767	1413	1599
	Mean	1890	1550	1768	1902	1804	1395	1890	1582	1723
Okla. '39	1	176	166	175	144	142	78	128	76	136
	2	196	162	176	196	114	122	106	76	144
	Mean	186	164	176	170	128	100	117	76	140
'40	1	755	852	939	852	810	708	726	618	782
	2	870	840	822	807	795	696	714	594	767
	Mean	812	846	880	830	802	702	720	606	774
'41	1	1014	753	873	924	540	255	231	405	624
	2	978	807	1035	915	390	171	204	294	599
	Mean	996	780	954	920	465	213	218	350	612

1/ Treatment one is "high fertilization" for the station in question. Treatment two is "low fertilization" for the station in question. These levels are consistent except for Lawton, Okla., where 10" between rows and 20" between rows were the treatments.

Table 3: Analyses of Variance for Yield of Seed Cotton (pounds per acre) by Stations and Years.[1]

Station	Source	d.f.	Actual F Ratio			Tabulated F	
			1939	1940	1941	.05	.01
Florence, S. C.	Blocks	3	2.86	6.79**	5.91**	3.3	5.6
	Varieties	3	6.61**	5.89**	21.72**	3.3	5.6
	Dates	1	1.29	2.88	873.81**	4.6	8.9
	Treatments	1	250.93**	29.19**	33.46**	4.6	8.9
	V x D	3	1.30	#	6.68**	3.3	5.6
	V x T	3	1.30	1.04	3.62*	3.3	5.6
	D x T	1	#	#	10.28**	4.6	8.9
	Error M.S.	14	10903.406	6510.696	3179.531		
	Coef. of Var. (%)		5.5	5.3	12.4		
Griffin, Ga.	Blocks	3	2.08	1.31	#	3.3	5.6
	Varieties	3	3.65	#	3.09	3.3	5.6
	Dates	1	118.85	7.95*	8.30*	4.6	8.9
	Treatments	1	#	1.40	3.27	4.6	8.9
	V x D	3	2.98	1.80	#	3.3	5.6
	V x T	3	#	1.27	#	3.3	5.6
	D x T	1	1.36	0.00	#	4.6	8.9
	Error M.S.	14	10179.482	59394.214	63397.562		
	Coef. of Var. (%)		10.7	22.5	35.6		
Stoneville, Miss.	Blocks	3	4.88*	1.62	#	3.3	5.6
	Varieties	3	71.40**	41.81**	2.68	3.3	5.6
	Dates	1	254.17**	214.19**	109.46**	4.6	8.9
	Treatments	1	7.20*	#	2.09	4.6	8.9
	V x D	3	6.13**	#	2.91	3.3	5.6
	V x T	3	4.43*	1.33	#	3.3	5.6
	D x T	1	8.64*	#	1.41	4.6	8.9
	Error M.S.	14	9638.430	23053.054	53365.464		
	Coef. of Var. (%)		5.5	9.8	12.9		

---

[1] In each analysis of variance, the second order interaction V x D x T with two degrees of freedom has been omitted. This term has no valid meaning for this study, nor is it a true error term.

\# Cases in which error M. S. greater than the M. S. being tested.

** Significant at 1% level.

* Significant at 5% level.

Table 3: (Continued)

Station	Source	d.f.	Actual F Ratio			Tabulated F	
			1939	1940	1941	.05	.01
Marianna, Ark.	Blocks	3	4.80	10.83**	#	3.3	5.6
	Varieties	3	9.79**	40.47**	6.72**	3.3	5.6
	Dates	1	4.01	61.27**	2.70	4.6	8.9
	Treatments	1	#	15.20**	13.90**	4.6	8.9
	V x D	3	#	3.16	1.88	3.3	5.6
	V x T	3	1.60	2.83	#	3.3	5.6
	D x T	1	#	#	#	4.6	8.9
	Error M.S.	14	20563.982	18804.214	35481.536		
	Coef. of Var. (%)		9.5	7.4	10.9		
Lawton, Okla.	Blocks	3	#	2.31	2.77	3.3	5.6
	Varieties	3	1.63	3.08	6.40**	3.3	5.6
	Dates	1	32.25**	33.58**	242.90**	4.6	8.9
	Treatments	1	#	#	#	4.6	8.9
	V x D	3	1.02	3.74*	1.42	3.3	5.6
	V x T	3	#	1.09	#	3.3	5.6
	D x T	1	1.60	#	#	4.6	8.9
	Error M.S.	14	964.000	4309.679	11901.054		
	Coef. of Var. (%)		16.0	8.5	17.8		

Cases in which error M.S. greater than the M.S. being tested

** Significant at 1% level.

* Significant at 5% level.

Table 4: Test of homogeneity of error variances for yield of seed cotton between stations by years.

5 Stations

	X^2	Probability of a greater X^2
1939	10.07	.03
1940	60.62	$\ll$.0001
1941	30.26	$\ll$.0001

4 Stations, Lawton omitted

	X^2	Probability of a greater X^2
1939	2.78	.45
1940	16.82	<.001
1941	24.29	<.0001

< means "Less than".

$\ll$ means "Much less than".

Table 5: Analysis of Variance for Yield of Seed Cotton (pounds per acre) Combining Five Stations for Three Years.

Source	d.f.	Mean Square
Stations	4	25,650,902**
Years	2	3,278,068**
S x Y	8	5,006,294**
Reps in Stations and Years	15	132,734
Dates	1	14,744,133**
Varieties	3	1,051,302**
Treatments	1	1,388,776**
D x V	3	153,922**
D x T	1	78,081*
V x T	3	67,468*
D x S	4	1,140,663**
D x Y	2	650,400**
D x S x Y	8	291,874**
V x S	12	235,310**
V x Y	6	457,428**
V x S x Y	24	102,366**
T x S	4	302,114**
T x Y	2	17,465 N.S.
T x S x Y	8	168,030 **
D x V x S	12	35,152**
D x V x Y	6	12,710 N.S.
D x V x S x Y	24	134,444 **
D x T x S	4	30,754 N.S.
D x T x Y	2	7,972 N.S.
D x T x S x Y	8	11,740 N.S.
V x T x S	12	22,460 N.S.
V x T x Y	6	7,182 N.S.
V x T x S x Y	24	26,805 **
Blocks in Reps in Stations and Years	30	23,855
Pooled Error	240	13,304
Total	479	--

** Significant at 1% level.

* Significant at 5% level.

N.S. Statistically non-significant.

Table 6: Table Showing Regression Models, Coefficients, Tests of Significance, and Multiple Correlation Coefficients.

Code: T = weekly mean temperature
E = weekly mean evaporation
SM = weekly mean soil moisture
t = week from planting date

Regression 1: $Y = F(T,E,SM)$ $R = .67$ $R^2 = .44$

Model: $Y = Ao + ao\Sigma T + a_1\Sigma Tt + a_2\Sigma Tt^2 + bo\Sigma E + b_1\Sigma Et + b_2\Sigma Et^2 + co\Sigma SM + c_1\Sigma SMt + c_2\Sigma SMt^2$

Coefficients

A_o	a_o	a_1	a_2	b_o	b_1	b_2	c_o	c_1	c_2
-6,581.85		6.937	1.783	3.304	-1.098	-1.016	5.701	.541	-.566

T test for above coefficients: Probability of a larger t.

.43		.44	.33	.50	.50	.48	.50	.50	.50

Regression 2: $Y = F(E,SM,E\text{-}SM)$ $R = .73$ $R^2 = .54$

Model: $Y = Ao + ao\Sigma E + a_1\Sigma Et + a_2\Sigma Et^2 + bo\Sigma SM + b_1\Sigma SMt + b_2\Sigma SMt^2 + co\Sigma ESM + c_1\Sigma ESMt + c_2\Sigma ESMt^2$

Coefficients

A_o	a_o	a_1	a_2	b_o	b_1	b_2	c_o	c_1	c_2
3,751.84	19.698	-9.414	-4.375	23.318	-2.254	-4.334	-.868	.159	.151

T test for above coefficients: Probability of a larger t.

.19	.24	.13	.10	.20	.50	.16	.22	.27	.12

Regression 3: $Y = F(T,SM,T\text{-}SM)$ $R = .43$ $R^2 = .18$

Model: $Y = Ao + ao\Sigma T + a_1\Sigma Tt + a_2\Sigma Tt^2 + b_o\Sigma SM + b_1\Sigma SMt + b_2\Sigma SMt^2 + co\Sigma TSM + c_1\Sigma TSMt + c_2\Sigma TSMt^2$

Coefficients

A_o	a_1	a_2	b_o	b_1	b_2	c_o	c_1	c_2
1,803.33	-.223			1.720		.097		-.066

T test for above coefficients: Probability of a larger t

.004	.50			.30		.39		.16

Table 7: Cotton Yields and Weekly Temperature, Evaporation, and Soil Water Variables.

#	ΣT	ΣTt	ΣTt^2	ΣE	ΣEt	ΣEt^2	ΣSM	ΣSMt	ΣSMt^2	ΣESM	ΣESMt	$\Sigma ESMt^2$	ΣTSM	ΣTSMt	$\Sigma TSMt^2$	Y_1	Y_2	Y_3	Y
1	730	22	4884	266	-107	1625	108	- 2	492	3299	-1030	13690	86873	- 114	39849	1979	2322	2375	2366
2	731	-11	4870	245	-177	1762	118	- 50	795	3335	-3584	23410	9495	-4205	64300	1451	2027	1388	2170
4	715	- 9	4686	269	-122	1772	206	- 45	864	6562	-3869	25987	16616	-3696	67830	1613	1220	1442	1611
5	709	47	4782	230	-162	1912	42	- 18	358	1237	-1469	12668	3309	-1134	28543	1308	1516	1606	812
6	716	10	4691	185	60	1193	52	70	432	1164	1813	9239	4155	5535	33790	1287	1467	1739	172
7	710	7	4697	257	- 56	1577	242	274	1689	6555	5551	42740	19136	21620	132842	1740	1794	1508	1230
8	702	-39	4651	213	-130	1498	316	160	1875	7078	- 358	43705	24569	11196	144114	1606	1898	1376	849
9	710	78	4620	248	55	1698	175	305	1435	5370	8956	42193	14150	25091	113248	1598	998	1702	1227
13	740	27	4920	245	- 70	1744	187	259	1410	4869	5806	37713	15481	21854	117132	1917	1664	1621	2515
15	715	53	4755	183	- 20	1163	89	- 46	770	1702	- 781	13682	6974	- 19	60313	1768	1763	1383	2344
16	724	36	4683	172	45	1082	156	250	1294	2963	5682	25770	12463	20506	102520	1741	1671	1671	1492
17	740	19	4913	238	-150	1795	194	99	1154	4808	- 104	31871	16015	8426	94948	1960	1921	1499	2378
18	741	-10	4898	196	66	1340	219	- 62	1349	4670	- 4	28550	18022	-5284	110353	1760	786	1146	884
19	732	-29	4875	285	36	1986	351	369	2068	11242	13251	72857	28374	29019	165852	1844	2079	1627	1735
20	733	3	4936	328	156	2433	443	198	2776	16490	15506	117757	36052	16473	228294	1766	1886	989	1634
21	702	15	4506	197	44	1316	182	258	1250	4205	6464	29803	14230	19835	94631	1305	1584	1758	2415
23	717	-29	4721	202	45	1283	273	248	1576	6419	6387	37248	21673	18677	144183	1761	1739	1497	1992
24	714	-43	4741	226	78	1640	307	175	1841	7946	7052	54345	24290	12461	143937	1386	1510	1401	2013
25	771	-24	5160	825	-269	5828	375	198	2265	33482	8085	214060	32028	16049	194343	578	65	1179	175
26	756	-36	4994	809	116	5691	440	34	2930	39742	8945	278039	36959	1065	243504	-233	92	624	128
27	726	7	4699	523	91	3253	292	241	1868	17031	15650	107261	23568	19185	147230	945	1037	1475	939

Legend: # 1 Florence, '39, Normal
2 " " Late
3 " '40 Normal
4 " " Late
5 " '41 Normal
6 " " Late
7 Griffin '39 Normal
8 " " Late
9 " '40 Normal
10 " " Late
11 " '41 Normal
12 " " Late
13 Stoneville '39 Normal
14 " " Late

#16 Stoneville, '40, Late
17 " '41 Normal
18 " " Late
19 Marianna, '39 Normal
20 " " Late
21 " '40 Normal
22 " " Late
23 " '41 Normal
24 " " Late
25 Lawton '39 Normal
26 " " Late
27 " '40 Normal
28 " " Late
29 " '41 Normal

E = Evaporation
T = Temperature
SM = Soil Moisture Tension
t = Week from planting date
Y_1 = Predicted yield regression #1
Y_2 = " " " #2
Y_3 = " " " #3
Y = Actual Yield Stoneville cotton

Standard Error Estimation:

Regression # 1	753.9
Regression # 2	718.2
Regression # 3	815.1

Table 8: Comparison of Various Combinations of Growth Periods by Three Scoring Methods.

Code:
S = Squaring
B = Blooming
SB = Setting Young Bolls
O = Opening

Scoring Methods

Period Omitted	1	2	3
S	.859	.765	.731
B	.758	.478	.666
SB	.817	.752	.715
O	.850	.765	.714
None	.829	.776	.723

The table shows linear correlations between yield and score using Stoneville variety, n = 19.

Rank of each combination according to least correlation; i.e., greatest loss by omission.

	1	2	3	Av. Rank
S	4	3.5	4	3.8
B	1	1	1	1
SB	2	2	3	2.3
O	3	3.5	2	2.8

Table 9: Code for Punch Card Data

Soil-Weather Project — Southern Cooperative Group

There is to be a card punched for each day's weather record at each station for the entire growing season. It has been suggested that the weather record be commenced two weeks preceding planting. All these variables will not be measured daily, but there should be provision for recording them on the card when they are taken.

		Columns
I. Identification:		
Station		1
Year		2–3
Month		4–5
Date		6–7
II. Climate:		
Maximum temp.		8–10
Minimum temp.		11–13
Mean temp.		14–16
Relative humidity, %		17–18
Precipitation (measure to hundredths)[1]		19–22
Intensity (am't/unit time)		23–25
Pyrheliometer 8 a.m.	÷ 100	26–28
12 noon		29–31
4 p.m.		32–34
Photo-tube integrator 8 a.m.	$\left(\frac{\text{difference}}{10}\right)$	35–37
Wind (mph. at 2 p.m.)[2]		38–39
24-hr. wind movement		40–41
III. Soil		
Temperature (2" depth; continuous record; range 0–150°F.)		
Maximum		42–44
Minimum		45–47
Temperature (8" depth; same as above)		
Maximum		48–50
Minimum		51–53
Moisture (transform to "tension" in cm. of H_2O; range 100 – 100,000[3]		
4" depth	(÷ 100)	54–57
8" depth		58–61
14" depth		62–65

---

[1] Punch "0000" for "no rain"; punch "9999" if it rained but the amount was not recorded (missing data).

[2] Punch "00" for no measureable wind; punch "99" if the reading was not taken (missing data).

[3] Punch "9999" for missing data.

FIGURE 1

GRAPHS COMPUTED FROM REGRESSION # 1

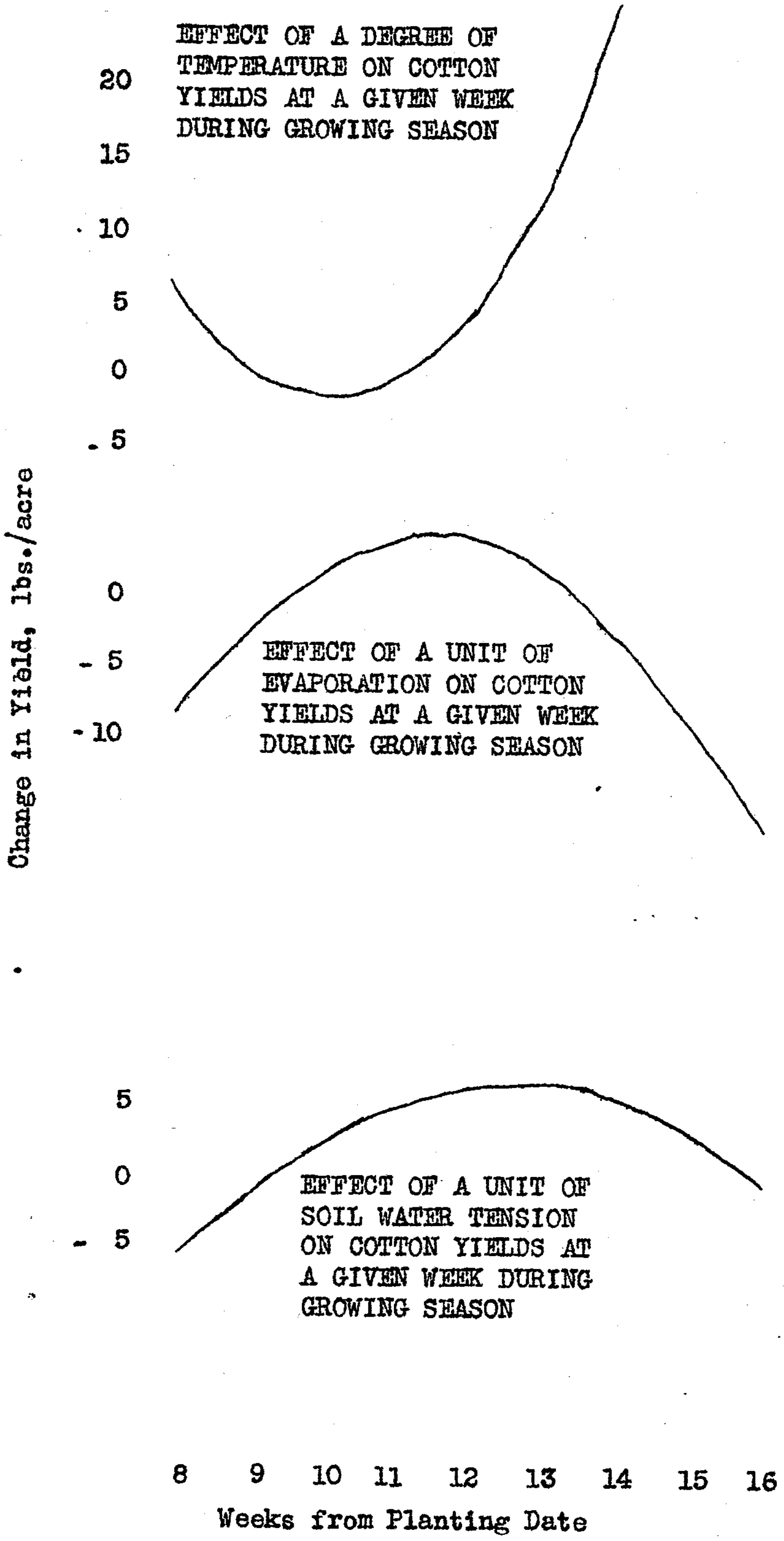
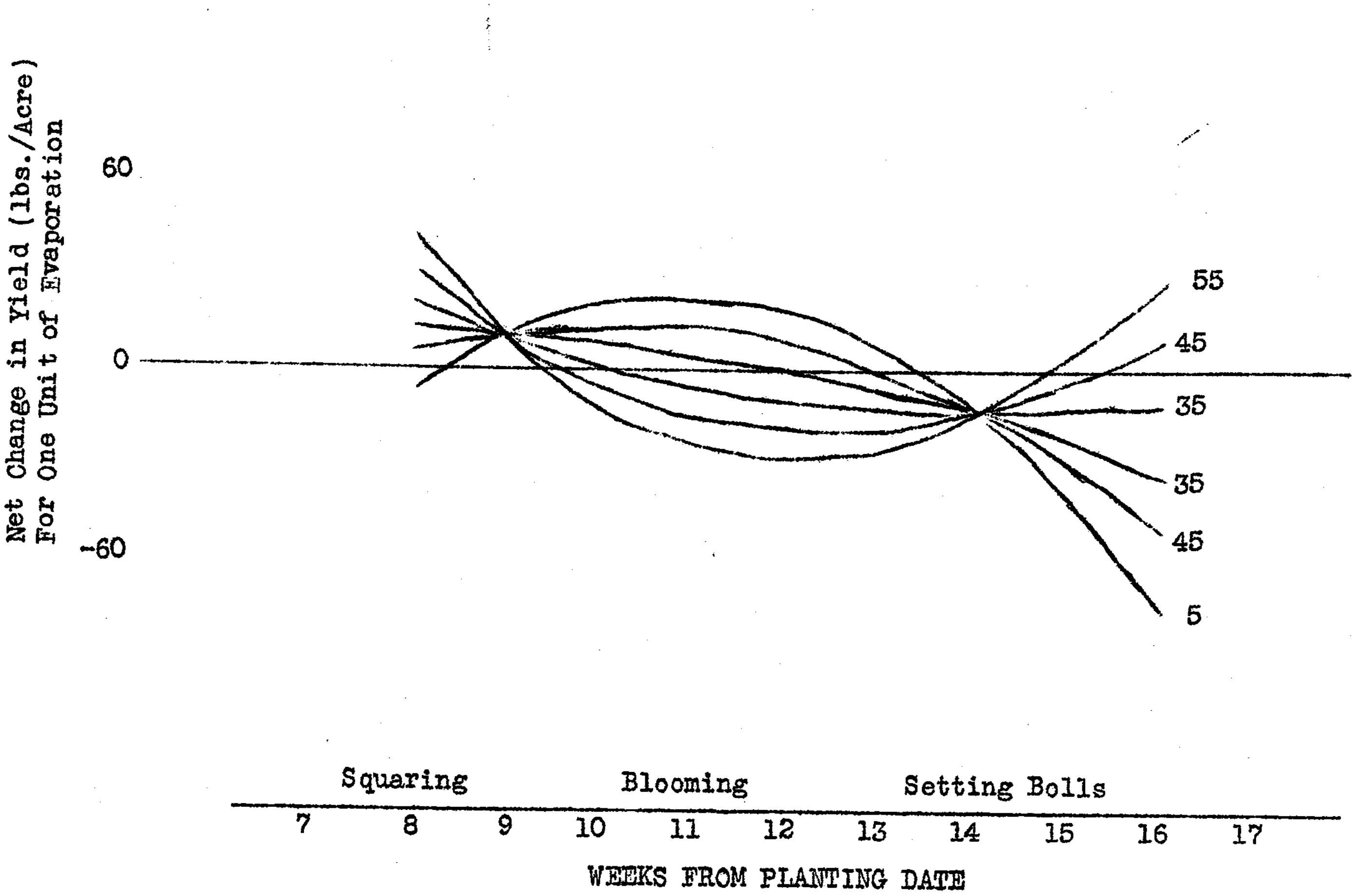
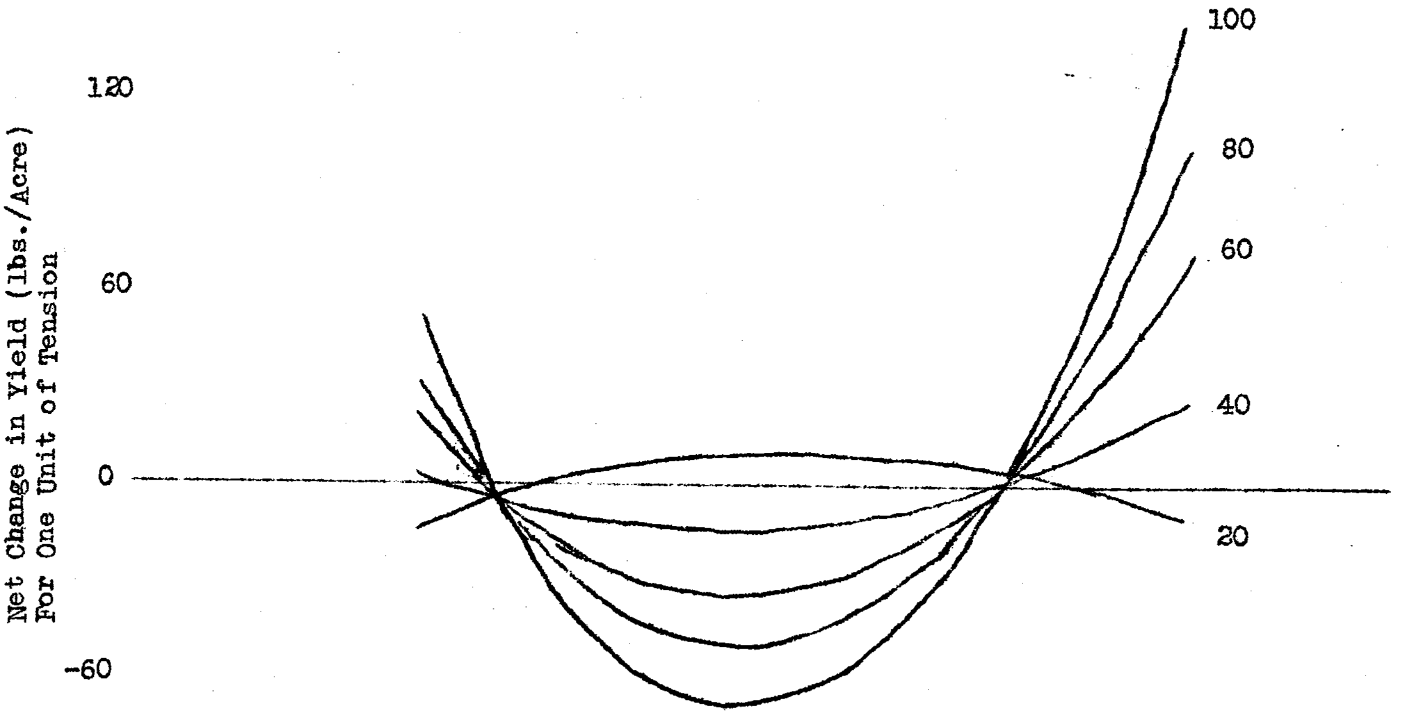


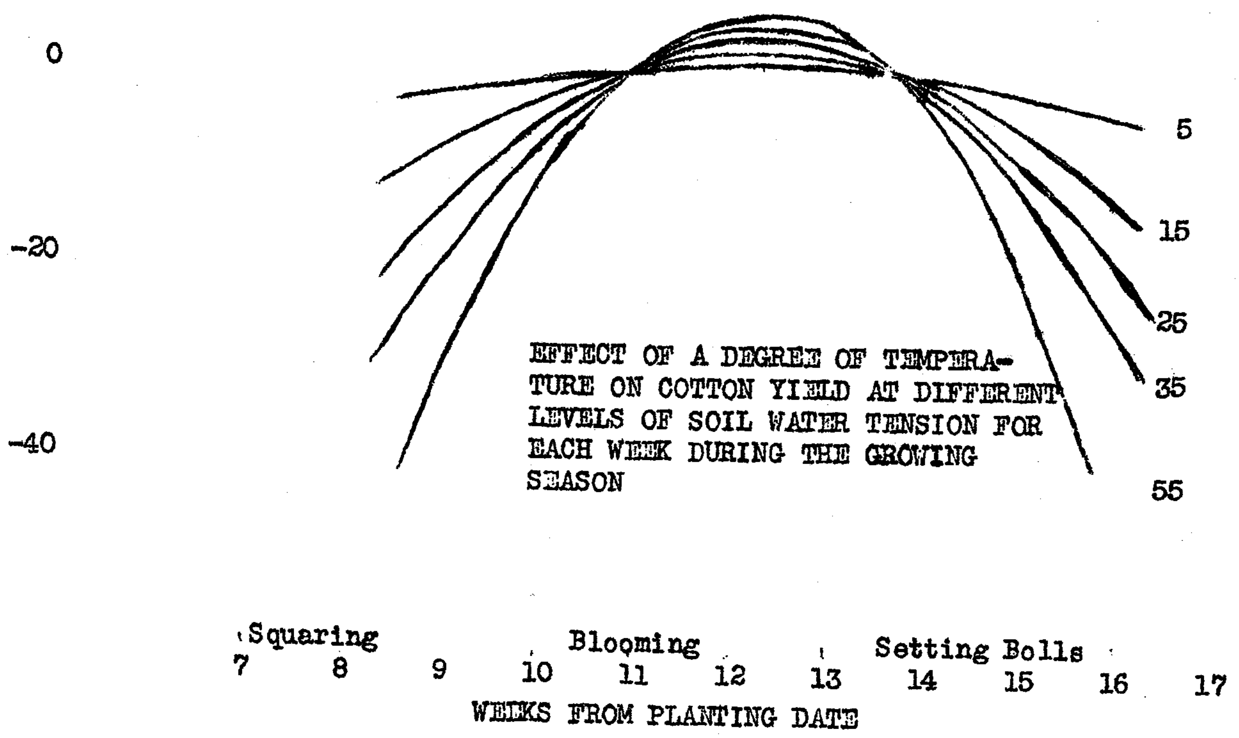
Figure #2: EFFECT OF A UNIT OF EVAPORATION ON COTTON YIELD AT DIFFERENT LEVELS OF SOIL WATER TENSION FOR EACH WEEK DURING GROWING SEASON



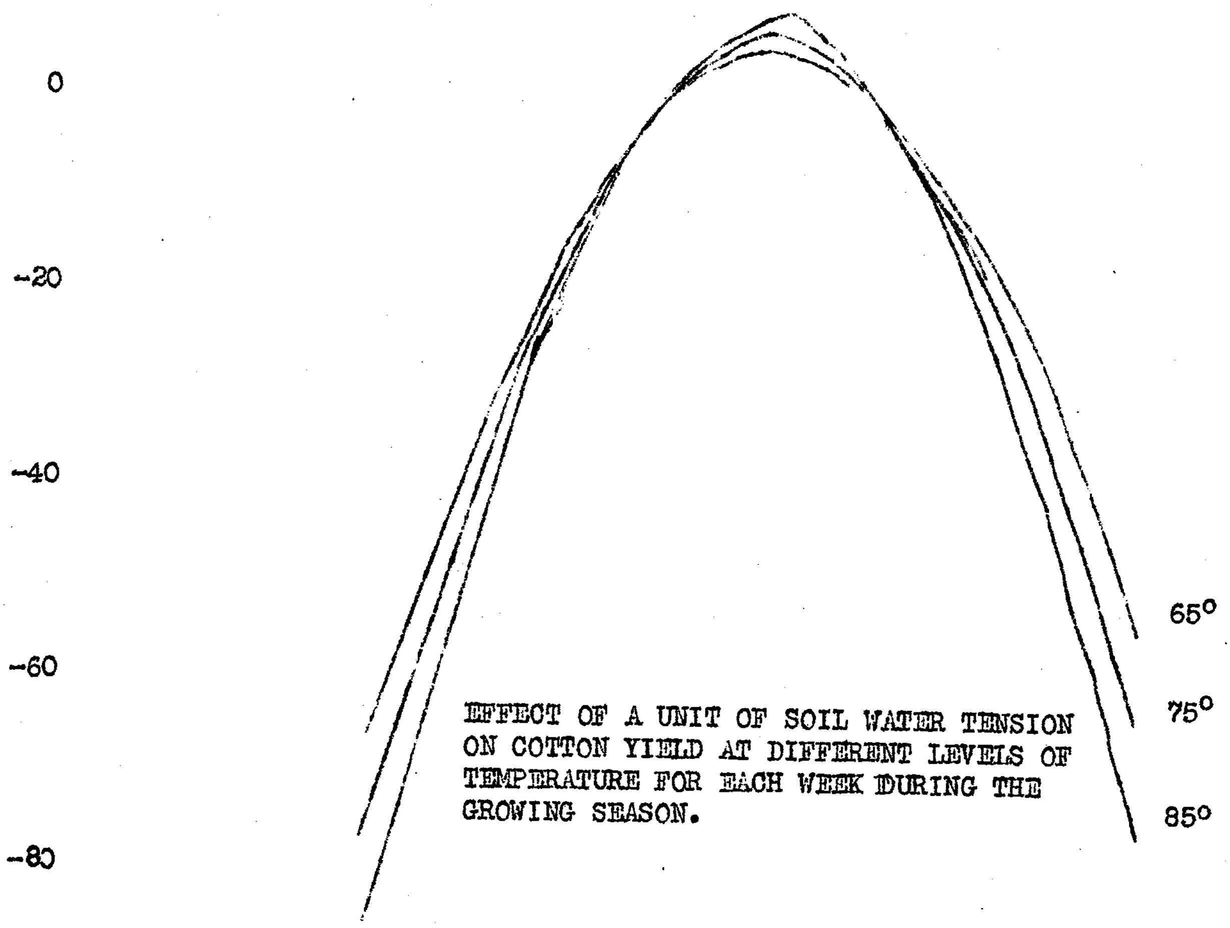


EFFECTS OF A UNIT OF SOIL WATER TENSION OF COTTON YIELD AT DIFFERENT LEVELS OF EVAPORATION FOR EACH WEEK DURING THE GROWING SEASON

Figure #3. GRAPHS COMPUTED FROM REGRESSION #3



EFFECT OF A DEGREE OF TEMPERATURE ON COTTON YIELD AT DIFFERENT LEVELS OF SOIL WATER TENSION FOR EACH WEEK DURING THE GROWING SEASON



EFFECT OF A UNIT OF SOIL WATER TENSION ON COTTON YIELD AT DIFFERENT LEVELS OF TEMPERATURE FOR EACH WEEK DURING THE GROWING SEASON.

Figure #4 SQUARING Weeks 6-9

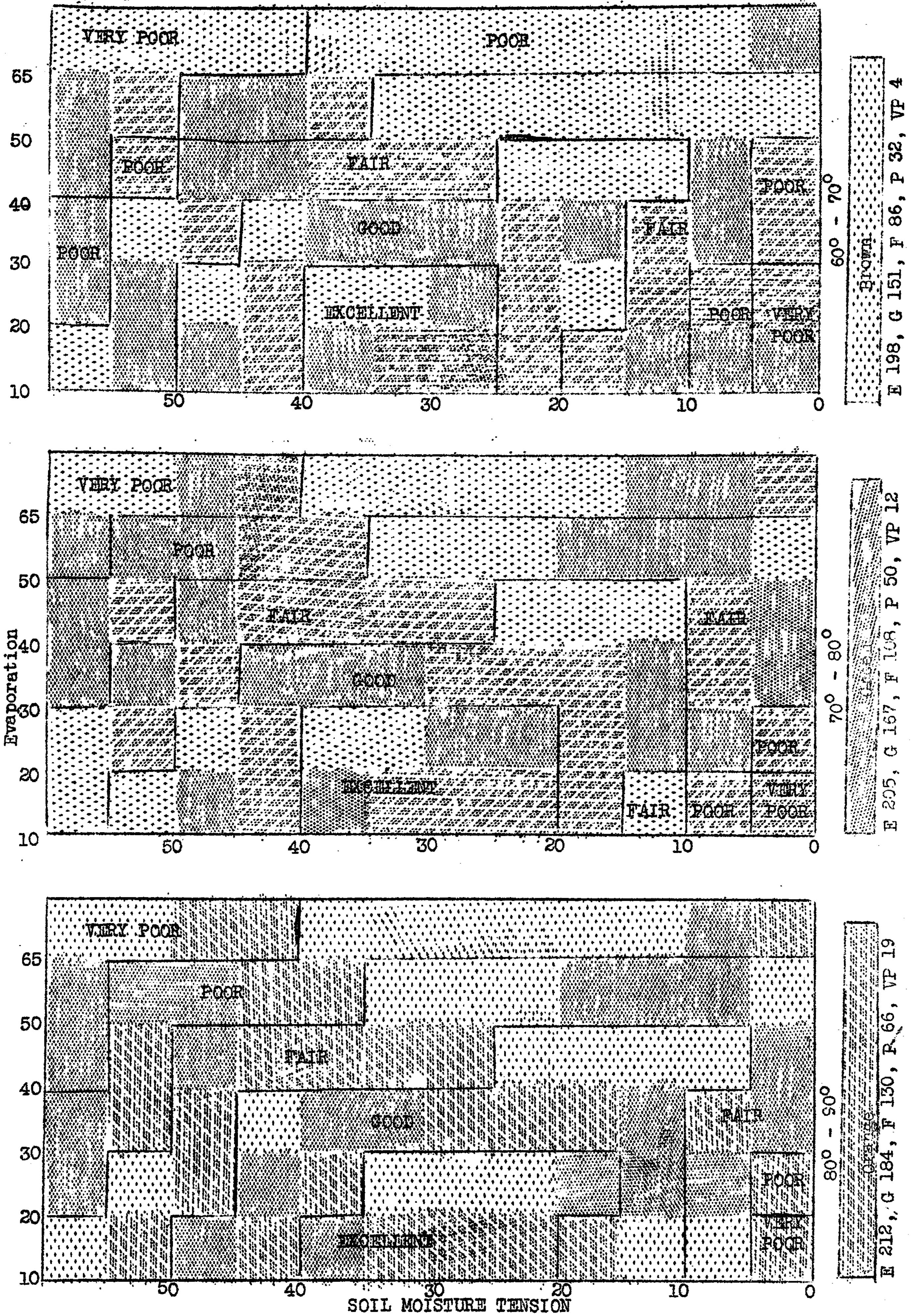


Figure #4. Cont'd. BLOOMING Weeks 10-12

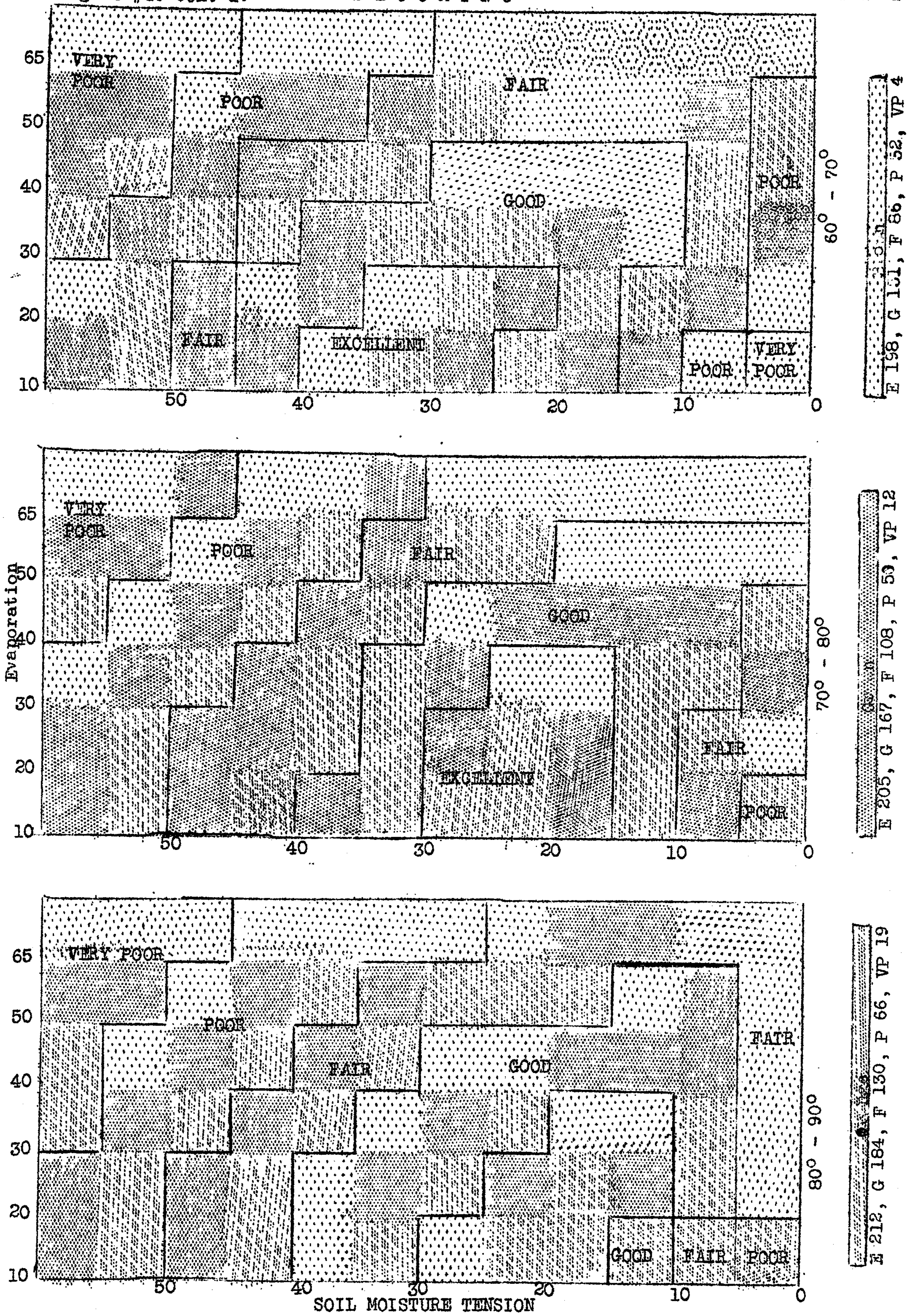


Figure #4 Con'd.

SETTING BOLLS

Weeks 13-16

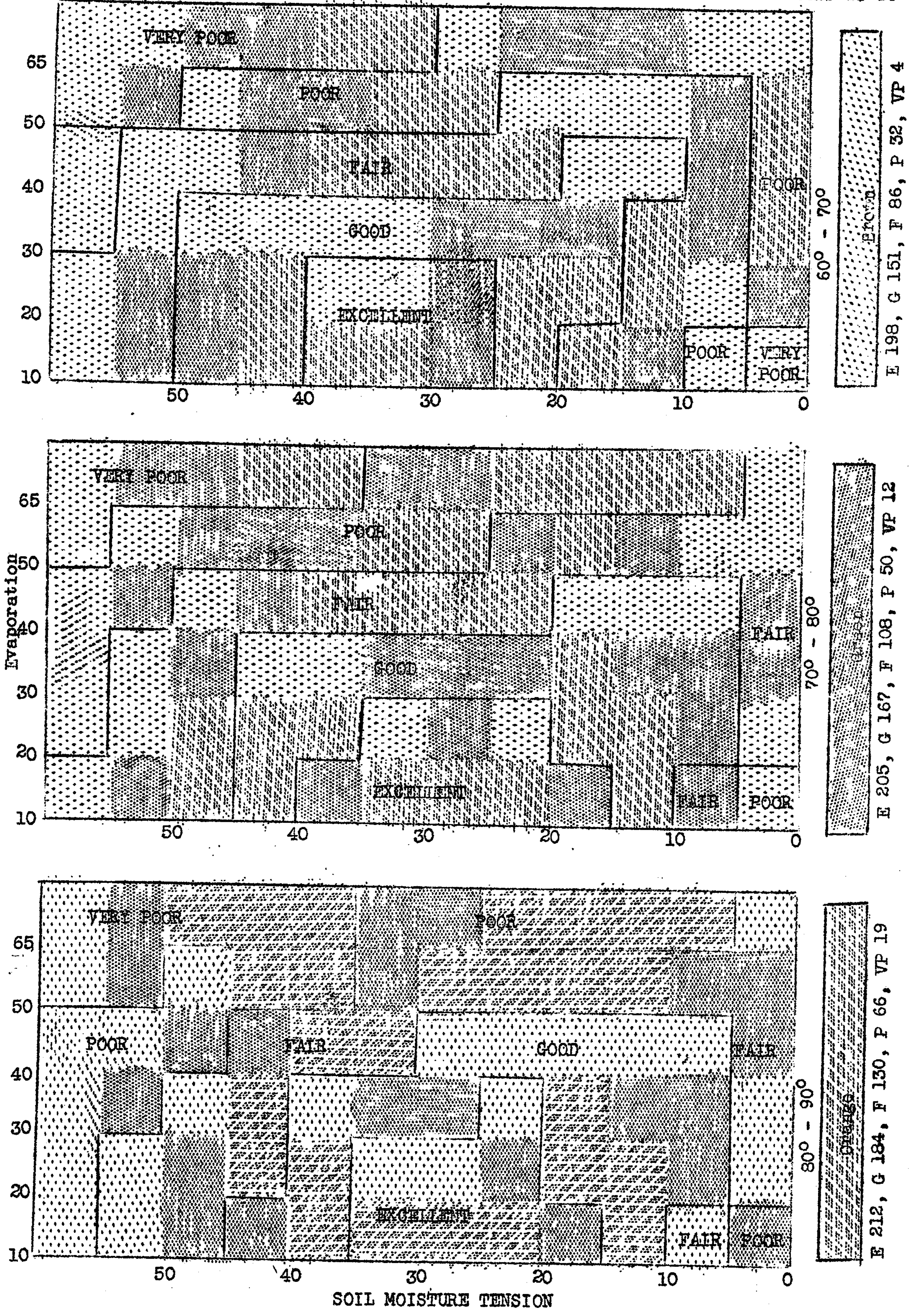


Figure #4. Cont'd. OPENING Weeks 16-19

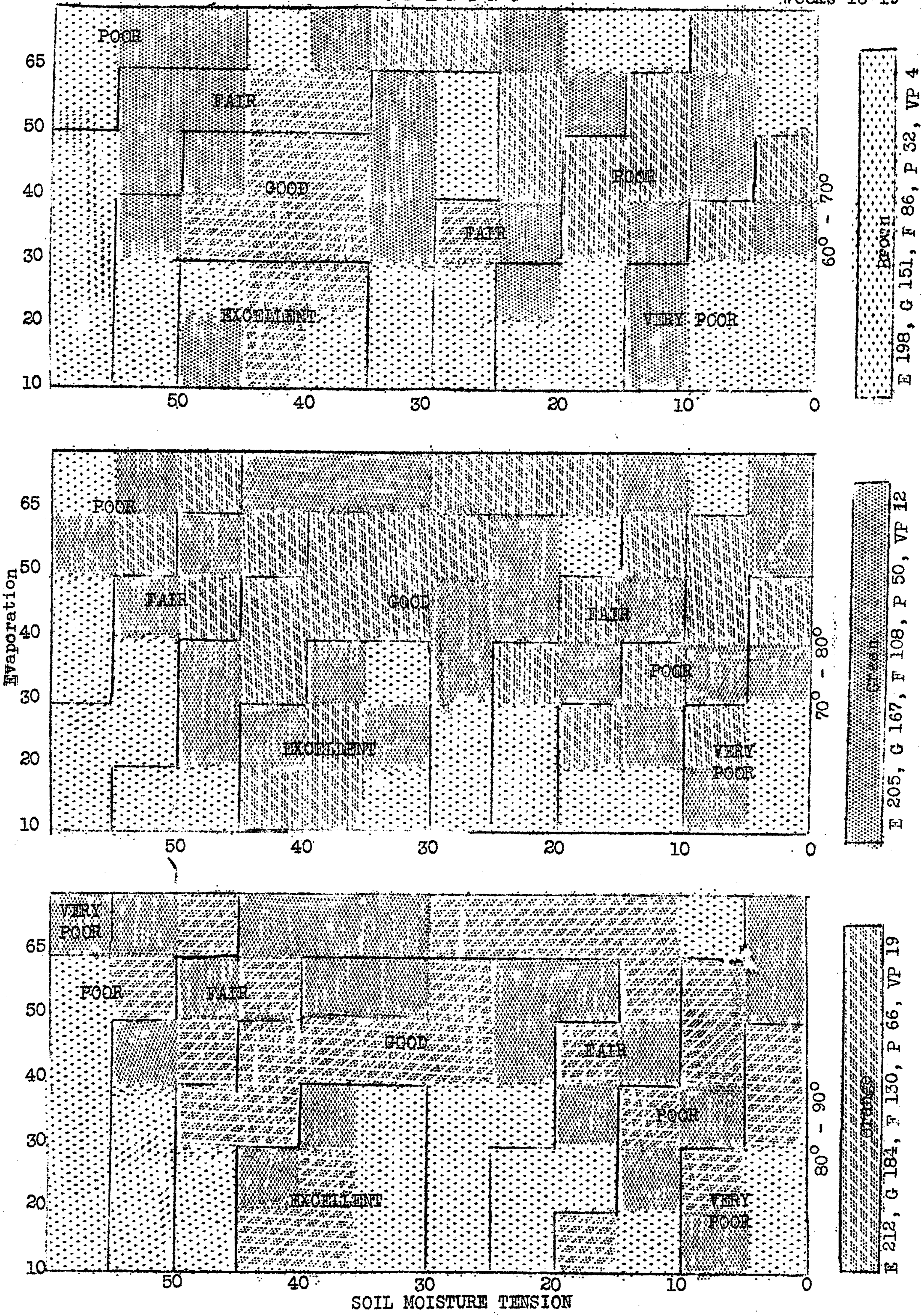


Figure #5 SCATTER DIAGRAMS: SCORE VS. YIELD 10-19 Weeks

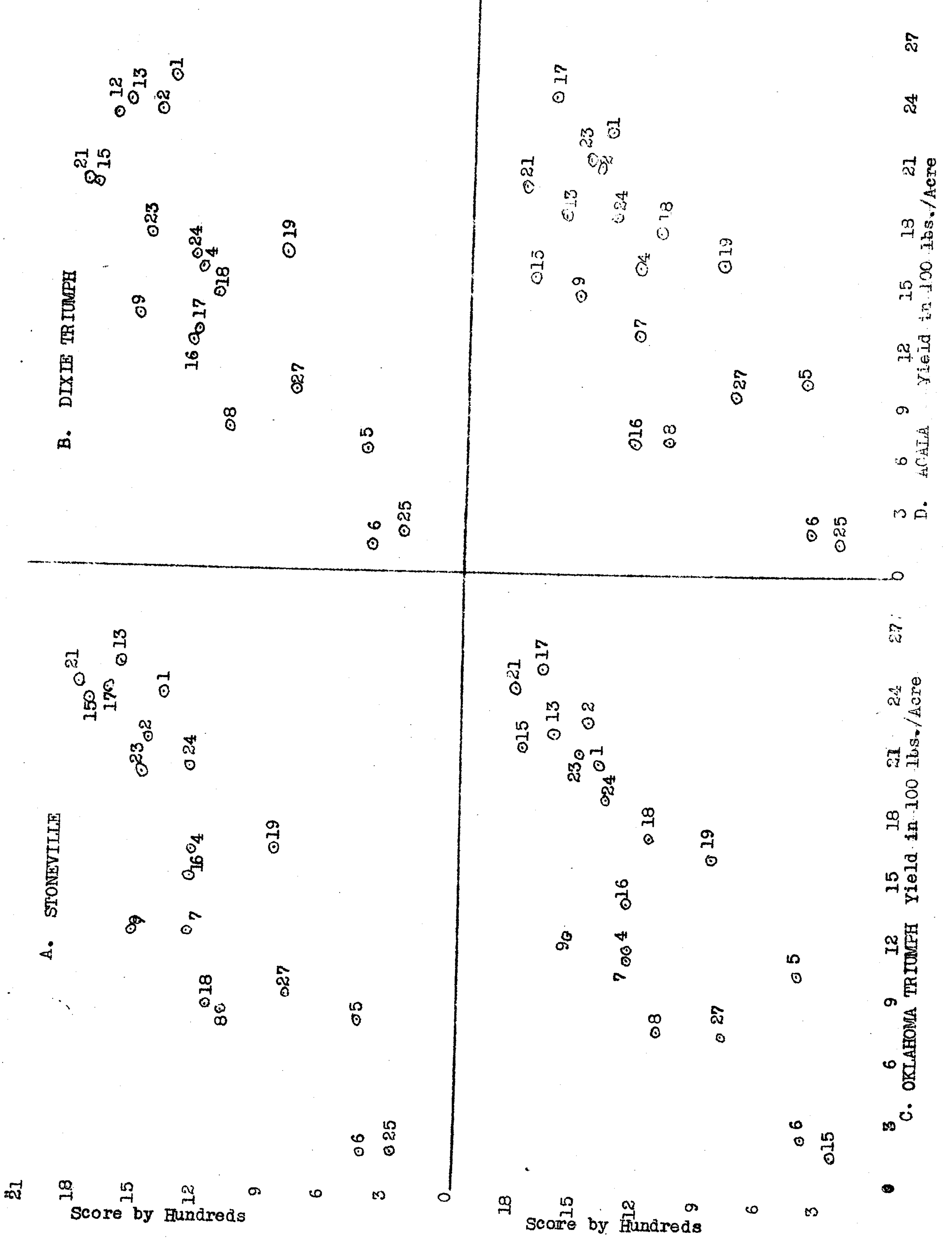


Figure #6

PREDICTION CURVES

10-19 Weeks

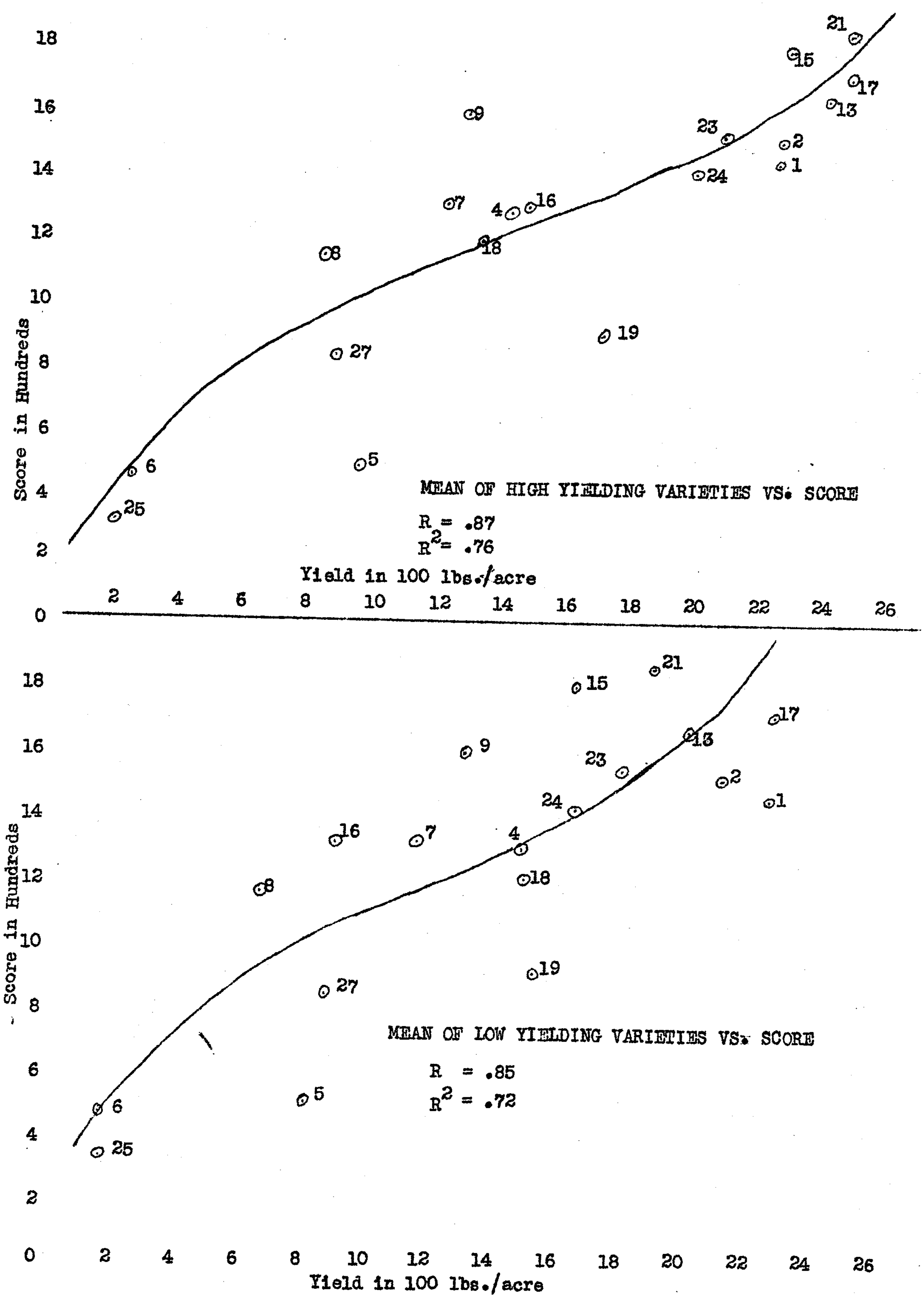


Figure #7 SCATTER DIAGRAMS: SCORE VS. YIELD 6-16 Weeks

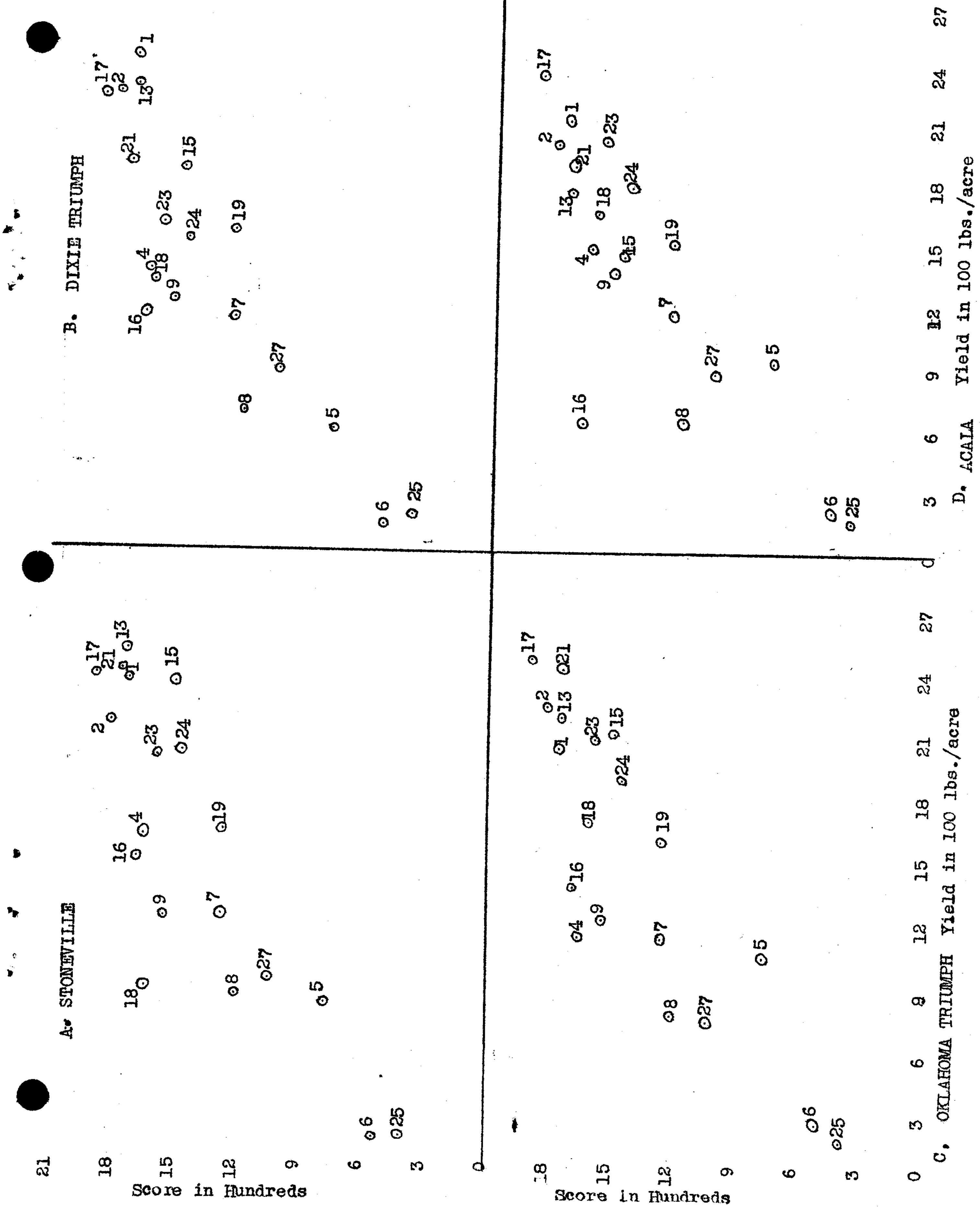
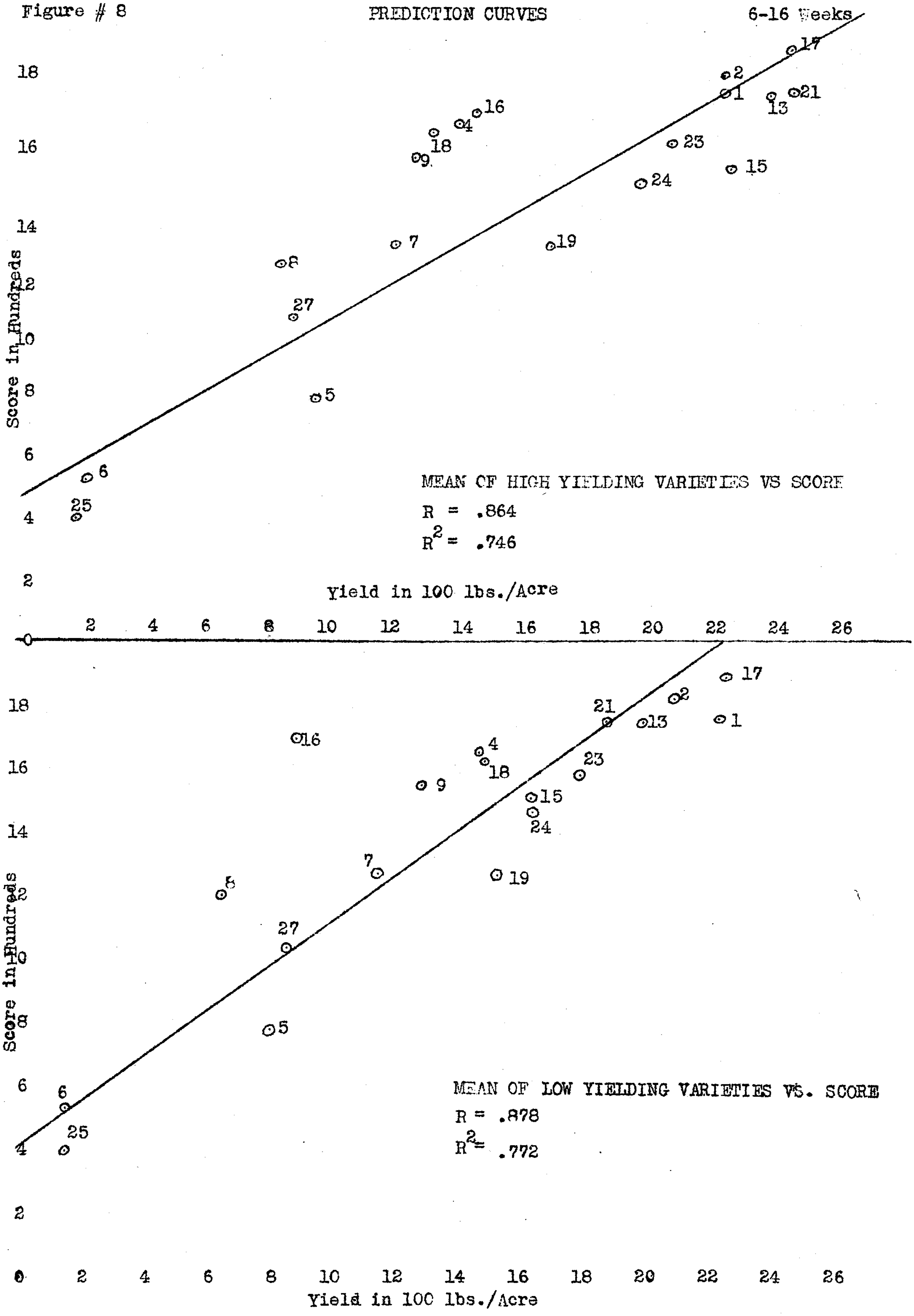


Figure # 8 PREDICTION CURVES 6-16 Weeks



MEAN OF HIGH YIELDING VARIETIES VS SCORE
R = .864
R^2 = .746

MEAN OF LOW YIELDING VARIETIES VS. SCORE
R = .878
R^2 = .772