Objective Forecasts of Cotton Yield

By Walter A. Hendricks and Harold F. Huddleston

An earlier paper[1] on this subject summarized results of studies on the 1954 crop in 10 Southern States. At that time three approaches to forecasting yield from plant observations were considered: (1) the multiple regression approach, (2) a "probability of survival" model, and and (3) an empirical approach. This paper describes a forecasting procedure which developed from these earlier studies as more data became available and as an appropriate way of making use of the fruiting habits of the cotton plant in a forecasting formula was better understood. The most noteworthy advance was the discovery of a simple device for estimating fruiting rate from a maturity classification of plants on August 1. This is of major importance for the August forecast because all of the fruit contributing to final yield is not yet formed by that date. After September 1, additional fruiting is no longer an important factor for the Southern region as a whole. This new approach is being used experimentally in 1956.

Another paper in this issue by Jack Fleischer reports on additional work being conducted on the "probability of survival" model.

COUNTS OF COTTON FRUIT during 1954 and 1955, and the corresponding yields derived from data on ginnings, provide the basis for the yield forecasting procedure described in this paper. That procedure makes use of known fruiting characteristics of the cotton plant and permits plant observations collected during the growing season to be translated into indications of final yield in logical fashion.

Data collected in 1954 and 1955 show that, for all practical purposes, a combined count of blooms, small bolls, large bolls, and open bolls as of September 1 represents the total yield potential for the season. Multiplying that count for small sampling units in sample fields by the weight of seed cotton per boll (as estimated from open cotton picked in the sample fields), noting that 37 percent of the seed cotton is lint, and multiplying by the appropriate expansion factor to convert the result to a pounds-per-acre level, gives that potential in terms of pounds of lint per acre.

The only unknown in the forecast as of September 1, or on succeeding dates, is the fraction of that potential that will go to the gin. Some fruit will fail to mature and some open cotton is missed in harvesting. In 1954, 9 percent of the September 1 potential was lost. In 1955 the loss was 9 percent. In both years almost exactly half of the loss was in the form of fruit that failed to mature; the other half was open cotton found in the fields after harvest which either was missed in picking or had opened after the farmer harvested the crop. Although 1954 was a dry season and 1955 was wet, growing conditions were not sufficiently adverse in either year to cause any unusual damage.

It appears that a loss of about 10 percent of the blooms and bolls present on September 1 can be regarded as a "normal" situation. Applying that deduction to the yield potential represented by the September 1 bloom and boll count and the observed weight of cotton per boll should provide a good forecast of yield as of that date.

For an August 1 forecast the situation is not so simple because a fruit count alone on that date does not tell the whole story. Plants do not yet have their full set of fruit and it is necessary to forecast the number of blooms and bolls still to be formed. Analysis of the 1954 and 1955 data indicates that the additional blooms and bolls that will appear between August 1 and September 1 can be forecast from observations on the stage of maturity of the plants as of August 1. From the maturity classification, it is possible to compute the rate at which plants are fruiting; that fruiting rate can be translated into a forecast of the blooms and bolls that will be added between August 1 and September 1.

On August 1 the weight of cotton per boll must also be predicted because little cotton is open by that date in most of the region. But, it appears likely that mature boll size is related to the fruiting potential already indicated by the August 1

[1] HENDRICKS, WALTER A., AND HUDDLESTON, HAROLD F. A FOUNDATION FOR OBJECTIVE FORECASTS OF COTTON YIELDS. Agricultural Economics Research. 7:108–111. 1955.

Reprinted from Agricultural Economics Research, Vol. IX, No. 1, January 1957 by the Standards and Research Division, Statistical Reporting Service.

plant observations. In 1954 the season was dry and the fruiting potential indicated by the August 1 plant observations was considerably lower than in 1955 when rainfall was more plentiful. The weight of seed cotton per boll was also lower. It appears logical that weather factors that are favorable for fruiting will also give rise to larger bolls, so that a positive correlation should exist between fruiting potential and boll size.

Forecasting procedures based on these observed relationships are being tested on the 1956 crop month by month as the season progresses. The procedures followed and the data on which they are based are described in the following sections.

The August 1 Forecast

The August 1 fruit counts in 200 sample fields for 1954, and in 400 sample fields for 1955, are shown in table 1.

Table 1.—*August 1 fruit counts per 10 feet of row*

Kind of fruit	1954	1955
	Number	*Number*
Squares	78.5	84.9
Blooms and small bolls	28.7	23.5
Large bolls	22.6	20.4
Total	129.8	128.8

In these counts a boll less than 1 inch in diameter was called a small boll. All other bolls, including those that had already opened, were called large bolls. The bloom count was combined with the small-boll count for purposes of analysis because a bloom lasts only a few days before becoming a small boll.

Although the total fruit count was almost identical in the 2 years, more fruit was in the early stages of development in 1955 because the crop got off to a later start that year. This must be taken into consideration along with the total fruit count in arriving at an explanation of the differences in final yields for the 2 years. The maturity factor can be introduced into the picture most conveniently by classifying all sample fields into three categories according to the kind of fruit found on the plants in the sample-row segments. That classification is shown in table 2.

Table 2.—*Classification of fields by maturity of plants on August 1*

Kind of fruit present	Fields 1954	Fields 1955
	Percent	*Percent*
Squares only	6.5	24.9
Squares, blooms, and small bolls	18.4	21.3
Large bolls	75.1	53.8
Total	100.0	100.0

Table 2 shows that in 1954 75.1 percent of the fields had reached the large-boll stage on August 1, but in 1955 only 53.8 percent were in that category. This is significant because the cotton plant has about the maximum fruit load it can carry by the time some of its fruit reaches the large-boll stage. If the total quantity of fruit in all categories on the cotton plant on any date is plotted against time, the resulting chart follows a sigmoid growth curve.

Fruiting increases rapidly during the first 3 weeks after squares begin to appear. About 3 weeks after the first squares are formed, blooms and small bolls begin to appear and the plant is adding fruit at its maximum rate. From that time on, the rate at which fruit is added begins to diminish and continues to decline for the next 3 weeks until large bolls begin to appear. When that stage is reached, the total quantity of fruit on the plant shows little or no further increase. As large bolls begin to appear about 6 weeks after squaring starts, the growth curve may be divided into clearly recognizable portions which can be related to the observable stage of maturity of the plants.

If the maximum fruit load is represented by A, that maximum is reached about 6 weeks after squaring begins. Because the growth curve is approximately symmetrical, the plant has around half its maximum load about 3 weeks after squaring starts; this is the stage at which blooms and small bolls begin to appear.

All plants in the first category of table 2 must be in a stage of maturity corresponding to a fruit load ranging anywhere from zero to A/2, depending upon whether squares are just beginning to appear or whether the plants already have squares that are almost ready to burst into bloom. The average plant in that category should have been squaring for about 1.5 weeks and should have about a fourth of its ultimate maximum load.

By the same token, plants that show squares, blooms, and small bolls, but no large bolls, should range from those on which blooms are just beginning to appear to those on which some small bolls are almost ready to graduate to the large-boll category. These plants will have been fruiting for 3 weeks to 6 weeks, with an average of 4.5 weeks. The fruit load on an average plant in that category should be halfway between $A/2$ and A or $3A/4$. The average plant in the second category in table 2 thus has about three-fourths of its total final load.

These characteristics of the cotton plant make it possible on August 1 to predict the total quantity of fruit that will be on the average plant by the time it has its total final load. In terms of A, the quantity of fruit on the average plant as of August 1, 1954 is $0.065(A/4)+0.184(3A/4)+0.751(A)=0.905A$. This means that the average plant on August 1, 1954 was carrying 90.5 percent of its total potential load.

Reference to table 1 shows that the total quantity of fruit in all categories on the average plant as of August 1, 1954 was about 130 units. This figure, of course, is in terms of fruit per 10 feet of row. It is clear from what has been said that $0.905A=130$ or $A=144$. This means that the total potential fruit load, as predicted from the August 1 data, is 144 units of fruit.

Similar computations may be made for the August 1, 1955 data. In terms of the maximum potential fruit load, the quantity of fruit already present on August 1 is $0.249(A/4)+0.213(3A/4)+0.538A=0.760A$. This means that on August 1, 1955 the average plant was carrying only 76 percent of its total potential load as compared with 90.5 percent in 1954.

Table 1 shows that on August 1, 1955 the average plant was carrying about 129 units of fruit, expressed in terms of 10 feet of row. The estimated total load is given by $0.760A=129$ or $A=170$. This shows quite clearly that on August 1, 1955 the average cotton plant already showed a much higher fruiting potential than was the case on August 1, 1954—170 as compared with 144.

But the main forecasting problem on August 1 refers to estimating the rate at which bolls are being formed on that date. An approximation to that rate can be derived by noting that the total quantity of fruit on the average plant increases from zero to A in about 6 weeks. The increase during the first 3 weeks is about the same as during the last 3 weeks. The average weekly rate of fruiting during the entire 6-week period is thus approximately $A/6$.

The 1954 data indicate that on August 1 the rate of fruiting on the average plant was approximately $0.065\,(A/6)+0.184\,(A/6)+0.751\,(0)=0.0415A=0.0415(144)$ or 6 units per week. On August 1, 1955, the corresponding average rate was $0.249\,(A/6)+0.213\,(A/6)+0.538\,(0)=0.0770A=0.0770(170)$ or 13.1 units of fruit per week. This means that on August 1, 1955 fruit was being added to the average plant at a rate 13.1/6.0 or 2.18 times as fast as on August 1, 1954. Table 3 shows that in 1954 17.4 bolls were added to 10 feet of row between August 1 and September 1. Therefore in 1955 the increase should have been (2.18) (17.4) or 37.9 bolls. The increase which actually took place was 39.3 bolls.

TABLE 3.—*Bloom and boll counts per 10 feet of row as of August 1 and September 1*

Kind of Fruit	1954		1955	
	Aug. 1	Sept. 1	Aug. 1	Sept. 1
	Number	*Number*	*Number*	*Number*
Blooms and small bolls	28.7	12.1	23.5	24.7
Large bolls	22.6	56.6	20.4	58.5
Total	51.3	68.7	43.9	83.2
Change from Aug. 1 to Sept. 1	+17.4		+39.3	

In these computations, bloom counts were combined with boll counts so that a bloom was counted as a boll. Apparently, data available as of August 1 can be used to forecast the number of bolls that will be found on September 1 when rate of fruiting is brought into the picture to supplement the August 1 fruit count. Rate of fruiting can be predicted from data reflecting the stage of maturity of the crop.

To make a forecast of the 1956 crop on August 1, it was convenient to compute a conversion factor for translating the weekly fruiting rate derived from the maturity classification into a forecast of the blooms and bolls to be formed between Augu

1 and September 1. Plotting the data for 1954 and 1955 on a chart indicated an almost perfect proportional relationship between the two variables. The slope of the proportional line is 2.97. In other words, the number of blooms and bolls still to be formed after August 1 is about 2.97 times the weekly fruiting rate computed from the maturity classification.

The August 1, 1956 fruit counts were as follows:

	Aug. 1 count per 10 ft. of row
Squares	85.5
Blooms and small bolls	29.1
Large bolls	21.3
Total	135.9

The classification of fields according to August 1 maturity stage was:

	Percent of fields
Squares only	8.1
Blooms and small bolls	31.0
Large bolls	60.9
Total	100.0

The estimated percent of a full load carried by the average plant on August 1 was (0.081) (25) + (0.310) (75) + (0.609) (100) = 86.2. The estimated total potential fruit load was 135.9/0.862 = 158. Assuming, as before, that 39.1 percent of the plants had a weekly fruiting rate of 1/6 of that quantity and that 60.9 percent were no longer adding fruit, the weekly fruiting rate of the average plant on August 1, 1956 was $\frac{(0.391)(158)}{6} = 10.3$. The number of blooms and bolls to be formed between August 1 and September 1 was forecast at (2.97) (10.3) = 30.6. The total number of blooms and bolls one would expect to find on an average 10 feet of row as of September 1, 1956 would then be 29.1 + 21.3 + 30.6 = 81.0; past experience indicates that about 90 percent of those will represent ginned cotton.

The quantity of cotton per boll must also be predicted for the August forecast because no actual weights are available until a month later. In 1954 and 1955 the average weight of seed cotton per boll was as follows, in relation to the fruiting potential per 10 feet of row computed from the August 1 data:

	Max. fruit load, number	*Seed cotton per boll, grams*
1954	144	4.62
1955	170	5.03

As the maximum fruit load in 1956 was computed to be 155 units, the average weight of seed cotton per boll should be about 4.84 grams. Applying that weight to the 81.0 bolls per 10 feet of row and multiplying by the appropriate expansion factor gives a forecast of 426 pounds of lint per acre. Assuming a normal loss of 10 percent, the yield forecast in terms of cotton to be ginned is 383 pounds per acre.

These computations seem reasonable enough. But some inaccuracy is introduced into forecast of fruit to be formed between August 1 and September 1 by an implicit assumption that plants in an early stage of fruiting by August 1 and plants already more mature will have the same total fruit load in any one season. In dealing with a region as large as this, fields in an advanced stage of maturity on August 1 may be found where yields tend to differ decidedly from those less mature. In other words, there is a spatial correlation between stage of maturity and yielding ability.

Such a disturbing association was found even within Texas. Cotton in the State is generally much farther along by August 1 in lower areas than in the High Plains. Yields in the two areas also tend to differ. Pooling all data for the State, and performing the same computations as shown above, leads to an average yield forecast of 222 pounds of lint per acre. Separate computations for lower Texas and for the High Plains give separate forecasts of 211 and 288 pounds per acre. As about half the cotton acreage in the State lies in each of the two areas, the average for the State computed on a stratified basis is 250 pounds per acre as compared with 222 on the other basis.

Because of such differences, yield forecasts were computed separately for each State in the 10-State region. The average of the State estimates was 357 pounds per acre for the region, which agrees extremely well with the August 1 forecast made by the Crop Reporting Board. It also differs appreciably from the forecast of 383 pounds per acre computed from the pooled data for the region as a whole. This is convincing evidence of need for stratification, at least by States. In a State in which conditions are like those in Texas, within-State stratification is also desirable.

The September 1 and October 1 Forecasts

As pointed out earlier, the blooms and bolls already on the plants by September 1 appear to determine the yield. It is necessary only to apply the observed weight of cotton per boll to the combined bloom and boll count and to make necessary deductions for losses that will occur between the forecast date and harvest time. Under "normal" conditions the loss appears to be about 10 percent. That figure can be determined more precisely as experience is gained. At present, we have no data to appraise losses that could occur in a year when weevils, disease, or unusually destructive weather factors are serious. Under such conditions losses would be heavier. But until growing seasons of that kind are actually encountered, and experience is gained as to the effects of adverse conditions, there is no basis for an objective computation.

September 1 weights of seed cotton per boll appear to give reliable indications of the average weight for the season. As part of the sampling program, open cotton in the selected sampling units was picked to ascertain the weight of seed cotton per boll in 1954 and 1955. This was weighed in the field and a small quantity taken to the office and reweighed after being air dried. In 1954 cotton was weighed only at the time of the September 1 visit and on the post-harvest survey. The September 1 weight of seed cotton per boll on an air-dry basis was approximately 4.6 grams. The weight on the post-harvest survey was much lower, but has no particular bearing on yield forecasts and estimates. In 1955 all open cotton found in the sampling units was picked and weighed on every visit to the sample fields. Table 4 shows the cumulative average weight of seed cotton per boll for the entire period covered by the observations, except the post-harvest survey.

All open cotton weighed at the time of the August 1 visit was found in South Texas only. In that area the average size of boll was considerably smaller than that of the 10-State region as a whole. By September 1 open cotton was found in a greater part of the region and weights on that date, when combined with August 1 weights, yielded an average which differs only slightly from the final cumulative average for all cotton picked during the season. It appears that the September 1 cumulative average is adequate for yield forecasting purposes. But in view of the fact that bolls opening early in the season may be of different size than those opening later, such differences must be taken into account in years when boll size is related to the date of maturity.

TABLE 4.—*Cumulative average weight of seed cotton per boll in 1955*

Month	Field weight	Dry weight
	Grams	*Grams*
Aug. 1	4.692	4.668
Sept. 1	5.067	5.024
Oct. 1	5.554	5.027

October 1 fruit counts were made for the first time in 1955. On that date 4.2 blooms and small bolls and 75.3 large bolls were found per 10 feet of row. The combined count of 79.5 is lower than the 83.2 present on September 1. Burrs were counted as bolls. The quantity of cotton already picked by August 1 is negligible for the region as a whole and presents no difficulties. By September 1 an appreciable quantity has been harvested, but almost all of that is in South Texas and represents a complete picking of individual fields then being diverted to other uses.

When no observations could be made in such fields, the August 1 bloom and boll count was accepted as the boll count that would have been found at harvest. But by October 1 many fields in various parts of the region h ve been partially picked, leaving additional bolls to be picked as they mature. There is reason to believe that counts in some of these fields were too low, particularly where cotton was snapped by hand, because burrs are not left on the plants. Then, too, bolls may have been knocked from plants if mechanical pickers were used.

Although the discrepancy is not alarmingly large and some mortality can be expected in a month's time, there is reason to suspect that this count may be too low—particularly for the large bolls. Field observations on cotton are complicated by farm harvesting practices during the season; cotton is being harvested somewhere in the 10-State region during almost the entire period covered by observations on the sample fields. The procedure followed in making fruit counts under this project attempted to arrive at counts each month so as to include bolls that had been picked. In other words, counts were designed to correspond to total final production, not merely to the portion remaining for harvest as of any date.

This difficulty was not encountered in 1954 because no October 1 counts were made in that year.

It may not be serious. but it is a problem that will require additional study. One solution might be to count only bolls still on the plants at the time of each visit to the sample fields and to relate that count to cotton to be picked only after that date. Current ginning data up to that date could be added to the forecast of cotton still to be harvested, as computed from fruit still on the plants, to provide a forecast of the total crop. Such a procedure might not be satisfactory if there is any appreciable lag between the time the cotton is picked and when it is taken to the gin.

If such an approach is impractical, an alternative would be to devise a procedure for getting complete fruit counts on sample plants that are left undisturbed in farm harvesting operations. This would be difficult to put into practice on sample fields when cotton is not picked by hand. Even in fields picked by hand, there is no assurance that farmers would always leave designated sample plants undisturbed, although specifically requested to do so.

A third possibility would be to use fruit counts of plants upon which fruit had been tagged on an earlier visit. Disappearance of tagged fruit would provide an estimate of burrs and bolls lost during harvesting operations. Under this scheme farmers could be instructed to treat tagged plants the same way as those in the rest of the field. But the presence of tags might encourage some farmers to pass up the tagged plants under the impression that they were doing the sampler a favor, even though they had been assured previously that such plants required no special treatment.