The *Journal of Cotton Science 9:115–120 (2005)*
http://journal.cotton.org, © The Cotton Foundation 2005

ECONOMICS AND MARKETING

Effect of Harvest Season Rainfall on Cotton Yield

D. W. Parvin*, S. W. Martin, F. Cooke, Jr., and B.B. Freeland, Jr.

ABSTRACT

The unusually poor weather that occurred during the harvest season of 2002 in the Midsouth rejuvenated an interest in the relationship between cotton yield and harvest date and rainfall. In 2002, a Mississippi cotton grower experienced a yield reduction of 705 kg of lint per hectare during harvest. To determine the number of hectares per picker, growers need to know how the yield of commercial cotton declines during harvest. To estimate yield losses experienced by a grower, economists need to know the relationship between yield and harvest season rainfall. Regression results based on 3 yr of data from research plots (1991, 1992, and 1993) and data from a commercial cotton operation in 2002 indicated that yield declines 2.35 kg of lint per day and 4.09 kg of lint per centimeter of accumulated rainfall.

Conventional wisdom insists that as the cotton harvest season progresses, yield declines with time and that the adverse effect of a given amount of rainfall increases with time. Unusually poor weather conditions during the 2002 cotton harvest in the Midsouth extended the harvest period, reduced yield and quality, and rejuvenated interest in estimates of the losses that occur during harvest. In 2002, tropical moisture from the Gulf of Mexico developed into 11 rain events resulting in greater than 50.8 cm of precipitation during the harvest season at many locations in the Midsouth (Freeland et al., 2004). The normal rainfall accumulation during harvest season (25 September through 25 October) at Stoneville, Mississippi, is 5.92 cm (NOAA, 2001).

Starting on 19 Sept. 2002, a tropical depression dropped 8.05 cm of precipitation at the National Weather Service station located at Stoneville, Mississippi. Hurricane Isidore deposited 10.21 cm at Stoneville on 25 and 26 September, and Hurricane Lili added 5.94 cm on 3 and 4 October (MASS, 2002). The longest period without rain was 8 d in November, approximately one month after the usual completion of harvest.

In 2002, two research tests located at the Stoneville research station had some plots that were harvested in early September and some of the plots were harvested in November. Freeland et al. (2004) reported yield reductions of 19 and 35% with an average loss in yield and quality of $593 per hectare. From 1985 through 1988, Parvin (1990) collected hand-harvested data on the relationship between yield and time in commercial cotton at 22 locations in the Delta area of Mississippi. Average loss per week increased from 3.91% in the middle of September to 5.57% by the end of October. The loss estimates ranged from 1.58 to 10.95% per week. Spurlock and Parvin (1988) suggested that yield increases at 2% per day for the first week of harvest, reaches a maximum in week 2, then declines at 0.5% per day during weeks 3 and 4, at 1% per day during weeks 5 and 6, and at 1.5% per day thereafter. Parvin and Cooke (1990) reported a yield index of 100, 96, 92, 86, 81, 76, 71, and 65 for weeks 1 through 8 of harvest.

Most of the earlier papers on the effect of harvest weather on yield and quality were developed to address the then emotional issue of whether the losses due to insect induced delayed maturity should result in reduced insect treatment thresholds (Parvin et al., 1985; Scott et al., 1985; Sheng and Hopper, 1988) and were generally discounted or ignored by the majority of university and USDA research and extension professionals. An exception was the research reported by Williford et al. (1995) from research plots at Stoneville, Mississippi, that examined replicated weekly harvest treatments for reductions in yield and quality during 1991, 1992, and 1993. The 3 yr provided different environments that were reflected in production. Yield in 1991 was above average, 1992 yield was average, and yield in 1993 was below average. The average yield was 1528, 1110, and 909 kg of lint per hectare for 1991, 1992, 1993, respectively.

D.W. Parvin, Department of Agricultural Economics, Mississippi State University, P.O. Box 5187, Mississippi State, MS 39762; Steven W. Martin and F. Cooke, Jr., Delta Research & Extension Center, Mississippi State University, P.O. Box 69, Stoneville, MS 38776; Bart B. Freeland, Jr., USDA, Office of the Chief Economist, World Agricultural Outlook Board, Stoneville, MS 38776
*Corresponding author: parvin@agecon.msstate.edu

Rainfall during the harvest period ranged from 14.86 to 37.64 cm and the yield reduction varied from 9% to 17%. Harvest date (time) significantly influenced yield and quality. A yield loss of 2.09 kg of lint per centimeter of accumulated rainfall was reported, but the authors noted that rainfall during late season seemed to result in a greater yield reduction than the same amount of rainfall during early season.

In the Midsouth, current cotton cultivars are ready for harvest about 25 September (Parvin et al., 1989; Williford et al., 1995). Most Midsouth producers have enough equipment to complete harvest in 20 to 22 working days, which usually requires 30 to 35 calendar days and involves two rainfall events that delay harvest (Martin et al., 2003; Parvin et al., 1985). If the producer owns excessive harvesting capacity, harvesting fixed cost is high. If harvesting capacity is low, the harvesting period is extended and the probability of yield and quality losses is increased (Parvin et al., 1987; Williford et al., 1995). Additional information is needed on the reduction in yield of commercial cotton as harvest is delayed or extended. Growers need to know the relationship between the value of a hectare of cotton and time, given average weather (rainfall), to determine the number of pickers per farm (hectares per picker). Economists need to know the relationship between value and rainfall to estimate losses experienced by a grower in a given year. This paper examines the relationship between yield and time and rainfall during the harvest period.

MATERIALS AND METHODS

The experience of a Mississippi cotton farm in 2002 comprised of four locations with similar expected yields is summarized in Table 1. The expected farm yield is approximately 6.18 bales per hectare, and the average yield by location has varied by less than 50 kg of lint per hectare since 1996. Harvest began at the first location on 23 September and the 357 hectares averaged 1485 kg of lint per hectare. By the end of harvest in mid-November, average yield for locations three and four had declined approximately 450 kg per hectare.

All bales were identified by harvest date. Daily production (kilograms of lint) and hectares harvested were employed to estimate yield for each harvest date. Daily rainfall was available at each location. The rainfall variable for each location was defined as its accumulated rainfall since the initial harvest date of 23 September for the farm. Initial defoliation date, which varied by location, was utilized to construct a time variable for each harvest date. The time variable for each location was defined as days since initial defoliation date plus 14 d. For example, initial defoliation date of 9 September plus 14 d is 23 September for location one, so the time variables for location one for 24 September was coded as 1.

Table 2 lists the 20 daily observations of the three variables, which describe the 2002 harvest, as follows:

Y = yield (kilograms of lint per hectare).
T = time (days since initial defoliation date plus 14d).
R = rainfall (accumulative centimeters of rainfall since 23 September).

During the 52-day harvest season, yield declined 705 kg, or 43 %, from 1639 to 934 kg of lint per hectare. The grower supplemented his usually sufficient number of harvesting units (pickers plus boll buggies and module builders) with fully supported custom pickers in October; otherwise, the harvest period would have been extended and yield would likely have continued to decline. Regression analyses were used to estimate the relationship between yield as the dependent variable and time and rainfall as independent variables. It was hypothesized that decreases in the yield variable could be explained by increases in the time and rainfall variables, and the time and rainfall interaction would be significant; therefore, the sign of the estimated coefficients for time, rainfall, and their interaction were hypothesized

Table 1. Number of hectares, cotton cultivars, initial defoliation dates, harvest dates, and average yields (kg of lint /ha) for a Mississippi cotton farm in 2002

Location	Hectares	Cultivar	Initial defoliation date	Initial harvest date	Ending harvest date	Average yield (kg/ha)
1	357	DPL1218BR	9 Sept	23 Sept	14 Oct	1485
2	83	DPL1218BR	28 Sept	19 Oct	1 Nov	1224
3	116	DPL215BR	29 Sept	1 Nov	9 Nov	1012
4	293	DPL215BR	2 Oct	2 Nov	13 Nov	1042

to be negative, and one centimeter of rainfall late in the harvest season was expected to decrease yield more than one centimeter of rainfall during early harvest season. Estimated coefficients were subjected to student's *t*-test, sign considered, with $P \leq 0.05$ (Steel and Torrie, 1960).

Table 2. Yield, days since initial defoliation date plus 14 d, and rainfall since 23 September for a Mississippi cotton farm in 2002

Location	Harvest date	Yield (kg/ha)	Days since initial defoliation date+14 d	Rainfall (cm)
1	23 Sept	1639	0	0.0
	24 Sept	1592	1	0.0
	30 Sept	1539	7	11.4
	1 Oct	1403	8	11.4
	2 Oct	1413	9	11.4
	14 Oct	1084	21	33.8
2	19 Oct	1291	7	35.8
	31 Oct	1264	19	47.2
	1 Nov	1166	20	47.2
3	1 Nov	1030	19	49.0
	2 Nov	1033	20	49.0
	8 Nov	987	26	61.5
	9 Nov	994	27	61.5
4	2 Nov	1100	17	39.1
	7 Nov	1085	22	48.5
	8 Nov	1069	23	48.5
	9 Nov	1041	24	48.5
	10 Nov	1046	25	48.5
	12 Nov	1002	27	49.5
	13 Nov	934	28	49.5

RESULTS AND DISCUSSION

One assumption of the classical linear regression model is that there are no exact linear relationships among the observed values of the regressors (time and rainfall in this study). In practice an exact linear relationship is highly improbable, but the general interdependence of many biological and economic variables can cause a near linear relationship in time series of regressors resulting in a statistical problem referred to as multicollinearity or intercorrelation (Goldberger, 1964).

Multicollinearity can produce large standard errors of the coefficients. While the estimated coefficients often reproduce the sample extremely well (large R^2), they often will not reproduce other sets of similar data. In alternative models, estimates of the coefficients can be highly variable or unstable and the researcher will be uncertain of their population values. With stepwise regression, while R^2 increases, standard errors often explode. The researcher may have to accept that reliable estimates of the coefficients may not be made from the sample itself and supplement the estimates with a priori information.

In practice, researchers seldom formulate and estimate a relationship, test a few hypotheses, and stop (Goldberger, 1964). It is common to explore a range of alternative models. This is especially true when problems of multicollinearity exist. After estimating a linear relationship, researchers often investigate the effects of adding or dropping a regressor. If the relationship between the correlated independent variables or regressors is known, improved results may be obtained from combining the regressors into a new variable. When the relationship between the regressors is not known, occasionally it is helpful to substitute their interaction for one of the intercorrelated independent variables. Increased sample size may improve the problem associated with multicollinearity. Because of multicollinearity or intercorrelation problems between the regressors or independent variables, time and rainfall, a series of six models was estimated.

Yield. Table 3 summarizes the regression results of six models that employed 20 observations from 2002 to explain yield as a function of time and rainfall. Model 4 was the model of most interest. It was the only model to contain estimated coefficients for time, rainfall, and the time by rainfall interaction. Because of expected statistical problems with multicollinearity, estimates of Models 1, 2, and 3 were obtained. The 20 observations on time and rainfall in the 2002 data had a correlation coefficient of 0.92. Models 5 and 6 were an effort to circumvent problems associated with multicollinearity. The intercept was statistically significant in all six models and the estimated coefficient for time and rainfall in Models 1 and 2 have the expected or correct sign and are significant. When the signed *t*-test is employed, an estimated coefficient cannot be statistically significant unless its sign is correct, but when a series of models is estimated from a single data set and a given estimated coefficient is not significant, it may be helpful to discuss its sign. Time explains 91% of the variability in yield and rainfall explains 87% of the variability in yield, so together they explain 93% of the variability in yield (Model 3). The estimated coefficients for time and rainfall in Model 3 have the correct sign and are statistically significant, but they indicate that a delay of one day causes a larger reduction in yield than an additional 2.54 cm of rainfall. The relative size of the two coefficients is not consistent with a priori information. In Model 4, the sign of the estimated coefficient for the interaction term is not correct, biasing upward the size of the time and/or rainfall coefficients. In Models 5 and 6, all of the estimated coefficients have the correct sign, but only the coefficient for rainfall in Model 6 was significant. Problems of multicollinearity indicate that the level of statistical confidence associated with any of the estimated coefficients in Table 3 and any prediction of yield based on them is unknown.

Table 4 compares yield and rainfall variables for 2002 with those reported by Williford et al. (1995) for 1991 to 1993. In 2002, average yield was 1242 kg per hectare compared with a range of 909 kg to 1528 kg for 1991 to 1993. The yield loss in 2002 was 705 kg compared with a range of 83 kg to 284 kg for 1991 to 1993. Rainfall during the harvest season in 2002 was 49.5 cm, compared with 32.0 cm in 1991, 7.9 cm in 1992, and 13.0 cm in 1993. The ratio of loss to rainfall in 2002 was 14.2 kg/cm compared with a range of 6.4 kg/cm to 14.1 kg/cm for 1993 and 1992. To increase sample size, increase the range in yield and rainfall variables, and improve the problem with multicollinearity, the 21 observations from the 1991 to 1993 study were combined with the 20 observations for 2002, and the six-yield model re-estimated with 41 observations. The correlation between time and rainfall dropped to 0.09 and was not significant.

Because yield varied considerably by years, a different intercept was estimated for each year. The results are summarized in Table 5. The estimated coefficient for time in Model 1 has the correct sign, is significant, and its size is consistent with estimates provided by Parvin (1990), Parvin and Cooke (1990), and Spurlock and Parvin (1988). It indicates that the yield of cotton declines 40.74 kg of lint or 3.4% per week given the average rainfall (amount and distribution) that occurred in 1991, 1992, 1993, and 2002. The estimated coefficient for rainfall in Model 2 has the correct sign and is significant. It is almost twice as large as the 13.11 kg per hectare estimate reported by Williford et al. (1995). It was expected to be larger since the loss experienced in 2002 (almost 50% of the observations) was 705 kg of lint per hectare or 4.4 times the average loss reflected by the 21 observations for 1991, 1992, and 1993.

Table 3. Results of the regression of yield as a function of time (T) and rainfall (R) from four locations on a Mississippi cotton farm in 2002

Statistic	Model[z]					
	1	2	3	4	5	6
R^2	.91	.87	.93	.94	.91	.88
Regressors	T	R	T, R	T, R, TR	T, TR	R, TR
Intercept	1596*	1579*	1606*	1644*	1599*	1559*
Estimated coefficients						
Time (T)	-23.47*	-	-14.89*	-20.87*	-24.05	-
Rainfall (R)	-	-10.49*	-4.24*	-5.81*	-	-7.39*
TR	-	-	-	0.15	-.01	-0.12

[z] Analysis based on 20 observations for 2002. Values followed by an asterisk are significant at $P \leq 0.05$ according to student's *t*-test.

Table 4. Average, maximum, and minimum yield, yield loss, harvest season rainfall, and loss divided by rainfall

Year	Yield (kg of lint/ha)				Harvest season rainfall (cm)	Loss/ rainfall (kg/cm)
	Average	Maximum	Minimum	Loss		
2002	1241.9	1638.7	933.7	705.0	49.5	14.2
1991	1527.7	1652.1	1368.6	283.6	32.0	8.9
1992	1109.6	1151.1	1040.2	111.0	7.9	14.1
1993	909.0	963.9	881.0	82.9	13.0	6.4

Table 5. Results for regression of yield as a function of time (T) and rainfall (R)

Statistic	Model[z]					
	1	2	3	4	5	6
R^2	.73	.94	94	.94	.85	.94
Regressors	T	R	T, R	T, R, TR	T, TR	R, TR
Intercept 1991	1741*	1628*	1619*	1622*	1627*	1622*
Intercept 1992	1265*	1130*	1130*	1130*	1090*	1130*
Intercept 1993	1098*	970*	970*	968*	936*	968*
Intercept 2002	1287*	1565*	1565*	1573*	1379*	1573*
Estimated coefficients						
Time (T)	-5.82*	-	.291	.01	1.73	-
Rainfall (R)	-	-10.10*	-10.29*	-10.74*	-	-10.74*
TR	-	-	-	0.02	-0.28*	0.02

[z] Analysis based on 41 observation from 1991 through 1993 and from 2002. Values followed by an asterisk are significant at $P \leq 0.05$ according to student's *t*-test.

The estimated coefficient for rainfall in Model 3 is biased downward (more negative) because the estimated coefficient for time has the incorrect sign. The bias is even larger in Model 4 since the estimated coefficients for time and the interaction term have the incorrect sign. Models 3 to 6 contain time or the time by rainfall interaction term. In each model the estimated coefficient for time has the incorrect sign, and was not statistically significant. In all models that contain the rainfall variable, the estimated coefficient for rainfall is significant. Models 4, 5, and 6 contain the time by rainfall interaction term. When the rainfall variable is present (Models 4 and 6), the estimated coefficient for the interaction term has the incorrect sign and is not significant. In Model 5, when the rainfall term is absent, the interaction term has the correct sign and is significant.

These results indicate that rainfall is more important than time for explaining the deterioration in cotton yield as the harvest season progresses. The least square estimate is 10.10 kg of lint per centimeter of rainfall.

CONCLUSIONS

In a given year, accumulated rainfall and time are obviously correlated. When time and rainfall were employed as independent variables in regression analysis based on one year of data (2002), statistical problems with multicollinearity prevented reliable estimates of their relationship with yield. When 4 yr of data were employed, results indicated that yield declines 10.10 kg of lint per centimeter of accumulative rainfall. Failure to include data from years like 2002 in studies designed to estimate the impact of harvest season weather on the yield of cotton may bias the estimates obtained. In this study, the reported estimate of the relationship between yield and rainfall of minus 10.10 may be biased downward (more negative), because the year with the highest rainfall and the largest yield reduction (2002) was weighted more than the other 3 yr. It is likely biased upward (less negative) because custom harvest shortened the harvest season in 2002. A priori information (Parvin, 1990; Parvin and Cooke, 1990; Spurlock and Parvin, 1988) indicates that the estimated coefficient for time of minus 5.82 kg of lint per day (3.4% per week) is a satisfactory estimate of the daily decline in yield given average weather.

DISCLAIMER

Mention of a trademark, warranty, proprietary product or vendor does not constitute a guarantee by Mississippi State University and does not imply approval or recommendation of the product to the exclusion of others that may be suitable.

REFERENCES

Freeland,T.B., S.M. Martin, M.W. Ebelhar, and W.R. Meredith. 2004. Yield, quality, and economic impacts of 2002 harvest season rainfall in the Mississippi Delta. p. 600-608. *In* Proc. Beltwide Cotton Conf., San Antonio, TX. 4-8 Jan. 2004. Natl. Cotton Counc. Am., Memphis, TN.

Goldberger, A.S. 1964. Econometric Theory. 3rd ed. John Wiley and Sons, Inc., New York.

Martin, S.W., F.T. Cooke, and D.W. Parvin. 2003. Organization and structure of Mississippi Delta cotton farms. Mississippi Agric. and Forestry Exp. Stn. Bulletin 397. Mississippi State University, Mississippi State, MS.

Mississippi Agricultural Statistical Service (MASS). 2002. USDA, MASS Weekly weather crop reports: 9 September 2002 to 25 November 2002.

National Oceanic and Atmospheric Administration (NOAA). 2001. Daily station normals of temperature, precipitation, and heating and cooling degree days, Station Number 228445. Climatograph of the United States Report 84. 1971-2000. Natl. Environ. Satellite,Data, and Information Serv. (NESDIS), MCDC. Asheville, NC.

Parvin, D.W. 1990. The weight loss associated with defoliated cotton under Mississippi conditions. Staff Paper Series 92. Dep. Agric. Economics, Mississippi Agric. and Forestry Exp. Stn. Mississippi State University, Mississippi State, MS.

Parvin, D.W. and F.T. Cooke. 1990. The mathematics of cotton harvest. p. 119-121. *In* Proc. Beltwide Cotton Conf., Las Vegas, NV. 9-14 Jan. 1990. Natl Cotton Counc. Am., Memphis, TN.

Parvin, D.W., F.T. Cooke, W.P. Scott, and J.W. Smith. 1985. The economics of cotton harvesting in the Midsouth. Staff Paper 69. Dep. Agric. Economics, Mississippi Agric. and Forestry Exp. Stn. Mississippi State University, Mississippi State, MS.

Parvin, D.W., F.T. Cooke, and S. R. Spurlock. 1989. A cotton harvesting model. p. 462-464. *In* Proc. Beltwide Cotton Conf., Nashville, TN. 2-7 Jan. 1989. Natl. Cotton Counc. Am., Memphis, TN.

Parvin, D.W., J.W. Smith, and F.T. Cooke. 1987. Cost effectiveness of harvesting at the right time. p. 39-40. *In* Proc. Beltwide Cotton Conf., Dallas, TX. 4-8 Jan. 1987. Natl. Cotton Counc. Am., Memphis, TN.

Scott, W.P., J.W. Smith, and D.W. Parvin. 1985. The economics of early season insect control. p. 256-259. *In* Proc. Beltwide Cotton Conf., New Orleans, LA. 6-11 Jan. 1985. Natl. Cotton Counc. Am., Memphis, TN

Sheng, C.F. and K.R. Hopper. 1988. Harvesting models and pest management in cotton. Environ. Entomol. 17(5): 755-763.

Spurlock, S.R. and D.W. Parvin. 1988. A simulation model for cotton harvesting decisions. Staff Paper 84. Dep. Agric. Economics, Mississippi Agric. and Forestry Exp. Stn. Mississippi State University, Mississippi State, MS.

Steel, R.G. and J.H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill Book Company, Inc., New York.

Williford, J.R., F.T. Cooke, D.F. Caillouet, and S. Anthony. 1995. Effect of harvest timing on cotton yield and quality. p. 633-635. *In* Proc. Beltwide Cotton Conf., San Antonio, TX. 4-7 Jan. 1995. Natl. Cotton Counc. Am., Memphis, TN.