The Confidence Intervals of Preharvest State Corn Yield Forecasts

by

Daniel O'Brien and S. Elwynn Taylor

THE CONFIDENCE INTERVALS OF PREHARVEST STATE CORN YIELD FORECASTS

Daniel O'Brien and S. Elwynn Taylor*

This paper examines the accuracy of preharvest corn yield forecasts from crop-weather models for major U.S. Corn Belt states. Cross section time series models using dummy variables to represent state crop reporting districts are estimated via ordinary least squares for Iowa, Illinois and Indiana. The models use normalized rainfall and temperature data and measures of crop maturity for the 1972-1991 period. Four successive corn crop-weather models are estimated for each state using information available on July 1, August 1, September 1 and October 1, respectively. Crop reporting district level yield forecasts and forecast errors are derived via unconditional forecast error calculations. These forecasts and forecast errors are then aggregated together on a monthly basis throughout the growing season to form state yield forecasts and confidence intervals. During the 1992-1994 period, forecast accuracy was poor when weather conditions were abnormal compared to the average conditions during the period of model estimation. This is illustrated by inaccurate forecasts for 1992 and 1993 for Iowa caused by abnormally wet and cool conditions. Future crop yield forecasting efforts should focus on the use of resource capture models, which hold promise of producing more accurate forecasts than crop-weather models.

Introduction

The purpose of this paper is to examine the accuracy of preharvest corn yield forecasts from crop-weather models for major U.S. Corn Belt states. Measures of state corn yield forecast accuracy may be useful to grain market analysts and to the users of corn yield futures contracts.

Crop-weather models are typically designed to measure the effect of technology trends and weather factors on corn yields. Thompson, Westcott and others have used crop-weather models to make preharvest forecasts of corn yields. Extensive estimation of pre-growing season crop yield probability distributions has been carried out by Day, Gallagher, Fackler, Fackler and Young, Moss and Shonkwiler, and Moss and Boggess. However, crop-weather models have not been used to forecast the probability distribution of corn yields conditional on weather up to a point in time during the growing season. Plant process models have been used for this purpose by Krog and Kunkel. Whereas crop-weather models assume that yields are functions of technology trends and deviations from normal weather conditions, plant process models build their yield projections from a zero yield base, assuming that the various processes of a plant work together in a cumulative manner to bring about a final crop yield. Agronomists and climatologists generally prefer using crop-weather models for forecasting yields because a 10% forecast error for PPMs,

* Extension Ag Economist-NW and Assistant Professor, Dept. of Agricultural Economics, Kansas State University; and Extension Climatologist and Professor, Dept. of Agronomy, Iowa State University.

starting from zero on a yield scale, is much greater than a 10% prediction error for a crop growth model whose beginning yield estimate consists of a point determined by the combination of a constant and a technology trend. Agronomists indicate that resource capture models are an alternative crop modeling source that has the potential to provide forecasts of greater accuracy than either crop-weather or plant process models.

The procedure used here will be to first estimate crop-weather models for major U.S. Corn Belt states. These state models will then be used to forecast corn yields and to derive appropriate forecast confidence intervals. Four successive crop-weather models are estimated for each state at monthly intervals using 1972-1991 data. The forecast dates coincide with successive USDA crop production forecasts throughout the U.S. corn growing season (i.e., on July 1, August 1, September 1 and October 1). Crop reporting district level yield forecasts and forecast errors are then estimated for 1992, 1993 and 1994 using unconditional forecast error calculations. These yield forecasts and forecast errors are aggregated together on a monthly basis throughout the growing season to form state level yield forecasts and forecast confidence intervals.

Thompson's Crop-Weather Model

The crop-weather models utilized here follow from Thompson. Thompson estimated corn crop-weather models for the 1891 to 1983 time period for five major Corn Belt states -- Illinois, Indiana, Iowa, Missouri, and Ohio. His purposes were to determine the impact of changes in climate and weather variability on corn production and to estimate the effects of departures from normal weather on corn yields. Thompson's multivariate quadratic equation is:

$$Y = a + bj * Dj * TREND + c * X(i) - d * X(i)^2$$

where:

Y = Corn yield
Dj = Dummy variable for technological trends during different time periods
(j = 1 => 1930-1959; j = 2 => 1960-1972; j = 3 => 1973-1983)
TREND = Technological trend
X(i) = Weather variables (departures from normal)
i = 1: Preseason precipitation Sept-June
= 2: June temperature
= 3: July rainfall
= 4: July temperature
= 5: August rainfall
= 6: August temperature

Thompson estimated three separate technical trends for the 1930-1959, 1960-1972, and 1973-1983 time periods to represent varying rates of technological change in corn yields. These state crop-weather models were estimated by ordinary least squares. The R^2 for the regression analyses were: Illinois, 0.97; Indiana, 0.96; Iowa, 0.96; Missouri, 0.93; and Ohio, 0.96. When using this model for forecasting during the growing season, Thompson assumed normal weather for the remainder of the year through harvest with normal weather defined as the average of conditions from 1891 to 1983.

In this study Thompson's crop-weather model is extended in three ways. First, instead of using whole month explanatory variables, disaggregated intramonth weather data is used for the key corn development months of July and August. The impact of weather conditions on corn yields during critical 10 and 20 day time periods in July and August may lead to increased model forecast accuracy. Second, separate crop-weather models are estimated at monthly intervals throughout the corn growing season, using only known weather information at the time of model estimation. Third, forecasts of corn yield probability distributions are calculated from crop-weather models for individual crop reporting districts, and then aggregated to form forecasts and forecast probability distributions for the states.

Weather and Yield Data

USDA corn yield and acreage data were collected by crop reporting district for 1972-1994. Monthly rainfall and temperature data were obtained from the climatological data base of the Midwest Climate Service in Champaign, Illinois (Kunkel, 1992). Rainfall and temperature data were also collected for July 1-10, 11-20, and 21-31, and for August 1-10 and 11-31 in an effort to improve the representation of weather conditions during critical corn development periods. USDA state survey estimates of the percentage of the corn crop planted by mid-May are used to represent planting progress. In the Iowa model alone, the date when 50% of the corn fields in a crop reporting district have reached 75% silking is used as an indicator of crop maturity. This information was not readily accessible and therefore not collected for other Corn Belt states.

The 1972-1991 time period was chosen for model estimation because the weather patterns and crop production technology of these years differed from earlier periods. Since the early 1970s there has been a marked increase in the variability of corn yields and growing season weather conditions. U.S. corn yields have continued to increase during the 1970s and 1980s, but at a more moderate rate than during the previous decade. Post-1972 corn yield variability was similar to that of the 1930s except that there were both unusually low and unusually high yields during the 70's and the 80's, while only unusually low yields occurred during the 1930s. Because of these factors, there is a higher likelihood of avoiding heteroskedastic corn yields across the Corn Belt by estimating the model for the 1972-1991 period than if earlier years were also included in model estimation.

The weather data was normalized for conditions within each crop reporting district. For example, normalization of June rainfall data for the 1972-1991 period for the west central Iowa crop and weather reporting district is carried out by subtracting the district's 1972-1991 average rainfall from the June rainfall for a specific year, and then dividing by the standard deviation of June rainfall for 1972-1991. Although weather data normality is not formally assumed here, normalization of the weather variables provides an approximate idea of the relative magnitude and expected frequency of weather deviations from normal. Normalization of the weather data also facilitates the use of special explanatory variables (i.e., aridity indices) to measure the effects of extreme weather conditions in model estimation.

Corn crop-weather models are estimated for the 1972-1991 time period at monthly intervals throughout the corn growing season. July 1, August 1, September 1, and October 1 models are estimated for Iowa, Illinois, Indiana and other Corn Belt states. In the western Corn Belt states crop-weather models are calculated separately for irrigated and nonirrigated corn production. The focus of this paper will be on the crop weather model estimates and minimum forecast error confidence intervals for Iowa, Illinois and Indiana. The dates for which the crop-weather models are estimated during the U.S. corn growing season (July 1, August 1, September 1, and October 1) coincide with the forecast dates represented by the USDA state corn yield and production estimates. USDA forecasts are released throughout the growing season 8 to 12 days after the first of each month. They represent the government's best prediction of corn yields and production given conditions up through the end of the previous month. However, the August 1 report is the first growing season corn crop estimate based on extensive field surveys.

Each successive monthly model is estimated using only data for weather conditions that have actually occurred up to that point in time during the growing season. This allows for the application of unconditional forecasting techniques and the assumption of normally distributed forecast errors. The explanatory variables for the successive monthly models are listed in Table 1 and definitions of the explanatory variables are included in Table 2. The yield affecting factors in these models are technological advancements (represented by trend), preseason rainfall accumulation, planting date, and growing season temperature and precipitation. Also, in Iowa crop maturity is represented by silking progress. The most critical stages of corn physiological development are tasseling and silking, both part of the process of corn pollination and seed set (Ritchie et al., 1989). These stages generally occur during July and early August. The intramonth rainfall and temperature effects are designed to represent yield impacting weather conditions during these critical times. After silking, the corn plant moves through the blister, milk, dough and dent stages on to physiological maturity. These stages typically occur from early August through late September or early October, till crop maturity or at the time of the first killing frost. The August and September explanatory variables are intended to measure the yield impact of weather conditions occurring during these later crop development stages.

Aridity indices are used to estimate the effects of temperature and rainfall extremes on corn yields. Either high temperatures or low rainfall alone may not have a dramatic effect on corn yields. However, when these weather conditions occur simultaneously, corn yields may be severely impacted. The criteria for weather extremes are defined in terms of higher positive and/or lower negative values for the normalized weather data. For example, HotJun measures the effect on yields of monthly average June temperatures which are greater than 1.5 normalized temperature units above average for the 1972-1991 period. For forecasting purposes, the yield impact of above average June temperatures would be represented by the coefficient of TempJun as long as the temperatures were less than or equal to 1.5 normalized units above average. However, if June temperatures rose more than 1.5 normalized units higher than average, then the yield impact would be represented by the coefficient of TempJun adjusted by the coefficent for HotJun. Note that for the HotDryJly, HotDryJlym, and HotDrySep aridity indices a (-1) is included in the aridity index equations in order to reverse the signs of the coefficients. Otherwise the effect on corn yields of combined hot and dry conditions during these time periods would be reported as positive.

Each state represents a grouping of crop reporting districts, each with twenty years of corn yield and weather data. The different crop reporting districts in a state can be thought of as cross sections of the state data set, with each district or cross section contributing a twenty year time series of yield and weather data. The number of observations used in model estimation for each state is equal to the number of crop reporting districts in the state multiplied by the number of annual yield and weather observations for each district. The cross section time series approach provides more degrees of freedom for model estimation.

In this pooled cross section time series data set structure, dummy variables are used to represent adjustments to the the model constant for each crop reporting district within a state. The intercept is the estimated fixed effect for an arbitrarily chosen district. Assuming there are n crop districts within a state, dummy variables are used to represent adjustments to the constant term for each of that state's n-1 other districts. This estimation approach assumes that the model constant may vary across CRDs within a state, but that there is no cross sectional variation in the yield effects of rainfall and temperature. If there is considerable variation in the effect of weather conditions upon corn yields across crop reporting districts within a state, this will limit yield forecasting accuracy. A full description of the dummy variable approach to estimating pooled cross section time series models is given on pp. 468-479 in Judge et al..

Results of Crop-Weather Model Estimation

Ordinary least squares estimates of the July 1, August 1, September 1 and October 1 crop-weather models for Iowa, Illinois and Indiana are given in Tables 3 and 4. Note that in Table 3 the Iowa crop-weather model is estimated for two time periods, 1972-1991 and 1972-1993. The was done to show the negative yield effects of the extraordinarily wet weather conditions and slow crop development during 1993. Significant yield effects for the July 1, August 1, September 1 and October 1 models are reported at the 5 percent and 1 percent levels (represented by "*" and "**", respectively). Overall, higher rainfall totals during key crop development periods had a positive effect on corn yields. However, preseason rainfall (RainOctMy) tended to have a negative yield impact, especially in Illinois and Indiana, and to a lesser degree in Iowa. Extremely wet conditions tended to negatively affect yields as shown by the significance, magnitude and signs of the aridity indices WetJly in Illinois and Indiana and in the 1972-1993 Iowa model.

In general, warmer than average temperatures during June, July and August had a negative impact on corn yields in these states. Higher May temperatures hurt yields in Iowa, helped yields in Indiana, and had little effect in Indiana. One exception is the significant positive yield impact of high early August temperatures in the 1972-1993 Iowa model. Combined hot and dry July conditions (HotDryJly) had a significant negative yield impact.

The yield trend is positive and significant in all of the monthly models for these states. Note the smaller annual trend yield increase in the Iowa 1972-1993 model as compared to the 1972-1991 version, reflecting extremely low 1993 Iowa corn yields. A higher than average percentage of the corn crop planted by mid-May (PlntMay) had a negative yields impact in Iowa, but was not consistently negative in Illinois and Indiana. Later silking dates (Silk*) had significant negative yield effects in both Iowa models. Later silking dates indicate that the crop is maturing slowly and at risk of not accumulating adequate heat units to reach full physiological maturity prior to the first killing frost. The model standard errors and R^2 values decrease and increase, respectively, throughout the growing season, as would be expected.

These results reaffirm the general idea that rainfall and temperature conditions during planting time and during July and August have critical impacts on corn yield. In addition, they indicate the potential for yield damage caused by excessive preseason rainfall, above normal June temperatures, and excessive amounts of rainfall in July and August. Yields can also be reduced by a combination of high temperatures and low rainfall during July and even September. The standard error and R^2 results imply that most of the increase in forecast accuracy occurs between the July 1 and the August 1 crop-weather models. After this initial gain in accuracy, model standard errors and R^2 measures improve to a lesser degree between the August 1 and September 1 models, and then only marginally between the September 1 and the October 1 models. In the forecast variance calculations, the standard forecast error for corn yields for a crop reporting district will always be at least slightly larger than the estimated standard error of the underlying crop-weather model. As a result, the monthly crop-weather model standard errors are major determinants of the potential accuracy of any forecasts derived from these models.

Unconditional Forecasts

If a yield forecast is made using weather information that is known at the time of the forecast, it is an unconditional forecast. For unconditional forecasts from the crop-weather models estimated via ordinary least squares, the yield predictions are unbiased and the forecast errors are normally distributed around the forecast of average yield. The formulas for calculating the unconditional forecast error for both the univariate and multivariate cases are given below (See Pindyck and Rubinfeld). Assuming that the crop-weather models can be represented in general form as $y_t = x_t * b + e_t$ in the univariate case, and $Y_t = X_t' * B + e_t$ in the multivariate case, with standard OLS assumptions for both, the respective unconditional forecast error equations are as follows.

Univariate case: $s_f^2 = s^2 * [1 + {}^1/_T + \{x_{T+1} - x^{BAR}\} / \Sigma\{x_t - x^{BAR}\}^2]$

Multivariate case: $s_f^2 = s^2 * [1 + X_k(X'X)^{-1}X_k']$

where,

s_f^2 = Forecast variance of the OLS model forecast

s^2 = Variance of the OLS model estimate

x = Univariate explanatory variable

x_{T+1} = Known univariate explanatory variable during the forecast period

x^{BAR} = Mean of univariate explanatory variable for estimation period

X = Multivariate explanatory variable matrix

X_k = Known vector of explanatory variables during the forecast period

T = Number of observations used for model estimation

b, B = Parameter estimates (univariate and multivariate, respectively)

e_t, e_t = Estimation errors (univariate and multivariate, respectively)

A key principle in unconditional forecasting is that the greater the difference between the value of a forecast period explanatory variable and the mean of that variable during the period of model estimation, the larger the forecast error is in relation to the standard error of the estimated econometric model. For example, if the amount of summer time rainfall received during an out-of-sample forecast year is approximately equal to the historic average rainfall during the model estimation period, then the difference between the standard error of the model and the forecast error of the yield estimate will be very small, depending mainly on the number of observations used in model estimation.

Deriving Forecasts, Forecast Variance, and Forecast Confidence Intervals

Crop reporting district level corn yield forecasts are obtained from the state corn yield models using known weather data for the forecast period. Crop reporting district level harvested acreage can be estimated during the growing season using historic planted to harvested acreage relationships. In this paper the actual harvested acreage for each crop reporting district is used to simplify the analysis. Forecast corn yields for a specific crop production district within a state and for the state are calculated as follows:

$$E[YLD_I^F] = \frac{\sum_i^n Ac_i \bullet E[Yld_i]}{\{\sum_i^n Ac_i\}} = \frac{\sum_i^n (Ac_i \bullet E[Yld_i])}{AC_I}$$

$$= \sum_i^n (Ac_i / AC_I) \bullet E[Yld_i]$$

where,

Ac_i	= Estimate of harvested acreage for district I (i = 1,2, ..n CRDs in state I)
$E[Yld_i]$	= Forecast Yield for district i
$E[YLD_I^F]$	= Forecast yield for state I (I = 1, 2, 3 for Iowa, Illinois and Indiana)
AC_I	= Total harvested acreage in state I

State level yield forecast variances are calculated as follows:

$$\sigma^2_{YLD_I^F} = \sum_i^n (Ac_i/AC_I)^2 \bullet \sigma^2_{Yld_i^F} + 2\sum_i^n \sum_{i<j}^{n-1} (Ac_i/AC_I)(Ac_j/AC_I)\mathrm{Cov}(Y_i F, Y_j F)$$

$$= \sum_i^n (Ac_i/AC_I)^2 \bullet \sigma^2_{Yld_i^F} \quad \text{if } \mathrm{Cov}(Y_i F, Y_j F) = 0 \Rightarrow (Y_i F, Y_j F \text{ are independent})$$

where,

$\sigma^2_{Yld_i^F}$	= Yield forecast variance for crop reporting district i
$\sigma^2_{YLD_I^F}$	= Yield forecast variance for state I
$\mathrm{Cov}(Y_i F, Y_j F)$	= Covariance of yield forecasts for crop reporting districts i and j

Whether $\mathrm{Cov}(Y_i F, Y_j F) = 0$ or not will impact the calculated state yield forecast variance and standard error. Independence will be assumed here because the factors that cause the covariance of yields across adjacent crop reporting districts (yield trend, rainfall, temperature, and other factors (Z)) are accounted for in the calculation of the district yield forecasts. Restated,

$$E[Yld_i] = Y_i F = f(Trend_i, Rain_i, Temp_i, Z_i) \quad \& \quad E[Yld_j] = Y_j F = f(Trend_j, Rain_j, Temp_j, Z_j)$$

Without question, there is considerable covariance of yields across crop reporting districts, i.e., $\mathrm{Cov}(Y_i, Y_j) \neq 0$. However, the simple covariance of crop reporting district yields over time is not the issue in this application. Rather, the concern here is with the covariance of trend and weather dependent yield forecasts.

$$\text{Cov}(Y_i, Y_j) \neq 0 \neq \text{Cov}(Y_i F(TRND_i, Rn_i, Tmp_i, Z_i), Y_j F(TRND_j, Rn_j, Tmp_j, Z_j))$$

The factors that cause similar year to year movement in forecast yields across adjacent crop reporting districts within a state are accounted for in the yield forecast equations. Since the primary causes of yield covariance are accounted for, independence of yield forecasts will be assumed, i.e., $\text{Cov}(Y_i F(*_i), Y_j F(*_j)) = 0$.

Crop reporting district level yield forecast confidence intervals can be estimated using forecast yield, the forecast error (s_f) and t-distribution values. Because the forecast error is normally distributed with mean = 0 and variance = σ^2, significance tests can be performed on y_0, the forecast value of Y, by calculating the normalized error. In practice, the parameters of a crop-weather model have to be estimated, so a t-distribution is used to represent the forecast error distribution.

Forecast error confidence intervals can be estimated around y_0, the forecast value of Y. The normalized error will have a t-distribution with T - 2 degrees of freedom. For example, a 95 percent forecast error confidence interval for y_0 can be constructed using the 1 - .95 = 0.05 or ν probability level of a t-distribution, where $0 < \nu < 1$. The resulting forecast confidence interval has a $1 - \nu$ or 95 percent probability of containing Y_0, the actual value of Y during the forecast period. The 95 percent confidence interval is calculated as follows:

$$y_0 - t_{n-k,.025} \bullet s_f \leq Y_0 \leq y_0 + t_{n-k,.025} \bullet s_f$$

where,

Y_0 = Actual value of Y for the forecast period
y_0 = Forecast value of Y for the forecast period
s_f = Standard error of the forecast

State Weather Conditions and Yield Forecasts for 1992-1994

Weather conditions during the 1992-1994 forecast time period was generally cooler and wetter than normal. Table 5 shows monthly temperature, precipitation and crop maturity data for 1992-1994 which is normalized to weather conditions during the 1972-1991 time period. The figures in Table 5 have been averaged across all crop reporting districts in each state. Table 5 shows that Iowa was particularly cool and wet during July of 1992 and 1993. Throughout the 1993 growing season, Iowa precipitation was abnormally high and temperatures were abnormally low (except during August). During the 1992-1994 period in Iowa, July temperatures were on average two standard deviations below normal while July precipitation was two standard deviations above normal. Slow crop development (as indicated by higher than normal values for Silk*) was particularly evident during 1993. This ultimately lead to widespread crop maturity problems in the state that year. In Illinois, wet June and moderately wet July conditions in 1993 were combined with slightly higher than normal July and August temperatures. In Iowa, Illinois and Indiana, no state-wide extreme temperature or rainfall problems are indicated during 1994.

The accuracy of this forecast procedure is shown for 1992, 1993 and 1994 for Iowa, Illinois and Indiana. July 1, August 1, September 1, and October 1 corn yield forecast probability distributions and actual yields are reported in Tables 6, 7 and 8. State corn production forecasts are also available for these same states and dates upon request from the authors. To derive the yield forecast probabilities in Tables 6, 7 and 8, the appropriate t value was multiplied by the yield forecast error for each month and either added or subtracted from the forecast yield for the month. Then the lower and upper bounds of the 99%, 95%, 90%, 67% and 50% confidence intervals were arranged in the form of a cumulative distribution of forecast yields. To illustrate, from Table 6, the August 1, 1994 Iowa crop-weather model yield forecast was 150.5 bu per acre. The 50% confidence interval for the 8/1/94 forecast includes yield forecasts from 146.7 bu (the 25% point on the cumulative distribution) to 154.3 bu (the 75% point in the distribution). Since the actual 1994 Iowa corn yield of 152.0 bushels per acre falls within this range, it is then said to be within the 50% confidence interval.

As shown in Table 6, the Iowa 1972-1991 model forecasts were "fair to poor" for 1992 and inaccurate for 1993. Considering the combination of abnormally low temperatures and high rainfall during July, 1992, the forecast was reasonably accurate. The final 1992 Iowa yield was within the 90% confidence interval on August 1 and the 99% confidence interval on September 1. However, during 1993, the model predicted very high yields, not capturing the yield reduction caused by slow crop development, even with the Silk* variable included. The high rainfall totals were interpreted by the model as increasing yields without adequate consideration of yield damage from combined cool and wet conditions. To better account for these factors, the 1972-1993 Iowa model was used to forecast 1994 Iowa yields. The accuracy of the 1994 forecast was much improved over those for 1992 and 1993, probably because 1994 was more of a "normal" crop year. The actual 1994 forecast was within the 50% confidence interval for the August, September and October 1 forecasts. It should be noted that the 1994 Iowa forecast using the 1972-1993 model was markedly more accurate than that from the 1972-1991 Iowa model. Overall, major changes in Iowa corn yield forecasts tended to occur between the July 1 and August 1 forecasts, with little change occurring after that time.

Yield forecasts for both Illinois (Table 7) and Indiana (Table 8) were extremely accurate for 1992 and 1993, but less so in 1994. The actual 1992 Illinois yield forecast fell within the 90% confidence intervals on August 1, and within the 50% confidence intervals for the September and October 1 forecasts. The 1993 actual Illinois yield fell within the 90% forecast confidence intervals for July and August 1, the 95% confidence intervals on September 1, and the 50% confidence interval for October 1. The 1994 Illinois forecast was not accurate, perhaps due to optimal timeliness of rainfall or other factors. In examination of the Illinois intramonthly weather data, 1994 was a year of basically "normal" weather with few extremes. Actual 1994 yields were approximately 20 bushels higher than forecast by the crop-weather models. Actual 1992 yields in Indiana fell within the 99% confidence intervals for the July forecast, the 90% confidence intervals for August, and the 67% interval for both the September and October 1 forecasts. In 1993, actual yields for Indiana were within the 90%, 99%, 50% and 67% confidence intervals for the July, August, September and October 1 forecasts, respectively. For 1994, actual Indiana yields were within the 90%, 99% and 99% confidence intervals for the July, September and October 1 forecasts, respectively. Indiana received timely rainfall during the first 10 days of July and normal weather thereafter in 1994, which may have boosted yields more than indicated in these models.

Conclusions

This research illustrates both the promise and the limitations of efforts to model and forecast corn yields with crop-weather models. The main contribution of this work is in an improved understanding of how weather conditions at critical times during the growing season can influence corn yields. These findings affirm the importance of July rainfall and temperature conditions on midwestern U.S. corn yields. This work also points out the impact of simultaneous rainfall and temperature combinations, of weather extremes, and of delayed plant maturity upon midwest corn yields. Applied forecasters can use this yield effect information in their efforts accurately predict corn yields. However, more research is needed in measuring the yield impact of delayed crop maturity in these key corn producing areas.

These forecasting results also point out the limitations of crop-weather model forecasts. Forecast accuracy was poor when weather conditions were abnormal compared to the average conditions during the time period for which the forecast model was estimated. This was particularly true with the abnormally wet conditions in Iowa during the 1992 and 1993 growing seasons, resulting in poor forecasts. These results support the USDA practice of not making forecasts based on actual field surveys until August 1. The July 1 forecasts in these models were not very accurate. However, the inclusion of critical July weather conditions in the August 1 crop-weather models markedly increased the explanatory ability of the models and the subsequent accuracy of their forecasts.

As Iowa corn yield futures become established it may be useful to compare the yield forecast variance implied by these models with actual variation over time in the yield futures contract. According to agronomists, resource capture models have the potential to more accurately model and forecast crop yields than either crop-weather or plant process models. Future research should focus on the forecast accuracy of these models at various stages of crop growth.

References

Day, R. (1965). Probability Distributions of Field Crop Yields. Journal of Farm Economics, 47, 713-741.

Fackler, P. (April, 1989). Modeling Trend and Higher Moment Properties of U.S. Corn Yields. Proceedings of Southern Regional Project S-232, Quantifying Long Run Agriclutural Risks and Evaluating Farmer Responses to Risk (pp. 65-80).

Fackler, P. and Young, D. (March, 1991). Notes on Modeling Regional Crop Yields. Proceedings of Southern Regional Project S-232, Quantifying Long Run Agricultural Risks and Evaluating Farmer Responses to Risk (pp. 89-106).

Gallagher, P. (1987). U.S. Soybean Yields: Estimation and Forecasting with Nonsymmetric Disturbances. American Journal of Agricultural Economics, 69, 796-803.

Judge, G. G., Hill, R. C., Griffiths, W. E., Lutkepohl, H. and Lee, T. (1988). Introduction to the Theory and Practice of Econometrics. New York: John Wiley and Sons, Inc.

Krog, D. R. (1988). Plant process model corn yield forecasts for Iowa (Doctoral dissertation, Iowa State University, 1988).

Kunkel, K. (1992). Midwestern Climate Information System. Champaign, Illinois.

Moss, C. B. and Boggess, W. G. (1992). New Procedures in Modeling Risk: Nihil Novum Sub Sol Est. (Staff Paper SP92-4). Institute of Food and Agricultural Sciences, University of Florida.

Moss, C. B. and Shonkwiler, J. S. (1991). Modeling Yield With Stochastic Trend and Nonnormal Errors (Staff Paper 91-17). Institute of Food and Agricultural Sciences, University of Florida.

Pindyck, R. S. and Rubenfield, D. L. (1991). Econometric Models and Economic Forecasts (3rd ed.). New York: McGraw-Hill, Inc.

Ritchie, S., Hanway, J. and Benson, G. (1989). How a Corn Plant Develops (Special Report No. 48, Reprinted July 1989). Iowa State University Cooperative Extension Service.

Thompson, L. M. (1986). Climatic Change, Weather Variability, and Corn Production. Agronomy Journal, 78, 649-653.

Westcott, P. C. (1989). An Analysis of Factors Influencing Corn Yields. United States Department of Agriculture - Economic Research Service, Feed Situation and Outlook, May, 16-22.

Table 1. Corn Crop-Weather Models

July 1 Model

Yield = f (Constant, Trend, $Rain_{OctMy}$, $Plnt_{May}$, $Temp_{May}$, $Temp_{Jun}$, Hot_{Jun}, $Rain_{Jun}$, D##@)

August 1 Model

Yield = f (Constant, Trend, $Rain_{OctMy}$, $Plnt_{May}$, $Temp_{May}$, $Temp_{Jun}$, Hot_{Jun}, $Rain_{Jun}$, $Temp_{EJly}$, $Rain_{EJly}$, $Temp_{MJly}$, $Rain_{MJly}$, Silk*, $Temp_{LJly}$, $Rain_{LJly}$, $HotDry_{Jly}$, Wet_{Jly}, D##@)

September 1 Model

Yield = f (Constant, Trend, $Rain_{OctMy}$, $Plnt_{May}$, $Temp_{May}$, $Temp_{Jun}$, Hot_{Jun}, $Rain_{Jun}$, $Temp_{EJly}$, $Rain_{EJly}$, $Temp_{MJly}$, $Rain_{MJly}$, Silk*, $Temp_{LJly}$, $Rain_{LJly}$, $HotDry_{Jly}$, Wet_{Jly}, $Temp_{EAgst}$, $Rain_{EAgst}$, $Temp_{LAgst}$, $Rain_{LAgst}$, Wet_{Aug}, D##@)

October 1 Model

Yield = f (Constant, Trend, $Rain_{OctMy}$, $Plnt_{May}$, $Temp_{May}$, $Temp_{Jun}$, Hot_{Jun}, $Rain_{Jun}$, $Temp_{EJly}$, $Rain_{EJly}$, $Temp_{MJly}$, $Rain_{MJly}$, Silk*, $Temp_{LJly}$, $Rain_{LJly}$, $HotDry_{Jly}$, Wet_{Jly}, $Temp_{EAgst}$, $Rain_{EAgst}$, $Temp_{LAgst}$, $Rain_{LAgst}$, Wet_{Aug}, $Temp_{Sep}$, $Rain_{Sep}$, $HotDry_{Sep}$, D##@)

Table 2. Definition of Crop-Weather Model Variables

1. Trend: Linear time trend, representing technical change
2. RainOctMy: Total precip. from October (previous year) to May (current year)
3. PlntMay: Percent of corn planted by mid-May (state level)
4. RainJun: Total June monthly precipitation
5. RainEJly: Total July 1 to 10 precipitation
6. RainMJly: Total July 11th to 20th precipitation
7. RainLJly: Total July 21 to 31 precipitation
8. RainEAgst: Total August 1 to 10 precip.
9. RainLAgst: Total August 11th to 31 precip.
10. RainSep: Total September monthly precipitation
11. TempMay: Avg May temperature
12. TempJun: Avg June temperature
13. TempEJly: Avg July 1 to 10 temperature
14. TempMJly: Avg July 11th to 20th temperature
15. Silk*: Date when corn is 75% Silked in Iowa (Iowa Crop-Weather Models only)
16. TempLJly: Avg July 21 to 31 temperature
17. TempEAgst: Avg August 1 to 10 temperature
18. TempLAgst: Avg August 11th to 31 temperature
19. TempSep: Avg September temperature
20. D##@ : Crop reporting district dummy variables, identified by region (##) and CRD (@) where @ is a letter representing the specific district

<u>Aridity Indices</u>

21. HotJun: Affect of June temps 1.5 normalized units (nzdu's) above average

Given: DHot = (If TempJun > 1.5, then = 1, otherwise = 0)

=> HotJun = DHot * TempJun

22. HotDryJly: Affect of combined high temps (>.5 nzdu's above avg) and low precip (>.5 nzdu's below avg) for July 11th to 31, intramonthly data

Given: DHot = (If [(TempMJly + TempLJly) / 2] > 0.5, then = 1, if not = 0)

DDry = (If [(RainMJly + RainLJly) / 2] < -0.5, then =1, if not = 0)

=> *HotDryJly = DHot * DDry * [(TempMJly + TempLJly) / 2] * (-1) * [(RainMJly + RainLJly) / 2]*

23. WetJly: Effect of precip >.5 nzdu's above avg for July 11-31, intramonthly data

Given: DWet = (If [(RainMJly + RainLJly) / 2] > 0.5, then = 1, if not = 0)

=> *WetJly = DWet * [(RainMJly + RainLJly) / 2]*

24. WetAgst: Effect of avg total precip for August 1-10 and August 11-31 being >.5 nzdu's above avg, intramonthly data

Given: DWet = (If [(RainEAgst + RainLAgst) / 2] > 0.5, then = 1, if not = 0)

=> *WetAgst = DWet * [(RainEAgst + RainLAgst) / 2]*

35. HotDrySep: Effect of combined high temps (>.75 nzdu's above avg) and low precip (>.75 nzdu's below avg)

Given: DHot = (If TempSep > 0.75, then = 1, if not = 0)

DDry = (If RainSep < -0.75, then = 1, if not = 0)

=> *HotDrySep = DHot * DDry * TempSep * (-1) * RainSep*

Table 3. Iowa Monthly Crop-Weather Models, 1972-1991 and 1972-1993

Variable	Iowa: 1972-1991				Iowa: 1972-1993			
	July 1	Aug. 1	Sept. 1	Oct.1	July 1	Aug. 1	Sept. 1	Oct.1
Constant	** 84.72	** 88.63	** 89.96	** 91.47	** 92.65	** 100.34	** 103.99	** 100.01
Trend	** 1.86	** 1.73	** 1.66	** 1.50	** 1.07	** 0.76	**0.64	** 0.78
RainOctMy	-2.34	* -2.42	-2.10	-1.46	0.11	0.31	-1.83	-0.98
PlntMay	** -6.41	* -4.11	** -4.76	* -3.34	1.52	1.05	-0.65	* -3.73
TempMay	1.75	** -4.64	** -8.29	** -8.86	-1.48	** -8.36	** -10.59	** -11.79
TempJun	-2.66	** -7.71	** -5.55	** -5.30	0.16	* -2.36	-1.86	** -6.41
HotJun	* -7.57	-0.93	* 1.21	2.19	* -7.71	-1.17	-1.05	2.69
RainJun	1.47	1.88	* 2.02	* 3.20	* -3.08	0.66	1.75	* 2.60
TempEJly		** -6.16	** -6.96	** -6.02		** -6.51	** -5.18	** -4.98
RainEJly		** 6.91	** 6.09	** 4.80		* 2.60	** 2.85	* 2.05
TempMJly		** -4.20	** -5.31	** -6.90		** -6.18	** -7.33	** -7.88
RainMJly.		1.56	1.46	0.66		* 1.96	* 3.23	-0.53
Silk*		** -11.78	** -10.42	** -9.30		** -15.02	** -11.75	** -13.19
TempLJly		- 1.00	-1.31	-1.03		* -2.48	0.72	0.54
RainLJly		** 4.96	* 3.40	1.66		* 2.95	* 4.15	-0.94
HotDryJly		** -11.35	* -7.40	* -7.36		** -12.46	* -8.57	*-10.26
WetJly		-1.69	-3.28	-0.37		-7.43	* -9.18	0.16
TempEAgst			-0.56	-0.85			* 3.03	* 2.60
RainEAgst			* 3.17	* 4.14			** 7.07	** 6.14
TempLAgst			** -4.77	** -5.60			**-6.42	** -8.75
RainLAgst			* 4.59	** 4.62			2.39	2.32
WetAgst			-4.92	-5.96			**-3.42	** -10.14
TempSep				* 2.78				** 6.70
RainSep				1.16				* 2.02
HotDrySep				-6.33				3.51
DNW	8.38	6.08	* 6.81	* 7.01	8.10	3.98	2.17	4.32
DNCent	**15.59	** 12.62	** 12.82	** 12.85	* 14.40	* 10.56	** 11.37	** 10.10
DNE	*11.61	* 9.30	** 9.54	** 9.43	* 10.76	* 7.75	* 8.94	* 7.35
DWCent	2.58	.64	0.98	1.00	3.00	-0.11	1.24	0.14
DCent	* 12.54	* 9.30	** 9.37	** 9.57	* 12.14	* 9.12	* 10.21	** 2.75
DECent	7.30	* 7.49	* 7.67	* 7.55	6.92	* 7.69	* 8.90	* 2.26
DSW	-1.29	-1.85	-1.33	-1.16	-0.81	-2.29	-0.59	-1.95
DSCent	* -11.04	** -11.18	** -11.12	** -11.23	* -11.46	** -11.92	** -10.85	** -11.88
Constant = CRD#	DSE	DSE	DSE	DSE	DSE	DSE	DSE	DSE
R^2	0.31	0.74	0.79	0.80	0.20	0.64	0.70	0.78
Adjusted R^2	0.24	0.69	0.75	0.76	0.13	0.59	0.65	0.73
Standard Error	18.72	11.88	10.73	10.56	21.39	14.80	13.52	11.88

Table 4. Illinois and Indiana Monthly Crop-Weather Models, 1972-1991

Variable	Illinois: 1972-1991				Indiana: 1972-1991			
	July 1	Aug. 1	Sept. 1	Oct. 1	July 1	Aug. 1	Sept. 1	Oct. 1
Constant	** 72.52	** 79.15	** 82.19	** 82.67	** 78.87	** 86.39	** 87.44	** 87.87
Trend	** 1.49	** 1.09	** 0.80	** 0.89	** 1.86	** 1.56	** 1.41	** 1.35
RainOctMy	* -4.04	-1.66	-1.20	-1.30	** -6.58	** -3.20	** -3.28	** -3.26
PlntMay	* -4.31	2.43	0.52	0.53	-0.17	0.41	0.22	1.26
TempMay	* 4.51	** 4.33	** 4.67	* 4.29	-1.83	0.54	1.14	0.26
TempJun	-1.64	-1.67	* 2.14	1.99	* -3.32	** -5.49	** -3.29	** -3.17
HotJun	-4.38	2.20	1.90	1.96	*11.10	* 5.85	* 6.08	* 5.86
RainJun	** 4.74	** 5.57	** 4.74	** 5.53	** 6.65	** 4.02	** 2.80	** 2.89
TempEJly		** -6.49	** -5.40	** -5.64		** -3.46	** -3.20	** -2.81
RainEJly		* 0.68	* 0.63	* 0.55		** 6.63	** 5.28	** 4.96
TempMJly		** -7.58	** -7.16	** -8.61		** -6.65	** -5.51	** -6.91
RainMJly		** 6.31	** 5.80	** 5.76		** 8.87	** 7.32	** 6.40
TempLJly		-0.90	-2.12	-1.59		-0.79	-0.08	-0.11
RainLJly		** 8.15	** 5.62	** 6.86		** 8.97	** 8.27	** 7.41
HotDryJly		-4.71	* -7.78	* 8.65		* -6.88	-3.10	-1.56
WetJly		** -8.97	* -8.63	** -9.54		** -18.82	** -17.77	** -16.20
TempEAgst			* -3.23	* -2.28			* -2.26	-1.41
RainEAgst			-0.26	1.43			-0.80	-0.08
TempLAgst			** -9.19	** -8.79			** -4.03	** -4.39
RainLAgst			* 2.94	* 3.49			* 2.13	* 2.26
WetAgst			-3.25	-6.21			-0.28	-1.51
TempSep				-0.33				1.58
RainSep				** 3.59				* 1.93
HotDrySep				-7.57				0.84
D(IL) / D(IN)								
DNW / DNW	** 25.19	** 24.93	** 23.81	** 24.10	* 8.43	** 8.31	** 8.38	** 8.47
DNE / DNCent	** 25.07	** 24.41	** 24.49	** 24.47	5.87	* 6.79	** 6.79	** 6.89
DNWCent / DNE	** 23.54	** 23.87	** 22.09	** 21.35	2.07	2.79	3.08	3.24
DNCent / DWCent	** 31.60	** 31.96	** 29.94	** 29.32	* 11.98	** 14.09	** 13.99	** 13.93
DNECent / DCent	** 26.64	** 26.22	** 25.27	** 24.67	** 15.71	** 16.49	** 16.57	** 16.65
DSWCent / DECent	** 28.80	** 28.37	** 27.22	** 26.96	6.56	** 8.23	** 8.19	** 8.19
DSECent / DSW	** 20.65	** 20.59	** 20.12	** 20.20	7.35	** 9.74	** 9.53	** 9.31
DSW / DSCent	-2.56	0.62	* 11.66	* 8.44	-3.34	-2.27	-2.31	-2.37
Constant = CRD#	DSE	DSE	DSE	DSE	DSE	DSE	DSE	DSE
R^2	0.38	0.74	0.82	0.84	0.40	0.82	0.85	0.85
Adjusted R^2	0.32	0.70	0.79	0.81	0.34	0.79	0.82	0.82
Standard Error	18.95	12.48	10.61	10.11	15.70	8.87	8.19	8.11

Table 5. Normalized Weather Conditions Averaged Across Crop Reporting Districts for Iowa, Illinois and Indiana for 1992-1994

State	Year	OMP	JNT	JNP	JLT	JLP	Silk*	AGT	AGP	SPT	SPP
Iowa	1992	0.28	-1.00	-1.20	-3.10	2.35	0.40	-2.10	-0.80	-1.00	0.98
	1993	1.02	-1.20	2.04	-1.30	3.51	1.68	0.06	1.51	-2.50	0.08
	1994	-1.51	0.63	1.30	-1.70	0.20	1.12	-0.90	-0.30	0.78	-0.10
	92-94	*-0.07*	*-0.54*	*0.70*	*-2.04*	*2.02*	*0.87*	*-0.99*	*0.14*	*-0.91*	*0.33*
Illinois	1992	-0.64	-1.20	-1.10	-1.30	1.89	NA	-1.90	-0.90	-0.80	0.82
	1993	0.57	-0.49	1.73	0.50	1.35	NA	0.56	0.55	-1.88	2.51
	1994	-0.64	0.95	0.43	-1.10	-0.60	NA	-1.00	0.08	-0.30	-0.50
	92-94	*-0.24*	*-0.25*	*0.34*	*-0.64*	*0.90*		*-0.80*	*-0.10*	*-0.99*	*0.95*
Indiana	1992	-0.78	-1.80	-0.40	-1.20	2.68	NA	-2.10	-0.60	-0.80	0.89
	1993	0.14	-0.60	1.31	0.85	0.25	NA	0.50	0.02	-1.70	1.78
	1994	-0.45	1.19	0.41	-0.50	0.15	NA	-1.20	-0.20	-0.30	-0.40
	92-94	*-0.36*	*-0.40*	*0.42*	*-0.28*	*1.03*		*-0.93*	*-0.27*	*-0.90*	*0.76*

Table 6, 1992-94 Iowa Corn Yield Forecast Probabilities

	1992 (1972-1991 Model)				1993 (1972-1991 Model)				1994 (1972-1993 Model)			
	7/1	8/1	9/1	10/1	7/1	8/1	9/1	10/1	7/1	8/1	9/1	10/1
Forecast YLD	123.8	154.3	158.1	159.2	148.3	156.2	158.9	150.0	121.6	150.5	155.0	154.8
Actual YLD	147.0	147.0	147.0	147.0	80.0	80.0	80.0	80.0	152.0	152.0	152.0	152.0
Forecast Error	6.91	4.82	4.44	4.66	7.16	5.01	4.91	5.90	7.85	5.64	5.24	4.68
Cumulative %												
0.5%	106.0	141.9	146.7	147.2	129.8	143.4	146.3	134.8	101.4	135.9	141.5	142.7
2.5%	110.3	144.8	149.4	150.1	134.2	146.4	149.3	138.4	106.2	139.4	144.7	145.6
5.0%	112.4	146.3	150.8	151.6	136.2	148.0	150.9	140.3	108.7	141.2	146.4	147.1
17.5%	117.2	149.7	153.9	154.8	141.5	151.5	154.3	144.4	114.1	145.1	150.0	150.3
25.0%	119.2	151.0	155.1	156.1	143.4	152.9	155.6	146.0	116.3	146.7	151.5	151.6
50.0%	123.8	154.3	158.1	159.2	148.3	156.2	158.9	150.0	121.6	150.5	155.0	154.8
75.0%	128.5	157.5	161.1	162.4	153.1	159.6	162.3	153.9	126.9	154.3	158.5	157.9
83.5%	130.4	158.9	162.4	163.7	155.1	161.0	163.6	155.6	129.1	155.8	160.0	159.2
95.0%	135.2	162.2	165.4	166.9	160.0	164.5	167.0	159.7	134.5	159.7	163.6	162.5
97.5%	137.4	163.7	166.8	168.4	162.3	166.1	168.6	161.5	137.0	161.5	165.3	163.9
99.5%	141.6	166.7	169.6	171.2	166.7	169.1	171.6	165.1	141.8	165.0	168.5	166.8

Table 7, 1992-94 Illinois Corn Yield Forecast Probabilities

	1992 (1972-1991 Model)				1993 (1972-1991 Model)				1994 (1972-1991 Model)			
	7/1	8/1	9/1	10/1	7/1	8/1	9/1	10/1	7/1	8/1	9/1	10/1
Forecast YLD	121.0	143.3	149.2	151.4	141.5	136.2	122.2	132.2	131.4	124.3	138.3	137.0
Actual YLD	149.0	149.0	149.0	149.0	130.0	130.0	130.0	130.0	156.0	156.0	156.0	156.0
Forecast Error	7.10	5.08	4.53	4.38	7.37	5.07	4.64	4.78	7.03	4.72	4.13	3.95
Cumulative %												
0.5%	102.7	130.2	137.5	140.2	122.5	123.1	110.3	119.8	113.3	112.1	127.6	126.8
2.5%	107.1	133.3	140.3	142.9	127.0	126.3	113.1	122.8	117.7	115.0	130.2	129.2
5.0%	109.3	134.9	141.7	144.2	129.4	127.9	114.6	124.3	119.9	116.5	131.5	130.5
17.5%	114.3	138.4	144.9	147.0	134.5	131.4	117.8	127.6	124.7	119.6	134.3	133.2
25.0%	116.2	139.8	146.1	148.5	136.5	132.8	119.1	128.9	126.7	121.1	135.3	134.3
50.0%	121.0	143.3	149.2	151.4	141.5	136.2	122.2	132.2	131.4	124.3	138.3	137.0
75.0%	125.8	146.7	152.2	154.4	146.5	139.6	125.4	135.4	136.2	127.4	141.0	139.6
83.5%	127.8	148.1	153.5	155.6	148.5	141.0	126.7	136.7	138.1	128.8	142.2	140.7
95.0%	132.7	151.6	156.6	158.6	153.6	144.5	129.9	140.0	143.0	132.0	145.1	143.5
97.5%	134.9	153.2	158.1	160.0	155.9	146.1	131.3	141.5	145.2	133.5	146.4	144.7
99.5%	139.3	156.4	160.8	162.7	160.5	149.3	134.2	144.5	149.5	136.4	148.9	147.1

Table 8, 1992-94 Indiana Corn Yield Forecast Probabilities

	1992 (1972-1991 Model)				1993 (1972-1991 Model)				1994 (1972-1991 Model)			
	7/1	8/1	9/1	10/1	7/1	8/1	9/1	10/1	7/1	8/1	9/1	10/1
Forecast YLD	134.0	151.6	149.6	149.5	138.2	140.8	131.5	129.3	136.0	132.2	135.7	135.5
Actual YLD	147.0	147.0	147.0	147.0	132.0	132.0	132.0	132.0	144.0	144.0	144.0	144.0
Forecast Error	6.21	3.72	3.51	3.54	6.30	3.62	3.49	3.70	6.17	3.62	3.40	3.37
Cumulative %												
0.5%	118.0	142.0	140.6	140.4	122.0	131.4	122.5	119.8	120.1	122.9	126.9	126.8
2.5%	121.9	144.3	142.7	142.5	125.9	133.7	124.7	122.1	123.9	125.1	129.0	128.9
5.0%	123.8	145.5	143.8	143.7	127.9	134.8	125.8	123.2	125.8	126.3	130.1	130.0
17.5%	128.1	148.1	146.3	146.1	132.2	137.3	128.2	125.8	130.1	128.8	132.4	132.3
25.0%	129.9	149.1	147.2	147.1	134.0	138.3	129.1	126.8	131.8	129.8	133.4	133.2
50.0%	134.0	151.6	149.6	149.5	138.2	140.8	131.5	129.3	136.0	132.2	135.7	135.5
75.0%	138.2	154.1	152.0	151.9	142.5	143.2	133.9	131.8	140.2	134.7	138.0	137.8
83.5%	140.0	155.1	152.9	152.9	144.2	144.2	134.8	132.9	141.9	135.7	138.9	138.7
95.0%	144.3	157.7	155.4	155.3	148.6	146.7	137.3	135.4	146.0	138.2	141.3	141.1
97.5%	146.2	158.9	156.5	156.4	150.6	147.9	138.4	136.6	148.1	139.3	142.3	142.1
99.5%	150.0	161.2	158.6	158.6	154.5	150.1	140.5	138.9	151.9	141.6	144.4	144.2