KY

295

Staff Papers

DEPARTMENT OF
AGRICULTURAL ECONOMICS
University of Kentucky • College of Agriculture
Lexington 40546

AAEA paper presented at its annual meetings,
Manhattan, KS, Aug 4-7, 1991

Staff Paper #295 June, 1991

COMBINING LONGER SERIES OF WEATHER AND CLIMATE DATA WITH SHORT SERIES OF YIELD DATA TO ENHANCE INFORMATION ABOUT YIELD RISK

Clemen Gonzales, Haiping Luo, Jerry Skees, and Mary Marchant

Clemen Gonzales and Haiping Luo are research assistants, Jerry Skees is a professor, and Mary Marchant is an assistant professor in the Department of Agricultural Economics, University of Kentucky Paper to be presented at the annual meeting of the American Agricultural Economics Association, Manhattan, KS , Aug 4-7, 1991 Staff papers are published without formal review.

COMBINING LONGER SERIES OF WEATHER AND CLIMATE DATA WITH SHORT SERIES OF YIELD DATA TO ENHANCE INFORMATION ABOUT YIELD RISK

ABSTRACT

This paper presents a methodology for adding longer series of weather data to short series on yield data for the purpose of improving knowledge about crop yield risk. Findings demonstrate that the weather during the past 33 years provide yield forecast with significantly less risk than the past 95 years.

COMBINING LONGER SERIES OF WEATHER AND CLIMATE DATA WITH SHORT SERIES OF YIELD DATA TO ENHANCE INFORMATION ABOUT YIELD RISK

Problem Statement and Relevance of Study

Improved understanding of crop yield risk is important for a variety of decisions Much of this understanding needs to be focused on the probability of yield risk with particular emphasis toward low yields (1) Farm managers must understand the likelihood of catastrophic yields in order to make a variety of risk management decisions. (Buzby, et al. (1990) demonstrated that farmers tend to forget their lowest yields) (2) Extension specialists working with risk management decisions must understand more about catastrophic yields (3) Federal policymakers must understand more about catastrophic yields for decisions about both Federal crop insurance and possible disaster assistance programs (4) The private insurance sector also needs improved information. This paper presents a methodology designed to fulfill these needs, i.e , to add to the understanding of catastrophic yield risk

Forecasting models that use weather and climate data are also important for a variety of uses. The United States Department of Agriculture (USDA) official estimates of local yields are developed by the National Agricultural Statistical Service (NASS) Typically, NASS data are not available until well past harvest (for fall crops, data are generally available in late spring). Weather and climate data are available on a more timely basis Therefore, forecasting models can be used to provide early estimates of local yields. There is a growing interest in providing crop insurance based on an area yield Farmers would be paid anytime the area (for example, a county) yield drops below a specified level (Miranda (1991), Barnaby and Skees (1990)) A major problem with this concept is the ability to make timely payments Crop forecasting models would provide the opportunity to make partial payments in a timely fashion In addition, the procedures presented in this paper are important for establishing rates and coverage levels for an insurance program based on area loss.

A major problem with any attempt to understand more about the probability density function for crop yields is the lack of long time series data on crop yields Short time series data simply cannot provide reliable estimates of yield risk However, longer time series are available on weather and climate data. This study presents a methodology for combining information from short time series data on yields with longer time series of weather and climate data for the purpose of acquiring a better understanding of catastrophic yield risk This paper develops models that incorporate weather risk elements to predict yields These models are then subjected to the longer data sets on weather in order to feed through the weather effects on yields and improve estimates of the yield risk inherent in weather events.

Research Objectives

The overall objective of this research is to increase decision makers' understanding of catastrophic yield risk due to weather events resulting in more informed management decisions. Specific research objectives are (1) to estimate a weather-based production function for corn in 9 climate divisions in Illinois using data from 1956-1988[1], (2) using the above estimated production function, to predict (ex-post) corn yields for two periods. (i) 1895 - 1988 and (ii) 1956 - 1988, and (3) to measure catastrophic yield risk by constructing cumulative distribution functions (CDFs) for these two time periods

Literature Review

Climate and crop yield relationships have been investigated using different perspectives and various methods. Climate has been viewed as (1) a production input, (2) a source of risk, and (3)

[1] Climate divisions and crop reporting districts are the same in Illinois.

an outcome predictor, depending on research objectives. Studies that treat climate as a production input for corn include Thompson (1969), Chang (1981), Offutt, et al. (1987) and French and Headley (1989). These studies used regression analysis and a gamut of indices to represent climate. Other studies used plant growth simulation models and/or dynamic programming to show input-output relationships between climate and crop yield [Reetz (1976), Mjelde (1985), and Mazzocco (1989)]

Investigations on climate as a source of risk dealt with downside climate-induced yield risk and variability. Parry and Carter (1985) assessed the risk of crop failure in Central England by constructing a 'risk-surface' map using accumulated temperature data. Studies that use various methods to deal with yield variability from rainfall and stochastic climatic variables for various cereals and countries are presented in Anderson and Hazell (1989).

Climate as an outcome predictor may be investigated by treating climate as a production input and/or as a risk source Crop yields can be predicted using expected values, probability distributions and prediction errors from stochastic climate variables. Researchers face the challenge of judging the reliability of climate scenarios as well as the predictive ability of their models. Liverman, et al. (1986) attempted to meet this challenge The authors developed a model, using climatic and environmental data, to predict grain corn yield in the North American Great Plains.

Overview

This study combines the three perspectives on climate and crop yield relationship. A production function is estimated using climate as a production input This function follows the conceptual scheme for sources of yield variability discussed in Anderson and Hazell (1989), i e , climate variables are uncontrollable from the farmers' point of view, but these should be included in the production function with other controllable inputs (e g , fertilizer) that contribute to yield variability.

From the estimated production function, corn yields were predicted and a CDF was derived to measure yield variability and risk caused by weather One expects that because of more theory-supported information, the risk measured by the production function approach provides more complete information to decision makers

Methodology and Data

The choice of the functional form of the production function was constrained by the small number of degrees of freedom available for this study (Fuss and McFadden (1978), pgs. 224-225)[2] Generally accepted for its theoretical properties and simplicity, the Cobb-Douglas production function and its underlying assumptions were adopted in this study Past studies and other sources of information guided the choice of explanatory variables for corn yield. Because of the similarity in weather effects among some or all of the 9 crop reporting districts, the possibility of contemporaneous correlation among error terms existed. Thus, to gain efficiency, the cross-section equations were estimated using SAS's seemingly unrelated regressions (SUR) procedure

The general specification of the function was

$$(1) \quad Ln(Y_{ij}) = \alpha + \sum_{k=1}^{6} Ln(X_{ij})_k \quad \textit{for each } i \textit{ and for each } j$$
$$\textit{where } i = 1,...,33 \ (1956\text{-}1988) \ \textit{and } j = 1,...,9 \ (\textit{climate divisions})$$

where Y_{ij} is the mean yield for crop reporting district 'j' in year 'i', α is the intercept, X1 is the temperature for July, X2 and X3 are rainfall in selected months, X4 is nitrogen application, X5 is the government set-aside policy, and X6 is a technological trend variable. Use of actual values for rainfall and temperature, instead of an index, to represent weather is widespread in the literature [e.g Babcock (1988), Byerlee and Anderson (1982), French, et al. (1985)] The

[2] The degrees of freedom equal 26, where the number of observations equal 33 for the years 1956-1988 and number of explanatory variables equal seven, including the intercept

coefficients for temperature and rainfall are expected to be negative and positive, respectively, from agronomic and plant physiology information Since a distinct positive trend was present, nitrogen application was developed as an index, reflecting deviations from the mean 1987 value Government policy on acreage set-aside was constructed to reflect percent of land used in production. A ratio of amount of land idled due to set-aside programs and total cropland for feed grains was used This variable was designed to test speculations that set-aside programs lead to (1) marginal lands being idled and (2) intensification, i.e., higher production rates, on the remaining area. A trend variable was used to represent technology changes over time. A priori expectations include positive signs for nitrogen and trend variable coefficients and a negative sign for the set-aside policy variable. Nitrogen and set-aside data were not available for climate divisions. Therefore, State aggregate data were used

Data on all variables except for nitrogen (Vroomen, 1989) and government set-aside (USDA, 1990), came from the Midwest Climate Center and National Agricultural Statistics Service. After estimating the production function, corn yields were predicted (ex-post) for two periods, (1) 1956-1988 and (2) 1895-1988, using coefficients for the weather variable with weather data, while setting the influence of all other variables at their respective mean values

Results and Discussion

Estimated Production Function

The estimated production function took the following form·

$$(2) \quad LN(\hat{Y}_i) = \hat{\alpha} + \hat{\beta}_{1i}\, LN(T7) + \hat{\beta}_{2i}\, LN(CR6) + \hat{\beta}_{3i}\, LN(R78) + \hat{\beta}_{4i}\, LN(DLSET) + \hat{\beta}_{5i}\, LN(NITR) + \hat{\beta}_{6i}\, LN(YR) \text{ for each } j$$

$$j = 1,\ldots,9 \text{ (climate district)} \quad i = 1,\ldots,33 \text{ (1956-1988)}$$

where Y is corn yield in bushels per acre, T7 is the temperature in July in degrees Fahrenheit, CR6 is cumulative rainfall from January to June, in inches, R78 is cumulative rainfall for July and August; DLSET is 1 - the ratio of the amount of land idled relative to total cropland for feedgrains, in percent; NITR is an index of residuals from detrended nitrogen data; YR is the year to reflect a technology trend over time.

The SUR parameter estimates for the 9 weather districts are summarized in Table 1 The SUR system has an adjusted R^2 of 0 676, compared with adjusted R^2 ranging from 0.635 to 0.802 for the ordinary least squares (OLS) estimations All SUR estimates were more efficient than OLS estimates, as shown by the variance-covariance matrices. All parameter estimates have the expected signs except for the set-aside variable (DLSET), which was insignificant at the 20% level for all districts except one Diagnostics were performed to test for autocorrelation, heteroskedasticity, and multicollinearity Though slight evidence was present for the existence of multicollinearity in two climate divisions, overall these types of problems were not significant

Temperature in July (T7), performed the best among the weather variables in all districts Since the coefficients presented in Table 1 are developed from a logarithmic production function, the coefficients are elasticities Thus, the coefficient for nitrogen is the marginal physical product (MPP) for nitrogen application These coefficients were consistently significant, with the expected sign, across the 9 climate districts For example, in crop reporting district 1, a 10 percent increase in nitrogen use results in a 3 4 percent increase in yield The trend variable suggests that technology (and other factors in the trend variable) causes higher yields as one goes from Northern to Southern Illinois (crop district 1 to crop district 9 in Figures 2 and 3). The presence of more marginal lands in districts 8 and 9 might be a factor in explaining this result.

Yield Variability, Risk and Skewness:

From the above estimated production function, corn yields were predicted for each crop reporting district 'j' and year 'i' using:

$$(3)\quad \hat{Y}_{ij} = (\hat{\alpha}_{ij} + \hat{\beta}_{4ij} + \hat{\beta}_{5ij} + \hat{\beta}_{6ij}) + \hat{\beta}_{1ij}\, LN(T7) + \hat{\beta}_{2ij}\, LN(CR6) + \hat{\beta}_{3ij}\, LN(R78)$$
$$j = (1,\ldots,9)\ (\textit{climate district})\ \textit{and}\ i = (1895,\ 1896,\ \ldots\ ,\ 1989)$$

The predicted corn yields and variability were influenced by the weather variables in the model and model measurement error These predicted yields were used to construct CDFs for two periods (1895 - 1988 and 1956 - 1988) for each district. The CDFs and other measures computed from the predicted corn yields, such as the mean, coefficient of variation and skewness, provide information on yield variability and risk Yield variability and risk due to weather were compared for the two periods after testing statistical significance

Figure 1 shows a plot of deviations from the mean predicted corn yields over 95 years for districts 1 and 9. The weather generated data is particularly troublesome when one focuses on the low yields prior to 1956 and when noting the extended period of low yields from 1925 to 1945. During this 20 year period, yields in climate division 1 were consistently below the mean. Table 2 shows measures of variability, risk and skewness for the 9 districts for the two periods Consistently, in all districts, the shorter time period has a higher mean yield and lower variability than the longer time period. Table 3 shows the years with the ten lowest yields generated from the weather data For example, upon comparing the lowest yield in the shorter time period with yields from the longer time period for crop district 2, eight yields in the longer period were even lower. This indicates that the "worst" year in recent history (1956-1988) is still "better" (for 8 years) when the longer term is examined (1895-1988)

The null hypothesis that the distributions for the two periods come from the same population distribution was tested (Cooper (1983), pg 320; and SAS Institute (1985): Inc., pg 607). For districts 1, 2, and 7, statistical tests for the two periods were significant at the ten percent level of significance, indicating that the null hypothesis can be rejected For each of these districts, the distribution for the longer time period had a lower mean and higher variance than the distribution for the shorter time period, indicating higher risk associated with the longer period. In general, all districts systematically exhibited similar results, although these results were not statistically significant at the ten percent significance level

To illustrate the measurement of yield risk, the CDFs for the two time periods in 8 districts are shown in Figures 2 and 3. The shorter period has a consistent stochastic dominance (first or second order) over the longer time period for all climate divisions The flatter left-side tail of the CDF for the longer time frame (1895 - 1988) indicates a higher probability for the downside climate-induced yield risk (i.e., the catastrophic yield risk is higher with the extended time frame). The yield risk can be affected by the skewness of the distribution. Table 2 shows that the distribution for both the long and short time periods tend to be negatively skewed, but Pearson's test of significance at 5% suggests that the distributions are normal for both periods

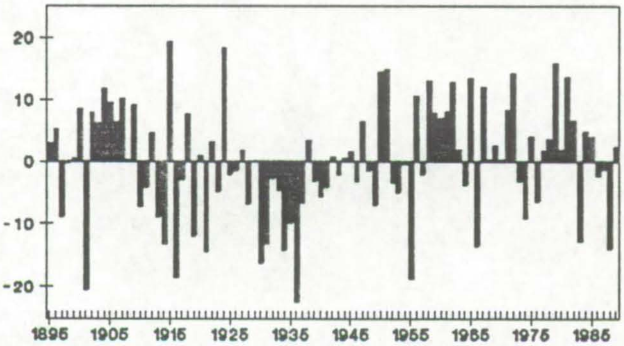
Conclusions

This study provides results regarding yield risk using weather events from two periods (1895 - 1988 and 1956 - 1988). The results and methods are promising. The results provide incentives to obtain more disaggregated climate data and to delve further into the plant physiology literature to learn more about weather events and timing For example, it is likely that different months of rainfall have different influences as one moves from Northern to Southern Illinois.

Results raise some concerns about the likelihood of low crop yields in years-to-come. In particular, weather events generate six or seven yields in the longer time period that are lower than any yields generated in the past 35 years There are also troublesome patterns that emerge in the autocorrelation of yields (see Figure 1) Patterns such as those between 1926 and 1938 suggest that we could have an extended period of low crop yields due to bad weather. These results alone should motivate further development and refinements of the methods presented in this paper.

This type of information can be useful for a variety of crop insurance programs. In particular, these types of models can contribute to design and implementation of a crop insurance program that is based on area loss Given that the variables in this model include weather data for several months prior to harvest, it is possible to develop forecasts that would be available to make area loss payments by harvest time. Further, since results demonstrate that the past 33 years were better weather years than the previous 95, this suggests that current methods which use short series of NASS yield data to establish rates and coverage will result in actuarial problems.

Figure 1
Deviations from the Mean of Predicted Yields for Climate Division 1



Deviations from the Mean of Predicted Yields for Climate Division 9

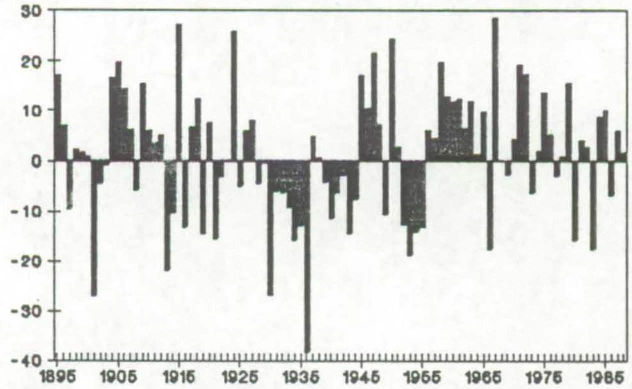


Fig. 2 -- CDFs for Two Time Periods

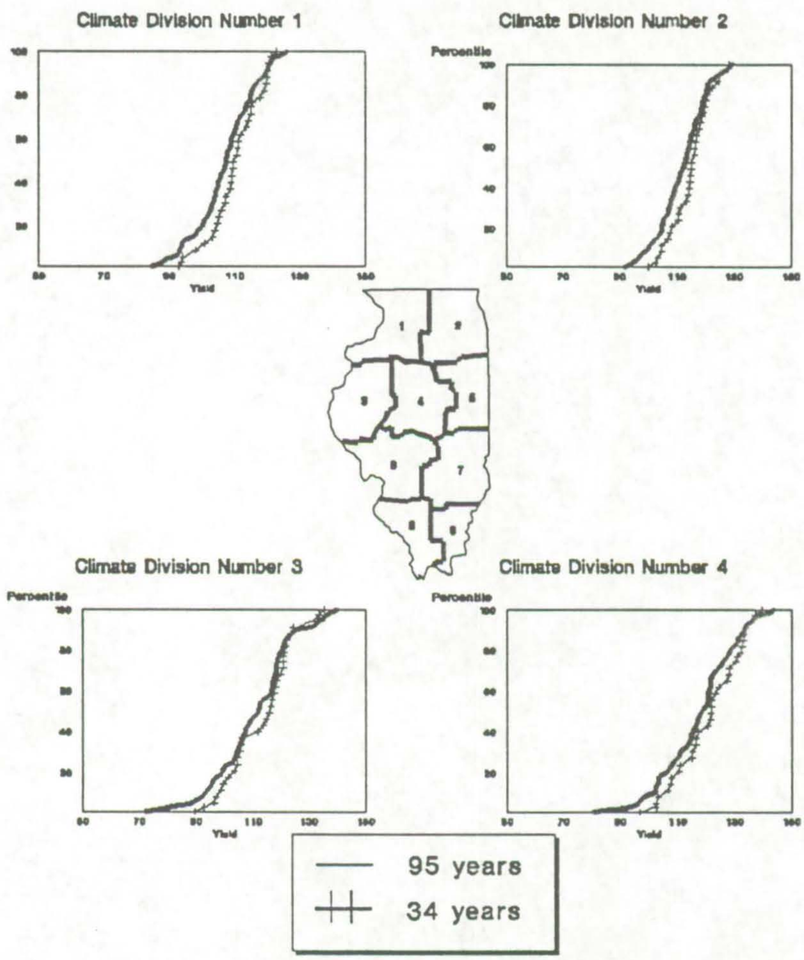


Fig. 3 -- CDFs for Two Time Periods

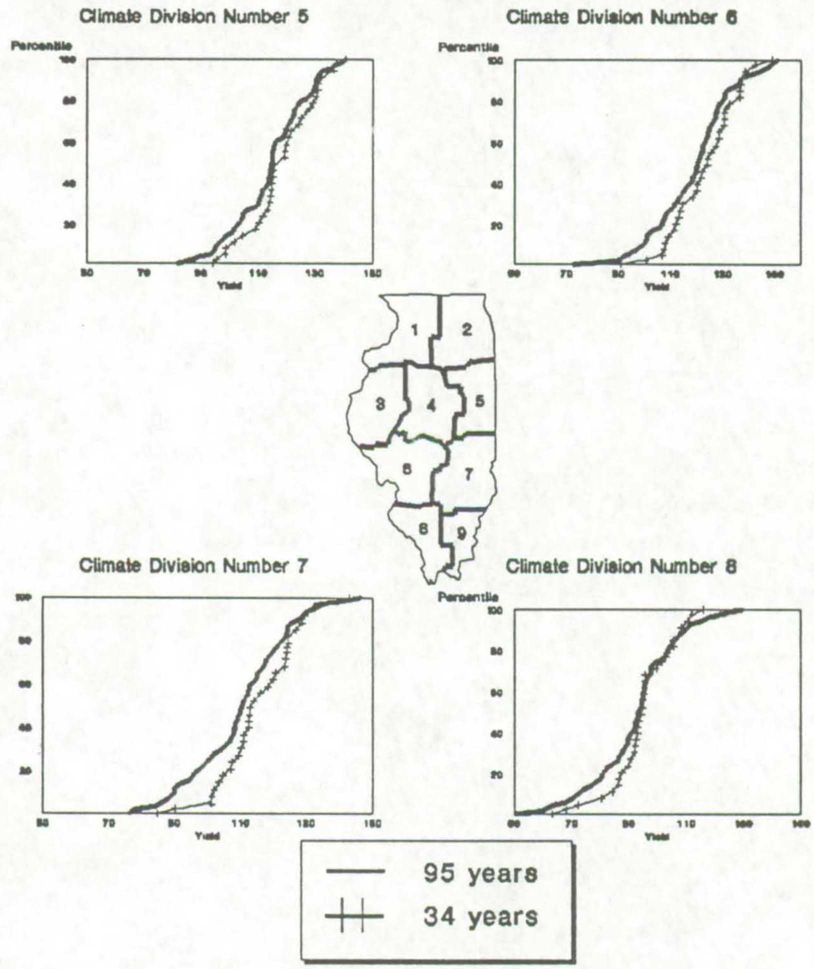


Table 1. Parameter Estimates from SUR Estimation for 9 Illinois Crop Districts

Independent Variables	Crop Reporting District 1	2	3	4	5	6	7	8	9
Intercept	8.210	5.989	11.791	11.933	11.995	9.483	9.111	12.293	13.410
	(2.692)[4]	(2.606)	(3.008)	(2.926)	(3.002)	(2.493)	(3.144)	(4.414)	(4.551)
Prob> |T|[3]	0.0025	0.015	0.0003	0.0002	0.00025	0.0004	0.0037	0.0049	0.003
T7	-2.252	-1.966	-3.268	-3.128	-3.264	-2.975	-2.988	-3.944	-4.100
	(0.648)	(0.618)	(0.103)	(0.676)	(0.109)	(0.568)	(0.735)	(1.037)	(1.074)
Prob> |T|	0.0009	0.002	0.00005	0.00005	0.00005	0.00005	0.0002	0.0004	0.0004
CUM R6	0.0266	0.009	-0.015	0.020	0.029	-0.045	-0.098	0.121	-0.025
	(0.045)	(0.046)	(0.057)	(0.051)	(0.058)	(0.049)	(0.056)	(0.094)	(0.100)
Prob> |T|	0.281	0.422	0.397	0.349	0.310	0.183	0.046	0.104	0.402
R78	0.114	0.140	0.068	0.077	0.103	0.158	0.242	0.224	0.163
	(0.027)	(0.030)	(0.038)	(0.036)	(0.033)	(0.033)	(0.041)	(0.063)	(0.065)
Prob> |T|	0.0001	0.00005	0.042	0.022	0.002	0.00005	0.00005	0.0008	0.009
DLSET	0.151	0.129	0.189	0.127	0.039	0.134	0.021	0.004	0.138
	(0.117)	(0.109)	(0.134)	(0.128)	(0.128)	(0.108)	(0.127)	(0.148)	(0.163)
Prob> |T|	0.104	0.125	0.085	0.159	0.380	0.113	0.434	0.488	0.201
NITRO	0.344	0.257	0.387	0.315	0.408	0.271	0.360	0.059	0.222
	(0.160)	(0.150)	(0.181)	(0.173)	(0.171)	(0.146)	(0.169)	(0.196)	(0.218)
Prob> |T|	0.020	0.049	0.021	0.040	0.0125	0.036	0.043	0.382	0.158
YR	0.975	1.292	1.186	1.068	1.055	1.536	1.513	1.867	1.739
	(0.209)	(0.195)	(0.237)	(0.225)	(0.226)	(0.191)	(0.223)	(0.263)	(0.285)
Prob> |T|	0.0001	0.00005	0.00005	0.00005	0.00005	0.00005	0.00005	0.0000	0.00005

[3] The probability values that test for statistical significance use a one tailed t-test.

[4] Numbers in parentheses are standard error estimates.

Table 2. Yield Variability and Skewness for Periods 1956-88 and 1895-1988, by Crop District

Crop Reporting District	1956-1988 (n=33)			1895-1988 (n=95)		
	C.V	Mean	Skewness	C V.	Mean	Skewness
1	7.458	110 049	-0 449	8.488	107.135	-0.236
2	6 263	114 638	-0.093	7.404	111.719	-0.189
3	10 477	113 507	-0.199	12.339	110.791	-0.257
4	9.649	121.04	-0.328	10 814	117.596	-0.339
5	10.485	118 372	-0 234	11 248	114 661	-0.253
6	10 623	123 177	-0.330	12 581	118.993	-0.251
7	11 189	115 767	-0 158	13 219	110 328	-0 127
8	12 508	94 662	-0 721	15 818	92 474	-0.234
9	10.999	94.759	-0.339	14.102	90 781	-0.264

Table 3: Implied Worst 10 Weather Years by Climate Division (1895-1988)

Climate Divisions:

Yield Order	1	2	3	4	5	6	7	8	9
94th[5]	1936	1916	1936	1936	1936	1936	1936	1936	1936
93th	1901	1934	1934	1901	1901	1901	1901	1901	1901
92nd	1955	1901	1901	1916	1934	1983	1930	1930	1930
91st	1916	1936	1916	1934	1916	1930	1983	1934	1913
90th	1930	1955	1955	1955	1921	1934	1913	1983	1953
89th	1921	1930	1983	1983	1988	1913	1935	1980	1983
88th	1934	1946	1930	1921	1955	1916	1943	1913	1966
87th	1988	1921	1914	1914	1933	1914	1919	1954	1980
88th	1966	1966	1966	1930	1966	1953	1933	1953	1934
87th	1931	1931	1921	1919	1930	1933	1955	1966	1921
Years since 1955[6]	2	1	2	1	2	1	1	3	3

[5] Weather generated data is ordered from the minimum yield, e g , the 94[th] lowest yield out of 94 years.

[6] Number of years since 1955 that were among the lowest yields from the weather generated data

REFERENCES

Anderson, J R and P. Hazell, eds Variability in Grain Yields: Implications for Agricultural Research and Policy in Developing Countries Maryland: The Johns Hopkins University Press, 1989.

Babcock, B J. Weather Information as a Free Public Good Is More Always Better? Working Paper, Dept of Econ and Business, North Carolina State University, 1988.

Barnaby, G.A., and Jerry Skees. "Public Policy for Catastrophic Yield Risk An Alternative Crop Insurance Program" Choices Second Quarter(1990) 7-9

Buzby, J C , P.L. Kenkel, J R Skees, J W Pease and F J Benson "A Comparison of Subjective and Historical Yield Distributions with Implications for Multiple Peril Crop Insurance" Paper presented at the annual meeting of the AAEA, Vancouver, Canada, Aug. 1990

Byerlee, Derek and Jock R Anderson. "Risk, Utility and the Value of Information" Review of Marketing and Agricultural Economics 50(1982) 231-246

Chang, J.H. "Corn Yield in Relation to Photoperiod, Night Temperature and Solar Radiation." Agricultural Meteorology 24(1981) 253-62

Cooper, R A., and A.J Weekes Data, Models and Statistical Analysis New Jersey Barnes & Noble Books, 1983.

French, J.B , D. Schroder, and J.C. Headley "Role of Weather and Technology on Crop Yield Risk for Corn, Soybeans and Winter Wheat in Selected Regions of the United States" Department of Agricultural Economics Paper No. 1985-3, University of Missouri, 1985.

French, J.B. and J C Headley "Influence of Technology and Weather on the Variability in US Maize and Wheat Yields." Variability in Grain Yields, ed Anderson, J R. and P. Hazell, pp. 270-86 Maryland: Johns Hopkins University Press, 1989.

Fuss, M , and D McFadden, eds Production Economics A Dual Approach to Theory and Application, Vol. 1. Amsterdam North Holland Publishing Co., 1978

Liverman, D M., W.H Terjung, J T Hayes and L O. Mearns "Climatic Change and Grain Corn Yields in the North American Great Plains" Climate Change. 9(1986):327-47.

Mazzocco, M A. "Valuing Climate Forecasts for Midwestern Grain Producers." Ph.D thesis, University of Illinois at Urbana-Champaign, 1989

Miranda, Mario J "Area-Yield Crop Insurance Reconsidered" American Journal of Agricultural Economics 73(1991):233-242

Mjelde, James W. "Dynamic Programming Model of the Corn Production Decision Process with Stochastic Climate Forecasts" Ph D thesis, University of Illinois at Urbana-Champaign, 1985.

Offutt, S E , P Garcia, and M Pinar. "Technological Advance, Weather, and Crop Yield Behavior" North Central Journal of Agricultural Economics 9(1987) 49-63

Parry, M.L. and Carter T R "The Effects of Climatic Variations on Agricultural Risk" Climate Change 8(1985):95-110

Reetz, H.F. Jr. "Corn Crops! A Physiology-Based Simulation of the Corn Crop." Ph.D. thesis, Purdue University, 1976.

SAS Institute, Inc. SAS User's Gride· Statistics, Version 5th ed NC: SAS Institute, Inc., 1985.

Thompson, L.M. "Weather and Technology in the Production of Corn in the US Corn Belt" Agronomy Journal 61(1969) 453-56

US Department of Agriculture Summary Statistics for Feed Grains Washington, D.C.· ASCS, 1990.

Vroomen, Harry. Fertilizer Use and Price Statistics USDA, 1989.