2.4 Relationship Between Normalized Corn Yields and Monthly Rainfall for the Midwestern United States

Nancy Westcott*, Steven E. Hollinger, and Kenneth E. Kunkel
Illinois State Water Survey
Champaign, Illinois

1. INTRODUCTION

The summertime climate of the central United States is characterized by high spatial variability in precipitation, resulting in high spatial variability in soil moisture, crop stress, and crop yield. Monitoring rainfall conditions and identifying areas of potential crop damage from deficient or excessive rainfall can be problematic because the existing network of precipitation observations is not of sufficient spatial resolution to identify small-scale variations in precipitation. Further, delay in the availability of quality controlled cooperative rainfall data prevents timely monitoring of precipitation. This study examines the relationship between high resolution (15-km) multi-sensor rainfall estimates obtained in near-real time from the National Centers for Environmental Prediction (NCEP), and county-level corn yields for a 9-state region (Fig 1).

2. DATA AND ANALYSIS

Rainfall data was collected from three sources for this study: 1) hourly gridded precipitation estimates based upon hourly gages and the WSR-88D radars obtained in near real-time from the National Centers for Environmental Prediction (NCEP), 2) daily quality-controlled National Weather Service (NWS) cooperative raingage (QC_Coop) data from NCDC, and 3) daily real-time NWS cooperative raingage (RT_Coop) data captured by the Midwestern Regional Climate Center. The analysis period covers the summers of 1997-1999 and 2001-2002. RT_Coop data, however, were archived beginning in 2001.

2.1 Gridded Precipitation Fields

Gridded (15-km) hourly multi-sensor precipitation estimates have been obtained in near real-time from NCEP since March 1997. These estimates are based upon a composite of data from the WSR-88D radars and upon hourly rain gage observations from the Hydrometeorological Automated Data System (HADS). The gridded data are downloaded on a daily basis, and are summed over the 24-hour period, (0700 – 0700 LST). County averages are computed for the radar (unadjusted), gage, and multi-sensor fields (currently created from the unadjusted radar and raingage fields) for the central Midwest, and are stored for analysis. On average, there are seven grids points per county.

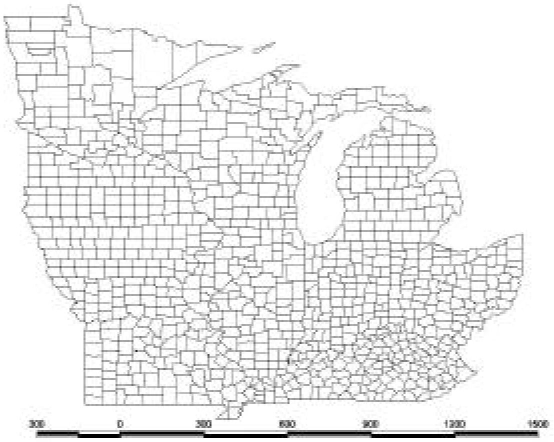

Figure 1. Counties included in study.

Approximately 30 radars and about 800 gages are located within the analysis region that are employed in the multi-sensor estimates. The gridded multi-sensor field was developed to account for spatial inhomogeneities in the rainfall estimates, under the assumption that the radar mean bias error has been removed (Fulton et al., 1998). A local adjustment is made to the rainfall estimates using a multivariate optimal estimation procedure that incorporates point gage data into the rainfall analysis. The weights for radar and gage estimates at each grid point are determined so that their linear combination minimizes the expected error variance of the estimate. A decreasing weight is placed on the gage as the distance increases from the gage, and an increasing weight is placed on the radar estimate (Fulton et al., 1998; Seo, 1998). This technique detailed by Seo (1998) attempts to account for within-storm variability

*Corresponding authors address: Nancy E. Westcott, Atmospheric Environment Section, Illinois State Water Survey, 2204 Griffith Dr, Champaign, IL 61820; e-mail: nan@uiuc.edu.

of rainfall and for variability due to the fractional coverage of rainfall (i.e. one instrument reports rainfall where the other does not).

During the winter of 2002, the NOAA's Office of Hydrology (OH) in conjunction with the NWS River Forecast Centers, implemented a Stage III/IV multi-sensor precipitation estimate (MPE) algorithm that includes provisions for quality-controlling gage data and incorporates a new method of bias-correction computation (Seo and Breidenbach, 2002). Preliminary examination of county-averaged monthly rainfall values for the summer of 2002 indicates that the newer MPE precipitation algorithm does provide improved rainfall estimates over the older multi-sensor algorithm. This improvement is due, at least in part, to the elimination of spurious gridded gage amounts.

2.2 Cooperative Raingage Estimates

The QC_Coop daily data, available some three months after the fact, were obtained from NCDC. The gages employed are the standard 8-inch non-recording gages (SNRG) or the Fisher Porter (FP) weighing bucket gages. Only gages having 90 percent or more data reported during the period were used. About 775 of the 858 counties in the study region contained at least one quality-controlled raingage. There were approximately 1500 cooperative gages reporting during this period. This resulted in an average of about two gages per county in counties with gages, or about one gage per 800 km^2. About 15% of these gages were part of the real-time gridded gage data set. Reporting times of the cooperative gages vary, with some at midnight, many between 05:00 and 9:00 LST in the morning, and a few at other times of the day. Observation times reported by the cooperative observers can vary from day to day. All gages, regardless of observation time, were employed in computing the QC_Coop monthly county averages.

The RT_Coop daily data were collected daily during the summer of 2001. About 725 gages were employed in this analysis, with about 66% reporting between 05:00 and 09:00 LST, 19% between 22:00 and 04:00 LST, and 15% reporting at other times. Those with observation times between 05:00 and 09:00 LST were largely from cooperative observers reporting SNRG amounts. Only about 530 of the 858 counties (61 %) had at least one raingage reporting during the summer of 2001, with an average of about 1.3 gages per county or about one gage per 1250 km^2. About 16% of these gages were part of the real-time gridded gage data set. ASOS/AWOS gages were not included in this sample.

For this study, monthly values were examined because of our primary interest in applications on climate timescales; also, this largely avoids issues related to differences in observation times of the cooperative gages. For all rainfall estimates, average monthly totals of 0.0 mm were eliminated from the analysis.

2.4 Crop Yields

County crop yields were obtained for the 9 state region for corn from the United States Department of Agriculture National Agricultural Statistics Service (http://www.usda.gov/nass/pubs/histdata.htm). The county corn yields were normalized each year using the preceding 5-year county average yield, because of the impact of differences in farming practices and soil and climate characteristic on crop growth.

2.5 Soil Characteristics

Soil characteristics (Table 1) were obtained from the Midwestern Regional Climate Center Soil database (Hollinger, 1995). These county characteristics are weighted averages of soil variables based upon the State Soil Geographic (STATSGO) database (Soil Conservation Service, 1993). The averages were obtained by weighting each soil variable by the fraction of total county area that each soil component occupied. Only the arable farmland (land easily tilled and comprising more than 0.01 percent of the county) was included in the average computation.

Table 1. Soil parameters included in the analysis.

Weighted Average Characteristics
Average slope of the soil surface, cm
Average depth of the water table, cm
Average depth of the bed rock, cm
Average root depth, cm
Average drainage class
For 0-50 cm soil layer
Sand content, %
Silt content , %
Clay content, %
Available (to roots) water content, cm / 50 cm
Bulk density (mass soil, air, water / unit vol), Mg/m3
Organic matter content, %
Permeability (rate water passes thru soil), mm/hr

3. RESULTS

3.1 Crop Yield vs. Rainfall

There was general agreement in the pattern of high and low rainfall amounts from the multi-sensor based and quality-controlled cooperative gage based data, for this region, with a correlation coefficient r of 0.79 for July of 1997-1999, 2001-2002.

It is generally known that high corn yields are associated with low to moderate May rainfall and moderate to high July rainfalls (e.g. Thompson, 1986). Table 2 presents the relationship of normalized corn yield to May and July rainfall for various rainfall estimation methods.

Table 2. Normalized Corn Yield (bu/acre) 1997- 1999, 2001-2002 vs. Monthly County Rainfall Estimates for Multi-sensor Estimate (MSE), the QC_Coop Estimate (QC), the RT_Coop Estimate (RT) and the new 2002 Multi-sensor Estimate (newMPE). Multiple Regression Correlation Coefficient (R) and Coefficient of Determination (R^2), sample N, standardized regression coefficients, B, significant at < 0.05 level, for monthly rain and year.

Summer Estimate	**R**	**R^2**	**N**	**May Rain B**	**July Rain B**	**Year B**
1997-02						
MSE	0.49	0.24	3737	*-0.21*	*0.41*	*0.12*
QC	0.47	0.22	3471	*-0.20*	*0.39*	*0.12*
1997						
MSE	0.61	0.38	714	*-0.40*	*0.33*	-
QC	0.63	0.39	707	*-0.33*	*0.44*	-
1998						
MSE	0.11	0.01	772	*-0.09*	-0.05	-
QC	0.16	0.02	709	-0.06	*-0.14*	-
1999						
MSE	0.43	0.18	750	-0.04	*0.43*	-
QC	0.43	0.18	695	-0.02	*0.43*	-
2001						
MSE	0.55	0.30	750	0.02	*0.55*	-
QC	0.54	0.30	679	-0.06	*0.54*	-
RT	0.41	0.17	448	*-0.20*	*0.42*	-
2002						
MSE	0.62	0.38	751	*-0.24*	*0.52*	-
newMSE	0.64	0.41	751	*-0.23*	*0.55*	-
QC	0.60	0.36	681	*-0.22*	*0.50*	-
RT	0.55	0.30	446	*-0.27*	*0.49*	-

Regression analysis incorporating May and July rainfall resulted in a multiple regression correlation coefficient (R) of about 0.49 between these factors and normalized corn yield for both data sets. The similar magnitude between yield and rain suggest that the real-time multi-sensor data is of comparable quality to the gage data for purposes of predicting crop yields. The remaining analysis will employ the Multi-sensor estimates only.

For large yields, May rainfall is generally less than 125 mm and July rain is generally greater than 50 mm. For low yields, July rainfall is generally less than 100 mm (Fig. 2). There is an area of overlap in the distributions when either very good or very poor yields can be found, for example, when May rains are less than 200 mm and July rains between 50 and 100 mm are observed.

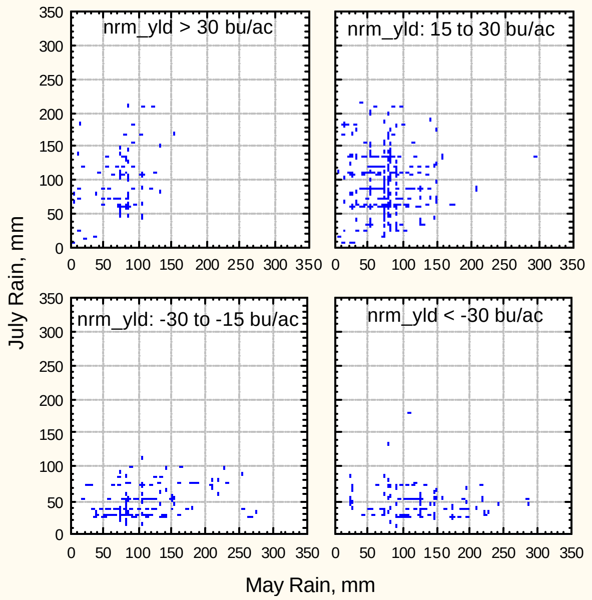

Figure 2. May and July Multi-sensor Rain estimates (mm) categorized by Normalized Corn Yield (bushels/acre).

3.2 Soil Parameters

To better predict good yields from bad, soil characteristics are incorporated into the analysis. Results of the multiple regressions analysis incorporating the soil parameter are presented in Table 3. Inclusion of drainage class, depth to the water table from the soil surface, and the percent of organic matter improve the relationship between normalized corn yield and rainfall.

The drainage class describes the occurrence of seasonal high water tables. Crops grown in well drained soils need more frequent rains than crops grown in poorly drained soils. Because poorly drained soils often result in extended periods of soil saturation, crops grown in these soils must be able to better tolerate wet root conditions. Most counties in MN, IA, IL, northern MO, OH, and IN are somewhat poorly drained. In WI, western MI, Southern MO, and KY, the soils are moderately well to well drained. In three of 5 years drainage class was positively correlated with normalized corn yield, the less well drained the better the yield.

The depth to the water table refers to the closest distance between the water table and the soil surface during the year. Throughout most of IL, IN, OH, MN, and northern MO, this depth is less than 100 cm. In the other regions the depths are generally between 100 and 150 cm. In 4 of 5 years, the depth to water table was positively correlated with normalized corn yield, the deeper the water table, the better the yield.

The third parameter, percent organic matter content, provides essential nutrients to crops and helps determine soil crop structure. Soils low in organic matter often display poor structure and are prone to severe compaction. In 3 of 5 years, organic matter was positively correlated with normalized corn yield, the higher the percent content, the greater the yield. In one year, 2001, organic matter was negatively correlated with normalized yield.

Examining all 5 years, the R value increased from 0.49 to 0.54. The improvement was greatest for the years 1998 and 1999. In 1998, the poorest normalized corn yields, particularly those with <-30 bushels / acre, were found in regions of low organic matter (<2 %), shallow water tables (<85 cm), a low silt content (35-45 %), and poorly drained soils (class 5 and 6). The poorest normalized corn yields were found in central Michigan where low May and July rains (<50 mm) occurred and in the southeastern tip of Missouri where very heavy rains (>150 mm) fell during both months. The highest normalized corn yields were found primarily in Minnesota, Wisconsin and Kentucky with May and July rains of 40-125 mm.

In 1999, the best yields tended to have higher sand content (15-50 %), and higher July rainfalls (60-180 mm) than the poor yields regions with sand contents of 5-20 % and July rains of 20-100 mm. Thus the poorest normalized corn yields, were found in regions with higher clay and silt content, and lower July rainfalls. The best yields were found in Wisconsin and central Michigan, and the poorest in Missouri, southern Ohio and Kentucky.

For any given year, the location of the rainfall with respect to given soil parameters is important. The response to heavy and light rains will vary depending on the location of the rain.

3.3 Location

Gradients in temperature and humidity that are found across the study area also result in differences in crop growth. Thus, latitude and longitude were added to the multiple regression analysis. Inclusion of longitude and in particular latitude, further improved the relationship between normalized corn yield and rainfall. The R values increased for every year and for all years. For 1997 to 2002 respectively the multiple correlation coefficients are: 0.75, 0.63, 0.67, 0.70, 0.76, and for all years: 0.56.

Table 3. Normalized Corn Yield (bu / acre), for five growing seasons. Multiple Regression Correlations, and standardized regression coefficients B (<0.05 significance level) with Weighted Average Soil Parameters and Monthly Rain.

Summer **Rain Estimate**	**1997** **MSE**	**1998** **MSE**	**1999** **MSE**	**2001** **MSE**	**2002** **MSE**	**ALL** **MSE**	**ALL** **QC**
R	0.70	0.55	0.58	0.67	0.72	0.54	0.52
R^2	0.48	0.30	0.34	0.45	0.52	0.28	0.27
Sample	712	766	747	748	749	3722	3455
Significant Parameters (0.05)							
Slope			-0.18	-0.11		-0.06	-0.06
Drainage Class	0.24		0.21		0.19	0.15	0.15
Depth to Water Table	0.27	0.20	0.25		0.20	0.17	0.17
Depth to Bed Rock	-0.10	-0.22	0.09				
Root Depth		-0.20	-0.20		0.19		-0.06
Sand Content, 0-50 cm	-0.32	-0.28	0.29		0.64		
Silt Content, 0-50 cm		-0.46		0.29	0.36		
Pot Plant Avail Water, 0-50 cm	-0.18	0.50					
Bulk Density , 0-50 cm	0.25		0.14				
Organic Matte, 0-50 cm	0.20	0.20		-0.10	0.16	0.11	0.13
Permeability, 0-50 cm							
Year	-	-	-	-	-	-0.12	-0.11
May Rain	-0.24	-0.20			-0.21	-0.18	-0.18
June Rain	-0.12			-0.21		0.04	0.05
July Rain	0.34		0.20	0.52	0.43	0.36	0.36

4. CONCLUSIONS

The relationship between normalized corn yields and monthly rainfall was examined at the county level. Monthly rainfall estimates were computed employing the gridded NCEP multi-sensor rainfall data, and the quality controlled and real-time NWS cooperative data. The similar magnitude in correlation between normalized corn yield and rain suggest that the real-time multi-sensor data is of comparable quality to the quality controlled gage data for purposes of predicting county crop yields.

For large yields, May rainfall is generally less than 125 mm and July rain is generally greater than 50 mm. For low yields, July rainfall is generally less than 100 mm. However, for moderate July rains, either positive or negative normalized yields could be found. Soil characteristics were incorporated into the analysis. For the two years, 1998 and 1999 when the correlation between yield and rain was least, the addition of soil information greatly improved the results. Different soil properties were important for these two years because of differences in the spatial distribution of rainfall with respect to soil features. Because of gradients in temperature and humidity across this large study region, latitude and longitude also were incorporated into the analysis. The addition, in particular of latitude further improved the relationship between crop yield and rainfall.

5. Acknowledgements

The authors wish to thank Jim Angel and Steve Hilberg for reviewing this manuscript. This work was performed under NOAA Cooperative Agreement, No. NA67RJ0146. The views expressed herein are those of the authors and do not necessarily reflect those of the funding agency or the Illinois State Water Survey.

6. REFERENCES

Fulton, R.A., J.P. Breidenback, D.-J. Seo, and D.A. Miller, 1998: The WSR-88D rainfall algorithm. *Wea. and Forecast.*, **13**, 377-395.

Hollinger, S.E., 1995: *Midwestern Climate Center Soils Atlas and Database.* Illinois State Water Survey Circular 179. 36p.

Seo, D.-J., 1998: Real-time estimation of rainfall fields using radar rainfall and rain gage data. *J. Hydrol.*, **208**, 25-36.

Seo, D.-J., and J. P. Breidenbach, 2002: Real-Time Correction of Spatially Nonuniform Bias in Radar Rainfall Data Using Rain Gauge Measurements. *J. Hydrometeor*:, **3**, 93–111.

Soil Conservation Service, 1993: State Soil Geographic Data Base (STATSGO) Data Users Guide. U.S. Department of Agriculture, Soil Conservation Service, Misc. Publ. 1492, p. 88.

Thompson, L.M. 1986: Climatic Change Weather Variability, and Corn Production. Agronomy J., **78**, 649-653.