Using the AIR Weather Index to Estimate the Contribution of Climate to Corn and Soybean Yields in the U.S.

Gerhard Zuba[1]
gzuba@air-worldwide.com

Oscar Vergara
overgara@air-worldwide.com

Tim Doggett
tdoggett@air-worldwide.com

Selected paper prepared for presentation at the
Southern Agricultural Economics Association Annual Meetings
Little Rock, Arkansas, February 5-9, 2005

Keywords: corn yield, soybean yield, technology trend, climate variability

Research in progress. Do not quote without authors' permission.

[1] Corresponding Author.
AIR Worldwide Corporation. 131 Dartmouth Street. Boston, MA 02116

Using the AIR Weather Index to Estimate the Contribution of Climate to Corn and Soybean Yields in the U.S.

Abstract

Using historical production data at the county level and statistical analysis, we investigate climate contributions to corn and soybean yields between 1974-2003. Crop yield trends are decomposed into two components: the technology-derived trend and the trend resulting from climate variability. Implications for agricultural risk management and farm policy are discussed.

1. Introduction

Climate is by far the most important source of uncertainty in the outcome of agricultural production, and weather-related perils remain the most important triggers of nationwide crop losses. Therefore, to perform any meaningful agricultural risk assessment, risk managers must be able to quantify the effect of climate on crop yields. This is often done by examining historical crop yield information.

However, there are several problems with using historical crop yield information, as follows:

- Technological advances produce a trend in crop yield histories that needs to be removed in order to develop appropriate yield distributions
- Variability in weather (especially the occurrence of extreme weather events) can produce significant crop yield variability that masks the technology trend

- Differences in farming practice, as well as geographical differences in soil, impact how much weather influences overall crop yield
- Time series of weather data and crop yield information can be incomplete or of inadequate duration.

Thus, de-trending historical crop yield information is very challenging.

This paper presents the AIR Weather Index (AWI) yield model, which offers a methodology for isolating climate contribution to yields from other factors. Specifically, we use historical production statistics at the county level and statistical analysis to investigate climate contributions to corn and soybean yields between 1974-2003. We decompose the yield trend into two components: the technology-derived trend and the trend resulting from climate variability.

Preliminary analysis indicates that yield distributions obtained from this method are more realistic since the true yield risk due to adverse weather effects is being objectively separated from the technology trend. These yield distributions provide risk assessments that better reflect the true changes in technology that have affected the most important crop producing regions of the United States during the last 30 years.

2. Background

As farmers have known through the ages, climate variability and corresponding favorable or adverse weather patterns affect the growing conditions of crops. Farming technology and management try to mitigate the negative impacts of adverse climate conditions, such as drought or excessive rainfall, on crop yields. Also, various researchers have attempted to develop risk assessment tools (e.g., Wu *et al*., 2004) that quantify the weather risk associated with growing a specific crop in a specific region. It is not only the extreme weather events such as droughts or floods that are of interest to scientists. Recent studies by Lobell and Asner (2003), Hu and Buyanovsky (2003) and Eitzinger *et al*. (2002) discuss gradual changes in temperature and amount of rainfall, as well as the gradual change in the frequency of severe events due to global climate changes. Gradual changes in climate may have contributed in a positive or negative way to corn and soybean yields observed in different regions of the United States during the last decades.

Improved farming practices and advances in technology are often credited for steady increases in crop productivity over time. On the other hand, the impacts of adverse climate conditions are reflected in yield histories. For example, Figure 1 shows the corn yield history of Nemaha County, Nebraska between 1974 and 2003. The negative impacts to corn yields due to droughts in the 1970s, 1980s, and 2002, as well as the flood in 1993 are evident. Nevertheless, an overall steady increase in corn yields is also observed.

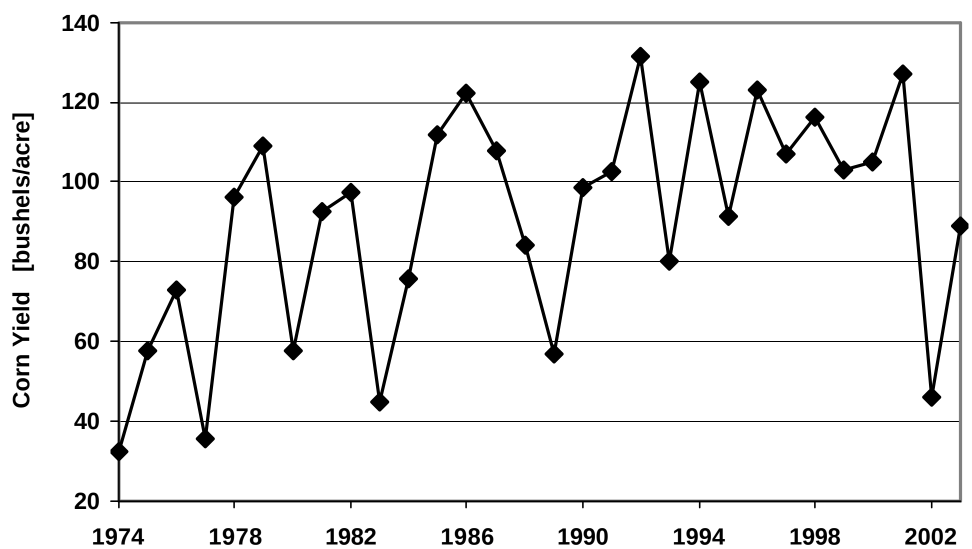

Figure 1: Observed corn yield in Nemaha County, Nebraska (Source: NASS)

A closer look at this increase reveals the following. A linear trend analysis for Nemaha County, Nebraska, results in a corn yield trend of 2.19 bushels/acre/year for the period 1974-2001. Yet a similar trend analysis results in a corn yield trend of 0.63 bushels/acre/year for the period 1982-2003. The difference in the two results is due to the fact that the occurrences of low yields are not equally distributed within the time series. In fact, if not for the very low yields in 2002 and 2003, the difference would not be as significant. From the yield history it can be seen that except for these two years the frequency of very low yields has declined starting in the early 1990s. Therefore we conclude that neither of these two values accurately reflects the true corn yield trend due to technology improvements for Nemaha County. Most importantly, this example demonstrates a general problem associated with any crop yield time series. Namely, there are two yield trends that interact simultaneously: the yield trend caused by technological improvements over time, and the yield trend that is due to changes in climate variability over time.

Agricultural economists have long recognized the problems posed by the effect of trends on yield time series, thus limiting their usefulness for modeling, simulation, or agricultural risk assessment. Because direct use of observed yields is inappropriate (Goodwin and Ker, 1998), a previous de-trending step is always involved when utilizing crop yield time series information (Challinor *et al*., 2004, Hao *et al*., 2004, Chen and Miranda, 2004, Clark *et al*., 2003 and Norwood *et al*. 2004).

The goal of the de-trending process is to eliminate the trend in yields due to technological improvements while preserving information about risk due to adverse climate conditions.

De-trending is a two-step process. First the technology trend is estimated and then this trend is removed from the crop yield time series so they can be used for risk analysis. The brief exercise above utilizing the yields in Nemaha County showed that estimating the technology trend cannot be accomplished by a simple linear trend analysis. This method is too dependent on the time period used. That means the pattern in time of climate variability has an influence on the trend estimation. Also, more complex methods such as log-linear de-trending or the LOESS procedure (Cleveland *et al*. 1988) are likely to remove both technology and trends in climate variability at the same time. The issue of correctly de-trending a yield time series when climate effects are intertwined with technology effects is important in order to accurately estimate yield anomalies or to construct a realistic crop yield distribution that can be used for risk analysis.

The example of yield histories in Figure 1 is only one of numerous examples where it is evident that changes in climate variability over time has an influence on yield trends and therefore has to be taken into account when de-trending yields. The objective of this paper is to describe an innovative method that is suitable to separate the climate contribution to yields from the technology contribution by using a novel de-trending procedure.

3. Data used

In order to model the climate contribution to corn and soybean yields, daily precipitation and temperature data between 1974-2003 were obtained from the National Center for Environmental Prediction (NCEP). The data are derived from daily observations of approximately 5,000 reporting weather stations within the U.S. and interpolated to a data grid. The grid has a spatial resolution of approximately 25 km for precipitation and 50 km for maximum and minimum temperature. These raw data were re-processed to form a county-level climate time series of daily maximum and minimum temperature and daily precipitation.

The base weather information was then used to compute growing degree-days and evapotranspiration datasets. Soil-related parameters (e.g., plant available water capacity, surface moisture, sub-surface moisture and runoff) were computed by integrating weather related data with the high-resolution USDA State Soil Geographic Database, administered by Penn State University.

County-level yield time series for corn and soybeans were obtained from the National Agricultural Statistics Service (NASS) database. A database of crop specific parameters (e.g., water requirements, crop phenological stages, planting date) was created using information from various academic sources, including Iowa State University, Purdue University, University of Minnesota, and Ohio State University.

For the following analyses, it is critical that the time series of weather data be homogeneous, i.e., there are no gaps in the data set. For example, missing a short period of rainfall data within a longer time span of dry conditions can have a significant impact on the resulting yield estimates. Furthermore, the data sets have to be processed so that the spatial coverage and resolution of all the data sets match each other. It is obvious, and our research confirmed, that the weather time series and the corresponding yield data have to be of comparable resolution in order to produce accurate results. To accomplish this, spatial interpolation of the time series of weather data has been performed utilizing standard meteorological interpolation techniques.

4. The AWI Yield Model

The proposed de-trending methodology separates the technology trend from the climate effects that are intertwined in a crop yield time series. This is done by regressing yields against a linear technology trend and a weather indicator that explains yield deviations from this trend, as follows:

$$\text{Yield}(t) = c_0 + m*t + c_1*\text{AWI}(t) + \varepsilon \qquad (1)$$

where c_0, m and c_1 are regression coefficients, t is time (year), m measures the technology trend, AWI (the AIR Weather Index) is the weather indicator and measures the weather effect on yields, and ε is the residual error. By including a weather component in the de-trending process, the combined effects of technology and weather on yields is accounted for under the assumption that the AWI is capable of modeling the deviations of yields from the mean technology trend.

The AWI model is constructed using time series of weather variables (e.g., temperature, precipitation), weather derived parameters (e.g., growing degree days, evapotranspiration), soil-related parameters (e.g., plant available water capacity, surface moisture, sub-surface moisture, runoff) and crop-specific parameters (e.g., water requirements, crop phenological stages, planting date). Compared to other crop growth models, the underlying methodology for the AWI favors simpler parameterization of yield-related crop growth and crop damage. Calibration of the model is done by adjusting a small number (3 to 4) coefficients used to optimally scale the effect of different weather perils on crop stand.

5. De-Trending Crop Yields Using the AWI Yield Model

The AWI is calculated from daily climate data by integrating observed crop conditions between planting and harvest and provides an estimate of potential yield at harvest.

For demonstration purposes, the yield history of corn yields in Nemaha County, Nebraska (Figure 1) is reanalyzed. Table 1 shows the results from de-trending for different time windows using the AWI yield model as compared to a simple linear de-trending method such as the one described by:

$$\text{Yield}(t) = c_0 + m{*}t + \varepsilon \qquad (2)$$

Large differences are expected between these two different de-trending methods because the simple linear de-trending is heavily affected by the low yield of 2002 due to drought, thus significantly decreasing the slope of the technology trend. On the other hand, the proposed method using Equation 1 (the AWI Yield Model equation) should be able to capture the low yield of 2002 in the AWI term of the equation. By modeling the deviation of observed yield from the mean technology trend for this year using climate variables, the effect on the technology trend (m) should be minimized. The results for the analysis are summarized in Table 1 below.

Time Window	m linear de-trending	r^2 linear de-trending	m AWI de-trending	r^2 AWI De-trending
1974-1990	2.59	0.22	1.87	0.79
1974-2001	2.19	0.41	1.77	0.80
1974-2003	1.49	0.21	1.38	0.78
1982-2003	0.63	0.03	1.30	0.70

Table 1: Estimation of the technology trend (m) in bushels/acre/year for different time windows using a simple linear de-trending (Equation 2) versus AWI de-trending (Equation1) on corn yield time series for Nemaha County, Nebraska

According to Table 1, the technology trend is clearly dependent on the time window used for the analysis. For the linear de-trending method, the technology trend variability exceeds 50% of its mean value (1.75 bushels/acre/year) between 1974-2003. On the other hand, the AWI de-trending method captures most of the yield variation due to weather effects for different time periods, as reflected by correlation coefficients in between 0.70 to 0.80. Also, the difference in the estimated technology trend for different time periods is smaller than 18% of the mean value (1.58 bushels/acre/year) between 1974-2003.

One could expect the same m value for all different time windows using the AWI de-trending methodology. The AWI is constructed to capture the deviations of yields from a linear trend caused by weather effects and therefore no dependency on the time window is expected. The different values can be explained as follows: Although the correlation coefficient is high, most but not all of the deviations from the assumed linear technology trend could be explained by the AWI method. Furthermore it seems that the decrease in technology trend for more recent time periods, represented by lower values for m, is due to the "stagnation" of corn yields in Nemaha County since the early nineties (see Figure 1). Therefore the assumption of a constant linear trend for the entire time period is only partly valid. In fact, this is one of the reasons why de-trending is not done on much longer time periods. Apart from the availability of long and accurate observed yield time series, the increase in yields due to technology is in general not linear over very long time periods.

A major advantage of using the AWI de-trending method is that it is less susceptible to variation of the time period chosen for the analysis. This is due to the fact that both the technology trend and the trend of climate variability are taken into consideration at the same time when de-trending the raw yield observations. Therefore, large deviations from the expected yield caused by adverse weather are captured by the AWI term in Equation 1, rather than affect the technology trend m.

Figure 2 and Figure 3 show additional examples of modeling the weather contribution for two different crop and county yield time series, as described by the AWI Yield Model. The high correlation coefficients between observed and modeled yields imply that the AWI model explains a significant amount of the weather effects on yields. More results on the skill measured by the correlation coefficient r^2 are provided in the next section.

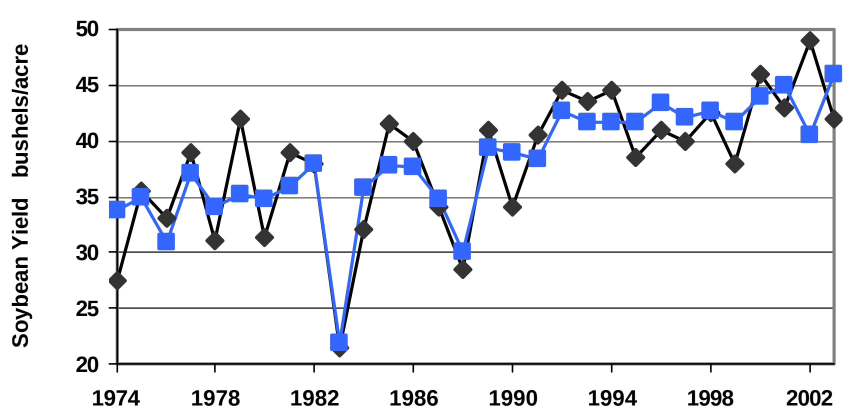

Figure 2: Observed (black diamonds) and modeled (blue squares) soybean yield for Macoupin County, Illinois. The correlation coefficient between observed and modeled yield is 0.70

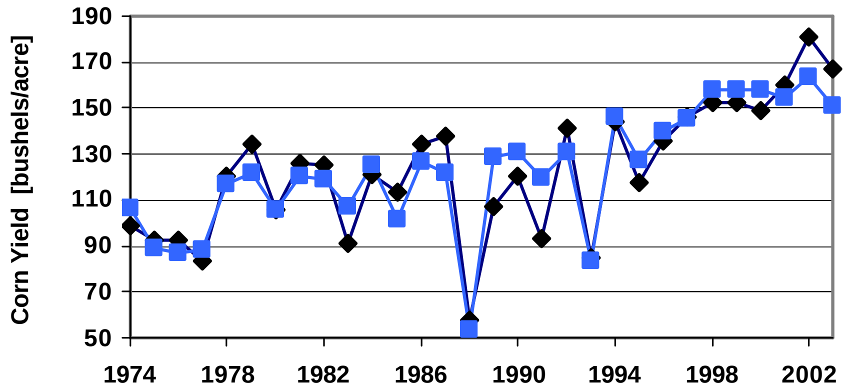

Figure 3: Observed (black diamonds) and modeled (blue squares) corn yield for Benton County, Iowa. The correlation coefficient between observed and modeled yield is 0.86

6. Application of AWI de-trending

Similar to the proposed AWI de-trending methodology, Lobell and Asner (2003) studied the technology component and the climate contribution to yield trends. Their starting point is the quality of the correlation between yield anomalies and the June to August average temperature anomalies on a county level for the specific time period 1982-1989. Only counties that have a significant negative correlation ($P < 0.01$, see Lobell and Asner, 2003) are used for the actual trend analysis that splits the total trend into separate technology and climate components.

In Section 5 we have shown that results from a linear trend analysis using Equation 2 and therefore yield anomalies calculated via this procedure depend on the chosen time period. We assume that this will affect the study of Lobell and Asner (2003) as well. To quantify this effect, we reproduced their correlation of June-August average temperature anomalies and corn yield anomalies for the period 1982-1998, and compared it to similar conditions for the period 1987-2003. As can be seen in Figure 4, the temperature-yield correlation differs when the time period is modified. Even though the overall pattern of the correlation coefficients is similar, there are significant localized differences (e.g., Iowa and Illinois). These differences can best be seen in Figure 5, which isolates counties with a temperature-yield correlation of P<0.01.

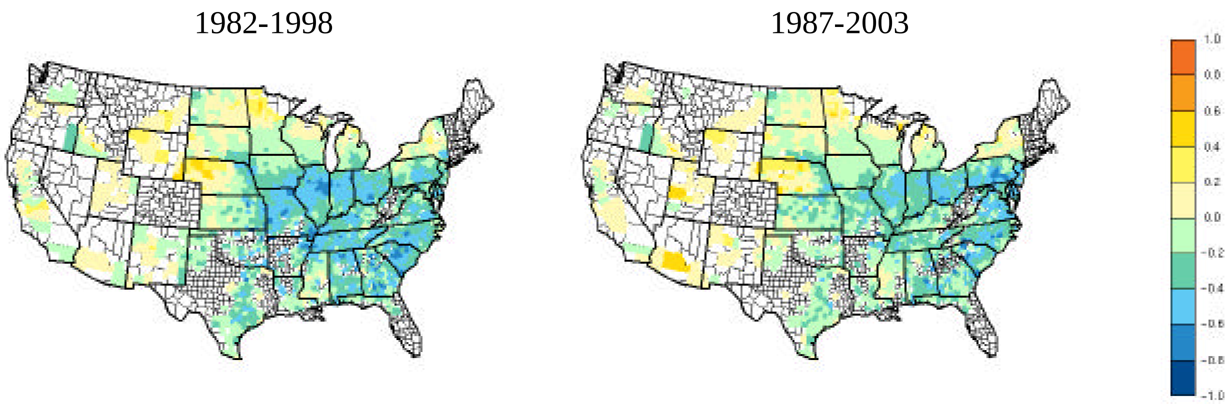

Figure 4: County-level comparison of correlation coefficients between June-August average temperature anomalies and corn yield anomalies for two different time periods.

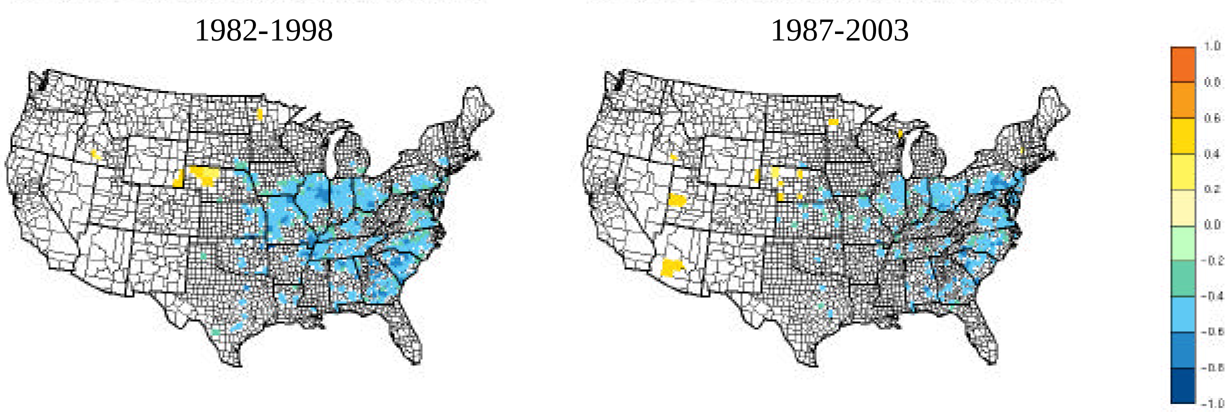

Figure 5: County-level comparison of correlation coefficients (P<0.01) between June-August average temperature anomalies and corn yield anomalies for two different time periods.

As Figure 5 clearly shows, by applying the P<0.01 criteria on the yield-temperature correlation different counties are selected for different time periods. This will have a clear affect on the result of any further analysis, especially given that almost all of Iowa—an important corn and soybean producing state—is spared from the analysis when using the period 1987-2003. We followed Lobell and Asner (2003) and calculated the non-climatic soybean trend for these two periods to 0.34 and 0.17 bushels/acre/year, respectively. Once again, these values are based on using Equation 2 for de-trending to calculate crop yield anomalies and June to August average temperature as a climate indicator.

Figure 5 and the just mentioned soybean example demonstrate once more the importance of separating yield trends due to technology improvements from yield time series in an objective and accurate way. A chosen time period will affect any de-trending methodology that does not take changes in frequency of adverse climate conditions into account.

The AWI de-trending method for estimating the technology-derived yield trend and the climate-derived yield trend eliminates the dependence of the results on the chosen time period of analysis. This can be accomplished because the AWI captures the yield variability due to adverse/favorable climate conditions and because the de-trending methodology according to AWI Yield Model (Equation 1) accounts for a technology trend and a contribution from climate variability at the same time. As shown in Table 1, this results in a more stable approximation of the value of the technology trend. The AWI de-trending method does this regardless of the time period used for analysis.

Application of the AWI de-trending method is shown in Figure 6. This figure visualizes the correlation coefficient of the regression when calculating the coefficient c_o, m and c_1 according to Equation 1 utilizing observed corn and soybean yields of counties in Illinois, Indiana, Iowa, Kansas, Minnesota, Missouri, Nebraska, Ohio, South Dakota and Wisconsin between 1974-2003 and corresponding values of the AWI.

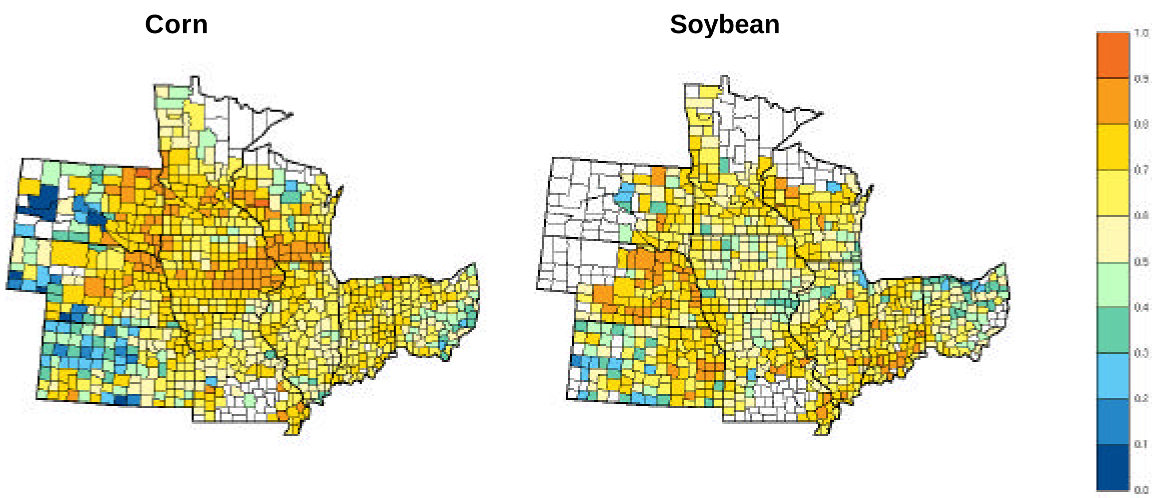

Figure 6: Correlation coefficient (observed versus modeled yield for corn and soybean yields using the AWI de-trending method (Equation 1) for a time window between 1974-2003.

The AWI explains approximately 70% of the yield variability due to climate variability in the Midwest within the last 30 years. The high r^2 values of correlations between modeled and observed yields shows the explanatory value of the AWI de-trending method and leads to excellent out-of-sample tests. This is due to the fact that the method only has 3 to 4 adjustable[2] parameters that are used to calibrate the AWI itself on a county-by-county basis. Furthermore this fulfills the requirement of predictability (Sinclair and Seligman, 2000) necessary to validate crop yield models.

[2] During the calibration process parameters used for constructing the AWI are limited to their natural physical constraints.

For example, the AWI de-trending method calibrated to the data of Figure 1 but only using the period from 1974-2001 and then using the calibrated model and the weather input for the years 2002 produces a county yield of 51 bushels/acre, which is apparently a very low yield. In fact, due to the severe drought that affected the county, the observed corn yield in 2002 was 46 bushels/acre. Extensive out-of-sample tests performed on the AWI de-trending methodology have proven its validity as an innovative climate and technology de-trending tool, as well as its skill in capturing the effect that climate variability has on crop yields.

The AWI de-trending procedure superimposes the effect of weather (AWI term in Equation1) on a linear technology-derived trend (term m*t in Equation 1). Therefore, the resulting AWI term of the equation can be used for analysis whether or not there is a trend in climate variability. The calibrated AWI's on a county-by-county basis were subjected to a linear fit over time according to Equation 3 to estimate a change in climate variability.

$$AWI(t)^{fit} = a + b*t + \varepsilon \qquad (3)$$

Figure 7 shows the results obtained from applying the AWI de-trending method to estimate the technological-derived trend (m) and applying Equation 3 on the resulting AWI's for estimating a trend in climate variability (b). It has to be mentioned that this trend can change from year to year depending on whether the added data corresponds to very good yields corresponding to favorable weather conditions or very low yields corresponding to the impacts of adverse weather. It is the technology-derived trend that stays more or less constant. For Figure 7, corn producing counties in Illinois, Indiana, Iowa, Kansas, Minnesota, Missouri, Nebraska, Ohio, South Dakota and Wisconsin between 1974-2003 were selected in the sample with a climate-yield correlation coefficient > 0.55 (see Figure 6).

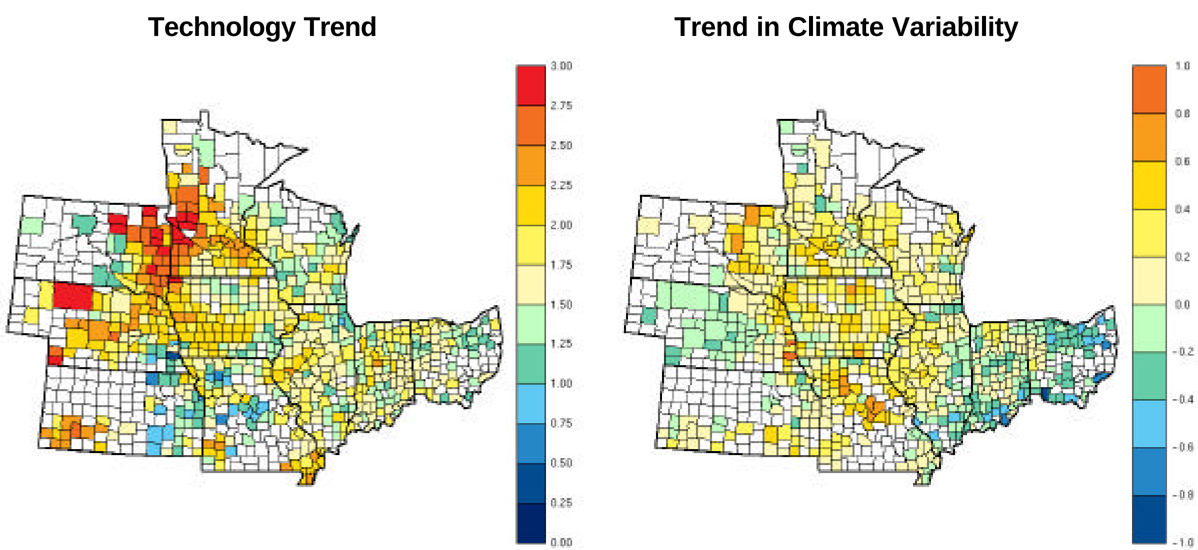

Figure 7: Technology and climate contributions (in bushels/acre/year) to corn yields derived from the AWI model for a time window between 1974-2003.

By isolating the effect of climate variations on yields, the AWI de-trending method allows for the quantification of the pure technology contribution to yields. As shown in Figure 7, in some areas of the Midwest, technology has increased yields by up to 3 bushels/acre/year. On the other hand, other areas seem to have reached a plateau with respect to marginal increases to yield potential derived from technology.

Figure 7 also shows the contribution to yields derived from climate effects observed between 1974-2003. Historically important corn producing regions have benefited from a production boost of approximately 1 bushel/acre/year due to favorable changes of climate variability. On the other hand, those same changes have been detrimental to corn productivity for other areas of the Corn Belt.

Besides de-trending yield time series the AWI Yield Model has further potential applications:

- By integrating crop conditions between planting and any date within a growing season, the AWI can be used as a real time monitoring tool to assess current crop conditions.
- For the current season the AWI can be used as an estimate of potential yield at harvest, which is available long before official NASS county yields are published.
- Because the AWI separates technology-derived from climate-derived effects on the yield time series it can be utilized to objectively determine APH yields for individual farms and therefore can be included in a procedure to mitigate declining yields due to successive low yields.

7. Conclusions

In order to assess crop risk, agricultural risk managers require an estimation of the contributions made by climate variability and weather-related perils to crop yields. However, developing an appropriate yield distribution for analyzing agricultural portfolio risk is a challenging task. Historical yield information needs to be de-trended because changes in technology have affected the most important crop producing regions of the United States during the last 30 years. The challenge lays in the fact that intra-year yield variability due to weather masks the contributions of technology improvements and farm management to the yield trend. Therefore, to de-trend the yield data correctly, the climate contribution to yield must be quantified.

The AWI de-trending method allows for the quantification of the climate contribution to yields, therefore allowing the true technology trend to be calculated and the risk due to climate variability to be isolated. Preliminary analysis indicates that yield distributions obtained from this method are more realistic since the true yield risk due to weather is objectively separated from the technology trend. The resulting distributions provide risk assessments that better reflect the true weather risk.

Implications of this method are many: By quantifying the effects that climate variability has on crop productivity, policymakers can develop the safety nets necessary to allow producers to remain farming when struck by disasters. Additionally, extension educators may find it useful to identify regions in which crop productivity still lags behind—despite favorable weather conditions—and adjust training and technology transfers accordingly. Other uses include the ability to monitor current crop conditions and estimate potential yields at harvest.

8. References

Challinor A.J., Slingo J.M., Wheeler T.R., Craufurd P.Q., Slingo J.M., Grimes D.I.F., 2004. "Design and optimization of a large-area process-based model for annual crops", *Agricultural and Forest Meteorology* 124, 99-120

Chen Shu-Ling, Miranda M.J., 2004. "Modeling Multivariate Crop Yield Densities with Frequent Extreme Events", *AEAA Annual Meeting*, Denver, Colorado, August 1-4, 2004

Clark N., Fleming R.A., Skees J.R., 2003. "Investigating the Relationship between Yield Risk and Agri-Environmental Indicators", *AAEA Annual Meeting*, Montreal, Canada, July 27-30, 2003

Cleveland W.S., Devlin S.J., Grosse E., 1988. "Regression By Local Fitting", *Journal of Econometrics* 37, 87-114

Eitzinger J., Štastná M., Žalud Z., Dubrovský M. 2002. “A simulation study of the effect of soil water balance on winter wheat production under different climate change scenarios”, *Agricultural Water Management* 61, 195-217

Goodwin, B., and A. P. Ker, 1998. “Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk Crop Insurance Contracts” *American Journal of Agricultural Economics* 80, 139-153

Hao J., Hartell J., Skees J., 2004. “Optimal Inter-Period Weighting of Cumulative Indices for Weather-Based Contingent Claims”, *AEAA Annual Meeting*, Denver, Colorado, August 1-4, 2004

Hu Q. and Buyanovsky G., 2003. “ClimateEffects on Corn Yield in Missouri”, *Journal of Applied Meteorology*, Vol. 42, pp 1626

Iowa State University: http://maize.agron.iastate.edu/corngrows.html

Lobell and Asner, 2003. “Climate and Management Contributions to Recent Trends in U.S. Agricultural Yields”, *Science*: 299 (2003), pp 1032

NASS: http://www.nass.usda.gov/QuickStats/

NCEP: http://www.cpc.ncep.noaa.gov/products/precip/realtime/us_precip.html, http://www.cpc.ncep.noaa.gov/products/tanal/accesspage.html

Norwood B., Roberts M.C., Lusk J.L., 2004. “Ranking Crop Yield Models Using Out-of-Sample Likelihood Functions”, *American Journal of Agricultural Economics* 86, 1023-1043

Ohio State University: http://www.ag.ohio-state.edu/~perf/sb2000/

Penn State University: http://www.essc.psu.edu/soil_info/index.cgi?soil_data&conus&data_cov

Purdue University: http://www.ces.purdue.edu/extmedia/NCH/NCH-18.html

Sinclair T.R. and Seligman No’am, 2000. Criteria for publishing papers on crop modeling, *Field Crops Research* 68, 165-172

University of Minnesota: http://www.soybeans.umn.edu/crop/growth/index.htm

Wu H., Hubbard K.J., Wilhite D.A., 2004. An Agricultural Drought Risk-Assessment Model for Corn and Soybeans, *International Journal of Climatology*, Vol 24, 723-741