# TREND YIELD ANALYSIS AND YIELD GROWTH ASSUMPTIONS

BY:

JOHN R. KRUSE

JUNE, 1999

FAPRI-UMC Report #06-99

FOOD AND AGRICULTURAL POLICY
RESEARCH INSTITUTE

UNIVERSITY OF MISSOURI
101 SOUTH FIFTH STREET
COLUMBIA, MO 65201

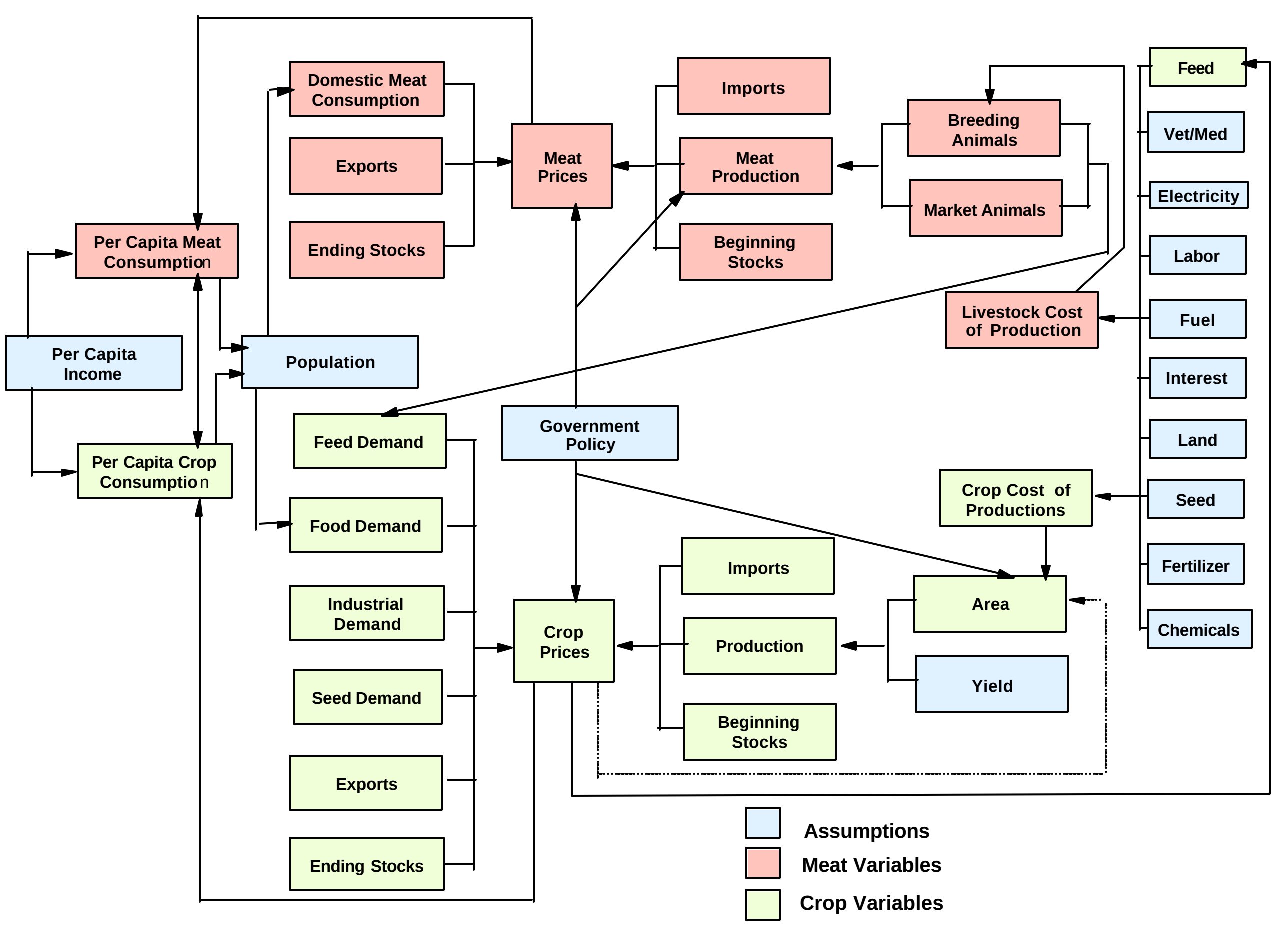
TEL: (573)-882-3576
FAX: [573]884-4688

WWW.FAPRI.MISSOURI.EDU

# TREND YIELD ANALYSIS
# AND YIELD GROWTH ASSUMPTIONS

In the process of building any baseline, forecast, or scenario one of the first steps is to decide on a core set of assumptions around those variables not specifically solved for in the econometric model. Consider the simplified representation of a country model below. As the flow diagram illustrates, the assumptions (blue boxes) pertain to income, population, government policy, input costs, and yields. Each of these assumptions has a critical role in determining the results of the analysis. Government policy regulates supply and demand through domestic supports and trade barriers. Population and income drive demand while cost of production and yields drive part of the supply sector. This paper will address the formulation of assumptions pertaining to yield growth in the crops sector.

## Country Model Flow Diagram



## Trend Yields

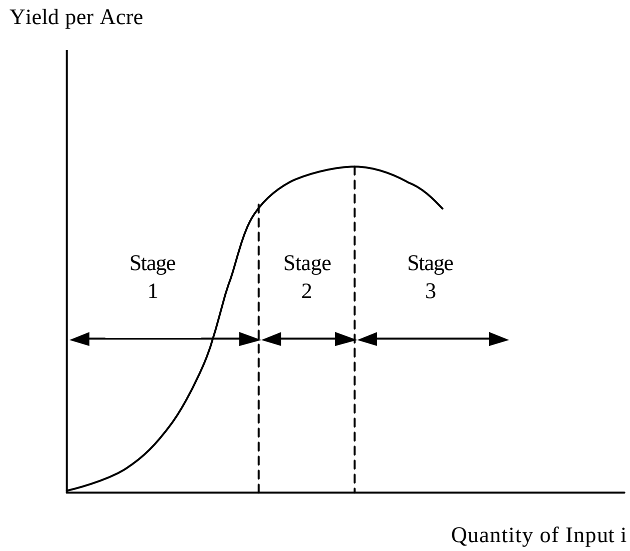
The yield assumptions used in any agricultural forecast are especially critical because U.S. agriculture is plagued with an over-capacity problem. While demand for agricultural commodities has been growing, yield growth in the U.S. and abroad has offset much of this new demand and has outpaced it in some cases. Rapid yield growth in other countries has created new competitors with the U.S. in export markets and caused some countries to quit importing.

Crop yields in any given year are determined by at least three major types of factors: weather, "autonomous" technology improvements, and economic factors. Unfortunately, only a small portion of these factors is known with certainty. Weather is a very large determinant of yields but remains very unpredictable. Technology improvements are sporadic and work differently in across regions creating some uncertainty in their impacts on a specific farm. Even economic factors are not known with complete certainty. Weather and changes in global supply and demand affect output prices and create a great deal of uncertainty around final prices. However, by assuming normal weather, average technology growth, and formulating a price expectation, farmer's can derive an optimal level of input application to maximize profits. This approach will be referred to as the traditional approach.

### *The Traditional Approach*

There has been a great deal of previous work focused on estimating crop yields as a function of the critical inputs such as fertilizer, chemicals, seeding rate, and technology, particularly in the U.S. where the data is more readily available. Traditionally agronomists have estimated production functions that resemble Figure 1 on the following page. While the graph doesn't illustrate the interaction among inputs, it is representative of the production function for a given input holding all other inputs fixed. In stage 1, the factor elasticity (the percentage change in output over the percentage change in inputs) is greater than or equal to one. This means the percentage growth in output increases faster than the percentage growth in inputs used. For producers in this stage, the return on an

Figure 1. Yield and Input Applied




additional input is obvious. As long as input prices do not get high relative to output prices, producer will not normally operate in Stage 1. However, returns to input use in the next two stages begin to slow down and eventually become negative in Stage 3. Clearly producers will not intentionally operate very far into stage 3 before they realize they are hurting their production level. However, there are other required inputs that are outside their control (such as weather) that may cause them to operate in stage 3.

Translating the production function into a farmer's bottom line is accomplished through the profit function. While the farmer may have a variety of goals other than profits, for simplification lets assume they want to maximize profits. The profit function (under pure competition) is given by:

$$p = P_o Y_o - \sum_{i=1}^{k} P_i Q_i$$

The point where profits are maximized is heavily dependent on the production function. There are many different functional forms of production functions that are documented throughout the literature. Among the simplest yet very realistic functional forms is the quadratic production function. Generally the farmer will operate in the latter part of stage 1 or early part of stage 2 of the production function. The quadratic production function is representative of stage 1 above the inflection point, stage 2 and stage 3. The quadratic production function is given by the following:

$$Yo = a + b_{11}Q_1 + b_{12}Q_1^2 + b_{21}Q_2 + b_{22}Q_2^2 + \cdots + b_{K1}Q_K + B_{K2}Q_K^2$$
$$+ d_{11}Q_1Q_2 + d_{12}Q_1Q_3 + \cdots + d_{k1}Q_1Q_k + d_{21}Q_2Q_3 + \cdots + d_{2K}Q_2Q_K + \cdots + d_{(K-1)K}Q_{(K-1)}Q_K$$
$$+ g_{11}Q_1Q_2Q_3 + \cdots + l_{11}Q_1Q_2Q_3 \cdots Q_K$$

where the variables are defined as:

$Y_o = Yield$

$Q_i =$ Quantity of Input i, for i = 1 to k

Of course the interaction terms may be of more or less importance depending on the inputs. However, agronomists often point out that much of the advances in hybridization the past 60 years have come through the interaction of increased chemical and fertilizer use and genetic enhancement.

Substituting any production function into the profit function and taking the first derivatives with respect to input i to obtain the first order conditions yields the following:

$$\mathrm{P_o} \frac{\partial \mathrm{Y_o}}{\partial \mathrm{Q_i}} = \mathrm{P_i}$$

The first order conditions reiterate the well know aspect of profit maximization that profits are maximized when marginal revenue equals marginal cost. Note that unless inputs are free, profits are not maximized when yield is maximized.

The difficulty in extending this methodology to predict yields stems from two sources. The first source is data availability. In many countries, data on the quantities of the critical inputs used per acre by crop is often not available. Even the basic data on fertilizer, seeding rate, and chemical use for herbicides and pesticides by crop is not available. In addition, inputs such as technology through genetics are not directly observable but interact strongly with other inputs to produce yield growth. Genetic improvements are sporadic, although to a certain extent known at least by the seed companies for the next few years. However, seed companies tend to consider their anticipated release of new genetics as proprietary information to protect their research investment. Even if one knew the precise release of new technologies, the rate of adoption by farmers is another difficult variable to predict.

The second source of problems stems from the fact that technology is often changing the shape of the production function, shifting it up and closer to the origin. For example, improvements in genetics could cause fertilizer to be more efficiently used by the plant such that less fertilizer is required to produce the same yield. At the extreme, the plant could become so efficient that the optimal level of fertilizer today could actually hurt yields 10 years from now when the technology is fully realized. This is not unrelated to situation today where many of the inputs, which were so significant in explaining yield growth in the past, have reached a plateau in use yet yields continue to grow. Trying to estimate a changing production function through time could lead to weak parameter estimates with ambiguous signs.

*An Alternative Approach*

Another possibility to estimate yields would be to consider historical yield growth and use it as a proxy for setting future trends. However, there are some limitations to remember when this is done. Among those is to first acknowledge the sporadic nature of yield growth due to improvement in genetics and management. Yields tend to make jumps over time rather than follow a smooth increase. Since future steps are very difficult to predict, the average historical trend yield growth will be used to project the future. This means that future yields will be overestimated in some years and underestimated in others.

In some cases historical trend yield growth may not be a good indicator particularly if there has been a large economic shock that has impacted input use. This is particularly true in the Former Soviet Union countries and Eastern Europe. The transition to a capitalistic economic structure has been slow and painful for these countries causing disruptions in input supplies and capital availability. After significant declines in yield since the reforms began in 1992, some countries are beginning to have some small recovery in yields. Projecting yield growth in these areas boils down to a guess that gets refined as more information becomes available.

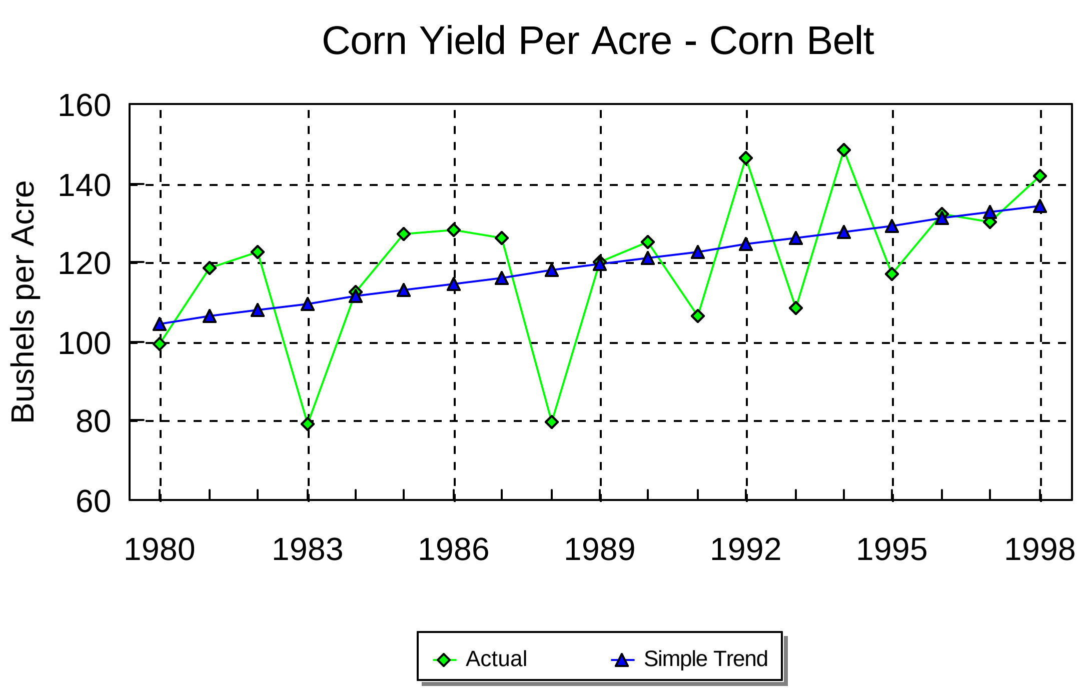
## Estimating Trend Yields

The functional form used to establish trend yield growth is also important. Research groups such as IFPRI, the World Watch Institute, and USDA have suggested that yield growth has begun to slow and subsequently many new model specifications use the log of a linear trend to project future yield growth. However, there seems to be little evidence of yield growth slowing around the world with the exception of the Former Soviet Union countries and Eastern Europe. But yields in these countries have declined from lack of inputs rather than reaching a yield plateau. An aggregation of these countries into a world yield average could lead to the incorrect conclusion that world yield growth is slowing.

Another way to look a yield growth is to consider the technical potential of the plant. Researchers at the University of California-Berkley have calculated the maximum possible vegetation that can be generated if a plant was given all possible inputs it needs. For corn and soybeans this translates into potential yields of 450 bushels per acre and 250 bushels per acre, respectively. Clearly current yields of 128 bushels per acre for corn and 40 bushels per acre for soybeans are no where close to the technical potential. Therefore, there appears to be no reason to suggest that yield growth is slowing.

Since weather forecasting remains inaccurate, it will be critical to evaluate those historical trend yields controlling for the effects of weather. A common observation about the effects of weather on yields is that there seems to be two categories: moderate and severe. For further illustration, consider the graph of corn yields in the U.S. Corn Belt below. The years of 1983 and 1988 were severe droughts and deviate strongly from the simple trend. However, the deviations from trend in years of 1991 and 1993 are more moderate, reflecting a moderate drought in 1991 and severely wet conditions in 1993. On the side of exceptional yields, the years of 1992 and 1994 stand out. Moderately good years of 1981-82 and 1985-86 also stand out. The graph also illustrates that the effects of weather are not symmetric for corn. Clearly the exceptionally bad years appear to deviate more from trend than the exceptionally good years. As one looks across crops, the lack of symmetry is most apparent in hybrid crops and less apparent in non-hybrid crops such as soybeans. The obvious effect of non-symmetric deviation in this case is to shift the simple trend line downward. To observe this, consider that 9 data points lie above the simple trend line while only 5 data points lie below it. Another more difficult effect to determine from the graph is any bias in the slope of the trend line.

Estimating trend yields correcting for weather effects requires a two-stage process. In the first stage, the yield for crop x in country j is estimated as a function of a linear trend using linear regression. The studentized residual errors resulting from that regression are then reviewed relative to their size. As a result of the observed moderate and severe weather effects and the lack of symmetry, four indicator (dummy) variables are created based on the value of the studentized residual errors. The estimation of the trend yield is



sensitive to the choice of the critical values used to create these indicator variables. The choice of these critical values should reflect the objective of removing moderate and severe weather effects. (An example of one possible choice is given below.) In case of the corn example above, one might look at the studentized residual errors associated with the years of moderate and severe deviation and use those as a guideline for setting the critical values. In the second stage of the process, the trend yield equation is re-estimated including the four indicator variables that remove the effects of weather. The model specification for the two stages is as follows:

Stage I:

$Yield_{xj} = f(Trend)$

Where the variables are define as:
$Yield_{xj}$: Yield for crop x in country j
Trend: Equals 1 in year 1, 2 in year 2, etc.

From the studentized residual for observation i the following indicator variables are created:

DMGD1 = 1 if studentized residual >.50 and <1.00, 0 elsewhere
DMGD2 = 1 if studentized residual >1.00, 0 elsewhere
DMBD1 = 1 if studentized residual <-0.75 and >-1.25, 0 elsewhere
DMBD2 = 1 if studentized residual <-1.25, 0 elsewhere

Stage II:

$Yield_{xj} = f(Trend, DMGD1, DMGD2, DMBD1, DMBD2)$

The selection of the critical values for creating the indicator variables is arbitrary in the specifications above and is for illustration only. The actual selection of critical values depends on the results from Stage I and are specific to crop and country.

**<u>Emperical Results</u>**

Continuing with the corn example discussed above, the actual results will be discussed and evaluated here. From the Stage I, the results of estimating corn yields in the Corn Belt as a function of trend are the following:

Corn Yield in the Corn Belt - Simple Trend Regression Results

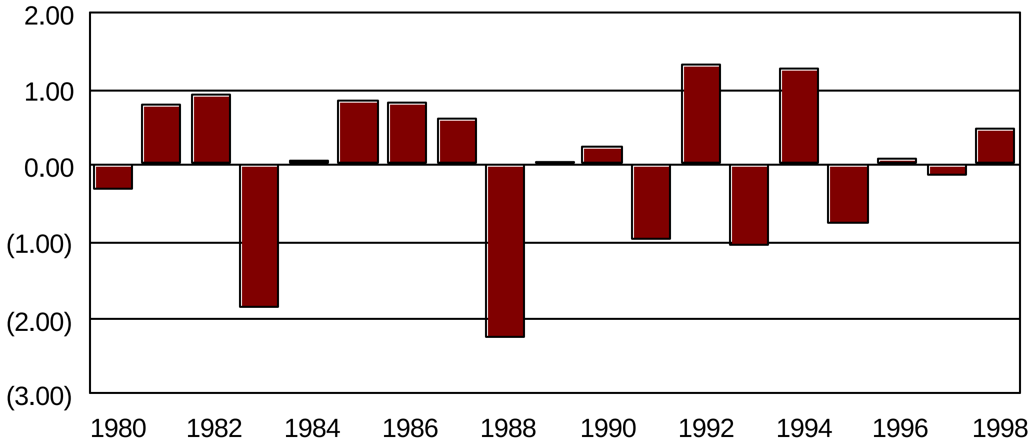
| Year | Trend | Actual Yield | Simple Trend Yield | Residual Errors | Standardized Errors | Studentized Errors |
|---|---|---|---|---|---|---|
| 1980 | 1 | 99.51 | 104.72 | (5.21) | (0.31) | (0.34) |
| 1981 | 2 | 118.66 | 106.36 | 12.30 | 0.74 | 0.79 |
| 1982 | 3 | 122.79 | 108.00 | 14.78 | 0.89 | 0.93 |
| 1983 | 4 | 79.14 | 109.65 | (30.51) | (1.83) | (1.89) |
| 1984 | 5 | 112.42 | 111.29 | 1.13 | 0.07 | 0.07 |
| 1985 | 6 | 127.08 | 112.93 | 14.15 | 0.85 | 0.86 |
| 1986 | 7 | 128.24 | 114.58 | 13.66 | 0.82 | 0.83 |
| 1987 | 8 | 126.13 | 116.22 | 9.91 | 0.60 | 0.60 |
| 1988 | 9 | 79.92 | 117.87 | (37.95) | (2.28) | (2.28) |
| 1989 | 10 | 120.15 | 119.51 | 0.64 | 0.04 | 0.04 |
| 1990 | 11 | 125.06 | 121.15 | 3.91 | 0.23 | 0.23 |
| 1991 | 12 | 106.40 | 122.80 | (16.40) | (0.99) | (0.99) |
| 1992 | 13 | 146.42 | 124.44 | 21.98 | 1.32 | 1.33 |
| 1993 | 14 | 108.47 | 126.09 | (17.61) | (1.06) | (1.07) |
| 1994 | 15 | 148.57 | 127.73 | 20.84 | 1.25 | 1.28 |
| 1995 | 16 | 116.92 | 129.37 | (12.45) | (0.75) | (0.77) |
| 1996 | 17 | 132.44 | 131.02 | 1.42 | 0.09 | 0.09 |
| 1997 | 18 | 130.38 | 132.66 | (2.28) | (0.14) | (0.15) |
| 1998 | 19 | 142.00 | 134.30 | 7.70 | 0.46 | 0.50 |

As expected, the results suggest a correspondence between the large residual errors and the years discussed above as very bad, bad, good, and exceptional (see graph on the following page). The severe drought years of 1983 and 1988 correspond to the studentized residuals greater than –1.75. The less severely affected years of 1991 and 1993 have studentized residuals close to –1. Likewise the exceptional years of 1992 and 1994 have studentized residuals greater than 1.3. Note that these exceptionally good years are not symmetric with the severe drought years. The moderately good years of 1981-82 and 1985-86 also stand out with studentized residuals ranging from .79 to .86. Based on these observations, the critical values that determine the values of the four indicator variables would be:

DMGD1 = 1 if studentized residual >.70 and <1.0, 0 elsewhere
DMGD2 = 1 if studentized residual >1.0, 0 elsewhere
DMBD1 = 1 if studentized residual <-0.75 and >-1.75, 0 elsewhere
DMBD2 = 1 if studentized residual <-1.75, 0 elsewhere

These critical values are still somewhat arbitrary but they reflect the natural breaks in the data. The years 1987 and 1998 also appear to be above average years but do not meet the critical values set above. The studentized residuals are 0.60 and 0.50 for 1987 and 1998,

Corn Yield in the Corn Belt
Simple Trend Regression Results
Studentized Errors




respectively. The studentized residuals for the other observations classified as moderately good years are clustered in the range of 0.79 to 0.93. There appears to be a natural break between the levels of the studentized residuals for those observations classified as moderately good years and the studentized residuals for 1987 and 1998.

Including the four new indicator variables in the regression produces significantly different results in the estimation of the trend. The graph on the following page illustrates the difference in the simple trend and the weather-adjusted trend. As expected the weather-adjusted trend has a slightly higher intercept than the simple trend. But what is surprising is the difference in the slope of the weather-adjusted trend when compared with the simple trend. The weather-adjusted trend has a significantly higher slope than the simple trend.

The estimated trend yields equations are the following:

Simple Trend Yield

COYHACB = 103.072 + 1.644 * Trend
(12.601) (2.291)

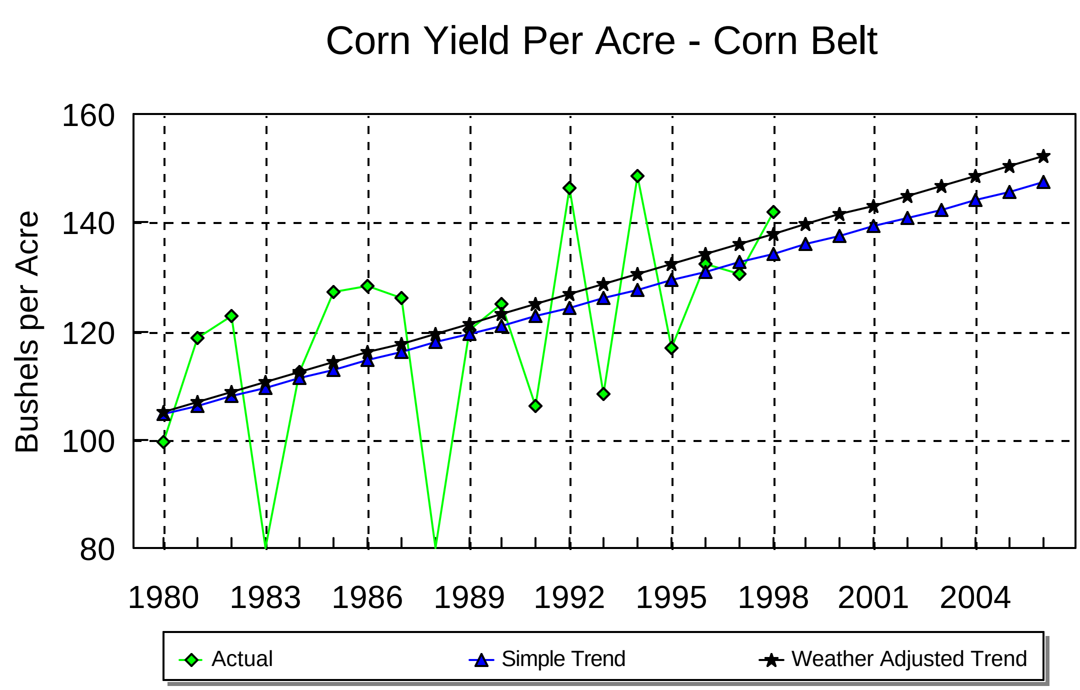
Weather Adjusted Trend Yield

COYHACB = 103.309 + 1.816 * Trend + 12.711 * DMGD1 + 18.761 * DMGD2
(36.7) (8.3) (4.5) (5.8)

-18.134 * DMBD1 – 35.580 * DMBD2
(-6.5) (-10.7)

These results further reinforce the earlier observation that the effects of weather are not symmetric for corn. In fact, the parameter estimates suggest that a severe drought has roughly twice the impact of exceptionally good weather.

Another convenient feature of the weather adjusted trend model is the ability to do



scenario analysis.  One could simulate each of the four possible weather situations simply by switching on the appropriate indicator variable.  This could be especially useful for assessing the risk to crop supplies and the policy implications of those risks.  In addition, the magnitude of the parameter estimates attached to the indicator variables provides information about the sensitivity of a particular crop to weather effects.  Comparing these parameter estimates across crops has implications for the level of yield risks across crops. These comparisons will be left to another paper.

**Evaluation**

Estimating the simple trend and weather adjusted trend lines does have some attractive aspects.  The trends are easy and methodical to estimate and require no additional data other than yields.  Estimates from the four indicator variables in the weather-adjusted trend may be useful in scenario analysis.  The methodology used in estimating the weather adjusted trend also removes the potential influence of outliers in the early or late years of the estimation that could bias the overall slope of trend line.

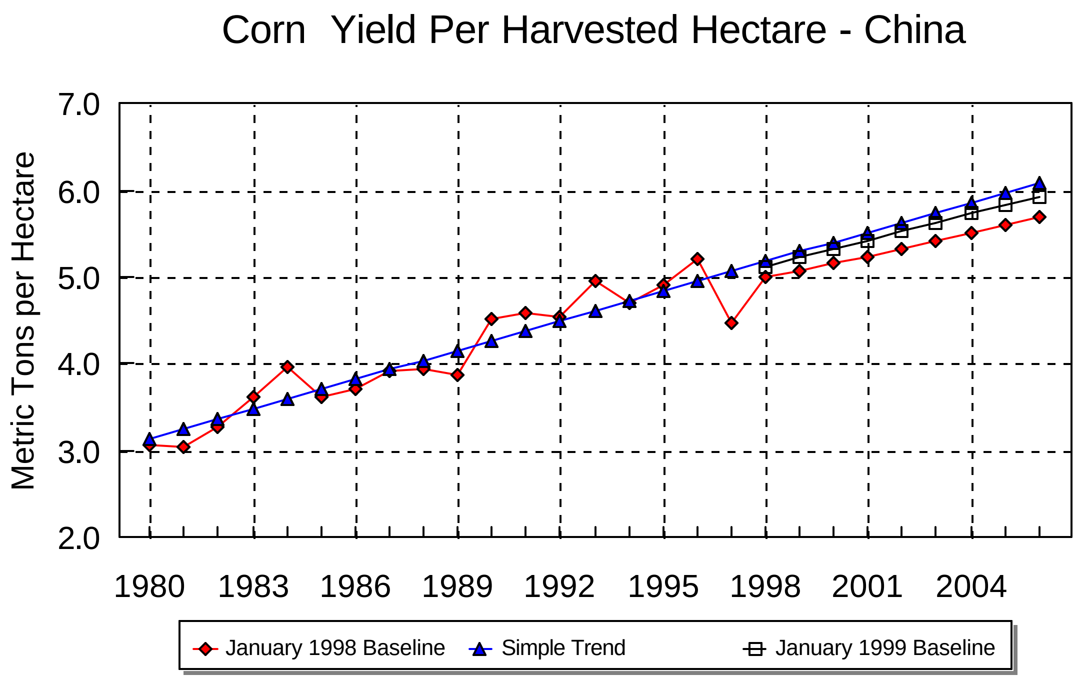
However, it is important to point out that the methodology of estimating the simple trend and weather adjusted trend is not foolproof.  The weather-adjusted trend is very sensitive to the selection of the critical values for creating the indicator variables. In addition, the lack of symmetry with respect to the good and bad weather indicators may be of concern. In the corn example presented, the weather-adjusted trend had a larger slope than the simple trend suggesting that corn yields have a greater downside variance.  In projecting the trend yield one year out, the weather-adjusted trend is the most appropriate indicator with no knowledge of weather events.  However, if one is considering a period of projections, such as a ten-year projection, the simple trend may be a better indicator because it factors in the effects of larger downside variance of yields.  Not using the simple trend could cause yields to be overestimated on average over the whole ten-year period.  Another way to solve this problem would be to make the indicator variable in the weather adjusted trend symmetric.   This could be easily done by creating two indicator variables that equal negative one in adverse weather years and positive one in good weather years.

## Summary

The method of estimating trend yields in a two-stage process to remove the effects of weather can significantly change the intercept and slope of the trend yield equation. The degree of this change depends upon how symmetric the effects of weather are on yield. In the perfectly symmetric case, removing the effects of weather will have little or no influence on the intercept and slope of the trend line. However, estimating weather-adjusted trends may be useful for scenario analysis.

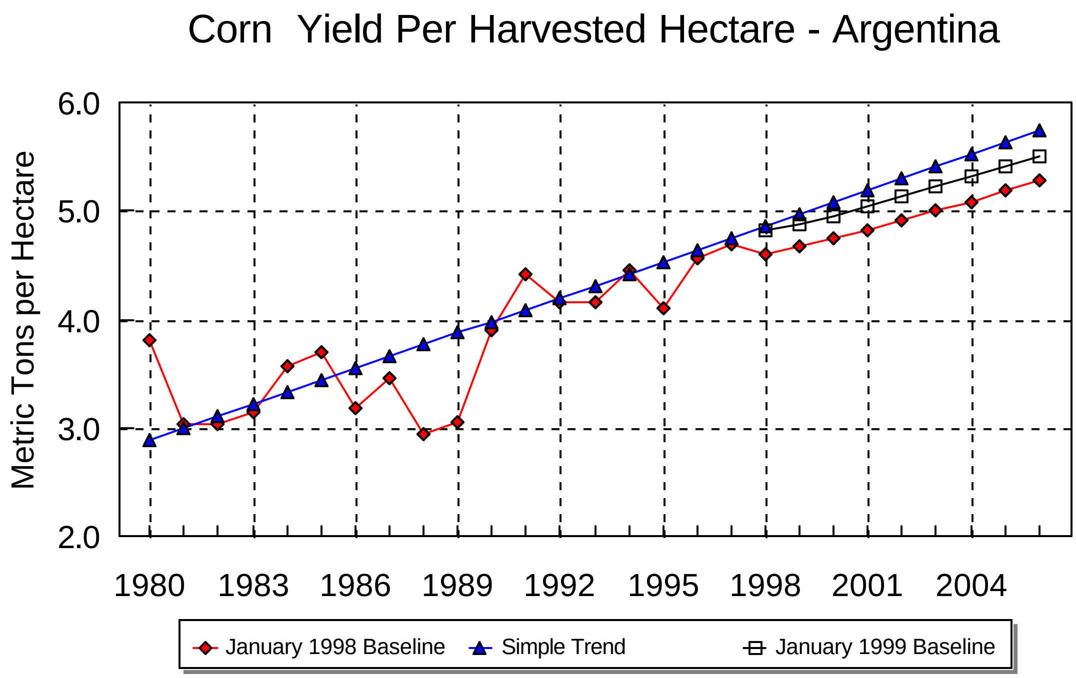
## Implications for the FAPRI Baseline

In the November 1998 and the January 1999 FAPRI Baselines the simple trend lines for all crops across all countries were revisited. Weather adjusted trends were not estimated in the formal manner discussed above. Instead a common set of critical values were used across all crops and countries. The result of not individualizing these critical values to each crop and country made the weather adjusted trends very similar to the simple trends. Thus, the simple trends were used to adjust the yield assumptions in most cases. Even so, some dramatic changes in the yield assumptions occurred. A few of the more significant changes are discussed in the following paragraphs.

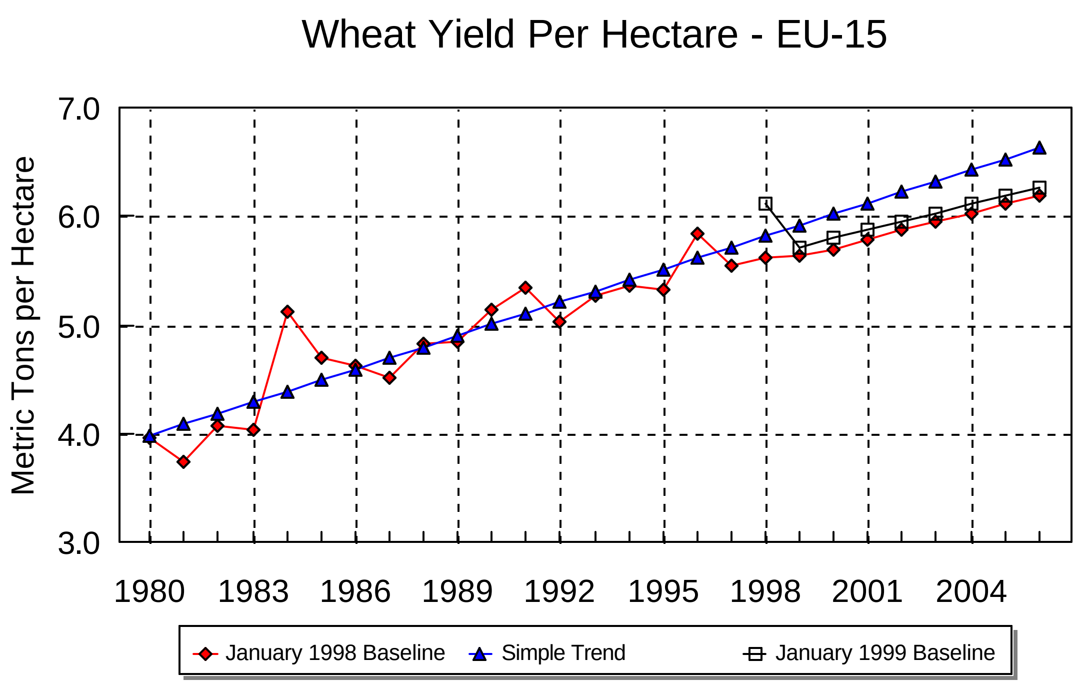



The China corn trend yield assumption is one of the most dramatic changes from the old January 1998 baseline. As the graph below illustrates, by 2006 there is a difference of 0.39 metric tons per hectare in yield assumptions based on the new simple trend. While this may not seem like a lot, China harvests about 24 million hectares of corn per year and that small yield change results in 9.4 million metric tons more corn production. Because the change was so different from last year's baseline, the trend yield assumption was not fully adjusted to the simple trend. However, the new assumption is very close to the simple trend. Despite their rapid growth in corn demand, this extra production keeps China from becoming a significant net importer of corn.

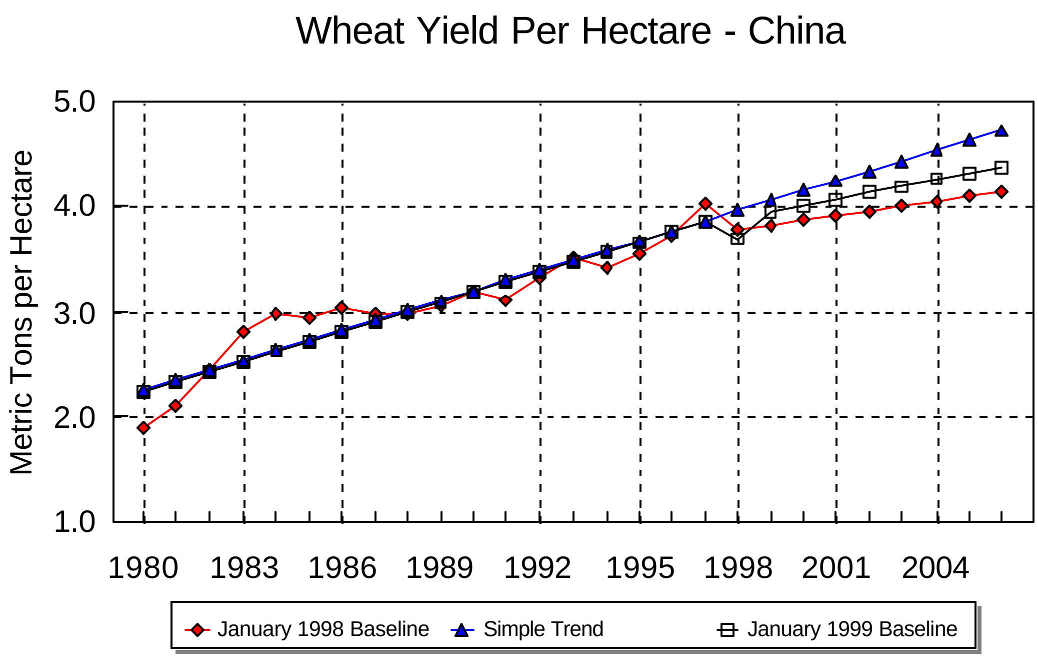
The Argentina corn trend yield assumption was also significantly changed, although only partially adjusted to the estimated simple trend. The increase in trend yield increases Argentina's corn production by 0.6 million metric tons.



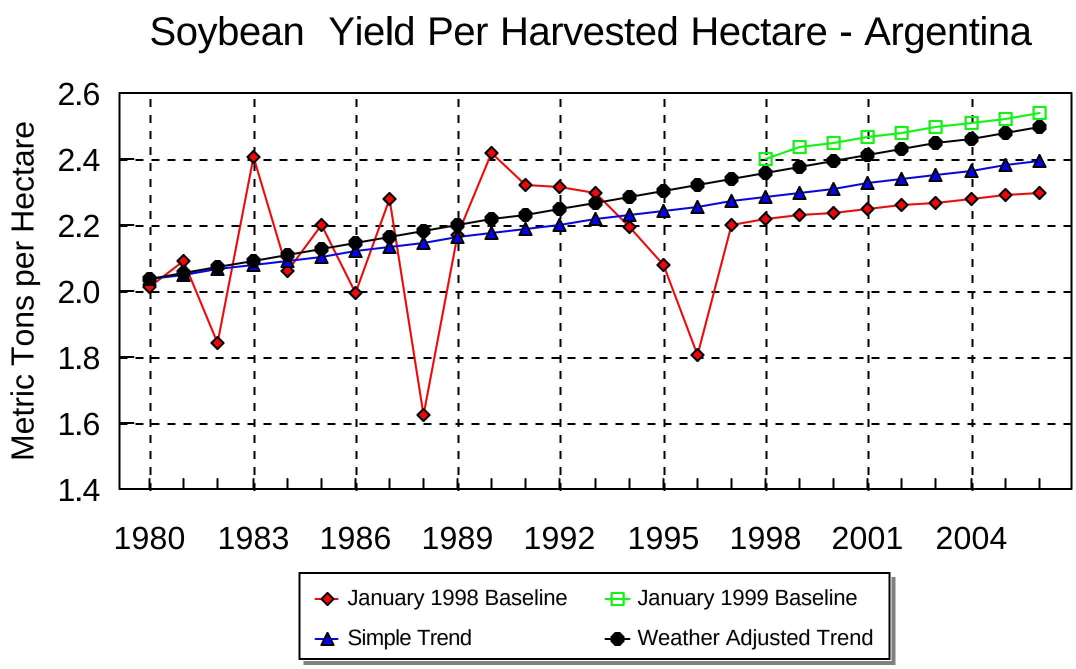

In the European Union, the simple trend also implied a significant revision in the wheat trend yield assumption. However, yields were only adjusted slightly upward in the new baseline. The increased wheat production implied by yields in the EU would be offset by higher set-aside rates so the net effect on production, and thus exports, is minor.



In China, the simple trend also implied a significant revision in the wheat trend yield assumption. As the graph illustrates, only part of the adjustment was taken here as well.




For soybeans, the most significant trend yield adjustments were made to Argentina (see graph on the following page.) Argentina yields were adjusted slightly higher than the weather-adjusted trend. The change in the soybean yield assumption increased soybean production by 1.8 million metric tons.




Trend yield assumptions for every crop in every country covered by the FAPRI system were reviewed. Some of the changes with the largest impacts were discussed above. Changes in the other yield assumptions can be found by comparing the FAPRI 1-99 and 2-99 Staff Reports with last year's versions.