Understanding the WAOB Crop Weather Model for Soybeans

Scott Irwin and **Todd Hubbs**

Department of Agricultural and Consumer Economics
University of Illinois

One of the key early assessments of the U.S. average soybean yield is the projection made by the USDA's World Agricultural Outlook Board (WAOB). Since 1993, those projections have been used in the May, June, and July WASDE reports to make supply, ending stocks, and price projections for the upcoming marketing year. While it is clear that the WAOB yield forecasts are perceived by market participants as containing important new information, these forecasts appear to be poorly understood by many and often confused with later forecasts released by the National Agricultural Statistics Service (NASS) of the USDA. The purpose of today's article is to describe the crop weather model that the WAOB uses to make soybean yield forecasts and demonstrate specifically how the 2020 forecasts were generated.

Analysis

To begin, it is important to understand that WAOB and NASS soybean yield forecasts are based on entirely different procedures. WAOB forecasts of national soybean yield are "model-based" and released as part of the May, June, and July WASDE reports each year. From 1993-2012, the WAOB forecasts were based on relatively simple trend analysis of historical yields, sometimes modified by planting progress (*farmdoc daily*, May 7, 2015). Since 2013, WAOB soybean yield forecasts have been based on a crop weather model as described in Westcott and Jewison (2013). NASS forecasts of state and national soybean yield are based on large scale farmer surveys and field measurement surveys. These forecasts are released in monthly crop production reports from August through November each year, with the final yield estimates released in January after harvest. The NASS soybean yield forecasts are released simultaneously with WASDE reports and the forecasts are used without adjustment in WASDE supply projections.

With that background, we can turn our attention to describing the crop weather model used to generate WAOB soybean yield forecasts. As detailed in Westcott and Jewison (2013), the WAOB crop weather model is a "Thompson-style" regression model. Specifically, the model of U.S. average soybean yield was originally estimated using data for 1988 through 2012 and included the following explanatory variables: i) a time trend variable to represent technological change, ii) June precipitation shortfall, iii) July-August average precipitation, and iv) July-August average temperature. The WAOB soybean model

does not include planting progress, different than the WAOB corn model (*farmdoc daily*, July 2, 2020). The weather data were collected for seven key soybean-producing States (Iowa, Illinois, Indiana, Ohio, Missouri, Minnesota, and Nebraska). These eight states typically represented about 70 percent of U.S. soybean production when the model was originally developed. An aggregate measure for the seven states was constructed using harvested soybean acres to weight state-specific observations.

Since extreme weather deviations from normal in June can have large impacts, as seen in 2012 and in 1988, the model uses a measure of the June precipitation shortfall from average in years when June precipitation is in the lowest 10-percent tail of its historical distribution. Since much of soybean reproduction occurs in July-August, both temperature and precipitation from those months were included in the model. Note that temperature and precipitation are averaged across July and August.

We collected yield, planting progress, and weather data for 1988-2019 following as closely as possible the definitions discussed above. We then computed acreage-weighted aggregates of the weather variables. Finally, we estimated the model using linear regression over 1988-2012 in an attempt to replicate the original model results found in Table 7 of Westcott and Jewison (2013). Our estimated results are compared to the original estimates provided by Westcott and Jewison in Table 1. We were unable to replicate exactly the original estimates but were able to come very close. The differences are fairly trivial and probably reflect small differences between the data sets used in the two regression estimations.

Table 1. Replication Results for WAOB Crop Weather Model for U.S. Average Soybean Yield, 1988-2012

	Model Coefficients	
	WAOB	**Estimated**
Intercept	60.100	62.659
Trend	0.447	0.452
June Precipitation Shortfall	-1.279	-1.283
July-August Temperature	-0.514	-0.526
July-August Precipitation	5.083	4.306
July-August Precipitation Squared	-0.619	-0.535
R2	0.800	0.799

Note: the WAOB model coefficients are drawn from Westcott and Jewison (2013).

The next step was to update the crop weather model estimates through 2019. Our conversations with WAOB staff indicate that an additional variable was added to the soybean model since the original estimation presented in Westcott and Jewison (2013). In particular, a dummy variable was added for 2003 to represent the yield loss due to aphids that year. The estimation results including this 2003 dummy variable are shown in Table 2. Even with the addition of the dummy variable, the coefficient estimates for the other variables over 1988-2019 do not differ greatly from those over 1988-2012 shown in Table 1, perhaps with the exception of the June shortfall variable. The coefficient estimates shown in Table 2 are presumably close to those used by the WAOB to generate the yield projection of 49.8 bushels per acre released in the May 2020 WASDE report. We now turn to the task of replicating this yield forecast using the estimation results shown in Table 2.

Table 2. Estimation Results for WAOB Crop Weather Model for U.S. Average Soybean Yield, 1988-2019

Regression Statistics	
Multiple R	0.952
R Square	0.906
Adjusted R Square	0.884
Standard Error	1.964
Observations	32

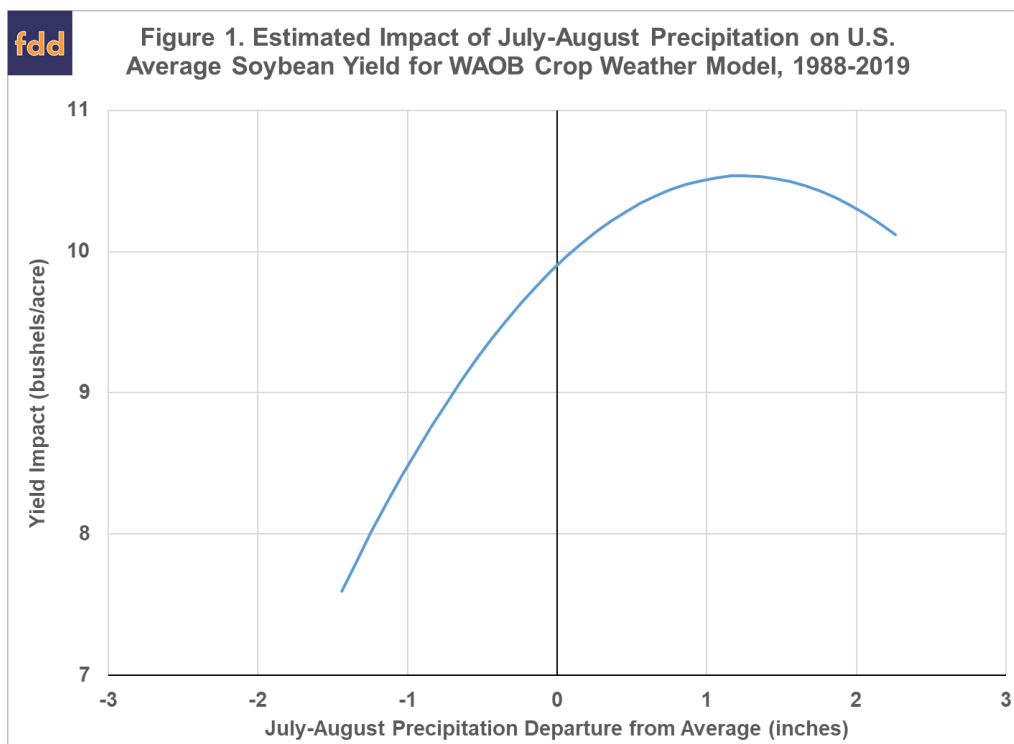|	*Coefficients*	*Standard Error*	*t Stat*	*P-value*
Intercept	59.173	17.431	3.395	0.002
Trend	0.523	0.039	13.461	0.000
2003 Dummy	-5.884	2.018	-2.916	0.007
June Precipitation Shortfall	-0.867	0.510	-1.699	0.102
July-August Temperature	-0.503	0.196	-2.560	0.017
July-August Precipitation	4.142	3.344	1.238	0.227
July-August Precipitation Squared	-0.407	0.389	-1.045	0.306

It would seem straightforward to estimate a 2020 soybean yield using the WAOB crop weather model. Simply take the coefficient variables found in Table 2 times the sample averages for each variable, sum the results, and generate the projection. Note that in the case of the June shortfall variable, the sample average is not literally zero but this zero is typically assumed unless a severe drought like 1988 or 2012 is expected. This procedure generates the following 2020 yield forecast:

	(1)	(2)	(3)=(1)X(2)
	Coefficient	**Sample Average**	**Product**
Intercept	59.173	1	59.173
Trend	0.523	33	17.274
2003 Dummy	-5.884	0	0.000
June Precipitation Shortfall	-0.867	0	0.000
July-August Temperature	-0.503	72.896	-36.659
July-August Precipitation	4.142	3.841	15.909
July-August Precipitation Squared	-0.407	14.752	-6.005
		Sum	**49.7**

The computed forecast is indeed very close the published WASDE forecast in May and June of this year of 49.8 bushels. Regardless, this is not exactly not how the WAOB computes its soybean projection.

A complication is introduced because the July-August precipitation variable is specified as a non-linear quadratic function. Figure 1 plots this estimated relationship using the coefficients presented in Table 2. For convenience, the x-axis is expressed in deviations from the sample average for July-August precipitation. The estimated "hill-shaped" relationship has the important property that the response to July-August average precipitation is asymmetrical. That is, large deviations below average reduce yield more than large deviations above average increase yield. This is the reason that one cannot plug average July-August precipitation into the model when making forecasts. The yield impact of average

July-August precipitation, zero in Figure 1, is not the average of all the yield outcomes plotted in the figure. The true average lies somewhere to the left of the zero line because poor July-August weather reduces yield more than good July-August weather helps yield.

Figure 1. Estimated Impact of July-August Precipitation on U.S. Average Soybean Yield for WAOB Crop Weather Model, 1988-2019

There are various methods to account for the non-linear impact of July-August precipitation when making yield projections with the crop weather model. One is to use Monte Carlo simulation, which is what we did last week in a different application of crop weather models (*farmdoc daily*, June 25, 2020). The WAOB instead determines the one-standard deviation range of July-August precipitation above and below the average, computes the yield prediction for all levels of July precipitation within this range, and then averages the resulting yield predictions. Westcott and Jewison (2013) indicate this procedure reduced the 2013 model yield projection by only 0.09 bushels per acre. They conclude that the soybean yield adjustment is relatively smaller than the similar adjustment for corn because soybean yields are less sensitive to weather conditions. This is a puzzling conclusion looking at Figure 1, which shows that soybean yield response to July-August precipitation is highly non-linear. An alternative explanation is that the particular adjustment procedure used by the WAOB does not account fully for extreme outcomes in the tails of the July-August precipitation distribution, and a larger adjustment will result if this is done.

We attempted to follow the WAOB non-linear adjustment for July-August precipitation as closely as possible. We did this by ranking July-August precipitation from low to high, determining the cutoffs for one-standard deviation above and below the average July-August precipitation, computing model predictions for each observation of July-August precipitation between the cutoffs, and then averaging all the resulting predictions. The revised computation for the 2020 soybean yield is shown below:

farmdoc daily   July 9, 2020

	(1) Coefficient	(2) Sample Average	(3)=(1)X(2) Product
Intercept	59.173	1	59.173
Trend	0.523	33	17.274
2003 Dummy	-5.884	0	0.000
June Precipitation Shortfall	-0.867	0	0.000
July-August Temperature	-0.503	72.896	-36.659
July-August Precipitation	4.142	3.841	15.909
July-August Precipitation Squared	-0.407	14.752	-6.005
July-August Precipitation Non-Linear Adjustment			-0.117
		Sum	**49.6**

Using these revised procedures we were not able to exactly replicate the published WASDE soybean yield forecast of 49.8 bushels released in May and June of this year. Our underlying data may differ slightly from that used by the WAOB and/or we may not have exactly replicated the non-linear adjustment for July-August precipitation used by the WAOB.

Implications

Soybean yield forecasts contained in May, June, and July WASDE reports from the USDA provide important information to market participants. The forecasts are based on a crop weather model maintained by the World Agricultural Outlook Board (WAOB). The model and procedures used by the WAOB to generate forecasts from the model appear to be poorly understood by many and often confused with later survey-based forecasts released by the National Agricultural Statistics Service (NASS) of the USDA. The WAOB crop weather model for soybeans is actually fairly simple, with trend, a 2003 dummy variable, and three weather variables. We collected the same yield, planting progress, and weather variables as listed in a 2013 report that first presented the WAOB model. Our model estimates using this data were reasonably close to the original estimates found in that report. We then went on to update the model estimates through 2019 and project the 2020 national average yield for soybean. After adjusting for the non-linear response of soybean yield to July precipitation, we were not able to exactly replicate the published forecast of 49.8 bushels per acre. However, our estimate of 49.6 bushels is very closer. We hope this exercise will increase understanding of the WAOB crop weather model and procedures used to generate these important and market-moving soybean yield forecasts.

References

Irwin, S., D. Good and D. Sanders. "Understanding and Evaluating WAOB/USDA Soybean Yield Forecasts." *farmdoc daily* (5):84, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, May 7, 2015.

Irwin, S. and T. Hubbs. "Understanding the WAOB Crop Weather Model for Corn." *farmdoc daily* (10):121, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, July 2, 2020.

Irwin, S. and T. Hubbs. "Unconditional vs. Conditional Trend Estimates for U.S. Soybean and Soybeans." *farmdoc daily* (10):116, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, June 25, 2020.

Westcott, P.C., and M. Jewison. "Weather Effects on Expected Soybean and Soybean Yields." FDS-13g-01, Economic Research Service, U.S. Department of Agriculture, July 2013. https://www.ers.usda.gov/publications/pub-details/?pubid=36652