



What Can We Learn About Corn and Soybean Yields From Crop Tours?

Scott Irwin, Darrel Good, and Gary Schnitkey

Department of Agricultural and Consumer Economics
University of Illinois

August 27, 2014

farmdoc daily (4):163

Recommended citation format: Irwin, S., D. Good, and G. Schnitkey. "[What Can We Learn About Corn and Soybean Yields From Crop Tours?](#)" *farmdoc daily* (4):163, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, August 27, 2014.

Permalink URL <http://farmdocdaily.illinois.edu/2014/08/learn-about-corn-and-soybean-yields-from-crop-tours.html>

It almost goes without saying that corn and soybean prices are heavily influenced by the size of the U.S. crop and that expectations about crop size are especially important for prices during the growing season. Market participants rely on a variety of information in order to form yield and production expectations. That information includes general growing season weather conditions, weekly USDA reports of crop conditions, and USDA survey-based forecasts beginning in August. Every year, there is no shortage of discussion of about evolving prospects for corn and soybean yields and 2014 has certainly not been an exception. Several recent *farmdoc daily* articles have been devoted to this topic ([July 9, 2014](#); [July 23, 2014](#); [August 20, 2014](#)).

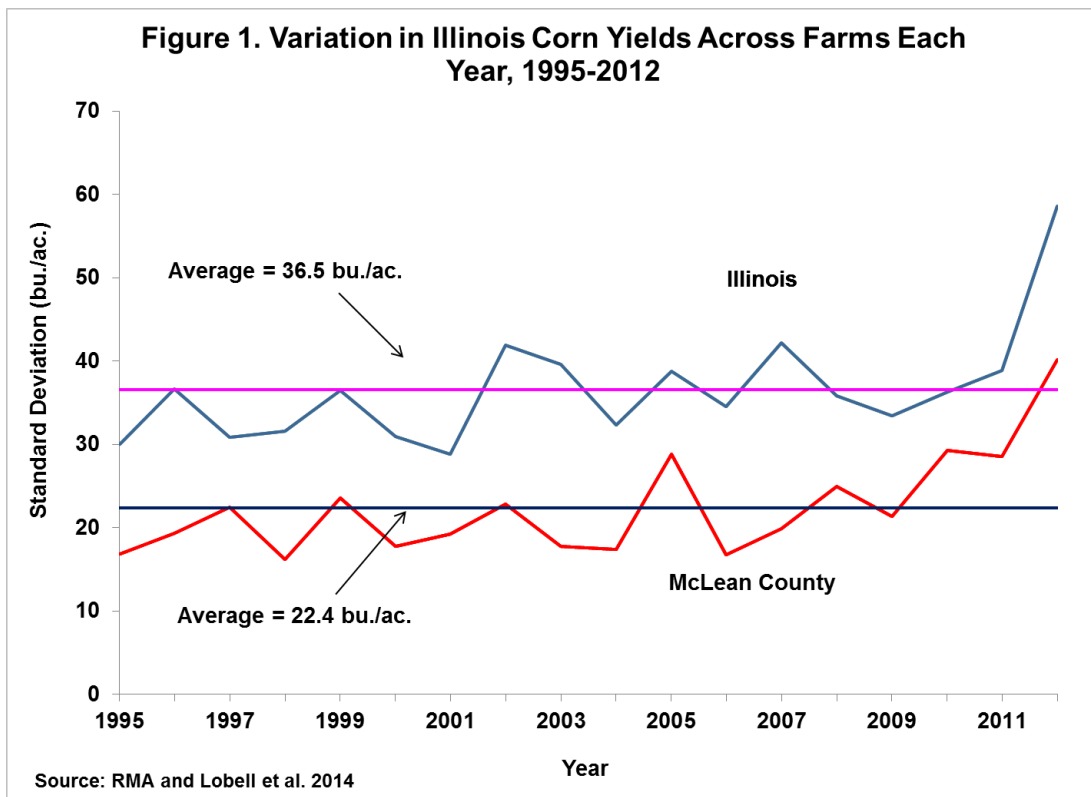
Late in the growing season, "crop tours" are conducted by a variety of entities to collect corn and soybean samples and make yield estimates. The scope of those tours varies considerably, with most focused on a relatively small geographic area such as a county. [A few of the tours](#) are for larger areas such as an entire state or even for the country as a whole and appear to have a notable market impact as results are released. The methods and procedures used in these tours also appear to vary considerably. For the most part, not much guidance is provided on how to interpret the results of these surveys. The purpose of this article is to provide some guidance about what can and cannot be learned about the corn and soybean yield estimates generated from crop tours.

We begin with an examination of one of the more important considerations in evaluating yield tour results—the size of the sample used in deriving the estimates. The sample size needs to be large enough to represent the region and to provide a reasonable level of confidence in the accuracy of the estimates. The size of the sample required, then, depends on the range in yields targeted, the variability of yields within the region, and the level of confidence that can be placed in the estimates. Larger sample sizes are required for narrower ranges of yield estimates and for areas with large yield variability. In addition, larger sample sizes are required for higher levels of confidence in the estimates.

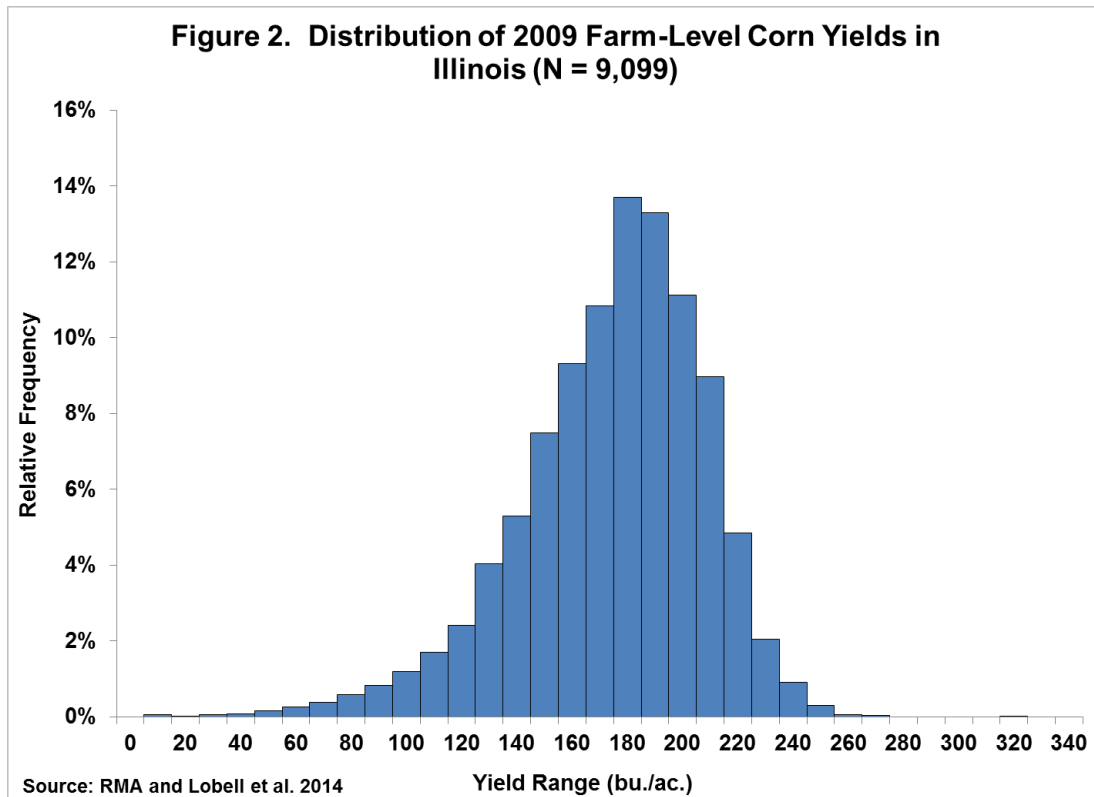
The statistical formula used to determine the [required sample size](#) for an estimation problem is actually quite simple. The main difficulty in applying the formula is specifying the correct population variance; in

We request all readers, electronic media and others follow our citation guidelines when re-posting articles from *farmdoc daily*. Guidelines are available [here](#). The *farmdoc daily* website falls under University of Illinois copyright and intellectual property rights. For a detailed statement, please see the University of Illinois Copyright Information and Policies [here](#).

this case, the variability of yields for a given year across the geographic area under consideration. Fortunately, Professor David Lobell of Stanford University and his co-authors recently made available to the public a [database on farm-level yields](#) from the USDA Risk Management Agency (RMA) that can be used to make reasonable estimates of the variability in corn yields across farms in Illinois. The database provides average corn yields for 100 randomly selected farms (insured units) for each county in Illinois from 1995 through 2012. Pooling across all observations in each year, the variation in farm yields in Illinois from 1995 through 2012 (as measured by [standard deviation](#)) ranged from 28.9 to 58.6 bushels per acre (Figure 1). The average variation was 36.5 bushels per acre. Similarly, the annual variation in farm yields in our example county, McLean County, Illinois, ranged from 16.3 to 40.1 bushels and averaged 22.4 bushels. It is no surprise that the variation in yields within a year is much larger for the state of Illinois than for a particular county, as the state is a much larger geographic area containing considerably more variation in soils and weather conditions than a county. It is also not surprising that yield variability increases substantially in years with poor crops, like 2012. Relatively small differences in weather conditions, particularly precipitation, can be the difference between poor yields and truly disastrous yields in such years.

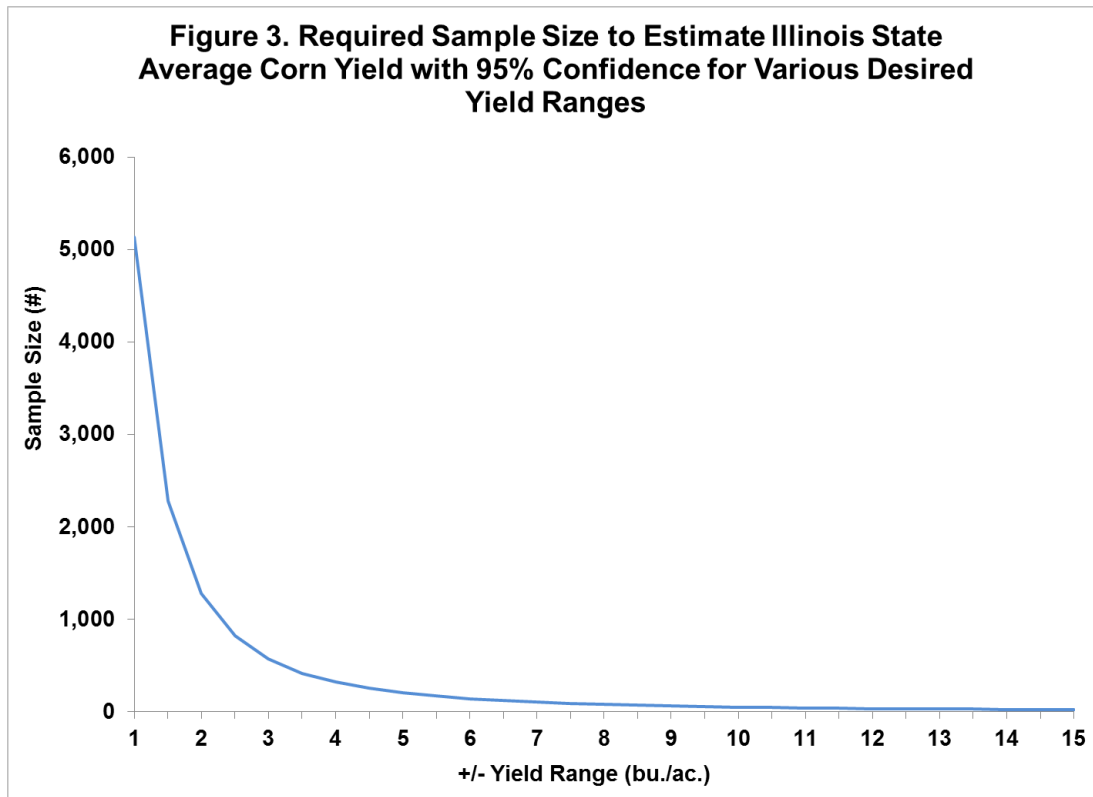


For any given year, the variability of yields shown in Figure 1 are likely greater than many expect. The full annual variation in yield is illustrated in Figure 2, which shows the frequency distribution of farm yields in Illinois during 2009. Note that about 50 percent of the yield observations in 2009 were below 150 bushels and above 190 bushels. A non-trivial percentage of the yields were below 100 bushels and above 230 bushels. The key point to take away is that the number of fields that need to be sampled is relatively large in order to ensure that the variability across the target region is adequately reflected.



As the next step in the analysis, we use these measures of yield variability to calculate the required sample size for alternative levels of desired accuracy. At both the state and county level, samples are assumed to be drawn from a yield distribution with a mean of 190 bushels. We assume that yield variability is the average standard deviation for the state (36.5 bushels) and McLean county (22.4 bushels) over 1995-2012. We also assume a target of 95 percent confidence that the actual yield for the region is within +/- 3.8 bushels (or 2 percent) of the estimated yield. At the state level under these assumptions, 355 yield samples would be required. At the county level, with the smaller yield variability of 22.4 bushels, only 133 samples would be required to provide the desired level of accuracy. Of course, if the range of estimated yields is increased and all else is held constant, then fewer samples will be needed. For example, assume a target of 95 percent confidence that the yield estimate is within +/- 9.5 bushels (5 percent). The required sample size drops dramatically, with only 57 samples needed at the state level and 21 at the county level.

The effect on required sample size of the desired range of yield estimates, or yield accuracy, is illustrated in Figure 3 for the state of Illinois (assumed yield variability of 36.5 bushels). The range of yield estimates for a 95 percent confidence level is plotted on the horizontal axis and the required sample size is plotted on the vertical axis. For example, if a relatively narrow yield range is specified, such as +/- two bushels, the required sample size to assure a 95 confidence level is 1,283. The required sample size increases exponentially with the level of desired accuracy. Reducing the desired range of yield estimates to +/- one bushels requires a huge sample size of 5,182.



What do these calculations tell us about interpreting yield estimates from crop tours? First, most crop tours do not provide information on the degree of accuracy targeted or the confidence level assumed. Typically, all that is provided is the number of samples and the average yield estimate from the samples. It is not clear from the published accounts of crop tours how much thought is put into these crucial issues. Interestingly, yield databases like that of Lobell et al. potentially make it a trivial exercise to provide information on confidence levels given a desired level of accuracy and the sample size. Second, there is an unavoidable trade-off between the desired accuracy of the estimates (yield range) and sample size all other things held constant. There is no “free lunch” when it comes to generating statistical survey estimates. We suspect that most observers would be surprised at either the low confidence levels, low levels of accuracy, or both, implied by the sample sizes used in many (but not all) crop tours.

The previous analysis and discussion assumes that standard statistical procedures are followed to draw the samples and best practices are followed in generating the actual measurements of yield. Again, it is difficult to know how well most crop tours measure up to these standards. For example, samples should be drawn randomly within the entire region and in a predetermined random area of the sample fields. If this is not the case, then tour estimates are even less reliable than implied by the previous calculations. In addition, the potential for non-sampling errors in the yield estimates should not be overlooked. The required steps and potential pitfalls of the yield estimation methods used in most crop tours for corn are well-described in this [recent article](#) by Professor R.L. Nielsen of Purdue University. Potential errors can occur if all measurements are not done correctly, if the appropriate factor is not used in converting kernel counts to grain weight, if yield estimates are based on various moisture levels, or if some adjustment for harvest loss is not made. Professor Nielsen cautions, “Remember that this method for estimating pre-harvest grain yield in corn indeed provides only an estimate. Since kernel size and weight will vary depending on hybrid and environment, this yield estimator should only be used to determine “ballpark” grain yields.” Furthermore, there is a [large literature](#) that formally evaluates the accuracy of so-called “crop cutting” yield estimates, and there is widespread evidence that such techniques have a tendency to over-estimate actual yields for a variety of reasons. This is borne out by the experience of the USDA with their objective yield surveys. According to a [recent study](#), USDA objective yield estimates for corn are

biased upwards by about 12 to 15 percent (it is not clear if the statement applies to state or national estimates and the bias is approximately offset by an opposite downward bias in farm operator yield estimates).

Implications

Late-summer corn and soybean yield estimates from crop tours have gained increasing prominence in recent years. Unfortunately, most crop tours do not provide information on the degree of accuracy targeted or the confidence level assumed; so, it is difficult to formally assess the reliability of their yield estimates. We suspect that most observers would be surprised at either the low confidence levels, low levels of accuracy, or both, implied by the sample sizes used in many (but not all) crop tours. In addition, the accuracy of crop tour estimates may be compromised by non-random sample selection and non-sample measurement errors. In view of these limitations, what can be learned about corn and soybean yields from crop tour estimates? Our position is that most crop tour estimates provide useful qualitative information about crop prospects (e.g., average, good, bad), but quantitative estimates (e.g., Illinois state average yield of 190 bushels) are not likely to have a high degree of reliability.

References

Irwin, S., D. Good, and J. Newton. "[Do Big Corn Crops Always Get Bigger?](#)" *farmdoc daily* (4):156, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, August 20, 2014.

Irwin, S. and D. Good. "[The 2014 U.S. Average Soybean Yield: Headed for a New Record?](#)" *farmdoc daily* (4):137, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, July 23, 2014.

Irwin, S. and D. Good. "[The 2014 U.S. Average Corn Yield: Big or Really Big?](#)" *farmdoc daily* (4):127, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, July 9, 2014.

Lobell, D. B., M. J. Roberts, W. Schlenker, N. Braun, B. B. Little, R. M. Rejesus, and G. L. Hammer. "Greater Sensitivity to Drought Accompanies Maize Yield Increase in the U.S. Midwest." *Science* 344(2014):516-519. Supplementary materials available online: <http://www.sciencemag.org/content/344/6183/516/suppl/DC1>.

Nandram, B., E. Berg, and W. Barboza. "A Hierarchical Bayesian Model for Forecasting State-Level Corn Yield." *Environmental and Ecological Statistics* 21(2014):507-530.

Newman, J. "Corn, Soybean Yields Shoot Higher: Harvests Are Shaping Up as the Biggest Ever, Pressuring Prices." *The Wall Street Journal*, August 20, 2014, Accessed August 27, 2014. <http://online.wsj.com/articles/corn-soybean-yields-shoot-higher-1408572785>.

Nielsen, R.L. "Estimating Corn Grain Yield Prior to Harvest." Corny News Network, Purdue Extension, August 2014. Online at: <http://www.agry.purdue.edu/ext/corn/news/timeless/yldestmethod.html>.

Poate, D. "A Review of Methods for Measuring Crop Production from Smallholder Producers." *Experimental Agriculture* 24(1988):1-14.

Wikipedia. "Sample size determination." Accessed August 27, 2014. http://en.wikipedia.org/wiki/Sample_size_determination.

Wikipedia. "Standard deviation." Accessed August 27, 2014. http://en.wikipedia.org/wiki/Standard_deviation.