Improving Accuracy of Yield Maps Through Calibration and Post-Harvest Data Processing

Collecting and mapping yield data can provide significant value to your operation by enhancing decision-making and improving efficiency. Studies analyzing USDA surveys on technology adoption have shown that producers who use yield monitors experience cost savings, with further benefits for those who adopt yield mapping. However, due to the nature of the sensors used in data collection, yield monitors must be regularly calibrated, and the data collected must be checked for errors (i.e., post-processed) to prevent flawed decision-making.

A sample of combines in Mississippi, during the 2023 and 2024 harvest seasons, found that calibration errors ranged between 6 percent and 58 percent. Furthermore, studies on post-processing have found that 13–27 percent of those data have errors that should be removed in post-processing. Collecting and mapping yield data allows you to analyze yield trends within fields and over time, which serves as the foundation for precision agriculture. Having accurate data is invaluable and makes calibrating and post-processing a worthwhile effort.

Understanding Yield Estimation and Potential Errors

Yield is estimated from multiple sensors and user inputs. Errors in any of those sensors/ inputs can lead to incorrect estimates of yield or errors in the location of the estimated yield. The yield at a point in the field is calculated based on the following equation and logged along with GPS coordinates:

$$\text{Yield}\left(\frac{\text{bu}}{\text{acre}}\right) = \frac{m \times t}{d \times w \times p} \times \frac{100 - \text{MC}_{\text{combine}}}{100 - \text{MC}_{\text{market}}} \times \frac{43{,}560\,\text{ft}^2}{\text{acre}}$$

where:

- **m** is mass flow rate (lb/s)
- **t** is logging interval (s)
- **d** is distance traveled in logging interval (ft)
- **w** is header width (ft)
- **p** is grain density (lb/bu)
- **MC** is moisture content (%) from the combine or marketable rates

The mass flow rate is determined by a sensor, with the most common type being impact plate sensors, which have been commercially available since the mid-1990s. The relationship between the force of the grain striking the plate and the weight of grain is found through calibration. This relationship can change over time due to wear and tear on the sensor and with changing grain properties (e.g., moisture, test weight, foreign material).

To determine the relationship between the sensor and the actual grain flow, at least two measurements of the combine-estimated weight and actual load weight are needed to capture the response of the sensor to different rates of crop flow. These changes in flow will happen during harvest due to changes in yield and/or speed. As yield or speed is reduced, the crop flow rate will also be reduced. Similarly, as yield or speed is increased, the crop flow rate will increase. Typically, the more points used in calibration, the more accurate the estimation of yield.

Figure 1 shows an example of the relationship between predicted and actual weight using two or multiple calibration points. In this example, the actual weight is underestimated when using only two points over most of the range of expected values (overestimates can also occur). However, some monitors do not allow for multipoint calibration, and some yield monitoring technology only requires single calibration points. Consult your operator's manual for the specifics of your yield monitor. There are other monitoring systems that rely on optical sensors or flow sensors, which may have different calibration processes outlined in their manuals.

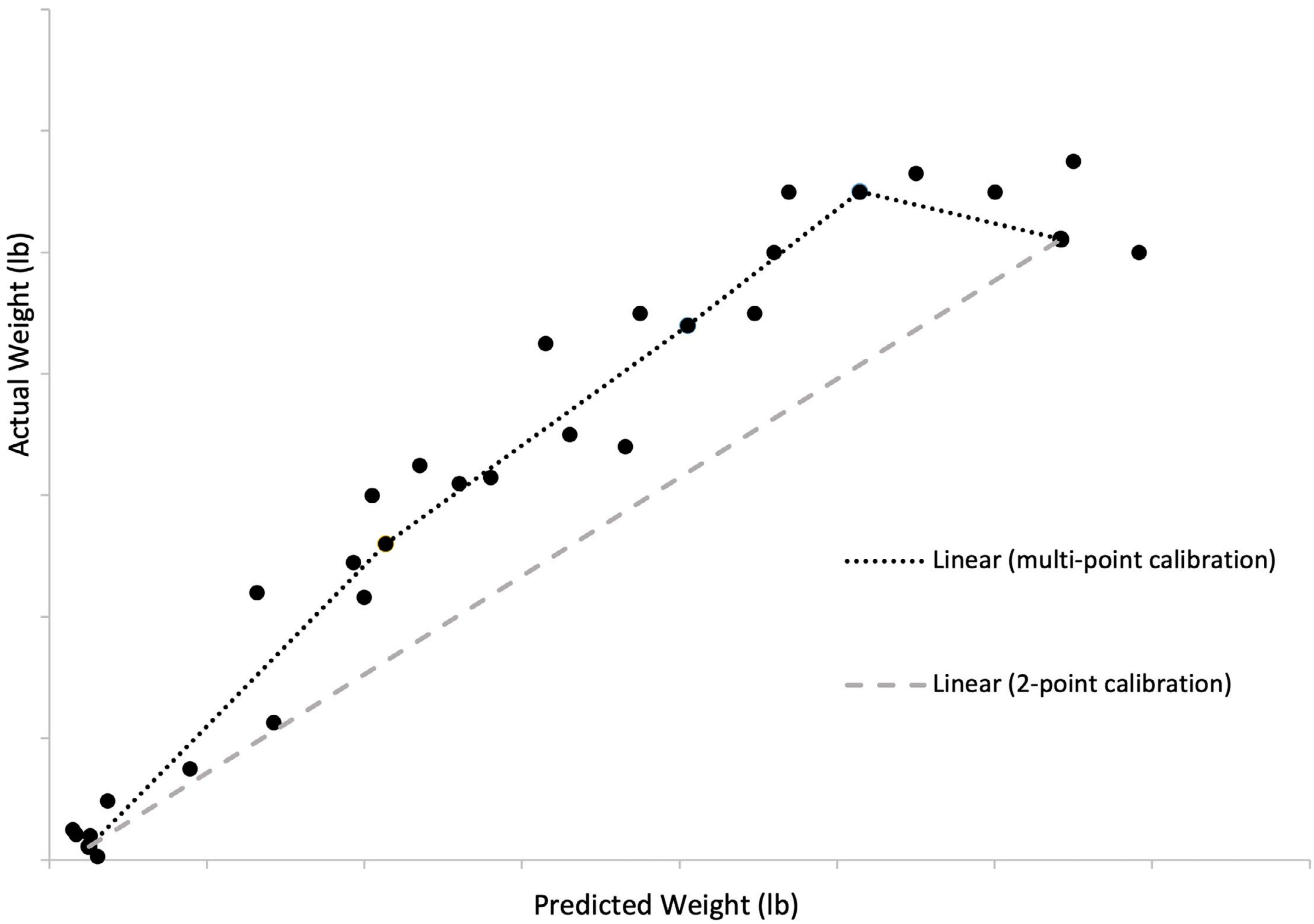

Figure 1. Example of calibration using two-point and multipoint methods. In this example, the actual weight is underestimated when using only two points over most of the range of expected values.

In addition to estimating the mass flow of grain, there are several other important values used in the estimation of yield that are collected by onboard sensors or entered by the operator. (See the equation in the section titled Understanding Yield Estimation and Potential Errors.)

- The logging interval is typically a default setting but may be adjusted by the user.
- The distance traveled is calculated from global navigation satellite system (GNSS) coordinates (or could be determined by a mechanical sensor as a backup). Rapid changes in travel distance can result in data errors as the travel distance becomes small.
- Header width is entered by the operator or, in some combines, may be automatically adjusted based on the overlap with adjacent passes. If header width is not correctly input, the resulting yield estimates will be in error. For example, yield can be underestimated in point rows and other locations where a full header width is not harvested.
- Similarly, errors in GNSS data collection can also result in inaccurate estimates of swath width or speed, which lead to incorrect estimates in yield. (See Purdue University's Wandering Swath Width Syndrome: Yield Monitor Errors.) This problem is most pronounced when not using differential global positioning system (DGPS) technology such as real-time kinematic positioning (RTK).
- Grain density is a conversion factor and is typically assumed at marketable values (e.g., 56 pounds per bushel for corn).
- Moisture content is typically measured via an onboard sensor that also requires calibration.

There are many potential sources of error in estimating yield that must be corrected via calibration or be removed during post-processing.

There are also errors that can occur based on how the data are collected. Typically, these types of errors need to be filtered out after data collection via specialized software; this is called “post-processing” or “cleaning” of yield data.

For example, due to the time it takes for grain to travel from the header to the impact plate at the top of the clean grain elevator, there is a data delay between where the crop was harvested and where it was measured. Also, very low speeds or fast braking can cause yield values to increase rapidly over a small area. Very high speeds may show the combine is traveling with the header down but not harvesting (i.e., deadheading). Data collected at the start and end of each pass are also more likely to have errors as the grain flow starts and stops.

Many of these issues are unavoidable during combine operation, making post-processing of data a critical step to ensure accuracy, even if the combine has been calibrated.

To ensure yield maps are accurate, both calibration (prior to and during harvest) and post-processing of data are needed. Without these steps, it is impossible to accurately compare yield data between combines or years to conduct on-farm research. The ability to compare treatments (e.g., seeding rate, varieties) on-farm can allow you to tailor practices to the conditions on your farm.

To ensure calibration does not slow down harvest, it is crucial to plan ahead and become familiar with the user manual for your yield-monitoring system. Also, be sure to regularly download and back up data to ensure your hard work is not lost. The following sections detail the process and offer some guidance on calibration and post-processing procedures.

Calibrating Yield Monitors

Careful planning before harvest will ensure the calibration process is efficient. The first step is to familiarize yourself with the user manual so that you can gather all the necessary equipment and plan for when, where, and how to perform the calibrations. In addition, consistently naming fields and having field boundaries in the monitor during harvests will aid in the analysis of data by ensuring data are only collected within field boundaries and are assigned to the correct field. Before beginning the calibration process and throughout the harvest season, it is important to inspect and clean all sensors. The following sensors should be calibrated following the user manual’s detailed procedures to ensure proper yield data collection. Calibration will ensure accurate yield estimations that can be used in future decision making.

Mass Flow Sensor

Calibrating the mass flow sensor is the most involved step in the process. Exact procedures vary by yield monitor, so it is important to consult your user manual. Some manufacturers have technology available to perform automated calibrations of the mass flow sensor, with the potential for manual offset adjustment following the procedures below. However, other sensors must still be manually calibrated following manufacturer guidelines.

Calibration is done by comparing the estimated weight of harvested grain to the weight measured by a scale or weigh wagon. If you plan to use a weigh wagon for calibration, you will also need to check the accuracy against a certified scale and plan to use the same wagon for the entire calibration process. During the calibration process, the weigh wagon should be parked in the same level spot for all calibration loads (Figure 2). If you do not have a weigh wagon, your local co-op or Extension office may be able to loan you one or help you find one.

Depending on the system, anywhere from one to six loads are recommended for calibration. Calibration loads are typically between 3,000 and 8,000 pounds, and each load should be a different flow rate of grain for multipoint calibrations. For single-point calibrations, the loads will be larger.

Figure 2. A combine unloads into a weigh cart as part of the calibration process.

The most common method to vary the flow rate of grain is changing the harvest speed, but the swath width can also be changed. As speed and swath width are increased, the flow rate of grain is also increased. Select speeds to cover the range of values you expect to use while harvesting.

When harvesting calibration loads, it is best to find level ground with low weed pressure. You also want to avoid starting and stopping as much as possible as this can introduce more errors.

At a minimum, calibration should be conducted for each crop at the start of each season. Calibration is also recommended when there are large differences in moisture content (e.g., high- versus low-moisture corn, a rain event), changes in test weight of more than a few points, or large swings in temperature. Many yield-monitoring systems will allow you to save multiple calibrations.

Temperature Sensor

This calibration should be done at the start of harvest on a combine that has been sitting for several hours out of direct sunlight. The air temperature should be measured at the combine location. Your local Extension agent can help you find a suitable temperature sensor. Correct temperature estimates will affect the accuracy of other sensors used to estimate yield.

Moisture Meter

The combine moisture meter should be compared against a high-accuracy, calibrated moisture meter. You can check the reported accuracy in the user manual, but note that this can drift over time, making the sensor less accurate. It may be worthwhile to bring some samples to the elevator or compare your meter against the certified elevator meter ahead of time to determine the relationship between your meter reading and the true moisture value.

Sample grain from several points in a calibration load and average them to get a representative sample. Since most yield monitors allow for only one point to be entered, the accuracy of this calibration can be poor and must be checked often. Check calibration at the start of the season and often throughout the season, especially between crops and with large changes in the grain moisture content.

Vibration

It is important to account for the effect of the combine's vibrations on estimating yield. This should be done for each header type, with the combine empty of grain and operating at full revolutions per minute (RPM) in operating position. Consult the user manual for specific details.

Header Height

To ensure that data are collected only when harvesting, it is important to ensure the header switch is correctly recording the up and down positions. The procedure will vary by combine, so consult the user manual.

GNSS Setup

All yield data are stored along with GNSS coordinates, so ensuring the system is set up correctly is a crucial factor in yield data collection. A common issue that can contribute to yield data error is improper measurements in the offsets of the GNSS receiver. This can cause a discrepancy in the true and recorded location of the estimated yield. Refer to the system user guide for information on measuring and entering offsets. Using DGPS technology, such as RTK, will lead to better estimates of the combine position during harvest.

Post-Processing Yield Data

Errors will still exist within the data even if the combine is calibrated and carefully operated. Software is available to post-process (also referred to as "clean") data to filter out and/or correct these errors based on known issues. The specific methods, number of options, and level of automation for post-processing will vary based on software type.

Many farm-management software packages also have a limited set of post-processing tools. USDA has a tool, Yield Editor, with a wider variety of specialized filters and tools for removing and visualizing errors in the data and correcting data delays.

There is no one-size-fits-all protocol for post-processing data. User experience and interpretation of the patterns seen in the map will guide the process. While this extra step can add time to the creation of yield maps, it has been shown to improve the quality of data and, thus, decision making. This task can be done yourself or by a professional service.

The following sections summarize the main sources of errors and methods to remove them.

Grain Flow Delay

The delay between when the grain enters the header and is measured by the sensor can be several seconds. This leads to a shift in the data between the yield map and the true value in the field. It can be especially noticeable across features such as waterways; a wave-like pattern can appear when the combine passes over areas with large changes in yield in different directions.

This delay value can be selected by trial and error or by automated features in some software. A default value of 12 seconds is typically applied, but this may not be the best value for all combines. As seen in Figure 2, the default delay still leads to the presence of a wavelike pattern in the data along a waterway.

Moisture Delay

There is also a delay in grain traveling to the moisture sensor. This should be corrected to improve the accuracy of yield prediction.

Travel Speed

Rapid changes in travel speed can cause errors in yield estimates. Braking rapidly, such as for an obstacle or clogged header, can cause spikes in yield estimates. Looking at the equation on page 1, you can see that, as the value of d decreases, the estimated yield increases.

Travel speed errors can be filtered by looking at high yield values, short travel distances, and/or rapid changes in speed. The filters used will depend on the capabilities of the software. An example of a yield spike surrounded by low yields is shown by the black circle in Figure 3.

Swath Width

If swath width is not correctly entered by the user or errors in the GNSS signal occur, the estimated yield will be affected. In areas such as point rows where the full header width is not harvested, yield will be underestimated if the full swath width is not corrected. Similarly, if a GNSS error leads to an incorrectly reported overlap in passes and the swath width is automatically adjusted lower, yield will be overestimated. In these cases, swath width could be manually adjusted using text editing or spreadsheet software, or the incorrect data points can be removed via filtering.

In Figure 3, the arrow indicates a swath of low yield, likely due to harvest of a partial header width. Without knowing the true swath width, the data cannot be corrected and should be removed from analysis. Ensuring correct swath widths are entered during harvest can preserve data and lead to more complete yield maps.

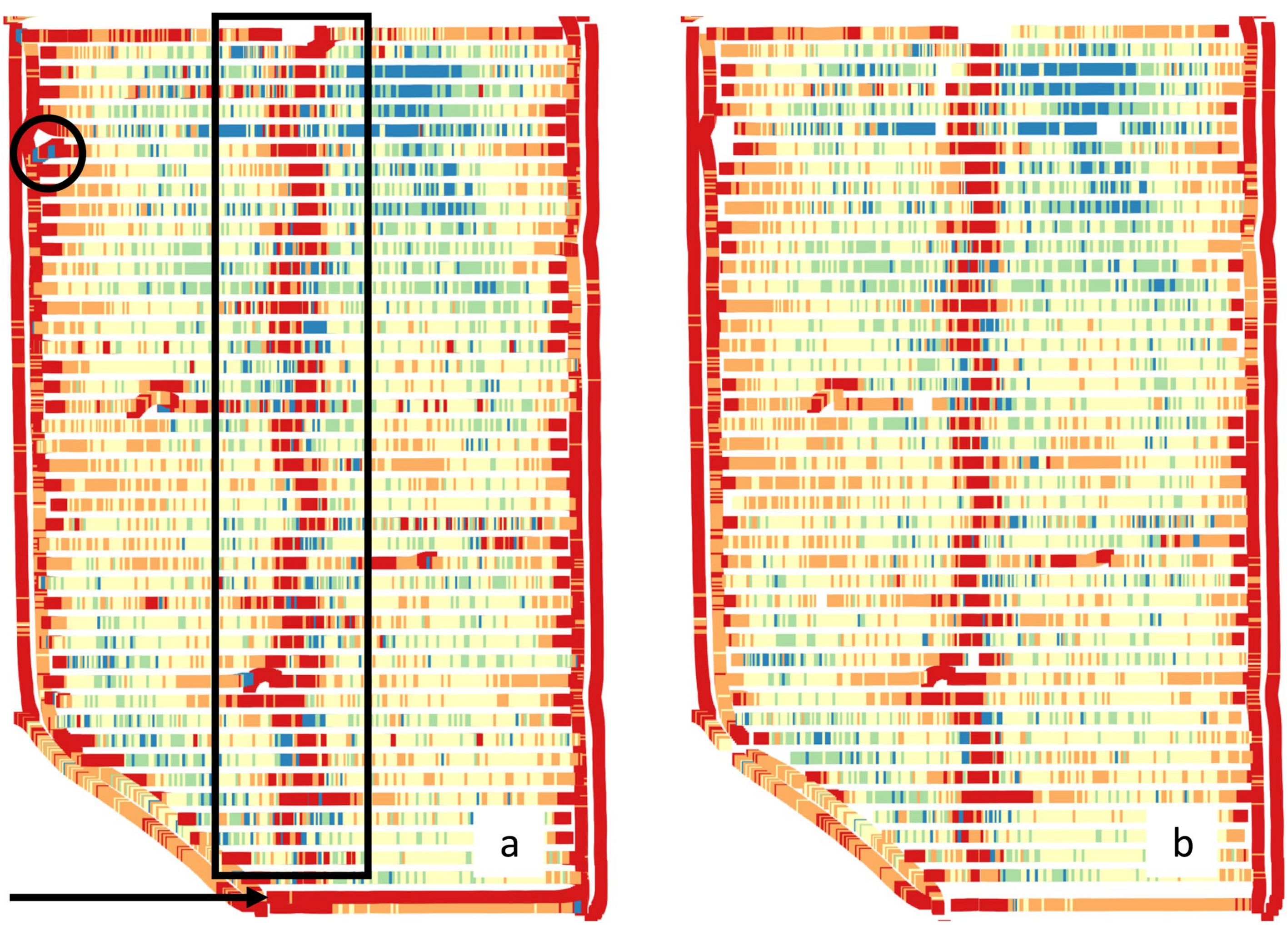

Figure 3. Example a) raw and b) post-processed yield map. Red values are lower yields, and blue values are higher yields. Errors in the raw map have been highlighted. The rectangle highlights an example of flow delay, the circle an increase in yield due to an abrupt change in travel speed, and the arrow a swath width error.

Header Position

If the header position sensor is not correctly calibrated or the combine is operated with the header down but not harvesting, very low yields will be estimated due to no crop flow. This can lower the overall yield estimate of the field. This may be seen at the field edges when turning or when deadheading passes across the field.

Other Filters

Other filters can be applied to remove more errors or questionable data. For example, filters can be applied to remove very high or low yield values, which are likely to be in error if the monitor was properly calibrated; however, care must be taken if the yield monitor was not calibrated.

Values at the start and end of each pass as grain flow rates rise and fall are likely to be erroneous and should be removed. Many software packages will automatically remove two to four points. In addition, most software will have the option to remove individual points at the user's discretion. Pay special attention to errors found outside of common locations, such as around the edge of the field or known obstacles, as these can have a larger influence on comparisons of yield between years.

Summary

Ensuring high-quality yield data requires both calibrating your combine and post-processing the data after harvest, as well as regular combine maintenance. Remember that calibration should take place multiple times during the harvest season when crop conditions (e.g., grain moisture, type, or variety) change. While calibration does take time, pre-harvest planning can speed up the process and ensure that all the steps in the process are addressed. Post-processing can be achieved using available software but requires attention to detail by the user. These practices will help ensure your yield maps are accurate and can be reliably used for on-farm decision making.

References

Kharel, T. P., Swink, S. N., Maresma, A., Youngerman, C., Kharel, D., Czymmek, K. J., & Ketterings, Q. M. (2019). Yield monitor data cleaning is essential for accurate corn grain and silage yield determination. *Agronomy Journal, 111*(2), 509–516.

McFadden, J., Njuki, E., & Griffin, T. (2023). *Precision agriculture in the digital era: Recent adoption on U.S. farms, EIB-248*. U.S. Department of Agriculture, Economic Research Service.

Schimmelpfennig, D., & Ebel, R. (2016). Sequential adoption and cost savings from precision agriculture. *Journal of Agricultural and Resource Economics, 41*(1), 97–115.

Sudduth, K. A., & Drummond, S. T. (2007). Yield Editor: Software for removing errors from crop yield maps. *Agronomy Journal, 99*(6), 1471–1482.

Additional Resources

Luck, J. D., & Fulton, J. P. (2014). Precision agriculture: Best management practices for collecting accurate yield data and avoiding errors during harvest, EC2004. University of Nebraska Lincoln Extension.

Nielsen, R. L. (2020, October). *Identify and eliminate "gremlins" from yield monitor data*. Corny News Network, Purdue University Extension.

Iowa State University Digital Ag Innovation Lab. (2025, March 14). *Monitor guide*. Research to Rows.

Publication 4133 (POD-09-25)

By **Jessica Drewry**, PhD, Assistant Professor, Agricultural and Biological Engineering; **John Lowe**, PhD, Assistant Professor, Agricultural and Biological Engineering; **Mike Mulvaney**, PhD, Associate Professor and Endowed Chair, Plant and Soil Sciences; and **Corey Bryant**, PhD, Assistant Professor, Delta Research and Extension Center.

Produced by Agricultural Communications.

Extension Service of Mississippi State University, cooperating with U.S. Department of Agriculture. Published in furtherance of Acts of Congress, May 8 and June 30, 1914. ANGUS L. CATCHOT JR., Director

MISSISSIPPI STATE UNIVERSITY EXTENSION