How and Why to Clean Corn Yield Monitor Data

Introduction

Corn grain and silage growers can evaluate management practices (variety selection, drainage, fertilizer, etc.) within fields and over time with data from yield monitors. Management decisions using yield data are less reliable when there are data errors. The process of correcting errors in raw yield data is called "data cleaning". Clean data form reliable maps with more accurate yield estimates (Figure 1).
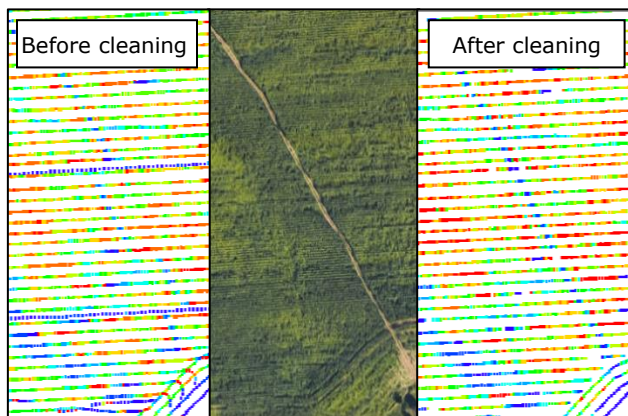

Figure 1: Yield data cleaned using Yield Editor compared to a Google Maps satellite image and raw yield monitor data.

Yield Editor is a freely available software application for yield data cleaning. A step-by-step data cleaning protocol was created to guide users through the process of fast and accurate data cleaning. This factsheet makes the case for data cleaning by addressing common misconceptions and outlines the data cleaning process.

Common Misconceptions

"My scale tickets match the yield monitor data, so data cleaning is not needed".
- Yield estimates from uncleaned data may be too high in some field sections and too low in others, causing the whole-field yield estimate to "balance out". Accurate whole-field yield estimates do not guarantee an accurate map.
- Within-field yield evaluation and practices like zone management (see Agronomy Fact Sheet #108) require a reliable yield estimate for every section of the field.

"Cleaning yield data is too hard/takes too long".
- Data cleaning does not need to be done one field at a time for large farms. Yield Editor users can manually clean data from 10 sample fields and then apply the average cleaning settings farm-wide. This is called "Batch Cleaning", and is both fast and accurate.
- Selecting sample fields for manual cleaning with known features (within-field tree lines, grass waterways, etc.) makes it easier to find the best cleaning settings (see Figure 1). The better you know your fields, the faster cleaning can be done.
- Average settings must be determined each harvest season as yield cleaning settings change with equipment, field conditions, and operator practices. With a bit of practice, data cleaning typically takes one to two hours per farm. No experience is necessary, but some computer skills are a big help.
- Properly calibrated equipment and good operator practices are essential for accurate and more easily cleaned yield data (Agronomy Fact Sheets #104 and 105).

Manual Cleaning with Yield Editor

There are only four cleaning settings that must be adjusted each harvest season. Yield Editor's other filters (e.g. removing overlapped passes) do not have to be adjusted in most cases.

Flow Delay → Period of time between the crop being cut and the grain/forage being massed.
- Many yield monitor systems are set to sense and record crop flow (biomass) every second during harvest.
- When biomass hits the flow sensor, the yield monitor records the pounds of crop and tags that information to the harvester's GPS location. This makes up one data point.
- If the corn takes 5 seconds to pass from the harvester's head to the flow sensor, the "flow delay" value is 5. Yield Editor shifts all flow data back 5 GPS points to create a more accurate yield map.
- Yield map trends are clearer with the correct flow delay applied (Figure 1).

Moisture Delay → Period of time between harvest and crop reaching the moisture sensor.
- Delay between harvest and the crop hitting the moisture sensor can be adjusted with a "Moisture Delay".
- Yield Editor's "Moisture" map will have clearer patterns with the correct moisture delay applied.

Start and End Pass Delay → Period of time when the harvester is slowing down or ramping up to speed.
- Yield estimates are often much higher or lower than the actual yield when the harvester is speeding up or slowing down.
- "Start Pass" and "End Pass" delays allow users to remove a set number of unreliable datapoints from the beginning and end of each harvester pass (Figure 2).
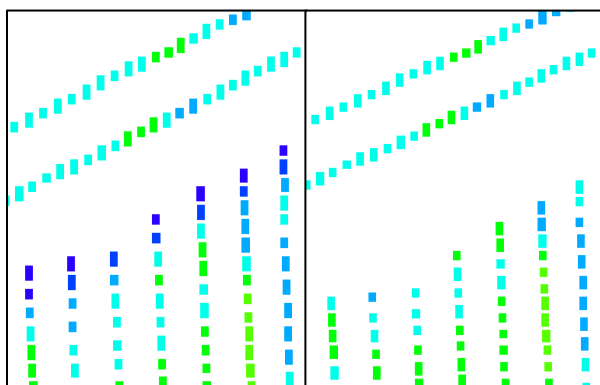

Figure 2: On the left, no start pass delay has been applied. On the right, start pass delay was set to 3.

- Applying the correct Start and End Pass Delays produces a more accurate yield estimate near field edges.
- Just like Flow and Moisture delays, Start and End Pass delays change annually with field conditions, equipment, and operator practices.

Batch Cleaning with Yield Editor
Farm-wide data cleaning can be completed quickly using the average delay values from the 10 manually cleaned fields. The process is called "Batch cleaning" and takes just a few clicks (Figure 3). After cleaning, yield data can be imported into AgLeader SMS or similar crop management software.

Batch cleaning saves a complete copy of the original uncleaned data. Compare the raw data with clean data to see how and where changes have occurred. Raw and cleaned data should be saved along with other essential farm records.
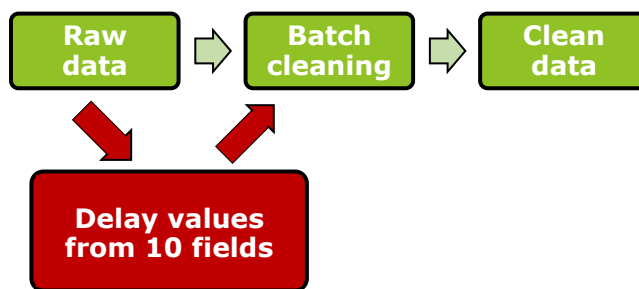

Figure 3: A whole year of yield data (all fields) can be cleaned in one "batch" submission, once the four main delay settings are determined.

General Summary
Reliable yield data are essential to make field and within-field management decisions and evaluate yield over time. Data cleaning can be done in one to two hours per farm per year by selecting ten fields with known within-field features, determining delay values for flow, moisture, start and end passes, and batch cleaning to correct errors in all harvested fields.

Additional Resources
- Nutrient Management Spear Program Fact Sheet Series: nmsp.cals.cornell.edu/guidelines/factsheets.html.
- Best Management Practices for Collecting Accurate Yield Data and Avoiding Errors during Harvest. 2004. University of Nebraska Lincoln. Bulletin EC2004. extensionpublications.unl.edu/assets/pdf/ec2004.pdf.
- Processing/Cleaning Corn Silage and Grain Yield Monitor Data for Standardized Yield Maps across Farms, Fields, and Years. 2018. Cornell University. Nutrient Management Spear Program (NMSP). nmsp.cals.cornell.edu/publications/extension/ProtocolYieldMonitorDataProcessing2_8_2018.pdf.
- Yield Editor software, Version 2.0.7. 2019. USDA-ARS. www.ars.usda.gov/research/software/download/?softwareid=370.

Disclaimer
This fact sheet reflects the current (and past) authors' best effort to interpret a complex body of scientific research, and to translate this into practical management options. Following the guidance provided in this fact sheet does not assure compliance with any applicable law, rule, regulation or standard, or the achievement of discharge levels from agricultural land.