2950 Niles Road, St. Joseph, MI 49085-9659, USA
269.429.0300 fax 269.429.3852 hq@asabe.org www.asabe.org

*An ASABE Meeting Presentation*

*Paper Number: 1110690*

# Using Yield Monitors to Assess On-Farm Test Plots

**Randal Taylor, Professor and Extension Ag Engineer**

Biosystems and Agricultural Engineering, Oklahoma State University.

**John Fulton, Associate Professor and Extension Specialist**

**Daniel Mullenix, Graduate Student**

Biosystems Engineering, Auburn University.

**Matthew Darr, Assistant Professor**

**Robert McNaull, Graduate Student**

Agricultural and Biosystems Engineering, Iowa State University.

**Lucas Haag, Graduate Student**

**Scott Staggenborg, Professor**

Agronomy, Kansas State University.

**Written for presentation at the**
**2011 ASABE Annual International Meeting**
**Sponsored by ASABE**
**Gault House**
**Louisville, Kentucky**
**August 7 – 10, 2011**

***Abstract.*** *Farmer test plots have become a staple for production agriculture. These plots can range from simple side-by-side demonstration plots to a replicated research study. The rush of harvest often creates a challenge for harvesting these plots. Yield monitor data were collected from field scale plots in multiple states to assess ability to measure on-farm research. Grain mass was also measured for each plot with a weigh wagon or certified scale. The variability of yield monitor error (standard deviation) was not correlated with the magnitude of the error (mean). Thus calibration in and of itself will likely not result in more consistent yield monitor error. Determining if treatments or observations from non-replicated studies are different will be challenging. Depending on the chosen probability level, this data indicate that distinguishing a 3 to 9 percent difference was possible. Statistical analysis of replicated trials results in similar conclusions with reference and yield monitor data. Mass flow rate is one factor impacting yield monitor error.*

**Keywords.** Weigh wagon, replication, on-farm research.

## Introduction

Grain yield monitors have been commercially available in the United States since the early 1990s and are becoming more commonly accepted by farmers. One of the reasons many farmers are buying yield monitors is to conduct what they consider to be research. This may be comparing varieties, fertilizer rates, herbicide treatments, or other cultural practices.

Krill (1997) compared weights from a yield monitor and weigh wagon on variety strip plots of popcorn, corn, and soybean in 1994-95. Plots were 15 feet wide and about 1000 feet long. Errors between yield monitor and weigh wagon results were typically less than five percent for corn and popcorn. Error for the soybean strips were typically less than 10 percent, but ranged as high as 20 percent. He found that the monitor seemed more accurate on strips of the same variety as opposed to strips of multiple varieties. The source of errors was not discussed, but the author indicated that at least 25 loads were used for calibration.

Horstmeier (1998) reported on 324 comparisons between yield monitors and weigh wagons at 14 sites. Plots were 15 feet wide and about 300 feet long. The highest average error at a site was over 18 percent. The lowest error was 0.3 percent. Correlation coefficients for weigh wagon and yield monitor data ranged from 0.55 to 0.99. The source of errors was attributed to poor calibration.

Grisso et al. (2002) studied the operation characteristics of yield monitors operated at the combine’s normal capacity and capacities that were above and below normal. They found errors between weigh wagon and yield monitor weights exceeding 10 percent at flow rates other than normal.

Al-Mahasneh and Colvin (2000) used an on-board scale to assess yield monitor performance. They found that correlation between the reference scale and yield monitor increased with longer harvest lengths in both oats and corn. Longer plots will directly result in more mass per plot.

The objective of this study was to evaluate the potential for using yield monitors to conduct on-farm research. Specifically,

1. Determine the differences that yield monitors can detect
2. Determine the source of errors
3. Make recommendations for optimal use of yield monitors for on-farm research

## Methods

Yield monitor data were compared to known reference weights for 29 different corn trials. Data were obtained from trials in Kansas, Alabama, and Iowa. Some data were collected from replicated on farm research trials and some were non-replicated hybrid trials. Data included yield monitor plot weight, reference plot weight, and moisture content for all trials. Reference weights were either determined from weigh wagons or commercial scales. Some trials also had the raw yield monitor data files.

For each test plot data set the error associated with the yield monitor estimate of yield was calculated by recording the weight difference between the yield monitor and the grain cart scale and dividing by the grain cart scale weight. It was assumed that the grain cart scale weight was calibrated before use. The following equation results in negative values when the yield monitor underestimates the actual weight.

$$Yield\ Monitor\ Error\ (\%) = \frac{Yield\ Monitor\ Weight - Reference\ Weight}{Reference\ Weight} * 100$$

A mean yield monitor error and standard deviation were determined for each trial. The mean reference yield and standard deviation were also determined for each trial.

To determine the ability of a yield monitor to distinguish statistically different distributions, the Chi-squared theorem was applied to these unique combine error distributions. The null hypothesis assumed that two unique yield measurements from adjacent plots are from the same distribution. The standard deviation and mean yield level for each trial was used to determine the threshold levels for minimum yield differences in individual plots that were discernable by the combine yield monitor. This procedure was also used to calculate the expected thresholds for 68 and 90 percent confidence intervals for two expected standard deviations.

Additional analysis was conducted in trials where the raw yield monitor data were available. Once the data were trimmed to the plot boundaries, the yield monitor data was used to determine the time required to harvest each individual plot. The reference mass was divided by this time to get an estimation of the average mass flow for each plot.

Each trial that had some experimental design and replications was treated as a case study to assess yield monitor relative to collecting data for research projects. The statistical analysis depends on the experimental design for each trial, but was conducted with reference and yield monitor yields. The results of these analyses were compared.

## Results

Average yield monitor error of all trials was -1.1 percent with an average error standard deviation of 3.6 percent. Since each yield monitor had a unique calibration and was operating in a unique set of conditions, average error provides a high level perspective on yield monitor accuracy but is not useful in assessing the impact of determining performance differences in test plot trials.

The error distribution of each individual combine provides additional insight into impact factors for using yield monitors to predict test plot winners. By plotting a distribution of error for each unique combine several key points were identified (Figure 1). Each combine had a unique calibration level as expected. The accuracy of the combine calibration is determined by the mean of each unique combine dataset. The combines analyzed as part of this research yielded yield monitor calibration accuracies between -27 to 15 percent, but generally fell between +/- 10 percent.

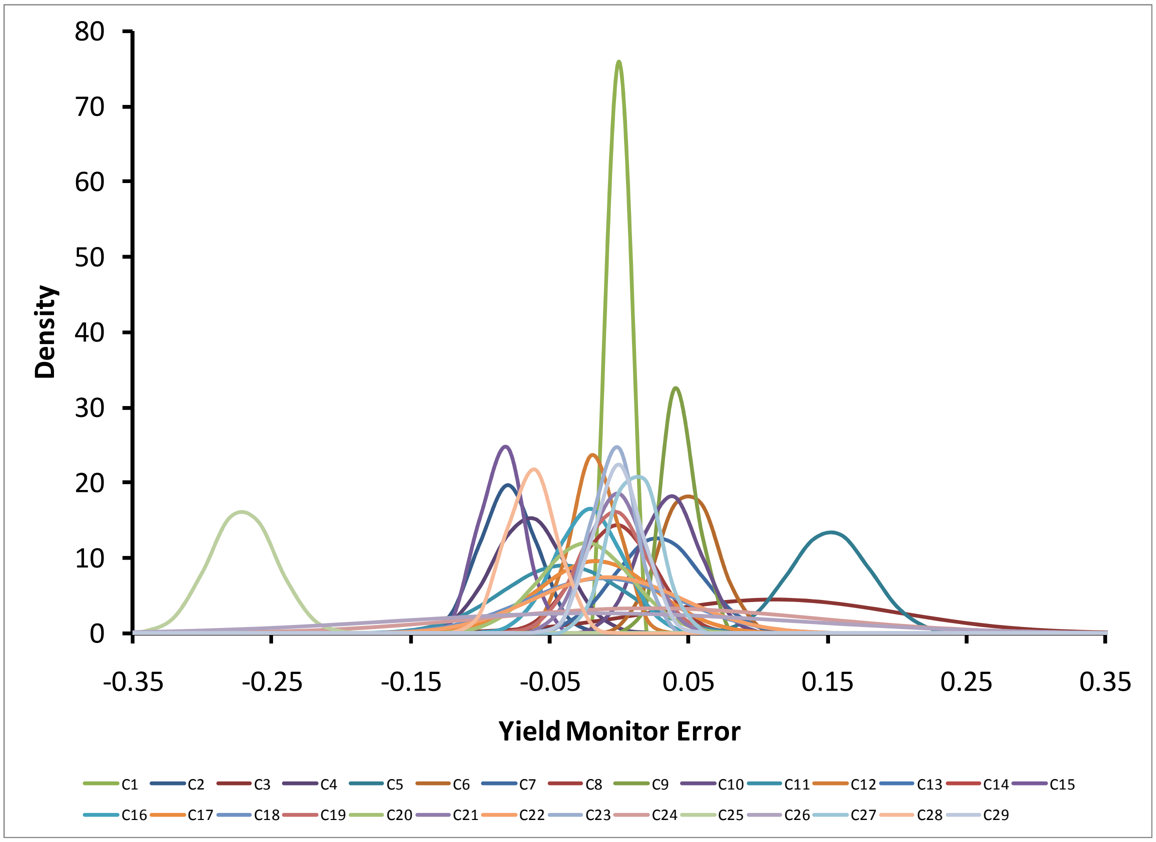

Figure 1. Distribution of yield monitor error for the 29 on-farm trials.

In general the standard deviations of each combine error distribution were similar with a few outliers. The range of standard deviations was from 0.5% to 14.7%.with a mean of 3.6%. The summary statistics for individual trials are shown in table 1 along with overall means and ranges.

Table 1. Summary statistics for individual trials.

| Trial | Mean Error | Error SD | N | Mean Mass | Mean Yield | Yield SD | Yield CV |
|---|---|---|---|---|---|---|---|
| C1 | -0.2% | 0.5% | 6 | 23000 | 141.0 | 10.8 | 7.6% |
| C2 | -8.0% | 2.0% | 24 | 4303 | 185.9 | 19.3 | 10.4% |
| C3 | 11.2% | 9.0% | 15 | 1506 | 100.1 | 26.3 | 26.3% |
| C4 | -6.5% | 2.6% | 10 | 8500 | 174.9 | 21.0 | 12.0% |
| C5 | 15.2% | 2.9% | 12 | 5469 | 160.1 | 18.3 | 11.4% |
| C6 | 5.0% | 2.1% | 13 |  | 206.4 | 8.8 | 4.3% |
| C7 | 2.8% | 3.1% | 10 |  | 194.2 | 10.4 | 5.3% |
| C8 | -0.2% | 2.8% | 18 |  | 217.2 | 10.7 | 4.9% |
| C9 | 4.4% | 1.2% | 5 |  | 243.3 | 15.8 | 6.5% |
| C10 | 3.7% | 2.2% | 12 |  | 232.6 | 16.5 | 7.1% |
| C11 | -4.0% | 4.5% | 4 |  | 234.6 | 13.4 | 5.7% |
| C12 | -1.7% | 1.7% | 8 |  | 224.9 | 16.2 | 7.2% |
| C13 | -2.1% | 2.4% | 10 |  | 188.8 | 12.6 | 6.7% |
| C14 | -1.6% | 4.2% | 11 |  | 192.2 | 9.9 | 5.2% |
| C15 | -8.4% | 1.6% | 10 |  | 202.1 | 9.0 | 4.5% |
| C16 | -2.1% | 2.4% | 10 |  | 188.8 | 12.6 | 6.7% |
| C17 | -1.6% | 4.2% | 11 |  | 192.2 | 9.9 | 5.2% |
| C18 | -1.1% | 5.5% | 8 |  | 224.7 | 18.4 | 8.2% |
| C19 | -0.4% | 2.4% | 6 |  | 198.8 | 7.1 | 3.6% |
| C20 | -2.4% | 3.3% | 11 |  | 169.7 | 12.6 | 7.4% |
| C21 | -0.2% | 2.1% | 25 |  | 201.8 | 11.2 | 5.6% |
| C22 | -0.8% | 5.4% | 11 |  | 192.2 | 21.4 | 11.1% |
| C23 | -0.4% | 1.6% | 39 | 4612 | 153.3 | 14.6 | 9.5% |
| C24 | 2.1% | 11.9% | 24 | 662 | 139.2 | 9.8 | 7.0% |
| C25 | -27.1% | 2.4% | 24 | 616 | 129.5 | 15.9 | 12.2% |
| C26 | -1.8% | 14.7% | 24 | 592 | 124.5 | 10.0 | 8.0% |
| C27 | 1.1% | 1.7% | 4 | 2523 | 118.8 | 2.6 | 5.0% |
| C28 | -6.3% | 1.8% | 6 | 413 | 117.2 | 7.2 | 5.0% |
| C29 | -0.1% | 1.8% | 14 | 10786 | 67.2 | 19.7 | 5.0% |
|  |  |  |  |  |  |  |  |
| Mean | -1.1% | 3.6% | 13 | 5248 | 176.4 | 13.5 | 7.7% |
| Min | -27.1% | 0.5% | 4 | 413 | 67.2 | 2.6 | 3.6% |
| Max | 15.2% | 14.7% | 39 | 23000 | 243.3 | 26.3 | 26.3% |

### ***Single Factor Comparisons***

When considering the use of yield monitors for quantifying non-replicated hybrid performance the standard deviation of the yield measurement is of greater importance than the absolute accuracy. For example, if a combine has a calibration error of 2%, then all yield readings would be biased by 2%, but the ranking of the hybrid performance would still be valid. On the other hand, a 3% standard deviation indicates that the actual yield may have a range of ± 6% at the

95% confidence level. Figure 2 shows the expected distinguishable differences as a function of yield for two probability levels (68 and 90 percent) and two standard deviations (3.6 and 2.7 percent). The 3.6 percent standard deviation error was the mean of the 29 trials whereas the 2.7 percent standard deviation error was the mean with a few extreme values removed. While there are two ways to be able to distinguish greater differences, reducing the probability level or assuming a lower standard deviation, based on the data presented here Figure 2 is a realistic representation.

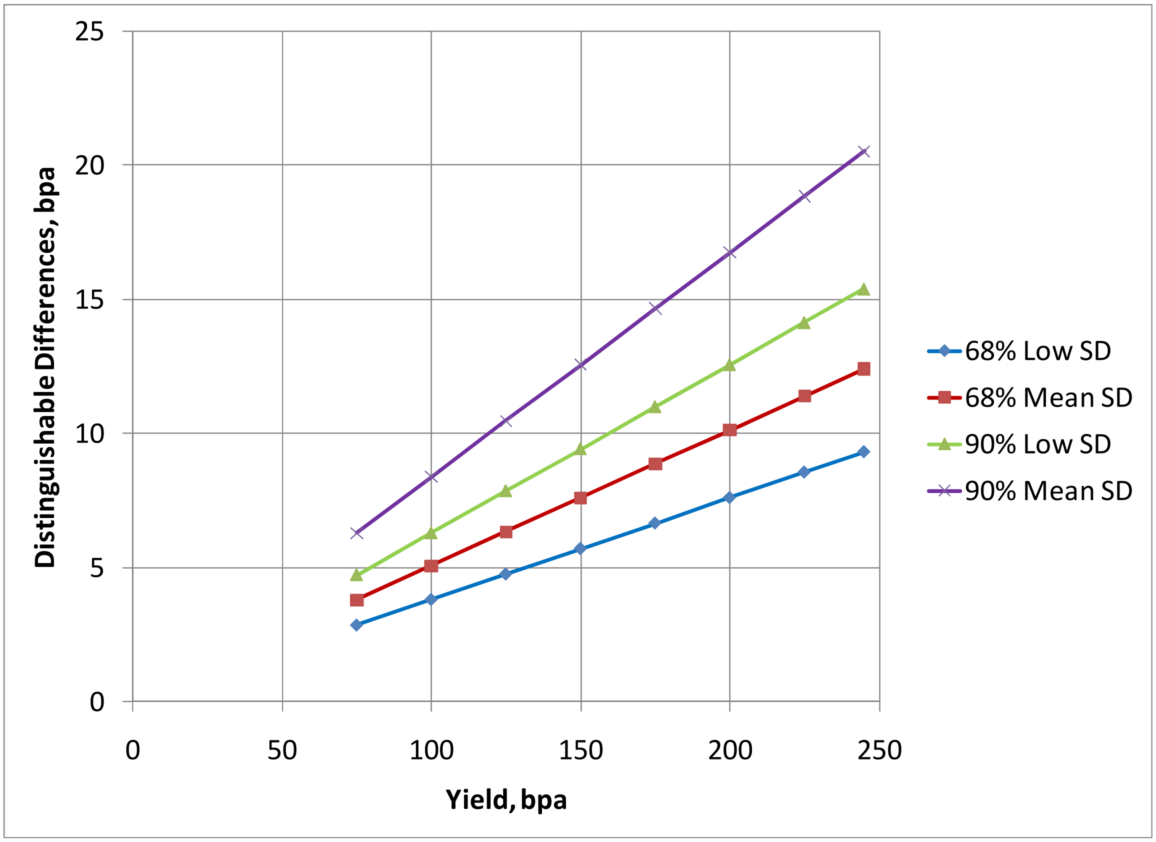

Figure 2. Minimum distinguishable yield monitor yield differences in test plots.

## *Replicated Trials*

Some of the trials presented in this study were replicated on farm research. These trials were analyzed using both the reference and yield monitor yields to determine the validity of using a yield monitor for replicated on farm research. Three of these projects are presented here.

Project C2 was conducted on a field in central Kansas to determine the effect of population and row spacing on corn yield. It consisted of five populations (32000, 42000, 52000, 62000, and 72000 seeds per acre) and three row spacings (7.5, 15, 30 inch). Plots were either planted with a row crop planter or air seeder. There were 22 plots used in the regression with each measuring 30 feet wide and 600 feet long. Plots were randomized, but this was not a factorial experiment. There were 13 treatments of which five were replicated at least twice. The grain from each plot was weighed in a weigh wagon at the edge of the field. The measured weight of grain harvested by the yield monitor was recorded after each plot was harvested. Yield, determined from the reference scale and yield monitor, was regressed as a linear function of seeding rate, row spacing, and seeder. Seeder was treated as an intercept shifter in the regression with drill as the default value (0). The regression model for the reference yield had an r-squared value of 0.83 and standard error of 10.0 bu ac$^{-1}$, while these values for the yield monitor model were 0.82 and 8.8 bu ac$^{-1}$. Regression results are shown in Table 2. Not only are the coefficients similar, but the same parameters are significant in both models. Both the reference and yield monitor yields lead to the same conclusion. It should be noted that the mean error on this trial was -8.0 percent with a standard deviation of 2.0 percent.

Table 2. Regression results for trial C2 using the reference scale and yield monitor yields.

| | Reference Scale | | Yield Monitor | |
|---|---|---|---|---|
| | Coefficients | P-value | Coefficients | P-value |
| Intercept | 83.17 | 3.82E-06 | 82.80 | 7.67E-07 |
| Seeding Rate | 1.28 | 2.33E-05 | 1.11 | 2.99E-05 |
| Spacing | 0.59 | 2.24E-01 | 0.34 | 4.19E-01 |
| Seeder | 14.57 | 3.45E-02 | 12.30 | 4.22E-02 |

Project C5 was conducted on a field in eastern Kansas to determine the effect nitrogen rate on corn yield. It consisted of four nitrogen rates (60, 90, 120, and 150 lbs N per acre) and was replicated three times There were 12 plots with each measuring 30 feet wide and 1100 feet long. The grain from each plot was weighed in a weigh wagon at the edge of the field. The measured weight of grain harvested by the yield monitor was recorded after each plot was harvested. A single factor analysis of variance was conducted on yield, determined from the reference scale and yield monitor, with nitrogen rate as the factor. The least significant differences for reference and yield monitor yields are shown in table 3 for three probability levels. The analysis shows the model is significant ($p<0.1$) for the yield monitor data, but not the reference yield measurement. The treatment means are shown in Table 4 and clearly show a similar trend. The yield difference between the 120 and 150 lbs N ac$^{-1}$ rates is 27.1 and 33.7 bu ac$^{-1}$ for the reference and yield monitor data respectively. From a practical view the analytical results are similar. The mean error on this trial was 15.2 percent with a standard deviation of 2.9 percent. Better yield monitor calibration would likely result in similar statistical significance.

Table 3. Analysis of variance results for trial C5 using reference and yield monitor yields.

| | Reference Scale | Yield Monitor |
|---|---|---|
| ANOVA P>F | 0.225 | 0.087 |
| LSD 0.10 | 24.3 | 25.6 |
| LSD 0.05 | 30.7 | 32.2 |
| LSD 0.01 | 46.4 | 48.8 |

Table 4. Treatment means for trial C5 using reference and yield monitor yields.

| | Reference Scale | Yield Monitor |
|---|---|---|
| 60 | 146.9 | 163.9 |
| 90 | 154.5 | 177.7 |
| 120 | 156.0 | 181.7 |
| 150 | 183.1 | 215.4 |

Project C29 was conducted on a field in southwest Kansas to determine the effect planting geometry on corn yield. It consisted of three treatments (conventional 30" rows, clumps in 30" rows, and drilled) and was replicated four times. There were 12 plots however the drilled treatment has some emergence issues and one rep was completely lost, thus only 3 reps of all 3 treatments were analyzed. The least significant differences for reference and yield monitor yields are shown in table 5 for three probability levels. The LSDs for the yield monitor data are greater at all probability levels, but still separate treatment means the same as the reference yield. The mean error on this trial was -0.1 percent with a standard deviation of 1.8 percent.

Table 5. Analysis of variance results for trial C29 using reference and yield monitor yields.

| | Reference Scale | Yield Monitor |
|---|---|---|
| ANOVA P>F | 0.0005 | 0.001 |
| LSD 0.10 | 7.1 | 9.1 |
| LSD 0.05 | 9.2 | 11.9 |
| LSD 0.01 | 15.3 | 19.7 |

## *Items Impacting Errors*

Table 6 contains the correlation coefficients for the measured parameters. These values were explored in an effort explain the variability in errors as described by the standard deviation. The greatest correlation with error standard deviation was the negative relationship with mean mass harvested from plots. This indicates that error decreases when mass harvested increases which is certainly intuitive. This relationship is shown in Figure 3 with an exponential regression. The regression is driven by a few points while most of the yield monitor error has a standard deviation of less than 3 percent.

Table 6. Correlation between variables.

| | Mean Error | Error SD | N | Mean Mass | Mean Yield | Yield SD | Yield CV |
|---|---|---|---|---|---|---|---|
| Mean Error | 1.000 | | | | | | |
| Error SD | 0.167 | 1.000 | | | | | |
| N | -0.182 | 0.322 | 1.000 | | | | |
| Mean Mass | 0.126 | -0.467 | -0.347 | 1.000 | | | |
| Mean Yield | 0.048 | -0.260 | -0.305 | 0.053 | 1.000 | | |
| Yield SD | 0.115 | 0.080 | 0.184 | 0.083 | -0.048 | 1.000 | |
| Yield CV | 0.147 | 0.320 | 0.274 | -0.176 | -0.341 | 0.746 | 1.000 |

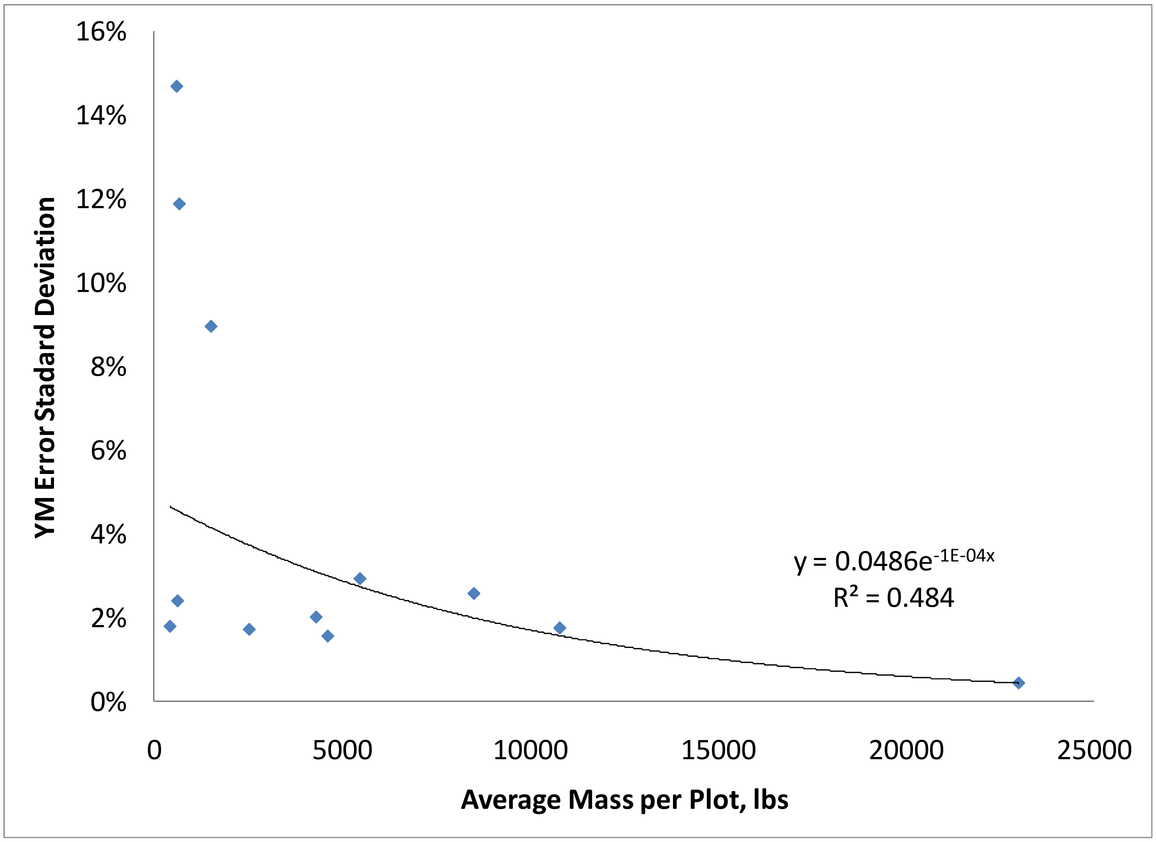

Figure 3. Error standard deviation as a function of mean mass harvested per plot.

Further examination of errors was conducted for the trials where raw yield monitor data were available. Given the non-linearity of mass flow sensors, it was expected that variation in mass flow rate could impact errors. Figure 4 shows yield monitor error as a function of mass flow for five trials. Mass flow was calculated from the reference mass and the time required to harvest the plot. Yield monitor error was affected to different degree in these five examples, but variation in mass flow across plots certainly impacts errors.

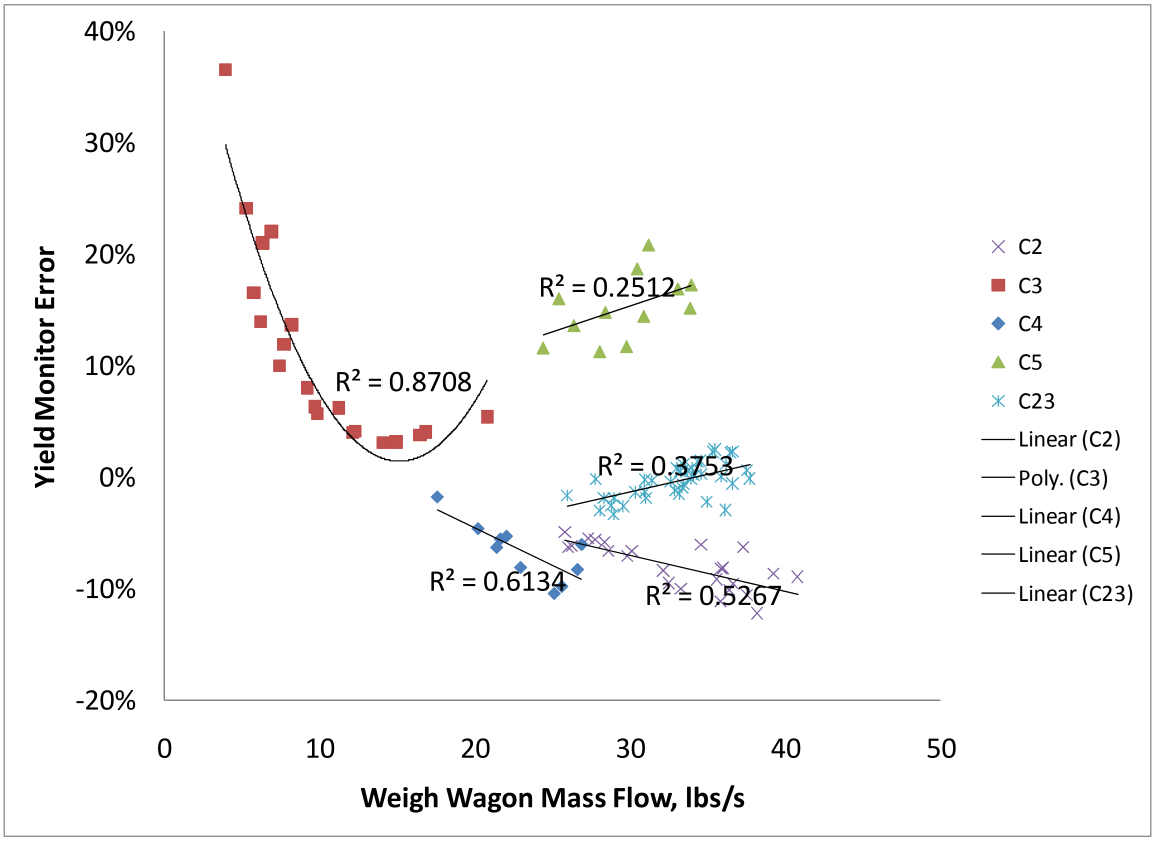

Figure 4. Yield monitor error as a function of the mean mass flow based on the reference mass.

Although crop factors, such as test weight, have been shown to influence yield monitor accuracy in past studies no significant correlation between yield monitor errors and test weight were identified during this work. The magnitude of previously reported errors from test weight are less in magnitude than the natural variation measured in this test data essentially pointing to test weight errors being incorporated within the noise of the overall sensor error.

## Conclusions

Based on the data presented here the following conclusions can be reached.

- The variability of yield monitor error (standard deviation) is not correlated with the magnitude of the error (mean). Thus calibration in and of itself will likely not result in more consistent yield monitor error.
- Determining if treatments or observations from non-replicated studies are different will be challenging. Depending on the chosen probability level, this data indicates that distinguishing a 3 to 9 percent difference is possible.
- Statistical analysis of replicated trials results in similar conclusions with reference and yield monitor data.
- Mass flow rate is one factor impacting yield monitor error.

Based on these results, a review of literature and field experiences of the investigators, several methods are suggested to improve yield monitor consistency and thus improve the distinguishable yield in test plot scenarios. These recommendations include:

- Calibrate the combine before beginning the test plot.
- Operate the combine at a consistent mass flow relative to the calibration.
- Use a large of a plot as feasible. Larger plots result in more mass harvested and more consistent results.
- Use replicated plots to improve the ability to distinguish treatment differences.
- Conduct rolling starts which means make sure the combine is moving as it enters the crop rather than entering the crop from a stationary position.
- Avoid significant variations in moisture and understand that moisture changes of 5% or more will require re-calibration. A 1% error in moisture content will result in a 2.5 bu/ac yield error.
- Avoid highly sloped or rough terrain. Combine slope will impact yield monitor accuracy.
- Maintain an accurate and constant header width when harvesting with a platform head in an effort to keep material other than grain consistent.

## References

Al-Mahasneh, M. A. and T. S. Colvin. 2000. Verification of yield monitor performance for on-the-go measurement of yield with on in-board electronic scale. *Trans. ASAE* 43(4): 801-807.

Horstmeier, G.D. 1998. A test of test plots. *Farm Journal, February, 1998.*

Grisso, R.D., P.J. Jasa, M.A. Schroeder, and J.C. Wilcox. 2002. Yield monitor accuracy: Successful farming magazine case study. *Applied Eng. in Agric.* 18(2):147–151.

Krill, T.L. 1995. Effectiveness of AgLeader® yield monitor for evaluation of varietal strip data. p. 819-825 *In* P.C. Robert et al., (ed.) Proceedings of the 3rd International Conference on Precision Agriculture. ASA, CSSA, SSSA, Madison,WI.