

EVALUATING FIELD TRIAL DATA

This article has been reprinted from Southwest Farm Press Vol 25, Number 11, April 9, 1998.

Field Trials can provide helpful information to producers as they compare products and practices for their operations. However, field trials must be evaluated carefully to make sure results are scientifically sound, not misleading and indicate realistic expectations for on-farm performance.

This fact sheet is designed to give you the tools to help you determine whether data from a field trial is science fact or science fiction.

What are the best sources of field trial data?

Field trials are conducted by a broad range of individuals and institutions, including universities, ag input suppliers, chemical and seed companies and growers themselves. All are potentially good sources of information.

What are the common types of field trials?

Most field trials fall into one of two categories: side-by-side trials (often referred to as strip trials) or small-plot replicated trials. Side-by-side trials are the most common form of on-farm tests. As the name suggests, these trials involve testing practices or products against one another in plots arrayed across a field, often in strips the width of the harvesting equipment.

These strips should be replicated across the field or repeated at several locations to increase reliability. Small-plot replicated trials often are conducted by universities and companies at central locations because of the complexity of managing them and the special planting and harvesting equipment often required.

Replicated treatments increase the reliability of an experiment. They compare practices or products against one another multiple times under uniform growing conditions in several randomized small plots in the same field or location.

Small-plot replicated trials also may be conducted on farmers' fields where special conditions exist, for example, a weed infestation that does not occur on an experiment station.

Are side-by-side plots more valuable than small-plot replicated trials, or vice versa?

Both types of plots can provide good information. The key is to evaluate the reliability of the data. It is also important to consider the applicability of the trial to your farming operation.

When is plot data valid, and when isn't it?

There isn't a black-and-white answer to that questions. But there are good rules of thumb that can help guide you. Consider these three field trial scenarios:

Scenario 1:

A single on-farm side-by-side trial comparing 10 varieties. Each variety is planted in one strip the width of the harvesting equipment and is 250 to 300 feet long.

What you can learn:

This trial will allow you to get a general feel for each variety or hybrid in the test, including how it grows and develops during the season.

However, this trial, by itself, probably won't be able to reliably measure differences in yield. This is because variability within the field, even if it appears to be relatively uniform, may be large enough to cause yield variations that mask genetic difference among the varieties. Other varietal characteristics, such as maturity or micronaire in cotton, can also be masked by soil variation.

Scenario 2:

Yield data from side-by-side variety trials conducted on the same varieties on multiple farms in your region.

What you can learn:

When data from multiple side-by-side trials are considered together, reliability increases. In this case, the more trials comparing the same varieties, the better. As you go from three to five to 10 or more locations, the certainty goes up that yield differences represent genetic differences and not field variability. Be aware, however, that small differences between treatments (in this case varieties) may still be within the margin of random variability of the combined trial and may not indicate actual genetic differences. One treatment will almost always be numerically higher. Statistical analysis helps determine if differences are significant (consistent).

Scenario 3:

A university-style small-block replicated trial comparing the same 10 varieties.

What can you learn:

Data from such trials, if they are designed well and carried out precisely, generally are reliable. This is, the results generally determine the yield potential of crop varieties. However, it is still important to consider whether results are applicable to your farming operation and are consistent with other research.

How do I know whether differences in yield, for example, are real and not caused by field variability or sloppy research?

Scientists use statistical analysis to help determine whether differences are real or are the result of experimental error, such as field variation. The two most commonly used statistics are Least Significant Difference (LSD) and the Coefficient of Variation (CV), both of which can provide insight on the validity of trial data. If these values aren't provided with trial results, ask for them.

Least Significant Difference (LSD) is the minimum amount that two varieties must differ to be considered significantly different. Consider a trial where the LSD for yield is four bushels per acre. If one variety yields 45 bushels per acre and another yields 43 bushels per acre, the two are not statistically different in yield. The difference in their yields is due to normal field variation, not to their genetics. In this example, a variety that yields 45 bushels per acre is significantly better than those yielding less than 41 bushels per acre. In many research trials, LSDs are calculated at confidence level of 75 to 95 percent. For example, a confidence level of 95 percent means you can be 95 percent certain that yield differences greater than the LSD amount are due to genetics and not to plot variability.

Coefficient of Variation (CV) measures the relative amount of random experimental variability not accounted for in the design of a test. It is expressed as a percent of the overall average of the test.

For measuring yield differences, CV's of up to five percent are considered excellent; 5.1 to 10 percent are considered good; and 10.1 to 15 percent are fair.

A high CV means there must be larger differences among treatments to conclude that significant differences exist. The bottom line: When considering yield test data, be skeptical when the CV exceeds 15 percent.

Is a one-year test valid, or are several years of results necessary to know whether one product or practice is superior to another?

In an ideal world, having several years of tests to verify use of a practice or product is best. But where changes are rapid, such as with crop varieties, having university data from multiple years isn't always possible.

When multi-year university data aren't available, pay more careful attention to statistical measures like CV and LSD, and the number of locations and testing environments.

Multi-year data on yield and performance can also be requested from the developers of new products prior to university testing. In either case, be cautious about making major production changes and trying large acreages of a given variety based on one year's data.

How should I evaluate trial results that are markedly different from other research in my area?

When research results are at odds with the preponderance of scientific evidence, examine the new research with extra care.

Pay special attention to factors that might have influenced the outcome, such as soil type, planting date, soil moisture and other environmental conditions, and disease, insect and weed pressures. For example, was the growing season unusually wet or unusually dry? When was it dry or wet? What was the crop growth stage when it was wet or dry?

Was there a disease that affected one variety or hybrid more than another one? Were there insect problems? Could this have influenced the trial's outcome and its applicability to your operation? If you determine that unusual circumstances affected the outcome, be cautious about how you use the results.