Site-Specific Management Guidelines

T.A. Wittig and Z.W. Wicks III

SSMG-18

Simple On-Farm Comparisons

Summary

Many farmers continuously experiment with new farming approaches to optimize profitability. To determine if management changes make a difference, it is necessary to compare one treatment versus another. The selection of a method for conducting the experiment and making comparisons is critical to minimize incorrect interpretation of experimental results. It is not necessary to implement complicated experimental or statistical methods to determine if a change in management will improve production. However, if proper techniques are not used, results can be misleading, and a change in management might not lead to an increase in productivity. The purpose of this guideline is to provide a framework for conducting simple on-farm experiments. These techniques can be used for any comparisons where two factors need to be compared with statistical precision.

Typically, an on-farm experiment will be composed of four parts:

- Problem or question
- Experiment
- Information
- Analysis

Problem or Question

Research questions must be very specific in nature. General questions may be important, but are difficult to research. General questions such as “Which fertilizer regime is best for me?” or “Is there a more productive way of doing what I’m doing?” can be answered if the question is developed to be specific.

Examples of specific types of questions that can be answered by performing on-farm research include:

- Which of two varieties is higher yielding?
- Which of two tillage regimes is more productive?
- Which of two fertilizer regimes is best?

It is important that a simple, straightforward comparison of similar treatments be selected to investigate. The goal of comparing treatments should be to answer a specific question or solve a specific problem.

Experiment

The major issue in planning an on-farm experiment is to be sure you are comparing apples with apples. This means that when two factors are compared they should be compared under equal circumstances. Imagine what the results of an experiment comparing the yields of two wheat varieties would be if one variety was planted on an old feedlot while the other was planted on a stony hilltop. The one grown on the feedlot would have a yield advantage unrelated to the specific variety, and the information or data obtained would be virtually worthless. The point is that when experiments are planned, the two treatments being compared should be evaluated under similar conditions.

The second component is replication. All of us have seen a 10-year-old bowl a strike and a pro-bowler throw a gutter ball. It is only by watching each throw several times that we can tell who is the better bowler. The same holds true for experimentation. When planning to compare two treatments, we must make sure that each is independently represented a number of times. This can best be accomplished by the use of alternating strips based on equipment size. The first two points can be graphically represented on the lay of the land as shown in **Figures 1 and 2**.

The first experiment (**Figure 1**) would not be as good as the second experiment (**Figure 2**) because it has treatments being applied at different elevations while the second experiment has treatments applied across elevations. This simple example shows the importance planning, considering all factors that may influence the results of the experiment that can be controlled. When comparing treatments, it is critical to control as many variables as possible to reduce the amount of error in the data that are produced from the experiment. A well-planned experiment will consider all factors that may influence the

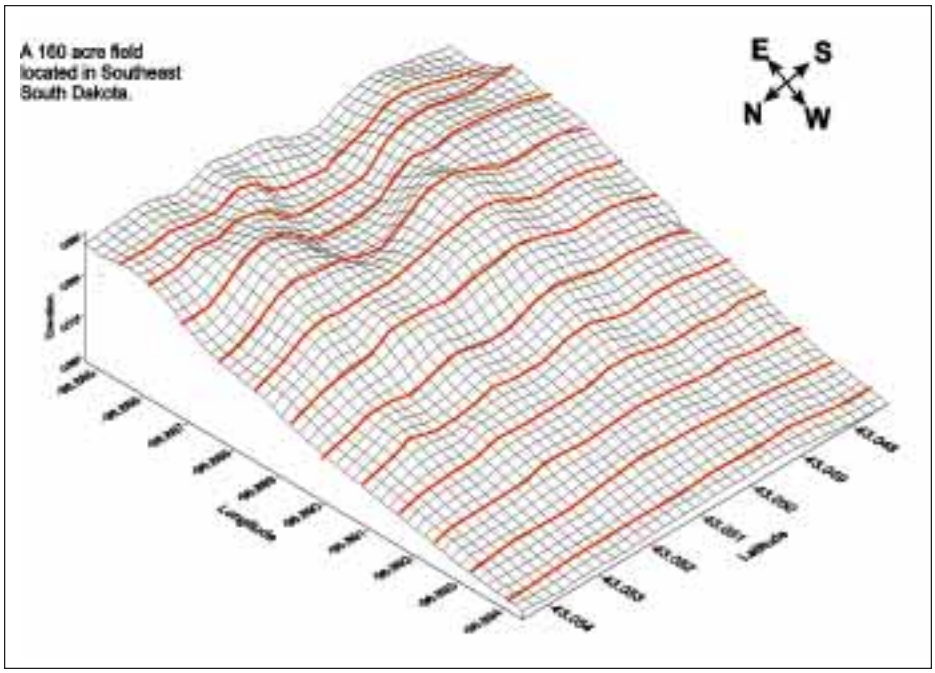

Figure 1. Experimental design in which treatment plots are laid out at different elevations.

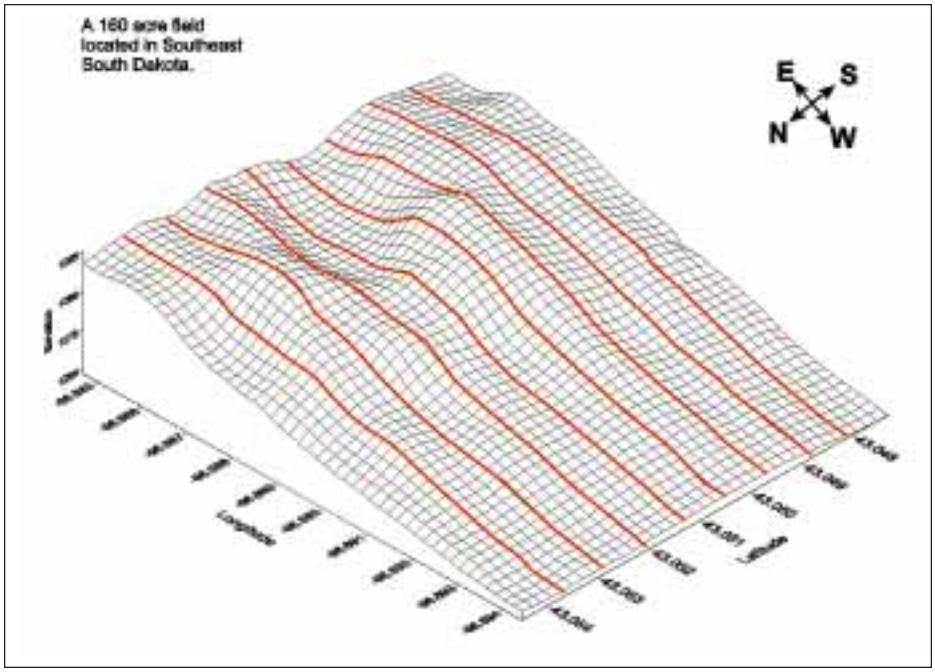

Figure 2. Experimental design in which treatment plots are laid out across elevations.

treatment being compared and attempt to control those factors as much as possible.

The simplest type of a "planned" experiment is referred to as a "paired t-test". This type of experiment would involve comparing two different "treatments" such as different varieties of crop seed, different fertilizers, or different rates of the same fertilizer. To perform such an experiment, it is recommended to have at least 10 to 15 pairs of plots. Each pair should be as comparable as possible with respect to any factor (other than the "treatment") that might be thought to influence the "response," (such as yield). It is not necessary that there be comparability between pairs of plots; only within the plot pair.

One plot in each pair receives one of the two treatments, the other plot the second treatment. The treatments should be randomly assigned to the plots in the pairs. This is most easily accomplished by tossing a coin. If the coin is "heads", use treatment A; if "tails", treatment B.

Information and Analysis

The next step is data collection. This means collecting equal amounts of information from each plot so the data can be equally compared. The information should be collected on a uniform plot length from comparable areas of the field. This means, for example, starting your measurements 50 to 60 ft. from the edge of the field and measuring performance over 100 to 200 ft. of field harvest. The size of the collected sample is determined by the size of the plot with the understanding that the more information from a uniform area, the better.

Once the responses have been measured (at the end of the experiment), the data could be summarized as follows:

	A	B	Difference
1	A_1	B_1	$D_1 = A_1 - B_1$
⋮	⋮	⋮	⋮
N	A_n	B_n	$D_n = A_n - B_n$

When comparing the results of the experiment if the differences (D's) are all (nearly) zero, there is no difference in the treatments. However if some (or all) of the D's are substantially different from zero, there is a difference in the treatments. If all other factors have been controlled to greatest extent possible, the difference between plots can be attributed to the treatments. Since the plots were treated as similarly as possible before the application of the tested treatment, the difference in response can only be attributed to the tested treatment. The number of plot pairs or replications is critical to reducing the amount of error within an experiment and is signified above as 'n'.

It is likely that most of the D's above will not equal zero. However, is the difference (D) between the two treatments large enough that in reality they are different? We need a method of deciding if they are sufficiently different from zero to say that there is a significant difference between the two treatments. Although this involves a few calculations, they can easily be carried out using a computer spreadsheet. Start with the original D's and calculate the squares of those numbers (D_i^2):

Observation, i	D_i	D_i^2
1	D_1	D_i^2
⋮	⋮	⋮
N	D_n	D_n^2
Sum	Sum of Ds	Sum of D^2s

Now calculate the following:

$$\overline{D} = \text{average of D's} = \frac{\text{Sum of D's}}{n}$$

$$S = \sqrt{\left(\frac{1}{n-1}\right)\left((\text{Sum of } D^2\text{'s}) - n\overline{D}^2\right)}$$

Look at the ratio $\frac{\overline{D}}{S/\sqrt{n}}$ (ignore whether this is positive or negative).

Now, we must make a decision (using this sample evidence) as to whether there is a difference in the treatments or not. Of course, we can always make a mistake. The idea is to try to control the rate at which mistakes are made. The mistake that is considered most crucial to control is deciding there is a difference in the treatments when in fact the treatments have equal effects. (This is akin to sending an innocent person to jail.) Suppose we let A represent the chances of deciding there is a difference in the treatments when, in fact, they have the same effect. The following gives a good (though rough) guide as to what decision to make.

We decide there is a difference in the treatments if (ignoring the sign) $\frac{\overline{D}}{S/\sqrt{n}}$ is larger than

	With a sample size of	A will be
2	30 or more	5%
2.5	10 – 15	
1.65	30 or more	10%
1.8	10 – 15	
0.8	30 or more	20%
0.88	10 – 15	

Example:

To evaluate the effects of a high level of copper (Cu) in their feed, six chicks were fed a standard basal diet to which 400 parts per million (ppm) of Cu was added. The following data show the feed efficiency ratio (g feed/g weight x $\frac{1}{100}$) at the end of three weeks.

Corn sample	0 ppm P	40 lb/A K	Difference	Squared difference
1	157	191	-34	1156
2	154	171	-17	289
3	165	155	10	100
4	157	167	-10	100
5	159	164	-5	25
6	158	167	-9	81
Total			-65	1751

Now $\overline{D} = \frac{-65}{6} = -10.83$ and

$$S = \sqrt{\left(\frac{1}{6-1}\right)1751 - 6\cdot(-10.83)^2} = 14.47$$

And so, $\frac{\overline{D}}{S/\sqrt{n}} = \frac{10.83}{14.47/\sqrt{6}} = 1.83$.

The sample size of six here is ridiculously small, but, this is only an illustration. Nonetheless, looking at the fourth line in the table given above (sample size of 10 to 15), since 1.83 is greater than 1.8, we will decide there is a difference in the treatments, knowing that we will be making a mistake by this decision 10 percent of the time.

As we stated in the beginning, this Guideline is proposed to compare treatments with statistical precision. If the ratio is not larger (in either case), then you don't have enough evidence to say, with statistical precision, there is a difference in treatments. ■

This Site-Specific Management Guideline was prepared by:

Dr. Timothy A. Wittig
Professor of Mathematics and Statistics
Statistics Coordinator
South Dakota State University
Box 2220
Brookings, SD 57007
Phone: (605) 688-6220
E-mail: wittigt@ur.sdstate.edu

Dr. Zeno W. Wicks III
Professor of Plant Science and Statistics
Northern Plains Biostress 0248B
South Dakota State University
Box 2140C
Brookings, SD 57007
Phone: (605) 688-5542
E-mail: zeno_wicks@sdstate.edu

Ref. # 99088/Item #10-0018