PACE INSIGHTS October, 2000 Volume 6 Number 10 page 1

A Guide to Testing Products and Management Practices Part I: Getting Started

by Wendy Gelernter, Ph.D. and Larry J. Stowell, Ph.D.

Bottom Line: Running your own testing program can help you to make technically sound decisions about which products and management practices will be the most beneficial at your golf course. In this first of a four part series on testing programs, we outline how to design and implement simple, on-site tests. In future issues, we will go more in-depth, with a look at experimental design, statistics and guidelines on the use of monitoring and testing equipment.

Today, the simple act of selecting the best products and practices for your golf course can be daunting. New pesticides, growth regulators, fertilizers, amendments and equipment are introduced each year. These products join the already wide selection of products available for use on golf courses. When you layer onto this the unique combination of soils, turf types and climate that characterize each golf course, it is difficult to predict how new products and practices will perform. Making responsible and effective choices in this area is an important part of the turf manager's job, but it is an area that grows more complex with every new product that is marketed.

To determine which products and management practices work best on your own golf course, consider setting up a testing program to supplement the more general efficacy data that manufacturers and suppliers should have available. The basic steps of running your own testing program are simple, and the rewards are great, giving you added confidence in your selection of products and practices. Test results can also indicate where changes are necessary to improve turf and playability conditions. Conducting your own tests will also allow you to get more out of technical articles on turf management. Another important benefit is that testing provides you with concrete backup to support the decisions and management changes you make.

This is the first in a series of four *PACE Insights* that describes how to set up a successful testing program on a golf course. The series includes:

Part 1: Getting Started. Covers the basic elements of testing programs, attitude, components of an experiment, record keeping.

Part 2: Experimental Design: Simple and efficient experiments, plot design, replication and randomization techniques.

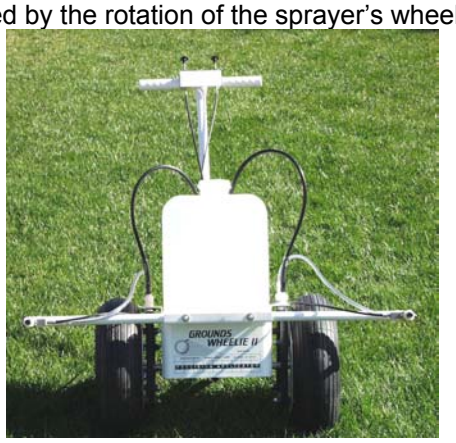
Part 3: Interpreting Results. Evaluation of test results, how to decipher scientific publications, and an introduction to basic statistics.

Part 4: Tools of the Trade. Useful equipment for monitoring and testing, and how to use.

You've got to have an attitude

Before you set up a testing program, perform an attitude check. If you are running a test just to prove that a product will work (or that it will fail), then you've got an attitude problem – one that may influence your results. One of the basic principles of scientific experimentation is to work to keep an open mind. If there isn't a level of suspense as you approach testing (how will the product perform? Will it be better or worse than what I'm using now?), then it's time to assess your objectivity. If your level of bias is low, your results are more valuable. Remember, with a proper unbiased attitude, you are performing tests to *compare, evaluate* and *demonstrate* – never to prove.

A motorless, self-calibrating, wheel-driven sprayer, such as the Wheelie Sprayer below, makes testing easier by allowing application volumes as small as 1 quart. The spray is delivered via peristaltic action powered by the rotation of the sprayer's wheels.



Your experiment

Each experiment is made up of four components:

- Objectives
- Materials and Methods
- Observations
- Discussion (summary)

Consider each of these components before and during the execution of each experiment. To start, put down the following information on paper:

Clearly state the **objective** of the test. Why is the test being conducted in the first place? How will the information be used?

List the **materials** to be used (products, sprayer, etc.) and the **methods** that will be used to implement the test (sprayer settings, product rates).

Once the experiment has started:

Begin recording **observations** for each product or process being tested. Observations can be descriptions of visual characteristics ("excellent", "good", "fair" are all visual descriptions), numerical ratings (weights of clippings, electrical conductivity readings, etc) or relative ratings (subjective performance estimates of quality, such as the 1 – 9 turf quality rating system).

At the end of the experiment:

Review your notes and write a **discussion**, or summary of your findings, why you think these observations occurred, how valid you think the test was, and how the information will be used. How will you incorporate these results into your management program? Are there any follow-up tests that might be useful?

These four components are essential to any testing program. If you omit any of them, you will find it difficult to determine what actually happened, and your time will be wasted.

An example

Figure 1. Problem area that is under investigation in this example. Note the green polka dot patterns of healthy turf, and the damaged surrounding turf.

A golf course testing program should not be limited to evaluation of products. You can use on-site testing to solve any number of problems that arise during the year. Let's run through an example and see how the four different components of an experimental plan are used to design a non-product-related experiment.

Imagine that you are a superintendent who has noticed the worrisome symptom illustrated in Figure 1: small green polka dots of turf, about one-inch in diameter, throughout the low, heavily trafficked areas of your bentgrass green, surrounded by chlorotic (yellow) turf. You suspect that your high soil salt levels, combined with a recent fall aeration may be responsible, since these conditions frequently cause water to preferentially channel down the sand-filled aeration holes, instead of uniformly flowing through the entire soil profile. How would you test the hypothesis that there is no difference between the salt content (electrical conductivity) of the aeration hole sand vs. the soil between the aeration holes?

Your experiment might look something like this:

Objective: Determine whether differences in soil salts (electrical conductivity) are responsible for the differences observed between high quality turf growing in aerification holes vs. low quality (chlorotic) turf growing in the surrounding areas.

Materials and Methods: Three cup-cutter samples will be taken from the stressed area where the green polka-dot pattern is most apparent. A knife will be used to dig out the sand under the green polka dots, and the sand will be placed into a clean plastic container that has been labeled "HOLES". Similar samples will be collected from between aeration holes for comparison and placed in plastic containers labeled "BETWEEN HOLES". Water will be added to each sample until it is saturated. A Cole-Parmer TDS-4 meter (see photo the right) will be used to measure electrical conductivity (EC), and values will be converted to saturated paste equivalent values according to Table 2 (see page 4).

Observations: The green polka dots of healthy turf occurred over aeration holes. The roots were white and more than an inch long. Under the chlorotic turf, the roots were short and discolored.

The ECs for the sand in the aeration holes (under the green turf) averaged 0.3 on the TDS-4 meter,

which converts to **1.6 dS/m** (see Table 2). This is an acceptable level for a bentgrass green, but you should keep an eye on salinity values to make sure they don't get any higher. For the soil collected from between the aeration holes, however, the average TDS-4 meter reading was 2.6, which converts to a whopping **7.8 dS/m**! Looking at Table 1, you can now identify the source of your damaged turf – salt levels are much higher than the 3 –6 dS/M range that is tolerated by creeping bentgrass.

Discussion: The salinity of the new sand in the aeration holes is lower than the surrounding older sand. This indicates that water may be preferentially channeling down the aeration holes (therefore leaching salts away from these areas) rather than percolating uniformly throughout the green. Because the ECs in the chlorotic areas are higher than the 3.0 dS/m that starts to put stress on bentgrass, the front of the green should be leached until the soil salinity is dropped to below 3.0 dS/m in the chlorotic areas in between the aeration holes. As a follow up action, you should continue monitoring ECs in the chlorotic areas to determine whether turf recovery and a reduction in EC is linked. In addition, traffic and compaction are probably making the problem worse, so golfers should be re-routed, using a rope barrier, at least three days per week.

The effects of these management actions will be evaluated 2 weeks and 4 weeks from now to determine whether they helped to improve the situation.

Table 1. Relative tolerance of turfgrasses to soil salinity. dS/m = decisiemens per meter, and is a measure of salinity, or electrical conductivity. **Blue type** = cool season turf; **Red type** = warm season turf.

Sensitive <3 dS/m	Moderate 3-6 dS/m	Somewhat tolerant 6-10 dS/m	Tolerant >10 dS/m
Annual bluegrass	Annual ryegrass	Bentgrass cv. Seaside	Alkaligrass
Colonial bentgrass	Chewings fescue	Perennial ryegrass	Bermudagrass
Kentucky bluegrass	Creeping bentgrass	Tall fescue	Seashore paspalum
Rough bluegrass	Hard fescue	Buffalograss	St. Augustinegrass
Centipedegrass	Bahiagrass	Zoysiagrass	

Data adapted from: Harivandi, M.A., Butler, J.D., Lin, W. 1992. Pp. 207-230 In "Turfgrass", Waddington, D.V., Carrow, R.N. and Shearman, R.C. eds. Monograph no. 32, American Society of Agronomy, Madison, WI.

This fictitious experiment provides an example of several of the benefits of testing that we have identified above. This simple test can be conducted in about one hour, but when the results are translated into action, turf quality can be significantly improved. And this improvement could be made without unnecessary pesticide or fertilizer applications. While these products are useful in many situations, doing your own testing can help you determine whether products are required, or whether a relatively simple change in management practices can provide the desired results.

The power of nothing

An introduction to experimentation wouldn't be complete without stressing the value of **nothing.** This is an important concept. For those of you familiar with experimentation, "nothing" is the same thing as a **non-treated** or **control** area, which is also sometimes called **the check**.

Once you have selected an area within a green, tee, fairway or rough that you want to apply a new treatment to (for example, a fertilizer, pesticide, leaching program, or cultivation practice), you will also need to select an adjacent area to serve as the **non-treated** plot. The non-treated plot should be managed exactly the same way as your treated area, with one exception. That is, you should not subject the non-treated area to the product or practice that you are evaluating.

By applying the concept of **nothing** to your test in this way, you will be able to use the non-treated control as a yardstick to measure any improvement (or damage) that results from the treatment under evaluation. If you fail to include a non-treated plot in your test, you really have no way of knowing how well the new product or practice is performing. For this reason, you should always be wary of manufacturer's data that doesn't include a non-treated or check plot.

The location of the non-treated plot is important, because it must be similar to the area being treated. If the non-treated area is located incorrectly (for example, if you place the non-treated area on low quality turf, and the treated area on good quality turf), the test will be biased. Results will be confusing, and you may make erroneous decisions based on your observations.

Where plywood is king

A piece of plywood, used properly, can save thousands of dollars in unneeded pesticide

applications. Do you believe this? Well, you should by the time you reach the end of this section.

Combined with our understanding of "the power of nothing" described above, plywood is one of the most effective and easy ways of creating "instant" non-treated plots. Let's say you have a fairway showing symptoms of brown patch, and you plan to evaluate the performance of a fungicide for its control. To make instant non-treated plots, simply place a piece of plywood (4 X 8 feet works well) on top of a few of the diseased areas, just before you are ready to begin spraying the test fungicide.

When you treat with the fungicide, the areas covered by the plywood will remain untreated. After the spray has dried, mark each of the plywood's four corners with turf paint. You can now remove the plywood, and the turf paint will allow you to locate the non-treated areas so that you can make your observations. If you observe the treated and non-treated areas daily for several days (recording your observations as described below), you should be able to determine whether the application improved, decreased, or had no effect on turf quality. If there is no visible difference between the treated and non-treated areas, the fungicide probably doesn't have much activity for that particular disease, and you will have saved yourself time, effort and expense by avoiding an unnecessary application. If, on the other hand, the treated area looks better than the non-treated area, then you can treat with the confidence that the product will produce the desired effects.

Record keeping

Without good record keeping, the effort put into a testing program will come to naught, because you will have no way of remembering how and why your results were obtained.

Before starting a testing program, buy several bound composition notebooks. These notebooks are inexpensive, and readily available at office supply and even drug stores. Leave a few pages blank at the front of the book to use as a table of contents, or index. Use only ballpoint pens with ink that is not water-soluble. Tape a business card on the inside cover, so that the book can be returned if it is misplaced.

Use the book to record your Objectives, Materials and Methods, Observations and Discussion, as well as any other thoughts you have about product performance, the reaction of golfers to a management practices, or any difficulty you experience handling or applying a material. Date each entry, and take notes carefully and legibly! Remember -- excess information is always better than insufficient information, so don't be stingy with your words. If you are able to take photos, tape them inside your experimental log; these can be invaluable in summarizing your results. If you are a good record keeper, you'll find that your notebooks will hold their value for years to come -- in resolving disputes about which practice or technique is best, where or how a product was applied, or the history of a problem area of turf.

Table 2. Conversion table for determining the saturated soil extract EC (Extract EC) from the direct TDS-4 saturated soil readings. All values are in dS/m (decisiemens/meter) which = mS/cm = mmhos/cm. Conversions are based on analysis of 32 soil samples that were analyzed as a saturated paste with the TDS-4 meter, and as a vacuum filtered extract using the standard EC measurement.

TDS-4	Extract EC	TDS-4	Extract EC	TDS-4	Extract EC
0.1	1.1	1.1	3.8	2.1	6.5
0.2	1.3	1.2	4.0	2.2	6.7
0.3	1.6	1.3	4.3	2.3	7.0
0.4	1.9	1.4	4.6	2.4	7.3
0.5	2.2	1.5	4.9	2.5	7.6
0.6	2.4	1.6	5.1	2.6	7.8
0.7	2.7	1.7	5.4	2.7	8.1
0.8	3.0	1.8	5.7	2.8	8.4
0.9	3.2	1.9	5.9	2.9	8.6
1.0	3.5	2.0	6.2	3.0	8.9

A Guide to Testing Products and Management Practices

Part II: The Nuts and Bolts of Experimental Design

by Wendy Gelernter, Ph.D. and Larry J. Stowell, Ph.D.

Bottom line: Properly designed tests can be invaluable in identifying the products and practices that are most useful at your specific location. But badly designed tests will yield poor information, will waste your time, and can lead to erroneous decisions. When designing a testing plan, simplicity should be your goal. Carefully choosing the treatments you want to test, understanding how to replicate and randomize for best results, and making sure that plots are placed and sized correctly are the keys to a clean, simple and effective experimental design.

In Part 1 of this series ("Getting Started", in the October, 2000 issue of *PACE Insights*), we discussed how a testing program can help you improve the technical decision making process on the golf course. In this installment of the four part series "How to Test Products and Practices", we will address some of the nuts and bolts of experimental design - how to determine which treatments to test, how many treatments to test, and how to lay out the test on your golf course. If you find that you enjoy conducting your own on-site testing program, and that it has been valuable, there are a few additional sources of information, listed in the "References" section at the end of this article, that also provide some useful information. While most of these references do not address turfgrass systems specifically, they do provide some basic information on testing strategies that will help round your library out.

The experiment pictured below was designed to evaluate 3 different rates of the growth regulator Proxy on poa greens. Each 5 foot X 10 foot rectangular area represents one test plot. Each rate of Proxy was replicated in three different plots. The plots are arranged in three tiers, with each tier representing one replicate.

Perform a background check

Spending quality time thinking and doing some background reading as you plan your experimental project will pay off in the end. Get hold of information on your topic of interest from trade journal articles, scientific publications, technical product literature, product labels, the internet, or data presented at meetings. This allows you to find out which questions have already been addressed, and which questions remain unanswered. You'll also learn which testing methods have worked (or haven't worked) and which pitfalls to avoid. At a minimum, this type of background check will save you time by helping you to avoid repeating the mistakes of others. And at its best, you'll get some new ideas on how to approach your own testing program.

Control the urge

A frequent urge when starting out your own testing program is to test all of your ideas at once. However, this strategy frequently leads to more questions than answers - something a good experimental design can help you to avoid. Remember that some of the most effective experiments are also some of the simplest, where only two treatments are tested -- one new product or cultural practice compared against your current practice, for example.

To be successful in answering your questions with technically sound answers that can be effectively put into practice on your golf course, limit the number of treatments to a manageable number (in test plot lingo, a "treatment" is anything you want to test the effects of - from a new turf variety, to a product rate or formulation, to a new aerification method).

A good rule of thumb is to limit the number of treatments to five or fewer, and not more than 10. There may be times when you will need to exceed these numbers, but be assured you will be more confident in your results when fewer treatments are evaluated in an experiment. Time spent culling out

unnecessary treatments before you start your test will be repaid many times over.

Break it down

If you have penciled out more than 10 treatments, consider breaking the experiment into its logical components. For example, if you are interested in determining the best timing and rates of application for fungicides labeled to control summer patch, you might at first choose to look at three different fungicides, each at the low and high labeled rates, and at two different application dates - preventative (before disease symptoms appear) and curative (after symptoms appear). That sounds like a fairly simple experiment, but in fact you would end up with 13 treatments! That would be:

3 (fungicides) X 2 (rates) X 2 (times of application)
= 12 treatments +1 non-treated control
= 13 treatments.

As you can see, adding extra factors can cause an experiment to blossom into a design that will be difficult to execute and will produce results that are hard to analyze. To simplify the experiment and to make the results easier to interpret, consider splitting the test up into two or more experiments.

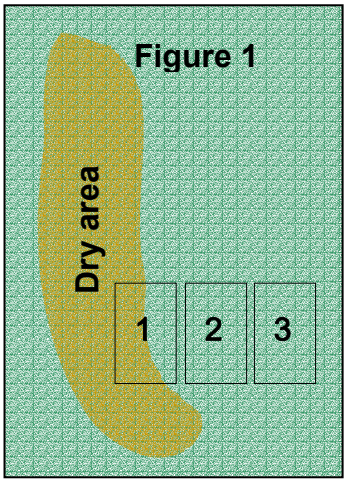
In the example above, the two main factors under investigation are rate of application and timing of application. Why not look at rates first, by keeping application timing the same for all treatments, but varying the rates? The three fungicides could then be tested at the low and high labeled rates, but the timing of application would be the same for all fungicides and all rates; this would be a seven treatment trial including the non-treated control. A second study could look at the effect of different application timings; this time, the rate would be kept the same for all treatments. By breaking the trial down, the execution of the trial and evaluation of the data at the end of the experiment are easier to handle.

Over and over: the beauty of replication

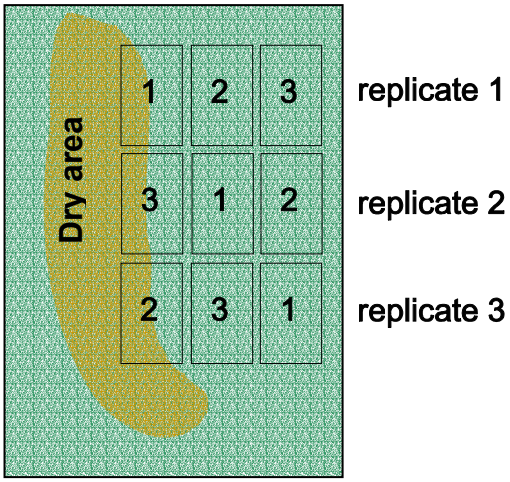
Despite our best efforts, the turf on a green or fairway is usually not homogeneous. There are differences in microclimate, moisture, turf quality and a host of other factors that result in variability that is beyond the control of the researcher. Without **replication** (repeating a treatment in two or more locations), this type of variability can lead us to draw the wrong conclusions from a research trial.

In the example illustrated in Figure 1, two fertilizers (labeled “1”and “2”) are being compared against an untreated control (treatment 3). However, a pre-existing dry spot is located on one side of the green, causing poorer quality turf in the dry area.

This lack of homogeneity in turf quality, which is common, causes an unfair bias against treatment 1. Because the trial in Figure 1 is unreplicated (each treatment is repeated only once), you run the risk of a serious error – wrongly concluding that fertilizer 1 didn’t work at all, and in fact made your turf look worse! Given the experimental design of only one replicate, it's impossible to tell if the results were due to a bad fertilizer, or due to some other factor, such as a pre-existing dry spot.



However, if each treatment is replicated in additional locations as in the illustration below, the chance of unfair bias against any one treatment is greatly reduced.



In most cases, three replications should be sufficient to separate out the good from the lousy treatments. If you read some of the scientific literature, you'll see that researchers frequently also use three replicates, but may use up to six replicates (and sometimes more) if they are trying to tell the difference between some very closely related treatments. However, for your purposes, where you are looking for treatments that are different enough to have a strong impact on turf quality, three replicates (as in the photograph on page 1) is plenty.

Size it right

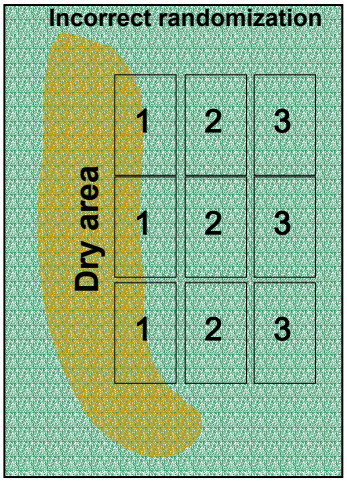
Most research trials are conducted using small plots, usually all placed within one green, one tee or one fairway. Despite the use of the term **small plots**, the last thing you will want to hear when conducting a field experiment is: "your plots are too small." We have found that the larger you make your test plots, the less likely it is that the whole plot will be destroyed by a mishap. For example, a hydraulic leak might damage half of the plot so that it is no longer usable. With larger plots however, the experiment can continue with the non-hydraulic-fluid-damaged areas of all plots being rated. Larger plots also ensure that a disease, insect or weed will be found in the test area. The smallest plots that we recommend for on-site testing are 4 ft X 4 ft (16 sq ft), but our usual small plot size is 5 ft X 10 ft (50 sq ft). For most small plot work, this is a convenient size for a sprayer that applies a 5 ft swath width.

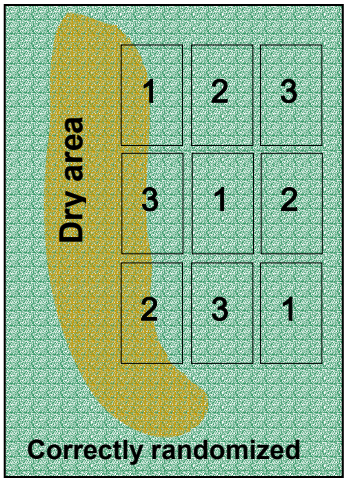
Split greens or macho plots

If you are not adapted to small plot work or just don't want to bother with specialized research equipment (see below), treat one-half or some other portion of a green or fairway using your standard equipment for applications. This is the best way to test a system prior to full adoption of a new cultural practice or new product. In this case, replication will probably have to take place on three separate greens or fairways due to size of the test area. A typical test would entail splitting the green in half and applying a new procedure to one half of a green and your standard treatment to the other half of the green.

Randomization, or rolling the dice

Replication (see above) helped us solve some of the problems presented by natural variability on turf. However, we need yet another tool - **randomization** - to attack this problem and to avoid bias in our design and results. The use of a randomized design helps us to properly arrange the treatments in the test area so that variability is minimized. Using the example of the fertilizer experiment illustrated in Figure 1 once again, there are several choices available to us in how we arrange the different treatments in each replicate. Problems can arise if the treatments are arranged in the same order in each replicate, as in the illustration showing incorrect randomization. Hopefully, you have noticed that treatment 1 is receiving an unfair amount of pressure, because the dry spot is concentrated in the treatment 1 plots. Using this design, would you be able to tell whether the poor performance of treatment 1 be due to the negative effects of fertilizer 1, or is it due to the our having placed treatment 1 plots where soil conditions are dry? You have no way to find this out, using a non-randomized design.





In contrast, treatments can be arranged randomly as in the illustration to the left. In this case, the randomization has been done correctly, and the negative effect of the dry area is more evenly spread over all of the treatments, giving you a fairer look at the performance of each treatment. This illustration represents just one way in which these plots could be randomized: can you think of any others?

A simple method for randomizing treatments within each replicate is to use a deck of cards. Remove the numbered cards that correspond to each of your treatment numbers (another reason to use only 10 treatments!). Shuffle the cards and lay them down to give you a random arrangement of treatment numbers -- it's that simple.

Avoid "the nursery effect"

An interesting phenomenon that occurs at most golf courses is the "nursery effect." As a result of lack of traffic, nurseries frequently survive without disease and stress damage when most of the greens in play are struggling to survive. Under the low-stress conditions seen on most nurseries, products and practices usually perform differently than on the rest of the course. For this reason, nursery greens should be avoided for experiments - unless you are investigating some really wild idea that is too risky to try on greens in play.

Measurements

There are many factors that will influence the outcome of an experiment and only a few of which you will be able to control. The accuracy and precision with which products and practices are applied to the turf is one area where you do have some control, and where cutting corners may result

in wasted time. The more care taken in precisely making measurements and calibrations, the more likely the results will be meaningful and repeatable.

You will have to accurately measure time using a stop watch, distance using a tape measure, volumes using graduated cylinders, syringes, pipettes, or precision flow meters, and weights using balances (scales) that can measure within 1 - 5% of the desired unit.

Try to measure all components with an accuracy of 1%. This means that to measure 100 grams of a product, the balance will have to have accuracy of 1 g. A standard triple beam balance will provide this level of accuracy for about $150. Volume measurements can be carried out using a variety of instruments including disposable pipettes with accuracy down to 0.01 ml for small volumes.

Equipment costs and your time

There are a variety of sources for testing equipment. Your existing equipment is the first place to start. However, if you are interested in small plot applications, Table 1 provides a list of suppliers and recommended items to assist in your efforts. Don't be fooled by the relatively low cost of the equipment needed to conduct testing programs. The investment of your time during experiment design, execution, observation and summary are far more costly than any equipment you might purchase. For that reason, a carefully designed experiment is one that will provide the greatest benefit at the least cost. Your golf course turf quality will benefit and your budget may drop, but be sure that you can afford the time needed to complete an experiment before you get started.

And, as a rule of thumb, if you think it will take half an hour to calibrate your sprayer, allot twice that time. For some strange and perverted reason, experiments always take at least twice as long as you think they will when you are sitting at your desk drafting up the objectives and materials and methods.

It's the law

Remember, it is illegal to use any pesticide that is not properly labeled, stored, and handled according its label. This extends to the use of labeled products on pests or application to sites that are not explicitly listed on the product label. Check with your County Agricultural Commissioner's office to find out what local and regional regulations must be complied with if you are testing a non-labeled product, or a non-labeled use of that product. Obtain the proper permits and certificates before conducting trials with products outside the constraints of the product label - it's the law. If you are uncertain, it's best to stick to experiments with labeled products.

References

Camper, N.D. ed., 1986. Research methods in weed science. Southern Weed Science Society, Champaign. 486 pp.

Hickey, K.D., 1986. Methods for evaluating pesticides for control of plant pathogens. APS Press, St. Paul. 312 pp.

Little, T.M., Hills, F.J., 1978. Agricultural experimentation. John Wiley and Sons, NY. 350 pp.

Table 1. Commonly used equipment in product testing. Prices are ballpark estimates for each item to provide a rough idea of the relatively low cost of equipment needed to test products on-site.

Source	Description	Cat. No.	Quantity	Price
A.M. Leonard 800-433-0633	36"Gandy spreader	36H12	1	252.00
Cole Parmer 800-323-4340	150 g Scale	H-11300-16	1	125.00
"	5000 g scale	H-1100-3-20	1	111.00
"	Container for 5000 g scale	H-11003-60	1	10.00
"	Graduated cylinder 500 ml	H-6137-90	2	15.00
"	Pipette pump	H-06221-03	1	21.00
"	Serological pipettes 1.0 ml	H-13000-06	1,000	146.00
"	Serological pipettes 10.0 ml	H-13000-36	500	170.00

A Guide to Testing Products and Management Practices

Part III: Using Statistics to Interpret Results

by Wendy Gelernter, Ph.D. and Larry J. Stowell, Ph.D.

Bottom line: Using some simple statistics is the final step in conducting a field experiment, and can sometime reveal what appear to be hidden truths embedded in the data. Try not to ignore this critical step that will help you get the most out of your experimental testing program.

Luck. An uneducated gambler in Las Vegas depends upon it, while the more experienced gambler carefully calculates the probability that they will win. Similarly, as a professional and as a superintendent, you prefer not to rely on lucky guesses when making management decisions on the golf course. Instead, you strive to make sure that your key decisions are based on factual information that allows you to accurately predict how new products and management practices will perform on the golf course.

When designed properly, a good testing program helps to support you in this effort. While you can never eliminate the possibility of unexpected results, you can surely reduce the possibility that you will be unpleasantly surprised by basing your decisions on data from a sound testing program.

Statistics: managing the game of chance

When a "fair" coin is tossed into the air, the likelihood that the coin lands with the heads facing up is 1/2, or 50%. This probability represents the number of heads on the coin (1) divided by the total number of sides on the coin (2, heads and tails). Probability theory tells us that there is a 50% chance that you will win a bet every time the coin is tossed regardless of whether you select heads or tails. Even if five tosses of the coin come up tails, the chance that the next toss of the coin will be heads is still 50% -- no more and no less than for any other toss of a fair coin.

When a field test is conducted, the odds are not so easily calculated as they are for a coin toss. Why is this? The answer is that the number of **variables**, or factors than can contribute to the outcome, are much higher for a field test than for a coin toss. In a field test, the turfgrass variety, turfgrass stress, soil type, traffic patterns, weather etc. can have a big effect on the performance of products and practices. In contrast, the number of variables contributing to the outcome of a coin toss are limited.

Because we cannot use guesswork, probability theory or any other system to predict how a product or practice will perform, field tests are conducted to give us information that can be used to make the best possible predictions and decisions.

Statistics is the tool that allows you, as objectively as possible, to analyze the information collected from field tests, and to predict, with as much confidence as possible, which products or practices will give you the best results. In the first two installments of this series, we described how to set up field tests, and how to collect the results. In this 3rd installment, we will describe the final steps -- how to analyze the results statistically, by calculating the **mean**, the **standard deviation**, and the **confidence interval.** In addition, we'll review methods that will allow you to clearly represent the results in the form of **line graphs, bar charts** and **data tables,** for use in your own records and for presentations to greens committees, general managers and others.

A real life example

To give our discussion of statistics some grounding in reality, we will use results from a field test conducted by the PACE Turfgrass Research Institute in 1997. This test was conducted with the assistance of Bill Gallegos, CGCS at Los Coyotes Country Club, and with financial support from Valent Corporation. The objective of the test was to look at the performance of 3 different rates of an experimental fungicide (procymidone, Valent Corporation) and to compare it to a standard fungicide, iprodione (Chipco 26019, Aventis), as well as no fungicide (the control treatment) for control of dollar spot, *Sclerotinia homeocarpa,* on a creeping bentgrass nursery. The five different treatments tested are listed in Table 1. Each treatment was replicated three times, and treatments were randomized. Results were collected two weeks after the fungicides were applied by making a visual estimate of percent turf damage due to dollar spot.

How do we use this data to make a decision on the best product and best rate to use for controlling dollar spot? By calculating the mean, the standard deviation and the confidence interval, as described below.

Calculating the mean

The first statistic to calculate is the **mean**, or average rating for each treatment. The mean is calculated by summing the values for each replicate of a given treatment, and then dividing by the number of replicates. For example, in our dollar spot experiment, the mean percent dollar spot in plots treated with procymidone at 0.5 oz active ingredient/1000 sq ft is 4.667 (rounded to 4.7 in Table 1):

sum of values (5 + 7 + 2) ÷ number of replicates (3) = mean (4.667)

To get the remaining mean values, this process is repeated for each treatment, as illustrated in Table 1. The mean can be easily calculated with pencil and paper. If you are using a calculator, the mean may be represented by the symbol $\bar{X}$. Although the mean is a powerful statistic, when used by itself, it can be misleading and can push you towards poor decisions. This is because the calculation of the mean doesn't take into account the **variability** of the results.

Variability: a complicating factor

There are many factors beyond our immediate control at work on a golf course, such as microclimate, moisture, turf quality, pest pressure, etc. As discussed in Part 2 of this series, these factors exert a powerful force on the way a product or practice performs, and how consistently it performs. The use of replication (repeating a treatment in more than one location) in designing your field test helps to minimize the effects of variability, but it can't erase them. As a result, it is extremely rare for a given treatment to produce the same result each time it is applied. In the dollar spot experiment, for example, procymidone applied at 0.5 oz active ingredient/1000 square feet produced three different disease incidence levels in each of three apparently identical plots -- 5%, 7% and 2% (Table 1).

How does variability affect your interpretation of the results? Let's assume that in the dollar spot experiment above, variability was much lower. In that case, the percent dollar spot values for the procymidone 0.5 oz treatment would be much more similar, for example, 4.6%, 4.7% and 4.7%. The mean for these hypothetical values would be identical to the mean calculated above -- 4.667, but the variability would be less.

Which set of data gives you a greater guarantee that the product will perform the same way the next time you apply it? Which data set gives you a greater sense of confidence? Statisticians tell us that the data set with the lowest variability gives us the best predictions for how products will perform. So, even when the means are the same for two data sets, we still want to know how variable the data was.

Measuring variability: the standard deviation and the confidence interval

There are a variety of statistics used to measure variability, but the most commonly used measure is the **standard deviation,** frequently represented by the symbol "**S**" on a hand calculator. A small standard deviation indicates that there is less variability associated with the mean, or that the data is more consistent than the same mean with a large standard deviation. In the dollar spot example presented in Table 1, the highest standard deviation (6.3) occurs in the non-treated check treatment, and the lowest standard deviation (0.6) occurs in two of the procymidone treatments -- the 1.5 oz and 2.5 oz rates.

Calculating the standard deviation is more complicated than calculating the mean, and we encourage you to purchase a calculator (most simple scientific calculators include standard deviation), or use a spread sheet program, such as Microsoft Excel, that performs the standard deviation function.

Table 1. Results of a fungicide trial for control of dollar spot on creeping bentgrass. Rates of fungicides are represented as ounces of active ingredient per 1000 sq ft. Check refers to the non-treated check plot.

	Procymidone rate oz ai/1000 sq ft			Iprodione rate oz ai/1000 sq ft	Non-treated Check
	0.5	1.5	2.5	2.0	----
% dollar spot (replicate 1)	5	2	2	2	13
% dollar spot (replicate 2)	7	2	1	4	25
% dollar spot (replicate 3)	2	1	1	1	20
TOTAL	14	5	4	7	58
Number of replicates	3	3	3	3	3
Mean (=total ÷replicates)	4.7	1.7	1.3	2.3	19.3
Standard Deviation (S)	2.5	0.6	0.6	1.5	6.3
Confidence Interval (mean —S) to (mean + S)	2.2 - 7.2	1.4 - 2.0	0.7 - 1.9	0.8 - 3.8	13.0 - 25.6

Looking at the means and standard deviations in Table 1, which treatment or treatments do you think gave the best dollar spot control? We still have one more calculation to perform before we can answer that question -- **the confidence interval.** The confidence interval is related to the standard deviation, and is an easy way to represent the interval, or range of values, or degree of variability associated with a mean. The

lower end of the interval is calculated by subtracting the standard deviation from the mean, and the higher end of the interval is calculated by adding the standard deviation to the mean. Staying with the example of procymidone at 0.5 oz, the confidence interval for this treatment would range from 2.2 (4.7 — 2.5) to 7.2 (4.7 + 2.5). In other words, we have a high level of confidence that the mean value for this treatment falls between 2.2% and 7.2%; and our best estimate for that mean is 4.7%.

To find out which treatments performed statistically differently from another, look for the treatments where the range of values of the confidence intervals do not overlap. For example, the non-treated check, with confidence limits of 13.0 - 25.6, is statistically different from all of the other treatments, whose confidence intervals never get as high as 13.0. In contrast, procymidone at 1.5 oz and 2.5 oz have overlapping confidence intervals. This means that, based upon the data from this trial, the treatments did not perform differently.

Once all of your calculations have been completed, make a **summary table** similar to that in Table 2. This table shows a letter following each mean value, something you will frequently encounter when reading scientific papers. These letters are a way of illustrating which confidence intervals overlap, and which don't. For example, values (such as 1.3%, 1.7% and 2.3%) followed by the letter "a" have overlapping confidence intervals and are therefore not statistically different from one another. In contrast, values that are followed by different letters ("b" or "c" in the case of Table 2), are statistically different from those followed by "a"s.

Table 2. Summary of dollar spot control results using tabular format. The numbers in the percent Dollar Spot column followed by the same letter are not significantly different using the standard deviation as the confidence interval. If the confidence intervals overlap, it is unlikely that the means are different. By this analysis, the top performing treatments can be identified, and are highlighted in blue.

Treatment (rate per 1,000 sq ft)	Mean % Dollar Spot	Confidence Interval
Procymidone 2.5 oz	**1.3 a**	**0.7 - 1.9**
Procymidone 1.5 oz	**1.7 a**	**1.4 - 2.0**
Iprodione 2.0 oz	**2.3 ab**	**0.8 - 3.8**
Procymidone 0.5 oz	4.7 b	2.2 - 7.2
Check 0.0	19.3 c	13.0 - 25.6

In fact, all of the information required to determine which treatments are best, which are worst, and which are the statistically considered to be the same, is contained in Table 2, but it's difficult for most of us to read tables. That's where graphs come in.

One picture is worth a thousand words

One of the best approaches towards interpreting results is to graph the information. There are two types of graphs that are used to illustrate data collected from field trials such as the fungicide trial described above -- **the bar chart** and **the line chart**.

For either type of chart, there are two axes, or lines, that define the chart -- the horizontal axis, also called the "X" axis, and the vertical axis, also known as the "Y" axis. Figure 1 illustrates the results of the dollar spot fungicide experiment presented in a bar chart. The X axis has no numerical units, just treatment names. The Y axis represents the mean percent dollar spot values presented in Table 1. Thus, the bar for the non-treated check is the tallest bar, registering at 19.3% dollar spot. The vertical lines extending above and below the tops of each bar are called **error bars** and represent the confidence intervals for each treatment mean.

Using the bar chart in Figure 1, note how quickly you are able to determine that all of the fungicide treatments performed better than the non-treated check. And how clear it is that the 2.5 oz procymidone treatment is the best overall performer, but is not statistically different from 1.5 oz procymidone or 2.0 oz iprodione. Finally, the 0.5 oz procymidone treatment did not perform as well as the other procymidone treatments. This bar chart would allow you to conclude that the two best treatments in this study were the 1.5 and 2.5 oz/1000 sq ft rates of procymidone. The statistical analysis tells us that the 1.5 oz rate is likely to perform as well as the 2.5 oz rate.

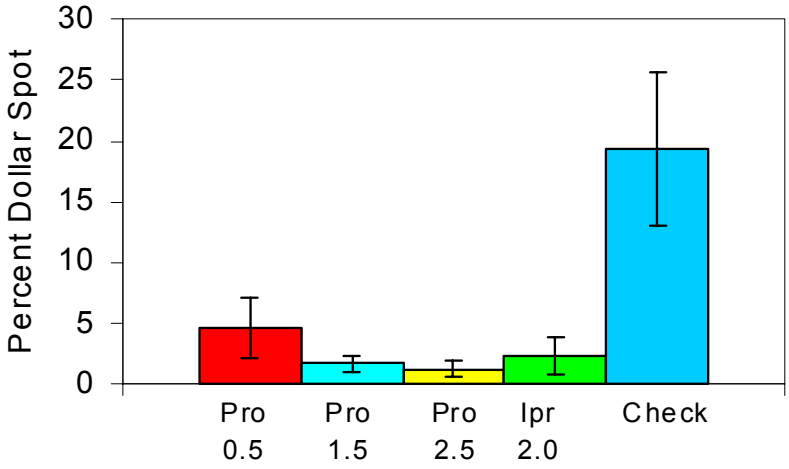


Figure 1. Bar graph of treatment categories and error bars representing one standard deviation above and below the mean. Pro refers to procymidone and Ipr represents iprodione. The numbers under Pro and Ipr are the rates applied in oz ai/M. Note that none of the fungicide treatment error bars overlap with the non-treated check error bar.

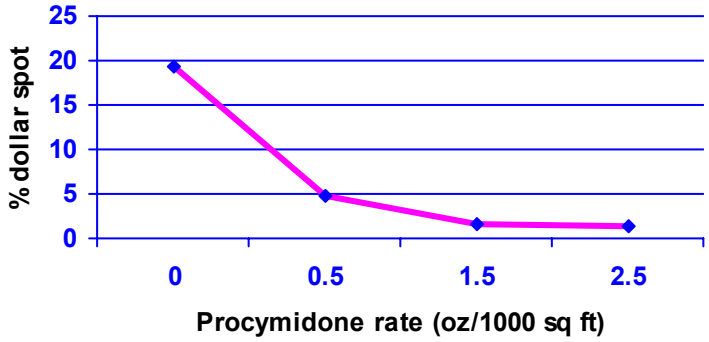


Figure 2. Line graph illustrating the effect of different procymidone rates on percent dollar spot incidence. The non-treated check is represented by the 0.0 oz ai/1000 sq ft treatment value.

Still using the dollar spot data from Table 1, a **line chart** can also be created, as in Figure 2. Note that the line drawn between each of the points allows us to do something we can't do on a bar chart -- to roughly estimate the percent dollar spot that might be produced by any rate of procymidone between 0.0 and 2.5 oz ai/1000 sq ft. It's best to use line charts when you are graphing **continuous variables** such as time, or as in the case of Figure 2, various rates of the same product. A bar chart, on the other hand, is most useful when you are comparing **discrete** (discontinuous) variables that fall into different categories. For example, the bar chart in Figure 1 compares the performance of two different products (procymidone and iprodione).

We suggest that you always try to graph your data. You can plot the results by hand using graph paper, or you can let a spreadsheet program on the computer do it for you, automatically.

Finale

This series of *PACE Insights* has so far described simple methods for designing a field testing program -- from developing an experiment plan with clear objectives, to executing an experiment, to analyzing the results. Although this is a cursory look at the scientific process, we hope it encourages you to begin, or if you have already started, to continue testing new ideas. Remember to take care to record your objectives, materials and methods, results, and conclusions. The next time someone asks you why you selected a particular practice or product, you may be able to pull a notebook from the shelf and point to a graph illustrating the advantages of your approach. Aside from personal pride, there is no better way to answer an agronomic practices question than to run a carefully designed, simple experiment.

References

For a good, relatively simple explanation of statistics, the following book is recommended.

Little, T.M., Hills, F.J., 1978. Agricultural Experimentation. John Wiley and Sons, NY. 350 pp.