

October 2002

The Status of Site-Specific Crop Response Estimation

By

Jess Lowenberg-DeBoer, Director

and

Rodolfo Bongiovanni, Ph.D.

Keywords: Spatial Statistics, Yield, Plot Trials, Economics

Fine-tuning crop management using site-specific crop response information was part of the original vision of precision agriculture. The hope was that producers could use yield monitor, soil test and other information to identify more accurately how crops respond to inputs on their farms. This vision has proven difficult to achieve and some even consider it impossible, but there has been some progress. The goal of this article is to outline the state of the art of site-specific crop response estimation.

Researchers in the U.S. and Argentina have estimated site-specific crop response curves from yield monitor data using various spatial statistics techniques. A study by Lambert and colleagues reviews those studies and the methods used (<http://www.agriculturadeprecision.org/ingles/indice.htm>). The preliminary conclusions coming out of these studies are:

- 1) Designed trials are essential with current technology – data mining yield maps for crop response has not been successful.
- 2) Crop responses can vary widely from place-to-place within fields and as a result the profit maximizing input choices also vary.
- 3) Ordinary Least Squares (OLS) regression can provide reasonably accurate estimates of the spatial variation, but the statistical reliability of these estimates is not good.
- 4) The various spatial regression methods seem to result in roughly the same numerical estimates and similar statistical tests. The choice between spatial regression methods may be influenced by ease of use and software availability.
- 5) Site-specific crop response can vary widely from year-to-year depending on the weather. Making input choices will require risk management.

Data Mining - almost all the successful examples of reliable crop response information from spatial crop data comes from designed trials. The term “data mining” in

this context means analysis that extracts useful information from observations on systems without any planned treatments. One of the key problems with data mining in crop production is that the range of naturally occurring variables may not be enough for estimation of robust response curves.

The study by R.J. Florax and fellow researchers in the Netherlands shows the pitfalls of data mining for crop data. They use hand harvested yield data from a farmer's field in the Republic of Niger, West Africa, to estimate statistically significant responses of millet to soil phosphorous and potassium test levels, but their estimated coefficients are unreasonably large, probably because the naturally occurring soil phosphorous and potassium levels cover only a narrow range of low test levels. For example, most phosphorous tests are below 6 ppm.

Spatial Variability of Response – Studies in Argentina and in Minnesota have shown that crop response can vary dramatically within fields. Both these studies were on nitrogen application on corn. The study in Argentina used a commercially available yield monitor and employs a strip trial design that could be used by any producer with a yield monitor and GPS. The Minnesota work relied on strips harvested by a plot combine on fifty-foot long segments. This would not be practical on a commercial scale, but it does provide an important test of the robustness of results with alternative ways of collecting the harvest data. The constraint is no longer data collection, but it is the software used to make the analysis. The Argentine data was analyzed with a statistical analysis tool called SpaceStat (www.SpaceStat.com). This is a DOS program with a good interface with ArcView.

OLS Coefficients – It should be expected that OLS and spatial regression coefficients have similar values. Statistical theory tells us that OLS coefficients are unbiased estimators even when there is spatial correlation in the data. The problem with OLS is that in the presence of spatial correlation the standard deviations are inflated and consequently many statistical tests are not significant. Spatial regression tends to show many more statistically significant variables than OLS. This facilitates decisionmaking because the estimated coefficients are more reliable

Alternative Approaches to Spatial Statistics – Spatial statistics have developed by several different scientific disciplines. For example, spatial econometrics comes out of geography and regional economics. Geostatistics was developed originally by geologists and mining engineers. Nearest neighbor analysis was initially proposed by an agronomist. It is not surprising that these different approaches make different assumptions. Geostatistics was originally used primarily for creating interpolated maps from sample sites, so it is logical that it assumes that spatial variability changes continuously. Regional economists and geographers often worked with data from political and administrative units of government (e.g. counties, cities, neighborhoods) and so they assumed that spatial correlation was a discrete relationship between these “polygons.” The work by

Lambert and colleagues suggests that all the spatial regression approaches result in the same decision for the Argentine nitrogen response data.

Risk Management – In the Argentine data nitrogen response in a wet year was greatest on the hilltop. In a dry year the response tended to be greatest in the lowland. In any given year the nitrogen response differed greatly from one area of the landscape to another, but ‘expected response’ was very similar everywhere. By ‘expected response’ we mean the curve that results when taking a weighted average of the estimated coefficients for each year. In the Argentine data the different responses in the wet and dry years effectively balance each other, resulting in an average response that differs little from one part of the field to another. While variable rate nitrogen was profitable “after the fact” for all fields, the expected response seldom favored variable rate application because there was little spatial difference in expected responses.

Maximizing the expected net value of crop response is a simple approach to dealing with crop variability, but is it the best way? Optimistic producers might prefer to always fertilizer for the wet year, while pessimistic producers might favor a very conservative approach of always fertilizing for the dry year. The best way to manage risk in spatial management of nitrogen is a current topic of research.

Conclusions

Estimation of site-specific crop responses from yield monitor data is possible with designed trials. It can be accomplished with data from simple strip trials. The major constraints to estimation of site-specific response curves is that the software is not very user friendly and the skills needed to interpret results are not widespread. The key to robust statistical estimation is that the spatial correlation of the observations must be explicitly modeled. Spatial correlation cannot be ignored.