ILLINOIS

Use Caution in Interpreting Clusters of Similar Values in Soil Fertility Maps

By Linda Anderson and Don Bullock

Global positioning satellites and associated technology have made variable rate technology (VRT) applications of fertilizer easier to perform. We can accurately record locations of data gathered from fields and then differentially treat areas of those fields. When correctly used, VRT increases fertilizer efficiency. The fertilizer is applied where it is needed and at the proper rate.

Illinois researchers use a simulation to demonstrate that clusters of similar values in a map don't automatically indicate the map is accurate.

An accurate map is needed for VRT fertilizer application. The map must indicate the real varying fertility levels of the field. Experience has taught us that this is much easier said than done and that in fact many maps used currently probably are not accurate and do not portray the actual fertility levels of the fields they represent. We are not suggesting that VRT does not work. Rather we are suggesting that if VRT is to work, then accurate maps must be produced.

To collect information for fertility maps, soil samples are collected – usually on some regular grid interval. In Illinois and much of the central Midwest this is a 330 ft. grid (2.5 acres). Maps are then produced with an implicit assumption that point samples from adjacent areas are representative of those areas and/or are correlated to adjacent and nearby sample points. It is assumed that when several samples of similar value (e.g. low Bray P-1 test) are clustered together in an area, that portion of the field has a low Bray P-1 value. The assumption may or may not be true. Certainly a cautious approach is warranted.

It should be understood that for any given field, similar soil samples can cluster together randomly and not be indicative of uniformity in a particular area of the field. For example, if two copies of the numbers one through fifty (100 pieces of paper) were put into a hat and drawn out one at a time, it would not be surprising if one were occasionally to sequentially draw out from the hat several numbers that are numerically similar (e.g., 46, 48, and 53). The problem is made even worse because most individuals will categorize such quantitative results into a small number of discrete intervals such as low, medium, and high. This results in an almost certain occurrence of a clustering of categories. Again it should be stressed that in this case the clustering would not be indicative of a true fertility area in the field. Rather it is a random, but anticipated, clustering.

An Example

Although taking soil tests is not drawing numbers from a hat, we performed a soil test simulation in the spirit of drawing numbers. Using the distribution (lognormal), mean (61.5), and variance (927) of an actual 80 acre research field, we produced a random data set on a 20 ft. grid (8,192 points) for a Bray P-1 test. Note that in this case the samples were generated randomly and are not correlated. Thus, the value of any given point gives absolutely no information regarding the value of an adjacent or nearby point. We then came back into this random data set and simulated the current standard 330 ft. (2.5 acre) sampling grid (**Figure 1**) and a 165 ft. (0.65 acre) sample grid (**Figure 2**). Clustering of varying fertility categories is evident in both. The clusters are indicative of nothing other than a random event.

It is critical to understand two major issues. First, the single measure for a given block, either the 2.5 acre (**Figure 1**) or the 0.65 acre (**Figure 2**) is not representative of the entire block although many would assume it to be. Second, the information obtained from a single point sample tells us nothing about nearby points. This second critical issue is absolutely true for this data set because it is random, but it is also true for actual field data if the sampling grid results in uncorrelated samples because they were taken at a distance greater than the range of correlation.

The range of correlation can and should be tested with geostatistical techniques, but commonly is not tested. Rather, an inverse distance interpolation is performed by most mapping programs and this assumes that real samples are correlated and that a weighted average of surrounding samples can be used to estimate any point that is not sampled.

This technique assumes that the similarity of points depends on the distance that separates them. Thus, the weights are proportional to the inverse of the distance (1/d), and the samples that are farther away are given less weight individually. Inverse distance squared ($1/d^2$), cubed ($1/d^3$), and to the fourth power ($1/d^4$), are also used. In these cases, distant samples are given even less weight than in inverse

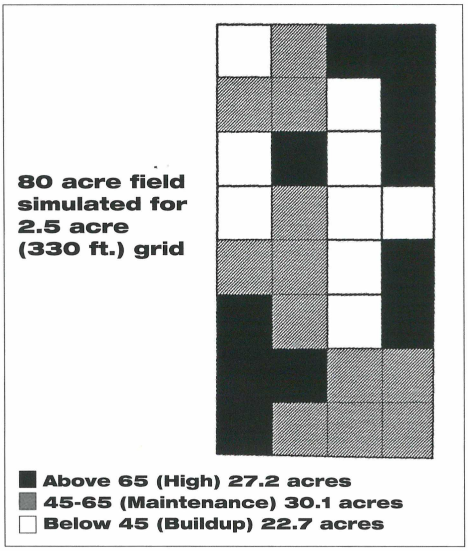

FIGURE 1. Field simulation of soil test P distribution using random numbers, 330 ft. grid.

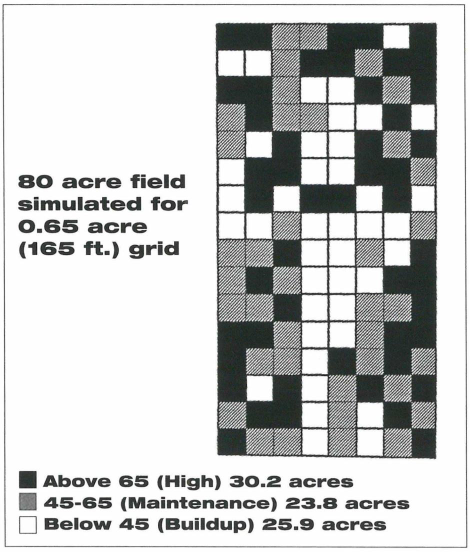

FIGURE 2. Field simulation of soil test P distribution using random numbers, 165 ft. grid.

distance, and the nearest samples are given most of the weight.

For example, given a 100 ft. interval, and a point to be estimated halfway between two samples, the closest samples would have weights proportional to 1/50, and the next 1/150, if inverse distance (1/d) is used. If $1/d^2$ is used, the weights would be proportional to $1/50^2$ (1/2,500) and $1/150^2$ (1/22,500). Thus, with inverse distance, the nearer samples would be weighted three times more than the farther samples. With inverse distance squared, the weights would be nine times greater, with inverse distance cubed, twenty-seven times greater. For $1/d^3$ or $1/d^4$, only the very nearest points are included in the estimate. This has the effect of simply drawing lines around all areas of similar soil test values. If the field is highly variable, very small areas are defined. Additionally, current algorithms will consider issues such as the number of neighbors and/or a minimum distance.

We used such a technique with the results of our generated random data set to produce soil fertility maps of the 330 ft. grid (1/d) (**Figure 3**); 165 ft. grid (1/d) (**Figure 4**) and 165 ft. grid ($1/d^3$) (**Figure 5**). The first two examples using the inverse distance produced convincing, if not similar maps. The last example, using the inverse distance cubed, produced a seemingly overly detailed map.

While the first two maps, in particular, look reasonable and fit well with our expectations for soil fertility variation, they are not correct and represent nothing other than a random clustering of values and the ingenuity of the programmers that developed the mapping software. We should note that such an endeavor predicts certain values for non-sampled points based on nearby points, but in fact the predictions are worthless in this case. If a nonsampled point does fall into the predicted category it is simply by chance.

Summary

We are not suggesting that VRT will not or does not work for fertilizer application. We believe sincerely that VRT can work and has much to contribute to the improvement of fertilizer use efficiency. VRT must have an accurate map, however, and that map must be based upon a

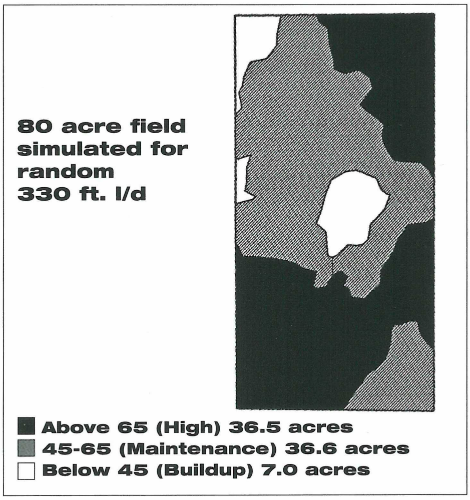

FIGURE 3. Field simulation of soil test P distribution, 330 ft. grid, inverse distance weighting.

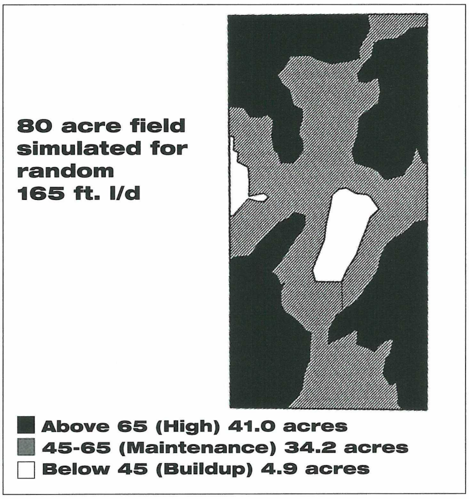

FIGURE 4. Field simulation of soil test P distribution, 165 ft. grid, inverse distance weighting.

sampling density which includes distances at which samples are correlated. We wish we could provide guidance to the required sampling density for all fields, but we and others are still researching that question. We have seen cases where it appears that 2.5 acre samples will work, but we have also seen many cases where a far more dense sampling regime must be used. We strongly argue that all fields be given more rigorous geostatistical consideration. We also believe that the best way to make VRT fertilizer application decisions for many fields is to base those decisions upon previous yield maps and nutrient removal calculations. BC

Linda Anderson is Research Specialist and Dr. Bullock is Associate Professor of Agronomy and Biometry, Biometry Group, Dept. of Crop Sciences, University of Illinois, Urbana, IL.

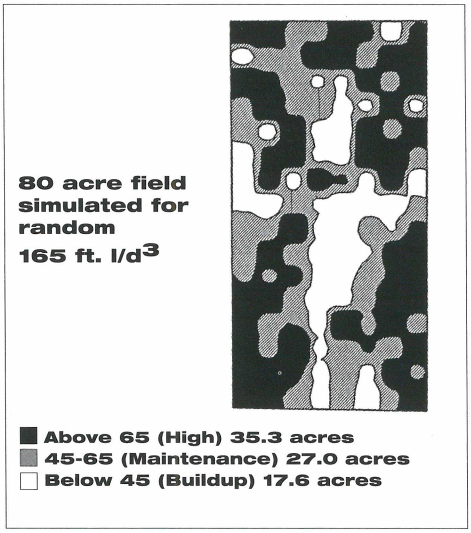

FIGURE 5. Field simulation of soil test P distribution, inverse distance cubed weighting.

Rice Yield... *(continued from page 25)*

would require a moderate addition of K fertilizer based on a composite sample assessment. However, the surface map of soil test K, **Figure 3**, was even more revealing. A large part of the field was shown to have K values at or below the critical level. The grid soil sampling and GIS evaluation plainly illustrated a compelling need for adequate K fertilization.

Summary

Rice production was almost certainly limited by P and K fertility as indicated by yield monitoring and soil test data. The most limited areas of P and K availability corresponded with the high yielding areas. Evidently, larger amounts of soil P and K were being removed where yields were placing the greatest demand. Rice in the less-demanding low-yielding areas was probably restricted by poor soil physical conditions and was not found to be limited by fertility considerations.

The combination of site-specific yield and soil sampling data provided a significant improvement in the quality of information available to make production assessments. While the expense of generating these types of site-specific data is significant, the increased insight and number of yield-improving options offer great promise. BC

Dr. Baker is Research Assistant Professor and S. D. Carroll is Research Specialist, University of Arkansas Soil Testing and Research Laboratory, P.O. Drawer 767, Marianna, AR 72360.

***Acknowledgment**: The authors wish to express their appreciation and recognize the investigative contribution, equipment use, and support by Archie Mason of Pro Ag Services located in Wynne, AR.*