2004

# EC04-704 Precision Agriculture: Listening to the Story Told by Yield Maps

Viacheslav I. Adamchuk
*University of Nebraska-Lincoln*, viacheslav.adamchuk@mcgill.ca

Achim Dobermann
*University of Nebraska - Lincoln*

Jianli Ping
*University of Nebraska - Lincoln*

Adamchuk, Viacheslav I.; Dobermann, Achim; and Ping, Jianli, "EC04-704 Precision Agriculture: Listening to the Story Told by Yield Maps" (2004). *Historical Materials from University of Nebraska-Lincoln Extension*. 710.
https://digitalcommons.unl.edu/extensionhist/710

University of Nebraska Cooperative Extension EC 04-704

Precision AGRICULTURE

# Listening to the Story Told by Yield Maps

Viacheslav I. Adamchuk, Extension Precision Agriculture Engineer
Achim Dobermann, Extension Soil Fertility and Nutrient Management Specialist
Jianli Ping, Post-Doctoral Research Scientist, NU Department of Agronomy and Horticulture

**RESOURCES**

For more information about precision agriculture research, education and demonstration programs at the University of Nebraska, visit the Web site at http://precisionagriculture.unl.edu/

When discussing current issues in precision agriculture, the terms "yield mapping" or "yield monitoring" often are used to refer to one of the most crucial components of the entire system for site-specific crop management. In fact, yield monitoring equipment was introduced in the early 1990s and is increasingly considered a conventional practice in modern agriculture. The pioneers of precision agriculture already have generated several years of yield history and have examined different ways of interpreting and processing these data. The goal of this publication is to review several common methods of yield data analysis and to discuss potential applications for the information obtained.

## Yield Mapping Concept

Yield mapping refers to the process of collecting georeferenced data on crop yield and characteristics, such as moisture content, while the crop is being harvested. Various methods, using a range of sensors, have been developed for mapping crop yields. The basic components of a grain yield mapping system are shown in *Figure 1*.

Grain harvested by the combine in a unit of time is usually determined with a grain flow sensor mounted at the exit of the clean grain elevator. A moisture sensor is used to compensate for grain moisture variability. A yield monitor display and a Global Positioning System (GPS) receiver are used to georeference and record data. A header position sensor is used to distinguish measurements logged during turns. A travel speed sensor determines the distance the combine travels during a certain logging interval. (Sometimes travel speed is measured with a GPS receiver or a radar or ultrasonic sensor.) Finally, the speed of the clean grain elevator is used by some yield mapping systems to improve the accuracy of grain flow measurements.

Each sensor has to be properly calibrated according to the operator's manual. Calibration converts the sensor's signal to physical parameters. A proprietary binary log file is created during harvest to record the output of all sensors as a function of time. This file

*Figure 1. Teaching display with the major components of a grain yield mapping system.*

IANR
Cooperative Extension
Institute of Agriculture and Natural Resources
University of Nebraska

can be converted to a text format or displayed as a map using the yield monitor vendor's software. The yield data file generated using an Ag Leader Technology, Inc. (Ames, Iowa) yield monitor (advanced export format) can serve as an example of a log file represented as text. Every line of such a file represents a single point in the map and consists of 17 comma-separated entries in the following order:

1. Geographic longitude in decimal degrees
2. Geographic latitude in decimal degrees
3. Grain flow (*Flow*) in pounds per second
4. GPS time in seconds
5. Logging interval (*Time*) in seconds
6. Distance traveled during the logging interval (*Length*) in inches
7. Swath width (*Width*) in inches
8. Grain moisture (*Moisture*) in percent
9. Header position
10. Pass number
11. Monitor serial number
12. Producer's Field ID
13. Producer's Load ID
14. Grain type
15. GPS status
16. Position dilution of precision (accuracy of geographic position)
17. Altitude (elevation)

A simple calculation can determine the harvested yield in convenient units such as bu/acre:

$$Yield = K \cdot \frac{Flow \cdot Time}{Width \cdot Length} \tag{1}$$

The unit conversion coefficient *K* equals *112011* for corn or sorghum and *104544* for soybeans or wheat. The yield obtained can then be compensated for moisture variation (in order to determine the grain yield at the reference moisture) using the following formula:

$$Yield_{compensated} = Yield \cdot \frac{100 - Moisture}{100 - Moisture_{reference}} \tag{2}$$

Standard reference moisture values are 15.5 percent for corn, 13 percent for soybeans and sorghum, and 12 percent for wheat.

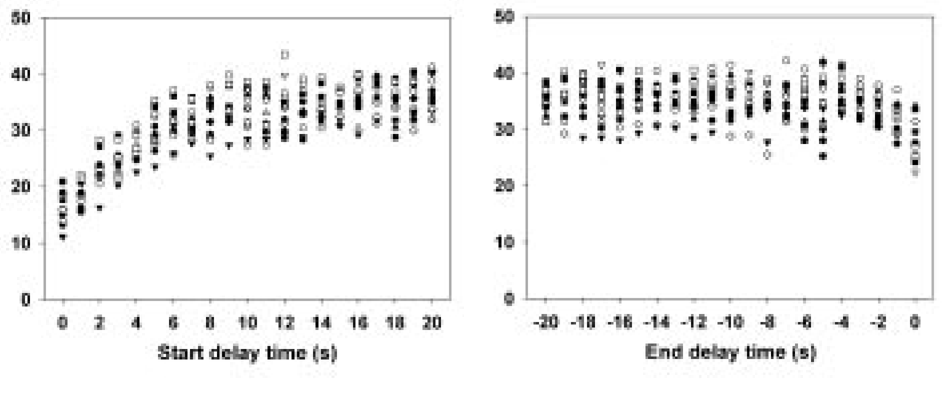

*Figure 2. Grain flows measured by the yield monitor near start (left) and end (right) of harvest passes. (The different symbols show different harvest passes in the same field.)*

## Processing Yield Maps

The yield calculated at each field location can be displayed on a map using a Geographic Information System (GIS) software package. The raw log file, however, contains points recorded during turns and the sensor measurements do not correspond to the exact harvest locations because grain flow through a combine is a delayed process (unless real-time correction is applied). To eliminate these obvious errors, the raw data is shifted to compensate for the combining delay, and the points corresponding to the header up position are removed. Settings for grain flow delay are combine- and sometimes even crop-specific, but typical values for grain crops range from about 10 to 12 seconds.

Usually a few points at the beginning and at the end of a pass should be removed as well. These are referred to as start- and end-pass delays. Start-pass delays occur when the combine starts harvesting the crop, but grain flow has not stabilized because the elevator is gradually filling up. Similarly, end-pass delays occur when the combine moves out of the crop and grain flow gradually declines to zero when the elevator is completely emptied. *Figure 2* illustrates start- and end-pass delays for a series of harvest passes in the same field. In this example, a start-pass delay of eight seconds and an end-pass delay of four seconds were the appropriate choices. Consult the manufacturer of your yield monitor for the most appropriate settings to use with your combine.

Shifting of raw data to correct for grain flow delay as well as deletion of points that represent header status up and start- and end-pass delays is the primary data filtering procedure built into software supplied with yield mapping systems. While these erroneous data points are the most common, other types of error often are found in yield map data. Several data filtering algorithms have been developed (and are incorporated into commercially available agricultural GIS products) to remove data points that fail predefined criteria, which include:

- points with yield values or individual sensor measurements exceeding the possible range;
- outliers based on descriptive statistics (outside the range of normal distribution);
- points with detectible misplacement (e.g. co-aligned points);
- points that do not agree with a predefined statistical estimate based on local neighborhood statistics (significantly different from their neighbors).

*Figure 3* shows an example of a yield map using the unfiltered raw data and a map of all data points that were removed as a result of six cleaning steps: 1) header status up, 2) start- and end-pass delays, 3) grain flow, distance traveled, and grain moisture outliers, 4) points exceeding minimum and maximum yield limits, 5) local neighborhood outliers, and 6) short segments and co-located points.

The order and complexity of the filtering algorithm defines the amount of data points removed; the maps obtained may differ slightly. Most filtering algorithms remove about 10 percent to 20 percent of the total yield data points, which typically improves the frequency distribution of the yield *(Figure 3)*.

Reasons for these errors can be numerous, but the most obvious are:

- varying the crop width which enters the header during harvest;
- changing lag time of the grain as it goes through the threshing mechanism;
- surging grain through the combine grain transport system;
- grain losses from the combine;
- travel speed, grain flow and grain moisture measurement errors; and
- the inherent 'wandering' error from the GPS.

Although not all listed errors can be eliminated with filtering algorithms, proper calibration of each sensor can significantly reduce the uncertainty of yield measurements. However, even if the manufacturer's specifications on sensor accuracy are met, the overall yield estimation error remains significant. For example, if *flow*, *time*, *width*, and *length* measurements *(Equation 1)* have 5 percent, 0 percent, 6 percent, and 4 percent uncertainty, respectively, the overall uncertainty of the calculated yield is 8.8 percent (found as the square root of the sum of the squared uncertainties of each measurement). This error does not include moisture compensation, combine dynamics and GPS position error.

If the above estimated uncertainty is random, averaging or smoothing the data over a larger area may improve the overall accuracy of the yield map. This can also be achieved by increasing the logging interval (from 1 to 3 seconds). The drawback of averaging is that erroneous values may not be as obvious and data filtering may be inefficient. Therefore, the best strategy would be to record data with the highest available frequency and apply smoothing after filtering.

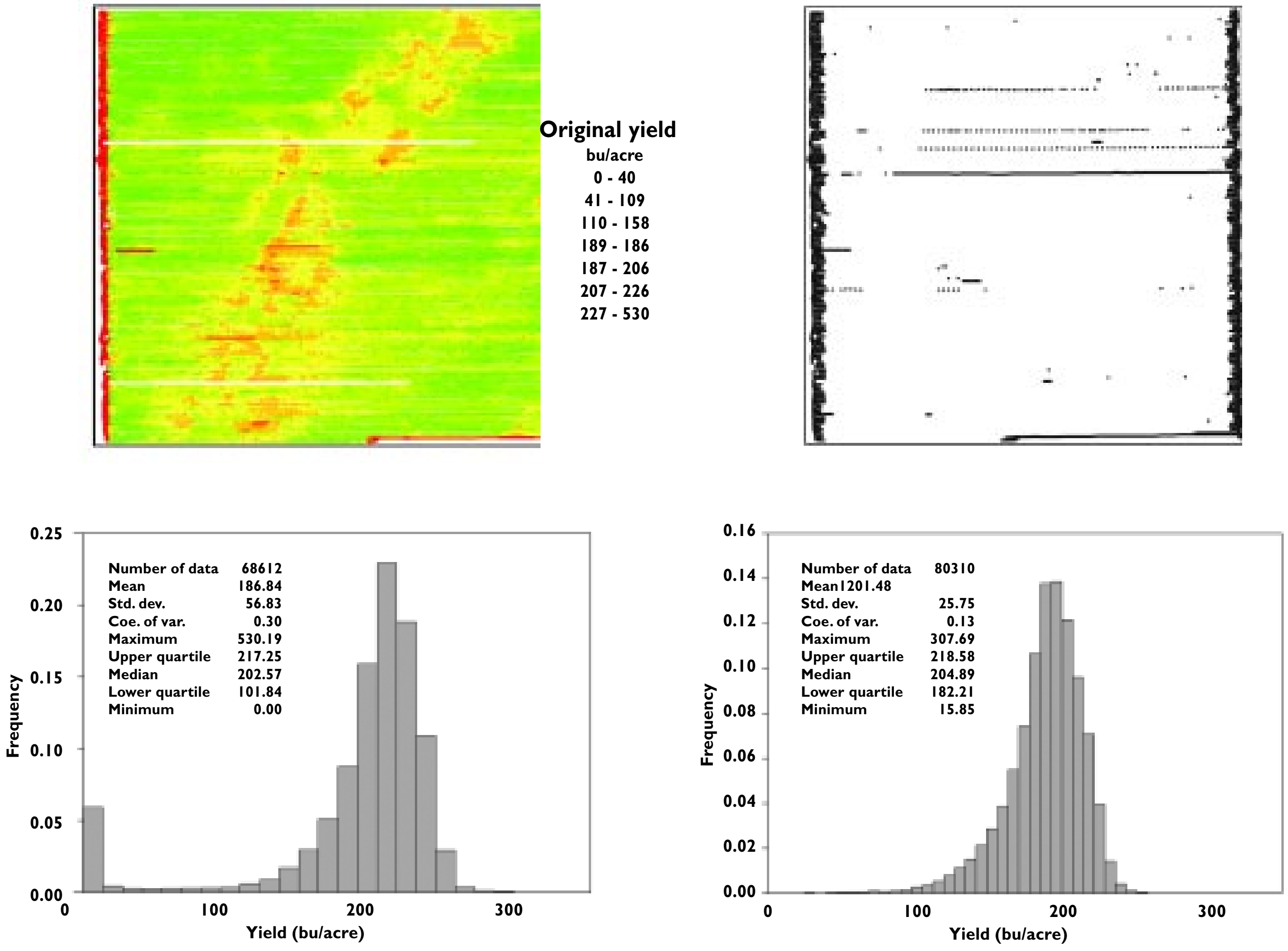

*Figure 3. Maps of yield monitor raw data* (top left) *and with all erroneous yield data removed* (top right) *and its effect on the frequency distribution of grain yield before* (bottom left) *and after screening* (bottom right).

Converting (averaging or interpolating) point yield data to a raster data layer is one of the principle data handling techniques available in numerous GIS packages. *Figure 4* shows a yield map (same as in *Figure 3*, but after data filtering) converted to a raster (grid) with 4 m (13.1ft) and 32 m (105 ft) resolution. In both cases, ordinary kriging interpolation was used (although many commercial packages simply average data points inside each raster). Note that interpolation to a coarse grid may result in loss of information and patterns that may not reflect the true spatial variation in crop yield. Larger grid sizes also result in more data averaging (smoothing), i.e., the overall range from minimum to maximum yields becomes narrower. The user can decide what size raster to use, based on the potential application for yield maps.

## Yield History Evaluation

Evaluating the temporal (year-to-year) variation of yield distribution within the field is an essential step in defining field areas with potentially high and low yields. Several approaches can be used to evaluate temporal effects on yield. One approach is to calculate the relative (normalized) yield for each point or grid cell. Normalized yield can be defined as the ratio of the actual yield to the field average:

$$Yield_{relative} = \frac{Yield_{actual}}{Average\ Yield_{year}} \tag{3}$$

When growing conditions in a field vary considerably, such as irrigated and dryland areas or different crops or varieties grown in different areas, normalization should be done separately for those areas, with the resulting relative yields recombined into one data file for the whole field. *Figure 5* shows a relative yield history for a field with corn (soybean in the southern half in 2000) grown using furrow-irrigation (until 2001) and center-pivot irrigation (in 2002).

Average (mean), standard deviation and the coefficient of variation are essential quantities that statistically describe many processes. The average relative yield over time is the sum of the relative yield values in each year divided by the number of years ($N$):

$$Average\ Yield_{relative} = \frac{Yield^{year1}_{relative} + Yield^{year2}_{relative} + \cdots + Yield^{year\ N}_{relative}}{N} \tag{4}$$

Similarly, standard deviation can be found as the sum of the squared difference between relative yields and their average divided by *N-1*:

$$StDevYield_{relative} = \sqrt{\frac{\left(Yield^{year1}_{relative} - AverageYield_{relative}\right)^2 + \cdots + \left(Yield^{year\ N}_{relative} - AverageYield_{relative}\right)^2}{N-1}} \tag{5}$$

The coefficient of variation (CV) expresses the relative yield variability among years and is calculated as:

$$CV[\%] = \frac{StDevYield_{relative}}{AverageYield_{relative}} \cdot 100 \tag{6}$$

Several simple approaches can be used to separate field locations that have high or low yield history from the rest of the field. For example, the difference between average relative yield and 1 (field average of relative yield) can be compared to the corresponding value of the standard deviation:

$$Class\ Yield = \begin{cases} Always\ High & if \quad AverageYield_{relative} - 1 > StDevYield_{relative} \\ Always\ Low & if \quad 1 - AverageYield_{relative} > StDevYield_{relative} \\ Variable\ and\ Average & \qquad Otherwise \end{cases} \tag{7}$$

In this case, field locations with highly variable yield (large standard deviation) and with yield close to the field average are considered as the same class (similar yield

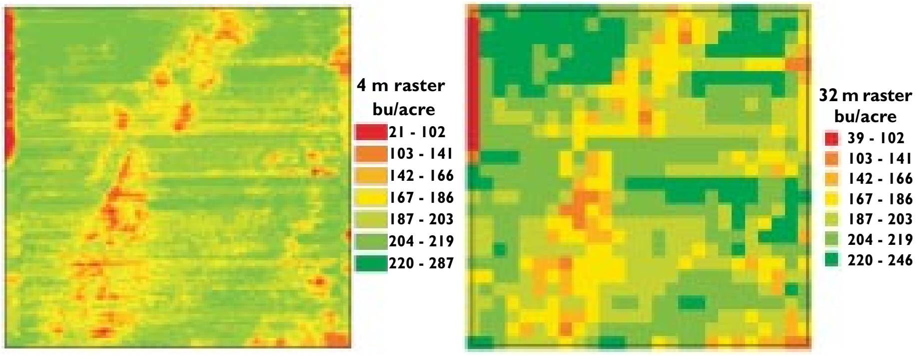

*Figure 4. Interpolated yield maps with 4 m raster (left) and 32 m raster (right).*

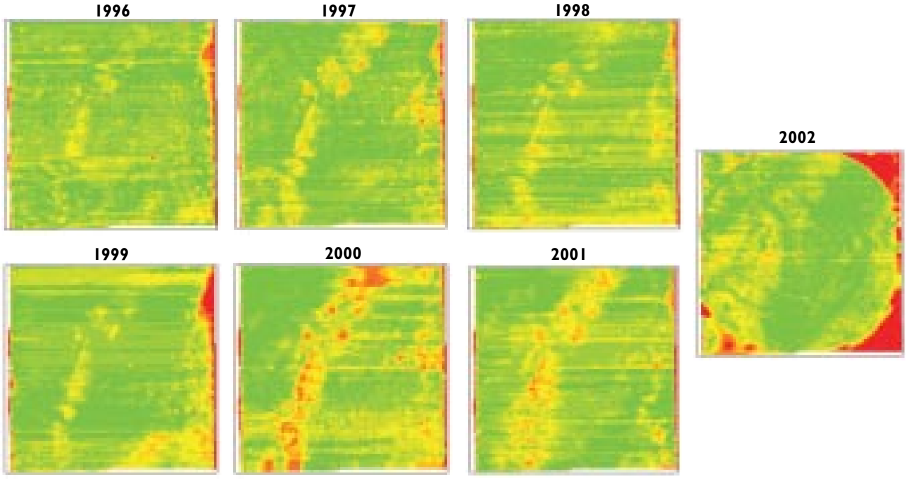

*Figure 5. Maps of relative yield of corn and soybean grown during a seven-year period (red indicates low-yielding areas and green indicates higher than average yields).*

potential). Alternatively, average relative yield *(Equation 4)* and its coefficient of variation (CV) *(Equation 6)* can be used to separate locations with stable and unstable yield using an arbitrary threshold value, such as a CV of 30 percent. Field locations with stable yield can be separated further into those with higher than average and lower than average yield:

$$\mathit{Class\ Yield} = \begin{cases} \mathit{High\ and\ Stable} & \mathit{if} \quad \mathit{AverageYield}_{relative} \geq \mathbf{1} \ \mathit{and}\ CV < 30\% \\ \mathit{Low\ and\ Stable} & \mathit{if} \quad \mathit{AverageYield}_{relative} < \mathbf{1} \ \mathit{and}\ CV < 30\% \\ \mathit{Unstable} & \mathit{if} \quad CV \geq 30\% \end{cases} \qquad (8)$$

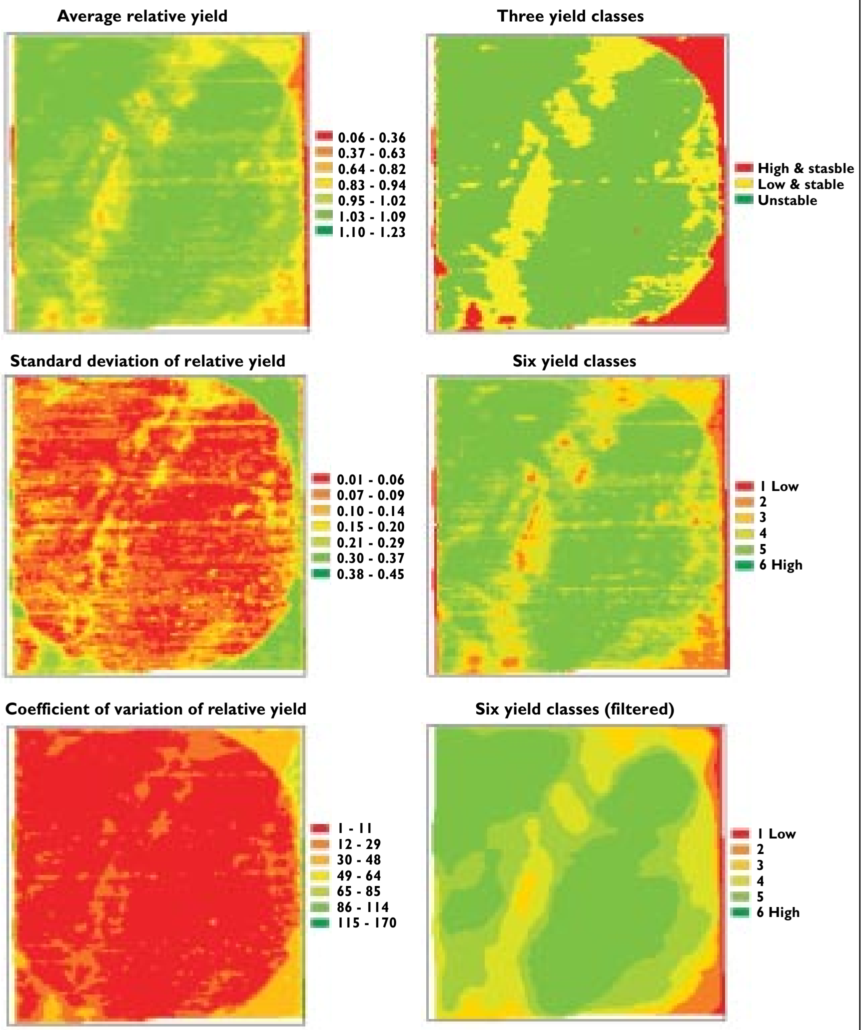

*Figure 6. Maps of average (mean), standard deviation (SD) and coefficient of variation (CV) of relative yield (left) compared to the yield classification maps obtained using* Equation 8 *(three classes), cluster analysis (six classes), and cluster analysis with contiguous zone filtering algorithm (right).*

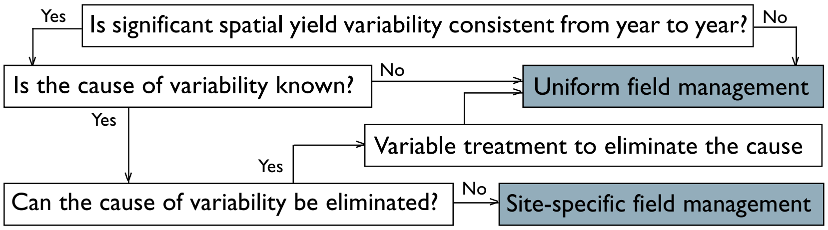

*Figure 7. Role of yield maps in decision-making strategy.*

Other, more complicated algorithms can separate yield into more than three classes to account for spatial structure (relationships with neighbors). *Figure 6* shows maps of average, standard deviation, and CV of relative yield measured during seven years *(Figure 5)*. Also, it illustrates maps of yield classes defined using *Equation 8* and a more advanced algorithm (cluster analysis and contiguous zones filtering). These maps can be used to determine field areas commonly described as management zones with different yield potential. The definition of management zones remains a researchable topic and commercial application of advanced algorithms in producer-oriented software is limited.

## Potential Applications

Yield maps represent the output of crop production. On one hand this information can be used to investigate the existence of spatially variable yield limiting factors. On the other hand, the yield history can be used to define spatially variable yield goals that may allow varying inputs according to expected field productivity.

*Figure 7* illustrates the process one might follow in deciding whether to invest in site-specific crop management, based on analysis of yield maps. If yield variability across the field cannot be explained by any spatially inconsistent field property, uniform management may be appropriate. Site-specific management becomes a promising strategy if yield patterns are consistent from year to year and can be correlated to one or more field properties (e.g. nutrient supply, topography, past management, etc.).

If the causes for yield variation are known and can be eliminated permanently, the entire area could be brought to similar growing conditions and managed uniformly thereafter. This concept was one of the earliest philosophies behind precision agriculture, but is likely only feasible for certain field properties. For example, variable rate liming can be used to correct acidic areas in a field. In this case, the yield map is used only to investigate whether low soil pH is a yield-limiting factor, and the soil map is used to prescribe variable application rates. Another example would be localized deep soil tillage to alleviate compaction in selected field areas.

Most yield limiting factors cannot be modified permanently through single measures because of economic or practical constraints. Consequently, site-specific crop management may be used to appropriately account for the existing spatial variability in attainable yield and/or soil properties.

Following are examples that show the potential use of the yield maps illustrated previously. *Figure 8* shows maps of various soil properties determined through a one-acre grid soil sampling and laboratory analysis (organic matter and phosphorus content) or using on-the-go sensors (elevation, slope, mechanical resistance, and apparent electrical conductivity). All of these maps have some resemblance to the general pattern of yield variability shown in *Figure 6*, suggesting that site-specific management of inputs, such as plant density and fertilizer, could increase the overall profitability of this field.

The examples illustrated in *Figure 9* show application maps that were developed using the yield classes *(Figure 6, bottom right)* and some of the soil properties shown in *Figure 8*. A variable rate nitrogen application map *(Figure 9, left)* was created by applying the UNL nitrogen algorithm for corn:

$$N_{rate}(lbs_N/acre)=35+1.2\cdot Yield_{goal}-8\cdot Soil_{NO_3\text{-}N}-0.14\cdot Yield_{goal}\cdot Soil_{OM} \quad (9)$$

The nitrogen fertilizer application rate ($N_{rate}$) is based on:

- $Yield_{goal}$, yield goal (bu/acre) calculated for each yield class *(Figure 6, bottom right)* as the average corn yield across seven years multiplied by 1.1 (to add 10 percent).
- $Soil_{OM}$, percent soil organic matter in the top eight inches of soil *(Figure 8, middle right)*.
- $Soil_{NO_3-N}$, field average nitrate nitrogen concentration in a four-foot deep soil profile (ppm).

As a result, prescribed nitrogen fertilizer rates were higher in areas with high yield potential, but lower in areas with low yield potential. Those included: headlands, dryland pivot corners, and the eroded ridge crossing the field.

According to current university recommendations, a phosphorus application map *(Figure 9, bottom)* was generated based on phosphorus content in the top

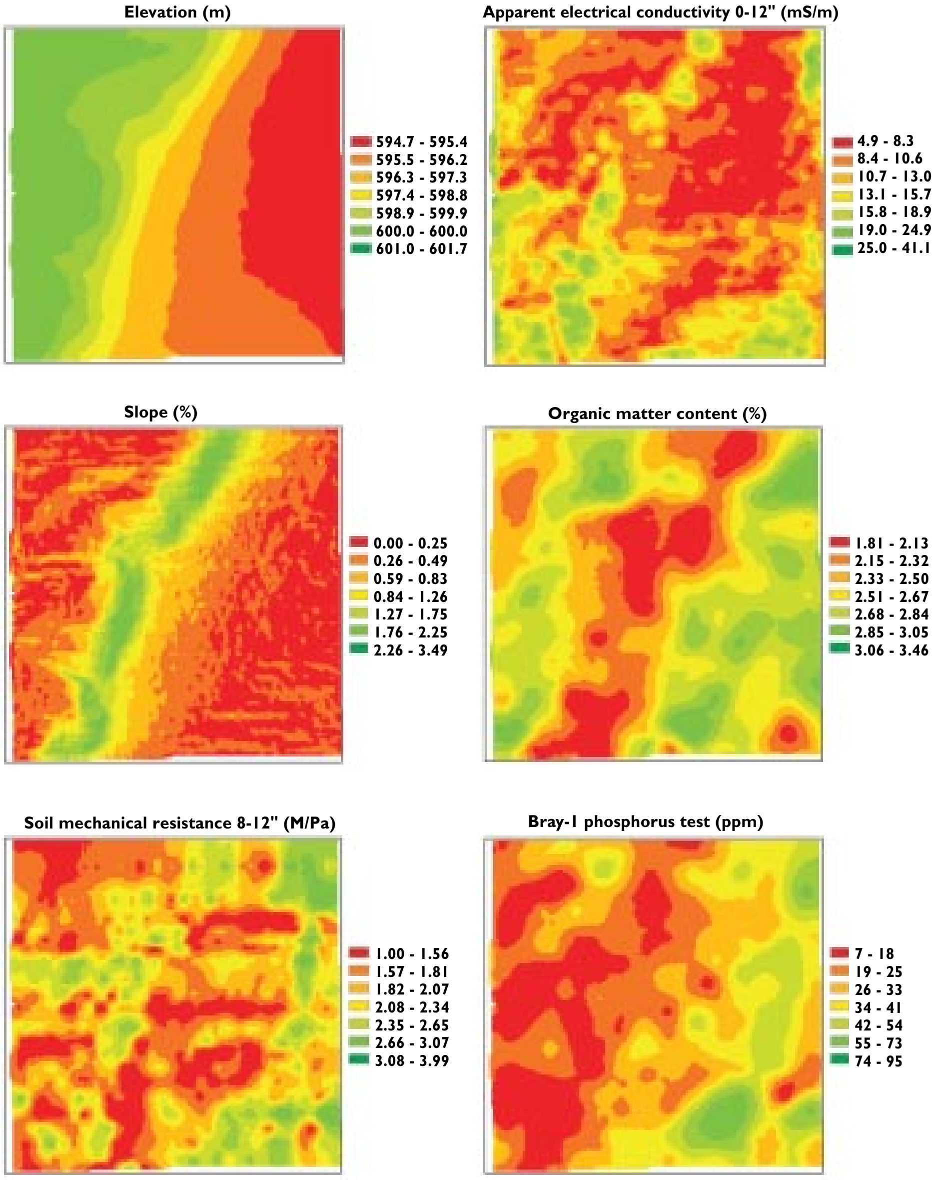

*Figure 8. Example of field maps that can be used to detect yield limiting factors and/or to prescribe variable rate application of agricultural inputs (red: low values; green: high values).*

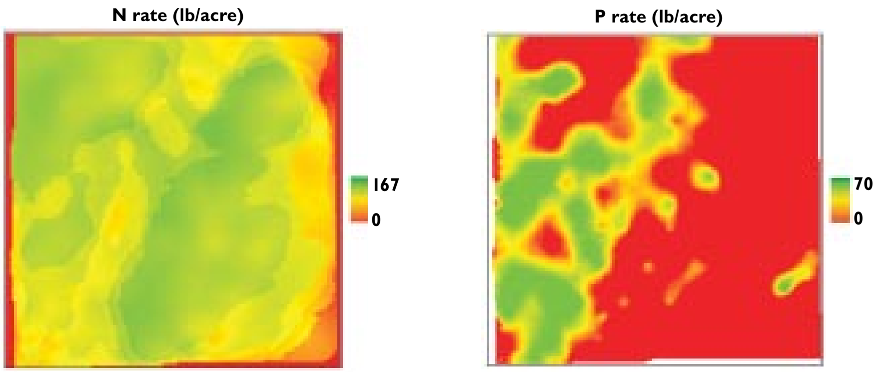

*Figure 9. Example of maps for variable rate application of agricultural inputs.*

eight inches of soil (ppm) measured using the Bray-1 soil test *(Figure 8, bottom right)*:

$$P_{rate}(lbs_{P_2O_2}/acre)=\begin{cases} 4\cdot(25-P_{Bray\text{-}1}) & \text{if } P_{Bray\text{-}1}<25 \\ 0 & \text{if } P_{Bray\text{-}1}\geq 25 \end{cases} \qquad (10)$$

As expected, the resulting phosphorous fertilizer application map ($P_{rate}$) reflects the map of soil test ($P_{Bray\text{-}1}$), but does not account for spatial variation in attainable yields within the field, which might be implemented through alternative fertilizer recommendations (currently under development).

## Summary

Yield maps are one of the most valuable sources of spatial data for precision agriculture. In developing these maps, it is essential to remove the data points that do not accurately represent the yield at a corresponding location. Map averaging or smoothing is usually done to aid data interpretation. A long yield history is essential to avoid drawing conclusions that are affected by the weather or other unpredictable factors during a particular year. Typically, at least five years of yield maps are desired. Processed yield maps can be used to investigate factors affecting the yield or to prescribe variable rate applications of agricultural inputs according to spatially variable yield goals (yield potential). Producers interested in precision farming should, however, always evaluate different management approaches to identify those that provide the greatest benefit at a particular site.

## Acknowledgement

Illustrations for this Extension Circular were developed based on spatial data collected from an irrigated corn field near Grand Island, Nebraska. The authors especially thank Arnie Hinkson, a Cairo, Nebraska farmer, for his ongoing contributions to UNL research and extension activities. He provided one of his fields with numerous collected spatial data layers for on-going research activities supported by the USDA-CSREES/NASA program on Application of Geospatial and Precision Technologies.

## Additional References

Arslan, S., and T.S. Colvin. 2002. Grain yield mapping: yield sensing, yield reconstruction, and errors. Precision Agriculture 3:135-154.

Dobermann, A., J.L. Ping, V.I. Adamchuk, G.C. Simbahan, and R.B. Ferguson. 2003. Classification of crop yield variability in irrigated production fields. Agronomy Journal 95:1105-1120.

Doerge, T.A. 1999. Yield map interpretation. Journal of Production Agriculture 12:54-61.

Lotz, L. 1997. Yield monitors and maps: making decisions. Extension Fact Sheet AEX -550-97, The Ohio State University, Columbus, Ohio.

Parsons, S., R. M. Strockland, R. Nielsen, and K. Morris. 2000. Yield monitoring and mapping. In: Precision Farming Profitability, ed. J. Lowenberg-DeBoer, Purdue University, West Lafayette, Indiana, pp. 62-74.

Ping, J.L., and A. Dobermann. 2003. Creating spatially contiguous yield classes for site-specific management. Agronomy Journal 95:1121-1131.

Taylor, R.K., G.J. Kluitenberg, M.D. Schrock, N. Zhang, J.P. Schmidt, and J.L. Havlin. 2001. Using yield monitor data to determine spatial crop production potential. Transactions of the ASAE 44(6):1409-1414.