EVALUATING TECHNIQUES FOR DEFINING MANAGEMENT ZONES IN THE SE COASTAL PLAIN.

Betsy K. Gerwig, E. John Sadler, and Dean E. Evans

Coastal Plains Soil, Water, and Plant Research Center
USDA-Agricultural Research Service
Florence, South Carolina

ABSTRACT

The SE US Coastal Plain has unique characteristics that require specialized techniques to explain yield variations and to develop management zones. This paper discusses several new methods to estimate yield variations for the development of management zones. Four techniques were developed based on the following: yield maps, black and white bare ground aerial photos, soil survey maps, and automated regular polygons. Two project fields were used for a detailed analysis of these techniques. Eight other fields were included for comparison. Results indicated that the amount of yield variation explained is related to the number of polygons used, regardless of the method used to generate the polygons. Therefore, an easily automated procedure based on regular polygons appears to be the least costly approach. Increasing the number of polygons per field reduces the size of each polygon; thus a limit will be reached at which regular polygons are not practical. Since the placement of regular polygons is arbitrary, the description of yield depends on where each polygon lands with respect to the yield variation. However, corn-based polygons showed more potential in explaining yield variation of other corn crops with fewer polygons (or fewer management zones). The prior-year corn yield maps were the preferred method of defining management zones, especially for corn followed by corn.

Keywords: precision farming, management zone, SE Coastal Plain

INTRODUCTION

Within site-specific farming, one goal is to determine the pattern of yield variation in the hope of developing effective management zones. In the Mid-West, much work has been done to explain yield variation within soil parameters and topography (Sudduth et al., 1996; Khakural et al., 1996). The yields of these areas are affected by topography, where in the SE Coastal Plain, topography has no consistent significant effect on yield as found in this research ($R^2 < 0.05$). Thus, this parameter can not be utilized in developing management zones. Other techniques for developing management zones must be evaluated for the SE Coastal Plain. The objectives of this research were to develop new techniques to estimate yield variations for the development of management zones and to critically evaluate the limitations of these techniques to describe yield.

METHODS AND MATERIALS

Project Description

A demonstration project titled "Management Practices to Reduce Nonpoint Source Pollution on a Watershed Basis", which is part of the Agricultural Systems for Environmental Quality (ASEQ) Project, was set up in Duplin County, North Carolina. The following cooperating agencies were participants in this USDA-CSREES funded project: Biological and Agricultural Engineering and Cooperative Extension Service, both of North Carolina State University; USDA-NRCS at the state, district, and county levels; USDA-ARS at Florence, SC; US Geological Survey; and several local farmer-cooperators. One objective of the ASEQ project was to improve and adopt precision farming as a best management practice. An overview of this objective was reported at this conference (Sadler et al., 2000).

To accomplish part of this objective, part of the demonstration project area, 172 ha in Duplin County, North Carolina, was chosen for this work. Ten fields were used to evaluate the techniques discussed below for developing management zones (Figure 1). Two project fields, F10 and F35, were chosen for detailed analysis. The eight other fields, F32, F33, F34, F37, F38, F39, F43, and F44, were used for further comparison with the project fields.

Two John Deere (Deere & Co., Moline, IL) combines, model 9500[1], were used to harvest corn, wheat and soybeans in 1997 and 1998. Both combines were equipped with the Green Star yield monitoring system with a GPS satellite link for differential correction. Data were collected for 26 yield events for all the fields listed above. Fields F34, F35, F37, and F38 had the following yield events: Corn 1997 and Corn 1998. Field F10 had the following yield events: Corn 1997, Wheat 1998, and Soybeans 1998. Fields F32, F33, F39, F43, and F44 had the following yield events: Wheat 1997, Soybeans 1997, and Corn 1998. All yield data were imported into AgLink Advanced (AGRIS Corp, Roswell, GA.) from the data cards. The yield data for each event were edited and field boundaries created. The

---

[1] Mention of trademarks is for information only. No endorsement implied by USDA-ARS or its cooperators.

Figure 1. Project area with the fields used in evaluation identified.

data were then exported as a shape file and imported into ArcInfo (ESRI, Redlands, CA) for analysis.

Soil Survey Polygons

Two conventional soil surveys were used to define management zones. The first was a digital version of the November 1996 Duplin County soil survey at a 1:24,000 scale. The second was a detailed soil survey completed by NRCS personnel at the Kenansville, NC, office. This survey resulted in approximately a 1:5000 scale. The detail survey was done only for F10 and F35, since it was both costly and time consuming.

Regular Polygons

Six regular polygon sets were created for F10 and F35. The polygon sizes were 50 x 50 m, 75 x 75 m, 100 x 100 m, 150 x 150 m, 200 x 200 m, and 300 x 300 m. The 100 x 100-m polygon set was based on the typical 1 ha (2.5 ac) soil sampling and extended in either direction for a range of sizes. All polygon sets had the same minimum x-y coordinates. Using ArcInfo, the RESAMPLE command was used to reduce the size of the largest polygon set to produce the other sets. For comparison, a 50 x 50-m polygon is 0.25 ha and a 300 x 300-m polygon is 9 ha. All polygons within a set, including partial or edge polygons, were used in the analyses. A 10 x 10 m and a 25 x 25 m polygons were included to further explain the limitations of describing yield with this technique.

Three sets of regular polygons associated with F35 were shifted to determine the effect of placement on the description capabilities. The polygon sets shifted were the 100 x 100 m, 200 x 200 m, and 300 x 300 m. The first two sets were moved up and/or right from point of origin by half the length of one polygon side, thus creating a total of 4 polygons sets for each size polygon. The 300 x 300 m polygon set was moved up and/or right from point of origin by 100 m twice to create a total of 9 polygon sets.

Aerial Photo Polygons

A black and white bare-ground aerial photo of the project area, taken on February 24, 1993, was obtained from USGS National Aerial Photography Program (NAPP). Using ArcInfo, the image was digitized and rectified at a 2-m pixel size. Using a printout of the project area, polygons were drawn on the printout to classify homogenous areas based on gray scale. These polygons were then digitized using Didger (Golden Software Inc, Golden, CO) and imported into ArcInfo. All polygons were cleaned and projected to the Universal Transverse Mercator (UTM) coordinate system. Two levels of detail were developed: one to identify major gray scale differences; the other to identify all visually discernable changes in gray scale. Both intensity levels were done to F10 and F35. In an attempt to test the repeatability of this skill, 8 other people repeated the process for these two fields with similar results. On the other fields, only the more detailed technique was applied.

Yield-based Polygons

The same technique and process described above was also applied to yield maps using the gray scale method. A comparison was done between the use of a color scale yield map and a gray scale yield map. There was no significant difference in the explanation of yield. Thus, the gray scale maps were used for two reasons. First, the eye was already trained to interpret gray scale images from the work done on the aerial photo. Second, there are only 256 shades of gray as compared to a multitude of shades in a color scale. The scale used for all yield maps ranged from white, equal to zero, to black, equal to maximum yield. The yield range for each crop was as follows: corn (0-10 Mg/ha), wheat (0-5 Mg/ha), and soybeans (0-3.8 Mg/ha). Yield-based polygons were evaluated against the yield event they were developed from (self) and the other yield events for that field. All results reflect data that does not include self-description except where noted.

Moving Window Smoothing Algorithm

A smoothing technique was developed to remove the noise in raw yield data. No commercial software was found to do this without first transforming the raw data to a grid, which itself imposes some smoothing. Thus, a Fortran program was written to interpolate a distance-weighted yield value for all input x-y coordinates. The weighted method used was a 2-D analog to the 1-D 1-3-5-3-1 moving average. The central point was weighted as 5 and points at a distance equal to the search radius were weighted 1, with linear interpolation between. Two search radii for neighbors were evaluated: 6-m and 12-m radius. This technique was applied to a subset of the original data set: F10 Corn 1998 and F35 Corn 1997. The input file included x-y and raw yield value. The output file contained the x-y (same as input file), raw yield value, and smoothed yield value.

Analysis

All data were imported into ArcInfo and projected to the UTM coordinate system. Data were extracted based on the x-y coordinates of a raw yield data file (point coverage). Each polygon set was extracted to a common file per field and event using the IDENTITY command. Then using the UNLOAD command, the data were converted to a text file to be used in SAS (SAS Institute, Cary, NC). In SAS, regression analysis was performed on yield vs. [polygon set], and the mean yield was found for each polygon within a set. The appropriate variable for comparison was R^2, which was compared to average polygon size and number of polygons per set. F10 and F35 yield events were compared to the following polygon types: the county soil survey, the detailed soil survey, the regular polygons, the photo-based polygons and the yield-based polygons. All other fields were compared to the same polygon sets except for the detailed soil survey and the regular polygons.

RESULTS AND DISCUSSION

Raw Yield Analysis

Figure 2 shows the comparison of the effectiveness in estimating yield with each polygon type associated with F10 and F35. As the number of polygons increased, the R^2 value increased. The regular polygons defined a nonlinear curve (trend line for regular polygons). Both the photo-based and soil survey polygon sets fall on this curve. Although most of the yield-based polygon sets fall on this curve, there is an indication of the potential for better explanation of yield by these types of polygons with fewer polygons. The primary difference between the 3 outlying yield polygon sets was that these describe corn yield by polygons based on corn yield.

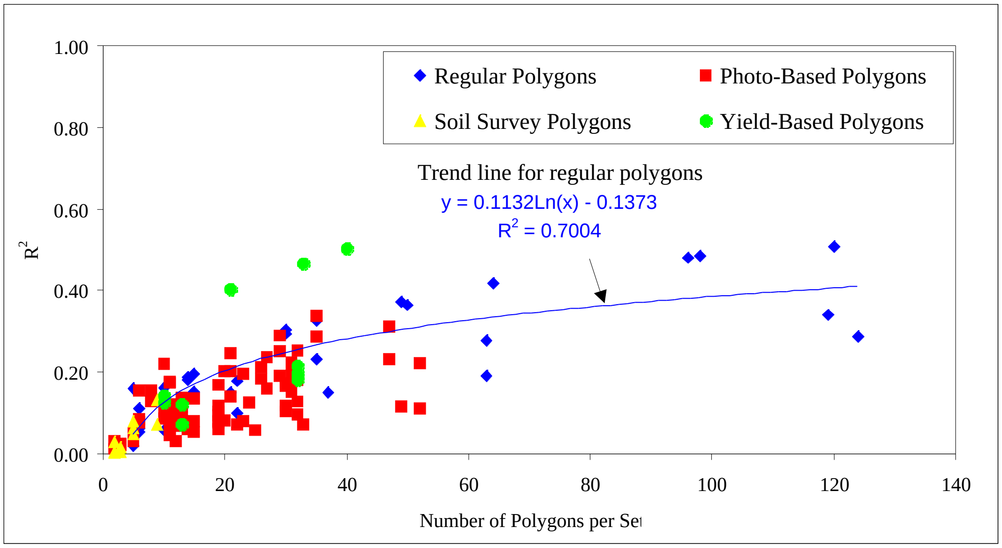

Figure 2. Yield explanation by polygon type for F10 and F35.

When identified by yield event, the comparison of R^2 to number of polygons indicates variation among crop types (Figure 3). By far, more variation was explained for corn yield. In the SE Coastal Plain, corn yield maps typically have distinct patterns that reoccur from year to year, whereas wheat or soybean yield maps show fewer patterns and are not as consistent from year to year (Sadler et al., 2000). This makes it difficult to define these patterns from year to year. This indicates that the development of management zones would better describe corn patterns.

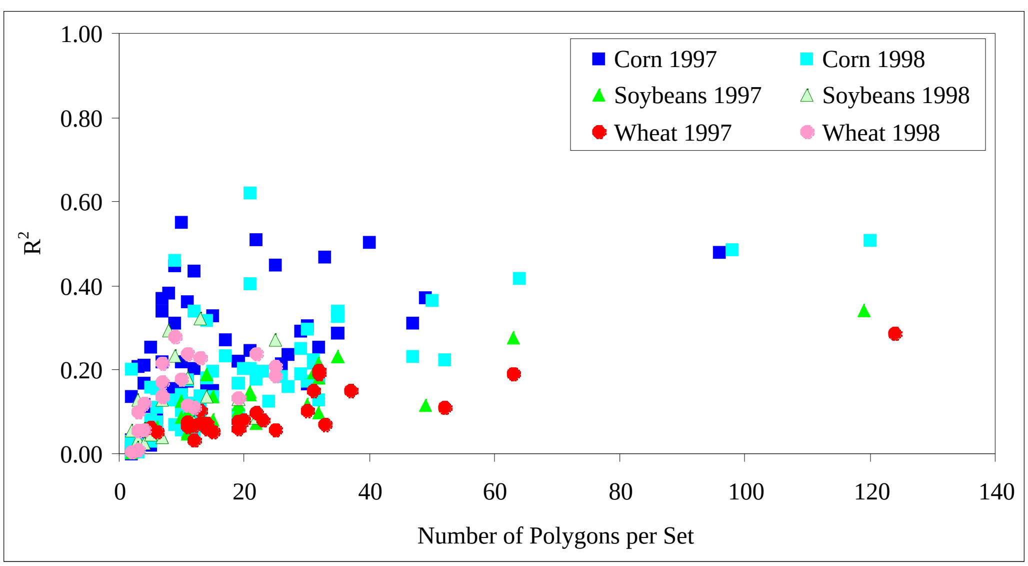

Figure 3. Yield explanation by yield event for all field data.

Smoothing Effect

Raw yield data contains random noise that cannot be corrected or removed during the editing process and is of particular interest when developing practical management zones. To eliminate some of this noise, a smoothing algorithm was applied to two yield events: F10 – Corn 1998 and F35 – Corn 1997. The original data were smoothed using a 6-m search radius and a 12-m search radius. The 6-m search radius had an average of 9 neighbors per yield point. The 12-m search radius had an average of 36 neighbors per yield point. Figure 4 shows the effect on the yield for F10 – Corn 1998. There is a noticeable reduction in the noise present in the yield maps from the raw to the 12-m search radius yield map.

As seen in Figure 5, an increase in the correlation of yield within polygons between the raw yield data and the smoothed yield data was apparent. The smoothing program removed approximately 10% of the noise with the 6-m search radius and 22% of the noise with the 12-m search radius. Also notice that the difference between the original and each search radius increased as the number of polygons increased. Thus, the smaller polygons were affected more by the noise than the larger polygons. Smoothing raw yield data will allow for more effective defining of management zones.

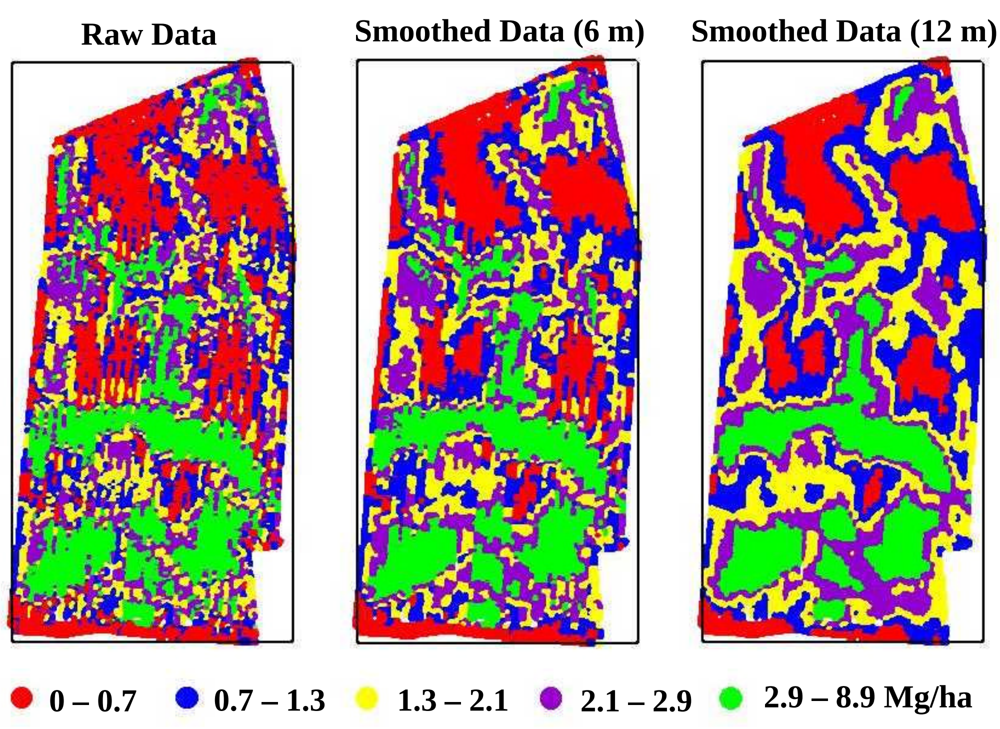

Figure 4. Effects of the smoothing algorithm on F10 – Corn 1998 yield data.

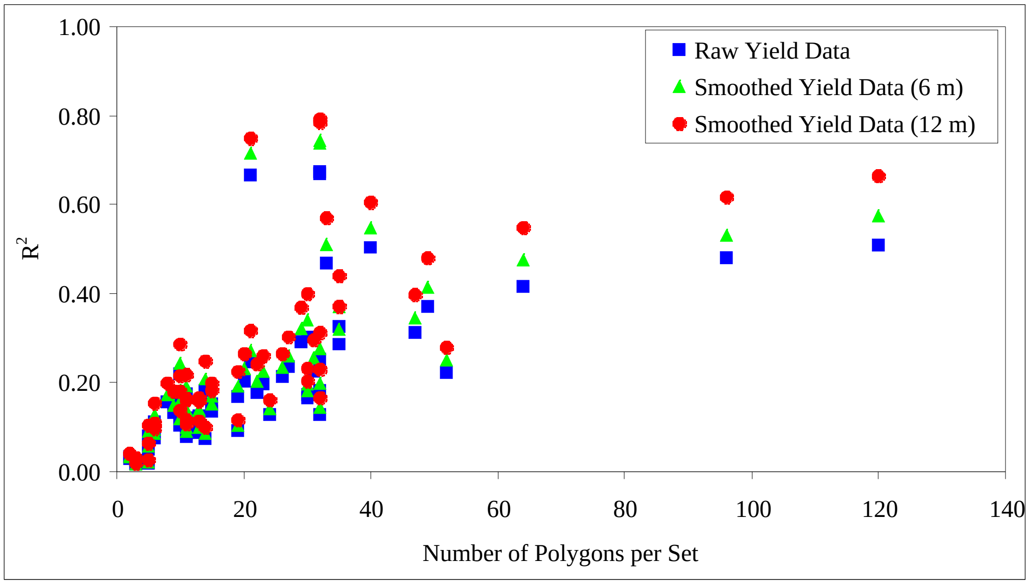

Figure 5. Comparison of original, 6-m, and 12-m smoothed yield data.

Regular Polygons

Since the process of developing regular polygons was completely automated, these polygons were much faster and easier to create. These polygons were also highly repeatable due to the advances in spatial software. A 10 x 10-m and a 25 x 25-m regular polygon set were included to determine the maximum description based on this polygon type. In Figure 6, the R^2 was compared to average polygon size. There was an inverse relationship between polygon size and R^2. This trend would continue until a 1:1 point-to-polygon ratio is reached. However, there was variability within each set of regular polygons, which can be primarily attributed to difference in the yield events. Using automation, this technique could be extended to a 1:1 ratio with perfect representation, but there would be no practical use for these data.

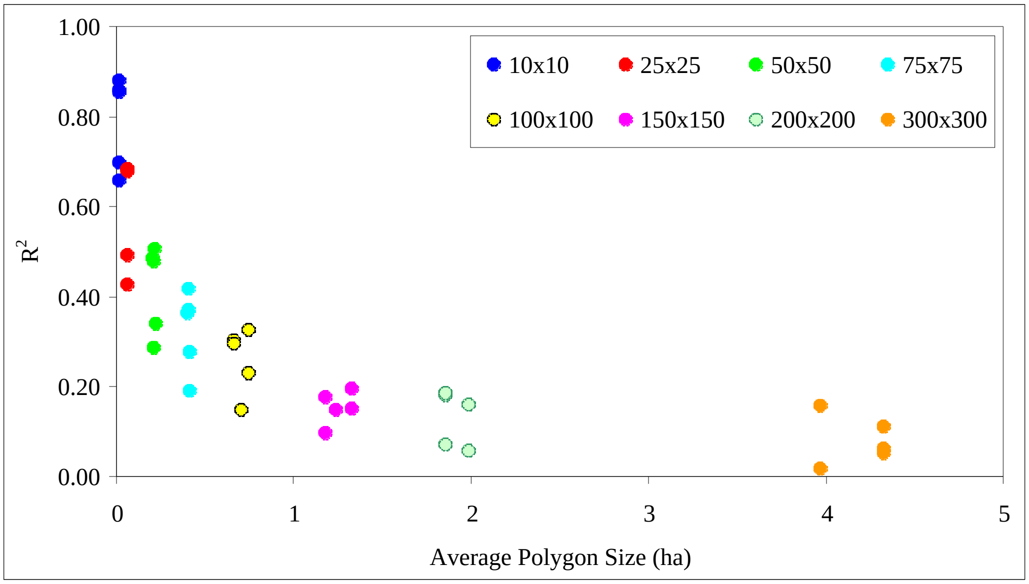

Figure 6. Yield explanation by regular polygons for F10 and F35.

To further investigate variability within each size regular polygon, three regular polygon sets associated with F35 were shifted. Results are shown in Figure 7. In addition to variation between yield events, shifting the polygons also resulted in variation within each event. As the polygon size decreases, the variability also decreases. Thus a smaller polygon size would result in less variability regardless of its spatial placement. Random placement of regular polygons can falsely inflate or deflate the apparent quality of yield estimation.

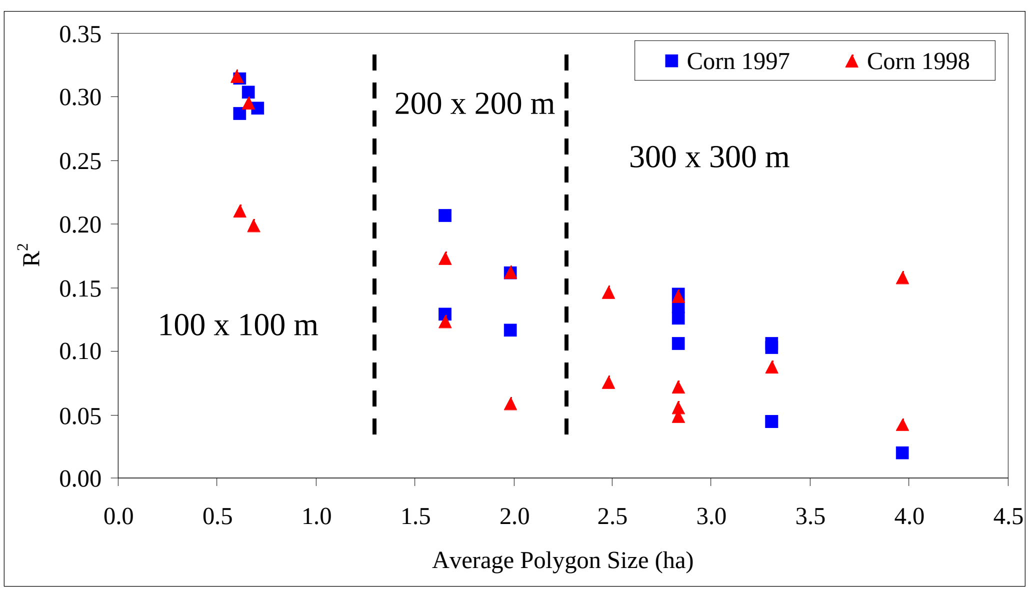

Figure 7. The effects of shifting regular polygons on F35.

Multiple Year Comparison

One use of management zones is to predict yield from year to year. This section discusses the comparison of the polygon sets for a corn-to-corn crop rotation. The fields included in the analysis were F34, F35, F37, and F38. Three polygon sets were used in this evaluation: 50 x 50-m regular polygons, photo-based polygons, and yield-based polygons. In Figure 8, the explanation of yield for some fields is shown. A definite trend is apparent between the mean yield of Corn 1997 and Corn 1998, regardless of polygon type. This indicates there were yield patterns that can be identified from year to year in corn. When the outliers noted were removed from the analysis, there was a 27% increase in the correlation from year to year. Also, there was correlation within each set of outliers.

As shown, there are 2 sets of outliers. Group 1 was identified to be edge polygons in F37 and F38, which were smaller than the average polygon size. Due to the small size of these polygons, any amount of variation within the polygon could effect the explanation of yield. These outliers could be attributed to field practices that were variable at the field edge, such as planting practices, fertilizer application, and harvest technique. These areas may also have poor physical and chemical properties. Outliers of this type should be included with a neighboring zone, but not included when defining the management zone.

Group 2 was identified as a "within-field" phenomenon attributed only to F35. Figure 9 identifies the area in question and some other possible contributors to the difference. This area was identified as primarily one soil type with a lower elevation. This was the only area where the yield did not follow the trend from 1997 to 1998. Areas, similar to this, should have more emphasis placed on factors other than yield, such as soil type and elevation.

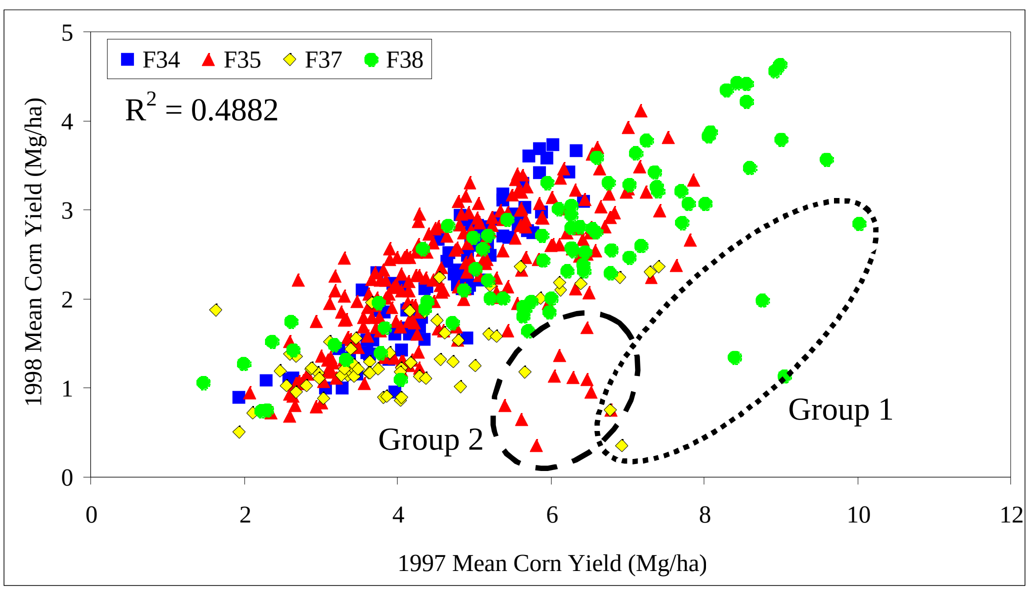

Figure 8. Multiple year comparison between Corn 1997 and Corn 1998.

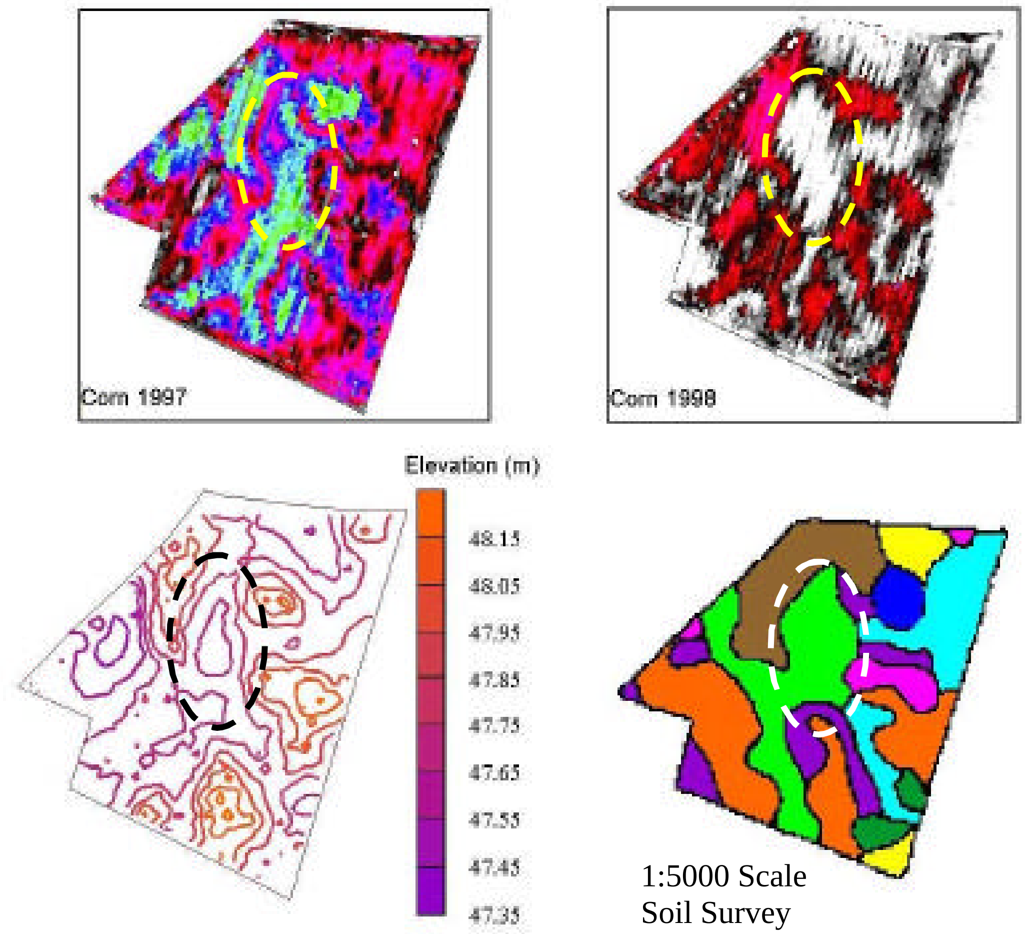

Figure 9. Evaluation of outliers in F35.

As seen below, and as tested, elevation had no correlation to yield ($R^2 < 0.05$). Elevation changes are so subtle that there is little effect on yield. Unlike the Mid-West region of the US, elevation is not a reliable source for explaining yield variation within the SE Coastal Plain.

Alternate Crop Comparison

Another alternative is to predict yield from crop to crop. This section discusses the comparison of the polygon sets with corn, wheat and soybean crops. The fields included in this analysis were F10, F32, F33, F39, F43, and F44. Alternate crop-based polygons had little correlation to other crop yields as evident in Table 1. Corn was better explained by wheat and soybean polygons. Corn was less effective in describing wheat and soybean yields. Yield maps of these crops show that corn and wheat yields have similar trends, for example: where corn yields are high, wheat yields are also high. However, corn and soybean yields show opposite trends, thus explaining why corn-based polygons do not explain soybean yield effectively. In general, zones based on one crop were not effective for other crops.

Table 1. Alternate crop comparison of R^2 values.

	Yield Event		
Polygon Sets	**Corn**	**Soybeans**	**Wheat**
Corn	--------	0.19134	0.17825
Soybeans	0.21335	--------	0.15002
Wheat	0.23799	0.11729	--------

Self-description

Any polygon with more than one data point will have variance caused by the differences among data within the polygon. As seen in the above discussion of local smoothing, 10% of variance in corn yield on F10 and F35 was attributed to variation within a 6-m radius, and 22% within a 12-m radius. This local variation will prevent any polygon set from explaining all yield variance in a field. This raises the question of how much can theoretically be explained. A yield-based polygon approach should be the best at explaining variance of the data set on which it was based. Thus R^2 for self-description can be useful in determining the maximum yield variation that could be explained by any other yield-based polygons.

For corn, self-describing yield-based polygons explained 62% of yield variation, with the remainder presumably acting at the local scale mentioned above (Table 2). The average of all other yield-based polygons (other year corn, any wheat, any soybeans) was about half. Similar results were obtained for wheat and soybean yields. Using a smoothing algorithm, an increase in the explanation of variance by self-describing and non-self-describing yield-based polygons would be expected.

Table 2. Self-description comparison of all yield-based polygons.

R^2	Corn	Soybeans	Wheat
Self-description	0.61991	0.28568	0.29590
Non-Self-description	0.31657	0.15716	0.16522

CONCLUSIONS

Overall, regular polygons and yield-based polygons showed potential in developing effective management zones. Regular polygons were enticing because the process of development was automated. However, the variation, due to spatial placement of polygons, may be extreme enough to inhibit the effectiveness of management zones. Being spatially dependent, yield-based polygons eliminate this factor from the effectiveness of management zones. Corn-based polygons were, by far, the most effective in describing other corn yield events. Management zones for corn would be most effective if based on a previous-year corn event.

Further research needs to be done to automate the development of yield-based polygons. These techniques should also be applied to maps of other spatial variables, such as nutrients, soil type, and organic matter. A comparison of photo-based polygons should be evaluated against soil survey maps. This would be beneficial in developing soil sampling plans. Computer simulations could be run to show possible improvements to yield by testing management options based on management zones developed using these methods

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the funding support given by USDA-CSREES project No. 95-4604. They also thank the three co-PI's on the project, Dr. Frank Humenik (NCSU Biological and Agricultural Engineering Dept., Raleigh, NC), Dr. Patrick Hunt (USDA-ARS, Florence, SC), and Mr. George Stem (USDA-NRCS, Raleigh, NC).

REFERENCES

Khakural, B. R., P. C. Robert, and D. J. Mulla. 1996. Relating corn/soybean yield to variability in soil and landscape characteristics. p. 117-128. *In* P. C. Robert et al. (ed.), Proc. of the 3rd Int. Conf. on Precision Agriculture. ASA, CSSA, and SSSA. Madison, WI.

Sadler, E. John, Betsy K. Gerwig, Joseph A. Millen, William Thomas, and Patrick Fussell. 2000. Experiences with Site-Specific Farming in a Demonstration Project in the SE Coastal Plain. Proc 5th International Conference on Precision Farming. July 16-19, 2000. Bloomington, MN. (this proceedings)

SAS. 1990. SAS Language: Reference, Version 6, First Edition. SAS Institute, Inc., Cary, NC. 1042 pp.

Sudduth, K. A., S. T. Drummond, S. J. Birrell, and N. R. Kitchen. 1996. Analysis of spatial factors influencing crop yield. p. 129-139. *In* P. C. Robert et al. (ed.), Proc. of the 3rd Int. Conf. on Precision Agriculture. ASA, CSSA, and SSSA. Madison, WI.

Back to the top of the manuscript.