Available online at www.sciencedirect.com
SCIENCE DIRECT
Computers and Electronics in Agriculture 46 (2005) 309–327
Computers and electronics in agriculture
www.elsevier.com/locate/compag

Identifying potential soybean management zones from multi-year yield data

Dan B. Jaynes*, Tom S. Colvin, Tom C. Kaspar

National Soil Tilth Laboratory, USDA-ARS, 2150 Pammel Dr., Ames, IA 50011, USA

Abstract

One approach for developing potential management zones for a variable-rate precision-agriculture system is to identify areas within a field exhibiting similar yield behavior. In this study, we applied cluster analysis of multi-year soybean (*Glycine max* [L.] Merr.) yield to partition a field into a few groups or clusters with similar temporal yield patterns and investigated the relationships between these yield clusters and the easily measured properties elevation (and the simple terrain attributes derived from elevation) and apparent soil electrical conductivity (EC_a). The analysis was applied to 5 years of soybean yield data collected from 224 plots arranged along eight transects spanning a 16-ha field. The partitioning phase of cluster analysis revealed that the 224 locations were best grouped into five clusters. These clusters were roughly aligned with landscape position and were characterized by the yield response to growing season precipitation above or below the 40-year average. Canonical discriminant functions constructed from the simple terrain attributes and EC_a predicted correct cluster membership for 80% of the plots. While not perfect, the discriminant functions were able to capture the major characteristics of the yield cluster distribution across the field, indicating that these easily measured variables are strongly related to soybean yield. Applying the functions with high-resolution terrain and EC_a attributes, we mapped soybean yield zones within the 16-ha field and an adjacent 16-ha field where multi-year yield data were not available. Cluster analysis of multi-year yield data and easily measured terrain and soil date may be useful in constructing effective management zones within fields and once developed can be applied to similar fields lacking detailed spatial yield data.
Published by Elsevier B.V.

Keywords: Yield zones; Clustering; Soybean; Yield variability; Terrain; Soil electrical conductivity

* Corresponding author. Tel.: +1 515 294 8243; fax: +1 515 294 8125.
E-mail address: jaynes@nstl.gov (D.B. Jaynes).

0168-1699/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.compag.2004.11.011

1. Introduction

One obstacle to applying precision agriculture practices to optimize crop production and environmental quality is identifying management zones—areas within the field by which inputs are managed to optimize economic return or environmental impact. Numerous methods have been used to construct potential management zones, but they essentially fall into three types. One approach is to use soil properties, e.g., soil series, water holding capacity, organic matter content, texture, depth to restricting layer, soil fertility test information—to construct potential management zones (Wibawa et al., 1993; Anderson and Bullock, 1998; Van Alphen and Stoorvogel, 2000; Ferguson et al., 2002). This approach assumes that the soil properties that control yield response to inputs are known and measurable. A major limitation of this approach is that extensive soil sampling, often in grid patterns, is required which can be costly and labor intensive.

The second approach is designed to circumvent the cost and time required to collect extensive soil data and instead uses surrogate variables to construct potential management zones. Typically, these surrogate variables include elevation and the simple terrain attributes that can be easily calculated from digital elevation data, such as slope and curvature, as these often account for much of the soil and yield variation observed within fields (Halvorson and Doll, 1991; Afyuni et al., 1993; Brubaker et al., 1993; Timlin et al., 1998; Yang et al., 1998; Kravchenko et al., 2000; Fraisse et al., 2001; Kaspar et al., 2003). Apparent soil electrical conductivity (EC_a) is another surrogate variable often used as it is has been found to be correlated with soil properties that affect yield (Rhoades and Corwin, 1981; Williams and Hoey, 1987; Kachanoski et al., 1988; McBride et al., 1990; Jaynes et al., 1995b) and has been found to be highly correlated with yield (Jaynes et al., 1995a). Apparent electrical conductivity can be measured for fields rapidly and easily using either electromagnetic induction instruments (Jaynes et al., 1993) or direct contacting equipment (Lund et al., 1999). While this approach circumvents the necessity to collect costly soil measurements, it assumes a strong correlation between the surrogate variables and the soil properties that control yield response to inputs.

The third approach makes no assumption regarding the interaction between yield and soil or landscape properties but instead uses the yield data directly to identify areas within a field where crops respond similarly over years (Lark and Stafford, 1997; Stafford et al., 1999; Lark, 2001). This approach makes the assumption that if yield patterns are similar over time then the areas must respond similarly to weather variability and management inputs and may function as effective management zones. Developing management zones from multi-year yield data is an intuitively attractive approach because it relies on direct observations to define yield zones rather than assuming a relationship between yield and soil or surrogate data. Of course, all of these approaches merely identify potential management zones. Further research is required to test whether or not the management zones identified by any approach do in fact function as effective management zones for the application of inputs.

Recently, Jaynes et al. (2003) took the third approach for developing potential management zones by applying unsupervised cluster analysis to multi-year corn yield data. They found that a 16-ha field could be partitioned into a few areas or zones where yield patterns over multiple years were similar. In the interpretation phase of cluster analysis, they found that these zones were largely determined by yield response to growing season precipitation