



Role of landscape and hydrologic attributes in developing and interpreting yield clusters

A. Bakhsh^{a,1}, R.S. Kanwar^{a,*}, R.W. Malone^b

^a Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, IA 50011, United States

^b National Soil Tilth Laboratory, USDA-ARS, 2150 Pammel Dr., Ames, IA 50011, United States

Abstract

Management of agricultural fields based on yield patterns may help farmers adopt environmentally friendly farming practices. Our objective was to investigate landscape and hydrologic attributes that affect spatial clusters of corn (*Zea mays* L.)–soybean (*Glycine max* L.) yields. The study was conducted at Iowa State University's northeastern research center near Nashua, Iowa, from 1993 to 1998. The yield data, normalized for annual climatic variability, were used in cluster and discriminant analysis, and the landscape and hydrologic data were overlain using ArcGIS software. Three clusters of low, medium and high categories were formed using 10 iterations with zero convergence options and satisfying the R^2 , pseudo F -statistic and cubic clustering criteria. The spatial clusters, however, varied greatly over space and time domain for the study period. The map overlay analysis using ArcGIS showed that high yield clusters were affected by soil and lower elevation levels in the below average precipitation year of 1994. The annual normalized subsurface drainage volume, nitrate leaching losses, soil type and topographic attributes of slope, aspect, and curvature were used in stepwise discriminant analysis to identify the variables significantly related to the clusters. Soil and topographic attributes of curvature and aspect contributed significantly in cluster formations for four of the six years at $P \leq 0.15$. The results suggest that cluster and discriminant analysis can be useful for identification of soil and topographic attributes affecting corn and soybean yield patterns, which can help in delineation of management zones for site specific management practices.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Yield; Soil; Topography; GIS; Subsurface drainage

1. Introduction

Crop yields are an outcome of the complex interaction among soil, topography, climate, and management practices and vary considerably within a field (Kanwar et al., 2005; Jaynes et al., 2005; Bakhsh and Kanwar, 2005; Reuter et al., 2005). Crop yield variability within a field may be due to intrinsic factors such as soil type, soil moisture, and nutrient potential as well as the extrinsic factors such as climate and management practices (Lamb et al., 1997; Bakhsh et al., 2000). Climatic variables are the most dominating factors that also cause temporal variability of crop yields within a field from year to year. The precipitation and temperature are the main two driving factors besides several others

in affecting the water and nutrient availability to plants (Mulla and Schepers, 1997; Bakhsh et al., 2002; Baker et al., 2005).

In addition to climatic effects, the soil moisture conditions within the soil profile are also affected by the landscape attributes and soil texture to move and retain water in the soil profile. Iqbal et al. (2005) reported that soil properties vary with the topographic settings and influence the redistribution of soil water content along the slope. Machado et al. (2002) reported that water, elevation and soil texture consistently influence the grain yields. Kravchenko and Bullock (2000) also examined the effects of topographic attributes and the derived hydrologic indices on variability in soil properties and crop yields. In other words soil characteristics and topography play an important role in varying the crop yield patterns within a field or watershed due to spatial variability effects (Afyuni et al., 1993; Fiez et al., 1994; Fraisse et al., 2001). The impact of topographic attributes on crop yield becomes more important especially for the soils having subsurface drainage 'tile' system (Bakhsh and Kanwar, 2004).

* Corresponding author. Tel.: +1 515 294 1434; fax: +1 515 294 6633.

E-mail address: rskanwar@iastate.edu (R.S. Kanwar).

¹ Visiting Scientist (on sabbatical leave, Associate Professor, Department of Irrigation and Drainage, University of Agriculture, Faisalabad, Pakistan).

Subsurface drainage is imperative for the soils in the mid-western parts of the United States to maintain their productivity potential by removing excess water from the root zone. The subsurface drainage systems, however, have also been reported to enhance the chemical transport from the bottom of the root zone to the edge of the field (Jaynes et al., 1999; Dinnes et al., 2002; Bakhsh and Kanwar, 2005). One such chemical, important for plant growth is nitrate-nitrogen ($\text{NO}_3\text{-N}$) which is soluble and non-adsorbent in nature, moves freely with the soil water and exits the system with subsurface drainage water (Baker et al., 1997). Leaching of $\text{NO}_3\text{-N}$ via subsurface drainage water has caused a serious environmental concern as well as economic loss to the farming community. The USEPA (1995) has identified the agriculture sector as one of the major contributors to soil and water pollution. The upper Mississippi river basin has been reported to be exporting 39% of the N delivered to the Gulf of Mexico which is suffering from the second largest hypoxic zone in the world (Alexander et al., 1995; USGS, 2005). Several studies have linked the hypoxic zones to the $\text{NO}_3\text{-N}$ loadings from the Mississippi river basin (Rabalais et al., 2002; Randall and Vetsch, 2005; Kanwar et al., 2005). Therefore to develop sustainable farming systems that minimize $\text{NO}_3\text{-N}$ leaching and maximize crop yield necessitate studying the role of soil and landscape features in transporting $\text{NO}_3\text{-N}$ from agricultural fields.

Leaching of $\text{NO}_3\text{-N}$ to subsurface drainage water depends on its concentrations in the soil profile at the time of water percolation below the root zone. The concentrations of $\text{NO}_3\text{-N}$ in the root zone have been reported to be affected by the controllable factors of management practices in addition to the uncontrollable factors of climatic variables (Kanwar et al., 1988; Bakhsh et al., 2002; Baker et al., 2005). Dinnes et al. (2002) concluded that N dynamics in agricultural fields in humid regions is affected by a number of factors including tillage, drainage, crop type, soil organic matter content and weather conditions. The interaction of climatic variables, soil and topography has caused the spatiotemporal variability among and within the fields despite having the same management treatments. Bakhsh and Kanwar (2005) have reported spatial variability effects on $\text{NO}_3\text{-N}$ leaching losses to subsurface drainage water for the field plots under the same management practices. They further recommended that site specific management of the soils needs to be made to reduce the offsite transport of $\text{NO}_3\text{-N}$ from agricultural fields. The spatial zones need to be delineated for site specific management practices.

One approach to address the spatial variability effects of the soils and topographic attributes can be grouping of the response variables into meaningful interpretable zones and using the map overlay capability of Geographic Information System (GIS) to study the spatial relationships (Bakhsh and Kanwar, 2004). GIS is a powerful tool for determining the integrated effects of the various soil and landscape data layers with crop yield patterns (Bakhsh et al., 2000). The effects of soil and topographic attributes on yield variation can be perceived better when data layers of these attributes are overlaid (Silva and Alexandre, 2005).

Quantification of the soil and landscape effects on crop yield is essential to agricultural decision making. For example, adoption of conservation practices requires weighing the benefits to the environment with the affects on crop production under the

site specific soil and climatic conditions. Malone et al. (2006, in this issue of *Geoderma*) quantified corn yield based on variable climate and N applications. As discussed above, soil and landscape attributes significantly affect crop yields, therefore, agricultural decision making tools should account for these effects.

Crop yield has also been considered a good indicator for delineation of stable management zones (Bakhsh et al., 2000). Cluster and discriminant analyses have been used to classify the crop yield data into meaningful groups and study the contribution of various soil and landscape attributes in discriminating these clusters, respectively. These approaches have been used in different disciplines by several researchers (Al-Sulaimi et al., 1997; Bakhsh and Kanwar, 2004; Kaspar et al., 2004; Jaynes et al., 2005; King et al., 2005). Bakhsh and Kanwar (2005) applied an integrated approach to study the offsite transport of $\text{NO}_3\text{-N}$ leaching losses by developing spatial clusters and seeking their spatial relationships with the soil and topographic attributes. They reported that spatial $\text{NO}_3\text{-N}$ leaching losses clusters were affected by the interaction of soil type and elevation levels. No study, however, has been conducted to investigate the spatial yield clusters for the soils having subsurface drainage and determining their relationships with the subsurface drainage flows, $\text{NO}_3\text{-N}$ leaching losses and the landscape attributes. The hypothesis of this study was that soil and landscape attributes can affect the crop yield patterns and have spatial relationships with the yield clusters. The specific objectives of the study were:

- Delineate spatial zones of corn–soybean yields using cluster analysis.
- Identify landscape and hydrologic attributes that contributed significantly in discriminating yield clusters using discriminant analysis.
- Integrate and overlay GIS data layers of the identified landscape and hydrologic attributes on crop yield clusters to establish cause–effect relationships.

2. Materials and methods

The field experimental data on corn–soybean yields from 1993 to 1998 were collected from 36 plots, each 0.4 ha in size, at Iowa State University's north-eastern research center near Nashua, Iowa. The soils at the site are Floyd loam (fine-loamy, mixed, mesic, Aquic Hapludolls), Kenyon silty-clay loam (fine loamy, mixed mesic, Typic Hapludolls) and Readlyn loam (fine-loamy, mixed, mesic, Aquic Hapludolls) (Kanwar et al., 1997). These soils are moderately well to poorly drained, lie over loamy glacial till with 3 to 4% organic matter and belong to the Kenyon–Clyde–Floyd soil association. Generally, pre-Illinoian glacial till overlies a carbonate aquifer but in some areas bed rock is almost near the surface.

These soils have seasonally fluctuating water tables and therefore need subsurface drainage to maintain the productivity level. Subsurface drains were installed at the site in 1979 at 1.2 m depth with 28.5 m spacing. Each plot has a separate drainage sump with flow meter for recording the subsurface drain flows and collecting composite water samples for chemical analysis. Drainage water sampling frequency averaged three times a week. Subsurface drain water samples were collected and

Table 1
Descriptive statistics for corn and soybean yields (Mg/ha) data from 36 experimental plots for six years (1993–98) at Nashua, Iowa

Corn (21 plots)	1993	1994	1995	1996	1997	1998	Average (1993–98)
Mean	5.48	7.32	5.57	8.44	9.19	8.96	7.49
Median	5.10	7.58	5.84	8.62	9.52	9.42	7.90
Standard deviation	1.61	0.96	0.87	0.78	0.91	1.23	0.86
Skewness	0.09	−0.67	−0.80	−1.02	−0.99	−1.13	−0.22
Kurtosis	−1.21	−0.32	0.47	0.59	0.73	−1.44	−1.53
Minimum	2.90	5.14	3.45	6.47	6.79	6.83	6.22
Maximum	7.97	8.60	6.74	9.42	10.24	10.96	8.74
Interquartile range	2.40	1.57	1.01	0.81	1.28	1.74	1.54
C.V. (%)	29.31	13.15	15.66	9.26	9.85	13.81	11.43
Normality test ^a							
Shapiro–Wilk	0.27	0.24	0.19	0.08	0.04	0.25	0.02
Kolmogorov–Smirnov	>0.15	>0.15	>0.15	>0.15	0.07	0.12	<0.01
Soybean (15 plots)							
Mean	2.64	3.31	3.21	3.99	3.66	3.99	3.47
Median	2.63	3.30	3.21	3.94	3.66	4.08	3.45
Standard deviation	0.08	0.23	0.08	0.15	0.14	0.24	0.10
Skewness	0.43	−0.17	0.92	0.33	−1.04	−0.64	−0.49
Kurtosis	0.83	−0.54	0.99	−1.26	3.60	−0.94	0.15
Minimum	2.48	2.88	3.11	3.79	3.29	3.53	3.25
Maximum	2.83	3.70	3.41	4.23	3.91	4.26	3.63
Interquartile range	0.10	0.36	0.13	0.23	0.14	0.40	0.15
C.V. (%)	3.31	7.03	2.57	3.70	3.73	5.94	2.91
Normality test ^a							
Shapiro–Wilk	0.89	0.97	0.22	0.25	0.10	0.09	0.63
Kolmogorov–Smirnov	>0.15	>0.15	0.13	>0.15	>0.15	0.09	>0.15
Normalized yields							
Mean	0.14	−0.08	−0.14	−0.03	−0.15	−0.25	−0.08
Median	0.00	0.00	0.00	0.00	0.00	0.00	0.02
Standard deviation	0.85	0.67	0.91	0.88	0.90	0.66	0.49
Skewness	0.32	−0.26	0.08	−0.90	−0.81	−0.29	−0.74
Kurtosis	0.85	−0.34	1.02	1.18	2.90	−1.06	−0.37
Minimum	−1.84	−1.55	−2.37	−2.65	−3.08	−1.49	−1.16
Maximum	2.54	1.31	2.35	1.35	2.08	0.89	0.62
Interquartile range	1.03	0.96	1.07	1.07	1.00	1.06	0.48
Normality test ^a							
Shapiro–Wilk	0.90	0.85	0.76	0.11	0.03	0.05	<0.01
Kolmogorov–Smirnov	>0.15	>0.15	>0.15	>0.15	>0.15	<0.01	<0.01
Growing season rainfall (mm)							
Rainfall	1030	750	800	680	750	980	771 ^b

C.V. = coefficient of variation in percent.

^a p -values < 0.05 = significant.

^b 30-yr growing season average.

refrigerated until chemical analyses were made. Further details on the subsurface drainage system installed at the site can be found in Kanwar et al. (1999).

The site has a total of 36,58.5 m by 67 m plots, with fully documented tillage and cropping records for the past 25 years.

They had been managed in a randomized complete block design with four tillage treatments and three crop rotations (chisel plow, ridge-tillage, moldboard plow, no-till, continuous corn, and both phases of a corn/soybean rotation) since 1978 (Bjorneberg et al., 1996). In 1993, new N-management practices were implemented for only two tillage systems of chisel plow and no-till with manure application, preplant single N as well as late spring soil test based N applications beneath continuous corn and corn–soybean rotation systems. More details on the experimental treatments applied at the site can be found in Bakhsh et al. (2002).

The same varieties of corn (Golden Harvest 2343) and soybean (Sands of Iowa) were grown in these plots during the six-year (1993–1998) study (Bakhsh et al., 2002). Corn, whether fertilized with preplant single or late-spring N applications, was planted in 750-mm rows into a seedbed prepared by fall chiseling and field cultivating in the spring. Soybean was drilled in 200-mm rows directly into corn stover from the previous year, and no fertilizer was applied. A single UAN application of 110 kg-N ha^{−1} was made before planting with a spoke injector, which injected UAN at about 200-mm intervals, 250-mm from corn rows (Baker et al., 1989). The late-spring UAN applications were determined based on the late-spring NO₃-N test (LSNT) developed for Iowa soils (Blackmer et al., 1989), in addition to 30 kg-N ha^{−1} applied with the corn planter (Bjorneberg et al., 1998). Based on LSNT, UAN was injected to increase the soil NO₃-N concentrations in the top 300 mm of the soil profile to 25 mg kg^{−1}. The amount of N applied for the LSNT treatment varied from 93 to 195 kg-N ha^{−1} during the 6-year period of this study. Corn and soybean yields were measured from each plot using a modified commercial combine with all stover left in the field. The yield data collected for each plot was tested for moisture content and adjusted to a constant water content of 155 g kg^{−1} (15.5%) for corn and 130 g kg^{−1} (13%) for soybean.

2.1. Data normalization

Descriptive statistics were calculated for corn and soybean yields separately using PROC UNIVARIATE procedure in SAS 9.1 (SAS, 2003) with the option of NORMAL PLOT to check the normal distribution of the data. The normal distribution of the data was checked using Shapiro–Wilk and Kolmogorov–Smirnov tests (Table 1). The data, however, did not violate the assumption of normality ($P < 0.05$) for corn and soybean and were normalized using the robust standardization technique (Jaynes et al., 2003). Also, to study the spatial and temporal patterns of data and compare it over the six years period (1993 to 1998), corn and soybean yield data were normalized separately on yearly basis. The yield data for each plot was normalized using the following relationship (Bakhsh and Kanwar, 2005):

$$z_j = \frac{y_j - y_j'}{S_j}$$

Where z_j is the normalized yield for each plot for j th year, y_j is the yield of each plot for j th year, y_j' is the median of yield for j th year and S_j is the estimate of yield variation for j th year. Similar approaches have been used by Jaynes and Hunsaker

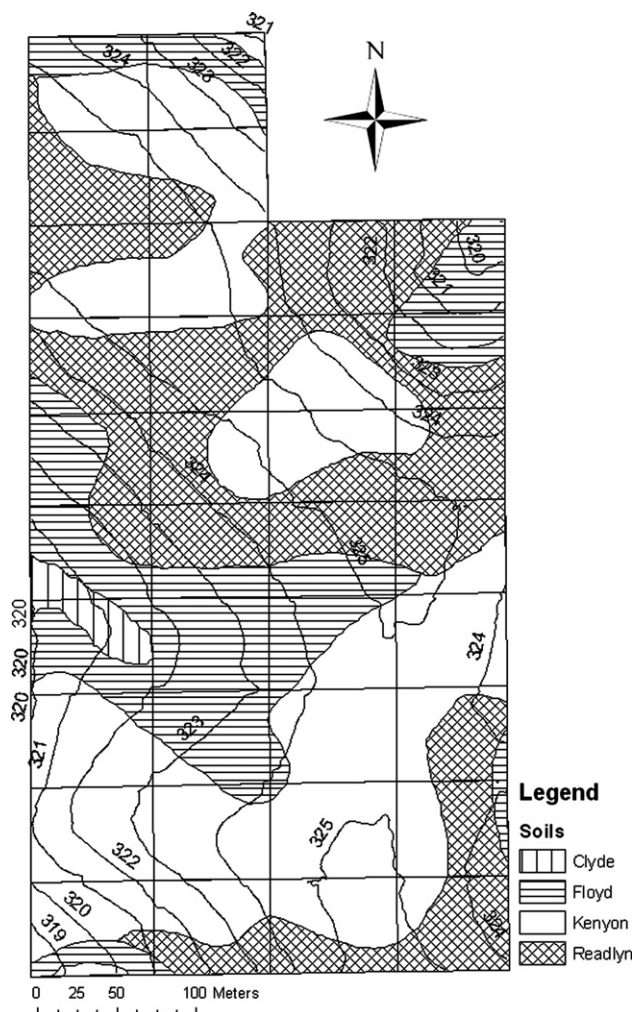


Fig. 1. Layout, soil type and topography of the study area. (contours show elevation in meters above mean sea level).

(1989) and Colvin et al. (1997) for normalization of their data. Median estimates were used for y_j' (Cressie, 1993) and interquartile range was used as an estimate of s_j . As robust estimators, the median and interquartile range reduce the impact of outliers and non-normality on the calculation of z_j (Colvin et al., 1997). Both the normalized data of corn and soybean had median of '0' and interquartile range of '1' and were combined. The normalized yield data were used during all the statistical as well as GIS analysis. Similarly, data on subsurface drainage volume and nitrate-nitrogen losses to subsurface drainage water were also normalized on annual basis for each plot. Normalization of data was necessary to remove the climatic effects on all the variables of yield, subsurface drainage and $\text{NO}_3\text{-N}$ leaching losses which were used in the subsequent cluster and discriminant analyses (Bakhsh et al., 2000).

2.2. Cluster analysis

PROC FASTCLUS procedure (SAS, 2003) with 10 iterations and zero convergence criteria was used to develop clusters based on normalized yield data from 36 plots for each year as well as over the six years. The FASTCLUS procedure in SAS combines

an effective method for finding initial clusters with a standard iterative algorithm for minimizing the sum-of-squared distances from the cluster means. The result is an efficient procedure for disjoint clustering of data sets with the option of specifying the number of clusters. This procedure uses nearest centroid sorting method and each observation is assigned to the nearest seed to form temporary clusters. The initial seeds are replaced by the cluster means and the process is repeated until no further changes occur in the clusters. More details on cluster algorithm can be found in SAS documentation (SAS, 2003). Different numbers of clusters were tried to get the best results and finally three clusters were selected based on the evaluation criteria of cluster's formation i.e., pseudo F -statistic; R -square and cubic clustering criterion (SAS, 2003). The cluster output was plotted showing the exact location of each member in the cluster when coordinate data of each member were given in the data set used in SAS.

2.3. Discriminant analysis

Stepwise discriminant procedure of PROC STEPDISC (SAS, 2003) was used to determine the contribution of the annual normalized subsurface drainage and $\text{NO}_3\text{-N}$ leaching loss, and the landscape attributes of elevation, slope, aspect, curvature and the

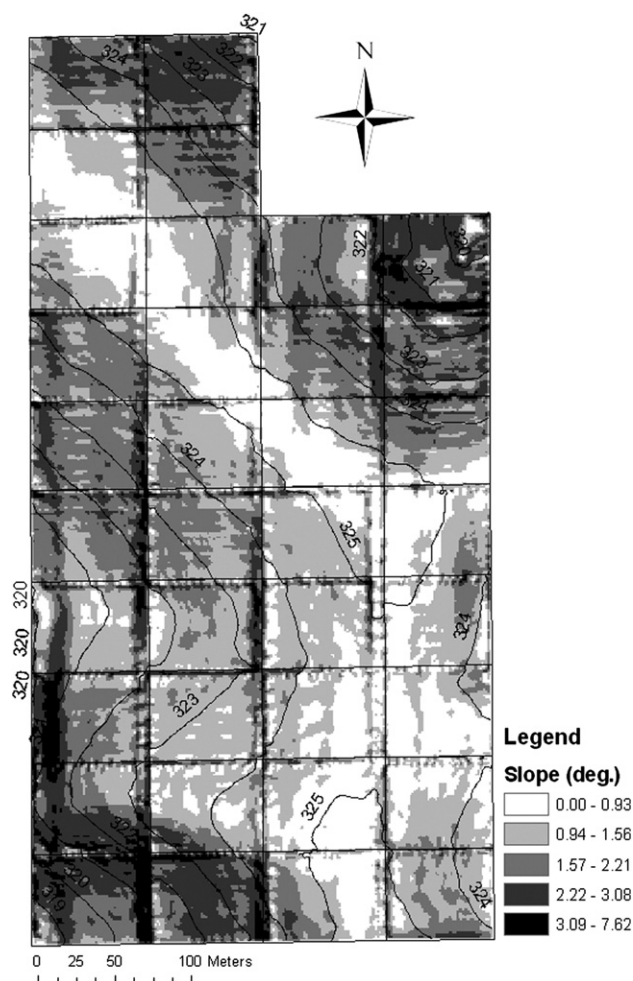


Fig. 2. Slope surface of the study area.

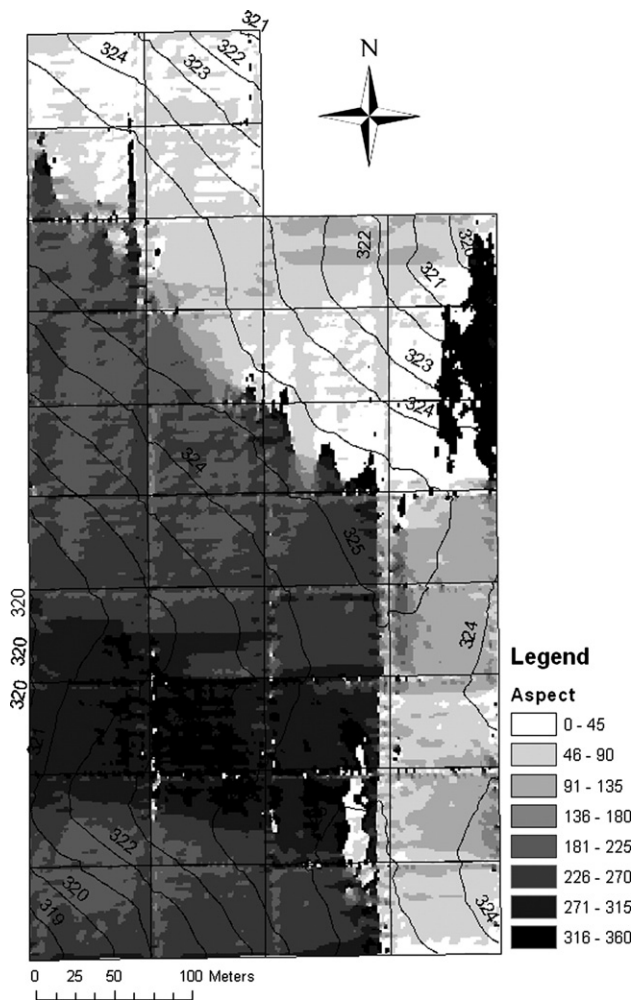


Fig. 3. Aspect surface of the study area in standard eight dimensions.

dominant soil type for each plot, in the formation of the normalized yield clusters. STEPDISC procedure performs a stepwise discriminant analysis to select a subset of the quantitative variables for use in discriminating these clusters. STEPDISC procedure is similar to the stepwise regression analysis and incorporates the variable into the model based on its significance level ($P < 0.15$). This procedure uses forward selection, backward elimination or stepwise selection technique. This procedure is useful to establish the cause–effect relationships for cluster spatial occurrence and may be helpful for applying management practices based on zones defined by the clusters. Verification of cluster formation was made using discriminant procedure i.e. PROC DISCRIM (SAS, 2000) to assess how accurately the clusters can be predicted. This procedure derives canonical variables (linear combinations of the quantitative variables) that summarize variations between clusters. More details about these procedures can be found in SAS documentation (SAS, 2003).

2.4. GIS data layers

The soil type map of the site was digitized after scanning the soil map and georeferencing it using ArcGIS (9.1), software. The ground control points selected from GPS survey were used

during the first order georeferencing procedure. A detailed topographic survey of the site was carried out using a Trimble 4700 survey grade dual frequency GPS receiver, with accuracy of 20 mm in x , y and z directions. The GPS unit was installed on the moving vehicle and about 6695 measurements were taken during side by side navigation tracks on all the 36 field plots. Elevation data along with coordinates were recorded. These elevation data were used to build a digital terrain model (DTM) for the site. ArcGIS Spatial Analyst extension was used to create the elevation surface with kriging procedure. A spherical model was used to interpolate the elevation surface (Bakhsh and Kanwar, 2004). From this DTM, slope, aspect and curvature data layers were derived for the site using Surface module in Spatial Analyst. Zonal function in the spatial analyst tool was used to compute an output table for the average elevation, slope, aspect and curvature data layers for 36 plots, which were used during discriminant analysis. The output of cluster analysis was used as input in ArcGIS to show cluster surface for low, medium, and high areas of yield and was overlaid by the soil and landscape attributes selected during stepwise discriminant analysis.

The attributes of various data layers which were identified as the variables that contributed significantly in discriminating the

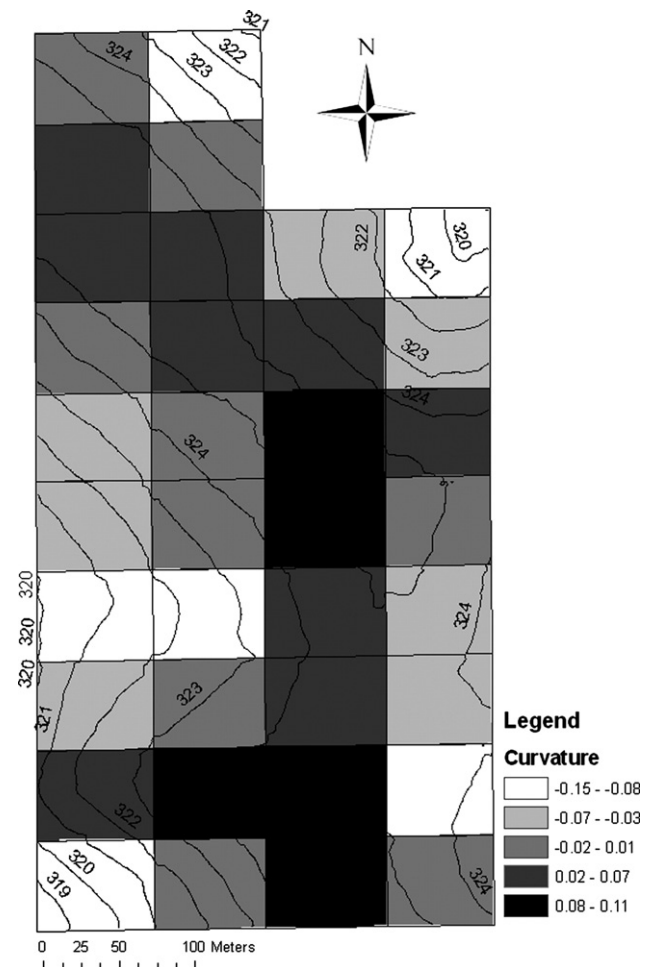


Fig. 4. Curvature surface of the study area showing concave (–ve) and convex (+ve) landforms.

Table 2
Cluster summary for normalized yield data

Years	Clusters	Frequency	Means	Standard deviation	Proportion	Overall R^2 / F -statistic / C.C.C. ^a
1993	1	9	−0.88	0.44	0.25	0.90 /
	2	25	0.36	0.48	0.70	40.14 /
	3	2	2.10	0.62	0.05	−5.32
1994	1	10	−0.95	0.32	0.27	0.90 /
	2	18	0.03	0.24	0.50	89.97 /
	3	8	0.77	0.28	0.22	−2.32
1995	1	10	−1.22	0.51	0.28	0.90 /
	2	19	−0.04	0.30	0.53	63.83 /
	3	7	1.12	0.58	0.19	−3.65
1996	1	5	−1.66	0.63	0.14	0.90 /
	2	17	−0.21	0.28	0.47	89.29 /
	3	14	0.77	0.32	0.38	−2.33
1997	1	2	−2.61	0.67	0.06	0.90 /
	2	20	−0.44	0.40	0.56	53.57 /
	3	14	0.62	0.49	0.38	−4.30
1998	1	14	−0.98	0.29	0.39	0.90 /
	2	16	0.08	0.19	0.44	120.51 /
	3	6	0.59	0.22	0.16	−1.02
Overall (1993–98)	1	8	−0.89	0.69	0.22	0.35 /
	2	9	0.09	0.59	0.25	14.62 /
	3	19	0.18	0.55	0.53	4.42

^a Cubic clustering criterion.

clusters were classified into three categories (Bakhsh et al., 2000). The range of these categories was defined as low (values < 1 standard deviation); medium (values between ± 1 standard deviation) and high (values > 1 standard deviation).

3. Results and discussion

3.1. Corn–soybean yields

Average corn grain yields varied greatly from 5.48 Mg ha^{−1} in 1993 to 9.19 Mg ha^{−1} in 1997 which show the effect of growing season rainfall variability because 1993 was a wet year (Table 1). The year 1993 had a rainfall of 1030 mm compared with 750, 800, 680, 750 and 980 mm for 1994, 1995, 1996, 1997, and 1998, respectively. This compares to the normal growing season rainfall of 771 mm for the study area (Voy, 1995). The amount and distribution of rainfall during the growing season are important for the soils having subsurface drainage systems. Heavy rainfall in 1993 flushed NO₃-N from the root zone and resulted in lower yields for that year (Bakhsh et al., 2002). Minimum corn grain yield of 2.90 Mg ha^{−1} was also observed in 1993 while maximum of 10.96 Mg ha^{−1} was found in 1998. Similarly hail damage in 1995 also resulted in lower average yields of 5.57 Mg ha^{−1} for that year (Bjorneberg et al., 1998). This shows that climate had a major impact on the corn grain yields. The CV (coefficient of variation) for corn grain yield ranged from 9% in 1996 to the maximum value of 29% in 1993. The CV for soybean yield ranged from 2.6% in 1995 to 7.0% in 1994. The minimum average soybean yield of 2.64 Mg ha^{−1} was observed in 1993 and maximum of 3.99 Mg ha^{−1} in 1996 and 1999 (Table 1). The yield data for both corn and soybean were also found to be skewed negatively and was

normalized. The normalized data had median of zero and interquartile range of about 1 because median and interquartile range were considered as the robust estimates of mean and standard deviation to reduce the impact of outliers (Jaynes and Colvin, 1997; Bakhsh et al., 2000).

3.2. GIS data layers of soil and topographic attributes

The study area had four soils of Clyde, Floyd, Kenyon, and Readlyn with percent area of 2, 21, 43, and 34 as determined during GIS analysis. Readlyn soils lie mostly at the higher elevation levels compared with the other soils. At the base of the slope, Floyd soil was present in the north east of the area and Clyde soil in the central west (Fig. 1). Topography plays an important role when fine soil particles are eroded from one place and are deposited at another, changing the texture for that zone (Li and Lindstrom, 2001). The spatial occurrence of these soils shows relationship with the topographic attributes as reported by other researchers (Conacher and Dalrymple, 1977; Si and Farrell, 2004; Jaynes et al., 2005; Iqbal et al., 2005).

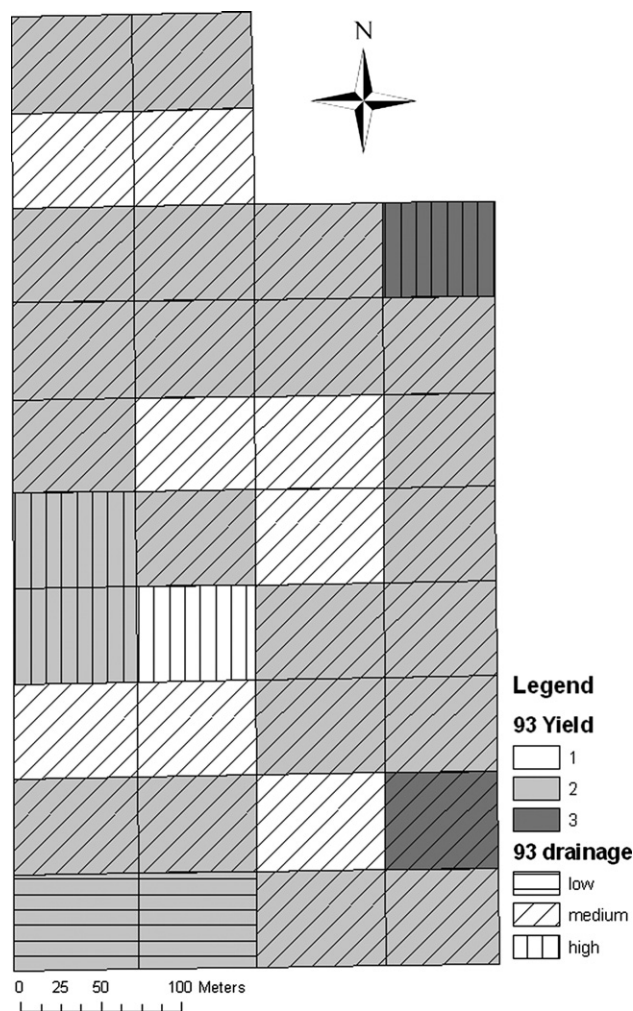


Fig. 5. Map overlay of the normalized yield clusters and subsurface drainage layers for 1993. (normalized yield cluster means for 1993: 1 = −0.88; 2 = 0.36; 3 = 2.10); normalized drainage (low < 1 SD; medium = ± 1 SD; high > 1 SD); SD = standard deviation = 1.09.

The topographic attributes of elevation, slope, aspect and curvatures showed distinct patterns for the northeast area and the southwest because a higher elevation ridge runs diagonally from south-east to north-west directions of the site. Minimum elevations were observed in the northeast and in the central west and south of the area (Fig. 1). Minimum slope or flat area was seen at the ridge running diagonally. Steeper slopes were observed at the base of the slope area in the north-east and south-west corners where elevation contour lines were relatively narrow (Fig. 2). A clear direction for the aspect attribute was seen for the area in the south west compared with that in the north east (Fig. 3). Aspect shows the slope direction with reference to true north of zero degrees. Another derived topographic attribute is the curvature. The negative curvature values correspond to concave surfaces and positive curvature shows convex surfaces. The curvature layer showed negative average value showing the overall concave surface. Greater crop yields have been reported at the concave position compared with the convex landforms (Kravchenko and Bullock, 2000; Pennock et al., 2001) (Fig. 4). Curvature surface has been reported to affect the infiltration and overland flow processes. These attributes affect the soil water movement, evapo-



Fig. 6. Map overlay of the normalized yield clusters and soils layers for 1994. (normalized yield cluster means for 1994:1 = -0.95; 2 = 0.03; 3 = 0.77).

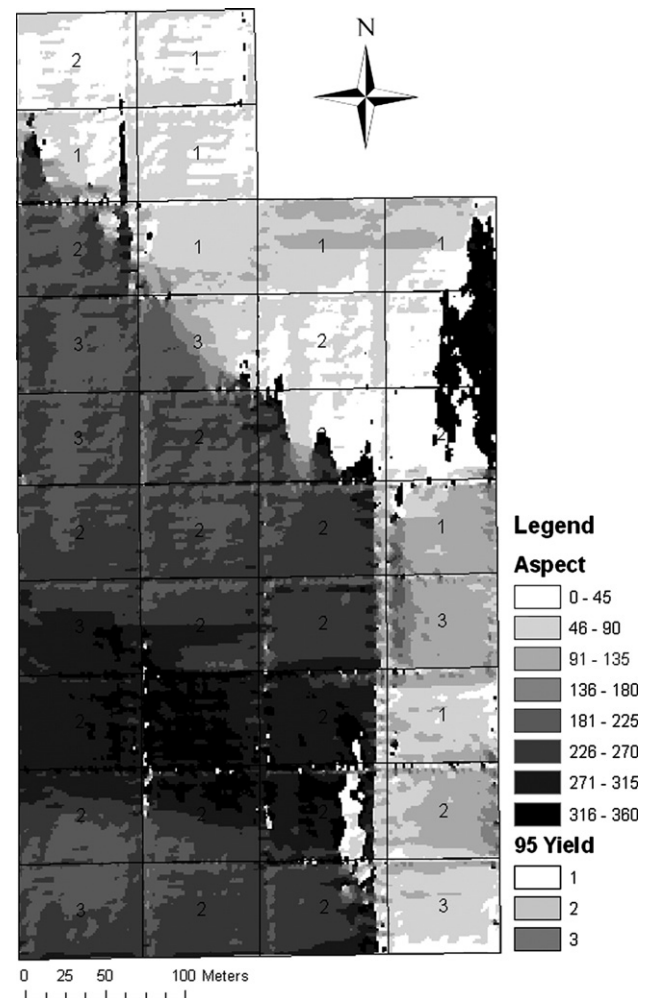


Fig. 7. Map overlay of the normalized yield clusters and aspect layers for 1995. (normalized yield cluster means for 1995:1 = -1.22; 2 = -0.04; 3 = 1.12).

transpiration rates and crop yields (Sinai et al., 1981; Changere and Lal, 1997; Kravchenko et al., 2002; Iqbal et al., 2005).

The interaction of the soils and topographic attributes with the precipitation induces variability in the soil moisture availability to the crops and results in yield variability across the fields (Mulla and Schepers, 1997). The normalized yield data were grouped into three classes using cluster analysis to study their relationships with the soil and topographic attributes effects.

3.3. Cluster and discriminant analysis

Cluster analysis was carried out to generate three spatial clusters of 1, 2, and 3, categorized as low, medium and high using normalized yields data on yearly as well as over the six years (1993–98). The cluster means, standard deviation, proportion and the cluster evaluation criteria are given in Table 2. The cluster formation was found to be satisfactory because of their indicators of R^2 , F -statistic and cubic clustering criteria (Table 2). These three cluster classes resemble three categories of values below 1 SD (standard deviation), mean \pm 1 SD, and values above 1 SD, as maximum number of the members was found in the medium cluster. The pattern of these clusters varied greatly over the years

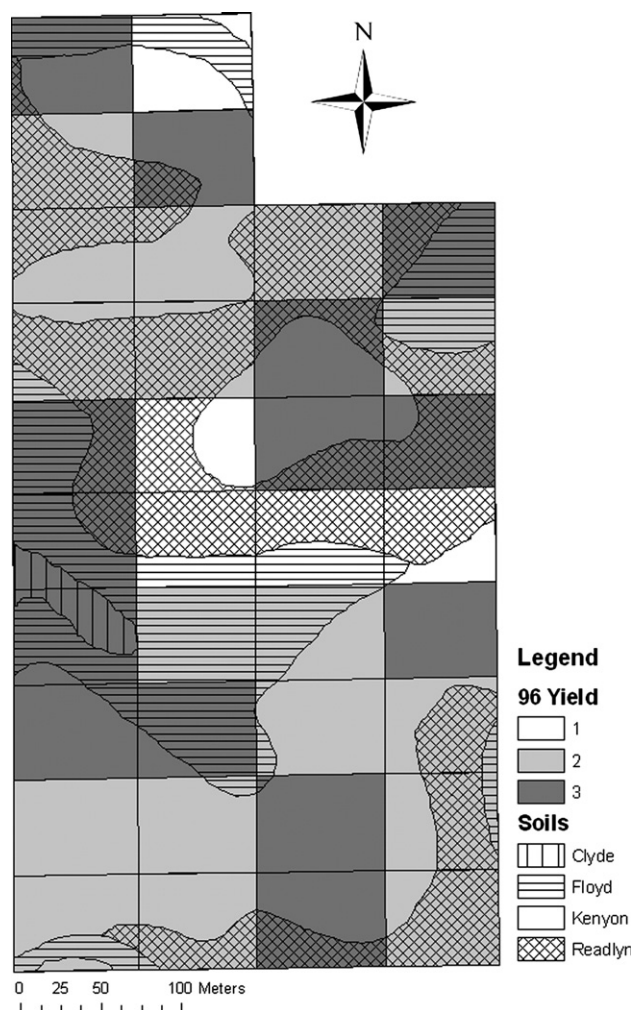


Fig. 8. Map overlay of the normalized yield clusters and soils layers for 1996. (normalized yield cluster means for 1996: 1 = -1.66; 2 = -0.21; 3 = 0.77).

as shown in Figs. 5–10 which can be associated with the yield variability over the years. Similarly the number of members assigned to cluster 1, 2 and 3 also varied over the years (Table 2).

The area represented by the low cluster varied from 14% in 1996 to the maximum value of 39% in 1998 (Table 2). Similarly medium and high yield clusters captured the proportional area in the range of 44% in 1998 to 70% in 1993, and 5% in 1993 to 38% in 1997, respectively. This shows the degree of variability in the cluster formation over the years. The generalized squared distance (Euclidean distance) was also found to vary over the years (Table 3). Clustering is done on the basis of Euclidean distances computed during analysis. Observations that are closer to each other are assigned to the same cluster, while observations that are far apart are assigned to different clusters (SAS, 2003). Maximum distance between clusters was observed between the low and high categories for most of the years (Table 3). The accuracy of cluster membership prediction was verified using the confusion matrix (Table 4). The error rate varied from 0.17 for 1993 to 0.34 for 1995. Clusters 1 (low), 2 (medium), and 3 (high) were predicted most accurately in 1997, 1994, and 1993, respectively. The overall analysis of six years of data showed an average error rate of 0.47, which can be

associated with the degree of variability and inconsistency of the yield clusters over the years.

The STEPDISC procedure identified the normalized drainage of 1993 and curvature as the most significant variables for the 1993 yield cluster formations (Table 5). The soil and curvature contributed significantly for the 1994 clusters. The aspect and soil variables contributed significantly for the 1995 and 1996 clusters, respectively. Aspect plays an important role when soil water content is limiting and may induce soil water stress (Iqbal et al., 2005). No variable qualified to enter the model for the years 1997 and 1998 at probability value of ≤ 0.15 as a threshold value (SAS, 2003). The topographic attributes of curvature and aspect have been reported to be important in crop production systems (Changere and Lal, 1997; Kravchenko and Bullock, 2000; Kravchenko et al., 2002; Jiang and Thelen, 2004). The overall analysis of the combined data for six years showed a distinct pattern from the rest of the years because of variability and no variable was identified as the one contributing significantly in cluster formation. The integration of these clusters as GIS data layers showed only one plot which

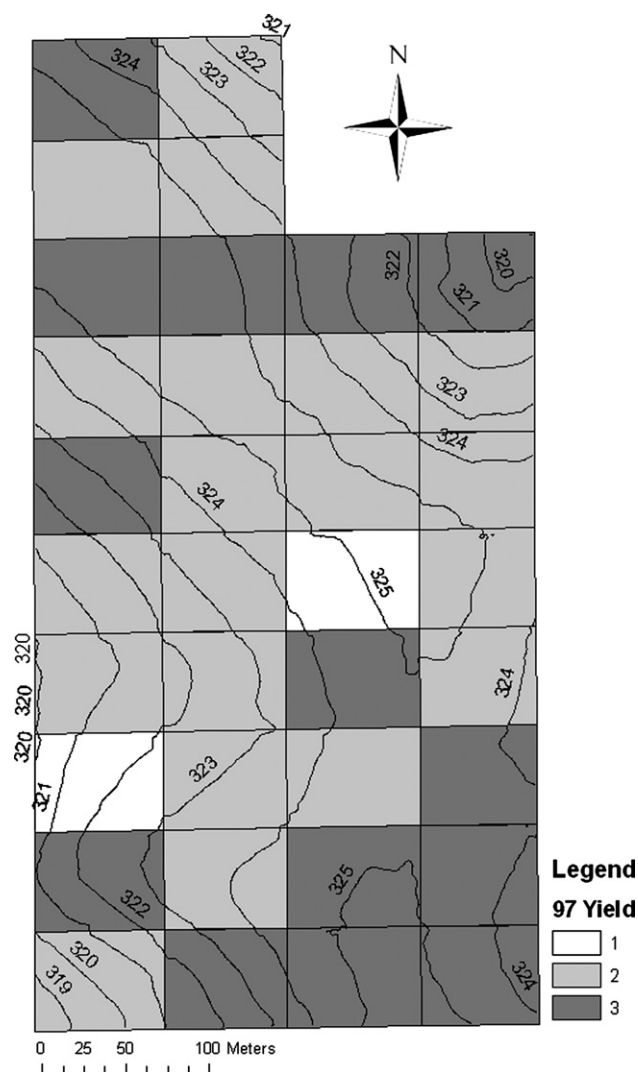


Fig. 9. Map overlay of 1997 normalized yield clusters and elevation contours. (normalized yield cluster means for 1997: 1 = -2.61; 2 = -0.44; 3 = 0.62).

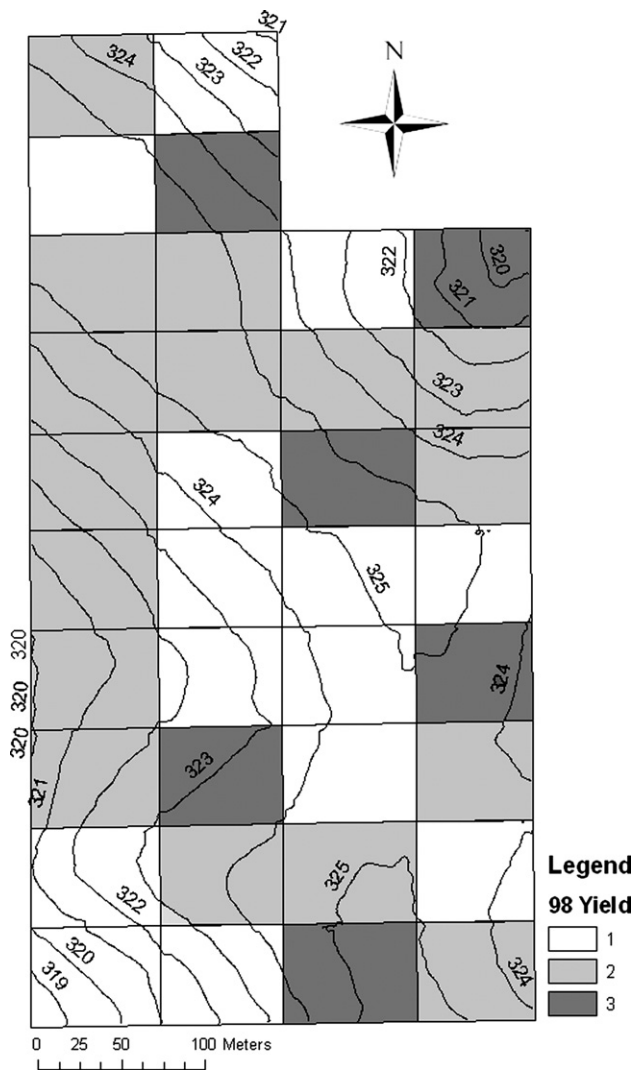


Fig. 10. Map overlay of 1998 normalized yield clusters and elevation contours. (normalized yield cluster means for 1998:1=−0.98;2=0.08;3=0.59).

was found to be consistent over the years in the medium category. No plot was found to be consistent in the low and high cluster categories over the six year study period (1993–98). This shows the degree of temporal variability effects on spatial yield clusters (Fig. 11).

3.4. GIS map overlay

The 1993 yield cluster data layer was overlain by the normalized drainage data layer of that year (Fig. 5) because it contributed significantly in cluster formation. Curvature was also identified as the significant variable because it affects the soil moisture distribution due to its surface nature (Sinai et al. (1981). Both of the variables probably affected the soil moisture and $\text{NO}_3\text{-N}$ availability to crops and resulted in yield variability. The majority of the clusters showing lower yields were observed along the ridge direction at relatively higher elevation levels (Fig. 5). Kravchenko and Bullock (2000) have reported that elevation had the most influence on yield, with higher yields consistently observed at lower landscape positions. Medium and

Table 3
Generalized square distance between clusters

Years	Clusters	1	2	3
1993	1	0	1.84	12.84
	2	1.84	0	11.87
	3	12.84	11.87	0
1994	1	0	2.72	2.42
	2	2.72	0	7.63
	3	7.42	7.63	0
1995	1	0	2.49	1.37
	2	2.49	0	0.93
	3	1.37	0.93	0
1996	1	0	2.03	5.30
	2	2.03	0	0.94
	3	5.30	0.94	0
1997	1	0	2.85	3.87
	2	2.85	0	0.28
	3	3.87	0.28	0
1998	1	0	0.93	4.23
	2	0.93	0	1.31
	3	4.23	1.31	0
Overall (1993–98)	1	0	3.13	2.87
	2	3.13	0	0.51
	3	2.87	0.51	0

Table 4
Confusion matrix for cluster membership prediction

From clusters	To clusters (number/percentage)			
	1	2	3	Total
1993				
1	7 (78)	2 (22)	0 (0)	9
2	6 (24)	18 (72)	1 (4)	25
3	0 (0)	0 (0)	2 (100)	2
Error rate	0.22	0.28	0	0.17
1994				
1	6 (60)	2 (20)	2 (20)	10
2	2 (11)	15 (83)	1 (6)	18
3	1 (13)	0 (0)	7 (87)	8
Error rate	0.40	0.17	0.13	0.23
1995				
1	8 (80)	1 (10)	1 (10)	10
2	3 (16)	14 (74)	2 (11)	19
3	1 (14)	3 (43)	3 (43)	7
Error rate	0.20	0.26	0.57	0.34
1996				
1	4 (80)	0 (0)	1 (20)	5
2	3 (18)	10 (59)	4 (23)	17
3	1 (7)	2 (14)	11 (78)	14
Error rate	0.20	0.41	0.21	0.27
1997				
1	2 (100)	0 (0)	0 (0)	2
2	3 (15)	10 (50)	7 (35)	20
3	2 (14)	3 (21)	9 (65)	14
Error rate	0.00	0.50	0.35	0.28
1998				
1	10 (72)	2 (14)	2 (14)	14
2	4 (25)	8 (50)	4 (25)	16
3	0 (0)	1 (17)	5 (83)	6
Error rate	0.28	0.50	0.17	0.32
Overall (1993–98)				
1	5 (63)	3 (37)	0 (0)	8
2	1 (11)	3 (33)	5 (56)	9
3	2 (10)	5 (26)	12 (63)	19
Error rate	0.37	0.67	0.37	0.47

higher yield clusters were observed at the medium and lower elevations levels which show the effect of soil moisture variation in 1993. This may also be explained by flushing of $\text{NO}_3\text{-N}$ from higher elevation areas towards lower which might have caused this spatial occurrence of yield clusters because of heavy rainfall in 1993. The year 1993 was a wet year and gave maximum tile flow for that year (Bakhsh and Kanwar, 2004). Similarly the convex and concave patterns of the landscape surface deals with the infiltration process. Several researchers have reported relationships between yield, topography and soil water content (Fiez et al., 1994; Schneider et al., 1997; Mulla and Schepers, 1997; Bakhsh et al., 2000).

In 1994, soil and curvature were identified as the effect variables for discriminating between clusters because it was a dry year (Sinai et al. (1981). Higher yield clusters for this year lie at the lower elevation levels because of better soil moisture availability (Fig. 6). Iqbal et al. (2005) have also reported that higher elevation areas generally yielded lower and may experience water stress earlier in the season, as compared with lower elevation areas. In 1995, aspect contributed significantly in cluster formation (Fig. 7). The map overlay of aspect and yield clusters showed that the majority of the low yield clusters were observed in the aspect zone of north-east directions (Fig. 7). Aspect can affect the evapotranspiration rates and crop yields (Yang et al., 1998; Kravchenko and Bullock, 2000).

Soil affected the cluster formation in 1996 (Table 5). The low yield clusters were mostly found at higher elevation levels and majority of the higher yield clusters was found at lower elevations for this year (Fig. 8) (Jiang and Thelen, 2004). The clusters formed in 1997 and 1998 did not show any relationship with the soil and topographic attributes. This could be associated with the rainfall amount and distribution during the growing season because the two years had the highest corn grain yields. The clusters formed during the analysis across the years also did not show the effect of any soil and landscape attributes because of variability over the years. This analysis shows that soil, curvature and aspect were the variables that contributed significantly for four of the six years. These variables have been reported to affect the soil moisture and nutrient availability to plants and crop yields (Sinai et al., 1981; Changere and Lal, 1997; Kravchenko and Bullock, 2000; Bakhsh et al., 2000). The analyses showed that no landscape variable contributed sig-

Table 5
Stepwise discriminant model for normalized yield based on clusters

Years	Variables	Partial R^2	F-value	Pr > F
1993	Normalized drainage 1993	0.17	3.25	0.05
	Curvature	0.29	6.66	<0.01
1994	Soil	0.17	3.37	0.05
	Curvature	0.27	5.98	<0.01
1995	Aspect	0.22	4.79	0.01
1996	Soil	0.25	5.43	<0.01
1997	–	–	–	–
1998	–	–	–	–
Overall	–	–	–	–

– No variable qualified to enter the model at $\text{Pr} \leq 0.15$.

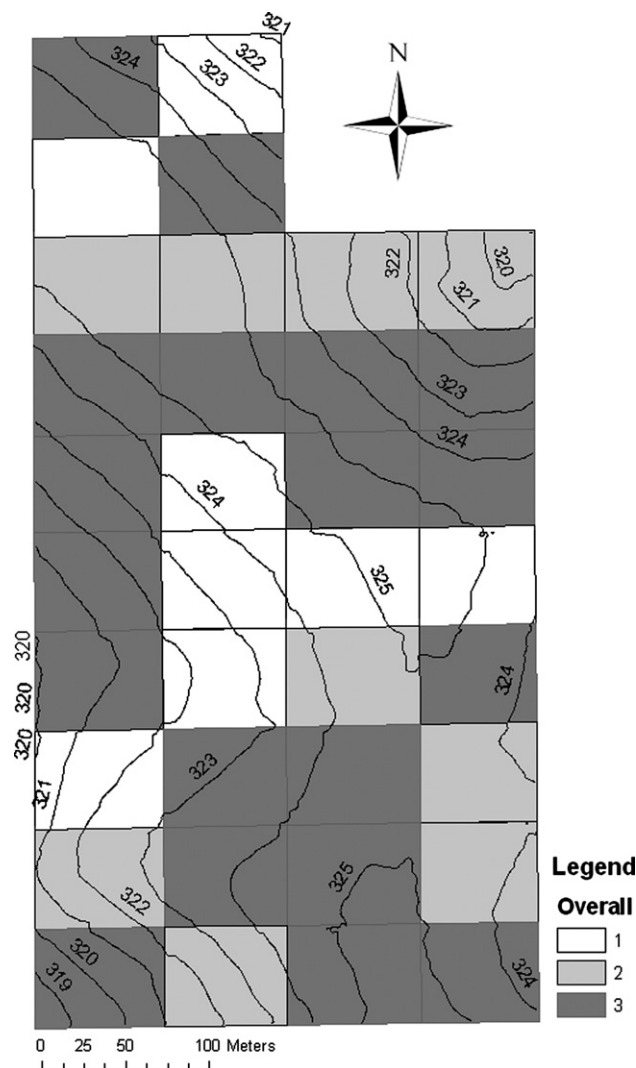


Fig. 11. Map overlay of the overall (1993–98) normalized yield clusters and elevation contours. (normalized yield cluster means for overall: 1 = -0.89; 2 = 0.09; 3 = 0.18).

nificantly for all the years because of inconsistent patterns of the yield clusters over the study period.

4. Summary

Grouping of response variables into meaningful classes can help manage their spatial occurrence. In this study, three classes of clusters namely low, medium, and high were generated for each year using corn–soybean yield data from 1993 to 1998. Cluster formation was found to be satisfactory using the evaluation criteria of R^2 , pseudo F -statistics and cubic clustering criteria. Corn–soybean grain yield clusters varied greatly over the years because of changing rainfall patterns. The distribution of the growing season rainfall and the cycles of wet and dry years affected soil moisture and nutrient availability to crops. The discriminant analysis revealed that landscape and hydrologic variables had significant effects on yields for the years having below and above average rainfall rather than those years having near average rainfall. The yield clusters, however, were found to be variable and inconsistent during the six year

study period. Soil and topographic attributes of curvature were identified as the variables that contributed significantly in cluster formations for most of the years. Both the variables have the potential to affect soil moisture and nutrient availability to crops especially for the soils having artificial subsurface drainage systems. Crop yield clusters were found to be affected by soil, topography and the growing season rainfall (Mulla and Schepers, 1997). The interactions between the soil and topographic attributes with the highly variable growing season rainfall affected spatial occurrence of clusters which were found to be inconsistent. The spatial zones delineated on the basis of the landscape and hydrologic attributes should be managed with suitable tillage and cropping systems to bring the long-term sustainability of watersheds. Similarly Fraisse et al. (2001) concluded that delineation of management zones based on topographic attributes and soil EC is a valid approach to capturing yield variability due to differences in plant water availability.

References

- Afyuni, M.M., Cassel, D.K., Robarge, W.P., 1993. Effect of landscape position on soil water and corn silage yield. *Soil Sci. Soc. Am. J.* 57, 1573–1580.
- Alexander, R.B., Smith, R.A., Schwarz, G.E., 1995. The regional transport of point and nonpoint source nitrogen to the Gulf of Mexico. In *Proc. Gulf of Mexico Hypoxia Conf.*, Kenner, LA. 5–6 Dec. 1995. USEPA Pub. 855R97001. USEPA Natl. Cent. for Environ. Pub. and Inf., Washington, DC.
- Al-Sulaimi, J., Khalaf, F.J., Mukhopadhyay, A., 1997. Geomorphological analysis of paleo drainage systems and their environmental implications in the desert of Kuwait. *Environ. Geol. (Berl.)* 29, 94–111.
- Baker, J.L., Colvin, T.S., Marley, S.J., Dawelbeit, M., 1989. A point-injector applicator to improve fertilizer management. *Appl. Eng. Agric.* 5, 334–338.
- Baker, J.L., Laffen, J.M., Schreiber, M.M., 1997. Potential for localized compaction to reduce leaching of injected anions. *J. Environ. Qual.* 26, 387–393.
- Baker, J.L., David, M.B., Lemke, D.W., 2005. Understanding nutrient fate and transport, including the importance of hydrology in determining losses, and potential implications on management systems to reduce those losses. *Proc. of Gulf Hypoxia and Local Water Quality Concerns Workshop*, held at Iowa State University on Sep. 26–28, 2005, pp. 11–25.
- Bakhsh, A., Kanwar, R.S., 2004. Using discriminant analysis and GIS to delineate subsurface drainage patterns. *Trans. ASAE* 47, 689–699.
- Bakhsh, A., Kanwar, R.S., 2005. Spatial clusters of subsurface drainage water NO₃-N leaching losses. *J. Am. Water Resour. Assoc.* 41, 333–341.
- Bakhsh, A., Colvin, T.S., Jaynes, D.B., Kanwar, R.S., Tim, U.S., 2000. Using soil attributes and GIS for interpretation of spatial variability in yield. *Trans. ASAE* 43, 819–828.
- Bakhsh, A., Kanwar, R.S., Bailey, T.B., Cambardella, C.A., Karlen, D.L., Colvin, T.S., 2002. Cropping system effects on NO₃-N losses with subsurface drainage water. *Trans. ASAE* 45, 1789–1797.
- Bjorneberg, D.L., Kanwar, R.S., Melvin, S.W., 1996. Seasonal changes in flow and nitrate-N loss from subsurface drains. *Trans. ASAE* 39, 961–976.
- Bjorneberg, D.L., Karlen, D.L., Kanwar, R.S., Cambardella, C.A., 1998. Alternative N fertilizer management strategies effects on subsurface drain effluent and N uptake. *Appl. Eng. Agric.* 41, 469–473.
- Blackmer, A., Pottker, M.D., Cerrato, M.E., Webb, J., 1989. Correlations between soil nitrate concentrations in late spring and corn yields in Iowa. *J. Prod. Agric.* 2, 103–109.
- Changere, A., Lal, R., 1997. Slope position and erosion effects on soil properties and corn production on a Mianian soil in central Ohio. *J. Sustain. Agric.* 11, 5–21.
- Colvin, T.S., Jaynes, D.B., Karlen, D.L., Laird, D.A., Ambuel, J.R., 1997. Yield variability within a central Iowa field. *Trans. ASAE* 40, 883–889.
- Conacher, A.J., Dalrymple, J.B., 1977. The nine unit landsurface model: an approach to pedogeomorphic research. *Geoderma* 18, 127–144.
- Cressie, N.A.C., 1993. *Statistics for Spatial Data*. John Wiley & Sons, New York NY.
- Dinnes, D.L., Karlen, D.L., Jaynes, D.B., Kaspar, T.C., Hatfield, J.L., Colvin, T.S., Cambardella, C.A., 2002. Nitrogen management strategies to reduce nitrate leaching in tile-drained Midwestern soils. *Agron. J.* 94, 153–171.
- Fiez, T.E., Miller, B.C., Pan, W.L., 1994. Winter wheat yield and grain protein across varied landscape positions. *Agron. J.* 86, 1026–1032.
- Fraisse, C.W., Sudduth, K.A., Kitchen, N.R., 2001. Delineation of site specific management zones by unsupervised classification of topographic attributes and soil electrical conductivity. *Trans. ASAE* 44, 155–166.
- Iqbal, J., Reed, J.R., Thomasson, A.J., Jenkins, J.N., 2005. Relationship between soil-landscape and dryland cotton lint yield. *Soil Sci. Soc. Am. J.* 69, 872–882.
- Jaynes, D.B., Colvin, T.S., 1997. Spatiotemporal variability of corn and soybean yield. *Agron. J.* 89, 30–37.
- Jaynes, D.B., Hunsaker, D.J., 1989. Spatial and temporal variability of water content and infiltration on a flood irrigated field. *Trans. ASAE* 32, 1229–1238.
- Jaynes, D.B., Hatfield, J.L., Meek, D.W., 1999. Water quality in Walnut Creek watershed: herbicides and nitrate in surface waters. *J. Environ. Qual.* 28, 45–59.
- Jaynes, D.B., Colvin, T.S., Kaspar, T.C., 2005. Identifying potential soybean management zones from multi-year yield data. *Comput. Electron. Agric.* 46, 309–327.
- Jiang, P., Thelen, K.D., 2004. Effect of soil and topographic properties on crop yield in a north-central corn–soybean cropping system. *Agron. J.* 96, 252–258.
- Kanwar, R.S., Baker, J.L., Baker, D.G., 1988. Tillage and split N-fertilizer effects on subsurface drainage water quality and crop yields. *Trans. ASAE* 31, 453–460.
- Kanwar, R.S., Colvin, T.S., Karlen, D.L., 1997. Ridge, moldboard, chisel, and no-till effects on subsurface drainage water quality beneath two cropping system. *J. Prod. Agric.* 10, 227–234.
- Kanwar, R.S., Bjorneberg, D., Baker, D., 1999. An automated system for monitoring the quality and quantity of subsurface drain flow. *J. Agric. Eng. Res.* 73, 123–129.
- Kanwar, R.S., Cruse, R., Ghaffarzadeh, M., Bakhsh, A., Karlen, D., Bailey, T., 2005. Corn–soybean and alternative cropping systems effects on NO₃-N leaching losses in subsurface drainage water. *Appl. Eng. Agric. ASAE* 21, 181–188.
- Kaspar, T.C., Pulido, D.J., Fenton, T.E., Colvin, T.S., Karlen, D.L., Jaynes, D.B., Meek, D.W., 2004. Relationship of corn and soybean yield to soil and terrain properties. *Agron. J.* 96, 700–709.
- King, J.A., Dampney, P.M.R., Lark, R.M., Wheeler, H.C., Bradley, R.I., Mayr, T.R., 2005. Mapping potential crop management zones within fields: use of yield-map series and patterns of soil physical properties identified by electromagnetic induction sensing. *Precis. Agric.* 6, 167–181.
- Kravchenko, A.N., Bullock, D.G., 2000. Correlation of corn and soybean grain yield with topography and soil properties. *Agron. J.* 92, 75–83.
- Kravchenko, A.N., Bollero, G.A., Omonode, R.A., Bullock, D.G., 2002. Quantitative mapping of soil drainage classes using topographical data and soil electrical conductivity. *Soil Sci. Soc. Am. J.* 66, 235–243.
- Lamb, J.A., Dowdy, R.H., Anderson, J.L., Rehm, G.W., 1997. Spatial and temporal variability of corn yields. *J. Prod. Agric.* 10, 410–414.
- Machado, S., Bynum, E.D., Archer, T.L., Bordovsky, J., Rosenow, D.T., Peterson, C., Bronson, K., Nesmith, D.M., Lascano, R.J., Wilson, L.T., Segarra, E., 2002. Spatial and temporal variability of sorghum grain yield: influence of soil, water, pests, and diseases relationships. *Precis. Agric.* 3, 389–406.
- Malone, R.W., Ma, L., Karlen, D., Meade, T., Meek, D., Heilman, P., Kanwar, R.S., Hatfield, J.L., 2007. Empirical analysis and prediction of nitrate loading and crop yield for corn–soybean rotations (in this issue of *GEODERMA*). *Geoderma* 140, 223–234.
- Mulla, D.J., Schepers, J., 1997. Key processes and properties for site-specific soil and crop management. *Proc. The State of Site Specific Management for Agriculture. ASA/CSSA/SSSA*, Madison, Wis, pp. 1–18.
- Pennock, D., Walley, F., Solohub, M., Si, B., Hnatowich, G., 2001. Topographically controlled yield response of canola to nitrogen fertilizer. *Soil Sci. Soc. Am. J.* 65, 1838–1845.
- Rabalais, N.N., Turner, R.E., Scavia, D., 2002. Beyond science into policy: Gulf of Mexico hypoxia and the Mississippi River. *Bioscience* 52, 129–142.
- Randall, G.W., Vetsch, J.A., 2005. Nitrate losses in subsurface drainage from a corn–soybean rotation as affected by fall and spring application of nitrogen and nitrapyrin. *J. Environ. Qual.* 34, 590–597.

- Reuter, H.I., Giebel, A., Wendroth, O., 2005. Can landform stratification improve our understanding of crop yield variability. *Precis. Agric.* 6, 521–537.
- SAS, 2003. SAS Online Doc, Version 9.1. SAS Institute, Inc., Cary, N.C.
- Schneider, S.M., Boydston, R.A., Han, S., Evans, R.G., Campbell, R.H., 1997. Mapping of potato yield and quality. In: Stafford, J.V. (Ed.), *Precision Agriculture '97: Proceedings of the First European Conference on Precision Agriculture*. BIOS Scientific Publishers, Oxford, UK, pp. 253–261.
- Si, B.C., Farrell, R.E., 2004. Scale-dependent relationship between wheat yield and topographic indices: a wavelet approach. *Soil Sci. Soc. Am. J.* 69, 872–882.
- Silva, J.R.M.D., Alexandre, C., 2005. Spatial variability of irrigated corn yield in relation to field topography and soil chemical characteristics. *Precis. Agric.* 6, 453–466.
- Sinai, G., Zaslavsky, D., Golany, P., 1981. The effect of soil surface curvature on moisture and yield — Beer Sheva observations. *Soil Sci.* 132, 367–375.
- USEPA, 1995. National Water Quality Inventory: 1994 Report to Congress. USEPA Office of Water, U.S. Government Printing Office, Washington, DC.
- USGS., 2005. http://toxics.usgs.gov/hypoxia/hypoxic_zone.html.
- Voy, K.D., 1995. Soil Survey of Floyd County, Iowa. USDA-SCS in cooperation with Iowa Agricultural and Home Economics Experiment Station. Iowa State University, Cooperative Extension Service, Ames, Iowa.
- Yang, C., Peterson, C.L., Shropshire, G.J., Otawa, T., 1998. Spatial variability of field topography and wheat yield in the Palouse region of the Pacific Northwest. *Trans. ASAE* 41, 17–27.