Computers and Electronics in Agriculture 117 (2015) 154–167

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Perspectives on delineating management zones for variable rate irrigation

Amir Haghverdi [a,*], Brian G. Leib [b], Robert A. Washington-Allen [c], Paul D. Ayers [b], Michael J. Buschermohle [b]

[a] Department of Biological Systems Engineering, University of Nebraska-Lincoln, Panhandle Research and Extension Center, 4502 Avenue I, Scottsbluff, NE 69361-4939, United States
[b] Department of Biosystems Engineering & Soil Science, University of Tennessee, 2506 E.J. Chapman Drive, Knoxville, TN 37996-4531, United States
[c] Department of Geography, University of Tennessee, Burchfiel Geography Building, Knoxville, TN 37996-0925, United States

ARTICLE INFO

Article history:
Received 27 April 2015
Received in revised form 20 June 2015
Accepted 23 June 2015
Available online 15 August 2015

Keywords:
Apparent electrical conductivity
Integer linear programming
Remote sensing
Soil water retention
Unsupervised clustering

ABSTRACT

This study aimed at investigating the performance of multiple irrigation zoning scenarios on a 73 ha irrigated field located in west Tennessee along the Mississippi river. Different clustering methods, including *k*-means, ISODATA and Gaussian Mixture, were selected. In addition, a new zoning method, based on integer linear programming, was designed and evaluated for center pivot irrigation systems with limited speed control capability. The soil available water content was used as the main attribute for zoning while soil apparent electrical conductivity (ECa), space-borne satellite images and yield data were required as ancillary data. A good agreement was observed among delineated zones by different clustering methods. The new zoning method explained up to 40% of available water content variance underneath center pivot irrigation systems. The ECa achieved the highest Kappa coefficient (=0.79) among ancillary attributes, hence exhibited a considerable potential for irrigation zoning.

1. Introduction

1.1. Precision farming and management zone delineation

In conventional agriculture each field is considered as a uniform unit, by purposely ignoring the heterogeneity across the field, thereby decision-making is based on an estimation of average conditions. The motivation for site-specific farming was first addressed by researchers during the late 80s and early 90s (Arslan and Colvin, 2002). As such, precision agriculture (PA) methodology is a way to look at field management by taking the within field variation into account and incorporating that variability into management decisions. Within-field heterogeneity is caused by both temporal and spatial variation of a variety of factors such as climate, topography and biologic activity (Córdoba et al., 2013).

A management zone (MZ) is a sub-region of a field that is relatively homogeneous with respect to soil-landscape attributes. It is expected that variable rate application across MZs will help by saving the resources and optimizing yield (Schepers et al., 2004). Protecting the environment and keeping agriculture sustainable may also be achievable through precision farming. Sensor-based and map-based approaches are two major methods to practice variable-rate application. In the sensor-based method, a real time decision on application rate is made using data collected via sensors and a pre-developed application algorithms. In the map-based method, application maps are prepared using site-specific information such as yield data and soil data prior to implementation. It is critical to select appropriate attribute(s) and method to delineate robust zones (Thöle et al., 2013).

A field can be zoned based on a single soil-crop variable or multiple attributes which are expected to affect yield (Khosla et al., 2010). Yield maps, topography, satellite photographs, canopy images and soil apparent electrical conductivity (ECa) are among suggested attributes to delineate MZs. Application of remote sensing is especially attractive because it is noninvasive and relatively inexpensive (Schepers et al., 2004). Yield maps are useful sources of information reflecting within-field variation. However, some difficulties have been reported to delineate zones solely by yield maps (Khosla et al., 2010). Temporal inconsistency among yield maps from year to year is probably the main reason causing this problem. Schepers et al. (2004) reported that temporal climate variability in an irrigated cornfield significantly affects yield spatial variability from year to year. Combining yield data with other ancillary information or averaging yield data over years can help explain spatial variation better and in turn can provide more trustable zones. Promising results have been reported by the studies that have utilized several years of yield data to create MZs.

* Corresponding author. Tel.: +1 (308) 632 1246.
E-mail address: ahaghverdi2@unl.edu (A. Haghverdi).

http://dx.doi.org/10.1016/j.compag.2015.06.019
0168-1699/

However, stability of such zones has to be tested for each individual field (Khosla et al., 2010). Calculating temporal variance can be helpful to verify zone stability (Basso et al., 2012).

There are some methods to screen potential attributes regarding their importance. Hornung et al. (2006) suggested assigning different weights to individual layers based on their importance to the variation in crop yield. Principal component analysis (PCA) can be utilized to linearly transfer original variables to new independent ones. Researchers apply this technique to understand the characteristics of a data set along with relative importance of each individual variable (Fraisse et al., 2001). Finally, the most important principal components will be used to delineate MZs.

There are several methods to delineate MZs. Applying unsupervised clustering techniques and zoning via user-defined thresholds are the main procedures. Clustering techniques group similar data points (cells), based on their inherent structure, into distinct classes. Methods such as *k*-means and fuzzy *k*-means have been widely used to identify MZs (Córdoba et al., 2013). These methods usually produce separated oval shape zones across a field. The software MZA was developed by Fridgen et al. (2004) for delineating MZs. The MZA uses fuzzy *k*-means for zoning and tests the result to evaluate how many zones to create in a given field. More recently, Zhang et al. (2010) developed a web-based decision support system for zoning using satellite imagery and field data. Moral et al. (2010) used regression kriging, PCA and fuzzy cluster classification to delineate MZs using soil texture information and electrical conductivity as ancillary data. Recently Cid-Garcia et al. (2013) proposed an integer linear programming MZ delineation method to make rectangular shaped zones which facilitates the work and operation of machinery.

1.2. Irrigation management zones delineation

Water has become the most valuable input for agriculture across the world. There has been a significant conversion from rainfed to irrigated production in humid regions as a safe guard against unpredictable severe drought periods within a cropping season which can cause yield reduction in such regions. At the same time, agriculture is under a great pressure to enhance its water use efficiency as other sectors such as industry and urban areas are demanding more water (Evans et al., 2013; Daccache et al., 2015). When within field soil spatial variation is significant, variable rate irrigation (VRI) comes into play as an engineering solution to manage spatial allocation of applied water through irrigation. Evidence exists in literature showing VRI could enhance water use efficiency and/or yield. Hedley and Yule (2009) mentioned up to 26% annual water saving by VRI comparing to conventional uniform irrigation. King et al. (2006) reported 4% higher potato tuber yield (statistically not significant) under VRI in comparison to uniform irrigation.

Precision irrigation center pivots have been commercially available for a while yet their adoption by farmers has been slow to develop. The prime drawbacks of VRI systems are high capital costs of implementation and management which is hard to rationalize considering energy and water saving at current price (Evans et al., 2013). Most of the available center pivots, however, have control panels with speed control module, meaning farmers can vary the irrigation across their fields to some extent at no extra cost, if it is needed. Changing the pivot travel speed enables putting some pie shape zones where speeding the system up provides less time to irrigation and slowing the system down increases the quantity of applied irrigation at the selected pies.

There is a critical need to dynamically develop irrigation MZs in an accurate and inexpensive manner (Evans et al., 2013). Delineating zones for irrigation management is challenging. Soil physical and hydraulic properties govern plant available water and thereby directly affect irrigation scheduling. Unfortunately hydraulic properties of soil are not readily available at a field-scale to be used for delineating irrigation MZs. There is evidence in the literature that easily collected ancillary data may be spatially correlated to soil physical and hydraulic properties, thus be useful for irrigation zone delineation. In irrigation scheduling, plant available water is calculated within the effective root zone which expands to deeper soil layers as the cropping season progresses. If in depth soil variation is significant at the field-scale, theoretically, the spatial arrangement of irrigation MZs may change during the growing season. This is not, however, in agreement with the usual practice of static MZs sometimes for years. Given the limitation in number, size and location of speed control pie shape zones, an optimization procedure is needed to find the best combination of pies across a field.

The objectives of this study were: (i) to evaluate the performance of some zoning algorithms on an irrigated field with significant spatial soil variation located in west Tennessee (ii), to investigate the usefulness of proximal data, including ECa and space-born satellite images, to delineate irrigation MZs, (iii) to analyze stability of irrigation MZs within growing season in respect to soil available water content, and (iv) to develop and evaluate an irrigation zoning method for center pivot systems with speed control capability using integer linear programing technique.

2. Material and methods

2.1. Field of study

An irrigated agriculture field, approximately 73 ha located in west Tennessee along the Mississippi river, was selected (Fig. 1). The field was located in a short season, semi-humid region with high rainfall potential during a cropping season but some dry periods when supplemental irrigation is used to fulfill the crop water requirement. There were two center pivot systems for irrigation covering majority of the field with some overlap in their coverage. Cotton was planted in 2012, 2013 and 2014 cropping seasons in that field. The spatial variation in soil physical and hydraulic properties was significant (Table 1 and Fig. 2) which made this field a perfect case for the purpose of our study.

Soil sampling was done on March 21 and 22, 2014. A total of 400 undisturbed samples were collected from 100 sampling locations at 4 different depths, 0–25 cm, 25–50 cm, 50–75 cm and 75–100 cm. The soil texture components as well as soil bulk density and volumetric water content were measured in lab. The soil water retention curve for samples and high resolution soil available water content (AWC) maps were predicted using pseudo continuous pedotransfer function (Haghverdi et al., 2012, 2014) and co-kriging. There was a decline in soil water holding capacity (i.e. water content at matric potential -33 kPa, θ_{33}) by depth (Table 1) which is in agreement with the expansion of the sandy regions across the field moving from surface to deeper layers.

2.2. Proximal data

It is appealing to delineate MZs by proximal data because they are low in cost, easy to collect and usually noninvasive. In this study soil ECa and space-borne satellite photos were considered as proximal information for zone delineation. The soil ECa measurements from shallow (i.e. ECS, 0–30 cm) and deep (i.e. ECD, 0–90 cm) layers were collected (about 4700 data points covering the entire field) using a Veris 3100 (Veris Technologies, Salina, KS) instrument on March 20, 2014. The Veris machine measures ECa using the principles of electrical resistivity. A small electrical current is introduced into the soil and the drop in voltage in two

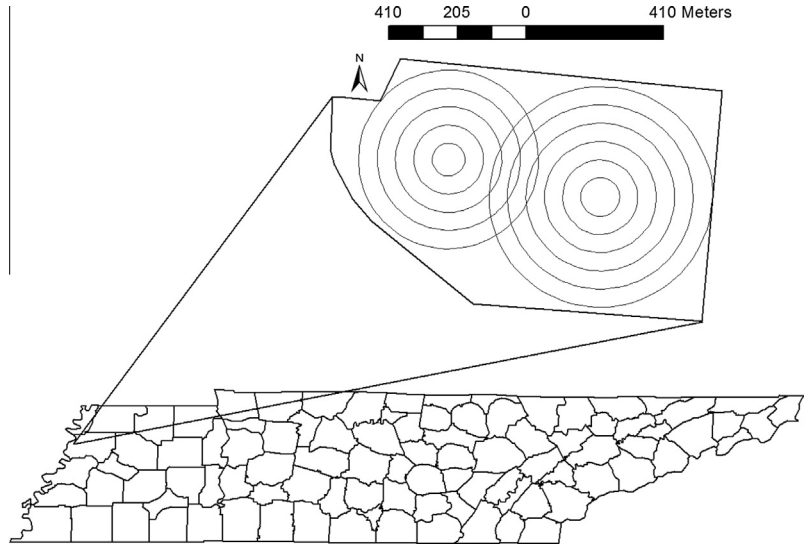

Fig. 1. Field of study within the state of Tennessee.

Table 1
Descriptive statistics for soil water content from different soil layers (Haghverdi, 2015).

Variable[a]	Depth	Min.	Max.	Mean	SD[b]
θ_{33}, %	0–25 cm	9.86	43.57	29.43	5.83
	25–50 cm	5.26	47.20	26.03	10.75
	50–75 cm	4.93	43.54	20.55	11.26
	75–100 cm	5.38	38.95	17.31	9.71
θ_{1500}, %	0–25 cm	7.72	29.61	19.69	4.28
	25–50 cm	5.23	35.49	16.98	7.22
	50–75 cm	4.57	29.32	13.14	6.47
	75–100 cm	4.75	24.52	11.13	5.09

[a] θ_{33}: water content at field capacity (−33 kPa); θ_{1500}: water content at permanent wilting point (−1500 kPa).
[b] SD: standard deviation.

depths is measured. The ECS data showed a normal distribution but ECD information were log transferred to have a roughly Gaussian distribution. Then, the ECS and ECD data were interpolated using kriging to cover the entire field (with 2 m spatial resolution) with the exception of a drainage pathway starting from west of the field and ending to the southeastern corner of the field.

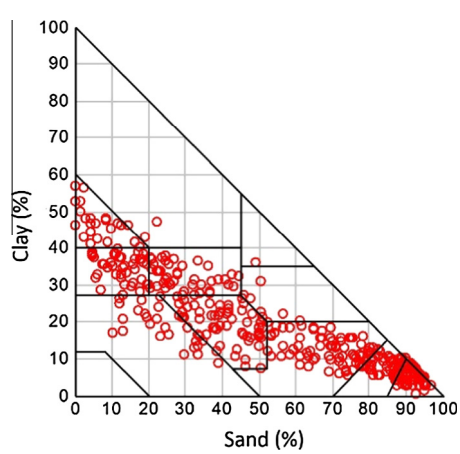

Fig. 2. Soil texture distribution for the samples ($n = 400$) collected from the field of study.

The satellite images were obtained from Landsat 8 Operational Land Imager (OLI) sensor during 2013 and 2014 (through http://earthexplorer.usgs.gov/). A total of 19 cloud-free images were selected for statistical analysis. The selected images contain the reflectance at multiple bands (i.e. quantized and calibrated scaled Digital Numbers (DN)). The reflectance of a canopy and bare soil will change during a cropping season as a plant goes through different growth stages and soil water status is changed. The band 8, panchromatic band, with wavelengths ranged from 0.50 to 0.68 μm for Landsat 8 has the highest spatial resolution (i.e. 15 m × 15 m spatial resolution) within available bands, thus was selected for this study. The field of study contained a total of 2320 cells.

The cotton yield data were available from 2010 to 2012 cropping seasons. First, the raw yield data were cleaned. Yield data that were greater or smaller than the mean ± 3 standard deviation were assumed to be outliers and removed from the analysis. Then, kriging was applied to interpolate yield data (with 2 m spatial resolution) across the field with the exception of the drainage path. Finally, the yield data were standardized for each year and averaged across years to make a single yield map.

2.3. Management zone delineation

Multiple zoning strategies were applied and evaluated through two phases (Table 2). In phase one, the entire field was zoned. The ECa and reflectance data (i.e. DN) were tuned to have a range between 0 and 1 by dividing each cell value by the overall maximum values across the field from all data points/cells. The ECa data were chosen as the prime attribute for zoning because they showed a good correlation with soil physical and hydraulic properties of the soil within the field of study (Haghverdi, 2015). Three unsupervised clustering techniques, i.e. *k*-means, ISODATA-maximum likelihood and Gaussian mixtures, with different combination of ECa data were examined. Unsupervised clustering techniques group the data based on their inherent structure as opposed to supervised clustering methods which require a priori knowledge on data for training. More information on each clustering technique is provided later in this paper. Given the performance of the zoning strategies and distribution of the digital reflectance data, only *k*-means was applied for zoning using satellite images.

Table 2
Summary of zoning algorithms and attributes.

	Clustering procedure[a]	Attributes[b]
Phase 1	*k*-means, ISODATA-ML, GM	ECS
	k-means, ISODATA-ML, GM	ECD
	k-means, ISODATA-ML, GM	ECS, ECD
	k-means	Satellite image
	User defined zones	AWC
Phase 2	ILP, *k*-means	ECD
	ILP, *k*-means	Satellite image
	ILP, *k*-means	AWC
	ILP, *k*-means	Yield

[a] GM: Gaussian Mixtures; ISODATA-ML: Iterative Self-Organizing Data Analysis Technique – Maximum Likelihood; ILP: Integer Linear Programming.
[b] AWC: Available Water Content; ECS: shallow ECa; ECD: deep ECa.

The within-season temporal stability of MZs was also investigated in phase one through a user-defined zoning strategy considering AWC as input. The effective root zone (0–100 cm) was defined as a depth in which crop absorbs most of its water requirement. To mimic the dynamic of effective root growth, four crop available water maps were created for 0–25, 0–50, 0–75 and 0–100 cm. Each map, then, was divided into user-defined MZs and the spatial arrangement of zones were compared against each other.

In phase two, we focused on the area underneath the center pivot irrigation systems for which we developed and evaluated a new zoning approach using integer linear programming (ILP). The yield data were also included to test how efficient this variable was in the zone delineation process. Most center pivot irrigators prefer to have less than 10 MZs in a field (Evans et al., 2013). Moreover, a majority of available center pivots in west Tennessee cannot put more than 10 pies. Therefore, clustering procedures were repeated to delineate from 2 to up to 10 MZs.

Two software products were used for clustering: Matlab R2015a (MathWorks, Inc., Natick, Mass.) and ArcGIS 10.2.2 (ESRI Inc., Redlands, California).

2.3.1. K-means

The *k*-means is one of the most widely used methods for clustering due to its simplicity and efficiency. The *k*-means, in an iterative process, tries to partition data into *k* groups so that the differences among the features in a group, over all groups, is minimized. It moves the observations between clusters and monitors the sum of distances from each observation to the center of its cluster until group membership stabilizes. The *k*-means++ algorithm was used to assign initial centers to the clusters. This algorithm chooses the cluster centers in a random manner yet favors spreading them out by a distance-based weighting process (Jain, 2010; Jain et al., 1999; ESRI, 2014). The Euclidean metric was used to obtain distances. No spatial constraint was considered.

2.3.2. ISODATA – Maximum likelihood

A well-known extension of *k*-means is called Iterative Self-Organizing Data Analysis Technique (ISODATA). The ISODATA starts with randomly assigning data into different classes. Then, over the course of an iterative process, the algorithm changes the membership of data points and tries to find the optimum clusters where the Euclidian distance between data points to cluster center is minimized. ISODATA applies some heuristics to adjust the number of clusters; clusters could be removed, divided or merged at the end of each iteration based on similarities between neighboring clusters and specified minimum class size (Jain, 2010; Jain et al., 1999; ESRI, 2014). In this study, functionality of ISODATA was combined with Maximum Likelihood which assigns cells to classes using Bayes' theorem. This method assumes that the distribution of a class sample is normal, hence uses mean vector and the covariance matrix to assign each cell to a class based on statistical probability.

2.3.3. Gaussian mixture model

A Gaussian Mixture (GM) model is a parametric probability density function which is often used for soft data clustering. If clusters have different sizes and correlation within them, GM may be more appropriate option for clustering than *k*-means. This method assumes data belong to a mixture of normal density components with unknown parameters. The GM parameters are estimated using an expectation maximization algorithm by assigning posterior probabilities to each component density. Each cluster corresponds to one of the Gaussian components, hence data are assigned to clusters such that posterior probability is maximized (Reynolds, 2009; Mathworks, 2014).

2.3.4. Integer linear programing

A linear programming (LP) problem is an optimization (maximization or minimization) problem of a linear objective function, subject to linear equality and/or inequality constraints. If some or all of the variables are limited to be integers, it is called an integer linear programming (ILP) problem. The LP has been a very popular technique to irrigation management and water resources system analysis and planning (Singh, 2012). Recently, the ILP has been successfully used for rectangular shape MZ delineation and crop planning optimization problems by Cid-Garcia et al. (2013, 2014).

The area underneath each center pivot irrigation system was divided into 360 pies each 1 degree wide. Then, a spatial join process was implemented to identify data points (cells) within each pie for each attribute. A total of 129,241 zones was created by combining 1 degree pies (Table 3). The standard deviation was calculated for data points within each zone with respect to the target attribute.

An integer linear programming model was designed to find the optimum spatial arrangement of pie-shape zones underneath each pivot. The general mathematical model was as follows:

$$\min \sum_{i=1}^{n}\sum_{k=1}^{m}(l \times a_k \times \sigma_{ik} \times x_i) \tag{1}$$

subject to

$$\sum c_{ij}x_i = 1 \tag{2}$$

$$\sum_{i=1}^{n} x_i \leqslant P_{\max} \tag{3}$$

$$lx \geqslant L_{\min} \tag{4}$$

Table 3
All of the pie shape zones underneath a center pivot irrigation system.

		Pie length			
		1	2	···	360
Start angle	0	0–1[a]	0–2	···	0–0
	1	1–2	1–3	···	1–1
	2	2–3	2–4	···	2–2
	.	.	.	···	.
	.	.	.	···	.
	.	.	.	···	.
	359	359–0	359–1	···	359–359

[a] *X–Y*: start angle-end angle.

$x_i \in \{0, 1\}$

where x_i is the decision variable for zone i, σ is the standard deviation for attribute k within zone i, a_k is a user-defined weighting factor based on the relative importance of an attribute in zoning process, m is the total number of attributes, n is the total number of the zones, l is the length of each zone in degrees, c_{ij} is a coefficient equal to 1 if angle j is covered by zone i otherwise equal to 0, $P_{\max}$ is the maximum number of desired zones.

Objective function (1) minimizes a weighted average of standard deviations across the zones. In this study we did not use more than one variable at each optimization attempt, thus k and a_k were set to one. Restriction (2) ensures that each degree is covered by only one zone and selected zones cover the entire area underneath the irrigation system. Constraint (3) ensures that total number of the pies is less than the maximum desired number which was set to 10 in this study. Constraint (4) ensures that the length of the pies is greater than the minimum desired value which was chosen to be 5 degree. For a pie, the decision variable of the model, x_i, was 1 if the pie was selected and 0 otherwise.

2.4. Performance evaluation

The overall within-zone variance of AWC was calculated to assess zoning strategies and find the optimum number of zones. The AWC variances for each zone were weighted considering the zone-area (Fraisse et al., 2001):

$$S_Z^2 = \frac{1}{n_Z}\sum_{i=1}^{n_Z}(\mathrm{AWC}_i - m)^2 \times \frac{n_Z}{n_T} \tag{5}$$

where S_z^2 is weighted variance for zone z, AWC is the soil available water content for cell i, m is mean of soil available water values in zone z, n_Z and n_T are number of cells in zone z and total number of cells across field, respectively.

Total within-zone AWC variance across field (S_T^2) was obtained as the sum of weighted within-zone AWC variances:

$$S_T^2 = S_1^2 + S_2^2 + \ldots + S_Z^2 \tag{6}$$

The better zoning scenario was selected as the one with minimum total AWC standard deviation. For the ILP method, the kappa coefficient (Cohen, 1960) was also calculated to measure interclassification agreement of different zoning strategies by ancillary data against optimum zones using AWC data. The kappa coefficient equals 0 and 1 when the agreement is due to chance alone or when there is perfect agreement, respectively. A positive coefficient indicates agreement exceeds chance and the magnitude of coefficient reflects the strength of agreement.

3. Results

3.1. Zoning using ECa and satellite images

Fig. 3 depicts the variance of AWC as well as average yield within zones for a different number of MZs and multiple clustering methods. The AWC variance across the field, not divided into zones, was chosen as a reference (i.e. one MZ with 100% variance). For k-means, up to 11%, 19% and 16% of variance was explained considering ECS, ECD and ECS plus ECD as input, respectively. Considering more than 4 zones only reduced the variance by an additional 1% (ECS input), 1% (ECD input) and 2% (ECS + ECD inputs). For ISODATA-ML, variance was reduced by 11%, 19% and 17% considering ECS, ECD and ECS plus ECD as input, respectively. Given more than 4 zones, the variance decreased by an additional 1% (ECS input), 1% (ECD input) and 4% (ECS + ECD inputs). For GM method, 10% (ECS input), 17% (ECD input) and 16% (ECS + ECD inputs) of the variance reduction occurred by dividing the field into 4 zones while further division to up to 10 zones only decreased the total variance by an additional 1% (ECS input), 2% (ECD input) and 5% (ECS + ECD inputs).

Adding more than 4–5 zones only slightly improved the clustering results. In fact, most of the AWC variance was explained by dividing the field into two MZs. This is in line with changes observed in within-zone yield average; when the number of zones increased to more than 4–5, one starts to see more zones with similar productivity. The clustering methods performed similar. The highest reduction in variance was observed for GM with 10 zones and ECS plus ECD as inputs. However, considering 4 as the optimum number of zones, k-mean and ISODAT-ML performed slightly better than GM. Noting this similarity among results of clustering methods and given the simplicity and ease of use of k-means method, it was decided to only use it for the rest of the unsupervised clustering tasks during this study. Aggregating ECS and ECD did not decrease the within-zone variance of AWC for k-means and ISODATA-ML procedures. For an optimum number of zones ($n = 4$), the lowest variance was observed when ECD was used as an input. Therefore, the ECD was considered an attribute of high importance for explaining the variability found in AWC for the field of study.

Fig. 4 illustrates the gray-scale panchromatic images taken by Landsat 8 during 2013 and 2014. Initial assessment of satellite images revealed the reflectance value of cells adjacent to the boundaries of the field were affected by the roads (located at north and west parts of the field) and trees (located at south and east parts of the field) and in turn influenced the spatial distribution of the entire field. Therefore, the border cells were filtered out from satellite images prior to running clustering analysis. The cells within drainage pathways were also removed from satellite images since they were missing from other attributes too. Moreover, they caused the same problem as border cells. The field was planted (harvested) in cotton on May 30, 2013 (December 2–3, 2013) and on May 5, 2014 (October 18–20, 2014), meaning some of the satellite images had been taken from bare soil but some from the cotton canopy in different growth stages. There were moderate temporal changes in spatial patterns of satellite photos. However, it was possible to visually associate brightness level to AWC spatial variation in a good number of images (i.e. photos taken on April 13, 2013; April 29, 2013; October 22, 2013; August 6, 2014; August 22, 2014; and September 23, 2014).

The coarse-textured soils with low AWC were mostly located over three spots across the field: close to the southern border of the field mostly outside of the outer pivot circles, (ii) underneath the eastern pivot approximately from the pivot point up to the right edge of the field, and (iii) a region in the northwestern part of the field above the drainage path adjacent to the northern border. In general, very bright cells (i.e. cells with high reflectance value/digital number) were mostly associated with coarse-textured soils with low AWC values while darker-colored cells (i.e. cells with low reflectance value/digital number) were associated with higher AWC soils. The photo acquired in August 6, 2014 revealed considerable spatial agreement with our understanding of AWC patterns. Factors such as ponding water due to rainfall (i.e. photos taken on May 15, 2013; December 25, 2013; February 7, 2014; March 31, 2014; April 16, 2014, May 2, 2014; May 18, 2014; October 25, 2014 and December 12, 2014) and wet soil due to an irrigation event (i.e. photo taken on July 21, 2014) affected the spatial arrangement of satellite photos.

To understand temporal changes, the distribution of standardized brightness values for the satellite photos are illustrated in Fig. 5 where similar distributions were roughly grouped by visual assessment. Coarse textured soils appeared as a long right tail in satellite photos grouped in panel A and D, while spots with

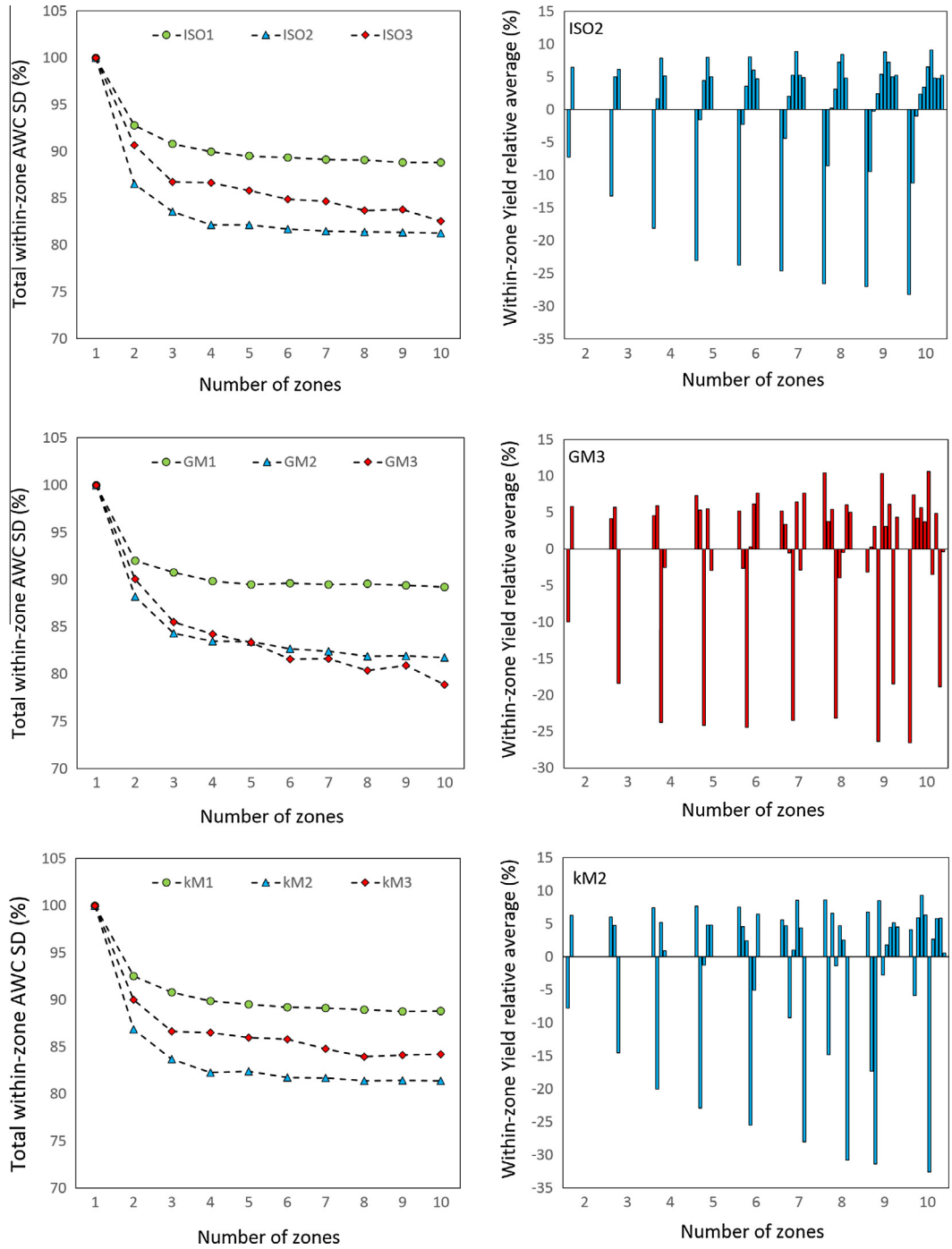

Fig. 3. Performance of clustering methods based on percent reduction in AWC variance and within-zone average yield. kM: *k*-means; GM: Gaussian mixture; ISO: ISODATA-maximum likelihood; 1, 2, and 3: ECS, ECD and ECS plus ECD were used for clustering.

ponding water formed a long left tail for images in panel C and caused negative skewness. The satellite images with less clear spatial pattern formed panel B. Images with no vivid spatial pattern showed high positive Kurtosis (panel E). The image taken April 16, 2014 did not follow the distributions observed with the other images; with negative kurtosis it depicted both ponding water and coarse soils but did not well represent the entire spatial arrangement of AWC across the field. The better satellite images seemed to follow a log normal distribution (panel D) which was the distribution of raw ECD data as well.

Fig. 6 shows the clustering results in terms of average within-zone AWC variance using satellite images as input. In 2014, images taken August 6 and September 23 showed the highest performance with 21% and 14% reduction in AWC variance, respectively. In 2013, images taken November 7 and April 29 exhibited the better result by explaining 12% and 10% of AWC variance, respectively. Identical to what we observed for ECa data, adding more than 4–5 zones only slightly improved explaining AWC variance. On average, for the abovementioned satellite images with acceptable performance, considering more than 4 zones only helped to explain an additional 2% of AWC variance. The zones delineated from rest of the images explained less than 10% of the AWC variance. The lowest performance belonged to images taken 19 July and 12 December 2014 which explained only 1% of the AWC variance. In

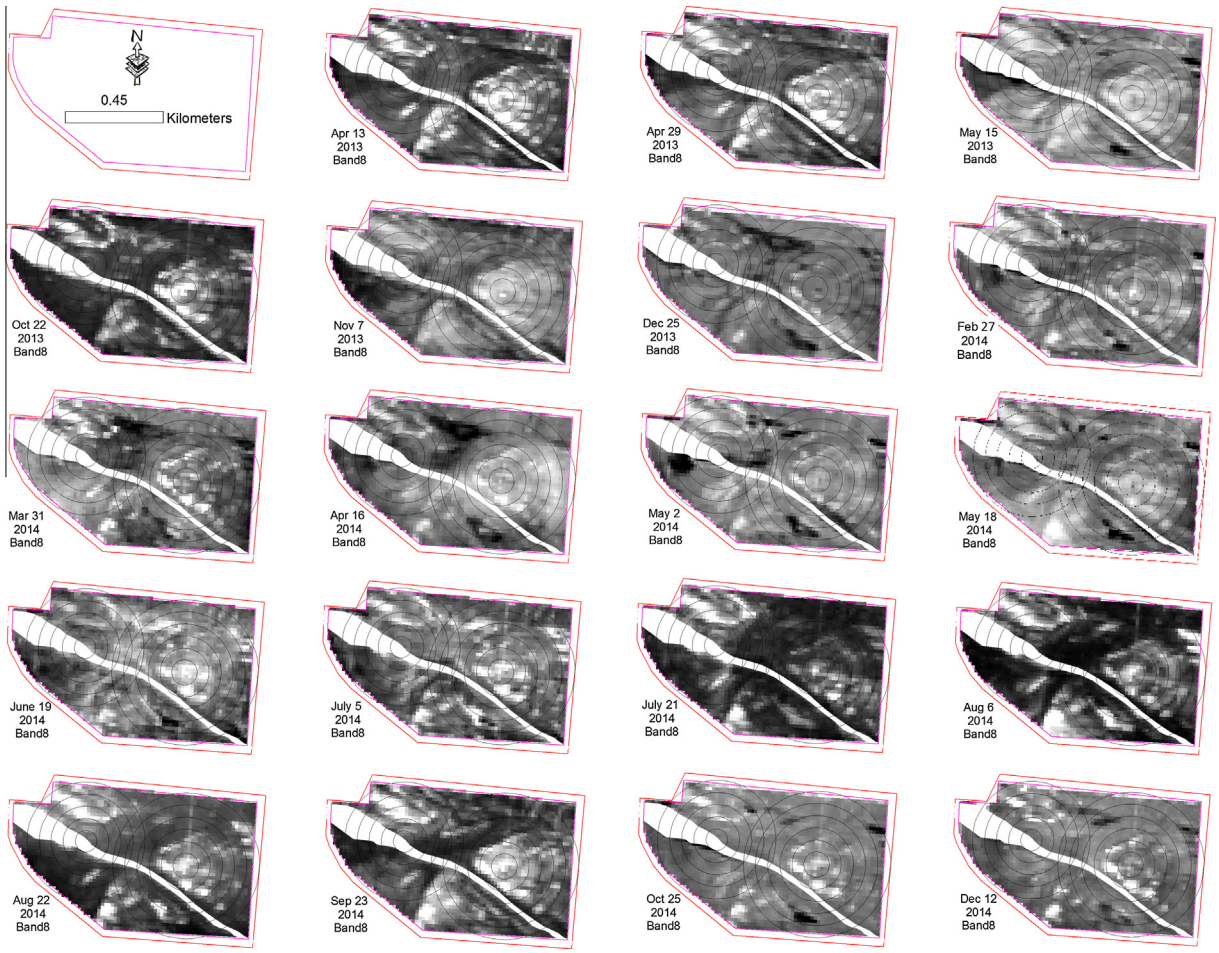

Fig. 4. Gray-scale maps of panchromatic spectral band during 2013–2014 from the field of study.

2013, images taken on 22 October and 25 December exhibited the lowest efficiency by explaining 4% and 3% of AWC variance, respectively. With a few exceptions, our visual assessment of reflectance distribution (Fig. 4) turned out to match the efficiency of images in zone delineation (Fig. 5). Overall, the highest performance was observed for images in groups D and A with an average 11% and 10% reduction in variance, respectively. In contrast, images in groups E and C showed the lowest potential for clustering by only explaining 3 and 4 percent of AWC variance, respectively. The higher number of cloud free images available contributed in better zoning results in 2014 compare with 2013. We also attribute this to difference in field conditions; in 2013 planting was delayed due to cold weather and wet soil which affected cotton growth and maturity and also diminished the yield contrast across the field which is usually expected to occur between soils with high and low water holding capacities.

Fig. 7 shows the optimum MZs ($n = 4$) using different clustering methods. Fig. 8 depicts the average standardized yield for optimum number of zones ($n = 4$) for k-means approach. There was a good similarity among MZs delineated by different algorithms which also matched the spatial arrangement of soil physical and hydraulic properties. The delineated zones by k-means and ISODATA-ML were similar which was expected due to similarity between these clustering methods. The dark blue (zone 4) and yellow (zone 1) colors represent soils with high and low AWC, respectively and the other two zones illustrates the transient between these extremes. The trend in average yield data across MZs matched the spatial arrangement of soil physical and hydraulic properties. For instance, higher yield belonged to soils with higher AWC and somewhat to the strip surrounded the drainage path which had good drainage condition hence did not suffer from excessive water content after heavy rainfall events which is likely to happen in west Tennessee. The yield on coarse-textured soil was 20% below average while soils with higher AWC produced 10% above average yield. In practice, small spots within clusters could easily be removed (for example by using a moving average window) to have more homogenous zones for variable rate applications.

3.2. Spatiotemporal changes in soil available water content

Temporal changes in spatial arrangement of user-defined irrigation MZs is shown in Fig. 9. The selection of breaking levels between zones was arbitrary. We chose 2 cm water within effective root zone as the target range. The highest spatial change took place from first thematic map (0–25 cm) to second one (0–50 cm). After that, there was a good agreement among thematic maps up to 100 cm. This trend suggests the possibility of having different irrigation zones early in the cropping season when the crop is only using the available water from surface soil.

3.3. Management zone delineation, phase two

Fig. 10 depicts the performance of ILP and k-means methods to delineate MZs underneath center pivots using different input attributes. The AWC variance for each pivot was considered as reference (100% level) which was equal to one MZ or uniform

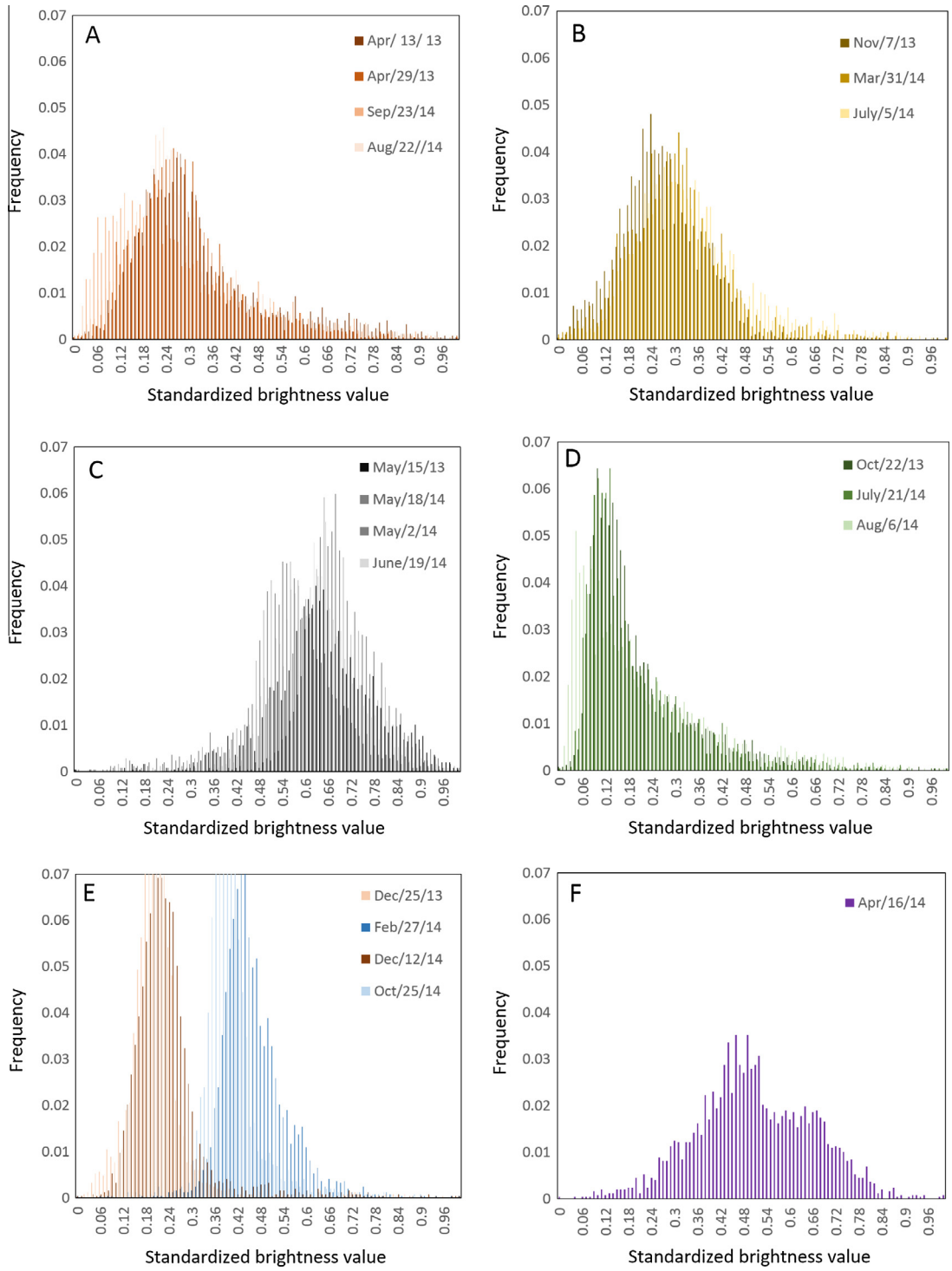

Fig. 5. Distribution of standardized brightness values for panchromatic spectral band during 2013–2014.

irrigation. Given AWC as input and *k*-means as clustering method, for both pivots within zone uniformity increased by adding more zones yet changes were minor as the number of zones became more than 4–5. The same trend was observed when ancillary data (i.e. ECa, reflectance value, and yield) were used for clustering but was less pronounced. This is in line with the observed trend when the entire field was clustered to MZs (i.e. phase one). Most of the reduction in AWC variance was obtained by dividing the pivot into two MZs. Given ECD and reflectance values as input, up to 30% and 15% of AWC variance was explained underneath the east and west pivots, respectively. Considering average yield as an input for clustering only helped to explain less than 10% of AWC variance. Given AWC as input and ILP as the zoning method, up to 30% and 40% of AWC variance was explained at the east and west pivots, respectively. In reality, this is the improvement expected to occur from uniform irrigation to VRI with limited speed control. Considering ILP as zoning method, both ECD and reflectance values were able to efficiently zone both pivots. The average yield, however, performed weaker than other inputs for clustering at the east pivot but was as effective for the west pivot. The difference between

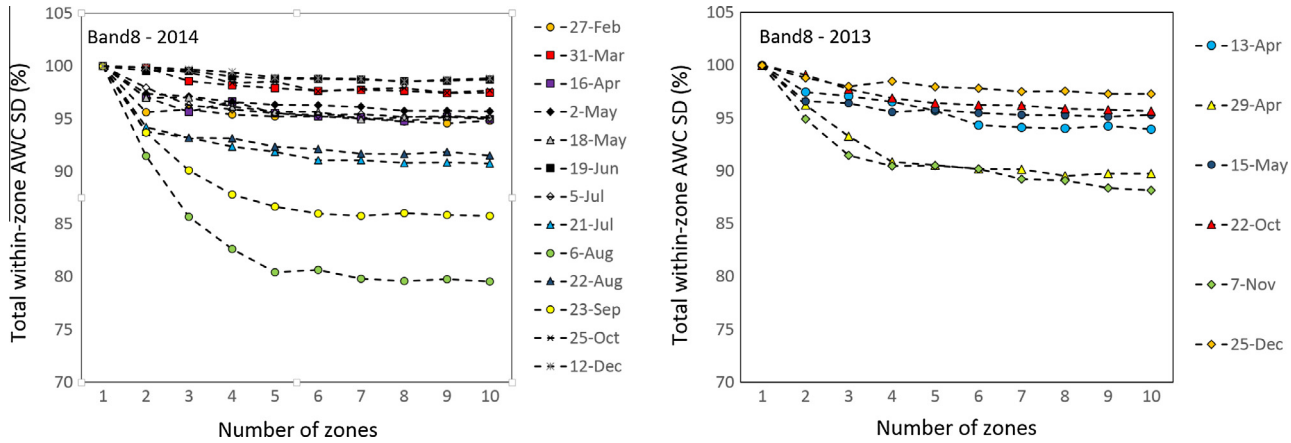

Fig. 6. AWC variance against number of zones using different satellite images. Band 8: reflectance value for panchromatic band from images taken by Landsat 8.

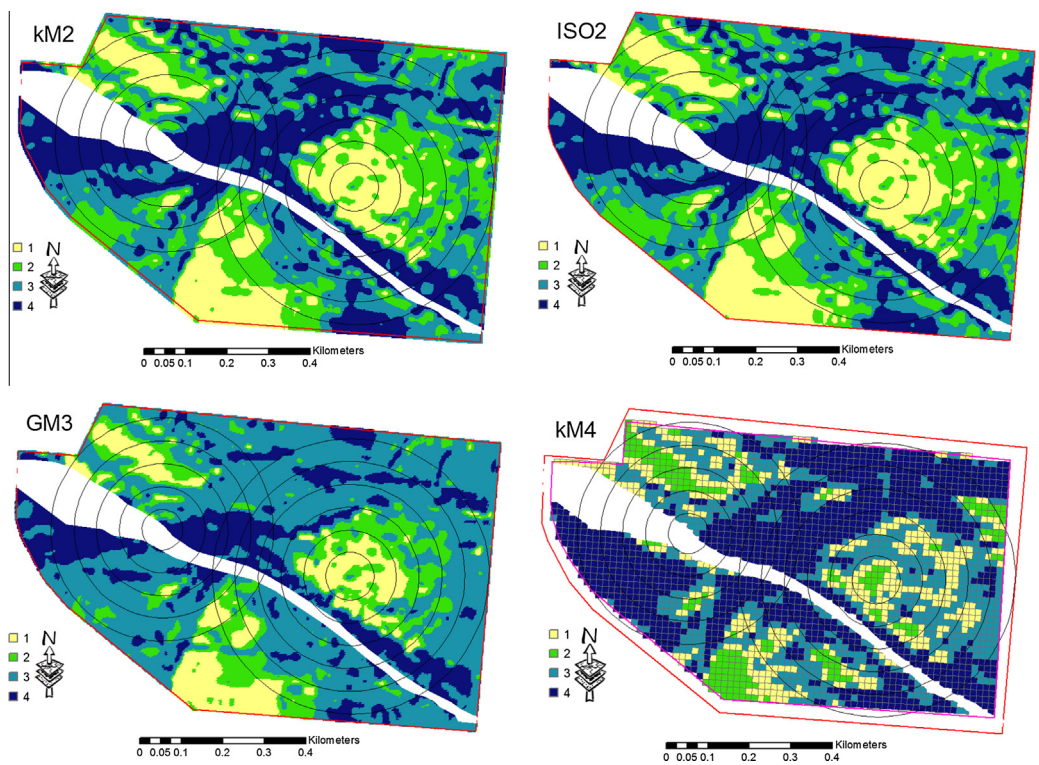

Fig. 7. Spatial arrangement of MZs. kM2: *k*-means with ECD as input; ISO2: ISO-ML with ECD as input; GM3: Gaussian mixture with ECS and ECD as inputs; kM4: *k*-means with satellite image (Landsat 8 panchromatic band) as input.

ILP and *k*-means for optimum number of zones ($n = 4$) was about 35% for both pivots in respect to reduction in AWC variance. This represents difference expected between VRI with limited speed control and a hypothetical VRI with ability to irrigate each 4 m^2 cell individually. The ability of a real precision irrigation system in reducing AWC variance is expected to be somewhere in between.

Fig. 11 represents the optimum irrigation MZs ($n = 4$) for both pivots considering different inputs and clustering methods. The user-defined MZs using AWC data were considered as the reference to evaluate the performance of other zoning strategies using *k*-means. The same was true for zoning by ILP model when zones delineated using AWC were considered to show the optimum arrangement of pies and other zoning configurations were compared against it. For ILP with AWC as input, (i) zones 1 and 4 and zones 3 and 4 covered mostly coarse soils with low AWC underneath east and west pivots, respectively; (ii) zones 3 and 2 underneath east and west pivots, respectively, mostly consisted of fine-textured soils with high AWC; (iii) other zones (zone 2 and 1 underneath east and west pivots, respectively) contained a mixture of soils with high and low AWC.

The spatial arrangement of optimum zones provided by ILP technique was not identical among models with different ancillary data as input (Fig. 11). By visual assessment, for the both pivots there is a good and moderate agreement between pie zones using ECD and Band 8, respectively and that of AWC but using yield for zoning resulted somewhat different zones. The Kappa coefficient (Fig. 12) confirmed our visual assessment. In the case of the east pivot ECD and yield data showed the highest and lowest kappa coefficient, respectively. The same trend was observed in the case of the west pivot, but there was a consistent reduction in Kappa

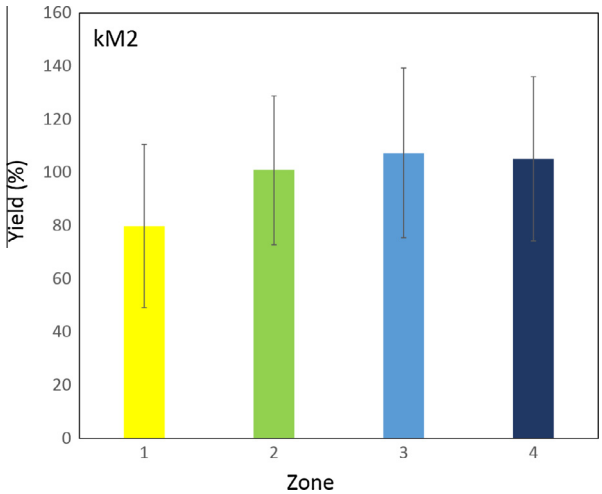

Fig. 8. Percent of average yield within MZs clustered with *k*-means considering ECD as input (bar colors are related to the zones in Fig. 7). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

coefficient for all attributes in comparison with that of the east pivot. It is clear that ECD and satellite images were good candidates for irrigation zone delineation as opposed to yield which exhibited less potential. This difference could be related to the spatial resolution of different soil attributes. The ECa data were the most exhaustive data, so carried information on soil spatial variation at a very fine scale. The satellite images were not point measurements but an integration of reflectance over 15 by 15 m cells. The AWC data were interpolated from soil cores.

4. Discussion

Recent research in precision agriculture has greatly focused on variable rate application and MZ as a method to enhance the efficiency of crop inputs usage. However, the phrase "management zones" remains uncertain unless additional information is included to clarify the goal in sub-dividing the field (Kitchen et al., 2005) which affects our choice of clustering technique and information used for zoning. The MZ delineation for variable rate application of fertilizer and seeding has extensively been studied in recent years. In contrast, there are only a handful of studies on zoning for variable rate irrigation. Given the ever increasing demand for fresh water, moving toward precise application of water for irrigation is unavoidable. The resulting zones should be simple, stable, accurate and inexpensive to identify, and enable within-field spatial variation to be managed (Khosla et al., 2010).

4.1. MZ delineation methods

Guastaferro et al. (2010) compared different algorithms for the delineation of management zones and mentioned different pros and cons for each method. We observed a good agreement among delineated zones using *k*-means, Gaussian mixture and ISODATA-ML methods. It seems for a field with high spatial structure, all well-known clustering methods are able to recognize an accurate pattern for irrigation MZs in a substantial way. The ILP zoning strategy, introduced in this study, also showed promising results. This method was designed to delineate pie shape zones for center pivots with limited speed control ability. There are several benefits associated with this zoning method. First, most of the available center pivots located in west Tennessee and other irrigated areas across US have suitable control panels to create pie shape zones, so there is no need to upgrade available pivots in order to make

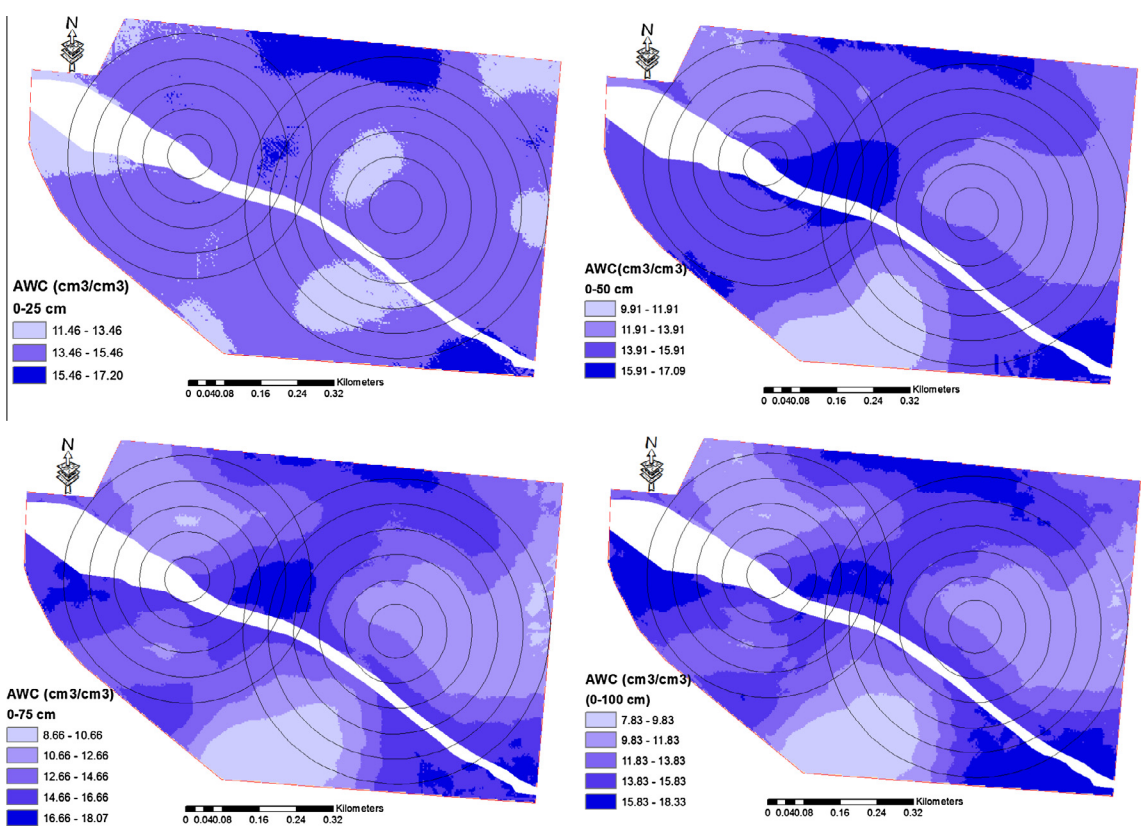

Fig. 9. Within-season temporal variation in user defined MZs considering available water content as input.

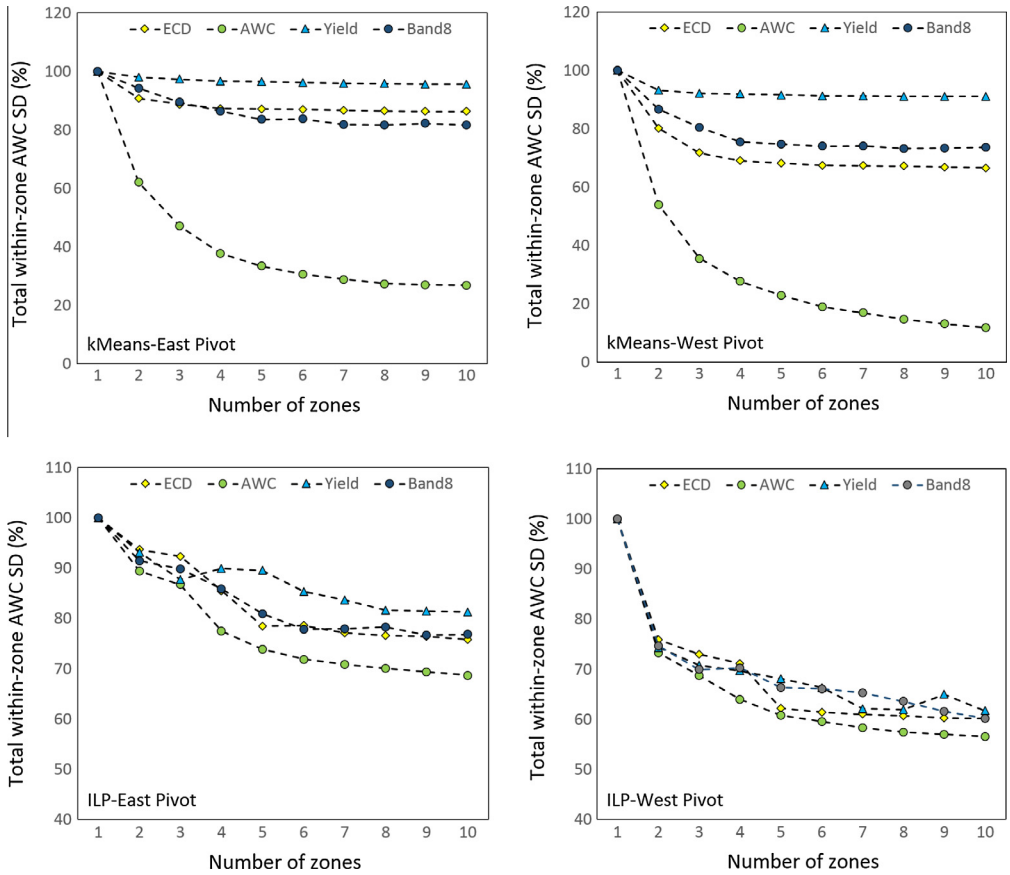

Fig. 10. Performance of irrigation MZ delineation methods based on AWC standard deviation reduction. ILP: integer linear programming; ECD: ECa deep; AWC: available water content; Band 8: reflectance value from satellite images taken by Landsat 8.

use of this method. Second, applying ILP showed a significant decline in within-zone AWC variance, meaning it will help make better site-specific irrigation decisions, and improve the choice of timing and amount of irrigation, which in turn enhances water use efficiency. In addition, it could promote VRI and act as a bridge to adopt more advanced precision irrigation systems if needed.

4.2. Optimum number of zones

It is extremely important to find the optimum number of irrigation MZs. We found 4–5 MZs to be sufficient to explain within field variability of AWC. By changing the number of zones from 1 to 10, the total within-zone variance drastically decreased but approached an asymptotic value slowly as the number of zones continued to increase. A similar trend for variance reduction was found by Zhang et al. (2010). Fraisse et al. (2001) delineated 2–6 MZs for two fields located near Centralia, Missouri, and monitored the within zone variance of yield to identify the optimum number of zones. They observed a minimum variance for 5 zones when up to 32% of the yield variance was explained.

Delineating irrigation zones is the first step toward site-specific irrigation management. Then, an irrigation prescription map needs to be developed before each irrigation event throughout the cropping season. One may only use zones as a guide to adjust the timing of supplemental irrigation in west Tennessee, i.e. to delay starting irrigation or terminate irrigation earlier in a specific zone (s). A more sophisticated scenario is to calculate the irrigation for each zone independently by sensing soil/crop water status/usage. Therefore, over estimating the optimum number of MZs makes irrigation scheduling more complicated, time-consuming and costly. On the other hand, under estimation of management zones diminishes the irrigating efficiency and may cause yield reduction. There is less freedom to explain variability with pie shape zones. It is most likely to have some zones consisting a mixture of soils with low and high AWCs. This is an unavoidable problem which is due to the inherent limitation of center pivots with limited speed control ability. The critical question is how to manage this variation or which soil to follow during irrigation. Answering this question needs quantitative information on crop-soil-water interaction which was outside of the scope of this study but is suggested for further investigation.

4.3. Application of proximal data

The AWC is a logical basis to delineate irrigation MZs since it directly affects both plant growth and yield, and irrigation scheduling. According to Fraisse et al. (2001) crop production potential is strongly related to plant available water so its productivity can be approximately determined on the basis of soil physical properties and topographic characteristics, when topography changes significantly in a given field. In the field of study more than 50% of cotton and soybean yield variability was explained by soil physical and hydraulic properties (Haghverdi, 2015). In practice, however, more easily obtained data should act as a surrogate for AWC information that is more difficult, time-consuming or expensive to obtain.

The ECa data exhibited a considerable potential for irrigation zoning. Moral et al. (2010) also found ECa to be a good proximal attribute for high resolution zone delineation. They reported ECD and percent of clay to be dominant factors explaining soil spatial

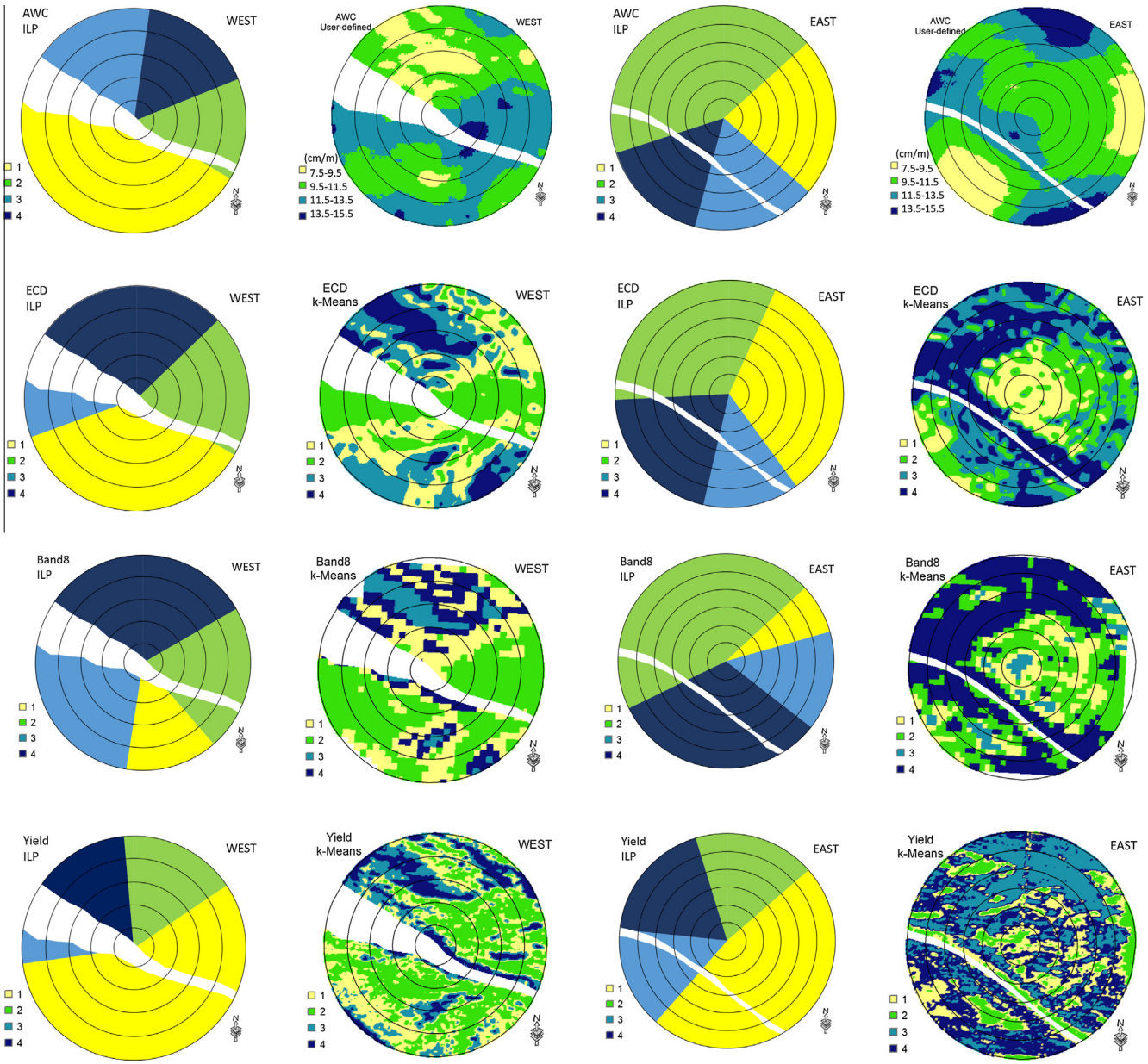

Fig. 11. Irrigation MZs for center pivot systems located at east and west parts of the field. ILP: integer linear programming; Band 8: reflectance values from panchromatic band taken by Landsat 8 satellite; ECD: deep ECa.

variability based on PCA. Kitchen et al. (2005) found ECa a reliable attribute for allocating productivity zones on claypan soil fields. Though less pronounced than ECa, the satellite images exhibited a good potential to be used for characterizing spatial variation in field-scale AWC. Our findings support the result reported by Guo et al. (2012) who observed a strong relationship between bare soil brightness and ECa with soil texture where low ECa and high brightness were associated with lower clay content. In another study, Song et al. (2009) used Quickbird commercial satellite data with 2.4 m resolution for MZ delineation and found it to be a reliable procedure. The main drawback of proximal sensor measurements is that they are complex and affected by multiple soil-crop properties (Corwin and Lesch, 2010). Therefore, they should show similar spatial distributions to those of soil physical and hydraulic characteristics, governing soil available water for crop, in order to be considered as an effective input to delineate irrigation zones. In addition, timing of capturing photos by satellite turned out to affect the practical application of satellite images for field-scale zone delineation. In fact, our results demonstrated that temporal variability alters the spatial pattern expressed by Landsat 8 panchromatic band and in turn its usefulness as an input to delineate MZs.

There were some similarities between zones delineated by ECa, satellite images and yield data, but yield data indicated less potential for zoning and exhibited some inconsistency. The reason is that different factors affect yield in a complex manner. In fact, it is not possible to distinguish between the variety of factors affecting crop growth and yield, such as physical and chemical properties of soil, irrigation regimes, pests and diseases and climate, using yield maps alone (Corwin and Lesch, 2010). We realized that even when visual assessment of yield maps indicates similar spatial arrangement as soil related maps the clustering result may be different. Moreover, for fields with no available yield map, using ECa and/or space-borne satellite images represents a great time saving.

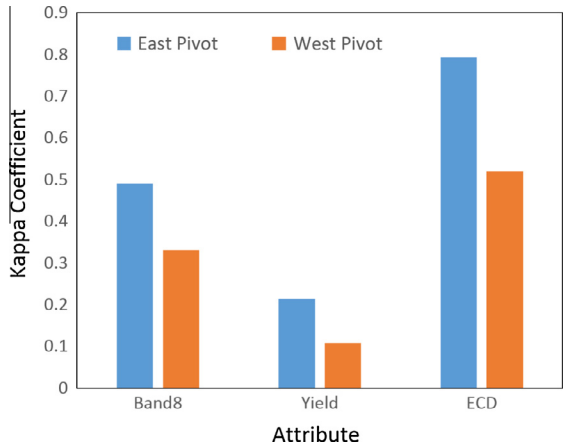

Fig. 12. Kappa coefficient for the zones delineated by the integer linear programming (ILP) method. Band 8: reflectance values from panchromatic band taken by Landsat 8 satellite; ECD: deep ECa.

Averaging yield maps across years, as we did in our study, could be a safe guard against temporal changes and an attempt to find more stable scheme of yield distribution. On the other hand, if high temporal variance in yield spatial distribution exists, averaging yield data across years may mask information and cause misleading interpretation. We averaged cotton yield data from 2012 to 2014. In 2012, the field was uniformly irrigated by the farmer but an irrigation study was performed in 2013 and 2014. Visual assessment of maps against each other suggested that spatial soil differences strongly affected yield patterns during those years, yet this impact was less pronounced for 2013 cropping season where delayed planting and temperature distribution over the growing season depressed yield and influenced its spatial distribution to some extent.

4.4. Temporal variability and role of crop

We examined the necessity of having multiple zoning schemes within a cropping season in our study site. The result revealed if in depth soil variation is significant, the spatial arrangement of optimum irrigation MZs may vary within a cropping season. There is evidence in literature showing the effect of season to season climate variability in spatial arrangement of MZs, hence supporting the idea of dynamic zoning. For instance, Schepers et al. (2004) showed that temporal climate variability even under irrigation may change the yield spatial variation. Temporal changes may affect the spatial patterns of yield such that no consistent high/low yield zones were observed (Guastaferro et al., 2010). Rainfall patterns and available heat unit are the prime factors affecting cotton growth and yield in west Tennessee with a short season humid environment. For instance, if dry periods at sensitive growth stages occur, soil with higher water holding capacity provides better conditions for crop to avoid undergoing water stress, hence higher yield is expected for such soil. In contrast, a wet year with unexpected early heavy rainfall may delay planting hence reducing the chance for the crop to accumulate enough heat units. In such scenario, soils with high AWC may suffer from excessive vegetative growth and unopened bolls by the time of harvesting. Fraisse et al. (2001) found that number of MZs for precision farming declined when either enough water was provided throughout the growing season or a drought tolerant variety was planned. The optimum number of MZs is also influenced by the crop planted (Fraisse et al., 2001). Further study is required to investigate the crop response to water level across different soil types as a means to provide crop-specific irrigation MZs.

5. Conclusion

Precision agriculture is a farming system which uses information technology to do site-specific crop management in which decisions on resource application are modified with regard to within field variation of components such as soil, water and crop (Whelan and Taylor, 2013). Field-scale spatiotemporal variation is significant in west Tennessee, thus variable rate irrigation may enhance water use efficiency and crop yield. Conventional irrigation management tries to answer when and how much to irrigate. Given within field soil variation, variable rate irrigation management ought to address where to irrigate as well. Delineating management areas within fields is an important part of a precision farming system where it is expected that applying identical treatment will cause significant yield differences. Methods for delineating MZs vary widely in the information used as well as the techniques for creating the zone boundaries. Use of on-the-go sensors and remote sensing technologies is appealing because it is easy to collect these data on a field-scale.

We evaluated the performance of several clustering methods for zone delineation. We also designed a new zoning strategy based on integer linear programming for center pivots with limited speed control ability. We used high resolution soil available water maps as standard input to zone the field. We also assessed the effectiveness of ECa, space-borne satellite images and yield data for zone delineation. The clustering methods performed similar in efficiently dividing the field into relatively homogeneous zones in respect to soil hydraulic properties. The introduced integer linear programming method offers the optimum zoning strategy for center pivots with limited speed control ability to match water input to soil spatial patterns. However, further research is essential to investigate the optimum irrigation strategy over pies with a mixture of soils. The results of this study suggested that ECa and satellite images may be used to determine site-specific irrigation MZs in west Tennessee if a high spatial similarity is observed between these ancillary data with soil hydraulic properties. The spatial and temporal resolution and precision varies among different types of ancillary data and affect their efficiency for zone delineation. Temporal variability in soil moisture altered expression of spatial variability in space-born satellite images. Yield data did not show a high potential for zone delineation. We attribute this to various factors that affect yield in a complex manner and also temporal differences in yield data. In practice, farmer knowledge on field conditions could be extremely useful to choose the most appropriate data set for zone delineation.

References

Arslan, S., Colvin, T.S., 2002. Grain yield mapping: yield sensing, yield reconstruction, and errors. Precision Agric. 3 (2), 135–154.
Basso, B., Fiorentino, C., Cammarano, D., Cafiero, G., Dardanelli, J., 2012. Analysis of rainfall distribution on spatial and temporal patterns of wheat yield in Mediterranean environment. Eur. J. Agron. 41, 52–65.
Cid-Garcia, N.M., Bravo-Lozano, A.G., Rios-Solis, Y.A., 2014. A crop planning and real-time irrigation method based on site-specific management zones and linear programming. Comput. Electron. Agric. 107, 20–28.
Cid-Garcia, N.M., Albornoz, V., Rios-Solis, Y.A., Ortega, R., 2013. Rectangular shape management zone delineation using integer linear programming. Comput. Electron. Agric. 93, 1–9.
Cohen, Jacob, 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Measur. 20 (1), 37–46.
Córdoba, M., Bruno, C., Costa, J., Balzarini, M., 2013. Subfield management class delineation using cluster analysis from spatial principal components of soil variables. Comput. Electron. Agric. 97, 6–14.
Corwin, D.L., Lesch, S.M., 2010. Delineating site-specific management units with proximal sensors. In: Geostatistical Applications for Precision Agriculture. Springer, Netherlands, pp. 139–165.

Daccache, A., Knox, J.W., Weatherhead, E.K., Daneshkhah, A., Hess, T.M., 2015. Implementing precision irrigation in a humid climate – recent experiences and on-going challenges. Agric. Water Manag. 147, 135–143.
Evans, R.G., LaRue, J., Stone, K.C., King, B.A., 2013. Adoption of site-specific variable rate sprinkler irrigation systems. Irrig. Sci. 31 (4), 871–887.
ESRI, 2014. ArcGIS 10.2.2 Help. Esri, Redlands, CA.
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 (1), 155–166.
Fridgen, J.J., Kitchen, N.R., Sudduth, K.A., Drummond, S.T., Wiebold, W.J., Fraisse, C. W., 2004. Management zone analyst (MZA). Agron. J. 96 (1), 100–108.
Guastaferro, F., Castrignanò, A., De Benedetto, D., Sollitto, D., Troccoli, A., Cafarelli, B., 2010. A comparison of different algorithms for the delineation of management zones. Precision Agric. 11 (6), 600–620.
Guo, W., Maas, S.J., Bronson, K.F., 2012. Relationship between cotton yield and soil electrical conductivity, topography, and Landsat imagery. Precision Agric. 13 (6), 678–692.
Haghverdi, A., 2015. A site-specific and dynamic modeling system for zoning and optimizing variable rate irrigation in cotton. PhD dissertation. University of Tennessee.
Haghverdi, A., Cornelis, W.M., Ghahraman, B., 2012. A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. J. Hydrol. 442, 46–54.
Haghverdi, A., Öztürk, H.S., Cornelis, W.M., 2014. Revisiting the pseudo continuous pedotransfer function concept: impact of data quality and data mining method. Geoderma 226, 31–38.
Hedley, C.B., Yule, I.J., 2009. Soil water status mapping and two variable-rate irrigation scenarios. Precision Agric. 10 (4), 342–355.
Hornung, A., Khosla, R., Reich, R., Inman, D., Westfall, D.G., 2006. Comparison of site specific management zones: soil color based and yield based. Agron. J. 98, 405–417.
Jain, A.K., 2010. Data clustering: 50 years beyond *K*-means. Pattern Recogn. Lett. 31 (8), 651–666.
Jain, A.K., Murty, M.N., Flynn, P.J., 1999. Data clustering: a review. ACM Comput. Surv. (CSUR) 31 (3), 264–323.
Khosla, R., Westfall, D.G., Reich, R.M., Mahal, J.S., Gangloff, W.J., 2010. Spatial variation and site-specific management zones. In: Geostatistical Applications for Precision Agriculture. Springer, Netherlands, pp. 195–219.
King, B.A., Stark, J.C., Wall, R.W., 2006. Comparison of site-specific and conventional uniform irrigation management for potatoes. Appl. Eng. Agric. 22 (5), 677–688.
Kitchen, N.R., Sudduth, K.A., Myers, D.B., Drummond, S.T., Hong, S.Y., 2005. Delineating productivity zones on claypan soil fields using apparent soil electrical conductivity. Comput. Electron. Agric. 46 (1), 285–308.
Mathworks, 2014. Statistics Toolbox™ User's Guide (R2014b). Retrieved November 10, 2011 from <http://www.mathworks.com/help/pdf_doc/stats/stats.pdf>.
Moral, F.J., Terrón, J.M., Silva, J.R., 2010. Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques. Soil Tillage Res. 106 (2), 335–343.
Reynolds, D., 2009. Gaussian mixture models. In: Encyclopedia of Biometrics. Springer, US, pp. 659–663.
Schepers, A.R., Shanahan, J.F., Liebig, M.A., Schepers, J.S., Johnson, S.H., Luchiari, A., 2004. Appropriateness of management zones for characterizing spatial variability of soil properties and irrigated corn yields across years. Agron. J. 96 (1), 195–203.
Singh, A., 2012. An overview of the optimization modelling applications. J. Hydrol. 466, 167–182.
Song, X., Wang, J., Huang, W., Liu, L., Yan, G., Pu, R., 2009. The delineation of agricultural management zones with high resolution remotely sensed data. Precision Agric. 10 (6), 471–487.
Thöle, H., Richter, C., Ehlert, D., 2013. Strategy of statistical model selection for precision farming on-farm experiments. Precision Agric., 1–16
Whelan, B., Taylor, J., 2013. Precision Agriculture for Grain Production Systems. CSIRO Publishing, Pages 208.
Zhang, X., Shi, L., Jia, X., Seielstad, G., Helgason, C., 2010. Zone mapping application for precision-farming: a decision support tool for variable rate application. Precision Agric. 11 (2), 103–114.