

Site-specific management zones based on geostatistical and fuzzy clustering approach in a coastal reclaimed area of abandoned salt pan

Wenquan Liu^{1*,2}, Fang Lu^{2,3}, Guanquan Chen^{1,2}, Xingyong Xu^{1,2}, and Hongjun Yu^{1,2}

¹Key Laboratory of Marine Geology and Metallogeny, Ministry of Natural Resources, Qingdao 266061, China. ^{*}Corresponding author (liuwq@fio.org.cn).

²Qingdao National Laboratory for Marine Science and Technology, Laboratory for Marine Geology, Qingdao 266061, China. ³Shandong University of Science and Technology, College of Ocean Science and Engineering, Qingdao 266590, China.

Received: 27 January 2021; Accepted: 19 April 2021; doi:10.4067/S0718-58392021000300420

ABSTRACT

Agricultural measures for cotton (*Gossypium* spp.) cultivation generally adopt a traditional and uniform management method in the reclamation area of the abandoned salt pan on the southern coast of Laizhou Bay, without considering spatial variability of the soil's physical environment. This method might lead to overfertilization in some high-nutrient plots and poor fertilization in other low-nutrient plots, which will cause high heterogeneity of crop yields and low profits. The objectives of this study were to delineate and to assess the suitability of the resulting site-specific management zones by describing spatial variability in soil attributes and crop yields. A 4.8 ha field of reclamation area in the coast was selected as the research object. A total of 12 soil attributes representing the physio-chemical attributes of soil (0-20 cm) and aboveground biomass (AGB) of cotton on 48 sampling locations were chosen from the study area. Principal component analysis and the fuzzy k-means clustering method were used to determine optimal clustering number of the study area and identify management zones. Results showed that the number of reasonable management zones were two, and confidence levels of soil organic matter (SOM) and AGB were significantly different at a probability level p < 0.01, while total N, available N, available K, cation exchange capacity, and soil salinity showed significant differences at p < 0.05. Management zones determined based on the selected soil attributes and crop AGB provide farmers with the opportunity to perform specific site management in cotton fields to achieve precision agriculture and provide a reference for formulation of a reclamation strategy for the salt pan.

Key words: Abandoned salt pan, coastal area, fuzzy clustering, geostatistical, Gossypium spp., management zones.

INTRODUCTION

Defining site-specific management zones is one of the popular research topics of precision agriculture (Hufkens et al., 2019). Management zones have been developed based on the idea of precision agriculture, that is, a way to manage soil and crops by varying inputs as needed. Understanding the spatial change relationship between soil attributes, field management measures, and crop productivity is the basis of defining management zones (Yao et al., 2014). Management zones are defined as the division of areas on the same land with similar topographic, soil, and crop characteristics (Gavioli et al., 2016). By identifying the management zones of farmland, fertilizer and irrigation input can be adjusted to improve the economic profit rate of crop production, reduce environmental risks, and achieve the purpose of "precise management". Sub-field areas are difficult to determine due to spatial variability and the dynamic interactions of the potential factors that may influence crop productivity (Fridgen et al., 2004; Moharana et al., 2019).

Furthermore, being represented areas of equal production potential, site-specific management zones have also been widely used in other fields of precision agriculture. Fleming et al. (2000) concluded that variable fertilization nutrient maps could be effectively drawn by producers based on management zones. Small and spatially uniform regions within fields may also help manage the spatio-temporal changes of soil attributes, increase fertilizer utilization efficiency (Shaddad et al., 2019) and crop yield, and facilitate crop model evaluation by linking the yield with soil and topographic parameters.

Numerous methods can be adopted for defining management zones within the field, but most methods rely on spatial information sources that are stable or predictable over time and related to crop yield (Gavioli et al., 2016). Remote sensing images (e.g., soil brightness, aerial photographs, and satellite images), topographic factors (e.g., gradient, slope length, and elevation), physio-chemical attributes of soil (e.g., soil texture, nutrition, pH, and electrical conductivity [EC]), and vield maps are considered the logical bases for defining homogeneous regions in the agricultural field (Schepers et al., 2004; Yao et al., 2014). These spatial information sources have been used to delineate subfield management zones by so many researchers. In recent years, the application of remote sensing technology in subfield management zones has become faster and more convenient with the improvement of the spatial, spectral, and temporal resolution. Rossi et al. (2018) established the relationship between the normalized difference vegetation index (NDVI) of alfalfa and proposed a management zones delineation method based on crop-soil relationship analysis. Breunig et al. (2020) used remote sensing data, such as surface reflectance and/or vegetation index (VI), to estimate the crop aboveground biomass (AGB) and then delineated management zones in agricultural fields. The topographical property and landscape location are the frequently-used factors for defining management zones, which have a good correlation with crop yield variation maps. Mieza et al. (2016) developed a simple and reliable method for delineating management zones based on the topographic position index (TPI). Spatial variations in soil physio-chemical attributes and crop productivity maps have also been used to determine consubstantial management zones. The variability of soil EC reflects the variability of soil salinity, which could be combined with other soil properties associated with crop productivity to generate maps for potential yield. Therefore, soil EC is commonly used to delineate management zones. Li et al. (2007) used soil EC, soil nutrient, cotton vield, and the NDVI value to define management zones of cotton vields in coastal cotton fields. Ortega-Blu and Molina-Roco (2016) used apparent soil electrical conductivity (EC_a), NDVI, and green vegetation index (GVI) to determine management zones in Chilean vineyards and concluded that the management zones based on the GVI performed better than those based on soil EC_a and NDVI.

The fuzzy clustering algorithm is an effective approach for detecting and analyzing the inherent structure mode of multi-source spatial data and enables researchers to understand the continuous variation phenomena occurring in nature better, making the classification results more reasonable. In recent years, the fuzzy k-means (FKM) clustering algorithm has been widely used in site-specific management zones because of its advantages, such as being a soft classification idea, ensuring a computable optimal classification number, and providing sufficient clustering information. Davatgar et al. (2012) indicated that FKM analysis provides an opportunity for farmers to adopt site-specific nutrient management by reducing within-zone variability. Yao et al. (2014) used the FKM clustering algorithm and 15 soil attributes to identify the site-specific management zones in a coastal reclaimed farmland. Ohana-Levi et al. (2019) proposed a method for determining management zones in a vineyard based on multivariate K-means clustering and machine learning.

The salt pan area has reached 420 km² on the southern coast of Laizhou Bay, North Shandong Province, China. In recent years, a large area of the salt pan has been left idle due to mining costs, market prices, and labor costs. Some salt pans were reclaimed and rented out to improve the economic benefits of enterprises. In this work, a coastal reclaimed area of abandoned salt pan on the southern coast of Laizhou Bay was selected as the study site. The soil quality and crop productivity are low in the study area, which are the typical characteristics of coastal reclamation saline soil in China (Liu et al., 2018). The study area implemented single-season cotton (*Gossypium* spp.) planting, which is also a typical planting mode in the northern coastal area of Shandong Province. Utilizing the FKM clustering algorithm, we selected 12 soil attributes and the cotton yield data to identify the crop productivity management zones. Recognizing the importance of the spatial variability of soil attributes in improving the spatial accuracy of management zones and for crop yield, the main objectives of this study were: (1) to ascertain the spatial variability of the soil attributes related to cotton yield, (2) to define management zones using the FKM clustering algorithm according to the spatial variability of the selected soil attributes, and (3) to assess the suitability of the resulting site-specific management zones.

MATERIALS AND METHODS

Overview of the study area

The study area is situated in the Zaohu saltworks reclamation area located northeast of Changyi City, North Shandong Province, China (Figure 1). The reclamation area is approximately 4.5 km away from the coastline of Bohai Sea in China, and the terrain is flat, with an average elevation of 1.7 m. The soil's parent material is dominated by sandy loam, consisting of marine sediments and alluviums. The climate of this area belongs to the warm temperate, semi-humid East Asian monsoon belt with four clear seasons. In spring, the wind is strong, and the area is prone to droughts. In summer, the temperatures are high, and the weather is rainy with occasional droughts. In autumn, the weather is mild and cool, and the area is prone to droughts at the end of the season. Finally, in winter, the weather is dry and cold, with sparse rain and snow. The mean annual precipitation is 660.1 mm, mean annual evaporation is 1859.4 mm, and mean evaporation-precipitation ratio is approximately 3.0. Before reclamation, this area was used for solar salt production throughout the year. Although a series of remediation measures have been performed after reclamation, the soil still contains a large amount of soluble salts due to long-term saline water infiltration. The high salinity and low fertility of the soil are the main limitations to the soil productivity of the region.

Reclamation history

The study area was reclaimed from the salt pan in 2012, with a length and a width of approximately 300 and 160 m, respectively. The study area was divided into four 40 m-wide zones by using salt drainage ditches. Before reclamation, the salinity of the soil exceeded 40 g kg⁻¹. In the first year of recovery, salt was leached by flood irrigation, and cotton (*Gossypium* spp.) was planted all the time. However, the yield has been meager, currently reaching 2.25 t ha⁻¹, and thus considered a low-yield field. The depth of groundwater in the study area is generally between 5.3 and 11.0 m.

Given that high salinity brine is buried in the underground phreatic layer, the layer cannot be adjusted in the rainy season due to the lack of water storage works, and the freshwater resources in the non-rainy season are scarce. Thus, the water (Cl: 887 mg L^{-1}) for farmland irrigation is mainly taken from the southern Liutuan town through an underground





pipe network. The drip irrigation under a mulch is used to meet the needs of crop growth, and the irrigation volume each time is between 450 and 600 m³ ha⁻¹. Compound fertilizers (525 kg ha⁻¹) and organic fertilizers (600 kg ha⁻¹) are used in the early cotton planting period. In 2017, the drip irrigation system did not operate due to the effects of rainfall and other factors (Figure 2).

Soil sampling

Using the mesh grid method, 48 sampling sites were determined (marked as sampling site in Figure 1) in the study area, with intervals of 15 m in the east-west direction and 50 m in the north-south direction. All sampling sites were georeferenced using a GPS with differential correction (Magellan, San Dimas, California, USA). Sampling was performed on 1 September 2017, in the early cotton harvest period. Soil samples were collected from the 0-20 cm surface layer using manual auger; 48 soil samples were collected and stored in plastic sealed bags, with the sampling site number indicated. At the same time, we collected soil bulk density of 0-10 cm in 48 soil sampling sites.

Soil analysis

The collected soil samples were naturally air-dried in the laboratory and passed through a 2 mm sieve to remove rock cuttings, intrusions and coarse organic matter before performing analyses for soil physio-chemical attributes. We identified 12 soil physical and chemical attributes related to soil productivity, and these soil attributes could help define the suitable site-specific management zones. The following methods were adopted for the measurement of various soil attributes.

The soil organic matter (SOM) was analyzed using the potassium dichromate-volumetric method. The soil salinity (EC_e) was measured by a conductivity meter (DDSJ-318, INESA Scientific Instrument, Shanghai, China) in a saturated soil paste extract. The total N (TN) was measured using the Kjeldahl method after dry combustion. The total P (TP) was analyzed by the sodium carbonate melting method. The analysis of total K (TK) was performed using the sodium hydroxide melting method. The alkaline hydrolysis diffusion method was used to determine available N (AN). The analysis of available P (AP) was measured using sodium bicarbonate (NaHCO₃) extraction and subsequent colorimetric method. Available K (AK) was determined using ammonium acetate extraction followed by flame photometry method. The cation exchange capacity (CEC) was analyzed through the ammonium replacement method. Pansu and Gautheyrou (2003) described in detail the measurement methods of the above soil attribute indicators.

The soil bulk density (ρ_b) was determined using the volumetric ring method (Grossman and Reinsch, 2002). The soil volume water content (θ_v) was calculated through the soil water content, which was measured by the gravimetric method and multiplied by the soil bulk density (Oldoni et al., 2019). The Bouyoucos Hydrometer method was used to analyze content of sand (0.02-2.00 mm), silt (0.002-0.02 mm), and clay particles (< 0.002 mm).



Figure 2. Average monthly precipitation and evaporation in the study area in 2017.

Aboveground biomass determination

The aboveground biomass (AGB) of cotton was measured 2 d before harvest. Five cotton (*Gossypium* spp.) plants were randomly selected in the 1 m \times 1 m area centered at each sampling site and cut at the bottom of the plant with scissors. Then, AGB of each cotton plant was obtained. A total of 240 AGB of cotton samples were collected. This work was completed on 1 September 2017, and the AGB of cotton was measured using oven drying at 105 °C.

Fuzzy k-means clustering algorithm

The basic idea of the fuzzy k-means (FKM) clustering method is to divide one dataset into k categories to find the iterative minimization of the objective function. The advantage of this method lies in the fact that it may provide analysis indicators with the best classification number. The objective function of the FKM method is as follows:

$$J_{f} = \sum_{i=1}^{n} \sum_{j=1}^{k} \mu_{ij}^{b} d_{ij}^{2}$$
(1)

where n denotes the number of data points; k denotes the number of categories; μ_{ij} $(1 \le i \le n, 1 \le j \le k)$ denotes the membership of the i-th soil sample x_i to the j-th clustering center C_j ; and b indicates the fuzzy weighted index $(b \ge 1)$, which controls the number of shared data among different categories. When soil data are classified, the b value between 1.2 and 1.5 may produce more reasonable classification results; d_{ij}^2 equals the square of the distance between x_i and the j-th clustering center C_j on a characteristic vector. μ_{ij} should meet the following constraints:

$$\sum_{j=1}^{n} \mu_{ij} - 1, \qquad i = 1, 2, 3, ..., n$$
⁽²⁾

If condition (3) is met, then the minimum value of Equation 2 may be calculated. If the partial derivative of J_f to m_j and μ_{ij} is 0, then the necessary conditions could be obtained.

$$m_{j} = \frac{\sum_{i=1}^{n} \mu_{ij}^{b} x_{i}}{\sum_{i=1}^{n} \mu_{ij}^{b}} \qquad j = 1, 2, 3, ..., k$$
(3)

$$\mu_{ij} = \frac{(1/||x_i - m_j||^2)^{1/(b-1)}}{\sum_{j=1}^{k} (1/||x_i - m_j||^2)^{1/(b-1)}} \qquad i = 1, 2, 3, ..., n$$
(4)

Equations 3 and 4 are solved using an iterative method, that is, the FKM algorithm (Chang et al., 2011).

The suitability of the number of management zones identified in a plot was determined by looking at the spatial variation and distribution characteristics of the plot itself. Two evaluation indicators were introduced to determine the optimal number of zones. The fuzzy performance index (FPI) represents the number of shared members among different categories and is used to measure the separation degree between k categories (Fridgen et al., 2004). The FPI ranges from 0 to 1. If the value is near 0, then different zones share few members, that is, the different management zones have large differences. If the value is near 1, then the zones share many members, that is, the different management zones have small differences.

$$FPI = 1 - \frac{k}{k-1} \left[1 - \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{k} \mu_{ij}^{2} \right]$$
(5)

The modified partition entropy (MPE) indicates the destruction degree of data organizations (or similarities) caused by dividing a dataset into different categories. The MPE ranges from 0 to 1. The smaller the MPE the higher the similarity between the soil attribute values within each management zone after the entire plot is divided, that is, the lower the variance inside the same area.

$$MPE = -\frac{1}{n-k} \sum_{i=1}^{n} \sum_{j=1}^{k} \mu_{ij} lg(\mu_{ij})$$
(6)

Data statistics and analysis

Given that the information reflected by the soil attributes was superimposed to some extent, the attributes to be included must be determined before applying the FKM clustering analysis. The principal component analysis (PCA) method was adopted in this study to group the selected attributes as statistical factors. The statistical factors with a characteristic root over 1 and a cumulative contribution rate over 85% were retained to perform the FKM clustering algorithm to divide the

management zones. The Fuzzy ME 3.0 software was used to calculate the FKM clustering value (Minasny and McBratney, 2002). The maximum number of iterations was 300, the convergence threshold was 0.001, and the fuzzy weighted index was 1.5. To find the optimum number of zones, 2, 3, 4, 5, 6, 7, 8, 9, and 10 categories were produced, and the FPI and MPE values of each category were calculated and compared. The classical statistical method was used to calculate the description statistics of all soil attributes. The normal distribution of each soil attribute was tested by the single-sample Kolmogorov-Smirnov (KS) method. The multiple regression model of cotton AGB and soil properties was fitted by stepwise multivariate regression analysis method. We evaluated the suitability of the management zones and made a statistical comparison of the soil attributes of the management zones using ANOVA. The SPSS 19.0 (IBM, Armonk, New York, USA) software was used to perform statistical analysis. The GS+10.0 software package was employed to fit the semi-variance map of the soil attributes (Robertson, 2008). The spatial distribution map of soil attributes was drawn by ordinary kriging method in ArcGIS 10.2 software (Esri, Redlands, California, USA).

RESULTS

Descriptive statistics analysis of soil attributes and aboveground biomass

Table 1 shows the descriptive statistical results of soil physio-chemical attributes and the crop AGB from 48 sampling sites in the study area. The soil in the study area is characterized by low organic matter, low nutrient, low water content and high salinity. The SOM plays a vital role in improving agglomerate stability, water permeability, and nutrient effectiveness (Bogunovic et al., 2018). The SOM content of the study area was only 7.53 g kg⁻¹, which is consistent with the research results of Li et al. (2007) in similar areas. The TN, TP, TK, AN, AP, AK contents and CEC of soil were 0.48 g kg⁻¹, 0.73 g kg⁻¹, 9.13 g kg⁻¹, 73.75 mg kg⁻¹, 33.98 mg kg⁻¹, 483.96 mg kg⁻¹, and 6.42 cmol kg⁻¹, respectively, which were far below the recommended level by the Chinese National Soil Survey Office (CNSSO, 1998), except for the TK and AK. Similar results were also obtained by Yao et al. (2014). Although the study area has been reclaimed for more than 5 yr and is still in the process of soil desalination evolution, soil EC_e was as high as 243.63 mS m⁻¹, exceeding the universal acceptance limit for most crops. The soil θ_v was only 11.27%, which is closely related to the soil texture. The study area is dominated by sandy loam, which has good air permeability and can easily lose water. The clay content and the ρ_b were 0.19% and 1.65 g cm⁻³, respectively.

Concerning the coefficient of variation (CV), TK and ρ_b showed a weak variability, and the CV values are less than 10%, while the CV values of the other soil attributes were between 10% and 100%. The one-sample KS normality test (p < 0.05, two-tailed) results show that all the soil attributes matched the normal distribution fairly well. Kriging interpolation provides the best estimate value only when the soil attributes are under a normal distribution (Zovko et al., 2018).

Variables	Min.	Mean	Median	Max.	SD	CV, %	Skewness	Kurtosis	K-Sp
AGB, g plant ⁻¹	52.00	134.00	131.00	262.00	41.20	30.75	0.71	1.07	0.70
SOM, g kg ⁻¹	4.42	7.53	7.21	11.70	1.65	21.93	0.54	-0.02	0.40
TN, g kg ⁻¹	0.30	0.48	0.48	0.62	0.07	15.13	-0.50	-0.07	0.53
TP, g kg ⁻¹	0.49	0.73	0.75	1.07	0.13	17.11	0.17	-0.09	0.97
TK, g kg ⁻¹	7.05	9.13	9.11	10.90	0.83	9.11	-0.22	0.13	0.53
AN, mg kg ⁻¹	35.30	73.75	75.07	113.00	16.87	22.88	-0.13	0.04	0.93
AP, mg kg-1	15.60	33.98	34.70	66.00	12.82	37.74	0.69	0.21	0.42
AK, mg kg ⁻¹	264.00	483.96	490.50	767.00	118.04	24.39	0.22	-0.39	0.98
CEC, cmol kg-1	4.23	6.42	6.51	8.96	1.05	16.29	-0.17	-0.24	0.82
ECe, mS m ⁻¹	186.00	243.63	244.50	325.00	28.03	11.50	0.35	0.40	0.92
$\theta_{\rm v}, \%$	7.13	11.27	11.44	15.20	2.15	19.05	-0.23	-0.73	0.67
ρ_b , g cm ⁻³	1.32	1.65	1.66	1.79	0.09	5.70	-1.18	2.25	0.60
Clay, %	0.07	0.19	0.21	0.41	0.07	34.17	0.58	1.10	0.43

Table 1. Descriptive statistics of soil physio-chemical attributes and cotton yield.

Min.: Minimum; Max.: maximum; SD: standard deviation; CV: coefficient of variation; K-Sp: significance (two-tailed) of one-sample Kolmogorov-Smirnov test; AGB: aboveground biomass; SOM: soil organic matter; TN: total N; TP: total P; TK: total K; AN: available N; AP: available P; AK: available K; CEC: cation exchange capacity; EC_e: electrical conductivity of saturated paste extract; θ_v : volumetric water content; ρ_b : bulk density; Clay: content of clay particles.

Table 2 presents the correlation analysis of the 12 soil physio-chemical attributes and the cotton yield. The results show that the cotton AGB was significantly, positively correlated with SOM, TN, AN, and CEC and a significant negative correlation with EC_e and the ρ_b , indicating that these are the main factors limiting cotton growth. The soil EC_e was negatively correlated with SOM and most of the other nutrient indicators. The CEC was significantly, positively correlated with SOM, TN, and the other nutrient indicators. Furthermore, a positive correlation existed between the nutrient attributes.

Multiple regression analysis of crop AGB and soil physio-chemical attributes

To explore the effect of soil attributes on the cotton yield, regression analysis was performed between the soil attributes and the crop AGB. Multiple regression analysis was conducted using the 12 soil attributes as independent variables and the cotton AGB as a dependent variable. Since there was no information on the correlation between soil properties and crop AGB in this area, stepwise multiple regression analysis was carried out. A multiple regression model for the cotton AGB and the soil attributes in the study area was developed.

 $AGB - 9.042 + 1.448SOM - 0.602EC_e + 0.765AN + 0.772CEC - 1.335\rho_b + 0.102AK (N = 48, r^2 = 0.71)$ (7) As can be seen from Equation 7, the r² is 0.71, which means that the model only uses six soil attributes to explain more than 70% of the total variation of cotton AGB. Apparently, the values of standardized coefficients of soil attributes showed that the largest negative effect on cotton AGB was EC_e, reaching -0.409, and the largest positive effect on cotton AGB was AN, reaching 0.313. The SOM, AN, CEC, and AK contributed positively to the cotton AGB, and an increase in these attributes would contribute to the promotion of cotton yield in the salt pan reclamation area. The soil EC_e and ρ_b had a negative effect, implying that their increase would greatly reduce the crop yield. Yao et al. (2014) and Behera et al. (2018) also found that soil salinity has a significant negative correlation with the crop yield in coastal farmlands.

Table 2. Correlation among measured soil physical-chemical attributes and cotton yield in the study area (n = 48).

	AGB	SOM	TN	TP	TK	AN	AP	AK	CEC	ECe	$\theta_{\rm v}$	$\rho_{\rm b}$	Clay
AGB	1												
SOM	0.439**	1											
TN	0.406**	0.400**	1										
TP	0.220	0.150	0.337*	1									
TK	0.238	0.120	0.205	0.046	1								
AN	0.349*	0.368*	0.209	0.038	0.274	1							
AP	0.246	0.211	0.184	0.368*	0.070	0.133	1						
AK	0.216	0.279	0.069	0.102	0.463**	0.390**	0.243	1					
CEC	0.364*	0.294*	0.358*	0.256	0.241	0.013	0.010	0.0606	1				
ECe	-0.407**	-0.336*	-0.206	-0.060	-0.150	-0.173	-0.151	0.117	-0.1700	1			
$\theta_{\rm v}$	-0.071	-0.102	-0.099	-0.020	0.130	-0.278	0.048	0.039	0.195	0.012	1		
$\rho_{\rm b}$	-0.366*	-0.058	-0.127	-0.166	-0.027	-0.231	-0.291*	0.212	-0.046	0.308*	0.178	1	
Clay	-0.257	-0.009	-0.153	0.074	0.114	0.183	0.191	0.254	-0.009	0.223	0.050	0.018	1

*, **Significant at P < 0.05 and P < 0.01, respectively.

AGB: Aboveground biomass; SOM: soil organic matter; TN: total N; TP: total P; TK: total K; AN: available N; AP: available P; AK: available K; CEC: cation exchange capacity; EC_e: electrical conductivity of saturated paste extract; θ_v : volumetric water content; ρ_b : bulk density; Clay: content of clay particles.

Table 3. Regression model	parameters	between soil a	attributes and	cotton al	boveground	biomass	(AGB)	(n = 48)	3).
---------------------------	------------	----------------	----------------	-----------	------------	---------	-------	----------	-----

Indicators	Unstandardi	zed coefficients	Standardized coefficients	t (df =47)	Significance level
	α	Std. error	β		
Constant	9.042	2.782	-	3.368	0.002**
SOM	1.448	0.545	0.258	2.142	0.038*
ECe	-0.602	0.178	-0.409	-3.416	0.001**
AN	0.765	0.294	0.313	2.608	0.012*
CEC	0.772	0.303	0.248	2.155	0.037*
ρ _b	-1.335	0.536	-0.396	-3.491	0.001**
AK	0.102	0.041	0.293	2.396	0.021*

*, **Significant at P < 0.05 and P < 0.01, respectively.

SOM: Soil organic matter; EC_e : electrical conductivity of saturated paste extract; AN: available N; CEC: cation exchange capacity; ρ_b : bulk density; AK: available K.

Geostatistical analysis and interpolation

Figure 3 shows the semivariograms of the six soil attributes and their fitting model parameters. The Gaussian model fitted the semivariance of the SOM, EC_e , AN, and AK well, while the exponential model fitted the semivariance of CEC and ρ_b . The nugget-to-sill ratio (C_o/C), expressed in percentage, can represent the spatial variation extent of the regionalized variables. Strong spatial dependency was observed for all the six soil attributes in Figure 3. The fitting semi-variance model of each attribute was examined by the determination coefficient (r^2) and the residual sum of squares (RSS). That is, the optimal fitting model has a r^2 near 1 and a RSS near 0. The results show that the fitted semivariogram of all the soil attributes is quite satisfactory.

Figure 4 shows the six soil attributes distribution map generated by ordinary kriging method according to the parameters in Figure 3. It can be seen from Figure 4 that the spatial distribution of soil attributes is zonal; EC_e and ρ_b exhibit consistent spatial distribution characteristics, with high content in the southern and northern parts and a low content in the middle of the study area. The SOM is low in the northeastern, northwestern, and southwestern corners. The AK is low in the northeast part, AN is high in the western region, and CEC is high in the south.

Principal component analysis

Due to the correlation between the raster data used for cluster analysis of the six soil attributes, the information reflected by each index was overlapped to a certain extent. Therefore, in this paper, the PCA was applied to concentrate the useful information in the original raster image into as few new principal component images as possible to enhance the raster image information, meanwhile the components with eigenvalues greater than 1 will be retained for subsequent analysis. Table 4 illustrates the PCA results of the six soil attributes.

Table 4 presents the three principal components (PCs) identified, explaining 75.027% of the total variance. They represent 75.027% of the original information that reflects the variation in cotton growth and soil attributes; thus, their credibility is relatively higher. The commonality indicates that the original variable variance can be explained by common components. The greater the commonality is, the higher the variance of the variable that can be explained by the components will be. According to the Table 4, commonalities show that the three PCs explained 67.7% of the variance in SOM, 79.5% of the variance in AN, 78.4% of the variance in AK, 77% of the variance in CEC, 63% of the variance in EC_e, and 72.5% of the variance in ρ_b . According to Table 4, the first PC explains 31.81% of the variance, and the proportion coefficients of SOM and AN are significantly higher than those of the other components, reaching 0.784 and 0.702, respectively. Thus, the first PC could be identified as the "organic nutrient supply component". The second

Figure 3. Semivariogram and model parameters for six soil attributes. Where Co is nugget value, Co+C is sill value, Co/(Co+C) is nugget-to-sill ratio and RSS is residual sum of squares.



SOM: Soil organic matter; AN: available N; AK: available K; CEC: cation exchange capacity; EC_e: Electrical conductivity of saturated paste extract; ρ_b : bulk density.

PC explains 23.616% of the variance, and the proportions of AK, ρ_b , and EC_e are the largest, indicating that the second PC is a comprehensive reflection of the AK, the ρ_b , and the EC_e. The third PC explains 19.601% of the variance, and the proportion of CEC is the largest, indicating that the third PC is a comprehensive reflection of the CEC. The factor score of each principal component can be obtained by Equations 8-10:

$$PC1 = 0.567SOM + 0.508AN + 0.300AK + 0.303CEC - 0.415EC_{e} - 0.257\rho_{b}$$
(8)

$$PC2 = 0.103SOM + 0.201AN + 0.653AK - 0.079CEC + 0.458EC_{e} + 0.554\rho_{b}$$
(9)

$$PC3 = 0.214SOM - 0.482AN - 0.087AK + 0.745CEC - 0.061EC_e + 0.394\rho_b$$
(10)

In summary, three principal components were extracted, and the cumulative contribution rate was 75%. They represent the original information reflecting the variation of soil characteristics, with high reliability. Figure 5 shows the spatial distribution of the three PCs. According to Figure 4, SOM and AN are dominant in the overall spatial variation of PC1. Variables AK, ρ_b and EC_e accounted for the majority of the spatial variation of PC2, and PC3 was the main representative of the spatial variation of CEC.

Figure 4. Spatial distribution map of six soil attributes. Where SOM is soil organic matter, AN is available N, AK is available K, CEC is cation exchange capacity, EC_e is electrical conductivity of saturated paste extract, and ρ_b is bulk density.



Table 4. Principal component analysis of the measured soil attributes.

	PC1	PC2	PC3	Communalities
SOM	0.784	0.123	0.220	0.677
AN	0.702	0.239	-0.495	0.795
AK	0.415	0.777	-0.089	0.784
CEC	0.418	-0.094	0.766	0.770
ECe	-0.573	0.545	-0.063	0.630
ρ_{b}	-0.355	0.660	0.405	0.725
				Total
Eigenvalue	1.909	1.417	1.056	
% Variance explained	31.81	23.616	19.601	75.027

SOM: Soil organic matter; AN: available N; AK: available K; CEC: cation exchange capacity; EC_e: electrical conductivity of saturated paste extract; ρ_b : bulk density.

Figure 5. Spatial distribution maps of the three principal components (PC) in the study area.



FKM clustering analysis

In order to facilitate the calculation of FKM algorithm, we first convert the data format, the three principal component images in Figure 5 were converted from raster data to vector point data in the ArcGIS 10.2 software. The FKM algorithm was used to analyze the point data obtained and determine the appropriate class. In order to find out the most suitable number of management zones, 2, 3, 4, 5, 6, 7, 8, 9 and 10 categories were generated respectively. The values of FPI and MPE of each class were calculated and compared. The smaller the two indexes, the better the partition results. Figure 6 shows the results of the cluster numbers according to the two indexes.

Figure 6 shows that the FPI and the MPE did not monotonically increase or decrease with the increase in the number of clusters, but followed a similar variation trend. It can be seen from Figure 6, the generated FPI and MPE values were the smallest when number of clusters is two, indicating that the variance within the zone was the lowest, and the variation between the zones was the largest. Therefore, the optimal number of clusters identified is two in the study area. In individual cases, when the trends of the FPI and the MPE are inconsistent with the number of clusters, the final decision on the number of clusters to be used to create management zones requires additional verification. Such as in the situation of developing productivity zones, the cluster numbers might be verified by comparing the within-zone yield variation as one increases the number of clusters (Fridgen et al., 2004).

Figure 6. Calculated fuzzy performance index (FPI) and modified partition entropy (MPE) for study area.



Figure 7 presents the result of management zones in the study area (left graph), which was divided into two zones. For the convenience of statistics and comparison, the standardized cotton AGB data interpolation was generated into a $4 \text{ m} \times 4 \text{ m}$ grid unit using the ordinary kriging method to examine whether these potential management zones could accurately reflect the cotton yield levels in the study area (right graph). Apparently, the spatial distribution pattern of the management zones map obtained by aggregating the information of the six soil attributes is very similar to that of the cotton yield. The productivity in the northeastern and southwestern parts of the study area is generally higher than that in the other areas.

Evaluating the suitability of management zones

In order to evaluate whether the two management zones determined can accurately reflect the spatial variability of soil attributes and the AGB of cotton, all soil sampling points and crop yield data points in the study area were divided into two independent management zones according to Figure 7. A total of 26 out of the 48 sampling sites in the study area were assigned to management zone 1, while the remaining 22 were assigned to management zone 2. ANOVA was used to compare the differences in all soil attributes and the AGB of cotton between the management zones. Table 5 summarizes the ANOVA descriptive statistics results of all the soil attributes and the AGB of cotton between the management zones.





Table 5. Descriptive statistics of soil attributes and cotton aboveground biomass (AGB) for two management zones.

Management zone	Variables	SOM	TN	TP	TK	AN	AP	AK	CEC	EC _e	$\theta_{\rm v}$	$ ho_{b}$	Clay	AGB
	Min.	4.42	0.30	0.49	7.05	38.20	15.60	287.00	4.48	208.00	7.13	1.44	0.11	52.00
	Max.	10.40	0.56	1.07	10.90	109.00	66.00	767.00	8.96	325.00	15.20	1.78	0.41	172.00
Zone 1 (n = 26)	Mean	6.90	0.46	0.71	9.17	69.82	33.87	489.73	6.29	249.38	11.68	1.66	0.21	110.46
	SD	1.35	0.06	0.15	0.89	15.12	13.45	110.08	1.18	26.00	2.33	0.09	0.06	25.33
	CV, %	19.54	14.08	20.43	9.73	21.66	39.71	22.48	18.78	10.42	19.93	5.47	31.20	22.93
Zone 2 (n = 22)	Min.	5.15	0.33	0.58	7.73	35.30	16.10	264.00	4.23	186.00	7.32	1.32	0.07	76.00
	Max.	11.70	0.62	0.92	10.42	113.00	62.20	722.00	7.75	286.00	13.26	1.79	0.34	262.00
	Mean	8.27	0.58	0.76	9.07	78.39	34.11	470.14	6.98	226.82	10.79	1.63	0.18	161.82
	SD	1.69	0.08	0.09	0.77	17.98	12.36	129.11	0.86	29.39	1.85	0.10	0.07	39.23
	CV, %	20.44	13.79	12.33	8.48	22.93	36.22	27.46	12.38	12.96	17.16	5.93	38.24	24.24
Significance level	р	0.00**	0.03*	0.05*	0.69	0.04*	0.25	0.03*	0.04*	0.04*	0.16	0.34	0.30	0.00**

*, **Significant at P < 0.05 and P < 0.01, respectively.

Min.: Minimum; Max.: maximum; SD: standard deviation; CV: coefficient of variation; SOM: soil organic matter; TN: total N; TP: total P; TK: total K; AN: available N; AP: available P; AK: available K; CEC: cation exchange capacity; EC_e: electrical conductivity of saturated paste extract; θ_v : volumetric water content; ρ_b : bulk density; Clay: content of clay particles; AGB: aboveground biomass.

The results in Table 5 show that the confidence levels of SOM and AGB of cotton in different management zones were significantly different at a probability level of p < 0.01, while TN, AN, AK, CEC, and EC_e showed significant differences at p < 0.05. The comparison of the mean values of the soil attributes and the AGB of cotton in the two management zones indicates that management zone 2 had the highest soil fertility level and crop production capacity, the highest AGB of cotton, and the lowest salt content. The indicators reflecting soil nutrient and fertility levels, such as SOM, TN, TP, AN, and AP, were significantly lower in the management zone 1 than those in the management zone 2, while EC_e and the soil clay content, which negatively affect the cotton yield, were lower. Furthermore, the AK content in the management zone 2 was lower than that in the management zone 1. The statistical results in Table 5 showed that there were significant differences in soil attributes and cotton AGB values between the two management zones. Therefore, different soil attributes, for example SOM, AN, AK, and EC_e, can be used to determine management zones as they represent the spatial variability of soil physio-chemical attributes and crop yield base on the FKM clustering analysis.

DISCUSSION

Spatial information sources play a decisive role in the determination of management zones

The delineation of site-specific management zones relies on various spatial information, such as soil physio-chemical attributes, topographic attributes, remote sensing images, crop canopy images, and yield maps (Schepers et al., 2000). All the spatial information has the characteristics of stability or predictability over space and time and a certain correlation with crop yield. Previous studies have proposed dozens of spatial information to delineate management zones (Yao et al., 2014; Oldoni et al., 2019; Shaddad et al., 2019). Based on such studies, PCA and FKM clustering approach were adopted in combination with the coastal characteristics of the study area to establish a database, including 12 soil physiochemical attributes. The delineation of site-specific management zones was developed by using this database in the salt pan reclamation area. In this paper, most of the selected soil and crop attributes were suggested by Fridgen et al. (2004), Li et al. (2007), and Yao et al. (2014). Meanwhile, some commonly used data sources were not included, such as elevation data. The study area is located on a low coastal plain. The terrain is flat, and the area is small. Long-term field cultivation management has made the interior study area free of fluctuations. Moreover, considering the small extension of the study area and the uniformity in soil texture and soil development, soil brightness or aerial photos were not used as data sources in our study, which have advantages in distinguishing soil sequences. The three PCs explaining 75.027% of the total variance were identified from the selected soil attributes (i.e., SOM, AK, AN, CEC, EC_e and ρ_b) as organic nutrients, soil salinization, and soil compaction. Furthermore, the proportion of explainable variations in crop AGB variation (75.027%) is lower than the universally adaptive value (85%), showing that some soil attributes that may be related to crop yield were not considered.

Effect of management zone on selected soil attributes and crop AGB

Our study found that the SOM, TN, AN, AK, CEC, EC_e and AGB of the management zones was significantly different. However, these differences were not observed in practical field fertilization and farming management. In particular, the content of SOM, which plays an important role in crop yield, showed a significant difference between the two management zones. The comparison of the spatial distribution of SOM in Figure 4 with that of the AGB of cotton in Figure 6 obviously indicates that they had a similar spatial distribution characteristic, also proving the suitability of the division of the area into management zones. The values of EC_e in the management zone 2 were significantly lower than those in the management zone 1 because EC_e had negative impacts on cotton yield (Shaddad et al., 2019), and possibly because the soil texture (clay content) and microtopography (low-lying terrain) in the management zone 1 affected its magnitude. The soil fertility level indicators, such as TN, AN, and AK, were lower in the management zone 1 than in the management zone 2. These might be related to the pore structure, soil texture, stability of the agglomerates, and even to the soil moisture condition in the two zones under the same field management (Shaver et al., 2013). The AK has a negative effect on crop AGB and a positive correlation with the soil salinity content (Yao et al., 2014), which is significantly higher in the management zone 1 than in the management zone 2; this finding is consistent with the results of Fu et al. (2010). In this paper, six soil attributes and crop AGB data were collected to determine the management zones in the study area. However, different factors have complex interactions, such as soil attributes, microtopography, weather conditions, and field management, and the crop yield may vary greatly in the same field over time. Therefore, soil attributes, meteorological, and crop productivity data for multiple years are needed to further determine the most appropriate management zones.

Implications for field management of reclamation area

Management zones are one of the key means to implementing precise agricultural variables management. Farmers can formulate scientific within-field management strategies for different management zones and perform precise fertilization, irrigation, and cultivation measures accordingly. To ensure that management is adapted to local conditions, the management zones should have a simple development, complete functions, and economic feasibility, while a complex field assessment and data manipulation might be unreasonable from the perspective of time, benefits, or economy (Li et al., 2007). This study used 12 attributes, including soil attributes and the AGB of cotton, to delineate the management zones of a salt pan reclamation area based on geostatistics, the PCA method, and FKM cluster analysis and evaluated its appropriateness and effectiveness. The results show that the study area could be divided into two management zones, which could reflect the level of crop productivity potential. The delineation of precise management zones must often consider various factors or variables. In this study, the data on spatial variations in soil physical and chemical attributes and cotton yield were used to determine the management zones from a single year, without considering the temporal variability or stability of these variables. The characteristics of the management zones (spatial continuity or aggregation and scale size) are correlated not only to the spatial autocorrelation level of the soil or crop growth status but also to the spatial scale of the selected data sources (i.e., the size of the space unit). The impact of spatial scale on the management zones will be verified further in follow-up studies.

CONCLUSIONS

Our results indicated that the soil organic matter, available N, available P, available K, cation exchange capacity, soil conductivity and soil bulk density could be used to delineate management zones. The optimal number of management zones for the reclaimed area of abandoned salt pan was two. Variance analysis results showed that the most of soil attributes and cotton aboveground biomass presented statistical differences in each defined management zone.

ACKNOWLEDGEMENTS

This research was funded by the NSFC Shandong Joint Fund for Marine Science Research Centers (U1806212). We are very grateful to the staff of the key laboratory of marine geology and metallogeny in providing the measured data. We would also like to thank Zaohu Farm for the convenience of on-site sampling. Finally, we sincerely thank the anonymous reviewers for their constructive comments.

REFERENCES

- Behera, S.K., Mathur, R.K., Shukla, A.K., Suresh, K., and Prakash, C. 2018. Spatial variability of soil properties and delineation of soil management zones of oil palm plantations grown in a hot and humid tropical region of southern India. Catena 165:251-259.
- Bogunovic, I., Trevisani, S., Pereira, P., and Vukadinovic, V. 2018. Mapping soil organic matter in the Baranja region (Croatia): Geological and anthropic forcing parameters. Science of Total Environment 643:335-345.
- Breunig, F.M., Galvão, L.S., Dalagnol, R., Dauve, C.E., Parraga, A., Santi, A.L., et al. 2020. Delineation of management zones in agricultural fields using cover-crop biomass estimates from PlanetScope data. International Journal of Applied Earth Observation and Geoinformation 85:102004.
- CNSSO. 1998. China soils. Chinese National Soil Survey Office (CNSSO). China Agriculture Press, Beijing, China. (In Chinese).
- Chang, C.T., Lai, J.Z.C., and Jeng, M.D. 2011. A fuzzy k-means clustering algorithm using cluster center displacement. Journal of Information Science and Engineering 27:995-1009.
- Davatgar, N., Neishabouri, M.R., and Sepaskhah, A.R. 2012. Delineation of site-specific nutrient management zones for a paddy cultivated area based on soil fertility using fuzzy clustering. Geoderma 173-174:111-118.
- Fleming, K.L., Westfall, D.G., Wiens, D.W., and Brodah, M.C. 2000. Evaluating farmer developed management zone maps for variable rate fertilizer application. Precision Agriculture 2:201-215.

- Fridgen, J.J., Kitchen, N.R., Sudduth, K.A., Drummond, S.T., Wiebold, W.J., and Fraisse, C.W. 2004. Management Zone Analyst (MZA): Software for subfield management zone delineation. Agronomy Journal 96(1):100-108.
- Fu, Q., Wang, Z.L., and Jiang, Q.X. 2010. Delineating soil nutrient management zones based on fuzzy clustering optimized by PSO. Mathematical and Computer Modelling 51(11-12):1299-1305.
- Gavioli, A., Souza, E.G.D., Bazzi, C.L., Guedes, L.P.C., and Schenatto, K. 2016. Optimization of management zone delineation by using spatial principal components. Computers and Electronics in Agriculture 127:302-310.
- Grossman, R.B., and Reinsch, T.G. 2002. The solid phase: bulk density and linear extensibility. p. 201-228. In Dane, J.H. and Topp, G.C. (eds.) Methods of soil analysis: Physical methods. Soil Science Society of America, Madison, Wisconsin, USA.
- Hufkens, K., Melaas, E.K., Mann, M.L., Foster, T., Ceballos, F., Robles, M., et al. 2019. Monitoring crop phenology using a smartphone based near-surface remote sensing approach. Agricultural and Forest Meteorology 265:327-337.
- Li, Y., Shi, Z., Li, F., and Li, H.Y. 2007. Delineation of site-specific management zones using fuzzy clustering analysis in a coastal saline land. Computers and Electronics in Agriculture 56(2):174-186.
- Liu, W.Q., Xu, X.Y., Lu, F., Cao, J.R., Li, P., Fu, T.F., et al. 2018. Three-dimensional mapping of soil salinity in the southern coastal area of Laizhou Bay, China. Land Degradation and Development 29:3772-3782.
- Mieza, M.S., Cravero, W.R., Kovac, F.D., and Bargiano, P.G. 2016. Delineation of site-specific management units for operational applications using the topographic position index in La Pampa, Argentina. Computers and Electronics in Agriculture 127:158-167.
- Minasny, B., and McBratney, A.B. 2002. FuzME Version 3.0. Australian Centre for Precision Agriculture, The University of Sydney, Sydney, Australia.
- Moharana, P.C., Jena, R.K., Pradhan, U.K., Nogiya, M., Tailor, B.L., Singh, R.S., et al. 2019. Geostatistical and fuzzy clustering approach for delineation of site-specific management zones and yield-limiting factors in irrigated hot arid environment of India. Precision Agriculture 3:1-23.
- Ohana-Levi, N., Bahat, I., Peeters, A., Shtein, A., Netzer, Y., Cohen, Y., et al. 2019. A weighted multivariate spatial clustering model to determine irrigation management zones. Computers and Electronics in Agriculture 162:719-731.
- Oldoni, H., Terra, V.S.S., Timm, L.C., Júnior, C.R., and Monteiro, A.B. 2019. Delineation of management zones in a peach orchard using multivariate and geostatistical analyses. Soil and Tillage Research 191:1-10.
- Ortega-Blu, R., and Molina-Roco, M. 2016. Evaluation of vegetation indices and apparent soil electrical conductivity for site-specific vineyard management in Chile. Precision Agriculture 17(4):434-450.
- Pansu, M., and Gautheyrou, J. 2003. Handbook of soil analysis: Mineralogical, organic and inorganic methods. Springer-Verlag, Berlin Heidelberg, Germany.
- Robertson, G.P. 2008. GS+: Geostatistics for the environmental sciences. Gamma Design Software, Plainwell, Michigan, USA.
- Rossi, R., Pollice, A., Bitella, G., Labella, R., Bochicchio, R., and Amato, M. 2018. Modelling the non-linear relationship between soil resistivity and alfalfa NDVI: A basis for management zone delineation. Journal of Applied Geophysics 159:146-156.
- Schepers, J.S., Schlemmer, M.R., and Ferguson, R.B. 2000. Site-specific considerations for managing phosphorus. Journal of Environment Quality 29(1):125-130.
- Schepers, A.R., Shanahan, J.F., Liebig, M.A., Schepers, J.S., Johnson, S.H., and Luchiari, A. 2004. Appropriateness of management zones for characterizing spatial variability of soil properties and irrigated corn yields across years. Agronomy Journal 96(1):195-203.
- Shaddad, S.M., Buttafuoco, G., Elrys, A., and Castrignano, A. 2019. Site-specific management of salt affected soils: A case study from Egypt. Science of Total Environment 688:153-161.
- Shaver, T.M., Peterson, G.A., Ahuja, L.R., and Westfall, D.G. 2013. Soil sorptivity enhancement with crop residue accumulation in semiarid dryland no-till agroecosystems. Geoderma 192:254-258.
- Yao, R.J., Yang, J.S., Zhang, T.J., Gao, P., Wang, X.P., Hong, L.Z., et al. 2014. Determination of site-specific management zones using soil physico-chemical properties and crop yields in coastal reclaimed farmland. Geoderma 232-234:381-393.
- Zovko, M., Romić, D., Colombo, C., Di Iorio, E., Romić, M., Buttafuoco, G., et al. 2018. A geostatistical Vis-NIR spectroscopy index to assess the incipient soil salinization in the Neretva River valley, Croatia. Geoderma 332:60-72.