Article

Soil Properties Spatial Variability and Delineation of Site-Specific Management Zones Based on Soil Fertility Using Fuzzy Clustering in a Hilly Field in Jianyang, Sichuan, China

Mohamed S. Metwally 1,2, Sameh M. Shaddad 2, Manqiang Liu 1,*, Rong-Jiang Yao 3, Ahmed I. Abdo 2, Peng Li 1, Jiaoguo Jiao 1 and Xiaoyun Chen 1,*

1 Soil Ecology Lab, College of Resources and Environmental Sciences, Nanjing Agricultural University, Nanjing 210095, China; helalamera@yahoo.com (M.S.M.); liumq@njau.edu.cn (M.L.); 2017203002@njau.edu.cn (P.L.); jiaguojiao@njau.edu.cn (J.J.)
2 Soil Science Department, Faculty of Agriculture, Zagazig University, Zagazig 44511, Egypt; shaddadsm@gmail.com (S.M.S.); drahmedabdo_2010@yahoo.com (A.I.A.)
3 State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences (CAS), Nanjing 210008, China; rjyao@issas.ac.cn
* Correspondence: xychen@njau.edu.cn; Tel.: +86-25-84395104

Received: 17 October 2019; Accepted: 5 December 2019; Published: 11 December 2019

Abstract: Avoiding soil degradation and improving crop productivity could be achieved by performing sustainable soil nutrient management with an appropriate understanding of soil properties’ spatial variability. The present fertilizer recommendations for the region where the study area is located are typically symmetric for large regions. This leads to the under-application of fertilizers in zones with low nutrient contents and over-application in zones with high nutrient contents. Therefore, this study was conducted to assess soil management zones (MZs) in the study area for effective soil nutrient management and to evaluate soil properties’ spatial variability. A total of 100 geo-referenced soil samples were collected at a depth of 0–20 cm, processed and analyzed for pH, available nitrogen (AN), available phosphorus (AP), available potassium (AK), soil organic carbon (SOC), total nitrogen (TN) and total phosphorous (TP), while C:N, C:P and N:P ratios were calculated. Soil properties’ coefficients of variation (CVs) widely varied from low (1.132%) to moderate (45.748%). Ordinary kriging and semi-variogram analysis showed different spatial variability patterns for the studied soil properties with spatial dependence ranged from weak to strong. MZs were delineated by performing principal component analysis (PCA) and fuzzy K-means clustering. Four PCs with eigen values more than 1 dominated 84.44% of the total variance, so they were retained for clustering analysis. Three MZs were delineated based on the two criteria modified partition entropy (MPE) and fuzzy performance index (FPI). The studied soil properties differed significantly among MZs. Thus, the methodology used for MZ delineation could be used effectively for soil site-specific nutrient management for avoiding soil degradation concurrently with maximizing crop production in the study area.

Keywords: sustainable agriculture; land degradation; precision agriculture; geostatistical analysis; spatial variability; principal component analysis; clustering analysis; site-specific management zones

1. Introduction

Soil degradation due to inadequate agriculture practices has become a global problem [1], which is clearly appearing in China due to fertilization mismanagement and the over exploitation of agricultural
lands. China’s population has increased rapidly during the last 50 years, likewise the production of the main crops has increased. However, such increases in productivity have been mainly attained by using unsustainable agricultural practices and inadequate and imbalanced fertilization. On the other hand, a depletion in soil nutrients occurs due to increasing crop yield in the land unit area, which will cause detrimental impacts on agricultural productivity in the future [2]. Whereas, the topsoil degradation due to soil erosion, decreased soil organic matter and soil nutrients content which led to decreasing soil fertility and consequently declining crop production [3–7]. One of the most roughly eroded regions in the Upper Yangtze River Basin is Sichuan Basin which is dominated by hilly areas [8], where the study area is located. Therefore, intensifying agriculture using balanced fertilization and environmentally sustainable practices will be the best solution for China.

Soil has a heterogeneous structure and its status affects ecosystem processes which control nutrients cycling [9]. Avoiding soil degradation and improving soil health and fertility level could be achieved by performing sustainable soil management with an appropriate understanding of soil properties [10,11]. Soil properties spatial variability is affected by farming management practices, such as irrigation and fertilization, as well as soil formation factors, like soil parent materials [12]. Therefore, managing field areas as a uniform unit oftentimes causes the treating of high nutrient content areas with over application of inputs, and treating low nutrient content areas with insufficient input applications [13,14], which leads to soil degradation. The effects of soil nutrients spatial variation on crop yields is clearly observed in cultivated sloping lands. However, what is known about soil nutrients spatial variation mechanism is still little [8,15]. If the spatial variation mechanism of soil nutrients is better understood, it will be possible to control its negative effects on agriculture production by using suitable soil management practices [8].

Precision agriculture is an idea for handling natural resources and understanding sustainable agricultural development [16]. Soil variation must be characterized quantitatively and locally to achieve the objectives of precision agriculture, in which the perfect benefits of environmental protection and profitability could be attained by matching each of the agricultural practices and land use with the local conditions [17–19]. Precision Agriculture was founded on the evaluation of within-field variation; hence the delineation of management zones is crucial for facilitating variable managing among the different zones [20]. The evaluation of soil properties spatial distribution could be conducted by using geostatistical methods [21]. Values of different soil properties at un-sampled locations could be predicted using geostatistical estimation by accounting the spatial correlation between sampled and estimated points, which results in reducing involved costs and estimation error [22]. Hence, geostatistics is essential for sustainable agriculture, as it provides valuable information about soil properties; this information contributes to knowing what, when, where, and how much farming inputs will maintain soil productivity and minimize costs concurrently with decreasing the environmental impact [23]. So, an appropriate understanding of soil properties spatial distribution could be used for soil site-specific management for sustaining crop and soil productivity by variable rate addition of nutrients [24–27].

Understanding the relationship between soil characteristics and spatially varied fertility was urgently needed due to the public concern about maximizing soil productivity and crop inputs efficiency [28,29]. Therefore, it was necessary to delineate site-specific management zones [30]. The most general approach used to manage within-field spatial variability was management zones (MZs) [31] because MZs are symmetric sub-regions with similar characteristics affecting yield or with the same yield productivity [32]. Also, they could be defined as sub-areas with equal potential productivity. A soil sampling grid was optimized by Khosla and Alley [33] by using homogenous management zones. Also, nutrient maps for fertilizer application with variable rates were developed based on management zones by Fleming et al. [34]. Soil properties’ spatial variability might also be delineated by locating within field spatially coherent areas [35].

Principal component analysis (PCA) and fuzzy means classification were frequently used to delineate soil MZs in different agriculture ecosystems with different crops [26,36,37]. A related set of
data could be summarized into a few expressive components by performing principal component analysis (PCA) [38]. Kriging methods could be used in mapping principal component values [39]. MZs could be delineated by using the principal components (PCs) scores to perform the clustering analysis [40].

Soil chemical properties were influenced by soil organic carbon (SOC) and soil pH among soil traits, as they affect the soil nutrients’ availability [41]. Soil nutrients availability was affected by soil pH, while chemical reactions and the physical and biological environment in the soil were modified by SOC [37]. The ecosystem production capacity was directly affected by SOC [42]. Also, total nitrogen (TN) and total phosphorus (TP) were considered to be from the most important indicators of soil productivity and fertility [43]. For plant growth soil N and P were major nutrient elements, which influence the photosynthesis process and other processes related to plant production [44]. Moreover, soil N and P cycles were associated with the SOC cycle [45], and had the ability to alleviate global climate change effects [46–48]. Hence, for the sustainable development of soil OC, N, P and their stoichiometry variations must be better understood to assess the status of nutrients in soil ecosystems. Furthermore, soil C:N ratio is considered to be a useful indicator of soil organic matter decomposition [49,50], whereas, by adding fertilizers soil N:P ratio was changed, thus it has a potential diagnostic value [51]. Besides, according to Tian et al. [52], soil phosphorous suppling capacity depended on the soil content of TN and the parent material weathering stage, which was varied by spatial heterogeneities.

However, information concerning soil properties’ spatial variability and soil MZs in the Jianyang–Sichuan province is still limited, which leads to the uniform management of Sichuan soils, which may result in soil degradation due to the over and under application of agricultural inputs. Thus, the present study was carried out to (1) evaluate soil properties’ spatial distribution using geostatistics, in addition to (2) classifying the study area into MZs based on the status of soil nutrients using PCA and fuzzy K-mean cluster analysis.

2. Materials and Methods

2.1. The Study Area

The study was carried out in a field of 9.22 ha in Jianyang, Sichuan basin, Sichuan Province, China, situated at (30° 30′ 49.85″ N and 104° 38′ 14.91″ E) (Figure 1). The area’s climate was described as subtropical humid, receiving an average annual rainfall of 944 mm. Summers had a good deal of rainfall, while the winters had very little. The mean highest temperature (33 °C) was recorded in July while the mean lowest temperature (11 °C) was recorded in January. The mean highest relative humidity (77%) was obtained in September and the mean lowest relative humidity (47%) was obtained in December. The soil of the study area weathered from purple sandstone and mudstone, and according to the soil classification of FAO (Food and Agriculture Organization) it was classified as Regosols. Four texture classes were found in the study area including clay loam, silty clay loam, silt loam and loam. The study area was planted with corn at the sampling time.
Figure 1. Location of the study area (Jianyang, Sichuan, China) with sampling points
2.2. Collection and Analysis of Soil Samples

A total of 100 soil samples were collected at a depth of (0 to 20 cm) based on an almost regular grid of 30 × 30m within the study area. Three soil subsamples were gathered using hand auger from a radius of one meter from each sampling location. These subsamples were mixed to get a representative soil sample for each sampling location. Sampling locations were obtained by using a Global Positioning System (GPS) hand-held device, in which latitude and longitude geographical coordinates were recorded for each sampling point. The obtained soil samples were air-dried, while debris and stones were discarded. By using a wooden pestle, the soil samples were grinded, then each soil sample was passed out of a 2 mm sieve and then were stored in polyethylene bottles to estimate each of soil pH, AN, AK and AP. A part of the 2 mm screened samples were re-sieved through a 0.15 mm sieve and saved in polyethylene bottles to determine each of SOC, TN and TP.

Soil pH was estimated in a soil water suspension of 1:2.5 (w/v) [53]. The content of SOC was evaluated by the method of Walkley and Black [54]. TN was estimated by the Kjeldahl method [55]. By using a spectrophotometer, TP was estimated [56]. AN was evaluated using the Micro-Kjeldhal procedure [57] and AP was tested following the way of Olsen [58], while AK was delineated using Flame photometer [59]. C:N, C:P and N:P ratios were calculated by dividing SOC by TN, SOC by TP and TN by TP, respectively. The international pipette method [60] was used to delineate the soil texture in only 50 samples representing the study area.

2.3. Descriptive Statistics

Descriptive statistics, like mean, minimum, maximum, coefficient of variation (CV) and standard deviation (SD), of the studied soil properties were delineated by using IBM SPSS 23 statistics software, Armonk, New York, U.S.A. Skewness and kurtosis were calculated for the studied soil properties to perform the test of normality for each of them. Except for AP, all the other soil properties passed the test. So, before analyzing the data geostatically, AP scores were transformed, performing the natural logarithm to be distributed normally. Then, using a weighted back transformation technique, the data were back transformed. The relationships between the ten soil properties were evaluated by obtaining the values of Pearson’s correlation coefficient.

2.4. Geostatistical Analysis

For evaluating the spatial variation pattern of the studied soil properties, a semi-variogram was calculated for each soil property using Equation (1) by utilizing ArcGIS 10.4.1 software.

\[
\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [z(x_{\alpha} + h) - z(x_{\alpha})]^2
\]

(1)

where: \(\gamma(h)\) represents the semivariance for the lag distance (h), \(N(h)\) is the samples pairs number which separated by (h), \(z(x_{\alpha})\) is the sample’s measured value at the sampling location (\(x_{\alpha}\)) and \(z(x_{\alpha} + h)\) is the sample’s measured value at the location (\(h + x_{\alpha}\)). Semi-variogram models like stable, circular, Gaussian, exponential and spherical, were estimated to delineate the best fitted model for each soil property.

The cross-validation technique was performed to choose the best fitted semi-variogram model for each of the studied soil properties, that is, by comparing the estimated values which were kriged by using the semi-variogram model with the actual values. So, mean error (ME) was calculated for each model to delineate the best fitted one for each soil property, in which the best is the model with the lowest (ME) value, as it has the highest prediction accuracy.

\[
ME = \frac{\sum_{i=1}^{n}[z(x_i, y_i) - z(x_i, y_i)]}{n}
\]

(2)
In which \( n \) is each case number of observations, \( z(x_i, y_i) \) is the observed soil property, \( z \times (x_i, y_i) \) is the estimated soil property, and \( (x_i, y_i) \) is the soil sample coordinates.

The values of the different soil properties were estimated at the un-sampled locations for each property using the ordinary kriging (OK) method [61]. The OK technique was performed because it is the most reliable of all predicting techniques based on (ME) [62]. OK is also the best unbiased predicting method in cases in which the soil samples locations were selected randomly and sparse to predict the values of the soil properties at the un-sampled point. It also reduces the outliers’ impact as one of its most important benefits [63].

2.5. Principal Component Analysis

To summarize the principal sources of the data variation between the correlated variables, principal component analysis (PCA) was used, as it is a multivariate analysis method for dimension reduction which uses the correlated variables to recombine and identify the orthogonal linear of the variables. Instead of a covariance matrix a correlation matrix including the studied soil properties was inputted for the PCA, so that a normalized PCA resulted. The number of the principal components (PCs) must be equal to the number of the variables inserted to the analysis. Only PCs with high eigen values were considered to be the best to represent the studied properties [64]. In this study the management zones were delineated by using the scores of PCs with eigen values more than 1 in the clustering analysis.

2.6. Fuzzy k-Means Clustering Algorithm

The commonly used fuzzy k-means classification method was performed to divide the datasets into diverse clusters; each one has its common characteristics [30]. By using FuzMe software, the study area was parted into two to seven clusters (with settings of maximum zones = 7, minimum zones = 2, maximum iteration = 300, fuzziness exponent = 1.5 and stopping criterion = 0.0001) [65]. For a practical use of the MZs, the maximum number of clusters was considered to be seven.

To delineate the optimal number of classes, the two quantitative estimation criteria modified partition entropy (MPE) [66] and fuzziness performance index (FPI) [67] were used, and they could be calculated by using the following equations:

\[
FPI = 1 - \frac{N}{N-1} \left[ 1 - \sum_{i=1}^{N} \sum_{j=1}^{K} \left( \mu_{m}^{i,j} \right)^2 / k \right] 
\]

(3)

\[
MPE = -\frac{1}{N \log(k)} \sum_{i=1}^{N} \sum_{j=1}^{k} \mu_{m}^{i,j} \log(\mu_{m}^{i,j}). 
\]

(4)

In which \( N \) is the number of soil samples; \( k \) is the number of classes; \( m \) is the weighting exponent of fuzziness; \( \mu_{ij} \) is the fuzzy membership and log is the natural logarithm. MPE estimates the amount of defective created by precise classes and FPI measures the fuzziness degree. The MPE depicts the uncertainty (or certainty) of fuzzy k-classification in which MPE = 1 refers to the maximum uncertainty and MPE = 0 means the maximum certainty. The FPI defines the membership sharing between each couple of fuzzy clusters, in which FPI = 1 refers to the maximum fuzziness and FPI = 0 means non-fuzziness. The optimum number of clusters was obtained against the minimum values of MPE and FPI. By using t-test analysis, the variance of each studied soil property across the management zones were evaluated.

3. Results and Discussion

3.1. Variability of Soil Properties

The studied soil properties’ descriptive statistics are given in (Table 1). The soil of the study area tended to be a somewhat alkaline, with a pH ranging between 7.85 and 8.31. The mean SOC
and TN were 9.58 and 0.98 g/kg, respectively. According to the soil nutrient classification standard in China [68], the means of SOC and TN content were low in the study area, while the mean soil AK and AP content were medium with a value of 147.10 and 17.02 mg/kg, respectively. Whereas the mean AN content was 35.42 mg/kg, which indicated that the study area’s AN content was very low. This shortfall in SOC, TN and AN might be due to removing the topsoil which is caused by water erosion, resulting in reducing soil nutrients and organic matter [6].

Table 1. Descriptive statistics for the studied soil properties.

<table>
<thead>
<tr>
<th>Parameter *</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>CV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN</td>
<td>13.13</td>
<td>49.88</td>
<td>35.42</td>
<td>35.88</td>
<td>7.28</td>
<td>−0.49</td>
<td>0.28</td>
<td>20.55</td>
</tr>
<tr>
<td>AK</td>
<td>100.00</td>
<td>210.00</td>
<td>147.10</td>
<td>145.00</td>
<td>22.80</td>
<td>0.50</td>
<td>0.39</td>
<td>15.50</td>
</tr>
<tr>
<td>AP</td>
<td>5.80</td>
<td>38.29</td>
<td>17.02</td>
<td>15.71</td>
<td>7.79</td>
<td>1.13</td>
<td>0.91</td>
<td>45.75</td>
</tr>
<tr>
<td>pH</td>
<td>7.85</td>
<td>8.51</td>
<td>8.13</td>
<td>8.13</td>
<td>0.09</td>
<td>−0.33</td>
<td>0.24</td>
<td>1.13</td>
</tr>
<tr>
<td>TP</td>
<td>0.54</td>
<td>1.11</td>
<td>0.83</td>
<td>0.82</td>
<td>0.13</td>
<td>0.16</td>
<td>−0.59</td>
<td>15.59</td>
</tr>
<tr>
<td>TN</td>
<td>0.62</td>
<td>1.43</td>
<td>0.98</td>
<td>0.98</td>
<td>0.17</td>
<td>0.19</td>
<td>0.10</td>
<td>16.92</td>
</tr>
<tr>
<td>SOC</td>
<td>4.41</td>
<td>14.47</td>
<td>9.58</td>
<td>9.68</td>
<td>2.22</td>
<td>−0.26</td>
<td>−0.46</td>
<td>23.20</td>
</tr>
<tr>
<td>CN</td>
<td>4.28</td>
<td>14.83</td>
<td>9.83</td>
<td>9.86</td>
<td>1.89</td>
<td>−0.38</td>
<td>0.64</td>
<td>19.27</td>
</tr>
<tr>
<td>CP</td>
<td>4.46</td>
<td>20.48</td>
<td>11.65</td>
<td>11.82</td>
<td>2.80</td>
<td>0.08</td>
<td>1.00</td>
<td>24.00</td>
</tr>
<tr>
<td>NP</td>
<td>0.77</td>
<td>1.68</td>
<td>1.19</td>
<td>1.16</td>
<td>0.43</td>
<td>0.07</td>
<td>15.50</td>
<td></td>
</tr>
</tbody>
</table>

* AN, available nitrogen (mg/kg); AK, available potassium (mg/kg); AP, available phosphorous (mg/kg); TP, total phosphorus (g/kg); TN, total nitrogen (g/kg); SOC, soil organic carbon; CN, carbon nitrogen ratio; CP, carbon phosphorus ratio; NP, nitrogen phosphorus ratio; SD: standard deviation; CV: coefficient of variation.

The soil C:N, C:P, and N:P ratios were considered to be good indicators of the soil nutrients status during soil development [52]. The high C:N ratio (>25) indicated that soil organic matter accumulation was faster than its decomposition [69]. As shown in Table 1, soil C:N ratio ranged between 4.28 and 14.83 with a mean value of 9.83, indicating a complete breakdown of soil organic matter in the study area. Bui and Henderson [50] reported similar results. Also, our results showed that soil C:P ratio values ranged from 4.46 to 20.48, when the mean value was 11.65, implying a phosphorous net mineralization. In which Paul [49] recorded that C:P ratio <200 referred to a net mineralization, while C:P ratio >300 referred to a net immobilization, and a C:P ratio ranging from 200 to 300 implied little change in the concentrations of soil soluble P. Our results revealed that the N:P ratio ranged between 0.77 and 1.68, with a mean value of 1.19, which indicated high microbial activity, whereas due to Wang et al. [70] the N:P ratio was negatively correlated with microorganisms’ activity and biomass in which N:P > 2 indicated biomass decline.

All soil properties were normally distributed with non-significant skewness except AP, which skewed with a value of 1.13. The lowest CV (1.13%) was for soil pH, while the highest CV (45.75%) was for AP, which was in line with Wang et al. [71]. Also, Karaman et al. [72] reported that AP was usually more variable than most of the other soil properties. According to Jakobsen [73], the CV of the different soil properties ranged from low (<10%) to moderate (10 to 100%). High CVs for the remaining soil properties revealed considerable spatial variability and so it was suggested to use nutrient site-specific management to improve the study area soil productivity.

3.2. Correlations Between Soil Properties

The degree of correlation among the ten soil properties is shown in Table 2. Most of the soil properties were significantly positively and negatively correlated with each other, showing similar and opposite spatial distribution patterns, respectively. AN, AP, AK, TN and TP were negatively correlated with soil pH, indicating that increasing N, P and K availability could be significantly achieved by decreasing soil pH in the zones with low concentrations of these nutrients. Soil AN and AP were positively correlated with SOC, in which SOC was an important portion of the soil which affected soil chemical, physical and biological properties influencing soil nutrients’ availability [37].
Table 2. Pearson’s correlation coefficient values showing relationship among the study area soil properties.

<table>
<thead>
<tr>
<th></th>
<th>AN</th>
<th>AK</th>
<th>AP</th>
<th>pH</th>
<th>TP</th>
<th>TN</th>
<th>SOC</th>
<th>CN</th>
<th>CP</th>
<th>NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AK</td>
<td>0.376 *</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP</td>
<td>0.109</td>
<td>0.450 *</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pH</td>
<td>−0.271 **</td>
<td>−0.427 **</td>
<td>−0.399 **</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>0.500 **</td>
<td>0.372 **</td>
<td>0.372 **</td>
<td>−0.283 **</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TN</td>
<td>0.546 **</td>
<td>0.536 **</td>
<td>0.342 **</td>
<td>−0.311 **</td>
<td>0.562 **</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOC</td>
<td>0.357 **</td>
<td>0.193</td>
<td>0.238</td>
<td>−0.147</td>
<td>0.274 **</td>
<td>0.572 **</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CN</td>
<td>−0.026</td>
<td>−0.207 *</td>
<td>−0.021</td>
<td>0.087</td>
<td>−0.164</td>
<td>−0.167</td>
<td>0.705 **</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td>0.025</td>
<td>−0.043</td>
<td>−0.030</td>
<td>0.035</td>
<td>−0.380 **</td>
<td>0.185</td>
<td>0.768 **</td>
<td>0.770 **</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>NP</td>
<td>0.106</td>
<td>0.206 *</td>
<td>−0.026</td>
<td>−0.047</td>
<td>−0.398 **</td>
<td>0.520 **</td>
<td>0.336 **</td>
<td>−0.030</td>
<td>0.600 **</td>
<td>1.000</td>
</tr>
</tbody>
</table>

AN: available nitrogen; SOC: soil organic carbon; AP: available phosphorous; AK: available potassium; TP: total phosphorus; C:N:C:N ratio; C:P:C:P ratio; N:P:N:P ratio. ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

The concentrations of SOC, TN and TP had a significant positive correlation with each other, and the comparatively high correlation coefficients (0.572) for soil SOC and TN and (0.562) for TN and TP, indicated that the C:N and N:P ratios were highly constrained. Also, a comparatively constrained C:P ratio was founded on the correlation coefficients of 0.274 for the SOC and TP concentrations. Ouyang et al. [74] obtained the same results for C:N and C:P ratios, and a different result for N:P ratio. Correlations among these properties revealed that PCA should be used to summarize the principal sources of data variance.

3.3. Soil Properties Spatial Distribution

As shown in Table 1, distributions of all the studied variables were lightly skewed (skewness < 1), and their means were close to their medians, except soil AN which skewed with a value of 1.13, so that before performing geostatistical analysis its values were log-transformed. Table 3 and Figure 2 show the parameters of the studied soil properties semi-variogram. Soil TN, pH, AP, AK, C:N, C:P and N:P were best modeled by using spherical models. Also, several authors found that most of the soil properties were best modeled by using spherical models [75–77] while TC, TP and AN were fitted best with stable models.
Figure 2. Experimental semi-variograms and fitted models for AN, AP, AK, SOC, TN, TP, C:N, C:P, N:P ratios and pH.
Table 3. The study area soil properties Semi-variogram parameters.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model</th>
<th>Nugget</th>
<th>Partial Sill</th>
<th>Sill</th>
<th>Nugget/Sill</th>
<th>SDC</th>
<th>Range (m)</th>
<th>ME</th>
<th>MSSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOC</td>
<td>Stable</td>
<td>3.950</td>
<td>3.050</td>
<td>7.000</td>
<td>0.5643</td>
<td>Moderate</td>
<td>679.4</td>
<td>-0.013</td>
<td>1.046</td>
</tr>
<tr>
<td>TN</td>
<td>Spherical</td>
<td>0.021</td>
<td>0.008</td>
<td>0.029</td>
<td>0.7241</td>
<td>Moderate</td>
<td>276.2</td>
<td>-0.004</td>
<td>1.003</td>
</tr>
<tr>
<td>TP</td>
<td>Stable</td>
<td>0.015</td>
<td>0.001</td>
<td>0.016</td>
<td>0.9375</td>
<td>Weak</td>
<td>154.0</td>
<td>0.001</td>
<td>1.019</td>
</tr>
<tr>
<td>pH</td>
<td>Spherical</td>
<td>0.007</td>
<td>0.001</td>
<td>0.008</td>
<td>0.8750</td>
<td>Weak</td>
<td>679.4</td>
<td>0.003</td>
<td>1.064</td>
</tr>
<tr>
<td>AP</td>
<td>Spherical</td>
<td>0.158</td>
<td>0.057</td>
<td>0.215</td>
<td>0.7349</td>
<td>Moderate</td>
<td>289.3</td>
<td>-0.066</td>
<td>1.031</td>
</tr>
<tr>
<td>AN</td>
<td>Stable</td>
<td>0.000</td>
<td>49.620</td>
<td>49.620</td>
<td>0.0000</td>
<td>Strong</td>
<td>116.1</td>
<td>-0.008</td>
<td>1.027</td>
</tr>
<tr>
<td>AK</td>
<td>Spherical</td>
<td>202.500</td>
<td>292.870</td>
<td>495.370</td>
<td>0.4088</td>
<td>Moderate</td>
<td>48.7</td>
<td>-0.691</td>
<td>1.056</td>
</tr>
<tr>
<td>NP</td>
<td>Spherical</td>
<td>0.030</td>
<td>0.003</td>
<td>0.033</td>
<td>0.9091</td>
<td>Weak</td>
<td>138.0</td>
<td>0.030</td>
<td>1.135</td>
</tr>
<tr>
<td>CP</td>
<td>Spherical</td>
<td>6.476</td>
<td>2.825</td>
<td>9.301</td>
<td>0.6963</td>
<td>Moderate</td>
<td>679.4</td>
<td>-0.031</td>
<td>1.111</td>
</tr>
<tr>
<td>CN</td>
<td>Spherical</td>
<td>2.879</td>
<td>1.350</td>
<td>4.229</td>
<td>0.6808</td>
<td>Moderate</td>
<td>600.0</td>
<td>-0.009</td>
<td>1.026</td>
</tr>
</tbody>
</table>

AN: available nitrogen; SOC: soil organic carbon; AP: available phosphorus; AK: available potassium; TP: total phosphorus; C:N: C:N ratio; C:P: C:P ratio; N:P: N:P ratio; SDC: spatial dependency class; MSSE: mean square standardized error; ME: mean error.

Cambardella et al. [78] reported that nugget to sill ratio value <0.25 reveals strong spatial dependence due to the intrinsic (inherent) factors like soil texture and mineralogy, while when the ratio’s value was between 0.25 and 0.75, it referred to moderate spatial dependence due to the extrinsic and intrinsic factors, whereas a ratio’s value >0.75 referred to weak spatial dependence due to the extrinsic factors like fertilization and tillage. The soil properties of this study had diverse spatial dependence due to their nugget to sill ratios. Soil AN had strong spatial dependence, whereas SOC, TN, AP, AK, C:N ratio and C:P ratio had moderate spatial dependence. Soil pH, TP and N:P ratio had weak spatial dependence, as shown in Figure 2, which might be due to the weak spatial distribution of these properties and hence it was recommended to carry out extra research based on a large scale sampling design to capture the spatial distribution of these variables.

The maximum distance in which spatial dependence or autocorrelation exists was defined as the range value of semi-variogram. As shown in Table 3, the range values of soil properties in this study ranged between 48 m for AK and 679 m for SOC. Larger than the obtained range values, spatial dependence does not exist for these soil properties. Lopez-Granados et al. [75] reported that a large range value indicated that estimated soil properties were influenced by anthropogenic and natural factors over larger distances than the other soil properties which have smaller ranges.

The distance between soil samples should be below half the semi-variogram range value [79]. Hence, the obtained range values for soil properties in this study could be used for planning the future soil sampling in the study area for geostatistical research by taking samples at interval distances less than half the obtained range values of the studied soil properties.

Cross-validation technique was performed to get the most precise predictions with the lowest ME values for the studied soil properties as shown in Table 3. According to Shaddad et al. [80], the lowest ME values revealed that soil properties kriging predictions were closer to the estimated values, while the MSSE value for each of the studied soil properties should be one; however, if the MSSE value was different from one but still within the tolerance interval $1 \pm (2/N)^{1/2}$, in which N was the number of soil samples, the model was considered to be accurate. The tolerance interval had ranged from 0.576 to 1.424 and as shown in Table 3. MSSE values for all soil properties were within this range, which referred to the high prediction accuracy of the semi-variograms models used for all the soil properties.

Figure 3 showed the distribution maps of the studied soil properties. The eastern part of the study area had the lowest content of soil TN, which had a similar distribution pattern to soil AN, due to the presence of a strong positive correlation between soil TN and AN. Soil AK was low in most of the study area parts. Soil AP was low in the west part of the study area. It can be clearly noticed that high pH distributed in the east and small regions in the center and the west of the area. As shown in Table 2, soil pH had a significant negative correlation with each of AN, AP, AK, TN and TP, so that the distribution pattern of pH was opposite in the location with the patterns of these soil properties as shown in Figure 3. SOC, TN and TP had similar distribution patterns due to the positive correlation between them, with low content dominating the east part of the study area and high content dominated...
the middle and the west parts. The map of the C:N ratio showed high values in the middle and the west and a small part in the east, while the maps of C:P and N:P ratios showed low values in most of the area except small regions in the middle and the west of the study area, and that was because of the study area, low content of soil OC and TN and the high content of TP.

Figure 3. Soil properties ordinary kriged distribution maps of the study area.
Valuable information about nutrient content in the study area was obtained from the distribution maps which were produced by using the OK technique. These pieces of information could be helpful to give recommendations for soil site specific nutrient managing, for getting maximum output and increasing the income by reducing the cost of the inputs paired with the best management practices.

3.4. Principal Component Analysis

Table 2 showed that most of the studied soil properties were significantly correlated, and PCA was carried out to summarize and aggregate the variability in the studied ten variables. The number of the resulted principal components (PCs) must be equal to the number of the variables which were inserted into the analysis. PCs with eigen values more than 1 were kept for the final analysis, in which, according to Sharma [81], a PC with an eigen value more than 1 explains variance more than an individual attribute. Based on this principle, only the first four PCs described about 84% of the measured data total variability, as shown in Table 4. The OK method was carried out to interpolate the distribution maps of the 4 PCs as shown in Figure 4.

Table 4. Principal component analysis of soil properties and loading coefficient for the first four principal components.

<table>
<thead>
<tr>
<th>Principal Component</th>
<th>Eigen Values</th>
<th>Component Loading (%)</th>
<th>Cumulative Loading (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.328</td>
<td>33.285</td>
<td>33.285</td>
</tr>
<tr>
<td>2</td>
<td>2.686</td>
<td>25.864</td>
<td>60.149</td>
</tr>
<tr>
<td>3</td>
<td>1.384</td>
<td>13.824</td>
<td>73.991</td>
</tr>
<tr>
<td>4</td>
<td>1.045</td>
<td>10.450</td>
<td>84.441</td>
</tr>
<tr>
<td>5</td>
<td>0.677</td>
<td>6.771</td>
<td>91.212</td>
</tr>
<tr>
<td>6</td>
<td>0.476</td>
<td>4.762</td>
<td>95.974</td>
</tr>
<tr>
<td>7</td>
<td>0.373</td>
<td>3.729</td>
<td>99.703</td>
</tr>
<tr>
<td>8</td>
<td>0.018</td>
<td>0.182</td>
<td>99.885</td>
</tr>
<tr>
<td>9</td>
<td>0.007</td>
<td>0.073</td>
<td>99.957</td>
</tr>
<tr>
<td>10</td>
<td>0.004</td>
<td>0.043</td>
<td>100.000</td>
</tr>
</tbody>
</table>

Principal component loading for each variable

<table>
<thead>
<tr>
<th>Principal component</th>
<th>AN</th>
<th>AK</th>
<th>AP</th>
<th>pH</th>
<th>TP</th>
<th>TN</th>
<th>SOC</th>
<th>C:N ratio</th>
<th>C:P ratio</th>
<th>N:P ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1</td>
<td>0.651</td>
<td>0.670</td>
<td>0.332</td>
<td>-0.518</td>
<td>0.560</td>
<td>0.874</td>
<td>0.735</td>
<td>0.139</td>
<td>0.346</td>
<td>0.386</td>
</tr>
<tr>
<td>PC2</td>
<td>-0.185</td>
<td>-0.337</td>
<td>-0.267</td>
<td>0.287</td>
<td>-0.573</td>
<td>-0.096</td>
<td>0.580</td>
<td>0.786</td>
<td>0.929</td>
<td>0.481</td>
</tr>
<tr>
<td>PC3</td>
<td>0.073</td>
<td>-0.226</td>
<td>0.170</td>
<td>0.005</td>
<td>0.459</td>
<td>-0.262</td>
<td>0.299</td>
<td>0.571</td>
<td>-0.032</td>
<td>-0.776</td>
</tr>
<tr>
<td>PC4</td>
<td>0.500</td>
<td>-0.221</td>
<td>-0.602</td>
<td>0.477</td>
<td>0.268</td>
<td>0.244</td>
<td>0.085</td>
<td>-0.104</td>
<td>-0.085</td>
<td>0.019</td>
</tr>
</tbody>
</table>
Principal component 1 (PC1) explained 33.28% of the total variability, and it was dominated by all the soil properties except soil C:N, C:P and N:P ratios. SOC, TP, C:N ratio and C:P ratio influenced the second principal component 2 (PC2), which explained 26.86% of total variability. While the third principle component 3 (PC3) explained 13.84 % of the total variability and was controlled by soil C:N and C:P ratios. Principal component 4 (PC4) explained 10.45 % of the variability and was affected by soil AN and AP.

3.5. Management Zones Delineation Using Clustering Analysis

Scores of the first four PCs were inserted into FuzMe software to perform the fuzzy K-mean classification technique to cluster the four PCs into management zones (MZs). The two functions FPI and MPE were plotted versus the number of classes as shown in (Figure 5) to obtain the optimal number of MZs in the study area [82], in which, the minimum FPI met the minimum MPE against an optimum number of three clusters. The kriged map shown in (Figure 6) was describing three fertility management zones named as MZ1, MZ2 and MZ3. Analysis of variance was carried out using t-test to evaluate the efficiency of the spatial variability description of the studied soil properties over the delineated MZs by the combination of PCA and fuzzy k-means clustering. As shown in Table 5, the analysis of variance showed that the three delineated MZs were clearly diverse from each other. The same results were obtained by Tripathi et al. [83] and Shukla et al. [26].
The study area was in a dire need of nitrogen applications, in which AN concentrations were extremely low in MZ1, MZ2, and MZ3, respectively. Also, potassium application was needed in MZ2 and MZ3 which had medium concentrations of AK while MZ1 had high. SOC concentrations were low in MZ2 and MZ3, and very low in MZ1 and MZ2, while. While TN concentrations were medium, low and high in MZ2, MZ1, and MZ3, respectively. Also, potassium application was needed in MZ2 and MZ3 which had medium concentrations of AK while MZ1 had high concentrations of it. For AP MZ2 and MZ3 had medium concentrations while MZ3 was high. SOC concentrations were low in MZ2 and MZ3, and very low in MZ1 and MZ3, so that it was very necessary to apply a considerable amount of organic matter to the study area, in which it was considered the main source of organic carbon which affects soil chemical, physical and biological properties influencing soil nutrients’ availability [37]. In addition, soil C:N and C:P ratios were very narrow in the study area which meant that soil organic matter was decomposing rapidly so that applying organic matter would be very useful to compensate for the

Soil properties significantly varied across the three MZs. MZ2 had the highest soil pH value. The study area was in a dire need of nitrogen applications, in which AN concentrations were extremely low in MZ3, and very low in MZ1 and MZ2, while. While TN concentrations were medium, low and very low in MZ1, MZ2, and MZ3, respectively. Also, potassium application was needed in MZ2 and MZ3 which had medium concentrations of AK while MZ1 had high concentrations of it. For AP MZ1 and MZ2 had medium concentrations while MZ3 was high. SOC concentrations were low in MZ2 and MZ3, and very low in MZ1 and MZ3, so that it was very necessary to apply a considerable amount of organic matter to the study area, in which it was considered the main source of organic carbon which affects soil chemical, physical and biological properties influencing soil nutrients’ availability [37]. In addition, soil C:N and C:P ratios were very narrow in the study area which meant that soil organic matter was decomposing rapidly so that applying organic matter would be very useful to compensate for the

Figure 5. Fuzzy performance index (FPI) and modified partition entropy (MPE) calculated for delineating the optimum number of clusters for the study area.

Figure 6. Soil management zones of the study area.

Table 5. The mean values of soil properties in the three management zones.

<table>
<thead>
<tr>
<th>Management Zones</th>
<th>No. of Points</th>
<th>AN</th>
<th>AK</th>
<th>AP</th>
<th>pH</th>
<th>TP</th>
<th>TN</th>
<th>SOC</th>
<th>C:N Ratio</th>
<th>C:P Ratio</th>
<th>N:P Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31</td>
<td>39.36 a</td>
<td>155.66 a</td>
<td>15.31 b</td>
<td>8.11 ab</td>
<td>0.84 a</td>
<td>1.09 a</td>
<td>9.03 b</td>
<td>8.23 c</td>
<td>10.98 b</td>
<td>1.32 a</td>
</tr>
<tr>
<td>2</td>
<td>44</td>
<td>36.64 a</td>
<td>141.77 b</td>
<td>15.37 b</td>
<td>8.15 a</td>
<td>0.85 a</td>
<td>0.96 b</td>
<td>10.66 a</td>
<td>11.09 a</td>
<td>12.63 a</td>
<td>1.14 b</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>28.39 b</td>
<td>145.85 ab</td>
<td>22.04 a</td>
<td>8.09 b</td>
<td>0.79 a</td>
<td>0.87 c</td>
<td>8.37 b</td>
<td>9.57 b</td>
<td>10.74 b</td>
<td>1.11 b</td>
</tr>
</tbody>
</table>

Different letters (a, b and c) within each column indicate a significant difference between the management zones at the 0.05 level.
decomposed organic matter. Therefore, the appropriate agricultural practices must be performed to improve soil organic matter content.

It is therefore very clear how much the importance of using site-specific land management for avoiding land degradation, as well as sustaining soil productivity, in the study area. Thus, cluster analysis is a suitable choice for identifying the study area MZs, and performing site-specific land management techniques will improve soil productivity, in which the average values of the studied soil properties within the MZs could be used as a reference for variable rate applications [36].

4. Conclusions

The studied soil properties’ coefficients of variation revealed high spatial variability and indicated that site-specific nutrient management techniques need to be performed in the study area. Soil AN, AP, TN and TP were positively correlated with SOC. Geostatistical tools were used to quantify the spatial variability of the studied soil properties. Geostatistical analysis revealed spherical and stable best-fit semi-variogram models for the studied soil properties. The studied soil properties were characterized by spatial heterogeneity, with spatial dependence ranging from weak to strong. Three MZs were delineated by performing PCA and fuzzy k-means clustering algorithms, and the variance analysis revealed soil fertility heterogeneity among the three zones. Future soil sampling processes may be optimized by using MZs information. The MZs map could be the guide for the precise management of nitrogen, phosphorous, potassium and organic carbon for the different zones in the study area to improve the soil chemical, physical and ecological properties to combat soil degradation, which was the principle of sustainable agriculture, in which the mean values of the studied soil properties in each of the three delineated zones could be used as a reference for applying variable rates of fertilizers.

Author Contributions: Conceptualization, M.S.M.; methodology, M.S.M. and R.-J.Y.; software, M.S.M., S.M.S. and R.-J.Y.; validation, M.S.M., S.M.S. and A.I.A.; formal analysis, M.S.M. and S.M.S.; investigation, M.S.M.; resources, M.L., P.L., J.J. and X.C.; data curation, M.S.M.; writing—original draft preparation, M.S.M.; writing—review and editing, M.S.M., M.L. and X.C.; visualization, M.S.M.; supervision, M.L.; project administration, M.L.; funding acquisition, M.L. and X.C.

Funding: This work was supported by Special Fund for Agro-scientific Research in the Public Interest (No. 201503121), China Agriculture Research System-Green Manure (No. CARS-22-G-10) and the Fundamental Research Funds for the Central Universities (No. KYY201702).

Acknowledgments: We greatly thank those persons who offered valuable helps in the soil sampling and laboratory analysis, namely J.S., D.Y., J.Z., J.H., and C.Z.

Conflicts of Interest: The authors declare no conflict of interest

References

3. Seutloali, K.E.; Beckedahl, H.R. Understanding the factors influencing rill erosion on roadcuts in the south eastern region of South Africa. Solid Earth 2015, 6, 633–641. [CrossRef]


44. Liu, Z.P.; Shao, M.A.; Wang, Y.Q. Spatial patterns of soil total nitrogen and soil total phosphorus across the entire Loess Plateau region of China. *Geoderma* 2013, 197, 67–78. [CrossRef]


55. Bremner, J. Determination of nitrogen in soil by the Kjeldahl method. *J. Agric. Sci.* 1960, 55, 11–33. [CrossRef]


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).