Spatial variability and sampling optimization of soil organic carbon and total nitrogen for Minesoils of the Loess Plateau using geostatistics

Jinman Wang a, *, Ruixuan Yang b, Zhongke Bai c

a College of Land Science and Technology, China University of Geosciences, 29 Xueyuanlu, Haidian District, 100083 Beijing, People’s Republic of China
b College of Sciences, Shihezi University, 832000 Shihezi, Xinjiang, People’s Republic of China
c Key Laboratory of Land Consolidation and Rehabilitation, Ministry of Land and Resources, 100035 Beijing, People’s Republic of China

A R T I C L E   I N F O

Article history:
Received 17 November 2014
Received in revised form 3 April 2015
Accepted 28 April 2015
Available online 16 May 2015

Keywords:
Spatial variability
Soil organic carbon
Total nitrogen
Minesoils
Loess Plateau

A B S T R A C T

Drastically disturbed minesoils can result in a significant loss of soil organic carbon (SOC) and total nitrogen (TN). To assess the effect of mining activities on minesoils and to track the changes in reclaimed soil quality, the variability of SOC and TN concentrations in the Shansi Pinghuo Antaibao opencast coalmine inner dump after dumping and before reclamation was analyzed using geostatistics, and a number of soil monitoring points were evaluated after land reclamation. Soil samples were collected from depths of 0–20 cm, 20–40 cm, 40–60 cm and 60–80 cm at 78 sampling sites in the study area over an area of 0.44 km². The coefficient of variation (CV) for TN was the least at <15% for depths of 0–40 cm and 60–80 cm. For TN at a depth of 40–60 cm and the SOC at all depths, the CV was moderate at 15–35%. Interpolation using kriging displayed a high heterogeneity of TN and SOC, and the spatial structure of the original landform was partially or completely destroyed. Revegetation was an important measure for increasing the accretion of C and N compared to an unmined site. The kriging-interpolated maps were a very valuable tool in monitoring soil properties after land reclamation at the field scale, and RMSE can be used to determine the number of sampling point for soil properties.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

For economic development, China requires abundant natural resources, particularly from coal mines. Currently, opencast coalmine production in China has increased to 30% of total production (Li et al., 2012). Surface soil and vegetation were destroyed by opencast mining, leading to the destruction of the local ecological environment (Vymazal and Sklenicka, 2012). Mining activities can cause severe changes to terrestrial ecosystems, leading to soil degradation (McSweeney and Jansen, 1984). In most cases, there is a significant loss of soil organic carbon (SOC) and associated nutrients (Akala and Lal, 2001; Bodlak et al., 2012; Wang et al., 2014). SOC influences the chemical and physical properties of soil and can release its contained nutrients through mineralization in forms available to plants. In agricultural ecosystems, soil total nitrogen (TN) is a major determinant and indicator of soil fertility and quality, which are closely related to soil productivity. Thus, information about the spatial distribution of TN and SOC is necessary to evaluate potential vegetation productivity.

Geostatistics is a useful tool for analyzing spatial variability, interpolating between point observations, and ascertaining the interpolated values with a specified error using a minimum number of observations (Long et al., 2014). There have been several attempts to characterize the variability of TN and SOC using geostatistics (Nyamadzawo et al., 2008; Shukla et al., 2007; Wang et al., 2010). Most of these studies were conducted on agricultural fields, and only a few were performed with minesoils (Akala and Lal, 2001; Shukla and Lal, 2005; Shukla et al., 2004).

For reclaimed minesoils, the variability of SOC and TN over time has been reported (Akala and Lal, 2001; Nyamadzawo et al., 2008). However, studies on the spatial variability of SOC and TN in the minesoils after dumping and before reclamation are insufficient, particularly in the Loess Plateau area. There is also a general lack of much finer scale information on minesoils, particularly the spatial variability of SOC and TN with changes in soil sampling size. Therefore, the objectives of this study were to (i) assess the spatial variability of SOC and TN in minesoils after dumping and before reclamation using geostatistics and to analyze the effect of opencast mining activities on SOC and TN and to (ii) determine

http://dx.doi.org/10.1016/j.ecoleng.2015.04.103
0925-8374/ © 2015 Elsevier B.V. All rights reserved.
2. Materials and methods

2.1. Study site

The study area was an opencast coalmine in Shanxi Pingshuo, which is the largest opencast coal mining area in China, including the Antaibao, Anjialing and East opencast mines. The Pingshuo opencast coal mine is located along the border of the Shanxi Province, Shaanxi and Inner Mongolia in the east Loess Plateau, with geographic coordinates of 112° 10' 58" to 113° 30' E, 39° 37' N. This mining area has a typical temperate arid to semi-arid continental monsoon climate and a fragile ecological environment. The altitude is 1300–1400 m, and the terrain is loess hills with grass vegetation. The average annual rainfall is approximately 450 mm, with 65% falling from June to September. The average annual evaporation, however, is approximately 2160 mm, 4.6 times more than the rainfall. Its chestnut soils are characterized by low levels of organic matter and poor structure.

The specific study area is located in the inner dump of the Antaibao mine with an area of 0.44 km². The study site is on the top platform of the inner dump, with an altitude of 1474–1480 m. It was dumped in 2012, and no vegetation was planted. The soils of the original landform in this mining area consist of a thick topsoil with low soil fertility. Opencast mining activities, such as excavation, transport and dumping, have significantly disturbed the soils, and the soil profile has been greatly changed. The study site was highly disrupted and the coal was removed from the underlying layers, followed by dumping with topsoil application and no planting as of yet.

We hypothesize that similar climate and mineral organic conditions are good candidates for studying the variation of SOC and TN. Therefore, we selected an undisturbed site at Guangling in the Shanxi province as a reference site. The average annual precipitation for the two sites is approximately 400–450 mm. The two sites were located close enough to each other to have similar climatic conditions, and their mineralogy is fairly similar (Xiao et al., 2011).

2.2. Soil sampling and analysis

In June 2013, 312 equally spaced soil samples at 78 sampling sites were collected using an auger at depths of 0–20 cm, 20–40 cm, 40–60 cm and 60–80 cm. The sampling sites were randomly arranged within a distance of 60–80 m. All soil samples were air-dried, and the clods were broken using steel rolling pins in order for the soil to pass through a 2-mm mesh. SOC and TN concentrations were determined using the dry combustion method (Elementar, GmbH, Hanau, Germany) with approximately 1 g of soil (0.25 mm).

2.3. Statistical analysis

Descriptive statistics, including the mean, median, dispersion of variance, coefficient of variation (CV), maximum, minimum and Kolmogorov–Smirnov (K–S) test, were obtained for each measured soil variable using SPSS 19.0. Geostatistical methods were used to study the spatial variability of the soil properties. The geostatistics approach consists of two parts: one is the calculation of an experimental variogram from the data and model fitting and the second is a prediction at unsampled locations (Burgos et al., 2006). The semivariogram of each soil property was constructed using the following model:

\[
\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (Z(x_i) - Z(x_i + h))^2
\]

where \(\gamma(h)\) is the semivariance for the lag interval \(h\), \(Z(x_i)\) and \(Z(x_i + h)\) are variables at locations \(x_i\) and \(x_i + h\), and \(N(h)\) is the number of pairs separated by a distance \(h\). Based on the minimization of the sum of the squared deviations between the experimental and theoretical semivariograms, a spherical model (Eq. (2)), an exponential model (Eq. (3)) and a Gaussian model (Eq. (4)) were selected to further investigate the spatial structure:

Spherical semivariogram function is:

\[
\gamma(h) = C_0 + C_1 \left[ \frac{3h}{2a} - \frac{h^3}{2a^3} \right] \text{ for } h \leq a = C_0 + C_1 \text{ for } h > a
\]

Exponential semivariogram function is:

\[
\gamma(h) = 0 \text{ for } h = 0 = C_0 + C_1 (1 - e^{- \frac{h}{a}}) \text{ for } h > 0
\]

Gaussian semivariogram function is:

\[
\gamma(h) = 0 \text{ for } h = 0 = C_0 + C_1 (1 - e^{- \frac{h^2}{a^2}}) \text{ for } h > 0
\]

where \(h\) is lag distance, \(C_0\) is the nugget effect, which is the local variation occurring at scales finer than the sampling interval or fine scale variability, measurement or sampling error, \(C_0 + C_1\) is the sill or total variance, and \(a\) is the range of spatial dependence.

Continuous maps of individual attributes were generated by point kriging without drift, which estimates the values of the points at the grid nodes (Candela et al., 1988). Geostatistical analysis of the data was performed with the ArcGIS10.0 geostatistical analyst tool in this study.

---

Table 1
Descriptive statistics for TN and SOC.

<table>
<thead>
<tr>
<th>Soil source</th>
<th>Soil properties</th>
<th>Depth (cm)</th>
<th>Mean (g·kg⁻¹)</th>
<th>Dispersion of variance</th>
<th>Median (g·kg⁻¹)</th>
<th>Min (g·kg⁻¹)</th>
<th>Max (g·kg⁻¹)</th>
<th>CV (%)</th>
<th>( p_{K-S} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study area</td>
<td>TN</td>
<td>0–20</td>
<td>0.21</td>
<td>0.000</td>
<td>0.21</td>
<td>0.15</td>
<td>0.26</td>
<td>10.43</td>
<td>0.680</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20–40</td>
<td>0.22</td>
<td>0.001</td>
<td>0.22</td>
<td>0.14</td>
<td>0.34</td>
<td>13.79</td>
<td>0.750</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40–60</td>
<td>0.21</td>
<td>0.001</td>
<td>0.21</td>
<td>0.13</td>
<td>0.36</td>
<td>15.77</td>
<td>0.304</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60–80</td>
<td>0.21</td>
<td>0.001</td>
<td>0.21</td>
<td>0.15</td>
<td>0.29</td>
<td>13.51</td>
<td>0.679</td>
</tr>
<tr>
<td></td>
<td>SOC</td>
<td>0–20</td>
<td>2.96</td>
<td>0.320</td>
<td>2.96</td>
<td>1.73</td>
<td>4.61</td>
<td>19.14</td>
<td>0.995</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20–40</td>
<td>2.89</td>
<td>0.497</td>
<td>2.82</td>
<td>1.49</td>
<td>5.39</td>
<td>24.35</td>
<td>0.761</td>
</tr>
<tr>
<td></td>
<td></td>
<td>40–60</td>
<td>2.91</td>
<td>0.510</td>
<td>2.86</td>
<td>1.30</td>
<td>5.73</td>
<td>24.53</td>
<td>0.873</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60–80</td>
<td>2.81</td>
<td>0.401</td>
<td>2.79</td>
<td>1.57</td>
<td>4.44</td>
<td>22.55</td>
<td>0.434</td>
</tr>
<tr>
<td>Reference site</td>
<td>TN</td>
<td>0–15</td>
<td>0.47</td>
<td>0.01</td>
<td>0.48</td>
<td>--</td>
<td>--</td>
<td>2.52</td>
<td>0.472</td>
</tr>
<tr>
<td></td>
<td>SOC</td>
<td>0–15</td>
<td>6.62</td>
<td>3.85</td>
<td>6.07</td>
<td>--</td>
<td>--</td>
<td>4.41</td>
<td>0.051</td>
</tr>
</tbody>
</table>

Note: \( p_{K-S} \) is the significance level of Kolmogorov–Smirnov test; the data for the reference site are from the experimental results of Xiao et al. (2011).
2.4. Determination of sampling number for minesoils properties monitoring

Cross-validation is used to test the accuracy of the geostatistical method in this study, and the root-mean-square error (RMSE) is used to measure the accuracy of the kriging method using the following model:

$$\text{RMSE} = \left( \frac{1}{n} \sum_{i=1}^{n} [Y(x_i) - Y'(x_i)]^2 \right)^{1/2}$$ (5)

where $Y(x_i)$ is the measured value, $Y'(x_i)$ is the predicted value, and $n$ is the sampling number. The kriging interpolation is hypothesized to be the most accurate when the RMSE is at a minimum and is stable.

In this study, 10, 20, 30, 40, 50, 60 and 70 sampling points were randomly selected from a total of 78 sampling points to perform the kriging interpolation. By analyzing the prediction accuracy of the minesoil properties under different sampling points in the study area, the rational sampling number for minesoils properties monitoring was determined according to the minimum prediction error.

3. Results and discussion

3.1. Variability of soil properties

Descriptive statistics, including mean, dispersion of variance, CV, minimum, and maximum values, for SOC and TN in the minesoils are presented in Table 1. The mean and median values were used as primary estimates of the central tendency, and the dispersion of variance, CV, minimum, and maximum values were used as the estimates of variability for each site. Normality tests were conducted using the significance level of the K-S test, and all of the values of SOC and TN passed the normality test ($p > 0.05$).

The mean and the median values were mostly similar, with the majority of median values either equal to or smaller than the mean values for the two soil properties. This indicated that the outliers did not dominate the measures of central tendency. A similarity of means and median for several physical, chemical, and soil properties has also been reported in other studies (Cambardella et al., 1994; Nyamadzawo et al., 2008). TN concentration showed a low CV (<15%), except at a depth of 40–60 cm, where it was moderate (15–35%). SOC concentration showed a moderate variability (15–35%) for all depths.

Higher TN and SOC concentrations in the surface layer were obtained from reference site (Table 1) and the SOC in the reference site was 3.5 times higher than the study site, primarily due to the presence of plants in the reference site that add more C and N than in study site with no plants.

3.2. Spatial variability of soil properties

Soil properties may vary due to intrinsic or extrinsic sources of variability. Intrinsic variability is the natural variation in soils (Cambardella et al., 1994), and extrinsic variability is caused by factors imposed on a site (Rao and Wagener, 1985). Descriptive statistics cannot discriminate between these two sources of variability. Therefore, the spatial correlation structure of each property was further investigated. The experimental site studied here also displayed differences in its spatial dependence as determined by its semivariograms (Table 2). The semivariance ideally increases with the distance between a sample location or lag distance to a more or less constant value, the total sill. The distance that the semivariance attains after a constant value known as the range of spatial dependence (Cambardella et al., 1994). Samples separated by a distance closer than the range are spatially correlated, and those separated by a distance greater than the range are independent. Semivariance ranges depend on the spatial interaction of soil processes affecting each property at the sampling scale (Trangmar et al., 1985).

The experimental semivariograms for all measured soil properties for all four depths intervals exhibited spatial structure and were obtained to a lag of 118 m. The semivariogram models and best-fit model parameters are provided in Table 2. The existence of a positive nugget effect in some of the variables studied can be explained by sampling error, short range variability, and unexplained and inherent variability (Burgos et al., 2006). The nugget semivariance was generally low for TN concentration but high for SOC concentration for all depths. A higher nugget values tends to mask the spatial variability of the attributes. No definite trend for the nugget variance was obtained with increasing depth. All semivariograms are generally well structured with a small nugget effect, indicating that the sampling density is adequate to reveal the spatial structures (McGrath et al., 2004).

The nugget to sill ratio (NSR) was used to define distinct classes of spatial dependence. If the ratio was <25%, the variable was considered to be strongly spatially dependent, if the ratio was between 25% and 75%, the variable was considered to be moderately spatially dependent, and if the ratio was >75%, the variable was considered to be weakly spatially dependent (Cambardella et al., 1994). Using the NSR, the semivariograms indicated moderate spatial dependency for most of the parameters (Table 2). At a depth of 0–20 cm, a strong spatial dependency was obtained for the TN concentration, and the range values were only similar at this depth. The TN and SOC concentrations had similar ranges at the depth of 0–20 cm. At a depth of 20–40 cm, the TN and SOC concentrations had the greatest difference in values, which were 720 cm and 176.74 m, respectively. At a depth of 40–60 cm, the TN concentration was 378.44 m and the SOC concentration was 194.83 m. At a depth of 60–80 cm, the TN and SOC concentrations were 361.09 m and 170.59 m, respectively. The dispersion variance

<table>
<thead>
<tr>
<th>Soil properties</th>
<th>Depth (cm)</th>
<th>Best model</th>
<th>Nugget</th>
<th>Sill</th>
<th>Range (m)</th>
<th>NSR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TN (g kg⁻¹)</td>
<td>0–20</td>
<td>Exponential</td>
<td>0.0001</td>
<td>0.0005</td>
<td>331.81</td>
<td>24.11</td>
</tr>
<tr>
<td></td>
<td>20–40</td>
<td>Gaussian</td>
<td>0.0007</td>
<td>0.0008</td>
<td>720.00</td>
<td>92.34</td>
</tr>
<tr>
<td></td>
<td>40–60</td>
<td>Spherical</td>
<td>0.0009</td>
<td>0.0010</td>
<td>378.44</td>
<td>86.09</td>
</tr>
<tr>
<td></td>
<td>60–80</td>
<td>Gaussian</td>
<td>0.0007</td>
<td>0.0008</td>
<td>361.09</td>
<td>94.31</td>
</tr>
<tr>
<td>SOC (g kg⁻¹)</td>
<td>0–20</td>
<td>Spherical</td>
<td>0.1868</td>
<td>0.3485</td>
<td>334.47</td>
<td>53.60</td>
</tr>
<tr>
<td></td>
<td>20–40</td>
<td>Exponential</td>
<td>0.115</td>
<td>0.5805</td>
<td>176.74</td>
<td>19.21</td>
</tr>
<tr>
<td></td>
<td>40–60</td>
<td>Gaussian</td>
<td>0.4384</td>
<td>0.5302</td>
<td>194.83</td>
<td>82.67</td>
</tr>
<tr>
<td></td>
<td>60–80</td>
<td>Exponential</td>
<td>0.1124</td>
<td>0.4375</td>
<td>170.59</td>
<td>25.69</td>
</tr>
</tbody>
</table>
was very high for the SOC, and there was not clear trend with increasing depth. The large dispersion variance indicated the heterogeneity of the minesoils (Nyamadzwo et al., 2008). The heterogeneity in the dispersion variance with depth may be a result of the mining activities. Because the spatial distributions of the SOC and TN are consistent, the spatial variability may be caused by structural factors (Cheng et al., 2004; Zhang et al., 2010).

The range values of the measured parameters varied within the study site. The trends were not the same for one parameter for the topsoil to the bottom in the study site or in the reference site, which was previously reported (Jia et al., 2004; Xiao et al., 2009). In that study, the ranges in the reference site were larger than those at the study site at a depth of 0–20 cm. The variability in the range values manifested at large scales could arise from features, such as soil types, landscape position, and history of management (Haws et al., 2004). The range values were high for the two properties at all depths. Because the range values were always greater than the chosen grid size, the sampling design was adequate for the study site (Zanini and Bonifacio, 1992).

3.3. Spatial distribution of soil properties

Spherical, exponential and Gaussian models were used in the kriging analysis to calculate the optimum weights at each sampling site. The results of the spatial dependence enabled the plotting of kriged maps of the different variables. Fig. 1 shows the contour maps obtained by simple kriging for SOC and TN. Maps for each variable were plotted on the same scales and with the same contour intervals to allow for easier comparisons.

The TN changed significantly from 0–20 cm to 60–80 cm; it also changed significantly within each depth interval. The TN concentration had the greatest change at 0–20 cm. The TN concentration decreased from south to north in the study area at all four depths. The SOC concentration had the greatest change from 0–20 cm to 60–80 cm, and it is also changed significantly within each depth, similar to the behaviour of the TN concentration. The SOC showed a patchy distribution at all four depths at the study site. The distribution of the SOC concentration was similar to that of the TN concentration at depths of 0–20 cm and 40–60 cm, but for the other two depths, it did not have the same distribution. The distribution of the SOC concentration was not similar to that of the TN concentration at depths of 20–40 cm and 60–80 cm, which also indicated that the spatial variability did not arise from structural factors, including the topography, vegetation and parent materials. For a large-scale coal mine, the disturbance due to large mechanical recycling operations and humans has far-reaching impacts on soil development (Mukhopadhyay et al., 2014). The excavation, transport and dumping activities significantly disrupted the minesoils, which resulted in spatial variability in the soil properties. Severe soil compaction and vegetation destruction also led to the loss of SOC and TN.

![Fig. 1. Spatial distributions (contour map) of SOC and TN (unit: g kg⁻¹).](image-url)
3.4. Optimization of sampling number for monitoring soil properties

Because SOC and TN have high variability in minesoils, a large number of points is required to monitor the effect of land reclamation on soil properties (Yim et al., 2003). The rational sampling number for minesoil properties monitoring was determined according to the minimum prediction error of soil properties under different sampling points (de Souza et al., 2014; Delhomme, 1978). A total of 10, 20, 30, 40, 50, 60, 70, and 78 sample points were randomly selected from 78 sample points, (a) 0–20 cm TN; (b) 0–20 cm SOC; (c) 20–40 cm TN; (d) 20–40 cm SOC; (e) 40–60 cm TN; (f) 40–60 cm SOC; (g) 60–80 cm TN; (h) 60–80 cm SOC). The vertical line shows the choice of the optimal sampling number at different depths for TN and SOC.
and the RMSEs for all four depth intervals were calculated. For accurate results, the selection was repeated three times (i.e., three treatments). The RMSEs of the soil properties are shown in Fig. 2. The RMSE value of the TN concentration decreased with increasing sample size at the four depths and was stable with sample sizes of 30, 40, 40, and 40 for depths of 0–20 cm, 20–40 cm, 40–60 cm, 60–80 cm, respectively. The RMSE value of the SOC concentration had similar trends at the four depths and was stable at 40, 40, 40, and 40 sample sites for depths of 0–20 cm, 20–40 cm, 40–60 cm, 60–80 cm, respectively. It is not possible for the sampling number in the topsoil to be less than that of the bottom soil; therefore, 40 sample sites is the rational sampling number for the study site.

4. Conclusions

The following conclusions can be drawn from our findings:

(1) There was a moderate spatial variability in the TN and SOC concentrations in minesoils after dumping and before reclamation, according to statistical and geostatistical analyses. The variance showed no obvious trend with increasing depth for the TN and SOC concentrations.

(2) Geostatistical analysis was useful for estimating TN and SOC concentrations and for interpreting the spatial variability, and substantial heterogeneity of these variables was observed from the contour maps.

(3) Further monitoring of the SOC and TN is necessary to evaluate the effects of reclamation based on geostatistical methods, and 40 sample sites is the ideal number for the study site based on cross-validation.

Acknowledgments

This research was supported by National Natural Science Foundation of China (41271528) and Beijing Higher Education Young Elite Teacher Project.

References


