

Comparison of kriging interpolation precision between grid sampling scheme and simple random sampling scheme for precision agriculture

Jiang Houlong ^a, Wang Daibin ^a, Xu Chen ^{a,*}, Liu Shuduan ^b, Wang Hongfeng ^a,
Yang Chao ^a, Li Najia ^a, Chen Yiyin ^a, Geng Lina ^a

^a Chongqing Tobacco Science Research Institute, Chongqing, China

^b Beibei Tobacco Company of Chongqing, Chongqing, China

Abstract

Sampling methods are important factors that can potentially limit the accuracy of predictions of spatial distribution patterns. A 10 ha tobacco-planted field was selected to compare the accuracy in predicting the spatial distribution of soil properties by using ordinary kriging and cross validation methods between grid sampling and simple random sampling scheme (SRS). To achieve this objective, we collected soil samples from the topsoil (0-20 cm) in March 2012. Sample numbers of grid sampling and SRS were both 115 points each. Accuracies of spatial interpolation using the two sampling schemes were then evaluated based on validation samples (36 points) and deviations of the estimates. The results suggested that soil pH and nitrate-N (NO₃-N) had low variation, whereas all other soil properties exhibited medium variation. Soil pH, organic matter (OM), total nitrogen (TN), cation exchange capacity (CEC), total phosphorus (TP) and available phosphorus (AP) matched the spherical model, whereas the remaining variables fit an exponential model with both sampling methods. The interpolation error of soil pH, TP, and AP was the lowest in SRS. The errors of interpolation for OM, CEC, TN, available potassium (AK) and total potassium (TK) were the lowest for grid sampling. The interpolation precisions of the soil NO₃-N showed no significant differences between the two sampling schemes. Considering our data on interpolation precision and the importance of minerals for cultivation of flue-cured tobacco, the grid-sampling scheme should be used in tobacco-planted fields to determine the spatial distribution of soil properties. The grid-sampling method can be applied in a practical and cost-effective manner to facilitate soil sampling in tobacco-planted field.

Keywords: Grid sampling scheme, Interpolation precision, simple random sampling scheme, precision agriculture

Article Info

Received : 13.08.2015

Accepted : 18.09.2015

© 2016 Federation of Eurasian Soil Science Societies. All rights reserved

Introduction

Among nutrients necessary for plant growth, potassium, nitrogen and phosphorus are the most important and sensitive for yield and quality of flue-cured tobacco. Tobacco plants receive potassium, nitrogen and phosphorus not only from fertilizers, but also from soil minerals. However, soil is characterized by a high degree of spatial variability because of geological and soil-forming factors that operate at different intensities and scales (Goovaerts, 1998; Quine and Zhang, 2002). Therefore, uniform fertilizer application likely leads to over-application in areas with high nutrient levels and under-application in areas with low nutrient levels (Ferguson et al., 2002), which results in extremely uneven tobacco yield and quality. Variable-rate fertilizer application, which is an efficient method of solving such problems, is possible if spatial variation in nutrient contents across a field is known (Cahn et al., 1994). The spatial distribution of soil

* Corresponding author.

Chongqing Tobacco Science Research Institute, Chongqing 400715, China

Tel.: +86 023 68250791

e-ISSN: 2147-4249

E-mail address: jhl513@163.com

DOI: <http://dx.doi.org/10.18393/ejss.2016.1.062-073>

properties can be predicted accurately by soil spatial sampling (Zhu et al., 2005). In soil sampling, point information is used to estimate soil fertility levels at locations where samples are not obtained. Soil sampling schemes are important because of their effect on the quality of spatial interpolation maps, and the corresponding values to manage such maps depend on the accuracy of predicted values (Frogbrook, 1999). Therefore, a suitable sampling method is necessary to estimate correctly the values of soil properties in areas of specified sizes.

An optimal and time-effective sampling method in any study should provide maximum estimation precision with the least sample cost. The number of soil samples and sampling strategies required to represent field variability has been extensively studied since the 1920s (Lindsley and Bauer, 1929). This requirement is possibly based on economic values (Peck, 1990). Increased agricultural income from reduced input, increased yield, and improved quality should offset not only the costs of characterizing soil variability but also the technology required for variable application. Studies have indicated that appropriate sampling or a combination of systematic is effective when the density of rare or clustered populations is estimated (Thompson and Seber, 1996; Christman, 2000). The sample pattern and sample spacing are important factors that affect the accuracy of the kriging interpolation (Wollenhaupt et al., 1994; Gotway et al., 1996), which is a technology for estimating the soil property values in nonsampled areas and generating the spatial distribution (Goovaerts, 2000). Few previous studies evaluated the effect of sample design and intensity on the accuracy of assessment (McBratney et al., 1999; Caeiro et al., 2002; Zhu and Stein, 2006; Brus et al., 2006; Li et al., 2007; Kumar, 2009; Liu et al., 2010). That research provides a starting point in developing a sampling strategy for analyzing effectively the spatial variability of soil properties in tobacco-planted field.

Furthermore, sampling strategies should be carefully selected to represent fields in which known sources of variability are considered. Field observation has been traditionally based on discrete sampling procedures by using either grid sampling or SRS (Li et al., 2007). A grid pattern is commonly an optimum sampling scheme to ensure that an entire field is represented. However, sampling with a grid pattern can be laborious if large areas are examined (Li et al., 2007). Therefore, the present study compared kriging interpolation precision between the grid sampling method and the SRS.

There are more than one million ha of agricultural soil for flue-cured tobacco plantation that serve as an important source of income for tobacco growers as well as local administration in China. In order to improve the quality of flue-cured tobacco, the Tobacco Company of China has requested that farmers adopt variable-rate fertilizer application. The objectives of this study were: (i) to quantify the spatial variability of soil properties across flue-cured tobacco (*Nicotiana tabacum*) plantation fields, (ii) to identify the appropriate sampling scheme for nine soil properties in order to minimize cost and maximize evaluation accuracy, and (iii) to provide a theoretical basis for establishing a reasonable sampling scheme in tobacco-planting fields.

Material and Methods

Study site

The study was performed at an agricultural experiment station in Southeast Pengshui County (29°8'14'4672"N, 107°57'3081"E) of Chongqing City in Southwest China in 2012. The site is characterized by a subtropical moist monsoon climate with an average annual temperature of 17.5 °C, a mean annual potential evapotranspiration of 950.4 mm, and an annual precipitation of approximately 1104.2 mm. The main soil type at this site ranges from light clay (80.6%) to heavy loam (3.5%), and the remaining 15.9% is medium clay. In addition, the soil is slightly acidic (pH = 5.87).

Soil sampling and laboratory analysis

A 10-hectare area was selected for soil sampling, and the overview of the boundary of the study site is illustrated in Figure 1. This area is surrounded by a hill and the field was bare at the time of observation. Soil samples were collected in March 4, 2012. All of the samples were taken from the topsoil (0-20 cm), and a real-time kinematic global positioning system (GPS) survey was used to identify sampling locations. At each point, the values were recorded by differential GPS (DGPS) and then converted to coordinates (x, y; Fig. 1). The soil samples of the grid sampling scheme were collected using an approximately 32 m grid sampling design (n=115, Fig. 1). The soil samples of the regular simple random sampling scheme were regularly taken from spaces that ranged between approximately 16 m and 32 m (n=115, Fig. 1).

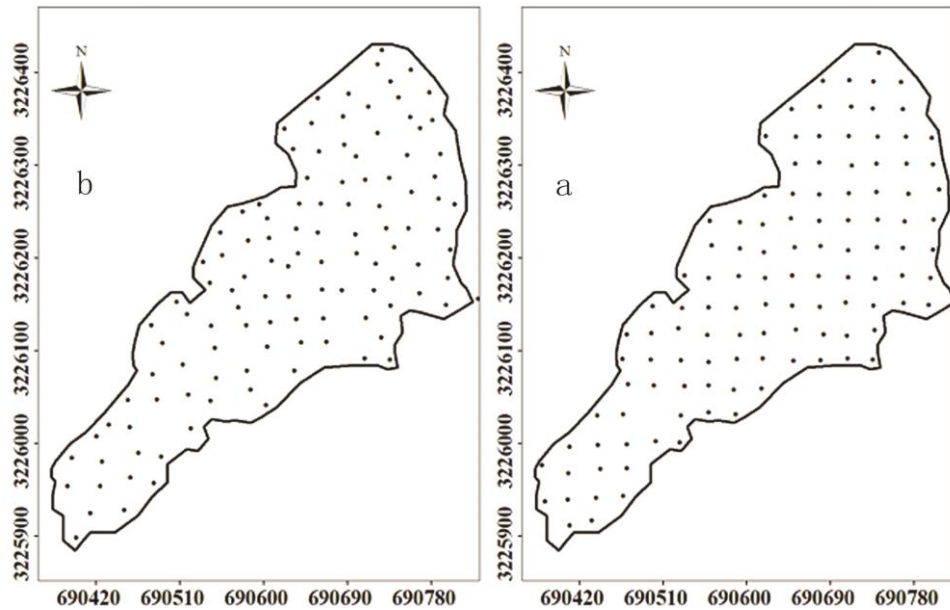


Figure 1. Soil sample distribution under two sampling schemes (a: grid sampling scheme, b: simple random sampling scheme) in the 10-ha area.

Soil samples were packed in plastic bags, air-dried, divided, and ground to a sufficient size to pass through a 2 mm sieve before analysis was conducted. The soil pH of the samples was subsequently measured using a pH meter with a glass electrode (soil-H₂O ratio = 1:2.5, W-V). The organic matter (OM) content was analyzed using the wet oxidation method of Walkley and Black (Nelson and Sommers, 1982). Cation exchange capacity (CEC) was determined by extraction using neutral sodium acetate (Chapman, 1965). Total nitrogen (TN) was determined using the Kjeldahl method (Bremner and Mulvaney, 1984). Nitrate nitrogen (NO₃-N) of a fresh sample was determined using a continuous flowing analyzer (Paramasivam et al., 2002). Total phosphorus (TP) was then determined by sulfate-perchlorate acid heating digestion-MoSb colorimetry (Lu, 1999). Available phosphorus (AP) was determined using the Olsen extraction method with alkaline sodium bicarbonate as an extractant at a ratio of 20:1 (Olsen et al., 1954). Total potassium (TK) was analyzed by inductively coupled plasma-atomic emission spectroscopy. Available potassium (AK) was determined using the neutral ammonium acetate method (Richards, 1954).

Evaluation method

In order to evaluate the accuracy of our estimates, 36 sites were selected by probability sampling for external validation. The interpolation values of these 36 points were compared with the actual measurements.

As an alternative method of evaluating the accuracy of our estimates, we determined the performance of each interpolation obtained using the two sampling methods. Deviations of the estimates from the measured data were then compared by cross-validation (Webster and Oliver, 2001). Comparison of performance of the two sampling methods was done using the following statistics:

In order to evaluate the accuracy of the estimates, the performance of each interpolation under different intervals was assessed by comparing the deviation of estimates from the measured data through cross-validation (Webster and Oliver, 2001). Comparison of performance between the two sampling schemes was done using the following statistics: mean absolute error (MAE), mean error (ME), mean square error (MSE), average standardized error (ASE), root mean square error (RMSE), and root mean square standardized error (RMSSE). The ME was used to determine the degree of bias in the estimates; MSE provided a measure of the size of MSE; ASE was used to identify the degree of ASE; RMSE provided a measure of the error size that it is sensitive to outliers; and RMSSE was used to determine the degree of RMSSE. The five error statistics of predictions were used in the cross-validation analysis. The equations are as follows (Johnston et al., 2001):

$$ME = \frac{1}{N} \sum_{i=1}^N [Z(x_i) - Z(\hat{x}_i)] \quad (1)$$

$$MSE = \frac{1}{N} \sum_{i=1}^N \frac{Z(x_i) - Z(\hat{x}_i)}{\sigma(i)} \quad (2)$$

$$ASE = \sqrt{\frac{1}{N} \sum_{i=1}^N \sigma(i)} \quad (3)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N [Z(x_i) - \hat{Z}(x_i)]^2} \quad (4)$$

$$RMSSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \left\{ \frac{Z(x_i) - \hat{Z}(x_i)}{\sigma(i)} \right\}^2} \quad (5)$$

where $\hat{Z}(x_i)$ is the predicted value, $Z(x_i)$ is the observed value, n is the number of values, and σ is the standard error for location i .

The criteria for cross-validation were as follows. ME and MSE indicate the degree of bias in the model prediction, and should be approximately zero. RMSE and ASE indicate the precision of prediction. Their values should be as small as possible. RMSSE compares the error variance with the kriging variance. RMSSE should be approximately 1. If RMSE equals ASE, then all errors are of small. If ASE > RMSE or RMSSE < 1, the model-predicted values are larger than the actual values (Tang and Yang, 2006). MAE indicates the precision of prediction, such that low MAE values indicate more accurate point-by-point predictions.

The measured variables in the data set were analyzed using descriptive statistical methods to obtain the mean, S.D., minimum, median, maximum, skewness, kurtosis, and coefficient of variation (CV) using SPSS 17.0 software. Distributions of these variables were evaluated to determine normality using Kolmogorov-Smirnov statistics. Semi-variance calculation, semi-variogram function model fitting, and cross validation were performed using the geostatistical software ArcGIS 9.3 for Windows.

Results

Descriptive statistics

Summary statistics of soil parameters is shown in Table 1. Distributions of all variables were only slightly skewed (skewness < 1), and their median values were close to their mean values. In addition, results showed that all data were normally distributed (Kolmogorov-Smirnov test, $P > 0.05$; Table 1).

Table 1. Descriptive statistics of soil properties ($n = 289$) in the topsoil (0-16cm) at different sampling dates.

Soil properties	Min.	Median	Mean	Max.	S.D.	# CV %	Skewne ss	Kurtosis	* P _{K-S}
pH	5.00	5.85	5.87	7.00	0.45	7.67	0.32	-0.25	0.59
OM, g kg ⁻¹	14.26	25.00	25.24	39.34	5.01	19.85	0.32	-0.11	0.59
CEC, cmol kg ⁻¹	6.23	17.19	17.65	31.09	5.34	30.25	0.35	-0.34	0.21
TN, g kg ⁻¹	0.34	0.79	0.81	1.35	0.20	24.69	0.09	-0.4	0.63
NO ₃ -N, mg kg ⁻¹	137.3	174.3	177.33	221.3	16.79	9.47	0.52	-0.23	0.08
TP, g kg ⁻¹	0.35	0.75	0.79	1.33	0.22	27.85	0.61	-0.12	0.01
AP, mg kg ⁻¹	13.54	19.95	20.46	29.51	3.42	16.72	0.57	-0.48	0.01
TK, g kg ⁻¹	10.85	17.26	17.22	24.57	2.93	17.02	0.27	-0.32	0.52
AK, mg kg ⁻¹	62.41	202.31	199.78	374.11	65.76	32.92	0.22	0.56	0.16

CV, coefficient of variation (%).

* K-S Test, Kolmogorov-Smirnov test was used to test the significance level of normality, all variables were normally distributed ($P > 0.05$).

The greatest variability was observed in AK (32.92%), whereas the least variability was found in pH (7.67%). The variability in pH and NO₃-N were low (CV < 15%), whereas all of the other soil variables exhibited moderate variation (CV: 15%-50%). AK also exhibits a localized pattern of enrichment according to the [National Soil Survey Office \(1993\)](#). In conclusion, the soil possesses high variability, which results in extremely uneven tobacco yield, quality, and agricultural benefits, which suggests that variable-rate fertilization application is necessary for improving tobacco quality in this area.

Geostatistical analysis

Semi-variograms were calculated to identify the possible spatial structure of different soil variables. Cross-validation was performed to compare the prediction performances of geostatistical interpolation algorithms with a particular sampling method. Cross-validation indicators and additional model parameters (nugget, sill, and range) helped determine the optimal model of the prediction maps for each soil property ([Issaks and Srivastava, 1989](#)). In this study, the optimal model that could describe specific spatial structures was identified. The result of the geostatistical analysis is shown in Table 2. In this analysis, several spatial distribution models were fitted, which revealed different spatial dependence levels of the soil variables based on the cross-validation model.

Our results suggested that NO₃-N, TK, and AK ideally matched an exponential model, whereas the remaining variables suitably fit the spherical model with the two sampling methods. The coefficient of determination (R²) of all variables, except for NO₃-N and AK, was greater than 0.90, indicating good fits. The spatial variability of soil nutrients may be affected by extrinsic (soil management practices such as fertilization and cultivation) and intrinsic factors (soil formation factors, such as soil parent materials). The strong spatial dependence of soil variables can typically be attributed to extrinsic factors ([Cambardella et al., 1994](#)). The value nugget % was used to qualitatively define spatial dependence values. Our results showed that pH, OM, CEC, TN, and NO₃-N exhibited <25% except NO₃-N in the SRS, which suggested strong spatial dependence. [Rodríguez et al. \(2009\)](#) found that spatial dependence of NO₃-N was high in sandy and loam soils. All of the other soil variables were moderately spatially dependent with nugget % ranging from 33.33% to 57.14%. This range could be interpreted as the zone of influence diameter. This result represented the average maximum distance at which the soil properties of the two sampling points are related. The ranges of spatial dependences exhibited large variation (from 95.78 m for OM of grid sampling to 558.41 m for AK of SRS, as shown in table 2).

Accuracy of validation samples

The spatial distributions of nine soil properties were completed with the optimal models using the two sampling schemes in the study area. The predicted values of evaluation samples (115 points) and validation samples (36 points) were obtained using the two sampling schemes.

The scatter diagrams, in which the predicted values are on the vertical axis and the measured values are on the horizontal axis, were constructed and a trend line was subsequently added (Fig. 2). The trend line was closer to the 45-degree line, indicating a more accurate prediction than other lines ([Li et al., 2008](#)). In the scatter diagram, the prediction accuracies of pH, TP, and AP in the SRS were bigger than that of the grid-sampling scheme, while the prediction accuracy of soil OM, CEC, TN, NO₃-N, TK and AK in the grid-sampling scheme was bigger than that of the SRS.

Analysis of interpolation results for SRS and Grid sampling

Results of cross-validation from the 115 points arranged in the SRS are shown in Table 3. These data may be more useful in comparing prediction errors. ME was always nearly equal to MSE, ASE was almost equal to RMSE, and RMSSE values were always close to 1. Consequently, kriging was shown to be accurate for soil pH, TP and AP by the five statistical criteria in the SRS, while the interpolation results of soil OM, CEC, TN, TK and AK in the SRS were poor than that in the grid-sampling scheme. The difference in accuracy of the soil NO₃-N between the two sampling schemes was not markedly significant.

For pH, TP, and AP, the absolute values of ME and MSE were slightly greater than zero, whereas the absolute value of RMSE was remarkably close to and less than that of ASE, and the RMSSE values were less than one, indicating that they were underestimated in the SRS. The absolute values of ME, RMSE, ASE, and MSE of TN and AK were very close to zero, and RMSE > ASE, RMSSE > 1, suggesting that they were slightly overestimated at the grid sampling scheme. However, it was underestimated by kriging for OM and TK in the grid sampling scheme (RMSE < ASE, RMSSE < 1).

Table 2. Characteristic of calculated semivariograms for all soil properties under two sampling methods in study area

Variable	Grid sampling scheme						Simple random sampling scheme							
	Model	Co	§Co+C	Nugget, %	Spatial dependence class	R ²	Range	Model	Co	Co+C	Nugget, %	Spatial dependence class	R ²	Range
pH	S	0.02	0.1	20	S	0.99	358.97	S	0.03	0.14	21.43	S	0.97	148.3
OM	S	4.17	17.15	24.31	S	0.95	95.78	S	3.02	18.21	16.58	S	0.94	107.2
CEC	S	4.05	22.33	18.14	S	0.97	154.43	S	4.09	17.57	23.28	S	0.97	139.95
TN	S	0	0.01	18.65	S	0.92	553.58	S	0	0.04	13.56	S	0.92	460.17
NO ₃ -N	E	24.78	143.31	17.29	S	0.87	325.42	E	64.15	165.41	38.78	M	0.87	439.56
TP	S	0.04	0.11	36.36	M	0.95	553.58	S	0.03	0.09	33.33	M	0.96	495.04
AP	S	6.2	16.01	38.73	M	0.92	478	S	4.82	13.78	34.98	M	0.93	558.41
TK	E	2.13	5	42.6	M	0.9	257.28	E	3.07	8.02	38.28	M	0.9	258.41
AK	E	833.7	2333.7	35.72	M	0.81	553.58	E	1409.2	4244.9	33.2	M	0.84	558.41

§ Co: nugget variance; E: exponential models.

*Nugget %: $Co / (Co + C) \times 100$. Nugget between 25% and 75%: moderate spatial dependence. Nugget < 25%: strong spatial dependence.

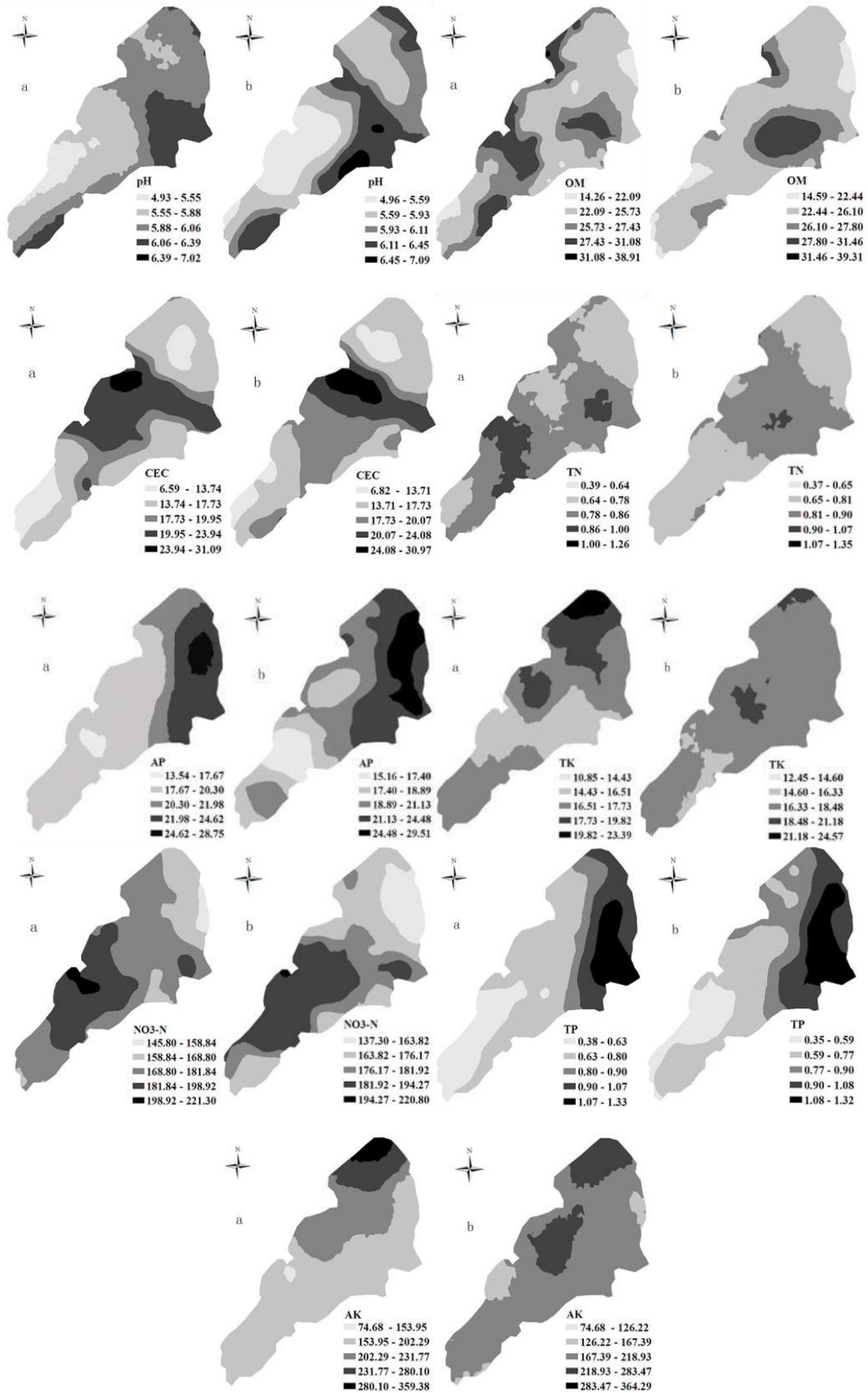


Figure 2. Distribution maps of soil pH, OM, CEC, TN, TP, TK, NO₃-N, AP, and AK by kriging interpolation under two sampling methods.

Thus, the results of cross-validation of soil OM, CEC, TN, TK and AK indicate that the interpolation results of the grid-sampling scheme were better than the regular simple random sampling scheme, and soil pH, TP and AP provided worse interpolation results in grid-sampling scheme. The interpolation results of soil NO₃-N have no markedly significant in the two sampling schemes.

Table 3. Results of cross-validation (ME, MSE, ASE, RMSE and RMSSE) from the 115 points, respectively, in the simple random sampling scheme and grid sampling scheme

	Item	pH	OM	CEC	TN	NO ₃ -N	TP	AP	TK	AK
Simple random sampling scheme	ME	0.002	0.039	0.057	0.004	-0.010	0.001	0.023	-0.001	0.671
	MSE	0.003	0.008	0.010	0.019	-0.001	0.004	0.006	0.000	0.009
	ASE	0.378	4.882	4.274	0.195	13.370	0.142	2.748	2.921	68.900
	RMSE	0.376	4.870	4.350	0.198	14.350	0.130	2.635	2.863	66.690
	RMSSE	0.996	0.997	1.017	1.015	1.071	0.892	0.907	0.981	0.968
Grid sampling scheme	ME	-0.005	-0.033	-0.018	0.002	0.187	-0.001	0.010	-0.005	-0.016
	MSE	-0.010	-0.003	-0.002	0.010	0.009	-0.001	0.004	-0.006	-0.001
	ASE	0.421	4.859	4.410	0.179	14.770	0.145	2.757	2.537	56.180
	RMSE	0.397	4.857	4.252	0.184	14.680	0.130	2.496	2.474	56.300
	RMSSE	0.944	0.998	0.969	1.028	0.999	0.914	0.966	0.978	1.006

Interpolation Maps

Kriging interpolation was performed to obtain a filled contour map using ArcGIS 9.3 with the two sampling methods to determine the spatial distribution and status of soil properties. The contour maps of the soil properties are shown in Figure 3. For each of the soil properties, the spatial distribution trend was similar using the two sampling schemes. Analysis of distributions of soil nutrients revealed that TP and AP have similar spatial distributions. High levels of TP and AP were detected at the east section and low distribution was observed in the southwest sections of the study area. The contour map of TK content showed the highest positional similarity with the krigged AK map. The distributions of pH, OM, and CEC were consistent with high values in the middle section of the field.

Discussion

Coefficients of variation (CVs) were very different among the variables, ranging from approximately 7.67% (pH) to nearly 32.92% (AP), indicating large heterogeneity in soil properties. Nitrogen (N), phosphorus (P) and potassium (K), which are the major minerals needed for tobacco growth and development. In addition, they had particularly high CVs in this study area. This is consistent with the findings of Wang et al. (2009). Our results differed from those reported by Gupta et al. (1999), who showed that soil NO₃-N exhibits moderate variation. This is mainly because of different planting patterns, uneven crop growth and non-uniform management practices, resulting in marked changes in topsoil over small distances. Therefore, site-specific management of nutrients may be necessary to achieve maximum economic and environmental benefits (Jiang et al., 2010). However, classical statistics do not show the spatial distribution of soil properties. The spatial distribution maps of soil nutrients based on geostatistical analysis provide a basis for variable-rate application.

It is important to determine the spatial dependence of soil properties because soil properties with strong spatial dependence are more readily managed (Jiang et al., 2010). The nugget-sill ratio can be used to classify the spatial dependence of soil properties (Jiang et al., 2011). The lower ratios for pH, OM, CEC, and TN were less than 25, indicating that the four soil variables had strong spatial dependence and the structural factors strongly influenced the spatial variability of these properties (Cambardella et al., 1994). Consistent with other reports, the spatial dependence of pH, TN, OM, and CEC were fairly strong (Jiang et al., 2010; Liu et al., 2008). Meanwhile, all other soil variables, except for NO₃-N of grid sampling, were moderately spatially variable, with the nugget/sill ratios ranging from 33.33% to 42.60%. Ranges of spatial dependence that varied from 95.78 m for OM of grid sampling to 558.41 m for AK of SRS were larger than the smallest sampling distance (16 m), suggesting that the two sampling schemes can satisfy the requirement of spatial variability structure analysis of soil properties in the study. Knowledge of the range of influence for various

soil variables allows for construction of independent data sets that can be used for classical statistical analysis. A small range indicated that a small sampling interval is required for OM.

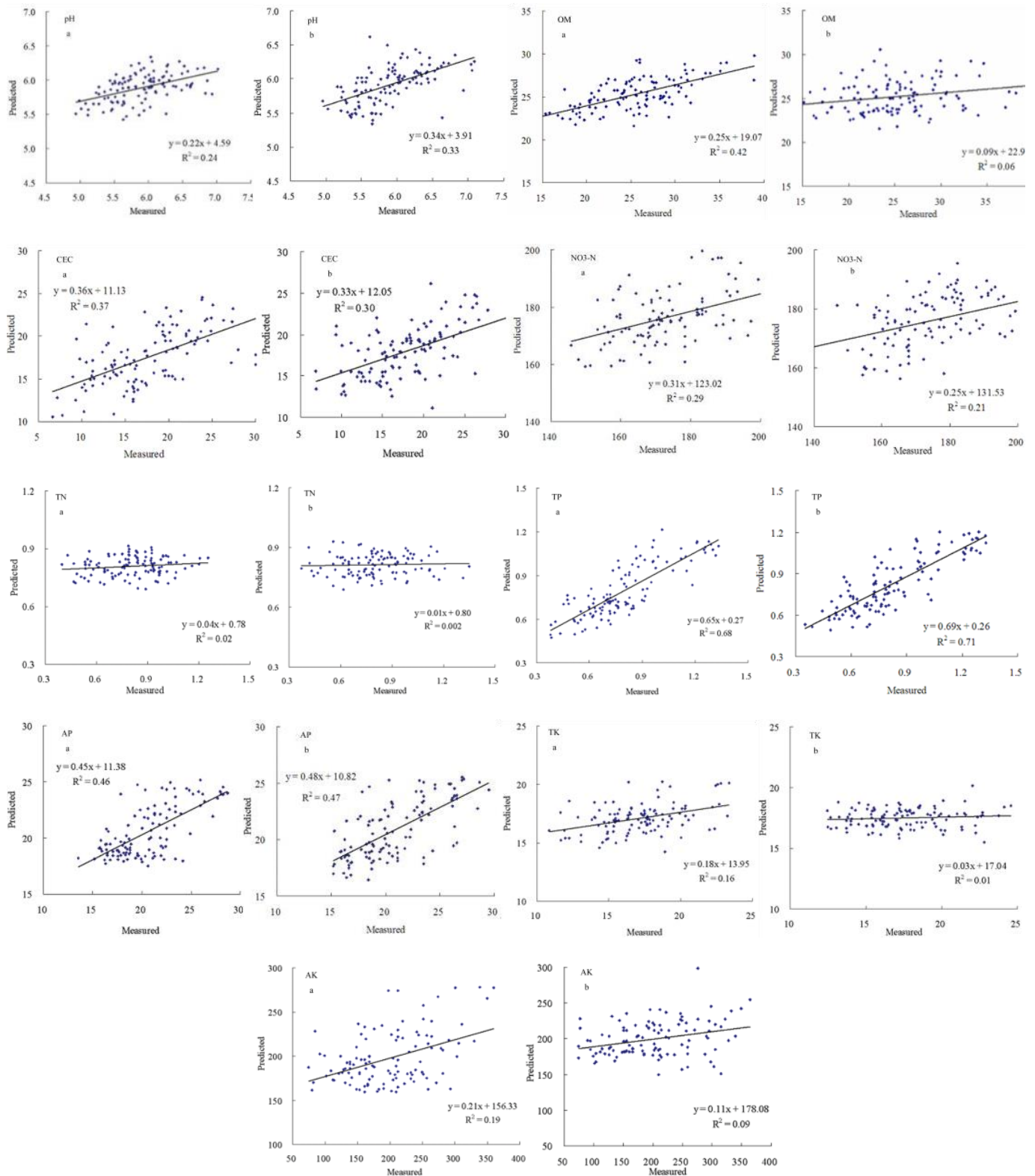


Figure 3. Smoothed contour maps produced by kriging for pH, OM, CEC, TN, TP, TK, NO₃-N, AP, AK under two sampling methods in the 10-ha area (a: grid sampling scheme, b: simple random sampling scheme).

For the same soil properties, the nugget-sill ratios of pH, CEC and NO₃-N of grid sampling were lower than those of SRS (Table 2). The other soil properties were just the opposite. Results of cross-validation of the soil properties indicated that a reasonable sampling scheme of pH, CEC and NO₃-N at the site was SRS, whereas

that for the remaining soil properties was grid sampling. Therefore, we can speculate that the sampling scheme with bigger nugget-sill ratio was optimal within a certain range of nugget-sill ratio. However, whether this is applicable to other soil properties or other fields must be verified further.

Variable-rate fertilizer application is only possible if the spatial variation in nutrient status across a field is known (Cahn et al., 1994). Geostatistics is an important tool for characterizing the spatial variability of soil properties, and it has been widely used in variable-rate fertilizer application. The accuracy of kriging interpolation also depends on the sample pattern and sample spacing (Gotway et al., 1996). Sampling design is also very important when the objective is to interpolate in an optimal fashion and to generate spatial distribution maps for soil properties within a region (Haining, 1990). Results of the present study revealed that the interpolation errors of soil OM, CEC, TN, TK and AK were the lowest with the grid sampling scheme. On the other hand, interpolation error of soil pH, TP, and AP was the lowest with the SRS. The interpolation results of soil TN, CEC, AK, and TK obtained by grid sampling were better than SRS, while the interpolation accuracy of soil NO₃-N was no markedly significant in the two sampling scheme. Our results differ from the findings reported by Mallarino and Wittry (2004), which found that grid sampling was the most effective for phosphorus, potassium, pH, and organic matter. Discrepancy between these results and those reported in previous studies may be attributed to inconsistencies in agricultural experimental treatments among different studies.

Potassium is one of the major minerals needed for tobacco growth and development (Liu, 2003). The potassium content of tobacco leaves is highly correlated to tobacco leaf quality and cigarette safety, and potassium content is an important index for measuring tobacco leaf quality (Yang et al., 2007). Potassium deficiency is the most common problem in tobacco planting fields in China, and is seen as a major constraint for improving tobacco quality. In this study, we found that the interpolation errors of soil potassium data in the topsoil layer were lowest in the grid sampling scheme. Similarly, It should be noted that the grid-sampling scheme has previously been shown to be more effective for potassium (Mallarino and Wittry, 2004). Thus, it can be deduced that the optimal spatial sampling scheme is grid sampling if the land was previously planted with flue-cured tobacco. Additionally, plants derive nitrogen and phosphorus, which are important for tobacco plant growth and development, not only from fertilizers, but also from soil minerals.

Our results clearly showed that the prediction accuracies of soil OM, CEC, TN, TK, and AK were smaller in grid-sampling scheme than SRS. The pH, TP, and AP had the smallest kriging errors in the SRS, the interpolation errors of NO₃-N were not markedly different between the two sampling schemes. We instead found that the grid sampling method is optimal for tobacco planting fields because soil properties, which are important for tobacco quality, are predicted with better accuracy, and an accurate site-specific fertilization scheme for precision farming can be more easily developed using this method.

This study provides a theoretical basis for the practice of precision agriculture in tobacco-planted field. For tobacco production, examining the distribution and abundance and shortage situation of key-nutrients is necessary, especially for the nutrient crucial for tobacco such as potassium (Liu, 2003). If we continue to use the experience based homogenization fertilization management, the variability of tobacco yield and quality will increase. Therefore, the key-nutrient must be closely considered by the producer when determining the sampling scheme. From the analytical results of this study, the interpolation error of AK and TK were the smallest, revealing that the best sampling scheme was grid sampling. In actual production, the grid-sampling scheme could be used as the basis for site-specific fertilizer management because potassium, which is key-nutrient for tobacco quality, is predicted with better accuracy. In the future, a reasonable sampling scheme should be selected in accordance with spatial autocorrelation, trend effect, anisotropy effect of soil properties, and the restrictions of monetary, time, workforce.

Conclusion

Classical statistical analysis showed considerable spatial variability of all soil properties. Soil pH and NO₃-N had low variation (CV, <15%), whereas all other soil properties exhibited medium variation (CV, 15%-15%). Geostatistical analysis of soil properties indicated that pH, OM, TN, CEC, TP and AP ideally matched the spherical model, whereas the remaining variables suitably fit the exponential model with two sampling methods. Classical statistical analysis and geostatistical analysis of soil properties revealed considerable spatial variability in the study area, suggesting that variable-rate fertilizer application is required. In addition, interpolation error analysis revealed that soil OM, CEC, TN, TK and AK had the smallest kriging

errors in the grid sampling scheme. The smallest interpolation error for pH, TP, AP was obtained by SRS. Moreover, the interpolation precisions of NO₃-N were not markedly different between the two sampling schemes, while the SRS was slightly better for soil pH, TP, AP, and grid sampling was slightly more appropriate for all other soil properties. Therefore, considering the fact that potassium is one of the major minerals needed for tobacco, the grid sampling scheme should be used for variable-rate fertilizer application to improve the quality of flue-cured tobacco.

References

- Bremner, J.M., Mulvaney, C.S., 1984. Total nitrogen In: Methods of Soil Analysis, page, A.L. (ED.). 2nd Edn. Agron. No. 9, Part 2: Chemical and microbiological properties. Am. Soc. Argon. Madison, WI. USA. pp. 595-624.
- Brus, D.J., de Gruijter, I.J., van Groenigen, J.W., 2006. Designing spatial coverage samples using the k-means clustering algorithm. In: Digital Soil Mapping: An Introductory Perspective. P. Lagachene, A. B. McBratney, M. Voltz (eds.). Elsevier, Amsterdam, Netherlands.
- Caeiro, S., Goovaerts, P., Painho, M., Costa, H., Sousa, S., 2002. Optimal spatial sampling design for mapping estuarine sediment management areas. AGILE Conference on Geographic Information Science, Palma, p. 1-6.
- Cahn, M.D., Hummel, J.W., Brouer, B.H., 1994. Spatial analysis of soil fertility for site-specific crop management. *Soil Science Society America Journal* 58: 1240-1248.
- Cambardella, C.A., Moorman, T.B., Parkin, T.B., Karlen, D.L., Novak, J.M., Truko, R.F., Konopka, A.E., 1994. Field-scale variability of soil properties in central Iowa soils. *Soil Science Society America Journal* 58: 1501-1511.
- Chapman, H.D., 1965. Cation exchange capacity. In: methods of soil analysis, black, C.A. (ED.). Part 2. Number 9 in the series agronomy: American institute agronomy, Madison, Wisconsin. pp. 891-901.
- Christman, M.C., 2000., A review of quadrat-based sampling of rare geographically clustered populations. *Journal of Agricultural Biological and Environmental Statistics* 5: 168-201.
- Ferguson, R.B., Hergert, G.W., Schepers, J.S., Gotway, C.A., Cahoon, J.E., Peterson, T.A., 2002. Site-specific nitrogen management of irrigated maize: Yield and soil residual nitrate effects. *Soil Science Society America Journal* 66: 544-553.
- Frogbrook, Z.L., 1999. The effect of sampling intensity on the reliability of predictions and maps of soil properties. In: Proceedings of the 2nd European Conference on Precision Agriculture. Sheffield Academic Press, Sheffield, UK. pp. 71-80.
- Goovaerts, P., 1998. Geostatistical tools for characterizing the spatial variability of microbiological and physico-chemical soil properties. *Biology and Fertility of Soils* 27: 315-334.
- Goovaerts, P., 2000. Estimation or simulation of soil properties? An optimization problem with conflicting criteria. *Geoderma* 97: 165-186.
- Gotway, C.A., Ferguson, R.B., Hergert, G.W., Peterson, T.A., 1996. Comparison of kriging and inverse-distance methods for mapping soil parameters. *Soil Science Society America Journal* 60: 1237-1247.
- Gupta, R.K., Mostaghimi, S., McClellan, P.W., Birch, J.B., Brann, D.E. 1999. Modeling spatial variability of soil chemical parameters for site-specific farming using stochastic. *Water Air and Soil Pollution* 110: 17-34.
- Haining, R. 1990. Spatial data analysis in the social and environmental sciences. Cambridge University Press. 409 pp.
- Issaks, E., Srivastava, R.M. 1989. An Introduction to Applied Geostatistics. Oxford University Press, New York.
- Jiang H.L., Liu G.S., Wang X.Z., Song, W.F., Zhang, R.N., Zhang, C.H., Hu, H.C., Li, Y.T., 2010. Spatial variability of soil properties in a long-term tobacco plantation in central China. *Soil Science* 175: 137-144.
- Jiang H.L., Liu G.S., Wang, R., Liu, S.D., Han, F.G., Yang, Y.F., Ye, X.F., Zhang, Q.J., Wang, X.J., Wang, Z.H., Hu, H.C., 2011. Delineating site-specific quality-based management zones for a tobacco field. *Soil Science* 176: 206-212.
- Johnston, K., Hoef, J.M.V., Krivoruchko, K., Lucas, N., 2001. Using ArcGIS Geostatistical Analysis. GIS User Manual by ESRI, New York.
- Kumar, N. 2009. An optimal spatial sampling design for intra-urban population exposure assessment. *Atmospheric Environment* 43: 1153-1155.
- Li, Q.Q., Wang, C.Q., Yue, T.X., Li, B., Yang, J., Shi, W.J., 2008. Error analysis of soil property spatial interpolation with RBF artificial neural network with different input methods. *Acta Pedologica Sinica* 459: 360-365.
- Li, Y., Shi, Z., Wu, C.F., Li, F., Li, H.Y., 2007. Optimised spatial sampling scheme for soil electrical conductivity based on Variance Quad-Tree (VQT) method. *Agricultural Sciences in China* 6: 1463-1471.
- Lindsley, C. M., Bauer, F. C. 1929. Test your soil for acidity. University of Illinois, College of Agriculture and Agricultural Experiment Station, Circular No. 346. USA.
- Liu, G.S., 2003. Tobacco Cultivation. China Agricultural Press, Beijing, China. (In Chinese)
- Liu, G.S., Wang, X.Z., Zhang, Z.Y., Zhang, C.H., 2008. Spatial variability of soil properties in a tobacco field of central China. *Soil Science* 173: 659-667.

- Liu, G.S., Jiang, H.L., Liu, S.D., Wang, X.Z., Yang, X.M., Hu, H.C., Liu, Q.H., Xie, D.P., Gu, J.G., Li, Y.T., 2010. Comparison of kriging interpolation precision with different soil sampling intervals for precision agriculture. *Soil Science* 175: 405-415.
- Lu, R.K., 1999. Chemical analysis of agricultural soil, Beijing, China. *Agricultural Science Technology Press*, p. 34–56. (In Chinese)
- Mallarino, A.P., Wittry, D.J., 2004. Efficacy of grid and zone soil sampling approaches for site-specific assessment of phosphorus, potassium, pH, and organic matter. *Precision Agriculture* 5: 131–144.
- McBratney, A.B., Whelan, B.M., Walvoort, D.J.J., Minasny, B., 1999. A purposive sampling scheme for precision agriculture. In: Proceedings of the 2nd European Conference on Precision Agriculture. Sheffield Academic Press, Sheffield, UK. p. 101-110.
- National Soil Survey Office. 1993. Soil series of China. Beijing, China Agriculture. Press.
- Nelson, D.W., Sommers, L.E., 1982. Total carbon, organic carbon and organic matter. In: methods of soil analysis. Part 2. Agron. Monogr. 9. ASA and SSSA, Madison, WI. P. 539-577.
- Olsen, S.R., Cole, C.V., Watanabe, F.S., Dean, L.A., 1954. Estimation of available phosphorus in soils by extraction with sodium bicarbonate. United States Department of Agriculture, Circular No. 939, Washington, DC. USA.
- Paramasivam, S., Alva, A.K., Fares, A., Sajwan, K.S., 2002. Fate of nitrate and bromide in an unsaturated zone of a sandy soil under citrus production. *Journal of Environmental Quality* 31: 671-681.
- Peck, T.R., 1990. Soil testing: Past, present and future. *Communication in Soil Science and Plant Analysis* 21: 1165-1186.
- Pooler, P.S., Smith, D.R., 2005. Optimal sampling design for estimating spatial distribution and abundance of a freshwater mussel population. *Journal of the North American Benthological Society* 24: 525–537.
- Quine, T.A., Zhang, Y., 2002. An investigation of spatial variation in soil erosion, soil properties and crop production within an agricultural field in Devon, UK. *Journal of Soil and Water Conservation* 57: 55-65.
- Richards, L.A., 1954. Diagnosis and improvement of saline and alkaline soil. Unites States Department of Agriculture, Agriculture HandBook. No. 60. Washington, DC. USA.
- Rodríguez, A., Durán, J., Fernández-Palacios, J. Gallardo, A., 2009. Spatial variability of soil properties under *Pinus canariensis* canopy in two contrasting soil textures. *Plant and Soil* 322: 139–150.
- Tang, G.A., Yang, X. 2006. ArcGIS experimental tutorial of spatial analysis GIS [In Chinese]. Science Press, Beijing, China, 384-385.
- Thompson, S.K., Seber, G.A.F., 1996. Adaptive sampling. Wiley, New York.
- Wang, X.Z., Liu, G.S., Hu, H.C., Wang, Z.H., Liu, Q.H., Liu, X.F., Hao, W.H., Li, Y.T., 2009. Determination of management zones for a tobacco field based on soil fertility. *Computers and Electronics in Agriculture* 65: 168-175.
- Webster, R., Oliver, M.A. 2001. Geostatistics for Environmental Science. John Wiley and Sons, LTD, Toronto, Ontario, Canada. pp. 271.
- Wollenhaupt, N.C., Wolkowski, R.P., Clayton, M.K., 1994. Mapping soil test phosphorus and potassium for variable-rate fertilizer application. *Journal of Production Agriculture* 7: 441-448.
- Yang, T.Z., Lu, L. M., Xia, W., Fan, J.H., 2007. Characteristics of potassium-enriched, flue-cured tobacco genotype in potassium absorption, accumulation, and in-ward potassium currents of root cortex. *Agricultural Sciences in China* 6: 1479-1486.
- Zhu, Z., Stein, M.L., 2006. Spatial sampling design for prediction with estimated parameters. *Journal of Agricultural, Biological & Environmental Statistics* 11: 24-44.
- Zhu, X.H., Yang, X.C., Cai, Y.L., 2005. Fractal and fractal dimension of spatial distribution of China soil system. *Acta Pedologica Sinica* 42: 881-888.