Pedosphere **30**(5): 628–637, 2020
doi:10.1016/S1002-0160(18)60045-1
ISSN 1002-0160/CN 32-1315/P

Published by Elsevier B.V. and Science Press

PEDOSPHERE

www.elsevier.com/locate/pedosphere

Spatial interpolation of orchard soil pH using soil type and planting duration as auxiliary information

Chuancheng FU[1,2], Haibo ZHANG[1,2], Chen TU[1], Lianzhen LI[1], Xinghua LIU[1,2] and Yongming LUO[1,2,*]

[1] *CAS Key Laboratory of Coastal Environmental Processes and Ecological Remediation, Yantai Institute of Coastal Zone Research, Chinese Academy of Sciences, Yantai 264003 (China)*
[2] *University of Chinese Academy of Sciences, Beijing 100049 (China)*

(Received March 8, 2017; revised May 25, 2017)

ABSTRACT

Intensified field management in orcahrds has resulted in significant and widespread acidification in the soils. However, effectively mapping the spatial patterns of soil pH aiming to support ecological management is impeded by its large variotions across soil types and planting durations. Kriging methods were used to integrate soil type and planting duration information for effective mapping of orchard soil pH in a case study in orchards of the Northeast Jiaodong Peninsula, East China. A total of 1 472 surface soil samples were collected, and the planting duration of each sampled orchard was acquired to generate a planting duration map *via* Voronoi tessellations. The performance of five kriging methods was compared, namely, ordinary kriging (OK), OK combined with soil type (OK_ST), OK combined with planting duration (OK_PD), cokriging combined with soil type and planting duration (OCK_STPD), and OK combined with soil type and planting duration (OK_STPD). Results showed that soil pH declined significantly with increasing planting duration and exhibited moderate spatial variability over the study area. Soil type and planting duration both had significant influence on the spatial distribution of soil pH. The OCK_STPD and OK_STPD methods showed better prediction efficiency than OK, OK_ST, or OK_PD. With regard to the predicted maps of soil pH, the OCK_STPD and OK_STPD methods highly reflected local variations associated with soil type and planting duration, but the OK method was poorly representative. Categorical soil type and planting duration information may be used as ancillary information to improve the mapping quality of orchard soil pH. The OCK_STPD and OK_STPD methods were practical and efficient methods for interpolating orchard soil pH in the study area. The resultant high-quality soil pH maps can contribute to improved site-specific management in the orchards.

Key Words: Jiaodong Peninsula, cokriging, kriging, orchard soil properties, soil acidification, soil map

Citation: Fu C C, Zhang H B, Tu C, Li L Z, Liu X H, Luo Y M. 2020. Spatial interpolation of orchard soil pH using soil type and planting duration as auxiliary information. *Pedosphere*. **30**(5): 628–637.

INTRODUCTION

Soil pH is a fundamental property that has important influences on soil physical, chemical, and biological processes, and is therefore considered to be a key soil variable (Liu *et al*., 2013; Ávila *et al*., 2017). Soil acidification can reduce the availability of some nutrients (Liu *et al*., 2013), but it can increase the availability of heavy metals to toxic levels that can render soils infertile and contaminate agricultural products (Li *et al*., 2014). Under natural conditions, soil acidification is a slow process occurring over hundreds to millions of years (Guo *et al*., 2010). However, because of intensive anthropogenic activities in recent decades, there has been significant and widespread acidification in Chinese agricultural lands (Guo *et al*., 2010; Xu, 2015). Effectively delineating the spatial patterns of soil pH and revealing the acidification conditions are thus crucial to agricultural, environmental, and ecological management of agricultural ecosystems.

Soil properties vary considerably depending on soil type, climate, parent materials, topography, vegetation, water conditions, and anthropogenic activities, all of which influence the distribution patterns of soils (Shi *et al*., 2009; Moitinho *et al*., 2015). The inherent accuracy of a spatial prediction can be improved by utilizing a suitable interpolation approach and integrating helpful auxiliary information in the interpolation. Considerable attention in pedometrics has been paid to utilizing spatially correlated ancillary information to improve the mapping quality of soil properties (Goovaerts and Journel, 1995; Hengl *et al*., 2004; Wu J P *et al*., 2006, Wu C F *et al*., 2009; Qu *et al*., 2013a; Mirzaee *et al*., 2016). Soil maps delineating the borders of soil types are available and combined categorical soil type information has proved to be useful in studies of many soil properties (Liu *et al*., 2006; Goovaerts, 2010; Zhang *et al*., 2010). Although the borders of soil types are not accurate enough to segment the geographical space, the categorical soil types can separate the soil property groups with clear differences in statistical inference (Brus *et al*., 1996). Spatial variation in soil properties comprises two parts: genesis processes and historical

*Corresponding author. E-mail: ymluo@yic.ac.cn.

practices. Variation caused by genesis processes is on a geological scale and is used to determine soil classification (Liu *et al*., 2006). Variation caused by historical practices generally occurs on a field scale within each soil type (Liu *et al*., 2006). Better predictions can be obtained by considering both of these sources of variance together. However, in many circumstances soil type may be one of many factors influencing the distribution pattern of soil properties (Hengl *et al*., 2004; Zhang Z Q *et al*., 2010; Zhang S W *et al*., 2012). Moreover, there are complex spatial changes at the field scale, even in the same soil type (Liu *et al*., 2006). To minimize interpolation variance, it is necessary to address the additional sources of variance as completely as possible (Liu *et al*., 2006).

Integrating more related information into the kriging method is considered helpful in improving the accuracy of interpolation of soil properties (Hengl *et al*., 2004; Zhang Z Q *et al*., 2010; Zhang S W *et al*., 2012). With regard to orchard soils, planting duration is a vital feature affecting environmental conditions and management practices, such as soil acidification and copper enrichment (Xue *et al*., 2006; Wang *et al*., 2009; Mackie *et al*., 2012; Li *et al*., 2014; Fu *et al*., 2018). However, whether planting duration can be used as a measurement to represent historical practices related to orchard soils remains unknown. No study has shown that planting duration can be used along with soil type to further improve interpolation accuracy. Therefore, in this study we considered whether incorporating planting duration along with soil type information into a (co)kriging method could increase the accuracy of the mapping of orchard soil properties.

The Northeast Jiaodong Peninsula has a long tradition of intensive horticultural crop production. Over the last few decades, large areas of farmland have been converted into orchards, and there has been widespread reclamation of hillsides, making it one of the largest apple and grape production areas in China. Soil acidification has become pronounced because of the intensification of field management in the orchards. A preliminary investigation in 2007–2009 reported that topsoil pH was less than 5.5 in 60.4% of 268 sites investigated and was less than 4.5 in 27.2% (Li *et al*., 2014). It is therefore urgent to investigate acidification in every orchard and to effectively map the spatial distribution of soil pH levels. The main aims of this study were to i) study the soil acidification status and its spatial variation on the Northeast Jiaodong Peninsula and ii) compare the performance and feasibility of (co)kriging combined with soil type and planting duration information in the mapping of orchard soil pH with ordinary kriging (OK).

MATERIALS AND METHODS

Study area

The study region (36°49′8″–37°48′32″ N, 120°30′40″–121°28′5″ E) covers an area of 3 620 km^2 on the Northeast Jiaodong Peninsula (Fig. 1). It comprises 21 municipal towns of Yantai City, a coastal city located on the rim of the Bohai and Yellow seas. The region is classified as a warm temperate zone with a marine climate, humid air, and ample sunlight with four distinct seasons (Li *et al*., 2014). The mean annual rainfall is 980 mm and the annual temperature averages 11.2–12.5 °C. Elevation in the orchards ranges from 0 to 766 m above sea level.

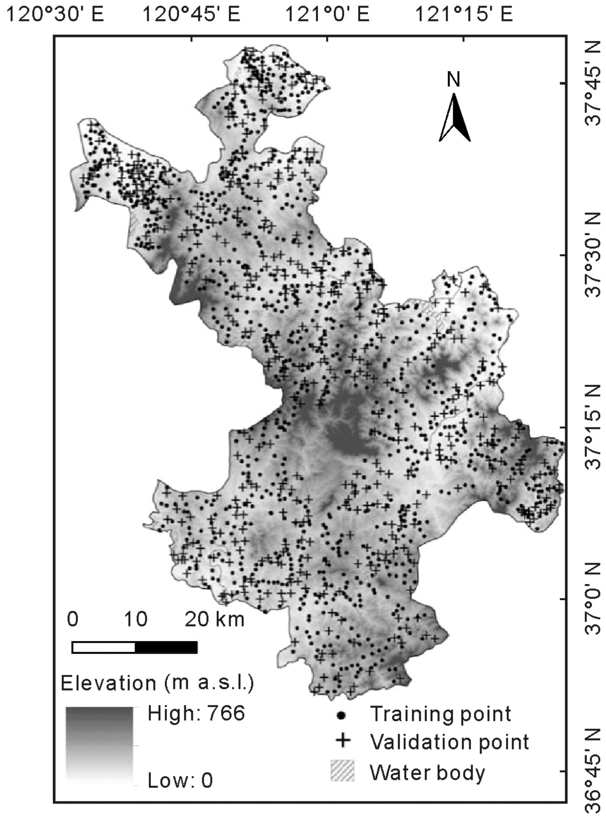

Fig. 1 Distribution of soil samples in the study area on the Northeast Jiaodong Peninsula. a.s.l. = above sea level.

Soil sampling and chemical analysis

The soil sampling sites were selected according to a soil map and based on the distribution of orchards in the study area. Soil samples were taken during April and June 2014. Five soil subsamples were collected at each sampling site within a 10-m radius using a shovel at a depth of 0–20 cm, and subsequently mixed thoroughly to obtain an accurate representative composite sample from the site. A global positioning system (GPSmap 60CSx, Garmin, Olathe, USA) was used to determine the locations of the sampling centers. A total of 1 472 samples were collected, and their locations are presented in Fig. 1. In addition, the age of each orchard was ascertained from the orchard manager or owner and classified into five groups, namely, < 5, 5–15, 15–25, 25–35, and > 35 years.

The soil samples were air-dried at room temperature.

After stones or other coarse debris were manually removed, the dried samples were then ground and sieved to ensure that the soil particles were < 2 mm in diameter. Soil pH was determined at a soil-to-water mass ratio of 1:2.5 using a pH meter (FiveEasy Plus FE20, Mettler Toledo, Greifensee, Switzerland). Detailed descriptions of the routine analytical methods used have been previously published (Lu, 1999).

Spatial interpolation

Ordinary kriging and kriging combined with auxiliary information methods were used for predicting the soil pH distribution of the study area. The semivariance and the interpolation principle of the OK method have been widely established by numerous studies (Lark and Webster, 2006; Wu *et al*., 2006; Qu *et al*., 2013a; Fu *et al*., 2018). Kriging combined with auxiliary information methods have been widely adopted in the geosciences and soil science (Goovaerts and Journel, 1995; Liu *et al*., 2006; Zhang *et al*., 2010, 2011; Qu *et al*., 2013a). Kriging combined with auxiliary information methods followed simple kriging with local means as presented by Goovaerts (1997). However, ordinary (co)kriging was conducted to interpolate the residual instead of simple kriging. These methods are by nature regression kriging (Goovaerts and Kerry, 2010).

Ordinary kriging combined with soil type or planting duration (OK_ST, OK_PD). Soil type is known to be a principal factor governing the spatial distribution of soil properties (Liu *et al*., 2006; Wu *et al*., 2008). Apart from the effect of soil type, soil acidification conditions generally exhibit a strong correlation with planting duration for orchard soils (Xue *et al*., 2006; Li *et al*., 2014). Soil samples within the same planting duration may not have similar mean pH values due to differences in soil type, but they may have a similar degree of acidification (Xue *et al*., 2006; Li *et al*., 2014). That is, mixing soil pH from different planting duration groups may increase its variability, as well as the interpolation uncertainty. Thus, the spatial variability of any particular soil attribute of orchard soils is partly due to the complex distribution of soil types and planting durations, not accounting for which can increase the uncertainty of interpolation prediction. To reduce this uncertainty, we suggested kriging combined with soil type (OK_ST) or planting duration (OK_PD) to incorporate their effects in the interpolation of soil properties. The average soil pH for each soil type or planting duration was calculated, then the pH value for each sample ($Z(x)$) was separated into two portions, the mean value corresponding to the soil type $\mu(T)$ or planting duration $\mu(D)$ and the associated residual $r_{\mathrm{ST}}(x)$ or $r_{\mathrm{PD}}(x)$:

$$Z(x) = \mu(T) + r_{\mathrm{ST}}(x) \tag{1}$$

$$Z(x) = \mu(D) + r_{\mathrm{PD}}(x) \tag{2}$$

where x is the location of the sample $Z(x)$ and T (or D) is the soil type (or planting duration) to which x belongs. The variance of the original $Z(x)$ is also divided into two portions, variance between different soil types (planting duration) and variance within a soil type (planting duration) (Liu *et al*., 2006; Zhang *et al*., 2011). The residuals $r_{\mathrm{ST}}(x)$ or $r_{\mathrm{PD}}(x)$ can be treated as a new stationary regionalized variable to be interpolated *via* OK (Liu *et al*., 2006; Qu *et al*., 2013a). The final result was obtained by adding together the mean pH value for the soil type or planting duration and the interpolated results of residuals.

Cokriging combined with soil type and planting duration (OCK_STPD). Cokriging is an extension of the kriging method that predicts the random variables simultaneously by utilizing the relationships and coregionalization of the variables (Odeh *et al*., 1995; Goovaerts, 1998). If coregionalization exists it is feasible to use the co-variable to improve the prediction of the target variable through cokriging (Wang *et al*., 2013). To describe the auto- and cross-semivariograms for use in cokriging, a linear model of coregionalization has to be fitted to these semivariograms, which consider the spatial dependence of the two variables and their interdependence simultaneously (Wang *et al*., 2013). Yates and Warrick (1987) found that cokriging is more effective in interpolation than kriging when sample correlations exceed 0.5. If strong relationships and spatial co-variability can be observed between $r_{\mathrm{ST}}(x)$ and $r_{\mathrm{PD}}(x)$, we can use $r_{\mathrm{PD}}(x)$ as auxiliary data to improve the prediction of $r_{\mathrm{ST}}(x)$ by ordinary cokriging. The final predicted soil pH map is the sum of the mean pH value for the soil type and the cokriging interpolated result of $r_{\mathrm{ST}}(x)$.

Ordinary kriging combined with soil type and planting duration (OK_STPD). Similar to OK_ST, the soil pH value of every unsampled point was estimated by adding the mean value of the categorical soil type and planting duration to the interpolated residual ($r_{\mathrm{STPD}}(x)$). The categorical soil type and planting duration can be obtained by overlapping the soil map and planting duration map. The mean values of the corresponding soil type and planting duration represent the typical acidification at each planting duration within each soil type, whereas the residual $r_{\mathrm{STPD}}(x)$ represents the variation caused by the limitations of soil classification and the differences in management intensities in different fields. The form of the equation follows that of OK_ST:

$$Z(x) = \mu(T, D) + r_{\mathrm{STPD}}(x) \tag{3}$$

The soil map utilized in the present study was obtained from the Yantai Soil and Fertilizer Station and digitized with ArcGIS 10.2 (ESRI Inc., Redlands, USA). The soil map (Fig. 2a) was based on the Genetic Soil Classification of China (GSCC) mapped in the Second National Soil Survey. It was difficult to obtain an accurate orchard planting duration map because of the absence of explicit boundaries. We therefore generated an approximation based on the sample points using Voronoi tessellation (Fig. 2b). Voronoi tessellation has been widely used to model the partition of space over widely disparate scales (Lark, 2009). The map shows abrupt changes at the borders between different planting duration groups, which can be utilized to group the sampled observations. However, the Voronoi map cannot reliably delineate the spatial distribution of planting durations because planting duration is influenced by more than distance. Using Voronoi tessellation to represent the planting duration is an assumption that was built into our method.

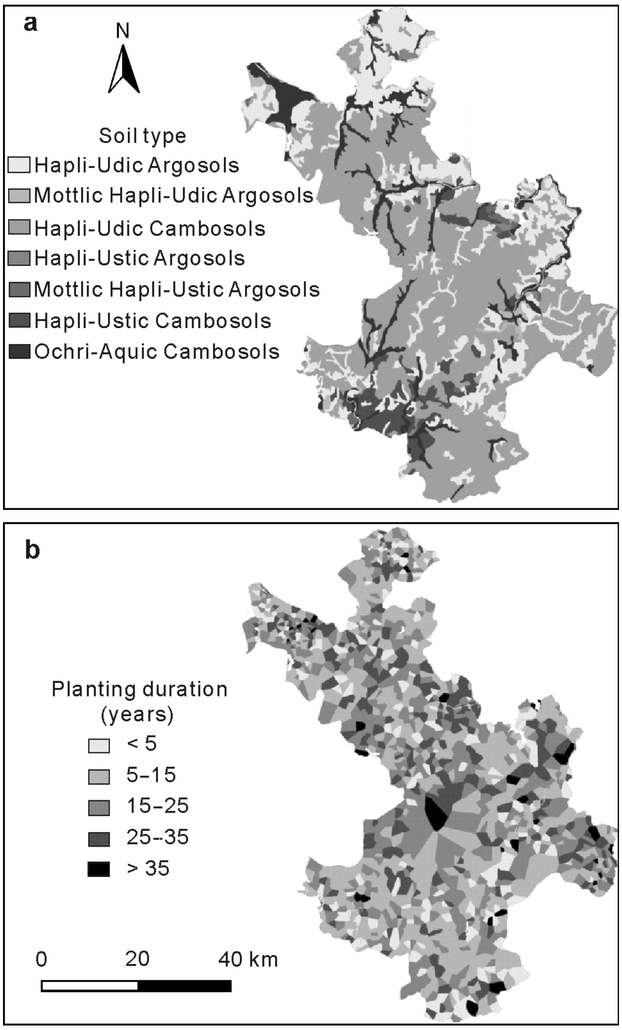

Fig. 2 Soil map (a) and planting duration map (b) of the study area on the Northeast Jiaodong Peninsula.

Validation

To evaluate the performance of the kriging methods the soil pH data for the 1 472 sites were randomly partitioned into two subsets using the Subset Features tool in the ArcGIS Geostatistical Analyst toolbox (ArcGIS 10.2), a training dataset (70%, $n = 1\ 030$) and a validation dataset (30%, $n = 442$). For each of the five methods, the correlation coefficients (r), mean errors (ME), and root mean square errors (RMSE) were calculated between the predicted maps and the validation data. Mean error and RMSE were calculated using the following equations:

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n}(P_i - O_i) \tag{4}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2} \tag{5}$$

where P_i is the predicted pH value of the ith location x_i, O_i is the observed pH value of the ith location x_i, and n is the number of validation points. The ME measures the interpolation bias and should ideally be close to 0, and RMSE measures the accuracy of the interpolation and should also be as small as possible.

Statistic analysis

All the results were stored in a Microsoft Excel (Microsoft Corp., Redmond, USA) spreadsheet. Statistical analysis was conducted using SPSS 20.0 (SPSS Inc, Chicago, USA). Variances in the soil pH between soil types and planting durations were determined using a linear mixed effect model (LMEM). The semivariograms were calculated and modeled *via* GS+ 9.0 (Gamma Design Software, Plainwell, USA), and kriging interpolations were conducted *via* ArcGIS 10.2. Scatter plot graphs were plotted using Origin Pro 8.0 (OriginLab Corp., Northampton, USA).

RESULTS AND DISCUSSION

Descriptive statistics and LMEM analysis

Summaries of the descriptive statistics for orchard soil pH are shown in Table I. Soil pH values ranged between 3.87 and 8.82 across the 1 472 samples, with a mean of 5.96 and a standard deviation of 0.91. Soil pH is often considered to be one of the less variable soil chemical properties (Sun *et al*., 2003; Liu Z P *et al*., 2013; Liu Y *et al*., 2016). In the present study, the coefficient of variation (CV) of all samples was 0.15, indicating medium variation among different soil types and planting durations. Judging from the skewness and kurtosis values, the soil pH data fitted an approximate normal distribution. Samples were classified into seven groups based on the soil types of their locations, of which the average pH ranged from 7.46 (Mottlic Hapli-Ustic Argosols) to 5.74 (Hapli-Udic Cambosols). The LMEM analysis results (Table II) showed that the variances in soil pH among the soil type categories were highly significant ($P < 0.001$). This indicates that soil type information will be a valuable input in kriging interpolation of soil pH to improve the accuracy of its estimation.

The average pH values for five groups of planting duration were in a clear descending order with increasing year, from 6.45 (< 5 years) to 5.18 (> 35 years). This indicates

TABLE I

Descriptive statistics of orchard soil pH across the study area on the Northeast Jiaodong Peninsula

Item	n	Mean	SD[a)]	CV[b)]	Minimum	Maximum	Skewness	Kurtosis
Soil type[c)]								
Hapli-Udic Argosols[d)]	490	6.00	0.91	0.15	3.93	8.40	0.28	−0.62
Mottlic Hapli-Udic Argosols[d)]	54	6.08	0.88	0.14	4.31	7.69	0.12	−0.76
Hapli-Udic Cambosols[d)]	588	5.74	0.89	0.16	3.87	8.21	0.48	−0.43
Hapli-Ustic Argosols[e)]	18	7.09	0.54	0.08	6.06	8.07	0.15	−0.22
Mottlic Hapli-Ustic Argosols[e)]	6	7.46	0.73	0.10	6.25	8.28	−0.88	0.46
Hapli-Ustic Cambosols[e)]	84	6.74	1.00	0.15	4.39	8.49	−0.27	−0.73
Ochri-Aquic Cambosols[f)]	232	6.02	0.95	0.16	4.03	8.82	0.25	−0.51
Planting duration (years)								
< 5	185	6.45	0.96	0.15	4.71	8.82	0.29	−0.67
5–15	567	6.00	0.93	0.16	4.34	8.36	0.36	−0.82
15–25	459	5.88	0.95	0.16	3.87	8.28	0.32	−0.64
25–35	230	5.78	0.84	0.15	4.04	7.68	0.21	−0.78
> 35	31	5.18	0.88	0.17	4.03	7.29	0.56	−0.28
Total	1 472	5.96	0.91	0.15	3.87	8.82	0.33	−0.62
Training	1 030	5.97	0.96	0.16	3.87	8.82	0.28	−0.66
Validation	442	5.94	0.93	0.16	4.19	8.49	0.45	−0.49

[a)] Standard deviation.
[b)] Coefficient of variation.
[c)] Soil type is based on the Genetic Soil Classification of China (GSCC) but referenced to the Chinese Soil Taxonomy (CST) (Gong *et al*., 2002; Shi *et al*., 2010).
[d)] Brown soil group.
[e)] Cinnamon soil group.
[f)] Fluvo-aquic soil group.

TABLE II

Linear mixed effect model for the effects of soil type and plant duration on the soil pH values

Effect	df[a)]	Sum of squares	Mean square	F value	Significance
Corrected model	29	131.119	4.521	5.531	0.000
Soil type	6	31.639	5.273	6.451	0.000
Planting duration	4	16.532	4.133	5.056	0.000
Soil type × planting duration	19	19.688	1.368	1.744	0.020

[a)] Degree of freedom.

that acidification in orchard soils was significantly related to planting duration and old orchards usually had lower soil pH than did the younger ones (Xue *et al*., 2006; Li *et al*., 2014). This is in agreement with Li *et al*. (2014) who reported that the soils on the Northeast Jiaodong Peninsula were significantly acidified in the adult (10–30 years) and old ($>$ 30 years) orchards (pH decreased by 0.53 and 1.90 units in adult and old orchard soils, respectively, compared to young orchard soils ($<$ 10 years)). Soil acidification can be caused by various factors, such as excessive application of chemical fertilizers, improper irrigation, acid deposition, and natural acidification (Zhang *et al*., 2013; Li *et al*., 2014; Xu, 2015). Chinese agriculture has intensified greatly since the early 1980s and the limited arable land has been expected to produce higher outputs through the application of nitrogen fertilizers and other resources (Guo *et al*., 2010). Wei and Jiang (2012) noted that approximately 612 kg fertilizer N ha^{-1} year^{-1} was used in orchards of the Jiaodong Peninsula, and the amount was still increasing. Goulding and Annis (1998) showed that 4 kmol H^+ can be generated from each 50 kg ha^{-1} year^{-1} of added ammonium-N, which requires approximately 500 kg $CaCO_3$ to neutralize it in field conditions. Moreover, large amounts of base cations, such as calcium and magnesium in the soils are lost through leaching and uptake by fruit trees (Rengel *et al*., 2000; Xu, 2015). Over a long planting duration, soil acidification might become more severe because of the constant application of ammonium-N fertilizers and a decline in exchangeable bases (Li *et al*., 2014). Planting duration significantly impacted soil pH, as can be seen in the LMEM analysis results (Table II). The variances in soil pH among different planting duration categories were highly significant ($P < 0.001$), indicating that planting duration is an important factor influencing the spatial distribution pattern of soil pH in the study area. Interactions between soil types and planting durations also exhibit a strong influence on soil pH ($P < 0.05$) (Table II). This indicates that planting duration is an effective measure of historical practices and has the potential to further improve the accuracy of interpolation of soil pH along with soil type.

Geostatistical analysis

Semivariograms can provide clear descriptions of the spatial structure of soil properties and insights into possible processes influencing their spatial distributions (Paz-Gonzalez *et al*., 2001; Qu *et al*., 2013b). As presented in Table III and Fig. 3, the five auto- and cross-semivariograms were all well-fitted by exponential models. It has been reported that if the R^2 of the semivariogram model is below 0.5, the spatial dependence of the variable is weak (Duffera *et al*., 2007). In the present study, the five fitted semivariograms all had moderate R^2 of 0.763, 0.702, 0.697, 0.693, and 0.659, showing that the spatial dependencies of the variables were moderate. After the mean values of each categorical group were eliminated, the parameters nugget (C_0)/sill ($C_0/(C_0+C)$), where C is the partial sill, and sill (C_0+C) as the ranges for the residuals, all decreased compared with those of the original soil pH data. The large differences between the semivariograms indicate that the mean soil pH values (local trend) in each group had substantial effects on the semivariances (Liu *et al*., 2006; Zhang *et al*., 2011). Not accounting for the local trend would increase the uncertainty of spatial prediction (Liu *et al*., 2006; Zhang *et al*., 2011). Thus, it was necessary to remove the drift to improve interpolation accuracy (Zhang *et al*., 2010; Shi *et al*., 2011). Among the model parameters for the four kriging combined with auxiliary information methods (Fig. 4), in general, the cross-(auto-)semivariograms of $r_{\mathrm{ST}}(x)\times r_{\mathrm{PD}}(x)$ and $r_{\mathrm{STPD}}(x)$ had smaller C_0 and C_0+C than $r_{\mathrm{ST}}(x)$ and $r_{\mathrm{PD}}(x)$. Thus, the local trend was further reduced when combining together soil type and planting duration information. The semivariograms between $r_{\mathrm{ST}}(x)$ and $r_{\mathrm{PD}}(x)$ show strong spatial co-variability, which echoes the strong relationship ($r=0.938$**, $P<0.01$) shown in their scatterplot in Fig. 4. A linear model of coregionalization was fitted to the auto- and cross-semivariograms ($r_{\mathrm{ST}}(x), r_{\mathrm{PD}}(x)$ and $r_{\mathrm{ST}}(x)\times r_{\mathrm{PD}}(x)$), which ensured the same spatial structure

TABLE III

Geostatistical parameters of models fitted to auto- and cross-semivariograms

Parameter[a)]	Model	Nugget (C_0)	Sill (C_0+C[b)])	C_0/C_0+C	Range	RSS[c)]	R^{2}[d)]
					m		
pH	Exponential	0.260	0.936	0.278	6 900	1.98×10^{-2}	0.763
$r_{\mathrm{ST}}(x)$	Exponential	0.180	0.835	0.216	4 800	1.18×10^{-2}	0.702
$r_{\mathrm{PD}}(x)$	Exponential	0.210	0.890	0.236	4 800	2.23×10^{-2}	0.697
$r_{\mathrm{ST}}(x)\times r_{\mathrm{PD}}(x)$	Exponential	0.120	0.816	0.147	4 800	1.13×10^{-2}	0.693
$r_{\mathrm{STPD}}(x)$	Exponential	0.146	0.798	0.183	3 900	6.51×10^{-3}	0.659

a) $r_{\mathrm{ST}}(x)$ = residual in the ordinary kriging combined with soil type method; $r_{\mathrm{PD}}(x)$ = residual in the ordinary kriging combined with planting duration method; $r_{\mathrm{STPD}}(x)$ = residual in the ordinary kriging combined with soil type and planting duration method.
b) Partial sill.
c) Residual sum of squares.
d) Coefficient of determination.

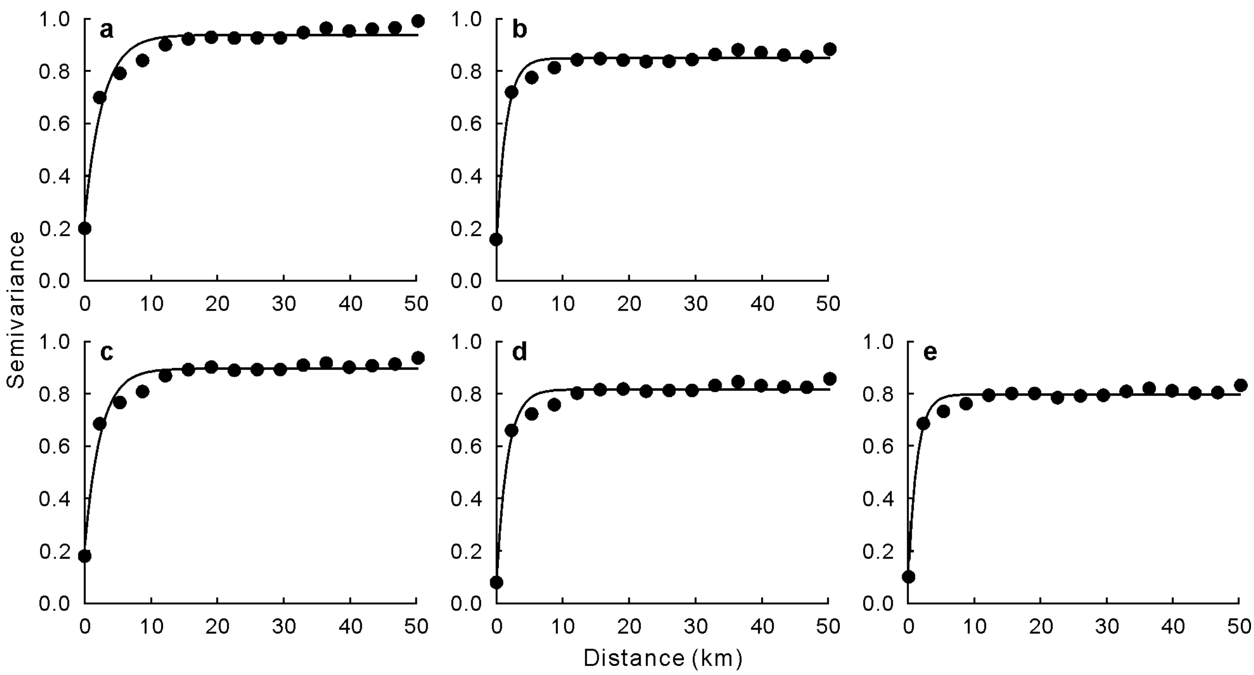

Fig. 3 Auto- and cross-semivariograms of soil pH and the associated residuals pH (a), residual in the ordinary kriging combined with soil type method ($r_{\mathrm{ST}}(x)$) (b), residual in the ordinary kriging combined with planting duration method ($r_{\mathrm{PD}}(x)$) (c), $r_{\mathrm{ST}}(x)\times r_{\mathrm{PD}}(x)$ (d), residual in the ordinary kriging combined with soil type and planting duration method ($r_{\mathrm{STPD}}(x)$) (e).

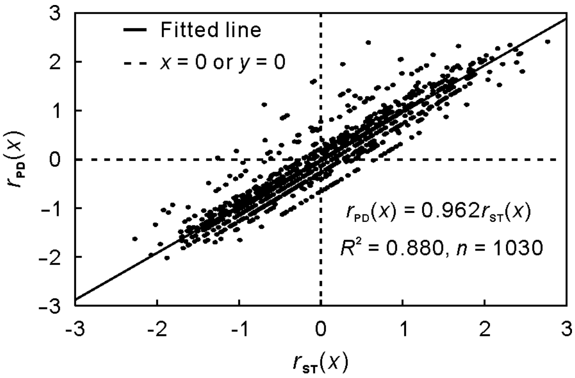

Fig. 4 Relationship between residual in the ordinary kriging combined with soil type method $r_{\mathrm{ST}}(x)$ and residual in the ordinary kriging combined with planting duration method $r_{\mathrm{PD}}(x)$. R^2 = coefficient of determination.

with equal ranges (only nuggets and sills were different) were utilized in the semivariograms (Bilgili *et al*. 2011; Wang *et al*. 2013). Therefore, $r_{\mathrm{PD}}(x)$ may be chosen as the auxiliary variable during the interpolation of $r_{\mathrm{ST}}(x)$.

Generally, strong spatial dependence of soil properties can be attributed to intrinsic factors (soil formation factors, such as soil parent materials and soil type), whereas weak spatial dependence can be attributed to extrinsic factors (soil management practices, such as fertilization and planting duration) (Cambardella *et al*., 1994; Wu *et al*., 2009). In the current study, the $C_0/(C_0 + C)$ of the soil pH was 0.278, which was between 0.25 and 0.75, and thus characteristic of moderate spatial dependence. This indicates that the spatial variability of soil pH might be affected by both intrinsic and extrinsic factors. The $C_0/(C_0 + C)$ for the residuals after removal of local means were all lower than 0.25, demonstrating strong spatial dependence of the residuals.

Comparison of the accuracy of prediction

The correlation coefficients r, ME, and RMSE of the five kriging methods are shown in Table IV. For comparison, statistics for the determined pH values at these same sites are included. Most of the MEs of soil pH (except for OK_PD) interpolated with auxiliary information were much closer to 0 than OK, which indicates that soil type and planting duration information may help the interpolator to be a more unbiased model. The RMSEs were generally in a descending order for all methods: OK > OK_ST > OK_PD > OK_STPD > OCK_STPD. The strengths of the correlations between the observed and estimated pH values obtained using different interpolation approaches were all significant ($P < 0.01$). The lowest r was from the OK method ($r = 0.457$, $P < 0.01$) and the largest was from the OK_STPD method ($r = 0.506$, $P < 0.01$). Compressions in the overall range of predicted soil pH values were ameliorated by integrating useful auxiliary information, although no method successfully predicted the largest and smallest soil pH values. Relative to the OK_ST and OK_PD methods, OCK_STPD and OK_STPD provided improvements in accuracy of prediction, implying that using planting duration along with soil type could improve the quality of pH predictions. As reported by Goovaerts (1999) and Miháliková *et al*. (2015), the contribution of the auxiliary information to the cokriging estimation depends not only on the correlation between primary and auxiliary variables, but also on their patterns of spatial continuity. The OCK_STPD method demonstrated a clear benefit from cokriging of the two residuals in the present study, thus the introduction of planting duration residuals helped improve the accuracy of prediction. The OK_STPD method generated a better result than OK_ST or OK_PD, consistent with previous studies in which soil organic content was interpreted using information on soil type and land use (Zhang *et al*., 2010; 2011). The OCK_STPD and OK_STPD methods exhibited similar efficiencies in mapping soil pH of the study area but the OK_STPD method was more effective in predicting the peak and low values. For the reasons mentioned above, planting duration was one measurement that represented historical practices related to orchard soils, and thus could be used along with soil type to further improve the accuracy of interpolation.

Differences in prediction errors between OK and kriging combined with auxiliary information might have been derived from a variety of sources (Zhang *et al*., 2015). Smoothing effects may be the major source of prediction

TABLE IV

Summary statistics for determined and estimated soil pH at 442 sites of the testing set as predicted by five kriging methods[a)]

Parameter[b)]	Determined	OK	OK_ST	OK_PD	OCK _STPD	OK_STPD
ME		−0.014	−0.012	0.019	0.008	0.010
RMSE		0.830	0.820	0.816	0.806	0.811
r		0.457**	0.473**	0.490**	0.494**	0.506**
Minimum	4.19	4.76	4.58	4.54	4.86	4.15
Maximum	8.49	7.80	7.72	7.79	7.91	8.06
Mean	5.94	5.94	5.93	5.96	5.94	5.95
Median	5.83	5.91	5.90	5.91	5.89	5.93
SD	0.93	0.52	0.54	0.57	0.52	0.57

**Significant at $P < 0.01$.

[a)] OK = ordinary kriging; OK_ST = OK combined with soil type; OK_PD = OK combined with planting duration; OCK_STPD = cokriging combined with soil type and planting duration; OK_STPD = OK combined with soil type and planting duration.

[b)] ME = mean error; RMSE = root mean square error; r = correlation coefficient; SD = standard deviation.

uncertainty because the least square principle is utilized to minimize the local error (Kerry and Oliver, 2007; Chai *et al*., 2008). Smoothing effects are a widely known characteristic of geostatistical interpolation techniques that lead to underestimation of high values and overestimation of low values (Lark and Webster, 2006; Xie *et al*., 2011; Fu *et al*., 2018). In the present study, without accounting for the difference among soil types or planting durations, the interpolated soil pH data *via* OK resulted in highest variation. This was primarily because the soil pH data between different soil types and (or) planting durations were significantly different. If the original soil pH data are interpolated with OK directly, a strong smoothing effect will be produced to reduce the prediction accuracy. However, the impact of the smoothing effect was gradually reduced when applying the OK_ST or OK_PD method, which accounted for the variation among soil types or planting durations. The effects of smoothing on the OCK_STPD and OK_STPD methods were still smaller, because account was taken of the variability caused by both soil type and planting duration, and the estimated soil pH values were more accurate over a larger range (Zhang *et al*., 2010).

Spatial distribution maps of orchard soil pH

The soil pH spatial distribution maps predicted by the OK, OK_ST, OK_PD, OCK_STPD, and OK_STPD methods are shown in Fig. 5. The maps are displayed with the same soil pH range classification scale to facilitate comparisons among the various methods. All maps showed the same distribution trends of soil pH values over the study area. Apart from this, the maps estimated and generated by diverse methods differed significantly. The soil pH range 5.00–6.00 was the largest in area, as expected from the sample statistics. The OK method generated a very smooth map. The predicted soil pH values were very flat and their differences among various soil types and (or) planting durations cannot be distinguished easily. However, the OK_ST, OK_PD, OCK_STPD, and OK_STPD methods generated maps with detailed structures, which had higher and lower values in local areas than OK. The polygons were more fragmented and with abrupt changes that reflected the variations in soil type and planting duration. Thus, kriging combined with auxiliary information methods exhibited better interpolation quality than the OK method. Among the four kriging combined with auxiliary

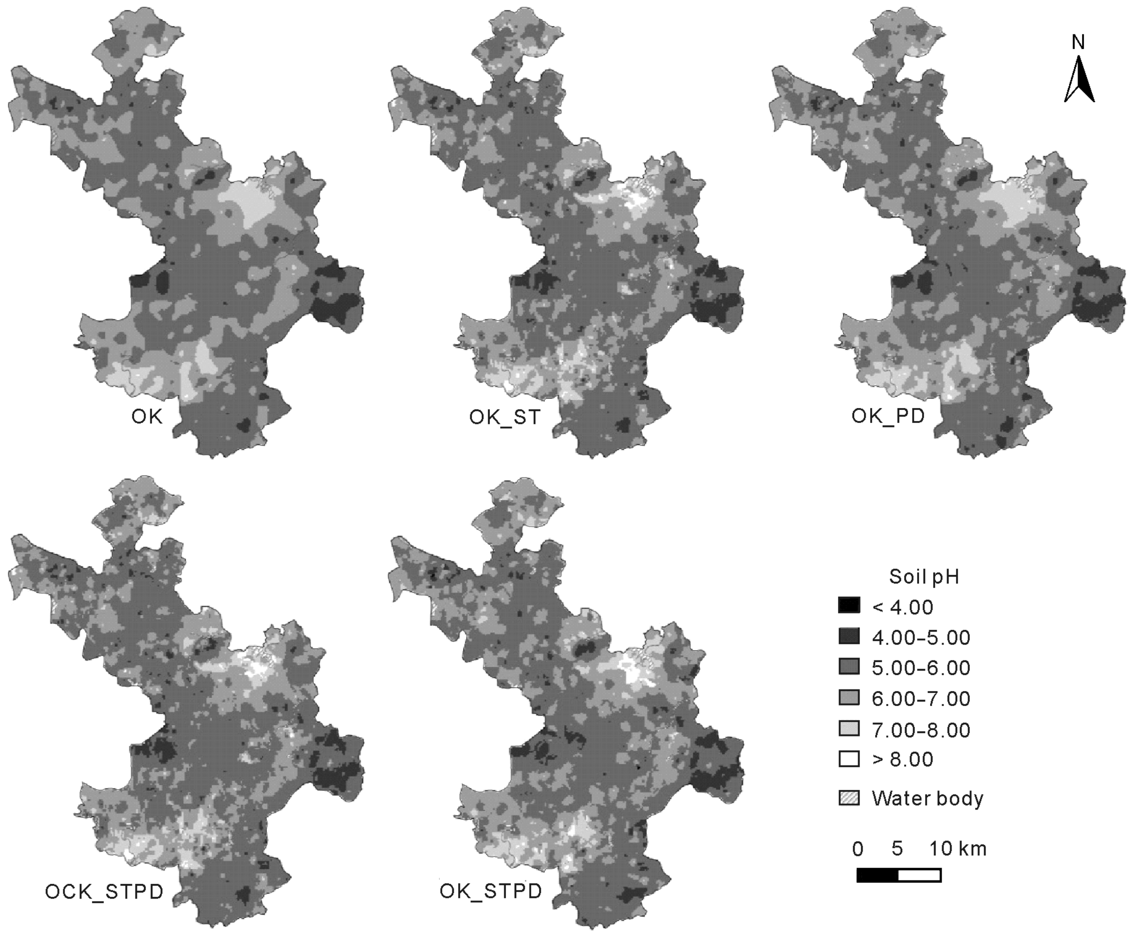

Fig. 5 Interpolated soil pH maps of the Northeast Jiaodong Peninsula generated using five kriging methods. OK = ordinary kriging; OK_ST = OK combined with soil type; OK_PD = OK combined with planting duration; OCK_STPD = cokriging combined with soil type and planting duration; OK_STPD = OK combined with soil type and planting duration.

information methods, to a lesser degree, the OCK_STPD and OK_STPD methods produced more detailed maps than OK_ST or OK_PD, indicating the advantage of integrating planting duration along with soil type information. However, extremely high and low soil pH values were rarely observed in the four interpolated maps, indicating that the smoothing effect was evident even in the kriging combined with auxiliary information methods. Only the OK_STPD method seemed effective in reducing the smoothing effect compared with other methods. In conclusion, the OCK_STPD and OK_STPD methods better modeled the local pH spatial variability and are recommended for interpolating soil pH, or even other soil properties, in the orchards. However, the delineation of planting duration based on Voronoi tessellations might be inaccurate because the shape and area of the Voronoi polygons did not match the outlines of the true orchards. More observations of planting duration are required to further improve the interpolation accuracy of soil pH.

CONCLUSIONS

On the Northeast Jiaodong Peninsula, orchards have recently been exhibiting soil acidification, and a clear correlation has been observed between soil acidification and planting duration. Soil pH exhibited wide spatial variability in the region owing to complicated variations in soil type and different planting durations. The OCK_STPD and OK_STPD methods were more promising spatial interpolation methods for improving the accuracy of prediction compared to OK, OK_ST, and OK_PD, clearly accounting for differences among soil types and planting durations. The results of this study suggest that planting duration may be used as a measurement that represents historical practices related to orchard soils. Categorical soil type and planting duration information may be used as ancillary information to improve the mapping quality of orchard soil pH. The OCK_STPD and OK_STPD methods were practical and efficient methods for spatial prediction of orchard soil pH on the Northeast Jiaodong Peninsula. These methods may be used to map the spatial variability of soil pH in other orchards in China and elsewhere in the world.

ACKNOWLEDGEMENTS

This research was funded by the National Key Research and Development Plan of China (No. 2016YFE0106400) and the National High Technology Research and Development Program (863 Program) of China (No. 2012AA06A204-4). The authors are grateful to Prof. Peter CHRISTIE for revision of the English and to the reviewers and subject editor of the journal for constructive comments on the manuscript.

REFERENCES

Ávila P F, da Silva E F, Candeias C. 2017. Health risk assessment through consumption of vegetables rich in heavy metals: The case study of the surrounding villages from Panasqueira mine, Central Portugal. *Environ Geochem Health*. **39**: 565–589.

Bilgili A V, Akbas F, van Es H M. 2011. Combined use of hyperspectral VNIR reflectance spectroscopy and kriging to predict soil variables spatially. *Precis Agric*. **12**: 395–420.

Brus D J, de Gruijter J J, Marsman B A, Visschers R, Bregt A K, Breeuwsma A, Bouma J. 1996. The performance of spatial interpolation methods and choropleth maps to estimate properties at points: A soil survey case study. *Environmetrics*. **7**: 1–16.

Cambardella C A, Moorman T B, Novak J M, Parkin T B, Karlen D L, Turco R F, Konopka A E. 1994. Field-scale variability of soil properties in central Iowa soils. *Soil Sci Soc Am J*. **58**: 1501–1511.

Chai X R, Shen C Y, Yuan X Y, Huang Y F. 2008. Spatial prediction of soil organic matter in the presence of different external trends with REML-EBLUP. *Geoderma*. **148**: 159–166.

Duffera M, White J G, Weisz R. 2007. Spatial variability of Southeastern U.S. Coastal Plain soil physical properties: Implications for site-specific management. *Geoderma*. **137**: 327–339.

Fu C C, Zhang H B, Tu C, Li L Z, Luo Y M. 2018. Geostatistical interpolation of available copper in orchard soil as influenced by planting duration. *Environ Sci Pollut Res*. **25**: 52–63.

Gong Z T, Zhang G L, Chen Z C, Luo G B, Zhao W J. 2002. Soil reference on the bases of Chinese Soil Taxonomy. *Chin J of Soil Sci* (in Chinese). **33**: 1–5.

Goovaerts P, Journel A G. 1995. Integrating soil map information in modelling the spatial variation of continuous soil properties. *Eur J Soil Sci*. **46**: 397–414.

Goovaerts P, Kerry R. 2010. Using ancillary data to improve prediction of soil and crop attributes in precision agriculture. *In* Oliver M A (ed.) Geostatistical Applications for Precision Agriculture. Springer Science, Netherlands. pp. 167–194.

Goovaerts P. 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York.

Goovaerts P. 1998. Ordinary cokriging revisited. *Math Geol*. **30**: 21–42.

Goovaerts P. 1999. Geostatistics in soil science: State-of-the-art and perspectives. *Geoderma*. **89**: 1–45.

Goovaerts P. 2010. Combining areal and point data in geostatistical interpolation: Applications to soil science and medical geography. *Math Geosci*. **42**: 535–554.

Goulding K W T, Annis B. 1998. Lime, liming and the management of soil acidity. *In* International Fertiliser Society (ed.) Proceedings-Fertiliser Society (United Kingdom) No. 410. p. 36.

Guo J H, Liu X J, Zhang Y, Shen J L, Han W X, Zhang W F, Christie P, Goulding K W T, Vitousek P M, Zhang F S. 2010. Significant acidification in major Chinese croplands. *Science*. **327**: 1008–1010.

Hengl T, Heuvelink G B M, Stein A. 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma*. **120**: 75–93.

Kerry R, Oliver M A. 2007. Comparing sampling needs for variograms of soil properties computed by the method of moments and residual maximum likelihood. *Geoderma*. **140**: 383–396.

Lark R M, Webster R. 2006. Geostatistical mapping of geomorphic variables in the presence of trend. *Earth Surf Processes Land*. **31**: 862–874.

Lark R M. 2009. A stochastic-geometric model of soil variation. *Eur J Soil Sci*. **60**: 706–719.

Li L Z, Wu H F, van Gestel C A M, Peijnenburg W J G M, Allen H E. 2014. Soil acidification increases metal extractability and bioavailability in old orchard soils of Northeast Jiaodong Peninsula in China. *Environ Pollut*. **188**: 144–152.

Liu T L, Juang K W, Lee D Y. 2006. Interpolating soil properties using kriging combined with categorical information of soil maps. *Soil Sci Soc Am J*. **70**: 1200–1209.

Liu Y, Wang H F, Zhang H, Liber K. 2016. A comprehensive support vector machine-based classification model for soil quality assessment. *Soil Till Res*. **155**: 19–26.

Liu Z P, Shao M A, Wang Y Q. 2013. Large-scale spatial interpolation of soil pH across the Loess Plateau, China. *Environ Earth Sci*. **69**: 2731–2741.

Lu R K. 1999. Soil and Agricultural Chemical Analysis Methods (in Chinese). China Agricultural Science and Technology Press, Beijing.

Mackie K A, Müller T, Kandeler E. 2012. Remediation of copper in vineyards—a mini review. *Environ Pollut*. **167**: 16–26.

Miháliková M, Başkan O, Dengiz O. 2015. Capability of different interpolation models and pedotransfer functions to estimate soil hydraulic properties in Büyükçay Watershed. *Environ Earth Sci*. **74**: 2425–2437.

Mirzaee S, Ghorbani–Dashtaki S, Mohammadi J, Asadi H, Asadzadeh F. 2016. Spatial variability of soil organic matter using remote sensing data. *Catena*. **145**: 118–127.

Moitinho M R, Padovan M P, Panosso A R, de Bortoli Teixeira D, Ferraudo A S, La Scala Jr N. 2015. On the spatial and temporal dependence of CO_2 emission on soil properties in sugarcane (*Saccharum* spp.) production. *Soil Till Res*. **148**: 127–132.

Odeh I O A, McBratney A B, Chittleborough D J. 1995. Further results on prediction of soil properties from terrain attributes: Heterotopic cokriging and regression-kriging. *Geoderma*. **67**: 215–226.

Paz-Gonzalez A, Taboada Castro M T, Vieira S R. 2001. Geostatistical analysis of heavy metals in a one-hectare plot under natural vegetation in a serpentine area. *Can J Soil Sci*. **81**: 469–479.

Qu M K, Li W D, Zhang C R, Zhao Y C, Huang B, Sun W X, Hu W Y. 2013a. Comparison of three methods for soil fertility quality spatial simulation with uncertainty assessment. *Soil Sci Soc Am J*. **77**: 2182–2191.

Qu M K, Li W D, Zhang C R. 2013b. Assessing the spatial uncertainty in soil nitrogen mapping through stochastic simulations with categorical land use information. *Ecol Inform*. **16**: 1–9.

Rengel Z, Tang C, Raphael C, Bowden J W. 2000. Understanding subsoil acidification: Effect of nitrogen transformation and nitrate leaching. *Aust J Soil Res*. **38**: 837–849.

Shi W J, Liu J Y, Du Z P, Song Y J, Chen C F, Yue T X. 2009. Surface modelling of soil pH. *Geoderma*. **150**: 113–119.

Shi W J, Liu J Y, Du Z P, Stein A, Yue T X. 2011. Surface modelling of soil properties based on land use information. *Geoderma*. **162**: 347–357.

Shi X Z, Yu D S, Xu S X, Warner E D, Wang H J, Sun W X, Zhao Y C, Gong Z T. 2010. Cross-reference for relating Genetic Soil Classification of China with WRB at different scales. *Geoderma*. **155**: 344–350.

Sun B, Zhou S L, Zhao Q G. 2003. Evaluation of spatial and temporal changes of soil quality based on geostatistical analysis in the hill region of subtropical China. *Geoderma*. **115**: 85–99.

Wang K, Zhang C R, Li W D. 2013. Predictive mapping of soil total nitrogen at a regional scale: A comparison between geographically weighted regression and cokriging. *Appl Geogr*. **42**: 73–85.

Wang Q Y, Zhou D M, Cang L. 2009. Microbial and enzyme properties of apple orchard soil as affected by long-term application of copper fungicide. *Soil Biol Biochem*. **41**: 1504–1509.

Wei S C, Jiang Y M. 2012. Investigation on fertilization status in apple orchards of Shandong Province. *Shandong Agr Sci* (in Chinese). **44**: 77–79.

Wu C F, Wu J P, Luo Y M, Zhang H B, Teng Y. 2008. Statistical and geoestatistical characterization of heavy metal concentrations in a contaminated area taking into account soil map units. *Geoderma*. **144**: 171–179.

Wu C F, Wu J P, Luo Y M, Zhang L M, DeGloria S D. 2009. Spatial estimation of soil total nitrogen using cokriging with predicted soil organic matter content. *Soil Sci Soc Am J*. **73**: 1676–1681.

Wu J P, Norvell W A, Welch R M. 2006. Kriging on highly skewed data for DTPA-extractable soil Zn with auxiliary information for pH and organic carbon. *Geoderma*. **134**: 187–199.

Xie Y F, Chen T B, Lei M, Yang J, Guo Q J, Song B, Zhou X Y. 2011. Spatial distribution of soil heavy metal pollution estimated by different interpolation methods: Accuracy and uncertainty analysis. *Chemosphere*. **82**: 468–476.

Xu R K. 2015. Research progresses in soil acidification and its control. *Soils* (in Chinese). **47**: 238–244.

Xue D, Yao H Y, Huang C Y. 2006. Microbial biomass, N mineralization and nitrification, enzyme activities, and microbial community diversity in tea orchard soils. *Plant Soil*. **288**: 319–331.

Yates S R, Warrick A W. 1987. Estimating soil water content using cokriging. *Soil Sci Soc Am J*. **51**: 23–30.

Zhang C P, Wu P, Tang C Y, Tao X Z, Han Z W, Sun J, Liu H. 2013. The study of soil acidification of paddy field influenced by acid mine drainage. *Environ Earth Sci*. **70**: 2931–2940.

Zhang S W, Huang Y F, Shen C Y, Ye H C, Du Y C. 2012. Spatial prediction of soil organic matter using terrain indices and categorical variables as auxiliary information. *Geoderma*. **171**-**172**: 35–43.

Zhang Z Q, Yu D S, Shi X Z, Wang N, Zhang G X. 2015. Priority selection rating of sampling density and interpolation method for detecting the spatial variability of soil organic carbon in China. *Environ Earth Sci*. **73**: 2287–2297.

Zhang Z Q, Yu D S, Shi X Z, Warner E, Ren H Y, Sun W X, Tan M Z, Wang H J. 2010. Application of categorical information in the spatial prediction of soil organic carbon in the red soil area of China. *Soil Sci Plant Nutr*. **56**: 307–318.

Zhang Z Q, Yu D S, Shi X Z, Weindorf D C, Sun W X, Wang H J, Zhao Y C. 2011. Effects of prediction methods for detecting the temporal evolution of soil organic carbon in the Hilly Red Soil Region, China. *Environ Earth Sci*. **64**: 319–328.