Soil texture distribution simulation and risk assessment using transition probability-based geostatistics

Xiaopeng Li[1,2], Jianli Liu[1]*, Jiabao Zhang[1], Weipeng Wang[1,3], and Wenwen Xin[1,3]

[1]Institute of Soil Science, Chinese Academy of Sciences, 71 East Beijing Road, Nanjing 210008, China
[2]Department of Renewable Resources, University of Alberta, 116 St. and 85 Ave., Edmonton, AB T6G 2E3, Canada
[3]University of Chinese Academy of Sciences, 19A Yuquan Road, Beijing 100049, China

A b s t r a c t. Three dimensional soil textural structure in a township was conditionally simulated using a transition probability-based indicator geostatistical method based on 270 soil texture samples from 27 profiles. Additionally the distribution of soil profiles lacking clay interlayers (indicating high irrigation water and nutrient leaching risk) was analyzed using 500 realizations from the simulation. The results indicated that the simulation could predict the soil texture distribution with low uncertainties using the existing data, and the predicted soil map (0-10 cm) formed by the maximum probable soil textures also exhibited a good agreement with the legacy soil survey map. For water and nutrient leaching risk analysis, the areas lacking clay interlayer could be located; however, their distribution was still highly uncertain if based only on the existing sampling data. That means supplementary sampling in future is required for the risk assessment, and the existing study can help to optimise the sampling points and their distribution. Generally, the transition probability-based geostatistical simulation, as a stochastic conditional simulation method, exhibited its potential in soil texture spatial reproduction and related risk assessment.

K e y w o r d s: soil texture, transition probability, geostatistics, conditional simulation, risk assessment

INTRODUCTION

As the 'skin' of the earth, soil ties in many aspects with the circulations of water, air and other inorganic/organic matters (Berner and Berner 2012; Condon *et al.,* 1992; Ek and Cuenca 1994; Knapp *et al.,* 1993). Many important ecological processes, such as evapotranspiration, soil water drainage, solute transport, and plant growth, are directly affected by soil properties and their spatial distribution (Golchin *et al.,* 1994; Kaye *et al.,* 2005). Due to the complexity of the pedogenic process affected by eluviation-illuviation, bioturbation and surface removal, erosion, deposition or inheritance, *etc*. (Phillips, 2001), soil texture usually shows apparent variations in space. Accurate characterization of the spatial distribution and variation of soil properties, especially the basic physical properties such as soil texture, has been regarded as an important subject for regional hydrological, agricultural and environmental studies (Racz *et al.,* 2012; Wang *et al.,* 2009, 2010). However, as restricted by labour and fund support, intensive sampling in a large area still seems currently impracticable in most cases. Therefore, introducing some new methods and strengthening the sampling data analysis to capture more spatial pattern information without extra labour and expense appears to be a way to solve the problem.

Taking 'everything is related to everything else, but near things are more related than distant things (Tobler, 1970)' as the fundamental principle, various kinds of spatial interpolation and stochastic simulation methods have been developed. In most studies, samples were usually collected at points, and then the point/borehole/profile sampling data were extended to surface or three dimensions using interpolation or stochastic simulation. Among the numerous data interpolation schemes *ie* nearest-neighbour, inverse distance weighted (IDW), radial basis function (RBF), spline and geostatistical (Kriging) methods (Cressie, 1993; Li and Heap 2011). Kriging interpolation is the most popular and widely used at present (Bocchi *et al.,* 2000; Gonzalez and Zak, 1994; Meul and Van Meirvenne 2003). However, the interpolation methods, including Kriging, actually have an apparently smoothing effect on the predicted results

---

*Corresponding author e-mail: jlliu@issas.ac.cn

in unsampled areas, and lead to an overall inaccuracy of the output (Goovaerts, 1999). Some stochastic simulation methods, such as variogram-based Sequential Gaussian Simulation (SGS) and Sequential Indicator Simulation (SIS), can generate realizations of the study area from point/borehole samples to overcome this shortcoming (Deutsch and Journel 1998; Juang *et al.,* 2004; Zhao *et al.,* 2005). However, there are still some problems when applying the variogram-based SGS and SIS methods to build a three dimensional soil structure and perform uncertainty analysis: first, sampling points are often too sparse to calculate the variogram in all directions, especially in the horizontal directions; second, the variogram-based methods cannot account for spatial relationship with strong asymmetric juxtapositional tendencies, which are actually very common in field; and furthermore, existing geological and subjective information cannot be directly infused into these methods to reduce the spatial uncertainty (Deutsch and Journel, 1998).

Transition probability-based indicator geostatistics uses transition probability instead of variogram to realistically characterize the three dimensional properties of the mediums (Carle and Fogg, 1997). In comparison to the variogram-based geostatistical methods, the transition probability-based approach 'improves consideration of spatial cross-correlations and facilitates the integration of geologic interpretation of facies architecture into the simulation'; and it is especially well-suited for typically sparse geological data sets (Carle, 1997). With the advantages of processing sparse data and anisotropy, transition probability-based indicator geostatistical simulation seems to be a good method for the prediction of spatial distribution of soil texture and related agricultural and environmental risk analysis.

The objective of this study was to apply the transition probability-based geostatistical simulation method for three-dimensional soil texture distribution prediction within a township scale in North China Plain, to analyse the uncertainty level based on the existing sampling data, and to assess the high water and nutrient leaching area (lack of clay interlayers) distribution using the simulation result.

MATERIALS AND METHODS

The sampling work was conducted in Pandian Township, Fengqiu County, Henan Province, China. It is only about 10 km away from the Yellow River in its midstream, and in the centre of the North China Plain (114.51°- 114.60°E, 34.98° -35.06°N), and covers an area of 32 km² (Fig. 1). The terrain is relatively flat in the area, with mean, minimum and maximum elevations of 64, 52, and 71 m, respectively. Most lands within the study area are productive farmlands and have been cultivated for over a thousand years. Nowadays, agriculture still plays a dominant role in the local economy.

Being situated close to the Yellow River, the soils were mainly formed by river alluviation. All the soil parent materials were sourced from the Yellow River deposition during the Holocene Epoch of Quaternary Period. Fluvo-aquic soil (Och-AquicCambisol) has the widest distribution and accounts for 98.32% of the topsoil in Fengqiu County. Aeolian sandy soil (Ust-SandicEntisol) which exists in the form of sand banks in the north part of the study area accounts for 1.68% of the topsoil in the county. Due to the fluctuation of water and sediment amount of the Yellow River, variation of local sedimentary environments, wind transportation and human influence, layering structures are very common and textural heterogeneity usually exists in both horizontal and vertical directions (Wang, 1987). The most common layering structure known by the local farmers is the sandy or sandy loam soil body with clay interlayer(s) at 40-80 cm depth in the profile. Since the clay interlayer limits the water and fertiliser leaching and contributes so significantly to water and nutrient conservation and crop yield, this kind of textural configuration was also named the 'Mengjin' soil by local farmers, which means 'gold-covered'. According to previous investigations, Mengjin soil accounts for 83.4% of the farmland area in Fengqiu county, and among these soils, had clay interlayer(s): 70.9% above 60 cm, and 12.5% below 60 cm (Li *et al.,* 2007; Wang, 1987). Figure 1 (right) shows the soil texture distribution deduced from a 1: 100 000 legacy soil map (Tian, 1982). For compatibility with the categories in the study, soils in the map were reclassified to four classes:

A – aeolian sandy soils;

B – light-textured fluvo-aquic soils (mostly sandy loam, loam, and silt loam);

C – light to middle-textured fluvo-aquic soils (mostly sandy clay loam, clay loam, and silt clay loam);

D – and other heavy-textured soils.

The complexity of the soil spatial variation in the study area may make geostatistical interpolation methods be faced with some difficulties to reconstruct the soil texture spatial patterns. However, the stochastic sequential indicator simulation based on transition probability should perform well and get some promising results due to its advancement in theory.

The soil sampling scheme was designed based on uniform grids and then adjusted according to their availability and accessibility. In total, 27 soil profiles up to 200 cm deep, positioned by a portable GPS (Garmin eTrex Vista, positioning error <15 m), were excavated within and around Pandian Township during 2010- 2011 (Fig. 2a). Soil samples were collected at every 20 cm in all profiles, and there were 270 samples in total. Soil particle size distributions were determined by a laser diffraction particle size analyser (Beckman Coulter LS 13-320) after organic matter removal and particle dispersion.

The 270 samples were classified into four groups with soil texture from light to heavy (Fig. 2b). According to the soil texture classification of International Soil Science Society (ISSS), class A represents the soils with clay less than 15% and sand more than 85% (*ie* sand or loamy sand),
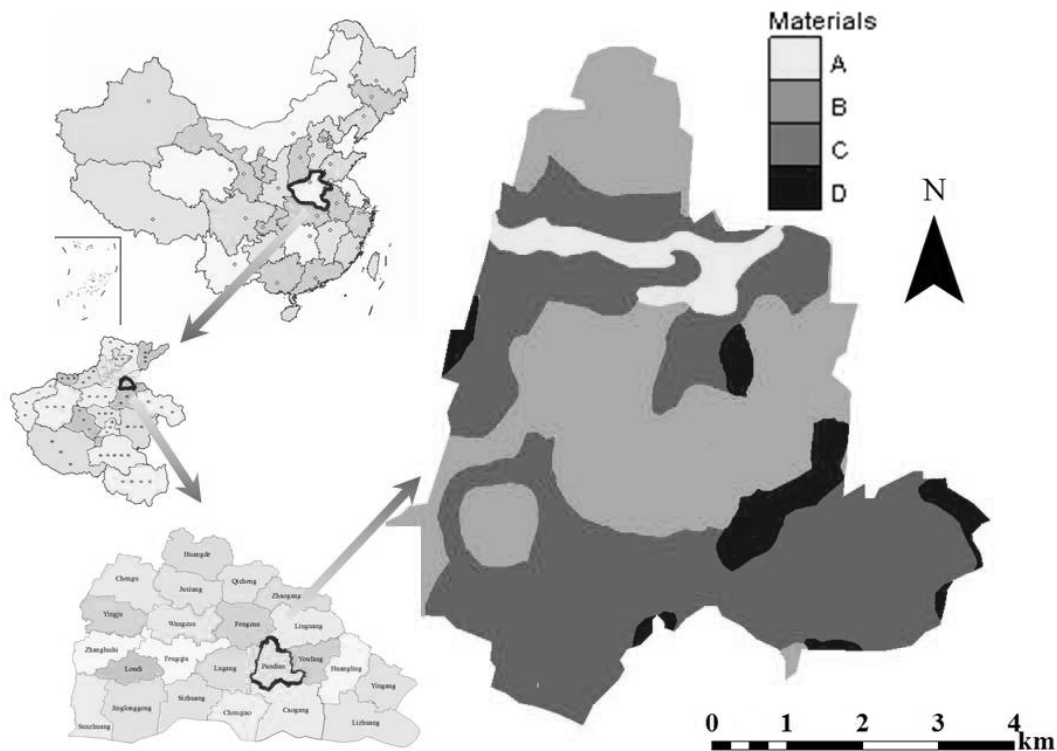
Fig. 1. The location of the study area and soil texture classification in topsoil from the legacy soil map. For compatibility with the simulation results, the map was reclassified: aeolian sandy soils as class A; light-textured fluvo-aquic soils (mostly sandy loam, loam, and silt loam) as class B; light to middle-textured fluvo-aquic soils (mostly sandy clay loam, clay loam, and silt clay loam) as class C; and other heavy-textured soils as class D.
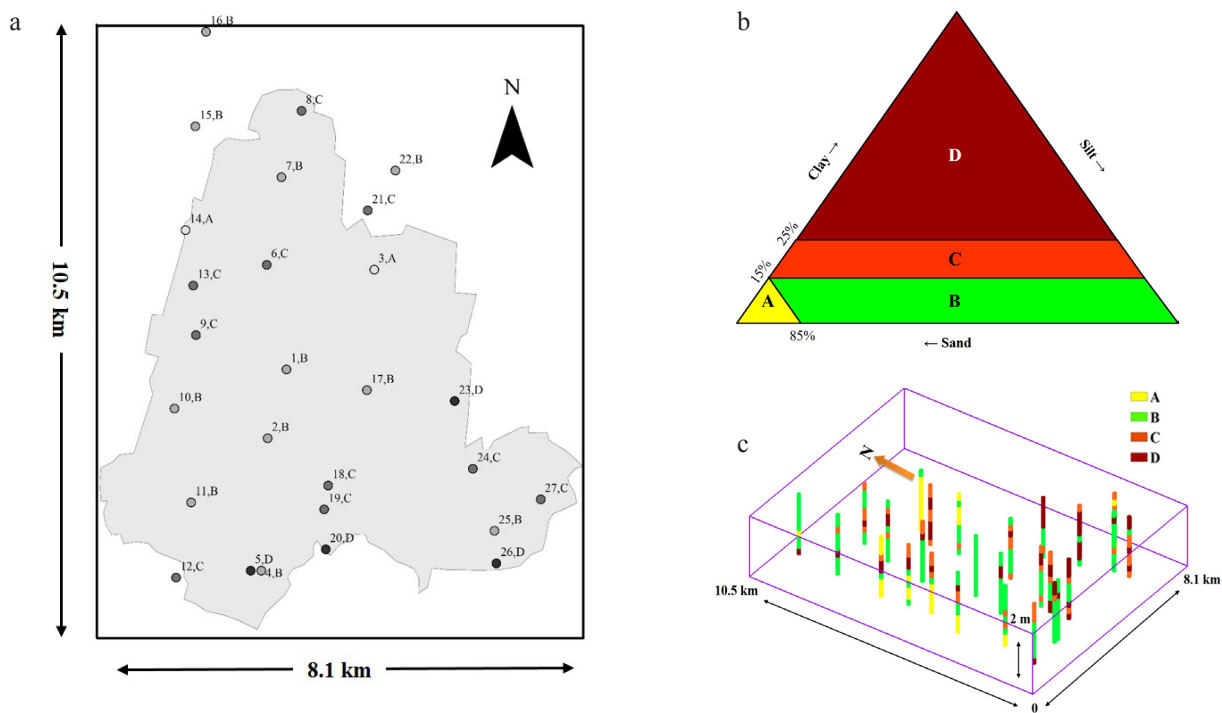

Fig. 2. (a) sampling sites (the label represents the name of the site and the soil texture class in the 0-10 cm layer, and the rectangle denotes the simulation area in Fig. 4), (b) four classes of soil textures according to international soil texture classification, and (c) soil texture classification in the sampled profiles (the frame denotes the simulation area in Fig. 4).

class B represents the soils with clay particles less than 15% and sand less than 85% (*ie* sandy loam, loam and silt loam soils), class C represents the soils with 15-25% of clay particles (*ie* sandy clay loam, clay loam, silt clay loam), and class D represents the other soils with high proportion of clay particles (*ie* sandy clay, loamy clay, silt clay, clay and heavy clay). By dividing the soils into these four classes, the hydraulic properties among the different soil classes became generally distinguishable, and the vertical soil texture configurations according to this classification at the sampled sites are exhibited in Fig. 2c.

The transition probability-based indicator geostatistical simulation method was firstly proposed by Carle and Fogg (1996). As an indicator Kriging-based sequential simulation method, its main features include indicator Kriging for estimator, transition probability as a measure for quantitatively describing the characters of spatial structure variation, SIS algorithm for stochastic simulation, and simulated quenching for post-processing (Deutsch and Journel 1998). In the method, Markov chain models are derived to fit the transition probability trends in directions for indicator Kriging estimation, and it offers an 'interpretable and mathematically simple yet powerful stochastic model for categorical variables' (Carle, 1999). Transition probability $t_{jk}(h)$ is defined as:

$$t_{jk}(h) = \text{Prob } (k \text{ occurs at } x + h \mid j \text{ occurs at } x) \quad (1)$$

where: x is spatial location, h is the lag (separation vector), and j, k denote mutually exclusive categories such as soil textures or geological units. Using the definition of conditional probability, the transition probability Prob $(B'|A)$ can be described as:

$$\text{Prob } (B' \mid A) = \frac{\text{Prob } (A \text{ and } B')}{\text{Prob } (A)} \quad , \quad (2)$$

where: A represents (j occurs at x) and B' represents (k occurs at $x + h$).

In the study, T-PROGS, the Transition PRObability-based indicator GeoStatistics software package developed by Carle and Fogg (1996), was used to generate stochastic realizations for the spatial patterns analysis of soil texture. T-PROGS provides an integrated technique with three components: transition probability model inference, transition probability-based sequential indicator simulation (with full indicator Kriging), and simulated quenching for post processing. The implementation of transition probability geostatistical method in T-PROGS mainly involves the following steps:

1. Calculating the measured vertical transition probabilities of the regionalized variables (*eg* soil texture) based on sampling data, by subprogram GAMEAS which can compute a set of transition probability curves as functions with lag distance at given sampling intervals.

2. Fitting the measured transition probability using Markov chains and formulating simulated transition trend equations. In the study, embedded Markov chains were employed since the soil layer interfaces often occur discretely.

3. Deriving the transition trends in horizontal directions from the vertical equations, using lens length ratios which represent the lens length of each material in horizontal directions vs. the lens thickness in vertical (Carle, 1999). Here, Walther Law, which states the lateral successions of a deposit can be represented by the vertical successions of deposited facies, applies, and then the transition probability models in horizontal directions can be derived from the vertical equations.

4. Generating three-dimensional stochastic realizations of material sets using subprogram TSIM based on the fitted transition probability models in all three primary directions. All realizations are conditioned to borehole data and possess the inherent transition tendencies in the measurements.

Uncertainty analysis is a part of risk analysis that focuses on the uncertainties in the data characteristics. Important components of uncertainty analysis include qualitative analysis to identify the uncertainties, quantitative analysis the uncertainties affecting the decision process, and communication of the uncertainty (Funtowicz and Ravetz 1990; Katz 2002; Regan *et al.,* 2002; Smith 2002). There are two kinds of uncertainties: local and spatial. Local uncertainty refers to an uncertainty involving a single location, and spatial uncertainty refers to an uncertainty related to multiple locations (Delbari *et al.,* 2009). Accordingly, the uncertainty analysis of the soil texture spatial patterns in the study includes the assessments of both local and spatial uncertainties.

Local uncertainty, the uncertainty of soil texture at a single location x', is modelled by a probabilistic model. The probability of soil type 'X' occurring at an unknown site x', Prob ($x' = X$), can be calculated as:

$$\text{Prob } (x' = X) = \frac{n \ (x' = X)}{n_{total}} \quad , \quad (3)$$

where: $x' = X$ represents a situation where the soil texture at location x' is 'X', n ($x' = X$) is the accumulated times where soil 'X' occurs at x' in all the realizations, and ntotal is the total number of the realizations.

Spatial uncertainty indicates the uncertainty of soil texture spatial distribution prevailing jointly at several locations or depths. It can also be described as the reliability that the area is delineated under a given condition 'Prob ($x' = X$) $\geq p_c$' (Goovaerts *et al.,* 1997). To assess the spatial uncertainty, the critical area under a given critical probability p_c should be firstly delineated, and then the joint probability is used to assess the spatial uncertainty of the critical area. The critical areas are delineated by the condition where the probabilities of soil X occurring at all locations are all greater than p_c. It means that, at all the locations of the critical areas, x_1, x_2, …, and x_m, Prob ($x_1 = X$),

Prob ($x_2 = X$), …, and Prob ($x_m = X$) are greater than p_c, where m is the total number of locations in the critical area. After the critical area is delimitated, the spatial uncertainty of this area can be assessed using joint probability. The joint probability that all the m locations in the critical areas in all realizations are soil 'X' is calculated as:

$$\text{Prob } (x_1 = X, \ x_2 = X, \ \dots, x_m = X) = \frac{n \, (x_1 = X, \ x_2 = X, \ \dots, x_m = X)}{n_{total}},$$

(4)

where: $x_1 = X$, $x_2 = X$, …, and $x_m = X$ represent the joint probability that soils at m locations x_1, x_2, …, and x_m are soil 'X'; $n \, (x_1 = X, x_2 = X, \dots, x_m = X)$ is the number of realizations that the simulated soil textures at m locations x_1, x_2, …, and x_m are all soil 'X'; and ntotal is the total number of the realizations.

Local uncertainty can be used to evaluate the reliability of the prediction at a certain unsampled location, or to locate the critical area according to certain critical probability p_c. Spatial uncertainty can be applied to assess the reliability of some multiple points related result, such as high water and nutrient leaching area (Delbari et al., 2009).

In the North China Plain where the study area is located, one of the big concerns for the farmers and the local government is the high leaching risk of irrigation water and nutrients due to poor soil water conservation properties and/or improper irrigation management. The hydraulic behaviours of clay and sandy soils are quite distinct. Different textured soils and their configuration in the soil body have large impacts on soil water and solute movements, and are crucial for soil water and nutrient conservation (Li et al., 2013; Zettl et al., 2011). According to previous studies, soils with higher leaching risk are usually found at locations lacking clay interlayers in the root zone ('golden cover' missing) (Li et al., 2010).

In the study, we considered the soil profiles without enough clay interlayers as high risk areas for water and nutrient leaching. From the geostatistical simulation results, the soil profiles where clay interlayers (soil D) were missing, or only of limited thickness, were delineated based on some selection criteria. It was formed by the loca-

tions where there was a high probability that the fraction of soil D within the profile was less than a certain thickness. We supposed that the 0-200 cm soil profile with less than 20 cm of clay had high water and nutrient leaching risk, and then two levels of clay texture deficiency were set: severe, the total thickness of soil D in 0-200 cm soil profile being less than 10 cm [Prob (D ≤ 5%)]; and slight, the thickness of soil D being more than 10 cm but less than 20 cm [Prob (5 < D ≤ 10%)].

In the study, the uncertainty of the simulation based on the existing sampling data was analyzed by local uncertainty analysis. The water and nutrient leaching risk areas were firstly located by local uncertainty, and then their reliability was assessed by spatial uncertainty analysis.

RESULTS AND DISCUSSION

Geostatistical Simulation Results and the Uncertainty in Soil Texture Distribution.

According to the general statistics of the samples (Table 1), sandy loam, loam and silty loam (soil B) were dominant in the study area, accounting for up to 48.9% at all sampled locations, and they occurred in 92.6% of all the sampled profiles, compared to 37.0, 81.5, and 77.8% for soil A, C, and D, respectively. The vertical lens length means the mean thickness of a soil layer in the vertical direction. Not surprisingly, soil B had a maximal mean thickness in the vertical as well. For these reasons, soil B was selected as the background soil type in this study. Soil A showed not only a relatively large vertical lens length but also a large standard deviation compared to its proportion, indicating that soil A was not evenly distributed but more probably distributed in clusters or strips. Additionally, the so-called 'Mengjin' soil, which had a clay interlayer (soil D) at 40-80 cm depth serving as water and nutrient conserving unit, took up to 51.9% of all the sampled profiles, and if taking the 0-200 cm profile into account, the soils with clay interlayer(s) in the profile increased to 77.8%.

The three-dimensional soil body of the study area was discretized into cubic cells before the simulation. In the study, the horizontal and vertical intervals were set as

T a b l e 1. General statistics of the sampled soil texture data in the study area

Soil class	Proportion (%)	P_e (%)	L_v (m)	$STDEV_{Lv}$ (m)	Soil texture
A	13.0	37.0	0.54	0.44	Sand and loamy sand
B	48.9	92.6	0.56	0.43	Sandy loam, loam and silty loam
C	19.3	81.5	0.27	0.12	Sandy clay loam, clay loam, silty clay loam
D	18.9	77.8	0.32	0.20	Sandy clay and clay

P_e – the probability of a soil class occurs in a soil profile; L_v – the mean lens length of a soil class in vertical direction; $STDEV_{Lv}$ – the standard deviation of the mean lens length in vertical direction.

100 m and 10 cm, respectively, and there were 105×81×20 cells in total. The soil properties, including texture within a cell, were thought to be homogeneous.

Experimental transition probabilities of soil texture in the vertical were calculated from the sampled profile (borehole) data, and then fitted by Markov Chain models which would be further used in the stochastic simulation to generate realizations. The Markov Chains can be defined by the proportions of the materials, lens lengths, and transitions rates. In the study, the 'maximum entropy factor' method was used to fit the Markov Chains to the measured transition probability curves. The maximum entropy factor represents the ratio of the transition rate to the maximum entropy transition rate, and can indicate the two categories that tend to occur next to each other when greater than 1.0 and the opposite when less than 1.0. Using this method, the proportions and lens lengths of the materials were chosen to use the experimental data, and the transition rates from one material to itself (diagonal terms) were determined by experimental means of lens lengths, and the transitions rates to others (off-diagonal terms) were determined by the maximum entropy factors. The experimental transition probabilities and their best-fitted curves revealing the spatial transition relationships between different soils are presented in Fig. 3. As can be seen from the diagrams, all the experimental transition probabilities tended to approach constants with the expansion of space lag, and most of them were well fitted by the Markov Chains. Only the experimental A → A and A → B transitions had some fluctuations at the lag spacing between 1.0 and 1.8 m. This is probably due to the relatively larger vertical lens length and greater variation of soil A compared to other soils (Table 1).

In Fig. 3, the diagonal terms indicate the transition probability from one category to itself, and the off-diagonal terms mean the transition probabilities from one category to another. According to the feature of transition probability, the four approached constants in the same column (*eg* A → A, A → B, A → C, and A → D) should be equivalent, and the value denoted the proportion or the 'marginal probability' of this soil type in all the soil types (Carle, 1999). The sum of the four marginal probabilities equalled one. As shown in Fig. 3, apparently asymmetrical transition trends of soil texture in the vertical direction were exhibited. For example, the transition probability of A → B (adjacent soil layers from above to below switching from soil A to B) was apparently greater than the probability of B → A, and the

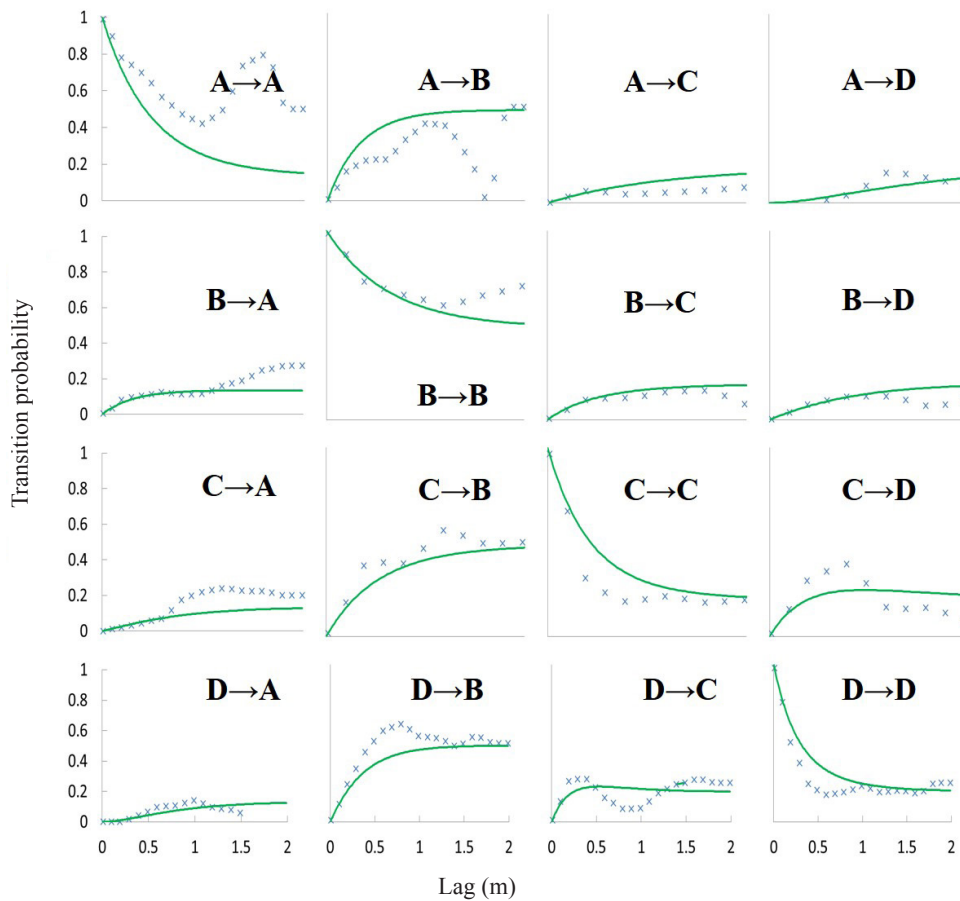**Fig. 3.** Matrix of transition probabilities measured (×) and modelled (—) in vertical direction.

probabilities of C → B and D → B were also larger than the probabilities of B → C and B → D. This kind of asymmetry was actually quite common in geological sediment and soil formation, however it cannot be expressed by the 'traditional' variogram-based geostatistical methods based on symmetric hypothesis (Deutsch and Journel, 1998). In addition, the transition probabilities of A → D and D → A were keeping quite low with the distance expansion, which meant that sandy or loamy sand soil, and clay or sandy clay soil, can scarcely co-exist in the same soil profile in field.

In theory, soil texture transition probabilities in horizontal directions can be directly calculated using experimental data. However, in most cases, the sampling data in the horizontal were usually too sparse to get a meaningful result because of the restrictions of sampling labour and funds. Therefore, we employed an indirect method to generate horizontal transition probabilities. According to Walther Law, the successions of geological units in lateral directions were consistent with their configuration in the vertical. The horizontal transition probabilities can be obtained by
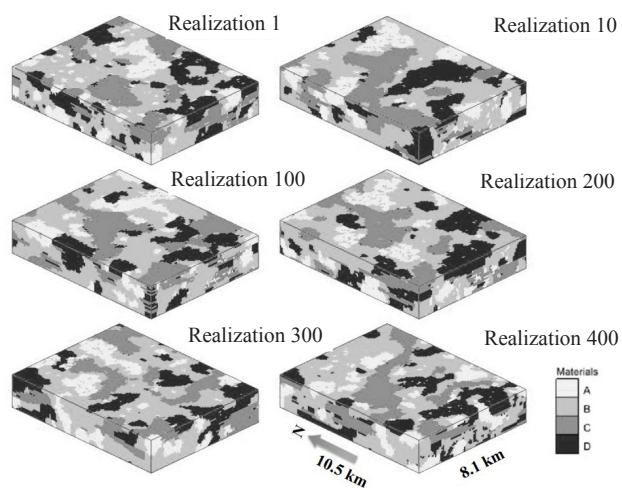

Fig. 4. Example realizations of soil texture distribution in the study area (including some adjacent area and vertical scale expanded by 1 000 times for visualization). Materials A, B, C, and D follow the same soil classification as that in Fig. 2.

the experimental lens length ratios (*ie* the ratio of the soil lens length in x or y direction with the length in z direction) and the vertical transition probabilities (Carle, 1999). In the study area, although some soils at certain depth (*eg* soil A at 0-10 cm depth) did show some horizontal anisotropy according to the soil survey map, when considering all depths and all soil types the anisotropic trends seemed not consistent. Therefore, we kept the assumption of horizontal isotropy in the study *ie* the lens length ratios in x and y directions were equivalent. The soil survey map (Tian, 1982) was used to deduce the lens length ratios by comparing the mean diameters of soil texture polygons with the measured lens lengths in the vertical direction, and the lens length ratios of soil A, B, C, and D were set as: 1 000:1, 1 200:1, 4 000:1, and 2 000:1, respectively.

Based on the Markov Chain models of transition probabilities in all three primary directions, 500 soil textural distribution realizations were simulated by SIS. As seen in Fig. 4, the stochastic realizations reproduced the overall soil texture spatial characteristics as the sampling data indicated. Soil B, which took up the largest proportion in the sampling data, was mostly interconnected in the realizations, and contrastingly, soil A, which only accounted for 13% in the sampling data, showed up as disconnected spots. A comparison of the statistic results of the 500 realizations with the measured data is listed in Table 2. If there were adequate realizations, the mean probability of a soil class occurring in all realizations (the simulated proportion) should be equivalent to the corresponding proportions in the sampling data; the predicted mean lens lengths should also approach their measured values. In the study, the simulated proportions of four soil classes and their mean lens lengths in the vertical direction all showed good agreements with the sampling data. It also confirms that the simulation captured the soil texture spatial patterns, and further studies can be carried out based on the realizations.

Using the 500 realizations, the occurrence probabilities of the four soil classes at all locations and depths were calculated, and Fig. 5 shows their distributions at surface (0-10 cm), moderate depth (90-100 cm), and bottom (190-200 cm). In general, the higher the probability, the more possible the occurrence of that soil texture at the location.

T a b l e 2. Comparison of statistic results from the simulated realizations and the field samples

Soil class	Proportion (%)	$Prob_s$ (%)	Lv_m (m)	Lv_s (m)
A	13.0	13.4	0.54	0.55
B	48.9	48.0	0.56	0.64
C	19.3	19.2	0.27	0.27
D	18.9	19.3	0.32	0.36

Proportion represents the proportion of each soil class from soil sampling; $Prob_s$ represents the mean probability of a soil class occurred in all realizations; and Lv_m and Lv_s are the measured and simulated average lens lengths, respectively.
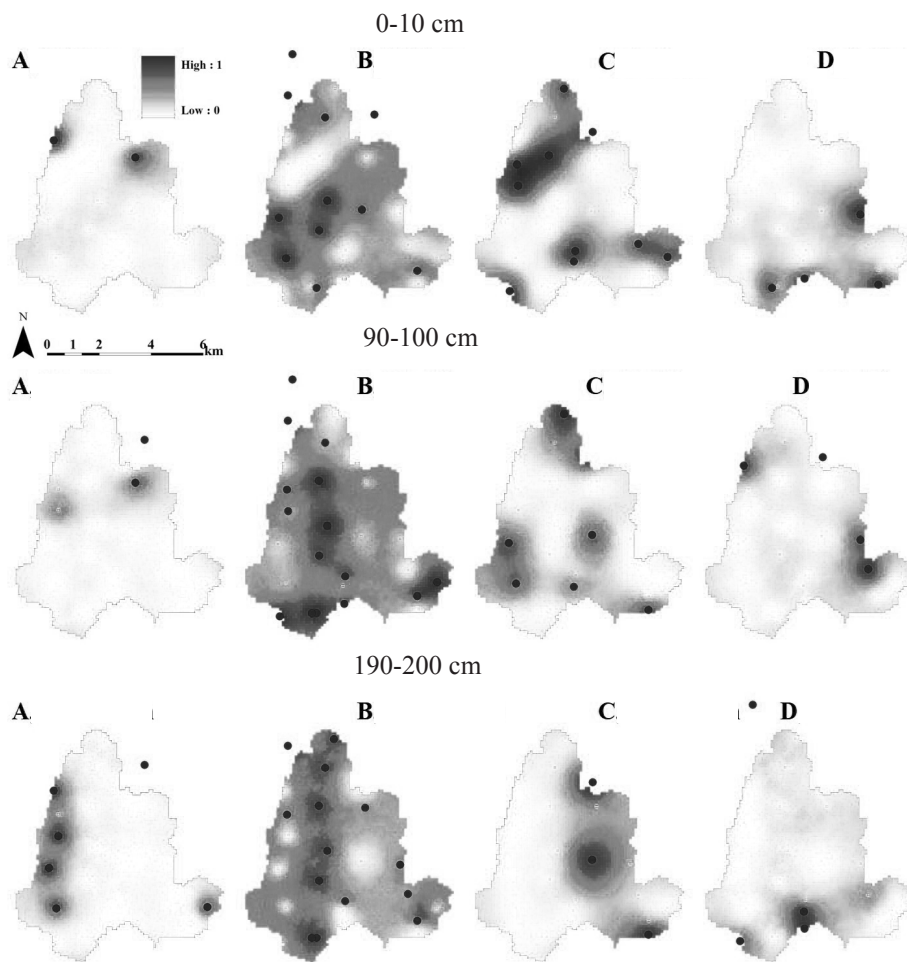
Fig. 5. Probability maps of various soil textures occurring at different depths (surface 0-10 cm, middle 90-100 cm, and bottom layer 190-200 cm).

It can be seen that the probability distributions at different depths exhibited apparent differences. For example, soil C at 200 cm depth had a high possibility of occurrence in the eastern part of the study area, and that was quite different from its probability distribution at the surface. Soil B, which was predominant in the sampling data, had large occurrence probabilities at most locations. Based on the probability maps at various depths, we found the soil textures with the highest occurrence probabilities at the locations in the study area (Fig. 6a) and their corresponding probabilities (Fig. 6b). As for the surface layer (0-10 cm), the simulated soil texture map (Fig. 6a, left) was quite consistent with the legacy soil map in Fig. 1. The highest occurrence probability (Fig. 6b) was represented by the determination/certainty level at the locations. If the highest probability at a location was low, it indicated that the soil type there was uncertain. In Fig. 6b, the highly uncertain areas are crossed by the 50% probability contours. It can be seen that the major parts of the study area had deterministic levels over 50%, and the high uncertainties, no matter at what depth, were always located at locations with too

sparse sampling sites or the sampled soil texture was in contrast with the others nearby. For example, in the southeastern part of the study area, the soil textures at the three exhibited depths all seemed relatively uncertain due to insufficient sampling sites. This result provides a basis for the supplementary sampling in future: more work should be conducted in these areas in order to reduce the overall uncertainty.

Figure 7 shows the probability distributions of clay interlayer deficiency (at two levels) calculated from 500 realizations. Areas with critical probabilities Prob (D ≤ 5%) (total clay proportion less than or equal to 10 cm) and Prob (D ≤ 10%) (total clay proportion less than or equal to 20 cm) were positioned, respectively. At the locations with high probability of clay interlayer deficiency, the water nutrient leaching risk was supposed to be correspondingly high. Generally, in the study area, soils with high water and nutrient leaching risk were clustered in a strip in the northwest-southeast direction. The soils with more than 50% probability of severe lack of water and nutrient conservation ability accounted for a sizeable proportion, and along
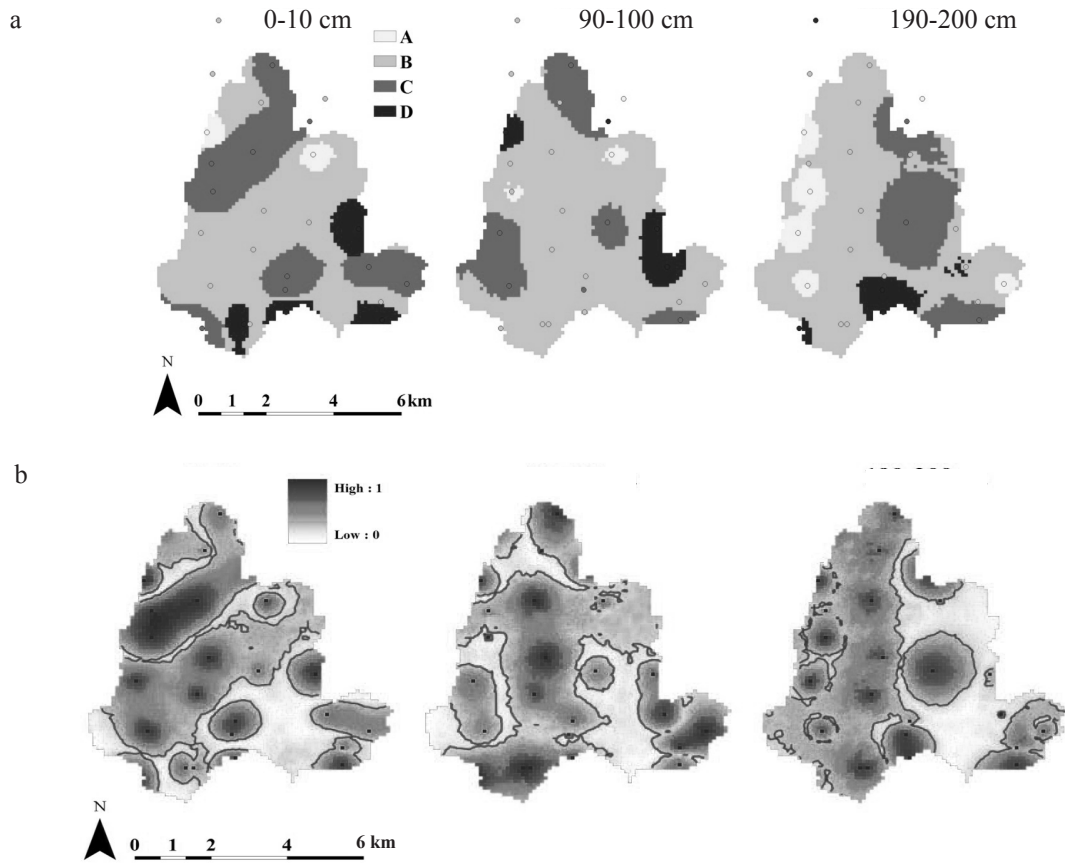
Fig. 6. (a) soil texture maps with the highest probability occurring at the location, and (b) the corresponding maximum probabilities (the solid lines represent the 50% contour).
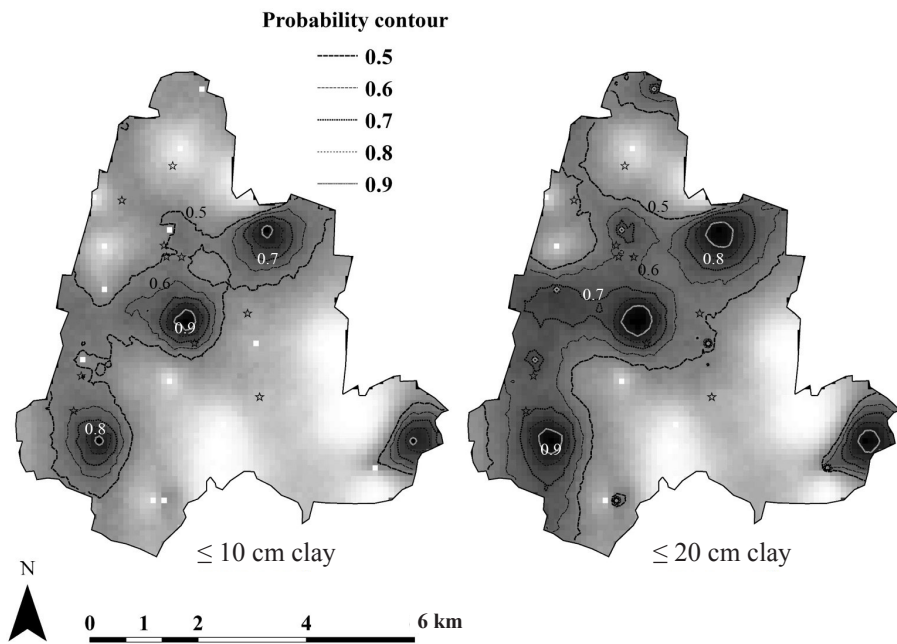

Fig. 7. Probability maps with the thickness of soil D being 0-10 cm (left) or 0-20 cm (right) in the 0-200 cm soil profile, and wheat yield sampling sites (✳).

with the slight level, it would make up to almost half of the study area, which should arouse the attention of the local agricultural managers. The distribution of the high risk area was confirmed by crop production data. According to the crop yield survey data in 2008 and 2009 (Fig. 7), the average yields at the sites with Prob (D = 0) > 50%, Prob (D $\leq$ 5%) > 50%, and Prob (D $\leq$ 10%) > 50% were of 5.74, 7.46, and 7.71 Mg ha^{-1}, respectively, compared to a mean wheat yield of 7.83 Mg ha^{-1} in the whole study area .

Figure 7 delineates the locations with high leaching risk and the corresponding determination level (local uncertainty). However, it provides no measure of spatial uncertainty to assess the reliability of these high risk areas because the conditional cumulative distribution functions (CCDFs) in the maps were only for single point rather than multi-point. In order to assess the reliability of the high risk areas, the joint probabilities at severe and slight levels were analyzed. Table 3 shows the joint probability variations with the increasing critical probabilities. A higher joint probability value meant lower uncertainty at unsampled locations and a more reliable delineation of the high risk areas. For the severe level, although the critical probability was over 80%, its joint probability was still less than 50%, indicating a high uncertainty/low reliability of the delineated high risk areas. The spatial uncertainty at the slight level became even higher because of the increasing acreage of the critical areas. A promising joint probability value could not be reached until the critical probability increased up to 90% (severe level) or 95% (slight level). The dilemma was that if we only used the existing sampling data for simulation, the high risk areas with high reliability would be only clustered close to the sampling sites. It meant that the existing sampling data could be used to successfully reproduce the soil texture spatial patterns, however, it still seemed insufficient

for confident determination of the high clay deficiency/water and nutrient leaching risk areas. More intensive sampling is required to make it reliable.

Another conclusion from Table 3 was that the spatial uncertainty levels were different even with the same sampling data when the critical areas were in different size. In other words, a larger critical area needed more sampling data to validate its reliability. This can also be used to estimate the approximate number of sampling and to guide the deployment of sampling density depending on research purpose.

CONCLUSIONS

1. Soil texture in the study area had significant spatial variations in both vertical and horizontal directions. The spatial pattern of soil texture in the study area was conditionally simulated using a transition probability-based indicator geostatistical method based on 270 soil texture samples from 27 profiles.

2. Using the existing sampling data, the probabilities for soil texture determination in the major parts of the study area were over 50%. The high uncertainty areas were always at the locations with too sparse sampling sites or very complex soil texture succession nearby. For example, the south-eastern part of the study area.

3. Although the simulation results based on the existing data could locate some critical areas with higher risk of water and nutrient leaching (areas where clay interlayers were missing or of limited thickness), the sampling seemed not sufficient to make the critical areas reliable because the spatial uncertainty remained high. However, this spatial uncertainty analysis results can still help estimate the supplementary sampling number and optimise the sampling distribution in future.

4. Generally, transition probability-based indicator geostatistical simulation can be an alternative upscaling method in addition to the conventional interpolation and variogram-based simulation methods, and it is also a useful tool for uncertainty and risk analyses.

T a b l e 3. Joint probabilities of soil D deficiency in the soil profile under the various critical probabilities

Critical probability $p_c(i)$ (%)	Joint probability (%)	
	0-10 cm clay	0-20 cm clay
50	0	0
60	0	0
70	0.2	0
80	1.8	0.4
90	62.4	19.4
95	91.2	66.8

0-10 cm clay and 0-20 cm clay represent the three soil D deficient levels with the thickness of soil D is less than 10 or 20 cm in the 0-200 cm soil profile, respectively.

REFERENCES

Berner E.K. and Berner R.A., 2012. Global Environment: Water, Air, and Geochemical Cycles. Princeton University Press, Princeton, NJ, USA.

Bocchi S., Castrignanò A., Fornaro F., and Maggiore T., 2000. Application of factorial kriging for mapping soil variation at field scale. Eur. J. Agron., 13, 295-308.

Carle S.F., 1997. Integration of geologic interpretation into geostatistical simulation. Annual Conf. Int. Association of Mathematical Geology (IAMG), September 22-27, Barcelona, Spain.

Carle S.F., 1999. T-PROGS: Transition Probability Geostatistical Software Version 2.1. Hydrologic Sciences Graduate Group, University of California, Davis, CA, USA.

Carle S.F. and Fogg G.E., 1996. Transition probability-based indicator geostatistics. Math. Geol., 28, 453-476.

Carle S.F. and Fogg G.E., 1997. Modeling spatial variability with one and multidimensional continuous-lag Markov chains. Math. Geol., 29, 891-918.

Condon A.G., Richards R.A., and Farquhar G.D., 1992. The effect of variation in soil water availability, vapour pressure deficit and nitrogen nutrition on carbon isotope discrimination in wheat. Aust. J. Agr. Res., 43, 935-947.

Cressie N.A.C., 1993. Statistics for spatial data. Wiley Press, New York, USA.

Delbari M., Afrasiab P., and Loiskandl W., 2009. Using sequential Gaussian simulation to assess the field-scale spatial uncertainty of soil water content. Catena, 79, 163-169.

Deutsch C.V. and Journel A.G., 1998. GSLIB, Geostatistical Software Library and User's Guide. Oxford University Press, New York, USA.

Ek M. and Cuenca R.H., 1994. Variation in soil parameters: Implications for modeling surface fluxes and atmospheric boundary-layer development. Bound-lay. Meteorol., 70, 369-383.

Funtowicz S.O. and Ravetz J.R., 1990. Uncertainty and Quality in Science for Policy. Kluwer Academic Publishers, Dordrecht, Boston, USA.

Golchin A., Oades J.M., Skjemstad J.O., and Clarke P., 1994. Soil structure and carbon cycling. Aust. J. Soil Res., 32, 1043.

Gonzalez O.J. and Zak D.R., 1994. Geostatistical analysis of soil properties in a secondary tropical dry forest, St. Lucia, West Indies. Plant Soil, 163, 45-54.

Goovaerts P., 1999. Geostatistics in soil science: state-of-the-art and perspectives. Geoderma, 89, 1-45.

Goovaerts P., Webster R., and Dubois J.P., 1997. Assessing the risk of soil contamination in the Swiss Jura using indicator geostatistics. Environ. Ecol. Stat., 4, 49-64.

Juang K.W., Chen Y.S., and Lee D.Y., 2004. Using sequential indicator simulation to assess the uncertainty of delineating heavy-metal contaminated soils. Environ. Pollut., 127, 229-238.

Katz R.W., 2002. Techniques for estimating uncertainty in climate change scenarios and impact studies. Climate Res., 20, 167-185.

Kaye J.P., McCulley R.L., and Burke I.C., 2005. Carbon fluxes, nitrogen cycling, and soil microbial communities in adjacent urban, native and agricultural ecosystems. Global. Change. Biol., 11, 575-587.

Knapp A.K., Fahnestock J.T., Hamburg S.P., Statland L.B., Seastedt T.R., and Schimel D.S., 1993. Landscape patterns in soil-plant water relations and primary production in tallgrass prairie. Ecology, 74, 549.

Li J. and Heap A.D., 2011. A review of comparative studies of spatial interpolation methods in environmental sciences: performance and impact factors. Ecological Informatics, 6, 228-241.

Li X., Chang S.X., and Salifu K.F., 2013. Soil texture and layering effects on water and salt dynamics in the presence of a water table: a review. Environ. Reviews, 21, 1-10.

Li X.P., Zhang J.B., Liu J.T., Liu J.L., Zhu A.N., Lv F., and Zhang C.Z., 2010. A modified checkbook irrigation method based on GIS-coupled model for regional irrigation scheduling. Irrigation. Sci., 29, 115-126.

Li Y., Chen D.L., White R.E., Zhu A.N., and Zhang J.B., 2007. Estimating soil hydraulic properties of Fengqiu County soils in the North China Plain using pedo-transfer functions. Geoderma, 138, 261-271.

Meul M. and Van Meirvenne M., 2003. Kriging soil texture under different types of nonstationarity. Geoderma, 112, 217-233.

Phillips J.D., 2001. Contingency and generalization in pedology, as exemplified by texture-contrast soils. Geoderma, 102, 347-370.

Racz A.J., Fisher A.T., Schmidt C.M., Lockwood B.S., and Los Huertos M., 2012. Spatial and temporal infiltration dynamics during managed aquifer recharge. Ground Water, 50, 562-570.

Regan H.M., Colyvan M., and Burgman M.A., 2002. A taxonomy and treatment of uncertainty for ecology and conservation biology. Ecol. Appl., 12, 618-628.

Smith E., 2002. Uncertainty analysis. In: Encyclopedia of Environmetrics (Eds A.H. El-Shaarawi, W.W. Piegorsch). Wiley Press, Chichester, UK.

Tian S.P., 1982. Soil map of Fengqiu County. Fengqiu County Bureau of Surveying and Mapping, Fengqiu, China.

Tobler W.R., 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography, 46, 234-240.

Wang E., Cresswell H., Bryan B., Glover M., and King D., 2009. Modelling farming systems performance at catchment and regional scales to support natural resource management. NJAS-Wageningen J. Life Sci., 57, 101-108.

Wang Y.Q., Shao M.A., and Liu Z.P., 2010. Large-scale spatial variability of dried soil layers and related factors across the entire Loess Plateau of China. Geoderma, 159, 99-108.

Wang Z.Q., 1987. Agriculture Development and Management Project in Tianranwenyanqu Basin (in Chinese). Sci. Press of China, Beijing, China.

Zettl J., Lee Barbour S., Huang M., Si B., and Leskiw L., 2011. Influence of textural layering on field capacity of coarse soils. Can. J. Soil Sci., 91, 133-147.

Zhao Y.C., Shi X.Z., Yu D.S., Wang H.J., and Sun W.X., 2005. Uncertainty assessment of spatial patterns of soil organic carbon density using sequential indicator simulation, a case study of Hebei province, China. Chemosphere, 59, 1527-1535.