Comparison of Three Methods for Soil Fertility Quality Spatial Simulation with Uncertainty Assessment

Spatial information of soil fertility quality is crucial for sustainable agriculture management. This study compared three methods based on sequential Gaussian simulation for spatial simulation of soil fertility quality index (SFQI) with uncertainty assessment through a case study in Hanchuan County, China. These methods are different in data utilization and simulation procedure: Method 1 first calculates SFQI sample data from the sample data of soil fertility indicators and then simulates the spatial distribution of SFQI with incorporation of land use information; Method 2 is similar to Method 1 except that land use information is not used; and Method 3 first simulates soil quality indicators using their sample data and then derives each SFQI realization from randomly grouped realizations of different soil quality indicators. Validation showed that Method 1 had the best performance in optimal prediction in terms of correlation coefficients between measured and predicted data, global prediction errors (mean error [ME], mean absolute error [MAEE], and root mean square error [RMSE]) and uncertainty modeling (accuracy plots and goodness statistic), and Method 3 was the worst. Methods 1 and 2 achieved, respectively, relative improvements of 32.27 and 12.72% in optimal prediction over Method 3. Probability maps estimated from simulated realizations using Method 1 showed that paddy fields generally had high quality grades while the fertility quality grades of dry farmlands were usually low in the study area. It is concluded that Method 1 may be a more effective method for SFQI prediction with uncertainty assessment and Method 2 may be adopted when land use information is unavailable.

Abbreviations: A-B, available boron; A-Cu, available copper; A-Fe, available iron; A-K, available potassium; A-Mn, available manganese; A-N, available nitrogen; A-P, available phosphorus; A-S, available sulfur; A-Zn, available zinc; CV, the coefficient of variation; IQI, integrated quality index; MAEE, mean absolute errors; ME, mean errors; r, correlation coefficient; RMSE, root mean square errors; RI, relative improvement; SF, scoring function; SFQI, soil fertility quality index; SGS, sequential Gaussian simulation; SOM, soil organic matter; TK, total potassium; TN, total nitrogen; TP, total phosphorus.

Soil is the basis of agro-ecosystems (Cao and Zhou, 2008). Better knowledge on the quality of a soil is important to sustainable land use management (McGrath and Zhang, 2003), early warning of adverse trends, and identification of problem areas (Bindraban et al., 2000). Since Doran and Parkin (1994) proposed the integrated soil quality index, this index has been often used for soil fertility quality assessment (Andrews et al., 2002b; Sun et al., 2003; Zhang et al., 2003; Sun et al., 2012a). However, the spatial heterogeneity of soil quality indicators implies that the SFQI, which is estimated based on the soil quality indicators, is also spatially heterogeneous, and consequently uncertainty in estimating the spatial distribution of SFQI is inevitable. Because the SFQI provides important information to soil or land managers about the general fertility quality or production potential of a soil, it is crucial to accurately predict the spatial distribution of
SFQI with uncertainty assessment at regional scales for sustainable agriculture and land management.

Geostatistical techniques have been often used in modeling the spatial variability of soil quality indicators and assessing their spatial uncertainty (Cambardella et al., 1994; Bourennane et al., 2007; Qu et al., 2012; Sun et al., 2012b). Although relatively rare, basic geostatistical interpolation techniques such as ordinary kriging have been used recently for mapping the spatial distribution of SFQI in a few of studies (see Sun et al., 2003; Zhang et al., 2004; Qi et al., 2009). However, the smoothing effect, commonly found in the maps generated by optimal spatial interpolations, results in less variation in estimated values than in observed values (Goovaerts, 1997; Juang et al., 2004). This problem causes low values to be overestimated and high values to be underestimated, thus impacting the delineation of low SFQI areas. Moreover, it is apparent that spatial uncertainty is an intrinsic characteristic in the estimation of the spatial distribution of soil quality indicators due to limited observation points, and such uncertainty consequently causes the spatial uncertainty in the spatial estimation of SFQI.

The spatial uncertainty information of SFQI should be valuable to decision making and risk cost estimation in soil management. However, stochastic simulation algorithms, which can provide a series of possible realizations of an unknown spatial distribution, do not aim to minimize local error variance (Zhao et al., 2009). Fluctuations between realizations provide a quantitative measure of the uncertainty about the underlying phenomenon. For these reasons, stochastic simulations are generally preferred to interpolations for applications in which the spatial variation of the measured field needs to be preserved and uncertainty assessment is required. In this study, conditional sequential simulation can provide not only the spatial distribution of SFQI, but also the uncertainty data of the simulation for quantitative evaluation.

Thus, it is more conducive to agricultural and land managers to make accurate judgments and scientific decision-making.

Studies in spatial uncertainty assessment of SFQI have been very rare in the literature so far. To the best of our knowledge, the only study is the one that was performed by Sun et al. (2012b). In that study, soil quality in a study area was evaluated based on the simulated maps of a series of soil quality indicators, and each simulated map of an indicator was randomly grouped with simulated maps of other indicators, one per indicator. Main drawbacks of this method include: (i) neglect of relations between soil indicators—for example, the simulated contents of soil total nitrogen (TN) may be larger than the simulated contents of soil organic matter (SOM) at some common positions, which is impractical in reality; (ii) heavy computation load and information loss—there are 500\(^{15}\) realizations of SFQI when 15 soil indicators are used and 500 realizations are generated for every soil indicator; however in that study, only 500 realizations of SFQI were randomly chosen for simplicity, which inevitably led to the loss of much information; and (iii) propagation of parameter errors from a large amount of simulations of a set of soil indicators. These drawbacks may be overcome or mitigated through first calculating SFQI at every sample location and then conducting geostatistical simulations directly on these SFQI data.

In the Jianghan Plain of Hubei Province, major crops are rice (Oryza sativa L.) and wheat (Triticum aestivum L.), which are cultivated under different land use types (i.e., paddy farmland and dry farmland, respectively). While wheat grows only in dry farmland, rice requires ample water for submerging paddy fields. These two different land use types usually cause differences in many soil properties such as the contents of SOM and N (Qu et al., 2012), and further affect the level of SFQI. Numerous studies have been performed in soil fertility quality changes induced by land-use alteration in recent decades (Wang and Gong, 1998; Mazzoncini et al., 2010; Xue et al., 2011). This means that incorporation of land use information into the spatial prediction and uncertainty assessment of SFQI would be desirable and also necessary.

In this study, we identified three different sequential Gaussian simulation (SGS) based methods for simulating the SFQI: Method 1 first calculates the SFQI data from the sample data of soil quality indicators using scoring functions (SFs) and the integrated quality index (IQI), and then simulates SFQI based on the calculated SFQI data and land use type information using the SGS-CI algorithm (i.e., SGS with categorical information being incorporated) (Qu et al., 2013). Method 2 is similar to Method 1 except that land use type information is not used. Method 3 first simulates the soil quality indicators based on their sample data using SGS, and then derives the SFQI realizations using SFs and IQI based on the simulated realization data of soil quality indicators. The main objectives of this study are: (i) to compare the three simulation methods of SFQI in terms of their performances in accounting for the spatial variability and uncertainty of SFQI, and (ii) to assess the associated spatial uncertainty in SFQI estimation of the topsoil in the study area. The ultimate goal is to find a more effective method for predictive mapping of SFQI and the associated spatial uncertainty, which can be valuable to precision farming and environmental management.

**MATERIALS AND METHODS**

**Study Area and Data**

The study was conducted in HanChuan County, an agricultural region in central China. The study area is located in Jianghan Plain, central Hubei Province, bounded by the longitudes of 113°2’ and 113°57’ E, and the latitudes of 30°22’ and 30°51’ N, with an area of 1659 km\(^2\). Within the northern subtropical monsoonal climate zone, it has a temperate-humid climate throughout a year and four distinct seasons, with an average annual temperature of 16.1°C and an average annual precipitation of approximately 1198 mm.

The land use data was collected from the county’s local agricultural department. In this region, the major land use types are paddy field and dry farmland. We further edited the land use map and simply classified the land use into four types (or categories): paddy field, dry farmland, water body, and others (i.e., other land use types). The spatial distribution of the land use types is shown in Fig. 1. Paddy fields (i.e., rice farmlands) and dry
farmlands use very different cultivation forms, which usually have strong impacts on soil properties that we use to evaluate the soil fertility quality in the study area. The others land use type refers to all those places without tillage, such as wood lands, wild lands, and built-up areas (villages and towns). Water bodies include rivers, lakes, and ponds. This land use type was first separated as a GIS layer and then overlaid to the simulated maps of soil quality indicators or SFQI, so it was not considered in the geostatistical simulation processes by different simulation methods. Paddy fields are the dominant land use type of the county.

In this study, the topsoil sampling points consist of prediction points \( n = 402 \) and validation points \( n = 135 \). Among the 402 prediction samples, 215 samples were taken from paddy fields, 130 samples were taken from dry farmlands, and the remaining samples were taken from other land use types (Table 1). The validation samples cover all of the three land use groups presented in the prediction samples, although they were collected with the consideration of randomness and homogeneity in the area. Sample locations were recorded using a hand-held global position system. All samples were taken in fall after harvest and before next cropping season to avoid the effect of fertilization during crop cultivation. When sampling, soils in top layers (0–15 cm) of six to eight points at each site within an area of approximately 0.01 ha were collected, and then mixed. A portion of 1 to 2 kg for each sample was delivered to a laboratory for analysis. All samples were air-dried at room temperature (20–22°C). After stones or other debris were removed, samples were then sieved to ensure the sizes of soil particles to be smaller than 2 mm. A portion of about 100 g for each sample was ground in an agate grinder and sieved through a 0.149-mm mesh.

All of the soil samples were analyzed for a series of soil properties including: pH (using the glass electrode method) (Lu, 2000); SOM (potassium dichromate colorimetry) (Lu, 2000); TN (Kjeldahl) (Bremmer and Mulvaney, 1982); available nitrogen (A-N) (alkaline potassium dichromate colorimetry) (Lu, 2000); total phosphorus (TP) (acid digestion) (Bremmer and Mulvaney, 1982); available phosphorus (A-P) (HCl-H\(_2\)SO\(_4\) extraction and colorimetry) (Lu, 2000); available sulfur (A-S) (calcium chloride extraction) (Lu, 2000); available potassium (A-K) (ammonium acetate extraction and colorimetry) (Lu, 2000); available potassium (A-Fe, A-Mn) (diethylentriaminepentaacetic acid and atomic absorption spectroscopy) (Lindsay and Norvell, 1978). Quality control was based on the use of certified samples (GBW 07413) and analysis duplicates. Statistical summary of these soil properties is shown in Table 1.

### Soil Fertility Quality Evaluation

In this study, soil fertility quality was evaluated based on the method of the total indicator set and the integrated quality index (IQI), which was evaluated as the most accurate qualitative soil quality evaluation method by Qi et al. (2009).

The selected indicators were scored and weighted following the approach adopted in Sun et al. (2003), Zhang et al. (2004), and Qi et al. (2009). In this approach, two standard scoring functions (SFs) were used to obtain the upper and peak limits, respectively (Table 2): (i) The upper limit function is:

\[
f(x) = \begin{cases} 
0.1, & x \leq L \\
0.9 \times \frac{x - L}{U - L} + 0.1, & L < x \leq U \\
1, & x > U
\end{cases}
\]

and (ii) the peak limit function is:

\[
f(x) = \begin{cases} 
0.1, & x \leq L, \text{ or } x > U_1 \\
0.9 \times \frac{x - L_1}{U_2 - L_1} + 0.1, & L_1 < x \leq L_2 \\
1, & L_2 < x \leq U_1 \\
0.9 \times \frac{U_1 - x}{U_1 - U_2} + 0.1, & U_2 < x \leq U_1
\end{cases}
\]

### Table 1. Statistical summary of soil properties from calibration and validation samples.

<table>
<thead>
<tr>
<th>Soil property</th>
<th>Calibration samples ( n = 402 )</th>
<th>Validation samples ( n = 135 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>Min.</td>
<td>Max.</td>
</tr>
<tr>
<td>SOM, g kg(^{-1})</td>
<td>2.84</td>
<td>57.90</td>
</tr>
<tr>
<td>TN, g kg(^{-1})</td>
<td>0.31</td>
<td>3.04</td>
</tr>
<tr>
<td>A-N, mg kg(^{-1})</td>
<td>32.00</td>
<td>299.00</td>
</tr>
<tr>
<td>TP, mg kg(^{-1})</td>
<td>0.11</td>
<td>1.84</td>
</tr>
<tr>
<td>A-P, mg kg(^{-1})</td>
<td>1.85</td>
<td>93.86</td>
</tr>
<tr>
<td>TK, mg kg(^{-1})</td>
<td>8.71</td>
<td>86.50</td>
</tr>
<tr>
<td>A-S, mg kg(^{-1})</td>
<td>26.10</td>
<td>485.00</td>
</tr>
<tr>
<td>A-S, g kg(^{-1})</td>
<td>16.57</td>
<td>285.10</td>
</tr>
<tr>
<td>A-B, g kg(^{-1})</td>
<td>0.06</td>
<td>1.35</td>
</tr>
<tr>
<td>A-Fe, mg kg(^{-1})</td>
<td>0.12</td>
<td>24.20</td>
</tr>
<tr>
<td>A-Mn, mg kg(^{-1})</td>
<td>0.29</td>
<td>14.33</td>
</tr>
<tr>
<td>A-P, mg kg(^{-1})</td>
<td>6.40</td>
<td>282.01</td>
</tr>
<tr>
<td>A-Mn, mg kg(^{-1})</td>
<td>4.44</td>
<td>127.30</td>
</tr>
</tbody>
</table>

Table 2. Soil quality indicators and their scoring functions, communalities, and weights.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Function type</th>
<th>Lower threshold</th>
<th>Upper threshold</th>
<th>Communalily</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>Peak limit</td>
<td>4.5, 5.5</td>
<td>7.5, 9.5</td>
<td>0.7373</td>
<td>0.0725</td>
</tr>
<tr>
<td>SOM†, g kg⁻¹</td>
<td>Upper limit</td>
<td>15.00</td>
<td>30.00</td>
<td>0.8531</td>
<td>0.0839</td>
</tr>
<tr>
<td>TN, g kg⁻¹</td>
<td>Upper limit</td>
<td>0.75</td>
<td>1.50</td>
<td>0.8778</td>
<td>0.0863</td>
</tr>
<tr>
<td>A-N, mg kg⁻¹</td>
<td>Upper limit</td>
<td>60.00</td>
<td>100.00</td>
<td>0.8328</td>
<td>0.0819</td>
</tr>
<tr>
<td>TP, mg kg⁻¹</td>
<td>Upper limit</td>
<td>0.40</td>
<td>1.00</td>
<td>0.6505</td>
<td>0.0640</td>
</tr>
<tr>
<td>A-P, mg kg⁻¹</td>
<td>Upper limit</td>
<td>5.00</td>
<td>15.00</td>
<td>0.8031</td>
<td>0.0790</td>
</tr>
<tr>
<td>TK, mg kg⁻¹</td>
<td>Upper limit</td>
<td>5.00</td>
<td>25.00</td>
<td>0.7037</td>
<td>0.0692</td>
</tr>
<tr>
<td>A-K, mg kg⁻¹</td>
<td>Upper limit</td>
<td>50.00</td>
<td>120.00</td>
<td>0.6894</td>
<td>0.0678</td>
</tr>
<tr>
<td>A-S, mg kg⁻¹</td>
<td>Upper limit</td>
<td>15.00</td>
<td>50.00</td>
<td>0.7430</td>
<td>0.0722</td>
</tr>
<tr>
<td>A-B, mg kg⁻¹</td>
<td>Upper limit</td>
<td>0.15</td>
<td>0.30</td>
<td>0.4936</td>
<td>0.0432</td>
</tr>
<tr>
<td>A-Cu, mg kg⁻¹</td>
<td>Upper limit</td>
<td>0.20</td>
<td>1.80</td>
<td>0.6765</td>
<td>0.0665</td>
</tr>
<tr>
<td>A-Zn, mg kg⁻¹</td>
<td>Upper limit</td>
<td>0.50</td>
<td>1.50</td>
<td>0.6895</td>
<td>0.0678</td>
</tr>
<tr>
<td>A-Fe, mg kg⁻¹</td>
<td>Upper limit</td>
<td>2.00</td>
<td>32.00</td>
<td>0.7229</td>
<td>0.0711</td>
</tr>
<tr>
<td>A-Mn, mg kg⁻¹</td>
<td>Upper limit</td>
<td>10.00</td>
<td>20.00</td>
<td>0.7608</td>
<td>0.0748</td>
</tr>
</tbody>
</table>


where \( f \) is the score, \( x \) is the soil quality indicator, \( L \) and \( U \) are the lower and upper threshold values, respectively. Soil quality tends to be standardized both nationally and internationally (Qi et al., 2009), therefore, the values of \( L \) and \( U \) in Table 2 were determined according to many other studies, for example, Sun et al. (2003), Zhang et al. (2004), and Qi et al. (2009), as well as local standards (ISMAPRC, 1996) and farming experiences of local farmers. The weight of each indicator (Table 2) was designated according to the indicator’s communality, which was derived from the standardized factor analysis on all indicator values of the calibration samples (Shukla et al., 2006; Cao and Zhou, 2008; Qi et al., 2009).

Based on the scores and weights obtained as explained above, SFQI was calculated at each site using the IQI method:

\[
SFQI = \sum_{i=1}^{n} W_i N_i
\]  

[3]

where \( W_i \) is the assigned weight of indicator \( i \), \( N_i \) is the indicator score of indicator \( i \), and \( n \) is the number of indicators.

The values of SFQI were divided into four grades according to their classification criteria (Table 3). The criteria were based on the classifications of region types and the fertility of cultivated lands in China (ISMAPRC, 1996). Grade I is considered to be the most suitable for plant growth, Grade II is suitable for plant growth but with some limitations, Grade III has severer limitations than Grade II, and for Grade IV the soil has the severest limitations for plant growth.

### Stochastic Simulation Methods Being Tested

Sequential Gaussian simulation is the most frequently used sequential simulation algorithm for simulating continuous variables. A detailed introduction on this method can be found in Goovaerts (1997). After a regularly spaced grid covering the study area is defined, SGS involves the following steps:

1. Transform the sample data of SFQI into standard normal data using the normal score transformation.

2. Compute and model the experimental variogram of the normal score transformed data.

3. Establish a random path through all of the grid nodes, in a way that each node is visited only once in each sequence.

4. At each node \( x_0 \):
   a. Estimate the parameters (mean and variance) of the Gaussian conditional cumulative distribution function (ccdf) of SFQI by simple kriging estimator with the normal score variogram model. The conditioning data includes both the original data and already simulated data within the defined neighborhood of the location to be simulated.
   b. Draw a simulated normal score value from the estimated ccdf and then add it to the conditional data set to be used for simulating other nodes.
   c. Proceed to the next grid node along a random path and repeat steps (a) and (b) until the entire set of grid nodes are simulated.
   d. Back-transform the simulated normal score values into the SFQI values in the original data space.

These sequential steps build up only the first realization, \( \{z^{(1)}(x_0), j = 1, ..., M\} \), which is only one model of the SFQI spatial distribution. To generate multiple, say \( K \), realizations, \( \{z^{(k)}(x_0), j = 1, ..., M, k = 1, ..., K\} \), Steps 3 to 5 should be repeated with different random paths passing through all nodes. In this study, geostatistical simulation used a grid having a pixel resolution of 400 × 400 m.

### Methods 1 and 2

Liu et al. (2006) suggested a kriging combined with soil map-delineation (KSMO) method for incorporating the effect of soil types in the interpolation of soil properties. The method separated an observation at a location into two components—the mean value over the soil type polygon and the residual, similar to Goovaerts and Journel (1995). The residual data were further interpolated using the ordinary kriging, rather than using the indicator kriging as done in Goovaerts and Journel (1995). Thus, the final estimate at an unsampled location was the sum of the mean value and the interpolated residual value. However, similar to other earlier studies, the KSMO method did not perform stochastic simulations. Recently, Qu et al. (2013) suggested a SGS-CI method, extending the KSMO method into a sequential simulation method based on the SGS with incorporation of related categorical information.

Table 3. Criteria for soil quality grade division.

<table>
<thead>
<tr>
<th>Grade</th>
<th>SFQI†</th>
<th>Grade</th>
<th>SFQI†</th>
<th>Grade</th>
<th>SFQI†</th>
<th>Grade</th>
<th>SFQI†</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>≥ 0.85</td>
<td>II</td>
<td>0.85 &gt; SFQI ≥ 0.7</td>
<td>III</td>
<td>0.7 &gt; SFQI ≥ 0.5</td>
<td>IV</td>
<td>SFQI &lt; 0.5</td>
</tr>
</tbody>
</table>

† SFQI, soil fertility quality index.
The spatial variability of SFQI data calculated from the sample data of soil fertility indicators using SFs and IQI is apparently partially caused by the complex distribution of land use types in the study region, which increases the uncertainty in SFQI prediction. To reduce this uncertainty, Method 1 divides the SFQI \( z(x_k) \) of each sample into two portions: the mean value \( u(x_k) \) under the land use type \( s_k \) to which the sample belongs, and the corresponding residual \( r(x_k) \), that is,

\[
z(x_k) = u(x_k) + r(x_k)
\]

where \( x_{kj} \) is the location of the sample \( z(x_{kj}) \). Then, the residual \( r(x_{kj}) \) can be treated as a new stationary regionalized variable and be simulated using SGS. A final simulated realization is the sum of a simulated residual realization and the mean values at all locations. Thus, a number of simulated realizations incorporating land use effect can be generated and the associated uncertainty with the optimal prediction (i.e., the E-type estimate) can be assessed.

Method 2 is similar to Method 1 except that the land use type information is not incorporated. That means that in Method 2, the SGS algorithm is directly applied to the SFQI data.

**Method 3**

In Method 3, soil quality indicators are first simulated based on soil quality indicator sample data using SGS with a number of realizations being produced for each soil indicator, and then each simulated realization of an indicator is randomly grouped with soils and the IQI method. A final simulated realization is the sum of a simulated residual realization and the mean values at all locations. Thus, a number of simulated realizations incorporating land use effect can be generated and the associated uncertainty with the optimal prediction (i.e., the E-type estimate) can be assessed.

Method 2 is similar to Method 1 except that the land use type information is not incorporated. That means that in Method 2, the SGS algorithm is directly applied to the SFQI data.

**Evaluation Criteria**

To evaluate the prediction performance, the Pearson’s correlation coefficients \( r \) between the 135 measured validation samples and the optimal prediction maps of Methods 1, 2, and 3 were computed. Mean errors, MAEE, RMSE, and relative improvement (RI) for results from the three methods were also computed. Here the optimal prediction maps refer to the “E-type” estimates, which are the point-by-point averages of the simulated realizations. The ME, MAEE, and RMSE can be calculated using the following equations

\[
ME = \frac{1}{N_v} \sum_{i=1}^{N_v} [z(x_i) - z^*(x_i)]
\]

\[
MAEE = \frac{1}{N_v} \sum_{i=1}^{N_v} [z(x_i) - z^*(x_i)]
\]

\[
RMSE = \sqrt{\frac{1}{N_v} \sum_{i=1}^{N_v} [z(x_i) - z^*(x_i)]^2}
\]

respectively. In above equations, \( N_v \) is the number of validation points, and \( z(x_i) \) and \( z^*(x_i) \) are the measured and predicted values at the validation point \( i \), respectively. The ME criterion is used to check the conditional bias, which consists in the assumption that prediction errors canceling out lead to an unbiased estimator over the whole range of values. The MAEE is a quantity used to measure how close predictions are to the eventual outcomes; it avoids the fact that the negative and positive biases tend to cancel each other. The RMSE is a quadratic scoring rule which measures the differences between values predicted by a model or an estimator and the values actually observed. These values should approach to zero for an optimal prediction. Greater \( r \) and lower ME, MAEE, and RMSE values indicate higher prediction accuracy.

The RI (\%) of a method over the other method is calculated using

\[
RI = 100 \times \frac{\text{RMSE}_R - \text{RMSE}_E}{\text{RMSE}_R}
\]

where \( \text{RMSE}_R \) and \( \text{RMSE}_E \) are the root mean square errors for the reference and evaluated methods, respectively. If RI is positive, the accuracy of the evaluated method is higher than that of the reference method, and vice versa (Zhang et al., 1992).

To evaluate the simulation performance, the accuracy and goodness of the reproduction of single-point conditional probability distributions by a set of realizations were examined. At any location \( x \), knowledge of the cdf \( F(x, z(a)) \) allows the computation of a series of symmetric \( p \)-probability intervals (PI) bounded by the \((1-p)/2 \) and \((1+p)/2 \) quantiles of the cdf. According to Goovaerts (2001), a probability distribution is considered accurate if the fraction of values of a validation dataset falling in the symmetric \( p \)-PI interval exceeds \( p \) for all \( p \in [0, 1] \). Considering a validation data set, \( z(x) \), the fraction of true values falling into a given symmetric \( p \)-PI interval can be computed as

\[
\xi(p) = \frac{1}{N_v} \sum_{j=1}^{N_v} \xi(x_j; p) \quad \forall p \in [0, 1]
\]

with

\[
\xi(x_j; p) = \begin{cases} 1, & \text{if } F^{-1}(x_j; (1-p)/2) \leq z(x_j) \leq F^{-1}(x_j; (1+p)/2) \\ 0, & \text{otherwise} \end{cases}
\]

The scattergram of the calculated fractions vs. the set of probabilities \( p \) is called “accuracy plot” (Goovaerts, 2001). The accuracy of a probabilistic model can be visually assessed through an accuracy plot. For an accurate case, most of the points must fall above the 45° line, that is, \( \xi(p) > p \) for most \( p \). Deutsch (1997) proposed to assess the closeness of the estimated and theoretical fractions using the following “goodness” statistics \( G \in [0, 1] \) as

\[
G = 1 - \int_0^1 \left[ 3a(p) \cdot 2 \right] \left[ \xi(p) - p \right] dp ,
\]

where the indicator function \( a(p) \) is defined as

\[
a(p) = \begin{cases} 1, & \text{if } \xi(p) \geq p \\ 0, & \text{otherwise} \end{cases}
\]
As shown in Eq. [11], twice more importance is given to deviations when \( T(p) < p \) (inaccurate case). The weight \( 3a(p) - 2 \) instead of 1 for the accurate case, that is, the case where the fraction of true values falling into the \( p \)-PI is larger than the expected. A greater \( G \) value indicates less prediction uncertainty.

### Uncertainty Evaluation

In this study, the local spatial uncertainty of SFQI was assessed based on the response realizations from the method which has the highest accuracy. The uncertainty of a soil fertility quality grade at a specific location \( x \) can be defined as the probability that unknown SFQI \( z(x) \) at location \( x \) falls within a given specific critical threshold interval. These probabilities can be calculated using the following equations:

\[
P\left[c \leq E_i(x) < c_e\right] = \frac{n(x)}{\text{T}}. \tag{13}
\]

where \( T \) is the total number of response realizations of the SFQI, which is 500 in this study, and \( n(x) \) is the number of response realizations whose SFQI response values at the location \( x \) are within the interval defined by a pair of thresholds with low value \( c \) and high value \( c_e \). Specific thresholds are listed in Table 3.

### RESULTS AND DISCUSSION

#### Descriptive Statistical Analysis

Due to the limitation of the soil sample data set in soil indicators, we chose 14 soil indicators (i.e., pH, SOM, total and available N, P, and K, available S, B, Cu, Zn, Fe, and Mn), which are mainly soil nutrient and nutrient-related properties, for estimating the SFQI. The descriptive statistics for SFQI data calculated from 402 calibration samples are given in Table 4. The mean SFQI of all samples is 0.7583, and the coefficient of variation (CV) is 18.44%. Samples were classified into three groups based on the land use types of their locations, of which the average SFQI values in descending order are 0.8148 for paddy fields, 0.7367 for other land use type (i.e., non-arable lands), and 0.6741 for dry farmlands. This indicates that land use types do affect soil quality. Such a result is consistent with previous studies (Wang and Gong, 1998; Mazzoncini et al., 2010; Xue et al., 2011). In general, the paddy fields have higher SFQI values, which may be attributed to more fertilizer input and organic matter accumulation under the persistent wet condition, whereas the dry farmlands have lower SFQI values, probably resulting from less fertilizer input and fast organic matter mineralization. The CVs with the three land use types are all relatively small.

The apparent correlation between the SFQI values of soil samples and land use types implies that categorical land use data should be valuable auxiliary information to the stochastic spatial simulation of SFQI and incorporating them may improve spatial prediction accuracy and reduce the prediction uncertainty.

#### Spatial Analysis of Soil Fertility Quality Index

Because no apparent anisotropy was found from the sample data, experimental variograms were estimated omnidirectionally for the normal score transformed data of each related variable at sample locations. These variables include 14 soil fertility indicators, SFQI and its residual data. The experimental variograms of TN, A-Zn, and A-Fe were fitted by spherical models and the others were all fitted by exponential models. Variogram model parameters are presented in Table 5. The \( C_0/(C_0+C) \) ratio may be used as a criterion for describing the spatial self-dependency of a variable (Cambardella et al., 1994). The \( C_0/(C_0+C) \) ratios of the variogram models for soil TP, A-P, and A-S are all below 25%, exhibiting strong spatial self-dependency which may be attributed to the intrinsic factors such as soil types. Other soil quality indicators exhibiting moderate spatial self-dependency (i.e., their variogram model \( C_0/(C_0+C) \) ratios are between 25 and 75%), which may be attributed to both the intrinsic factors...
such as other soil properties and the extrinsic factors such as human activities (e.g., land use).

Five hundred realizations of SFQI were simulated using each of the three methods, that is, Methods 1, 2, and 3. We selected one realization from each simulation and displayed them in Fig. 2. These realizations may represent realistic spatial distributions of SFQI without the smoothing effect under different knowledge and information input conditions. The difference between these realization maps generated respectively by Methods 1, 2, and 3 is visually not obvious, probably due to the complexity of land use polygons (see Fig. 1) and the moderate effect of land use types on SFQI in this case study (see Table 4). However, obvious differences can be found between the optimal prediction maps (i.e., E-type estimates) generated respectively by the three different methods (Fig. 3): (i) The optimal prediction map generated by Method 1 looks much rougher in texture compared with those generated by the other two methods due to its incorporation of the influence of different types of land use patches, and it shows the lower and higher contents more accurately as expected on different land use types. (ii) While both Methods 2 and 3 do not incorporate the effect of land use types, the optimal prediction map generated by Method 2 is relatively smoother than that generated by Method 3; the latter shows some abrupt changes, which do not coincide with the boundaries of land use patches. The reason should be that Method 2 directly simulated the spatial distribution of SFQI estimated from the sample data of soil fertility indicators, whereas Method 3 indirectly generated the SFQI optimal prediction map with the drawbacks mentioned in the Introduction section. Despite the differences, these optimal prediction maps show similar trends, with higher SFQI values mainly appearing in the Northwest region and lower values mostly occurring in the mid-South region of the county.

**Prediction Accuracy Analysis**

The results of the validation indices for all the methods, which are estimated using the validation sample data and the optimal prediction maps, are presented in Table 6. A larger $r$ value and smaller ME, MAEE, and RMSE values mean that a method generates a...
more accurate prediction map. The correlation coefficients (i.e., $r$ values) are 0.89, 0.79, and 0.71 for the optimal prediction maps generated by Methods 1, 2, and 3, respectively, which indicates a method performance sequence of Method 1 > Method 2 > Method 3. The increasing sequences of ME, MAEE, and RMSE values for the three methods are all Method 3 > Method 2 > Method 1; this again indicates that Method 1 is better than Method 2 whereas Method 2 is better than Method 3 in global prediction accuracies of SFQI. Methods 1 and 2 have relative improvements of 32.27 and 12.72% over Method 3, respectively. Method 3 is apparently the least accurate estimator here.

Accuracy plots of simulated results using the three methods were computed using Eq. [9] and depicted in Fig. 4. Compared with those of Methods 2 and 3, the accuracy plot of Method 1 has more points falling above the 45° line. This implies that Method 1 generated a higher accuracy in modeling uncertainty than the other two methods. When $p$ values are large ($p = 0.80~0.90$), Method 1 simulation also has a few of points below the 45° line, indicating that it has lower accuracy at some high probability intervals. Smaller (i.e., poorer) $G$ values are thus obtained for Methods 2 and 3 compared with that for Method 1. Again this reveals that the performance of Method 1 is the best whereas that of Method 3 is the worst for SFQI simulation in reducing estimation uncertainty. When no land use information is incorporated, Method 2 is a more accurate estimator here compared to Method 3. Main drawbacks of Method 3 were discussed above in the Introduction section and it is apparent that Method 2 overcomes these drawbacks.

### Evaluation of Soil Fertility Quality

The spatial distribution of corresponding soil fertility quality grades based on the optical prediction map of Method 1 is presented in Fig. 5. High soil fertility quality areas (Grades I and II) scatter around mainly in the paddy fields while low soil fertility quality areas (Grades III and IV) mainly appears in the middle of the study area along a river where dry farmlands dominate (also see Fig. 1). In fact, the lowest soil fertility grade (i.e., Grade IV) almost does not occur in Fig. 5 except for at some sampling locations. Probability maps of SFQI for different soil fertility quality grades were estimated from simulated realizations of SFQI using Method 1 and are presented in Fig. 6. One can see that Grade I has larger probability values in the paddy fields, while Grade III has larger probability values in dry farmlands. Grade II may occur almost everywhere, but Grade IV only has some low possibilities to occur in dry farmlands. This means that the fertility quality of paddy fields is generally higher than that of dry farmlands and that very poor quality lands in the country are rare, according to the SFQI evaluated in this study.

### Some Relevant Issues

We would like to point out that this study only aimed to compare the performances of the three different methods so that a relative better method could be recommended for simulating the spatial variability of SFQI and assessing associated uncertainty. Thus, many other issues concerning the estimation and use of SFQI were not addressed. The SFQI is a typical integrative index for representing the general fertility quality or land potential of a soil. Such an index may provide useful information for sustainable land use management. However, this index does not reflect the deficiency or sufficiency of any single nutrient. Thus, it may not be proper to use SFQI to direct the application of any fertilizer or to predict crop yield. Fertilization is an agricultural measure which is usually specific to the temporal status of a crop and a single nutrient in soil in a piece of farmland. Crop yield is impacted by many complex soil and non-soil factors, particularly field management.

### Table 6. Validation indices† of soil fertility quality index (SFQI) predicted using different methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>$r$</th>
<th>ME</th>
<th>MAEE</th>
<th>RMSE</th>
<th>RI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method 1</td>
<td>0.89</td>
<td>0.0083</td>
<td>0.0341</td>
<td>0.0703</td>
<td>32.27</td>
</tr>
<tr>
<td>Method 2</td>
<td>0.79</td>
<td>0.0252</td>
<td>0.0403</td>
<td>0.0906</td>
<td>12.72</td>
</tr>
<tr>
<td>Method 3</td>
<td>0.71</td>
<td>0.0361</td>
<td>0.0432</td>
<td>0.1038</td>
<td></td>
</tr>
</tbody>
</table>

† $r =$ correlation coefficient, ME = mean error; MAEE = mean absolute error, RMSE = root mean square error, RI = relative improvement.

Fig. 4. Accuracy plots and $G$ statistics for simulated results by (a) Method 1, (b) Method 2, and (c) Method 3.
measures. These cannot be quantitatively accounted for by an integrative index like SFQI.

The estimation of SFQI in this study followed the methodology established in earlier studies in soil science (Karlen and Scott, 1994; Andrews et al., 2002b; Qi et al., 2009; Sun et al., 2012b) and used some commonly used soil properties. The 14 soil indicators considered in this study are mainly soil chemical properties. These soil indicators may not be sufficient to reflect the integrative soil fertility quality due to the absence of related soil physical properties. However, this does not impact the conclusions made on the performances of the spatial statistical methods used for simulating the spatial variability of SFQI. Currently, there is still few standards on how the SFQI should be calculated and what soil indicators should be incorporated into SFQI. Different soil indicators have been used in earlier studies (e.g., Qi et al., 2009; Sun et al., 2012b). Special soil functions of interest and the defined management goals for the system should be considered in choosing soil indicators for SFQI calculation (Andrews et al., 2002a). Nevertheless, further studies are necessary to address these issues and explore some potentially better SFQI estimation methods, such as the scoring function used for soil health evaluation which gives the highest score to the lowest value and vice versa (Gugino et al., 2009).

CONCLUSIONS

The performances of three methods in simulating the spatial distribution of SFQI for uncertainty assessment were explored. The differences between these three methods are that Method 1 incorporates the effect of land use types while Method 2 does not, and that Method 3 indirectly simulates SFQI by deriving SFQI realizations from simulated realizations of soil quality indicators. It was found that Method 1 has the best performance and Method 3 is the worst for SFQI simulation in reducing estimation uncertainty. While the better performance of Method 1 over Method 2 reflects the effect of land use types on SFQI, the worst performance of Method 3 implies that the indirect SFQI simulation method has information loss or propagates larger errors in addition to heavy computation load.

Soil fertility quality grades were evaluated based on a number of simulated realization maps by Method 1. Results show that paddy fields generally have high soil fertility quality (Grades I and II), and most dry farmlands fall into the low soil fertility quality grades (Grades III and IV) although Grade IV rarely occurs. We can conclude that Method 1 may be a more effective method for SFQI prediction with uncertainty assessment, and when land use (or soil type) data is unavailable Method 2 is more appropriate to use than Method 3.

ACKNOWLEDGMENTS

This work was supported by the Knowledge Innovation Program (KZCX2-EW-QN404) and Strategic Priority Research Program (XDA05050503) of the Chinese Academy of Sciences, the National Natural Science Foundation of China (41071299, 41101491), the Special Research Foundation of the Public Natural Resource Management Department from Ministry of Environmental Protection of China (201109018), and the China Postdoctoral Science Foundation Funded Project (2013M530273).