

Article

Sampling and Kriging Spatial Means: Efficiency and Conditions

Jin-Feng Wang ^{1,*}, Lian-Fa Li ¹ and George Christakos ²

¹ Institute of Geographic Sciences & Nature Resources Research, Chinese Academy of Sciences, Beijing, China; E-Mail: lilf@lreis.ac.cn

² Department of Geography, San Diego State University, San Diego, CA, USA;
E-Mail: gchrista@mail.sdsu.edu

* Author to whom correspondence should be addressed; E-Mail: wangjf@lreis.ac.cn; Tel.: +86-10-64888965; Fax: +86-10-64889630

Received: 11 May 2009; in revised form: 16 June 2009/ Accepted: 29 June 2009 /

Published: 2 July 2009

Abstract: Sampling and estimation of geographical attributes that vary across space (e.g., area temperature, urban pollution level, provincial cultivated land, regional population mortality and state agricultural production) are common yet important constituents of many real-world applications. Spatial attribute estimation and the associated accuracy depend on the available sampling design and statistical inference modelling. In the present work, our concern is areal attribute estimation, in which the spatial sampling and Kriging means are compared in terms of mean values, variances of mean values, comparative efficiencies and underlying conditions. Both the theoretical analysis and the empirical study show that the mean Kriging technique outperforms other commonly-used techniques. Estimation techniques that account for spatial correlation (dependence) are more efficient than those that do not, whereas the comparative efficiencies of the various methods change with surface features. The mean Kriging technique can be applied to other spatially distributed attributes, as well.

Keywords: random field; mean Kriging; spatial dependence; GIS

1. Introduction

Spatial estimation techniques have many applications in the study of attributes, such as soil and land cultivation properties, water resource parameters, air pollution variables, population disease characteristics, regional poverty levels and agricultural production indices [1-3]. In addition, the assessment of the uncertainty associated with the generated estimates is as important as attribute estimation itself. E.g., if the accuracy of an attribute estimate is low (i.e., the uncertainty is high), the estimate is rather useless or even misleading. If, on the other hand, the accuracy is high, the estimated attribute value could be used in decision-making, such as the international negotiations of carbon emission reduction to address the global warming challenge.

In the GIS context, there are two main methodologies of area mean attribute estimation:

- a. One general methodology focuses on *spatial statistics*-based estimates using a set of observations/measurements across space [4-7]. This methodology includes objective analysis-, superpopulation model-, sampling- and design-based techniques [7-12].
- b. Another major methodology relies on *physical mechanism*-based estimates in addition to the datasets available [13-17]. This methodology includes Bayesian maximum entropy (BME) techniques [18-21], variational assimilation and Kalman techniques [22-26].

In theory, methodology *b* is superior to methodology *a* since, in addition to the available datasets, it can offer a more physically meaningful and informative analysis of the phenomenon of interest by accounting for valuable knowledge in the form of scientific theories, physical laws and primitive equations [27-29]. In practice, however, this kind of knowledge is often not available (or, if available, the computational procedures to account for it do not yet exist or are of limited use), in which case the efficiency of the techniques belonging to methodology *a* proves to be very useful. It is for this reason that a spatial statistics-based technique belonging to methodology *a* is considered in this work.

Spatial statistics-based techniques seek to account for uncertainty caused by gaps between the sampled sites [30]. A simple sample mean is an unbiased estimate of both the observable population and the superpopulation means, under the conditions of a second-order stationary object surface and a randomly distributed sample over space [31,32], but the variance of the estimate is not minimized. Spatial sampling techniques improve the efficiency of sampling and estimation by taking spatial correlation (dependence) into account [33], but that does not always guarantee that the estimation variance is minimized. Kriging leads to an unbiased estimate for unsampled values with the least variance [7], but the estimation of the mean attribute in terms of a summation of individual estimates at unsampled sites may also accumulate the errors of each individual estimate. Kriging the attribute mean across space yields an estimation of the area mean that is unbiased and has the minimum estimation variance. The technique has already existed in the literature for several decades – Kriging was originally developed in the context of Wiener-Kolmogorov estimation and objective analysis [6,10,34-36]. In the present work, our concern is twofold: the estimation of the spatial attribute mean over a specified area using the mean Kriging technique, and the study of the probability distribution of these estimates over the area of interest. Practical insight is gained in terms of a temperature dataset and a land use dataset distributed in space, in which the mean Kriging analysis is

compared with previous techniques, such as ordinary Kriging, spatial random sampling and simple random sampling techniques.

2. Spatial Random Field Representation of Attributes and Their Means

Let a geographical attribute be represented mathematically by the *spatial random field* (SRF), $Y(s)$ in the sense of Christakos [34]. The s denotes the spatial coordinates of location s and the SRF includes a family of spatially correlated (geographically dependent) random variables y_1, \dots, y_n at sample points s_1, \dots, s_n . A number of concepts of GIS interest can be defined in the SRF context, see below.

The *observed spatial population mean* (OSPM) over an area \mathfrak{R} of the attribute represented by the SRF $Y(s)$, also called the *observed area mean*, is defined as:

$$\bar{Y}_{\mathfrak{R}} = \frac{1}{|\mathfrak{R}|} \int_{\mathfrak{R}} Y(s) ds \quad (1)$$

where s varies within \mathfrak{R} . The $\bar{Y}_{\mathfrak{R}}$ is a random quantity, i.e., even when considering the same area \mathfrak{R} , one may get different results if the $\bar{Y}_{\mathfrak{R}}$ is computed over different realizations.

In the GIS context, the *superpopulation mean* (SPM) of the SRF at each location s , also called the *stochastic mean*, is defined as:

$$m(s) = E[Y(s)] = \int ds \psi(s) f_Y(s) \quad (2)$$

where $E[\cdot]$ denotes stochastic expectation, the $f_Y(s)$ is the probability density function (pdf) of the SRF $Y(s)$ and $\psi(s)$ is the SRF realization at s . The $m(s)$ is the average value of all SRF realizations at each s and is a non-random quantity. Note that it has to a single value m for all locations s , as long as the SRF is 1st-order stationary, i.e., $E[Y(s)] = \text{const.}$ for all s .

The *simple sample mean* (SSM) is defined as:

$$\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n y_i \quad (3)$$

where y_i are the corresponding random variables at locations s_i ($i = 1, \dots, n$) within the study area \mathfrak{R} . The \bar{Y}_n is a random quantity, since the random variables y_i can assume various values (realizations) and the n sample units can be drawn randomly across space. Equation (3) would be the best linear unbiased estimate of both the observable population mean and the superpopulation mean if the s_i ($i = 1, \dots, n$) are randomly distributed over space and the corresponding SRF is 1st-order spatial stationary; i.e., $E[\bar{Y}_n] = \frac{1}{n} \sum_{i=1}^n E[y_i] = \frac{1}{n} nm = m$.

The *weighted sample mean* (WSM) is defined as:

$$\bar{Y}_n^w = \sum_{i=1}^n w_i y_i \quad (4)$$

where w_i are weights assigned to the random variables y_i ($i = 1, \dots, n$). Again, \bar{Y}_n^w is a random quantity. Clearly, \bar{Y}_n is a special case of \bar{Y}_n^w when all weights are equal, i.e., $w_i = \frac{1}{n}$.

3. SSM of OSPM

The variance of SSM is given by:

$$\sigma_{\bar{Y}_n}^2 = E[\bar{Y}_n - \bar{Y}_{\mathfrak{R}}]^2 = E\left[\frac{1}{n}\sum_{i=1}^n y_i - \frac{1}{\mathfrak{R}} \int_{\mathfrak{R}} ds Y(s)\right]^2 = \sigma_p^2 F(n)$$

where σ_p^2 is dispersion variance of the population of the target area and $F(n)$ is a variance reduction factor and estimated by [37]:

$$\begin{cases} F(n) = 1/n & \text{simple random sampling} \\ F(n) = \{1 - E[r(s_i - s_j) | \mathfrak{R}]\} / n & \text{spatial random sampling} \\ F(n) = \{1 - E[r(s_i - s_j) | (\mathfrak{R}/k)]\} / n & \text{spatial stratified sampling} \end{cases} \quad (5)$$

where n is the number of sampling units and k is the number of strata; simple random sampling disregards spatial correlation, whereas the spatial random sampling and spatial stratified sampling take spatial correlation into account; the $r(s_i - s_j)$ expresses spatial dependence between any two sites s_i and s_j ; $E[r(s_i - s_j) | \cdot]$ is usually a positive quantity lying in the interval $[0, 1]$ and can be estimated directly from the observed $r(s_i - s_j)$ values and the probability distribution of distances over the study area \mathfrak{R} or the strata \mathfrak{R}/k [38].

Next we investigate the role of spatial correlation and sampling design on the sample mean variance. Let n_0 , n_r , and n_s denote the numbers of sample units for simple random sampling, spatial random sampling and spatial stratified sampling, respectively. To assure the required estimation accuracy $\sigma_{\bar{Y}_n}^2$, one finds from (5) that:

$$\frac{n_r}{n_0} = 1 - E[r(s_i - s_j) | \mathfrak{R}] \quad (6)$$

$$\frac{n_s}{n_0} = 1 - E[r(s_i - s_j) | (\mathfrak{R}/k)] \quad (7)$$

$$\frac{n_s}{n_r} = \frac{1 - E[r(s_i - s_j) | (\mathfrak{R}/k)]}{1 - E[r(s_i - s_j) | \mathfrak{R}]} \quad (8)$$

Because $0 \leq E[r(s_i - s_j) | \mathfrak{R}] \leq E[r(s_i - s_j) | \mathfrak{R}/k]$ by Tobler's first law of geography [39], which argues that nearby attribute values are more similar than those that are further apart; consequently, $n_s \leq n_r \leq n_0$ [from Equations (6)-(8)]. Similarly, given the same sample size n , one can compare the variances of the three sampling mean estimates and conclude that: $\text{Var}(\text{simple random sampling mean}) \geq \text{Var}(\text{spatial random sampling mean}) \geq \text{Var}(\text{spatial stratified sampling mean})$. The conclusion is that the stratified sampling is generally more efficient in reducing estimation variance than random sampling, and the sampling regarding spatial correlation is generally more efficient than that which neglects spatial autocorrelation. Efficiency refers to the fact that using fewer sample units leads to higher estimation accuracy. The SSM property of best linear unbiased estimation when sampling 1st-order spatial stationary SRF would not be retained when sampling 2nd-order stationary SRF, a drawback that can be overcome by WSM or mean Kriging.

4. Mean Kriging of OSPM

One can estimate the OSPM ($\bar{Y}_{\mathfrak{R}}$) by the WSM (\bar{Y}_n^w) using a Kriging technique (a presentation of the various Kriging techniques and their relation to other spatial estimation methods can be found in [34]). The WSM \bar{Y}_n^w satisfies two conditions: (a) it is an unbiased estimate of the OSPM $\bar{Y}_{\mathfrak{R}}$, and (b) it minimizes the mean squared estimation error. Condition (a) implies that:

$$E[\bar{Y}_n^w] = E[\bar{Y}_{\mathfrak{R}}] \text{ or } E[\sum_{i=1}^n w_i y_i] = E[\frac{1}{\mathfrak{R}} \int_{\mathfrak{R}} d\mathbf{s} Y(\mathbf{s})]$$

Since the SRF $Y(\mathbf{s})$ is 1st order spatial stationary $E[y_i] = E[Y(\mathbf{s})]$, that leads to:

$$\sum_{i=1}^n w_i = 1$$

The mean squared estimation variance is given by:

$$\begin{aligned} \sigma_{\bar{Y}_n^w}^2 &= E[\bar{Y}_n^w - \bar{Y}_{\mathfrak{R}}]^2 = E[\sum_{i=1}^n w_i y_i - \frac{1}{\mathfrak{R}} \int_{\mathfrak{R}} d\mathbf{s} Y(\mathbf{s})]^2 \\ &= \frac{1}{\mathfrak{R}^2} \int_{\mathfrak{R}} \int_{\mathfrak{R}} d\mathbf{s} d\mathbf{s}' c_Y(\mathbf{s}, \mathbf{s}') + \sum_{i=1}^n \sum_{j=1}^n w_i w_j c_Y^{ij} - \frac{2}{\mathfrak{R}} \int_{\mathfrak{R}} d\mathbf{s} \sum_{j=1}^n w_j c_Y(\mathbf{s}, \mathbf{s}_j) \end{aligned}$$

which, by condition (b), must be minimized with respect to the weights subject to $\sum_{i=1}^n w_i = 1$, that is, quantity $\sigma_{\bar{Y}_n^w}^2 + 2\theta[\sum_{i=1}^n w_i - 1]$ must be minimized with respect to the weights w_i and the Lagrange multiplier θ . This leads to the system of equations:

$$\sum_{i=1}^n \sum_{j=1}^n w_i w_j c_Y^{ij} = \frac{1}{\mathfrak{R}} \int_{\mathfrak{R}} d\mathbf{s} \sum_{j=1}^n w_j c_Y(\mathbf{s}_j, \mathbf{s}) - \sum_{j=1}^n w_j \mu = \frac{1}{\mathfrak{R}} \int_{\mathfrak{R}} d\mathbf{s} \sum_{j=1}^n w_j c_Y(\mathbf{s}_j, \mathbf{s}) - \theta \tag{9}$$

or in matrix form:

$$\begin{bmatrix} c_Y(\mathbf{s}_1, \mathbf{s}_1) \dots c_Y(\mathbf{s}_1, \mathbf{s}_n) & 1 \\ \vdots & \vdots \\ c_Y(\mathbf{s}_n, \mathbf{s}_1) \dots c_Y(\mathbf{s}_n, \mathbf{s}_n) & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \theta \end{bmatrix} = \begin{bmatrix} \frac{1}{\mathfrak{R}} \int_{\mathfrak{R}} d\mathbf{s} c_Y(\mathbf{s}_1, \mathbf{s}) \\ \vdots \\ \frac{1}{\mathfrak{R}} \int_{\mathfrak{R}} d\mathbf{s} c_Y(\mathbf{s}_n, \mathbf{s}) \\ 1 \end{bmatrix} \tag{10}$$

where C_Y denotes the corresponding covariances and θ is a Langrange multiplier that accounts for the estimation unbiasedness condition. Note that the Equation (10) above are essentially the block Kriging equations [35] but derived without the assumption of the identical dispersion variance. The integral is evaluated by a summation of the values at regularly discretized points over the area of interest. The integration error is incorporated in the estimation mean variance, see Equation (13) below. The domain boundary effect can be mitigated by drawing more samples around the intersection of the integration grid and the study area boundary.

After the weights w_i and the multiplier θ have been calculated from Equation (10), they are substituted back into Equation (4) to obtain the WSM, \bar{Y}_n^w . The corresponding minimum error estimation variance of the WSM is given by:

$$\sigma_{\bar{Y}_n^w}^2 = E[\bar{Y}_n^w - \bar{Y}_{\mathfrak{R}}]^2 = \frac{1}{\mathfrak{R}^2} \int_{\mathfrak{R}} \int_{\mathfrak{R}} d\mathbf{s} d\mathbf{s}' c_Y(\mathbf{s}, \mathbf{s}') - \frac{1}{\mathfrak{R}} \int_{\mathfrak{R}} d\mathbf{s} \sum_{j=1}^n w_j c_Y(\mathbf{s}, \mathbf{s}_j) - \theta \tag{11}$$

As we shall see below, the set of the mean Kriging Equations (4), (10) and (11) can be implemented with efficiency in the GIS environment.

Let $y = (y_1, \dots, y_k)$ be the random vector (family of random variables) of the SRF $Y(s)$ at points s_1, \dots, s_k . According to probability theory [40], if $(y_1, \dots, y_n) \sim N(\mathbf{m}, \mathbf{V})$, then from Equation (4) it is valid that $\bar{Y}_n^w \sim N(m, \sigma^2)$, where:

$$m = \sum_{i=1}^n w_i E[y_i] \quad (12)$$

[the weights w_i have been calculated from Equation (10)], the assumption of a spatially constant SRF mean still holds, and:

$$\sigma^2 = \sum_{i=1}^n w_i^2 \sigma_{y_i}^2 + 2 \sum_{i < j} w_i w_j c_Y(\mathbf{s}_i, \mathbf{s}_j) \quad (13)$$

In light of Equation (13), a confident interval of the mean Kriging can be calculated given a confidence level. E.g., with 95% confidence the value of WSM falls into the interval $m \pm 1.96\sigma$. Note that if y is shown to be skewed by the Kolmogorov-Smirnov statistics test or it turns out to be non-stationary, then detrending, square root transformation, lognormal transformation etc. may be used to transform y into a normal probability distribution [40].

In GIS practice to implement the Mean Kriging, one needs to calculate the spatial dependence functions, covariance and variogram, that are related as:

$$\left. \begin{aligned} c_Y^{ij} &= \sigma_Y^2 - \gamma_Y^{ij}, \quad i, j = 1, \dots, N \\ \sigma_Y^2 &= c_0 + c_1 \end{aligned} \right\} \quad (14)$$

where c_Y^{ij} ($i, j = 1, \dots, N$) is the covariance between the points s_i and s_j , γ_Y^{ij} is the corresponding variogram, σ_Y^2 is the variogram sill, c_0 is the nugget effect and c_1 is the partial sill. Usually, the variogram is first calculated experimentally, then a theoretical model is fitted to the experimental variogram, and finally the corresponding covariance is obtained using Equation (14). To the experimental variogram calculated on the basis of the dataset one can fit one of the available theoretical variogram models [41-44]. E.g., the spherical variogram model is used in the temperature case study considered in this work, see later.

5. Case Study I

Next we demonstrate the use of the mean Kriging technique in a GIS environment using a temperature dataset. This dataset includes temperature values (in °C) generated by the remotely sensed image of surface temperature over the study area. We then compare mean area temperature values estimated by simple random sampling, spatial random sampling and ordinary Kriging.

5.1. Study Area

The study area is the Shandong Province located in the eastern part of China, along the downstream of the Yellow River and bordering the Bohai Sea and Yellow Sea. Shandong lies in the temperate zone with a half-moisture monsoon climate, an annual average temperature of 12.7 °C and an average annual rainfall of 750 mm. Shandong Province is one of China's most important agricultural economic regions. The climate change has a significantly impact on the region's agriculture.

Figure 1 shows the MODIS image of ground temperature in the Laiyang county (Shandong province) obtained at 10:20 pm on May 14th, 2007. Each pixel of the MODIS image is regarded as a

candidate sample unit. Empirical sample datasets are readily obtained by randomly sampling the image with different proportions. The dataset shows that the temperature distribution is very close to the normal distribution (Figure 2). The skew statistics is $S = 0.038$ and the std error is $\sigma(s) = 0.188$ [i.e., $S \ll 2\sigma(s)$], in which case the skew value indicates that the distribution is almost normal although slightly positively skewed.

Figure 1. Temperature distribution (in °C), Laiyang county (Shandong, China) at 10:20 pm on May 14th, 2007 (MODIS).

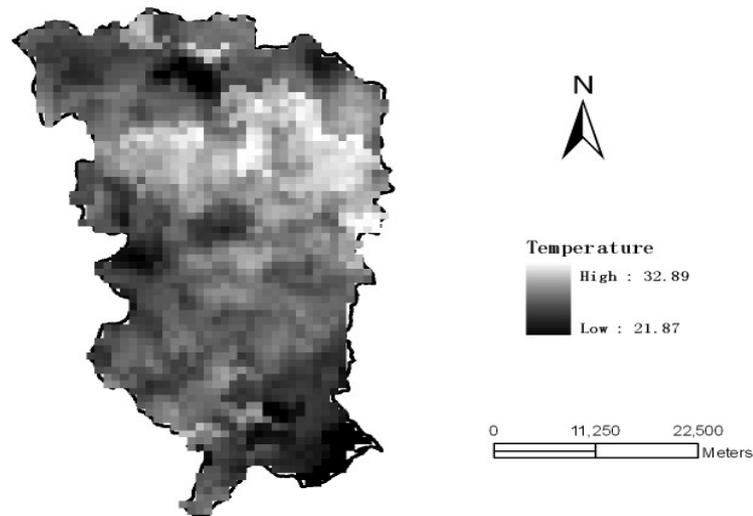
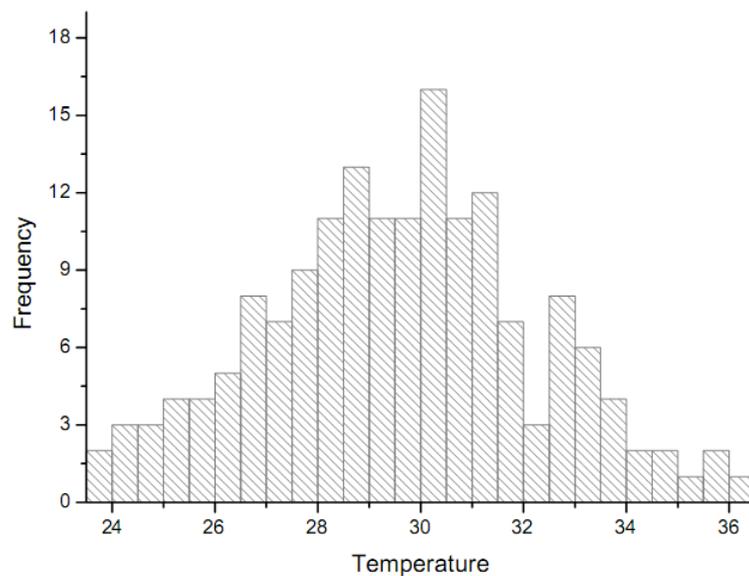


Figure 2. Histogram and simulated pdf of the normal distribution for the temperature dataset (in °C). The dataset belongs to a normal distribution with mean $m = 26.68$ °C and std deviation $\sigma = 1.403$ °C.



5.2. Variogram and Covariance Modeling

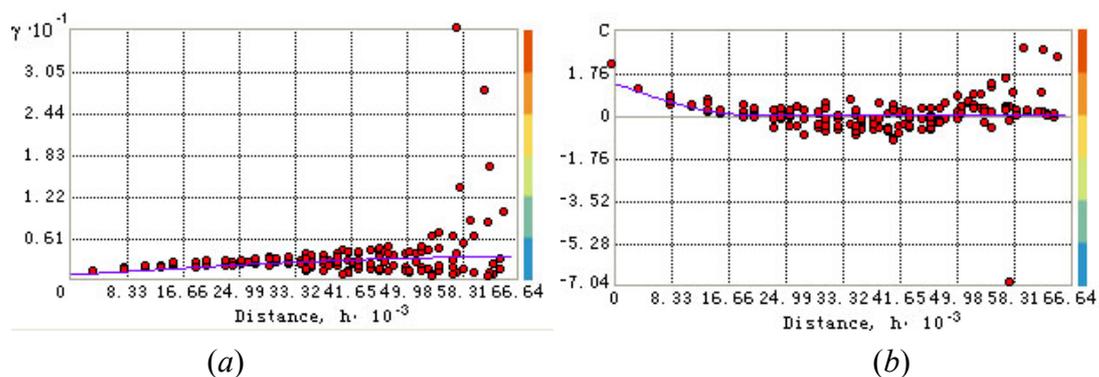
Pair-wise correlation (dependence) are calculated using MODIS image of ground surface. To the discrete (experimental) variogram we fitted the spherical variogram model (Figure 3a):

$$\gamma_Y^{ij} = \begin{cases} c_0 + c_1 \left[1.5 \frac{h_{ij}}{a} - 0.5 \left(\frac{h_{ij}}{a} \right)^3 \right] & \text{if } h_{ij} \leq a \\ c_0 + c_1 & \text{otherwise} \end{cases} \quad (15)$$

where $h_{ij} = s_i - s_j = \lambda h$ ($h = 5482.5$ meters is the lag and $\lambda = 1, \dots, 12$ is the lag number); $c_0 = 0.61848$ is the nugget effect, $c_1 = 2.667$ is the partial sill, and $a = 64985.6$ meters is the variogram range, the values are regressed from sample data. The corresponding covariance is as follows (Figure 3b):

$$c_Y^{ij} = \begin{cases} c_1 \left[1 - 1.5 \frac{h_{ij}}{a} + 0.5 \left(\frac{h_{ij}}{a} \right)^3 \right] & \text{if } h_{ij} \leq a \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

Figure 3. Fitting the spherical model to the experimental: (a) variogram and (b) covariance temperature values.



The variogram model (15) was chosen on the basis of experimentation. Several models were tested and the spherical variogram model offered a closer numerical fit to the observed data and also a simpler analytical form (Figure 3). Surely, the present analysis is tailored to the particular dataset of the case study. Hence, one can't say with certainty that the spherical model offers an ultimate representation of temperature variation. More tests are required to determine a spatial variogram that provides the closest match to regional temperature variation with specified environmental, geophysical and soil characteristics. The maximum dependence range was calculated from the experimental variogram plot. The weighted least square (WLS) technique performed better than the OLS technique in fitting the theoretical model to the experimental variogram; in particular, WLS obtains more accurate spatial continuity estimates than OLS close to the origin ($h = 0$) and it does not need the assumption of normal and independent-identically-distributed (iid) residuals. There is a certain level of model uncertainty in experimental variogram fit, and this has an impact on the mean kriging variance.

5.3. Spatial Temperature Mmeans Obtained by the Various Techniques

Table 1 lists the mean values, their variances and confidence intervals using the techniques of simple random sampling, spatial random sampling, ordinary Kriging and mean Kriging, under a

sampling proportion of 10%. Figures 4-6 display the means, their standard variances and confidence intervals estimated by the techniques under different sampling proportions.

Table 1. Spatial temperature means and their confidence intervals estimated by three techniques (10% sampling proportion).

Technique	Spatial mean (°C)	Standard deviation (°C)	95% confidence interval
Simple random sampling	29.70	1.194	[27.43, 31.97]
Spatial random sampling	29.84	1.19	[27.50, 32.17]
Ordinary Kriging	29.62	1.31	[27.06, 32.18]
Mean Kriging	29.84	1.16	[27.49, 32.18]

Figure 4. Estimates of the temperature OSPM (in °C) by various techniques.

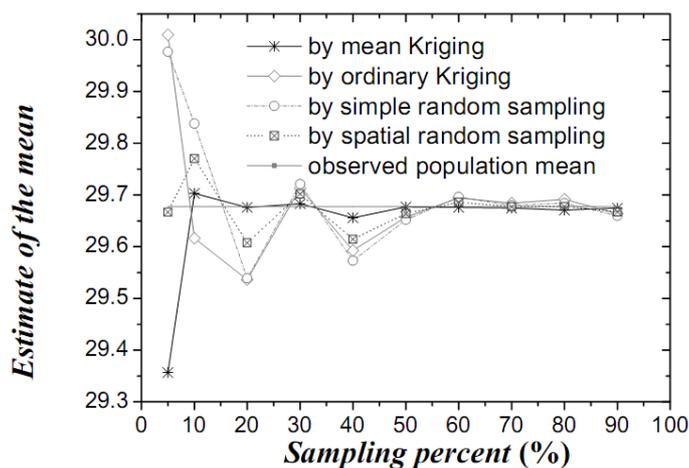
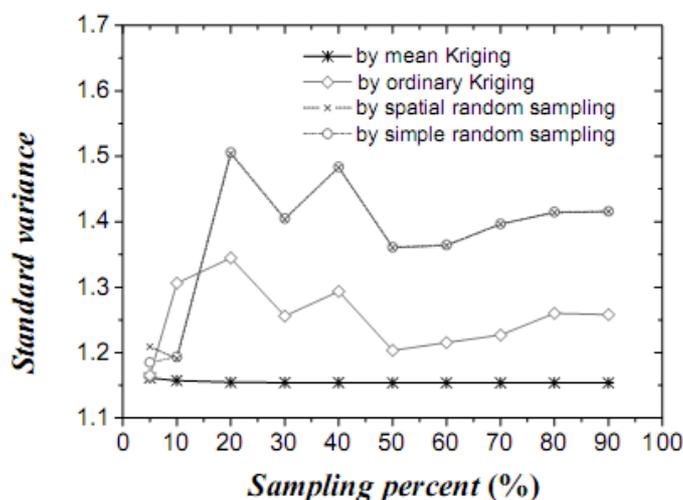


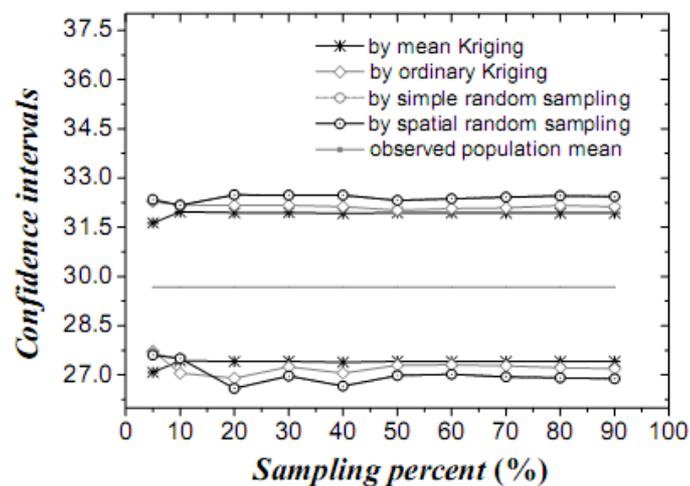
Figure 5. Standard variances of the estimated temperature (in °C) means by the three techniques.



The results obtained above show that mean Kriging has achieved a better effect, smaller variance and better accuracy of the temperature mean among the proportions ranging from 5% to 90% (5, 10, 20, 30, 40, 50, 60, 70, 80 and 90%). In Figure 4, the temperature mean estimated by mean Kriging is closer to the reference line of the observed temperature value than are the mean values obtained by

ordinary Kriging, spatial random sampling and simple random sampling; the sampling proportion varies from 5 to 90%. The relatively small change in “standard variance” with increasing sample size in the case of mean Kriging is linked to the apparently small sill in the modelled variogram (i.e., the data is highly homogeneous so that additional samples add little information about the SRF). Furthermore, Figure 5 shows that the std error variance of spatial mean estimation in terms of mean Kriging technique is minimized, which is not the case with ordinary Kriging, spatial random sampling and simple random sampling. Finally, Figure 6 shows that the confidence intervals obtained by mean Kriging are narrower than other techniques, a fact that indicates the higher accuracy of mean Kriging.

Figure 6. Confidence intervals of the estimated temperature means (in °C) by various techniques.



6. Case Study II

6.1. The Study Region

The study region is Shandong province (eastern China). Our aim is to obtain a survey of the proportion of cultivated land in the Shandong province. Actually, the cultivated land and the total territory have already been completely counted in the year 2000 by aerial photos (Figure 7). Table 2 gives a descriptive statistics of the enumerate survey.

Figure 7. Cultivated land enumerate survey by aerial photos in Shandong (China) in year 2000.

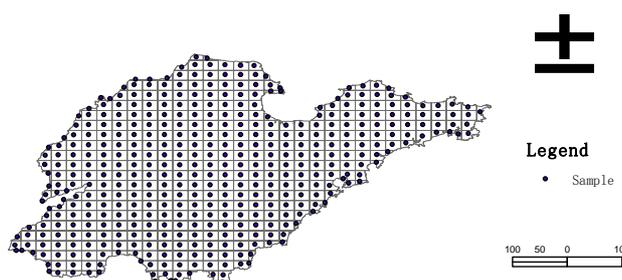


Table 2. Descriptive statistics.

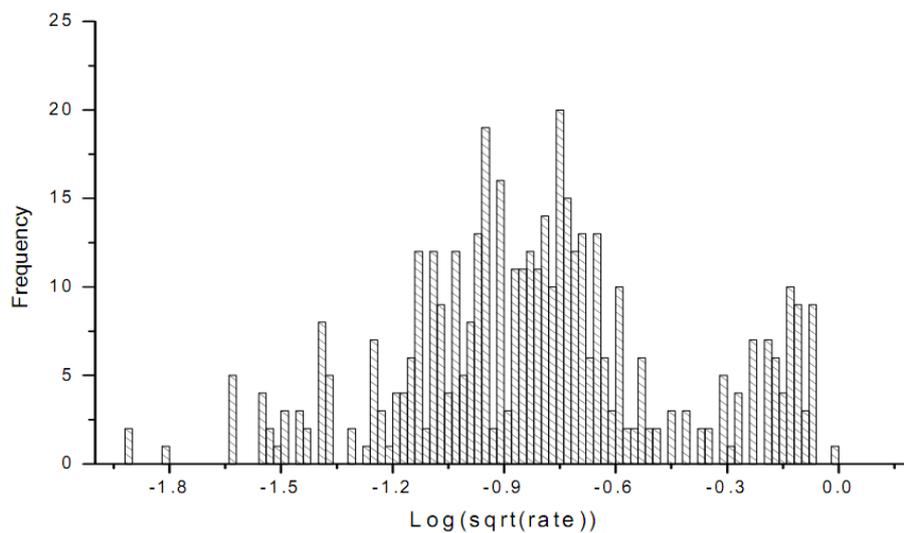
N*	Minimum	Maximum	Mean	Std. Deviation	Skewness	
					Statistic	Std. Error
438	.0218	.9977	.2003	.2171401	1.570	.117

* Note: the observed proportion of cultivated land of Shandong province is 0.265 via completed counting of the coverage.

6.2. Transformation of the Target Variable

Let the original target variable (the proportion of the cultivated land in Shandong province) be denoted by x . The x is found to be non-normally distributed, in which case the transform $y = \log(\sqrt{x})$ is conducted. The histogram of the transformed values is shown in Figure 8.

Figure 8. Transformed histogram of crude cultivated land values.



The skew statistics is $S = 0.09$ and the std error is $\sigma(s) = 0.3722$, in which case $S \ll 2\sigma(s)$. The skew value indicates that the distribution of the transformed sample attribute is almost normally distributed.

6.3. Modeling the Variogram

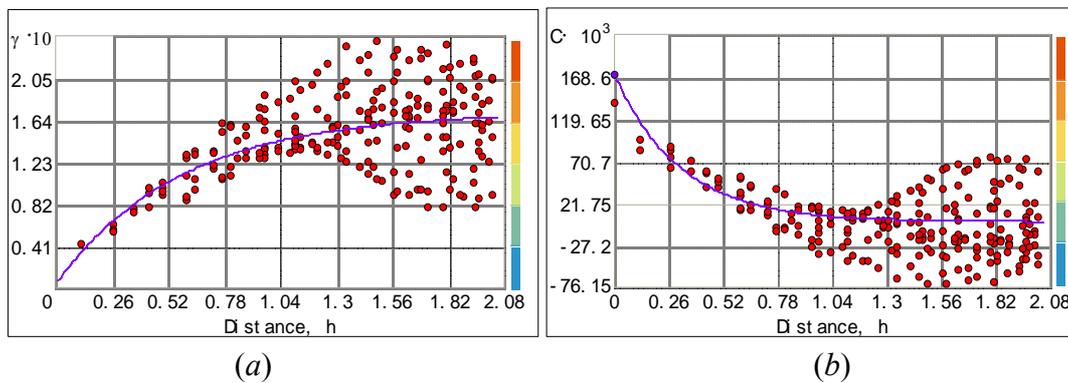
The experimental variogram and covariance are presented in Figure 9. By exploratory data analysis, we use the spherical variogram model below to simulate the data:

$$\gamma_Y^{ij} = \begin{cases} 0 & h = 0 \\ c_0 + c_1(1 - \exp(\frac{-3h_{ij}}{a})) & h > 0 \end{cases}$$

where the nugget effect is $c_0 = 0$, the partial sill is $c_1 = 0.15775$, the major range is $a = 146.437 \text{ Km}$; $h_{ij} = s_i - s_j = \lambda h$ ($h = 33.08 \text{ Km}$ is the lag and $\lambda = 1, \dots, 12$ is the lag number);

$$c(h_{ij}) = \begin{cases} 1 & h_{ij} = 0 \\ c_1 \exp(\frac{-3h_{ij}}{a}) & h_{ij} > 0 \end{cases}$$

Figure 9. Fitting the spherical model to the experimental: (a) variogram and (b) covariance.



6.4. Sample Estimates of the Rate of Cultivated Land

Table 3 presents estimates of the proportion of cultivated land by the various techniques all in 10% sampling proportion; and Figures 10-12 present the parameters, as Figures 4-6.

Table 3. Descriptive statistics.

Technique	Spatial mean	Standard variance	95% confidence interval
Simple random sampling	0.2040	0.20561	[-0.199, 0.607]
Spatial random sampling	0.2041	0.20083	[-0.199, 0.607]
Ordinary Kriging	0.1984	0.04226	[0.1154, 0.281]
Mean Kriging	0.1966	0.014218	[0.1687, 0.2245]

Figure 10. Estimates of the cultivated land OSPM by various techniques.

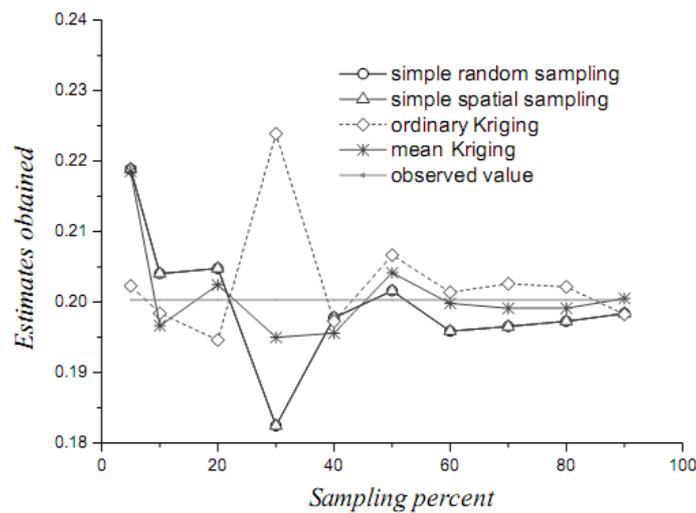


Figure 11. Standard variances of the estimated cultivated land proportion by the three techniques.

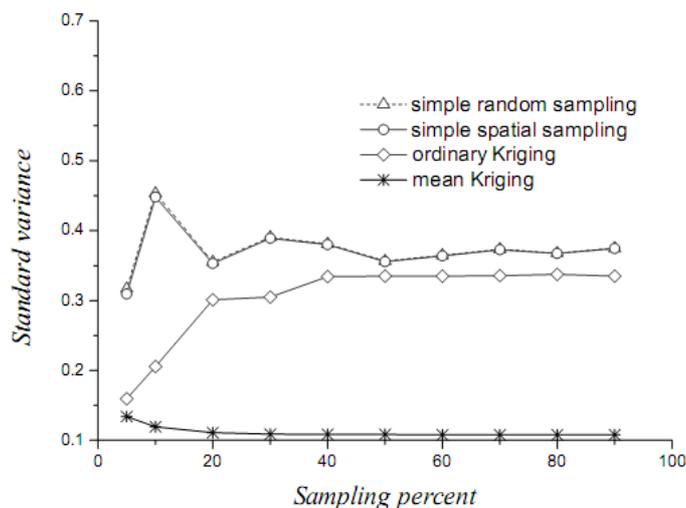
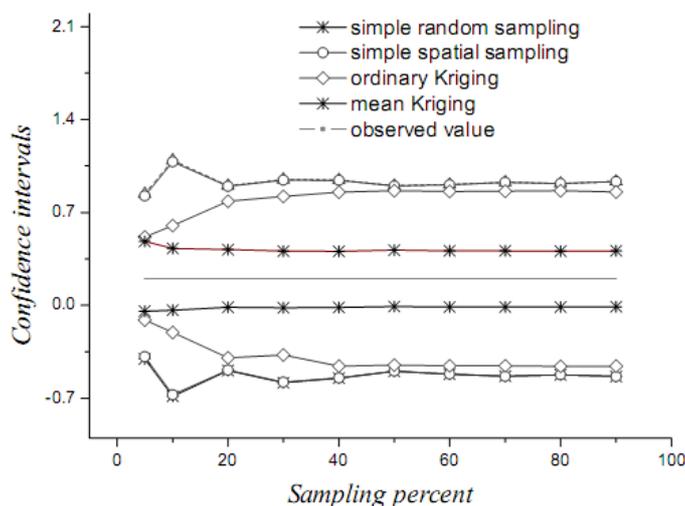


Figure 12. Confidence intervals of the estimated cultivated land proportion by various techniques.



7. Discussion and Conclusions

A method is discussed to estimate the OSPM in a GIS environment. Table 4 summarizes the formulas describing spatial means and the associated variances obtained by different techniques of estimating the OSPM. The estimation variances are ranked as: $0 = \text{variance of observable spatial population mean} < \text{variance of mean Kriging mean} < \text{variance of ordinary Kriging mean} < \text{variance of spatial random sampling mean} < \text{variance of simple random sampling mean}$. The second inequality is due to the fact that ordinary Kriging minimizes the variance at a single site but cannot guarantee minimization of the sample mean variance; the latter is guaranteed by mean Kriging. The comparative advantage of one method over another reduces when the studied area tends to be more homogeneous and less spatially-dependent. In practice, the randomness of empirical cases could lead to insignificant differences.

Table 4. Mean and variance formulas of four spatial estimation techniques. The OSPM is also shown.

	Mean value	Variance of mean value
Simple random sampling [31]	$\frac{1}{N} \sum_{i=1}^N y_i$	$\frac{1}{N} A$ where $A = \frac{1}{N} \sum_{i=1}^N \{y_i - E[y_i]\}^2$
Spatial random sampling [29]	$\frac{1}{N} \sum_{i=1}^N y_i$	$\frac{1}{N} \{A - E[c_Y(s_i, s_j)]\}$ where $c_Y(s_i, s_j) = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1}^{N-1} \{y_i - E[y_i]\} \{y_j - E[y_j]\}$
Ordinary Kriging [35]	$\sum_{i=1}^N w_i y_i$ and $\frac{1}{N} \sum_{i=1}^N y_i$	$A - \sum_{i=1}^N w_i E[c_Y(s_i, s_j)] - m$ and $\sum_{i=1}^N w_i \sigma_i^2$
Mean Kriging (this paper)	$\sum_{i=1}^N w_i y_i$	$\frac{1}{\mathfrak{R}^2} \int_{\mathfrak{R}} \int_{\mathfrak{R}} ds' c_Y(s, s') - \frac{1}{\mathfrak{R}} \int_{\mathfrak{R}} ds \sum_{j=1}^N w_j c_Y(s, s_j) - m$
Observable spatial population mean (OSPM)	$\frac{1}{\mathfrak{R}} \int_{\mathfrak{R}} ds Y(s)$	0

Although the calculation of a SSM is meaningful, straightforward and unbiased (it has the same expected value as the OSPM and SPM), its variance is not minimized and it suffers from the assumption of equal probability drawing. The OSPM can be estimated by a summation of both values at sampled sites and values at unsampled sites estimated by ordinary Kriging. The Kriging weights attached to different spatial locations within clusters are smaller than those of distanced points, so Kriging is a declustering technique. The mean squared estimation error obtained by kriging was considerably smaller than that of the unweighted sample mean. Although unbiased estimates were derived and their variances were minimized by ordinary Kriging estimation, the spatial mean estimation error (derived by the summation of Kriging values) may accumulate. In addition to the OSPM estimates, the probability distribution of these estimate for the region of interest were derived. These are best linear unbiased OSPM estimates and can be used in more relaxed GIS situations than the original block Kriging.

Using MODIS we generated ground temperature values in the Laiyang county, Shandong Province (China). It was shown that the mean Kriging technique outperformed techniques based on simple random sampling, spatial random sampling and ordinary Kriging in estimating the OSPM of the temperature. The mean Kriging not only accounted for the spatial correlation, as do the conventional spatial sampling techniques, but it also minimized the variance of the objective value as does Kriging. In this study, we focused on the sample estimation and its uncertainty due to sampling design and sample statistics. Since a sample unit is often not uncertainty-free, the sample uncertainty could finally propagate in the context of spatial mean and its variance, which is something that deserves further investigation.

Acknowledgements

This study was supported by the NSF China (40471111, 70571076, 40601077/D0120), MOST, China (2006AA12Z215, 2007AA12Z233), CAS, China (KZCX2-YW-308), and the California Air Resources Board, USA (55245A).

References

1. Christakos, G.; Bogaert, P.; Serre, M.L. *Temporal GIS*; Springer-Verlag: New York, USA, 2002.
2. Wu, B.F. China crop watch system with remote sensing. *J. Remote Sens.* (in Chinese) **2004**, *8*, 481-478.
3. Yang, X.; Zhu, B.; Wang, X.; Li, C.; Zhou, Z.; Chen, J.; Wang, X.; Yin, J.; Lu, X. Late Quaternary environmental changes and organic carbon density in the Hunshandake Sandy Land, eastern Inner Mongolia, China. *Glob. Planet. Change* **2008**, *61*, 70-78.
4. Arrow, K.; Dasgupta, P.; Maler, K.G. Evaluating projects and assessing sustainable development in imperfect economics. *Enviro. Resour. Econ.* **2003**, *26*, 647-685.
5. Christakos, G. Modern statistical analysis and optimal estimation of geotechnical data. *Eng. Geol.* **1985a**, *22*, 175-200.
6. Cressman G.P. An operational objective analysis system. *Mon. Weather Rev.* **1959**, *87*, 367-374.
7. Olea, R.A. *Geostatistics for Engineers and Earth Scientists*; Kluwer Academic Publishers: Boston, MA, USA, 1999.
8. Bayraktar, H.; Turalioglu, F.S. A Kriging-based approach for locating a sampling site-in the assessment of air quality. *Stoch. Environ. Res. Risk Assess.* **2005**, *19*, 301-305.
9. Brus, D.J.; Gruijter, J.J.D. Random sampling of geostatistical modeling? Choosing between design-based and model-based sampling strategies for soil (with discussion). *Geoderma* **1997**, *80*, 1-44.
10. Gandin, L. Objective Analysis of Meteorological Fields. *Gidrometeorologicheskoe Izdatel'stvo 1963*. (English translation. Israel Program for Scientific Translation: Jerusalem, Israel, 1965).
11. Haining, R. Estimating spatial means with an application to remote sensing data. *Commun. Stat.-Theory Methods* **1988**, *17*, 537-597.
12. Saito, H.; Zimmerman, D.A.; Coburn, T.C. Geostatistical interpolation of object counts collected from multiple strip transects: ordinary Kriging versus finite domain Kriging. *Stoch. Environ. Res. Risk Assess.* **2005**, *19*, 71-85.
13. Christakos, G. Recursive parameter estimation with applications in earth sciences. *Math. Geol.* **1985b**, *17*, 489-515.
14. Courtier, P.; Talagrand, O. Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations. *Tellus* **1990**, *42A*, 531-549.
15. Daley, R. *Atmospheric Data Analysis*; Cambridge University Press: Cambridge, UK, 1991.
16. Li, X.; Yeh, A.G. Modelling sustainable urban development by the integration of constrained cellular automata and GIS. *Int. J. Geogr. Inf. Sci.* **2000**, *14*, 131-152.
17. Serre, M.L.; Christakos, G. BME studies of stochastic differential equations representing physical laws-Part II. In *5th Annual Confer, Intern Assoc for Mathematical Geology*. Trondheim, Norway **1999**, pp. 93-98.
18. Christakos, G.A. Bayesian/maximum-entropy view to the spatial estimation problem. *Math. Geol.* **1990**, *22*, 763-776.
19. Christakos, G. Spatiotemporal information systems in soil and environmental sciences. *Geoderma* **1998**, *85*, 141-179.

20. Kolovos, A.; Christakos, G.; Serre, M.L.; Miller, C.T. Computational BME solution of a stochastic advection-reaction equation in the light of site-specific information. *Water Resour. Res.* **2002**, *38*, 1318-1334.
21. Papantonopoulos, G.; Modis, K. A BME solution of the stochastic three-dimensional Laplace equation representing a geothermal field subject to site-specific information. *Stoch. Environ. Res. Risk Assess.* **2006**, *20*, 23-32.
22. Courtier, P.J.; Derber, R.; Errico, J.; Louis, F.; Vukicevic, T. Important literature on the adjoint, variational methods and the Kalman filter in meteorology. *Tellus* **1993**, *45A*, 342-357.
23. Derber, J.C. A variational continuous assimilation technique. *Mon. Wea. Rev.* **1989**, *117*, 2437-2446.
24. Fillion, L.; Errico, R.M. Variational assimilation of precipitation data using moist-convective parameterization schemes: a 1D-VAR study. *Mon. Wea. Rev.* **1997**, *125*, 2917-2942.
25. Ott, W.; Hunt, B.R.; Szunyogh, I.; Zimin, A.V.; Kostelich, E.J.; Corazza, M.; Kalnay, E.; Patil, D.J.; Yorke, J.A. A local ensemble Kalman filter for atmospheric data assimilation. *Tellus* **2004**, *56*, 415-428.
26. Zhang, C.S.; Fay, D.; McGrath, D.; Grennan, E.; Carton, O.T. Use of trans-Gaussian kriging for national soil geochemical mapping in Ireland. *Geochemistry-Exploration, Environ. Anal.* **2008**, *8*, 255-265.
27. Christakos, G.; Kolovos, A.; Serre, M.L.; Vukovich, F. Total ozone mapping by integrating data bases from remote sensing instruments and empirical models. *IEEE Trans. Geosci. Remote Sensing* **2004**, *42*, 991-1008.
28. Goodman, J.M. *Space Weather and Telecommunications*. Springer : New York, NY, USA, 2005.
29. Ripley, B.D. *Spatial Statistics*. John Wiley & Sons: New York, NY, USA, 1981.
30. Zhu, A.X.; Hudson, B.; Burt, J.; Lubich, K.; Simonson, D. Soil mapping using GIS, expert knowledge, and fuzzy logic. *Soil Sci. Soc. Am. J.* **2001**, *65*, 1463-1472.
31. Cochran, W.G. *Sampling Techniques*. John Wiley & Sons: New York, NY, USA, 1977.
32. Griffith, D.A.; Haining, R.; Arbia, G. Heterogeneity of attribute sampling error in spatial data sets. *Geogr. Anal.* **1994**, *26*, 300-320.
33. Haining, R. *Spatial Data Analysis: Theory and Practice*. Cambridge University Press: Cambridge, UK, 2003.
34. Christakos, G. *Random Field Models in Earth Sciences*. Dover Publications Inc, Mineola, NY, USA, 2005.
35. Isaaks, E.H.; Srivastava, R.M. *An Introduction to Applied Geostatistics*. Oxford University Press: Oxford, UK, 1989.
36. Dingman, S.L.; Seely-Reynolds, D.M.; Reynolds, R.C., III. Application of Kriging to estimating mean annual precipitation in a region of orographic influence, *J. Am. Water Resour. Assoc.* **1988**, *24*, 329-339.
37. Rodriguez-Iturbe, I.; Mejia, J.M. The design of rainfall networks in time and space, *Water Resour. Res.* **1974**, *10*, 714-728.
38. Ghosh, B. Random distances within a rectangle and between two rectangles. *Bull. Calcutta Math. Soc.* **1951**, *43*, 17-24.

39. Tobler, W. A computer movie simulating urban growth in the Detroit region. *Econ. Geogr.* **1970**, *46*, 234-240.
40. Grimmett, G.R. *Probability and Random Processes*. Oxford Science Publications: New York, NY, USA, 1992.
41. Shi, W.Z.; Tian, T. A hybrid interpolation method for the refinement of regular grid digital elevation model. *Int. J. Geogr. Inf. Sci.* **2006**, *20*, 53-67.
42. Christakos, G. *Modern Spatiotemporal Geostatistics*. Oxford University Press: Oxford, UK, 2000.
43. Krejcar, P. Development of the Kriging method with application. *Appl. Math.* **2002**, *47*, 217-230.
44. Pardo-Igúzquiza, E.; Dowd, P.A. Empirical maximum likelihood Kriging: the general case. *Math. Geol.* **2005**, *37*, 477-492.

© 2009 by the authors; licensee Molecular Diversity Preservation International, Basel, Switzerland. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).