Uncertainty assessment of mapping topsoil DDT spatial distribution using multiple indicator kriging

Yong-Cun Zhao[A]

[A]Institute of Soil Science, Chinese Academy of Sciences, Nanjing, 210008, China, Email yczhao@issas.ac.cn

Abstract

Topsoil samples (n=544) were collected from the Zhangjiagang county of China. MultiGaussian (MG) approach and multiple indicator kriging (MIK) were applied for mapping the spatial distribution of DDT concentrations and the probability that DDT concentration exceeded a critical threshold concentration. High variability of DDT concentrations was observed. Areas with higher DDT concentrations are mainly distributed in the northeastern parts of the county where fluvo-aquic soils are the dominated soil types and historical heavy applications of DDT pesticide for cottons were identified during investigation. The DDT spatial distributions obtained by MG and MIK methods showed similar spatial patterns except that estimate of MG method is smoother. The areas with DDT concentrations that exceeded the background value of the Chinese Environmental Quality Standard for Soils (CEQS) using MIK method is larger than those obtained by MG method. The estimated conditional cumulative distribution functions (ccdf) using MG and MIK methods are accurate and the goodness statistics G both close to 1. However the MIK is of higher precision for modelling DDT uncertainty because the widths of probability intervals are narrower while including the expected proportions of true values.

Key Words

Uncertainty, multiple indicator kriging, DDT, spatial distribution.

Introduction

Information and assessment of spatial patterns of DDT contaminated areas are important for risk assessment, soil remediation, as well as effective management recommendations. Spatial prediction of soil contaminant, however, usually involves uncertainties that need to be considered when making decisions for future management of contaminated areas (Goovaerts 2001), because such uncertainties can be propagated into subsequent environment modelling and fundamentally impacts the ultimate results of the model (Lark and Bolam 1997). In practice, however, it is difficult to accurately characterize the spatial patterns of pollutant because two common features are often involved in the pollution data: 1) highly positively skewed histograms (hot-spots), and 2) presence of data below the detection limit (censored observations) (Saito and Goovaerts 2000), i.e. DDT concentration in soils. The multiGaussian (MG) and indicator kriging (IK) approaches provide the estimate of the ccdf as opposed to the traditional kriging approach. And IK offers a way to deal with classes or populations of both high values and values below the detection limit. As pointed out by Journel (1983) the primary advantages of the IK approach are: 1) it makes no distributional assumptions; 2) the procedure is resistant to outliers; and 3) the procedure can accommodate high connectivity of extreme values. The specific objectives of this study are 1) to map the spatial distribution of DDT concentrations using MG and MIK methods; 2) to assess the performances of MG and MIK method for modelling the uncertainties associated with the mapping process.

Methods

Study area

The study area, Zhangjiagang County, is situated on a flat alluvial plain in Yangtze River Delta region of China (Figure 1). The total area is 999 km^2. With a north sub-tropical monsoon climate, Zhangjiagang has a mean annual temperature of 15.2^oC and 1039.3 mm annual rainfall. The main soil types in the study area are fluvo-aquic soils (Aquic Cambosols) and paddy soils (Stagnic Anthrosols) (SSOSC 1984). There are different crop planting systems on each type of soils. On the fluvo-aquic soils, before the 1980’s, the rotation of cotton as summer crop and wheat as winter crop was dominant, however, after the 1980’s most of the cotton was increasingly substituted with rice. On paddy soils, the rotation of rice and wheat has always been the dominant planting system.

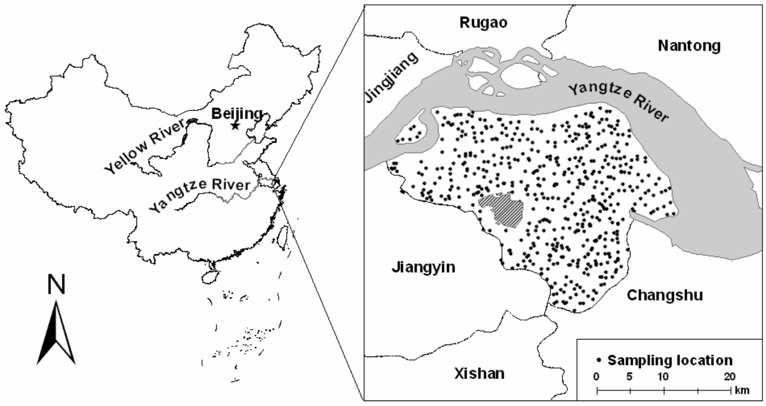

Figure 1. Location of Zhangjiagang soil sampling distribution patterns in 2004.

Soil sampling and chemical analysis

A total of 544 samples in the topsoil (0-15 cm) were collected in 2004 (Figure 1) according to soil types, land use, and thorough coverage of the study area. The collected soil samples were air-dried at room temperature and sieved to pass 70-mesh sieve for analysing DDD, DDE, and DDT using an Agilent gas chromatograph 6890 equipped with a Nickel 63 electron capture detector and a HP-5 column (30 m×0.25 mm inside diameter, 0.25 μm film thickness) (GB/T 14550-93) (SEPAC 1993).

MultiGaussian (MG) approach and multiple indicator kriging (MIK)

MG approach: If the random function (RF) $Z(\boldsymbol{u})$ is multivariate Gaussian, then the simple kriging estimate and variance identify the mean and variance of the posterior ccdf. Since that ccdf is Gaussian, it is fully determined by these two parameters (Deutsch and Journel 1998). MIK approach: The kriging algorithm applied to indicator data provides least-square estimates of the ccdf (Deutsch and Journel 1998).

Prediction precisions: The performances of the MG and MIK approaches were assessed using cross-validation (Goovaerts *et al.* 2004). The ability of different methods to estimate DDT concentration was quantified using the mean prediction error (MPE) and root mean squared prediction error (RMSPE). Local uncertainty: the p-probability intervals bounded by the $(1\text{-}p)/2$ and $(1+p)/2$ quartiles of the ccdf (PI) (Goovaerts 2001) and the "goodness" statistics G (Deutsch 1997) were used to assess the local uncertainty.

Results

Descriptive statistics

Table 1 showed that the range of ΣDDT concentrations is as large as 590 μg/kg, and *p,p'*-DDE is the major composition of DDT, the *p,p'*-DDE can account for 55% of the ΣDDT concentration averagely. Great variation exists in the DDT data, and the CV varied from 107 to 167%. Skewness of ΣDDT concentrations is 1.73, indicating that the distribution of ΣDDT data is strongly positively skewed, and suggested that higher ΣDDT concentration values existed and the spatial distribution of ΣDDT is not homogeneous.

Table 1. Descriptive statistics of DDT in soils of Zhangjiagang County, China (n=544).

	Mean *μg/kg*	*Minimum* *μg/kg*	*Median* *μg/kg*	*Maximum* *μg/kg*	*Standard deviation* *μg/kg*	*Skewness*	*Kurtosis*	CV %
p,p'-DDE	48.58	ND	23.11	344.30	61.08	1.98	6.76	126
p,p'-DDD	9.74	ND	6.93	86.81	11.31	3.35	18.30	116
p,p'-DDT	24.50	ND	14.74	384.66	35.41	4.93	38.60	145
o,p'-DDT	4.99	ND	0.78	61.17	8.35	2.45	10.01	167
ΣDDT	87.80	ND	50.60	590.10	93.86	1.73	6.31	107

ΣDDT represents the sum of four metabolites of DDT and the descriptive statistic were calculated assuming non-detect (ND) measurements were equal to one-half the detection limit.

Spatial distribution of DDT concentrations

Areas with higher DDT concentrations are mainly located in the northeastern parts of the county where fluvo-aquic soils are the dominated soil types (Figure 2). And areas with lower DDT concentrations are mainly distributed in southern parts where paddy soils are dominated. The DDT spatial patterns presented by the MIK method are more complex than those by MG method, especially in central and southern parts of the county. The local variability of DDT concentrations at central and southern parts can be intuitively observed when using MIK. However such local variability was filtered out by MG method due to stronger smoothing effect of MG method. Figure 3 presented the scatterplot of observed ΣDDT concentrations versus estimates at each of the 544 individual sampling sites. The MG method resulted in an average underestimation of 54.93 µg/kg and overestimation of 23.07 µg/kg. With respect to MIK method, the average underestimation and overestimation are 35.52 and 56.20 µg/kg, respectively. The MPE in MIK method is closer to zero than MG method, and the RMSPE in MIK method is lower than MG method, which highlights the MIK method is of higher precision for mapping the spatial distribution of DDT concentrations in the county.

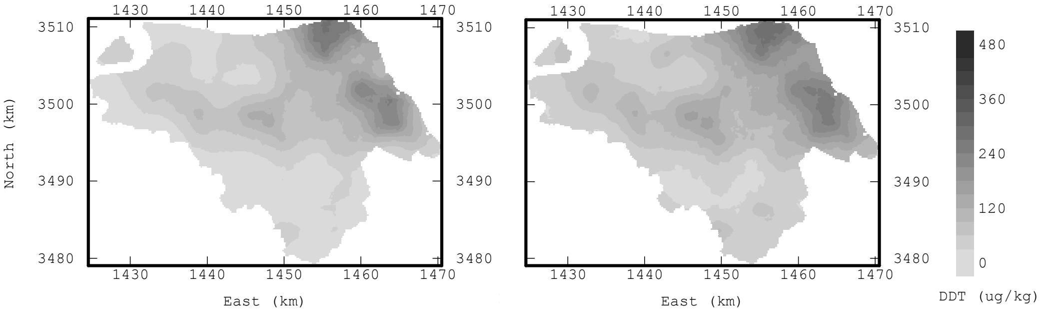

Figure 2. Estimates of DDT concentrations using multiGaussian (MG) method (left), and multiple indicator kriging (MIK) (right).

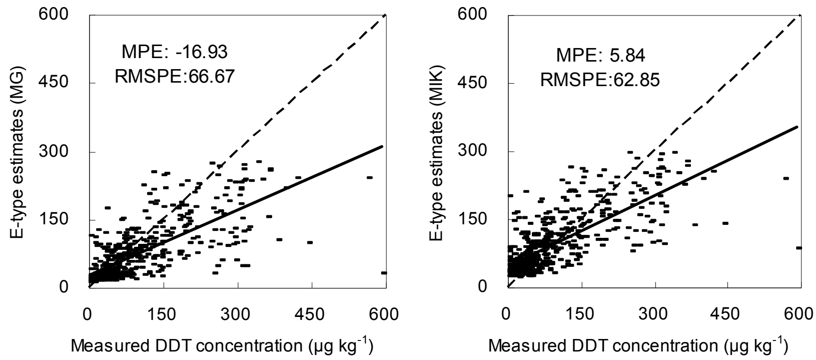

Figure 3. Cross-validation for E-type estimates of ΣDDT concentrations using multiGaussian (MG, left) method and multiple indicator kriging (MIK, right).

Performance for modelling uncertainty

The areas with DDT concentrations exceeded the background value of CEQS using MIK method is larger that those obtained by MG method (when critical probability 0.5 is used, areas with DDT concentrations exceeding 50 µg/kg using MG and MIK methods covered 47% and 49%, respectively) (Figure 4). When the critical limit is set as 500 µg/kg, MIK and MG method are of probabilities lower than 0.5, indicating that although soils in the county are not polluted by DDT based on threshold 500 µg/kg in CEQS, the potential risks of DDT residue in soils are still high because areas with DDT concentrations exceeded the background value 50 µg/kg covered nearly half of the county.

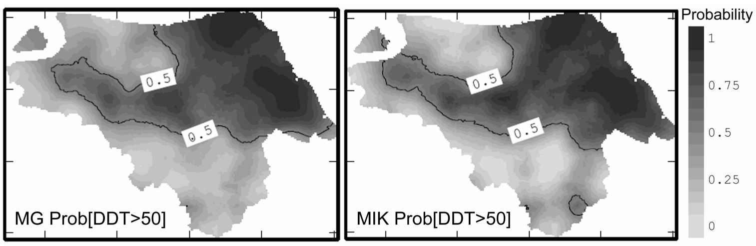

Figure 4. Probability for ΣDDT concentrations greater than the thresholds 50 µg/kg and 500 µg/kg.

The accuracy plots (the proportions of the measured data falling into different probability intervals) in Figure 5 indicated that the probability intervals of MG and MIK models of uncertainty contain a higher than expected proportion of true values. The goodness statistics (*G*) for cross validation were 0.92 for MG method and 0.94 for MIK, which is closer to the ideal value of 1.This suggests that the ccdf was quite accurately estimated. However, Figure 5 (right) showed that the best model of uncertainty is obtained by MIK method because the widths of PI are narrower (larger precision) while including the expected proportions of true values (large goodness statistics).

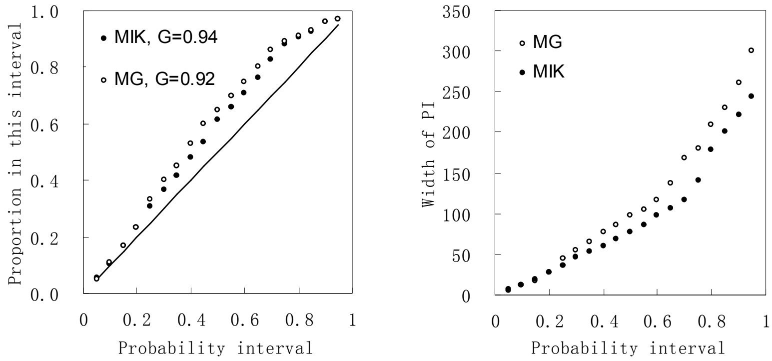

Figure 5. Accuracy plot (the proportion of measured ΣDDT values falling within theoretical probability intervals) and the width of these intervals versus the probability intervals.

Conclusion

Great variation exists in the DDTs data, and the CV varied from 107 to 167%. Areas with higher DDT concentrations are mainly located in the northeastern parts of Zhangjiagang county where fluvo-aquic soils are the dominated soil types, while areas with lower DDT concentrations are mainly distributed in southern parts. The DDT spatial distributions obtained by MG and MIK methods showed similar patterns except that estimate of MG method is smoother. The areas with DDT concentrations exceeded the background value of the CEQS using MIK method is larger that MG method. The estimated ccdf using MG and MIK was accurate and the goodness statistics *G* both close to 1. However the MIK is of higher precision for modelling DDT uncertainty.

References

Deutsch CV, Journel AG (1998) 'GSLIB: geostatistical software library and user's guide'. (Oxford University Press: New York)

Deutsch CV (1997) Direct assessment of local accuracy and precision. In 'Geostatistics Wollongong '96' (Eds EY Baafi, NA Schofield) pp. 115-125. (Kluwer Academic Publishers: Dordrecht)

Goovaerts P (2001) Geostatistical modelling of uncertainty in soil science. *Geoderma* **103**, 3-26.

Goovaerts P, AvRuskin G, Meliker J, Slotnick M, Jacquez GM, Nriagu J (2004) Modeling uncertainty about pollutant concentration and human exposure using geostatistics and a spacetime information system: Application to arsenic in groundwater of Southeast Michigan. In 'Accuracy 2004: Proceedings of the 6th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences' (Portland, Maine, USA)

Journel AG (1983) Nonparametric estimation of spatial distributions. *Mathematical Geology* **15**, 445–468.

Lark RM, Bolam HC (1997) Uncertainty in prediction and interpretation of spatially variable data on soils. *Geoderma* **77**, 263-282.

Saito H, Goovaerts P (2000) Geostatistical interpolation of positively skewed and censored data in a dioxin contaminated site. *Environmental Science & Technology* **34**, 4228– 4235.

Soil Survey Office of Shazhou County (SSOSC) (1984) 'Description of Soils in Shazhou County.' (Agriculture Bureau of Suzhou City, Jiangsu Provincial Office for Soil Survey: Suzhou). (in Chinese)

State Environmental Protection Administration of China (SEPAC) (1993) National Standard of People's Republic of China: Soil quality-Determination of BHC and DDT-Gas chromatography (GB/T 14550-1993). http://www.zhb.gov.cn/tech/hjbz/bzwb/trhj/trjcgfffbz/199401/t19940115_82035.htm