中國環境工程學刊　第十卷　第四期　（民國八十九年）
Journal of the Chinese Institute of Environmental Engineering, Vol. 10, No. 4, pp. 291-299 (2000)

APPLICATION OF DISJUNCTIVE KRIGING ON DELINEATION OF HEAVE-METAL CONTAMINATED SOILS

Kai-Wei Juang and Dar-Yuan Lee *

Graduate Institute of Agricultural Chemistry
National Taiwan University
Taipei, Taiwan, R.O.C.

Key Words : Kriging, non-linear geostatistical model, heavy-metal contamination, risk of decision-making

ABSTRACT

The spatial distribution of pollutant is an essential information to the delineation of a contaminated site. Recently, kriging has been used to estimate the spatial distribution of pollutant. Based on the spatial distribution, the hazardous areas needed for remediation can be determined. However, the kriging estimates have uncertainty due to the measurement errors and data variation. Therefore, there are risks in determining hazardous areas needed for remediation only based on the kriging estimates. Thus, for the delineation of a contaminated site, the risks of false delineation should be taken into account. Disjunctive kriging is a non-linear geostatistical approach. One can use the disjunctive kriging estimator not only to estimate the pollutant concentration but also to obtain the conditional probability of the pollutant concentration above a given cutoff value. The conditional probability of the pollutant concentration being above the cutoff value can be coupled with the estimated pollutant concentration to obtain the risks of false decision-making. The purpose of this study is to investigate the feasibility of disjunctive kriging on delineation of heavy-metal contaminated soils. A real data set of soil Cd concentrations in a contaminated site in Taoyuan County, Taiwan was studied. The results of this study demonstrate that the normalized transformation used in disjunctive kriging can prevent the spatial structure from the interference of great variation and high skewness of original data. Moreover, under a given cutoff value (for example, Cd = 10 mg/kg), the estimates of Cd concentration and the conditional probabilities of the Cd concentration, being less than or above 10 mg/kg at unsampled locations, obtained from the disjunctive kriging estimation were used to calculate the risks of false decision-making. The risks of false decision-making obtained by this approach are useful for decision-makers who need to delineate hazardous areas in a heavy metal contaminated site.

INTRODUCTION

Kriging, a spatial interpolation technique, has been used frequently for the estimation of the spatial distributions of pollutants in contaminated soils [1-7]. The final product of kriging estimation is often a contour map showing the spatial distribution of a pollutant. It can be used to delineate hazardous areas for risk assessment and remediation. Thus, kriging became an increasingly attractive method to help in decision-making for environmental management.

The linear kriging method like ordinary kriging is most frequently used in environmental science. Using the ordinary kriging in a contaminated site usually focuses on the best linear unbiased estimation of a trait of interest at an unsampled location [8-12]. However, in assessing soil contamination, knowledge of the spatial distribution of a pollutant is not enough for successful risk assessment and correct delineation. For some applications, particularly for environmental pollution, not only should the pollutant's concentration be estimated, but also the probability density of the pollutant's concentration $z(x_0)$ at location x_0 should be as well. This information can be used to determine the probability that $z(x_0)$ exceeds considered hazardous or undesirable levels of pollutant's concentrations and for environmental management. A technique proposed by Matheron [13], known as disjunctive kriging, is usually

---

* corresponding author.

used to estimate the probability density of a trait of interest at discrete spatial locations. This is a more thorough analysis in comparison with ordinary kriging [14].

Disjunctive kriging is one example of a non-linear approach. In disjunctive kriging, the variable of pollutant concentration $z(x)$ is non-linear transformed into the other variable $y(x)$, which follows a specified probability density function, like the standard normal distribution. The disjunctive kriging estimated value $z^*_{DK}(x_o)$ is decomposed into a sum of the Hermitian polynomial terms, $H^*_n[y(x_o)]$, which can be estimated by using the simple kriging estimator. According to $H^*_n[y(x_o)]$, a condition probability of the true value $z(x_o)$ being above a cutoff value z_c at a given location x_o is available [15-17]. Yates and Yates [18] have demonstrated how disjunctive kriging can be used as a management decision making tool. Moreover, they have emphasized that the conditional probability can be used as an input to a management decision-making model to provide a quantitative means for determining whether management actions are necessary. Recently, Oliver *et al*. [19] have also used disjunctive kriging for environmental management. Disjunctive kriging enables the uncertainty of estimation to be converted into the conditional probability of the true value $z(x_o)$ that exceeds the cutoff value.

In a contaminated site, the conditional probability of a pollutant's concentration being above a threshold can be coupled with the estimated pollutant's concentration to obtain the risks of false decision-making [20,21]. The false decision-making includes two types, false positives and false negatives. When the estimated value of the pollutant concentration is shown to be greater than the cutoff value, the probability of $z(x)$ being less than the cutoff value is called the risk of false positives. On the other hand, when the estimated value of the pollutant concentration is shown to be less than the cutoff value, the probability of $z(x)$ being greater than the cutoff value is call the risk of false negatives. Because the risks of false decision-making indicate the magnitude of confidence that decision-makers have in determining areas in need of remediation, it is very important to decision-makers [22]. In disjunctive kriging, the pollutant's concentration and the conditional probability of a pollutant's concentration being above a threshold at an unsampled location can be estimated simultaneously based on the standardized semivariogram. Moreover, the standardized semivariogram is usually calculated using the transformed data $y(x)$, preventing the kriging estimation from the interference of original data $z(x)$ with great variation and high skewness. Therefore, the purpose of this report is to demonstrate how disjunctive kriging can be used to assess the risks of false decision-making in delineating hazardous areas in a heavy-metal contaminated site. In this study, we proposed a means to obtain the probabilities of false positives and false negatives using both the conditional probability of a pollutant's concentration being above a cutoff value and the disjunctive kriging estimate of the pollutant's concentration. An application of this proposed method to a real data set of soil Cd concentrations in a contaminated site in Taiwan was used for illustration.

MATERIALS AND METHODS

Sampling and Measurements

This study area of about 5 ha was situated in Taoyuan county, Taiwan. According to soil taxonomy, the soil is Typic Paleudult. The soil was contaminated by heavy metals in the discharge from a chemical plant. Soil samples were regularly taken using a sampling interval of 25 m as shown in Fig. 1. The total number of sampling points was 78 in the study area. The topsoil (0-15 cm) was sampled for heavy metal measurements at each sampling location. Ten grams of soil (passed through 20 mesh) and 100ml 0.1M HCl were placed in a 250 ml flask. Each flask was shaken for 1hr, and then the soil suspensions were filtered through Whatman No. 42 filter paper. Cadmium in the filtrates was then determined by means of atomic absorption spectroscopy [23].

Disjunctive Kriging

Disjunctive kriging is described in detail in many studies [13-15,17,24]. Only a brief description of disjunctive kriging is presented herein. The disjunctive kriging method is based on the statement, the second order stationarity hypothesis. Consider a second order stationary random variable $z(x)$ with the spatial correlation, which can be described by a semivariogram. To obtain the disjunctive kriging estimator, the original variable $z(x)$ must be transformed into a new variable, $y(x)$, with a standard normal distribution. Matheron [13] has recommended the uses of Hermite polynominals and then $z(x)$ has the form:

$$z(x) = \Phi[y(x)] = \sum_{k=0}^{\infty} C_k H_k[y(x)] \qquad (1)$$

where $H_k[y(x)]$ and C_k are the Hermite polynomial and Hermitian coefficient of the order k. C_k is determined using the properties of orthogonality and evaluated numerically using

$$C_k = \frac{1}{k!\sqrt{2\pi}} \sum_{i=0}^{J} w_i \Phi(v_i) H_k(v_i) \exp(-v_i{}^2/2) \qquad (2)$$

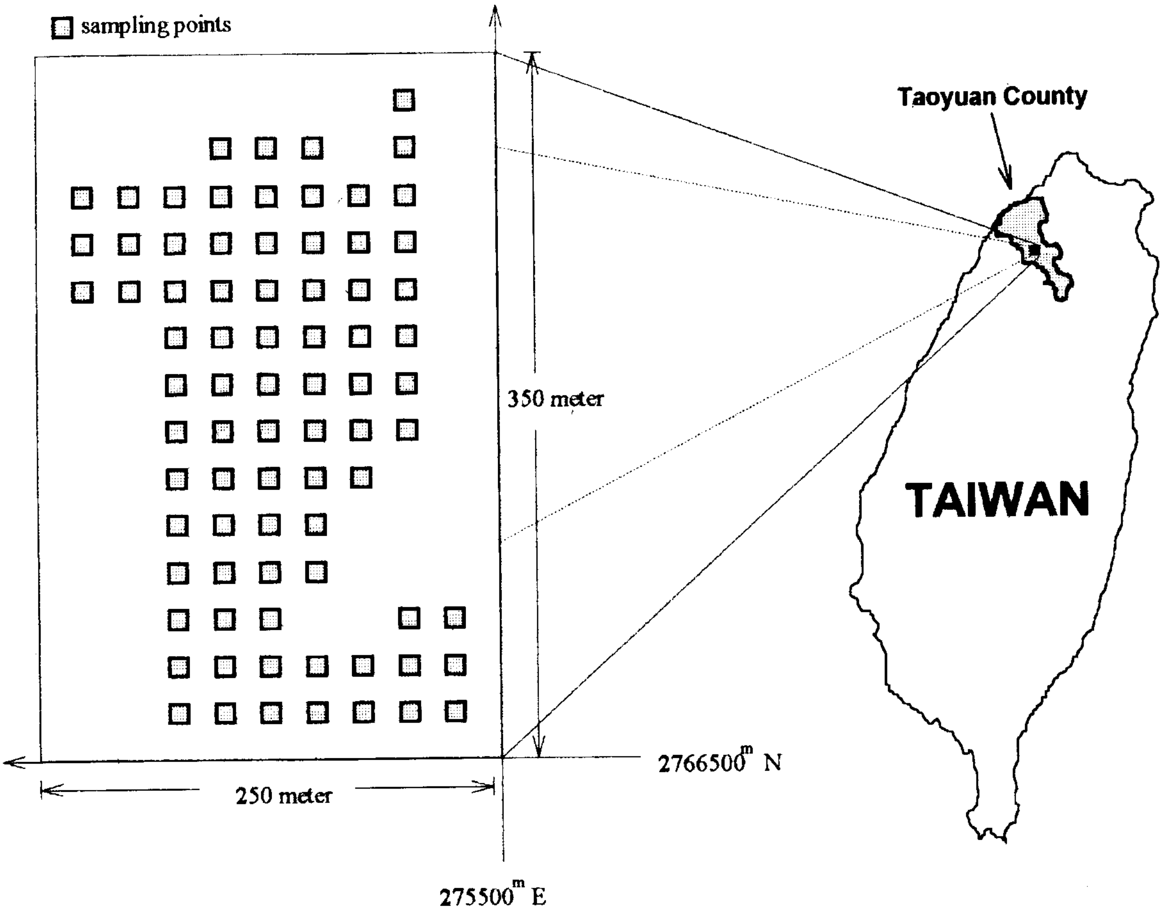

Fig.1. The study site and the sampling points.

where v_i are the specific values of $y(x)$ and w_i are the corresponding weights. Because the transformed variable $y(x)$ is a standard normal random function, the mean and variance values of $y(x)$ are equal to C_o and $\sum_{k=1}^{\infty} k!C_k^2$.

According to the relation between $z(x)$ and $y(x)$, shown as Eq. (1), the disjunctive kriging estimator, $z_{DK}^*(x_o)$, can be obtained from a combination of Hermite polynomials. The disjunctive kriging estimator is required to be unbiased and with a minimum estimation variance. It is shown as follows:

$$z_{DK}^*(x_o) = \sum_{k=0}^{K} C_k H_k^*[y(x_o)] \tag{3}$$

and

$$H_k^*[y(x)] = \sum_{i=0}^{n} \lambda_{ik} H_k^*[y(x_i)], \tag{4}$$

where the series in Eq. (3) is truncated to K terms and λ_{ik} are the disjunctive kriging weights. For each k, λ_{ik} must be found by solving the simple kriging system

$$\sum_{i=1}^{n} \lambda_{ik}(\rho_{ij}^k) = \rho_{oj}^k;\ j = 1,2,3,\cdots n. \tag{5}$$

ρ_{ij} and ρ_{oj} are autocorrelation coefficients obtained from the autocorrelation function $\rho(h)$, which can be written in terms of the semivariogram $\gamma(h)$ of $y(x)$ shown as

$$\rho(h) = 1 - \gamma(h)/r(\infty) \tag{6}$$

where $\gamma(\infty)$ is the sill value of the semivariogram. In this study, the experimental semivariogram and fitted model were obtained using the GS+ software [25].

Based on Eq. (4), this allows the estimation of the conditional probability of exceeding a cutoff value z_c to be written in terms of $H_k^*[y(x_o)]$ as

$$\begin{aligned} Prob^*[z(x_o) > z_c] &= Prob^*[y(x_o) > y_c] \\ &= 1 - G(y_c) + \sum_{k=1}^{K} g(y_c)H_{k-1}(y_c)H_k^*[y(x_o)]/k!, \end{aligned} \tag{7}$$

where y_c is the normalized value corresponding to z_c, and $G(y_c)$ and $g(y_c)$ are the cumulative distribution function (cdf) and probability density function (pdf), respectively, for the standard normal distributed variable, $y(x)$. The cutoff value z_c of Cd indicating a location hazardous was set to be 10 mg/kg in this study. The disjunctive kriging process was performed using the GEOPACK software program [26].

Risks of False Decision-Making

When the threshold, which is considered hazardous or undesirable, is known, one can assign each unsampled location to be hazardous or safe based on the comparison between the threshold and the disjunctive kriging estimate of the pollutant's content at each unsampled location, and then determined necessary management actions. However, each estimated value at an unsampled location has uncertainty; that is, the true value maybe exceeds the threshold even though the estimate is shown to be less, or the true value may be less than the threshold even though the estimate is shown to be higher. Therefore, false decision-making occurs under two conditions, called false positives and false negatives. Under whether the disjunctive kriging estimate is greater than the threshold or not, the risks of false positives $\alpha(x)$ and false negatives $\beta(x)$ can be written in terms of the conditional probability associated with the disjunctive kriging estimate. The risk of false positives $\alpha(x)$ and the risk of false negatives $\beta(x)$ are shown as follows [20]:

$$\begin{aligned}\alpha(x_o) &= Prob[z(x_o) \leqq z_c \,|\, z^*_{DK}(x_o) > z_c] \\ &= 1\text{-}Prob^*[z(x_o) > z_c] \qquad (8)\end{aligned}$$

and

$$\begin{aligned}\beta(x_o) &= Prob[z(x_o) > z_c \,|\, z^*_{DK}(x_o) \leqq z_c] \\ &= Prob^*[z(x_o) > z_c] \qquad (9)\end{aligned}$$

According to Eqs. (8) and (9), $\alpha(x_o)$ and $\beta(x_o)$ occur only on the basis of $z^*_{DK}(x_o) > z_c$ and $z^*_{DK}(x_o) \leqq z_c$, respectively. That is, if $z^*_{DK}(x_o) \leqq z_c$, then $\alpha(x)$ is equal to 0; on the other hand, if $z^*_{DK}(x_o) > z_c$, then $\beta(x)$ is equal to 0. The risks of false decision making can indicate the magnitude of confidence in determining management actions.

RESULTS AND DISCUSSION

Descriptive Statistics, Spatial Structures, and Transformation

Table 1 shows the descriptive statistics of the soil Cd concentrations in the contaminated site. The mean value of Cd concentrations is obviously greater than the background level, 0.05 mg/kg. The magnitudes of Cd concentrations range widely, and then the standard deviation (S. D.) and coefficient of variation (C. V.) are very large. Moreover, the data distribution is highly skewed. This indicates the heterogeneity of the spatial distributions of Cd in the site. There must be some extreme high observations, which usually mask the spatial dependence of the trait of interest [27]. Figure 2 shows the experimental semivariogram of the soil Cd concentrations. It indicates that the semivariogram has been influenced by the great variation and high skewness of the data set and then presents low spatial-dependence. Thus, the experimental semivariogram of soil Cd concentrations is so erratic that it is difficult to fit a reliable model for kriging estimation.

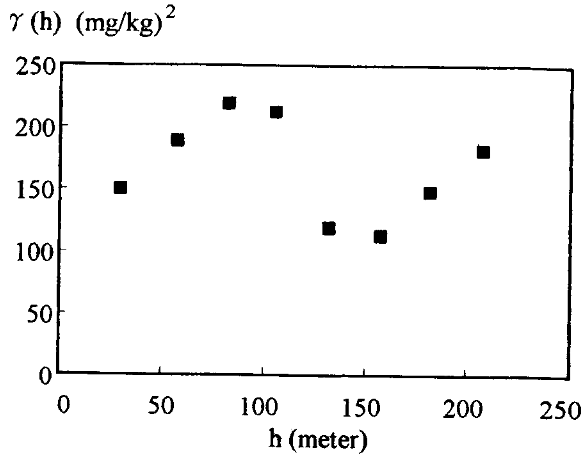

Fig. 2. The experimental semivariogram of Cd concentrations.

In the standard normal disjunctive kriging method, the original value should be transformed into the other scaled variable, which follows the standard normal distribution. The descriptive statistics of the normalized values of soil Cd concentrations are also shown in Table 1. The mean value and standard deviation (S.D.) are almost equal to 0 and 1. The normalized data are skewed slightly. The transform function is also shown in Fig. 3. The upward convex also indicates the positive skewness of original data. The normalized transformation can prevent the spatial structure from the interference of great variation and high skewness of the original data. Figure 4 shows the semivariogram of normalized values for soil Cd concentrations. The spatial structure of the normalized data shows the spatial dependence more obviously than that of original data, and then the reliable fitted model can be used in Eq. (6) for disjunctive kriging.

Choropleth Mapping of Disjunctive Kriging Estimates, Conditional Probabilities, and Risks of False Decision-Making

Figure 5(a) shows the choropleth map of the disjunctive kriging estimates of soil Cd concentrations. More than a half area of soils contain Cd concentrations exceeding the given cutoff value, 10 mg/kg. The contamination of Cd is serious on the site. Figures 5(b) and 5(c) show the spatial

Table 1. The descriptive statistics of the soil Cd concentations in the study site(n=78)

Original data	Mean	S.D.	Min	Max	C.V.(%)	Skew.
	(mg/kg)					
	9.14	12.28	0.00	83.70	134.35	3.78
Normalized data	Mean	S.D.	Min	Max	Skew.	
	0.04	0.93	-1.32	2.60	0.39	

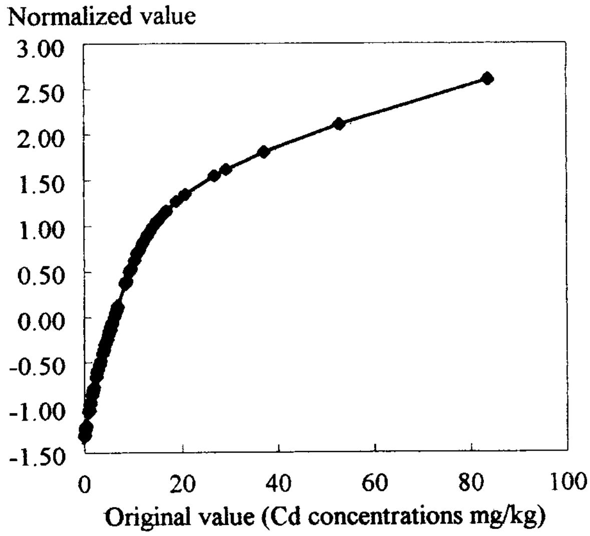

Fig. 3. The graph of transformation function for Cd concentrations.

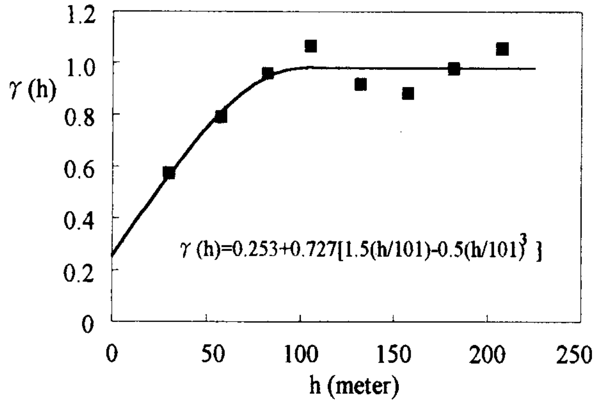

Fig. 4. The experimental semivariogram and fitted model of the normalized values for soil Cd concentrations.

distributions of the conditional probabilities of soil Cd concentrations being less than and above 10 mg/kg, respectively. Figures 5(b) and 5(c) can be coupled with Fig. 5(a), respectively, to determine the false decision-making risks.

According to the definitions of Eqs. (8) and (9), the false positive occurs, the disjunctive kriging estimate of the soil Cd concentration being above 10 mg/kg. The false negative occurs, the disjunctive kriging estimate of the soil Cd concentration being less than 10 mg/kg. Therefore, when the disjunctive kriging estimate of the soil Cd concentration at an unsample location is above 10 mg/kg, the associated conditional probability of the soil Cd concentration being less than 10 mg/kg could be assigned to be the risk of false positives (α). On the other hand, if the disjunctive kriging estimate of the soil Cd concentration at an unsample location is less than 10 mg/kg, the associated conditional probability of the soil Cd concentration being above 10 mg/kg could be assigned to be the risk of false negatives (β). Figures 6(a) and 6(b) show the choropleth maps of α and β, respectively. A false positive decision will result in the waste of resources used in remediation. If the budget for remediation is limited, Fig. 6(a) can be used to delineate hazardous areas in need of remediation based on a tolerable risk level of α. On the other hand, a false negative decision will result in hazards to human health and the environment. If heavy metal hazards to human health and the environment are of much greater concern than the cost of remediation, Fig. 6(b) can be used to delineate hazardous areas based on a tolerable risk level of β.

CONCLUSIONS

In this study, we used disjunctive kriging to produce the nonlinear estimates, which were combinations of Hermite polynomials, of soil Cd concentrations and the conditional probabilities of soil Cd concentrations being above 10 mg/kg at unsampled locations. In the standard normal disjunctive kriging procedure, the Cd concentration values were transformed into standard normal scaled values. The normalized transformation prevented the spatial structure from interference of great variation and high skewness of the original data. Moreover, based on whether the estimated concentration was greater than the considered hazardous or undesirable value or not, we obtained the false decision-making risks using the conditional probability. The obtained false decision-making risks are useful for decision-makers who need to delineate hazardous areas for reclamation in a heavy metal contaminated site.

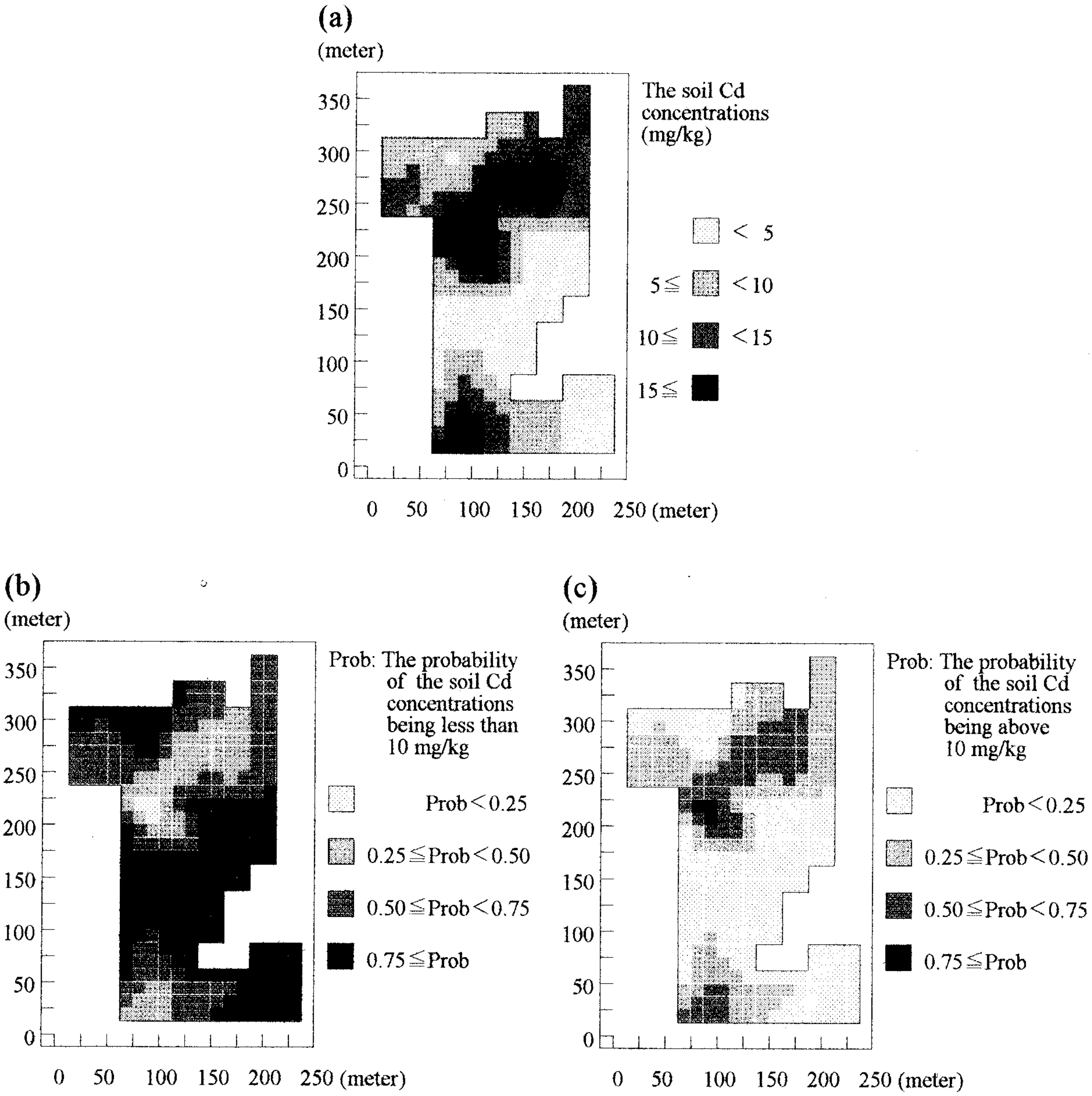

Fig. 5. Choropleth Maps of (a) the soil Cd concentrations, (b) the probability of soil Cd concentrations being above 10 mg/kg.

ACKNOWLEDGEMENTS

This research was sponsored by the National Science Council, Republic of China, under grants Nos. NSC-87-2621-P002-013 and NSC-88-2211-E-002-014.

REFERENCES

1. Stein, A., C. Varekamp, C. van Egmond, and R. van Zoest, "Zinc Concentrations in Groundwater at Different Scales," *J. Environ. Qual.*, 24, 1205-1214 (1995).
2. Arrouays, D., M. Mench, V. Amans, and A. Gomez, "Short-Range Variability of Fallout Pb in a Contaminated Soil," *Can. J. Soil Sci.*, 76, 73-81 (1996).
3. von Steiger, B., R. Webster, R. Schulin, and R. Lehmann, "Mapping Heavy Metals in Polluted Soil by Disjunctive Kriging," *Environ. Pollut.*, 94, 205-215 (1996).
4. Chiueh, P. T., S. L. Lo, and C. D. Lee, "Prototype SDSS for Using Probability Analysis in Soil Contamination," *J. Environ. Eng. (ASCE)*, 123, 514-519 (1997).
5. Juang, K. W. and D. Y. Lee, "A Comparison of Three Kriging Methods Using Auxiliary Variables in Heavy-Metal Contaminated Soils," *J. Environ. Qual.*, 27(2), 355-363

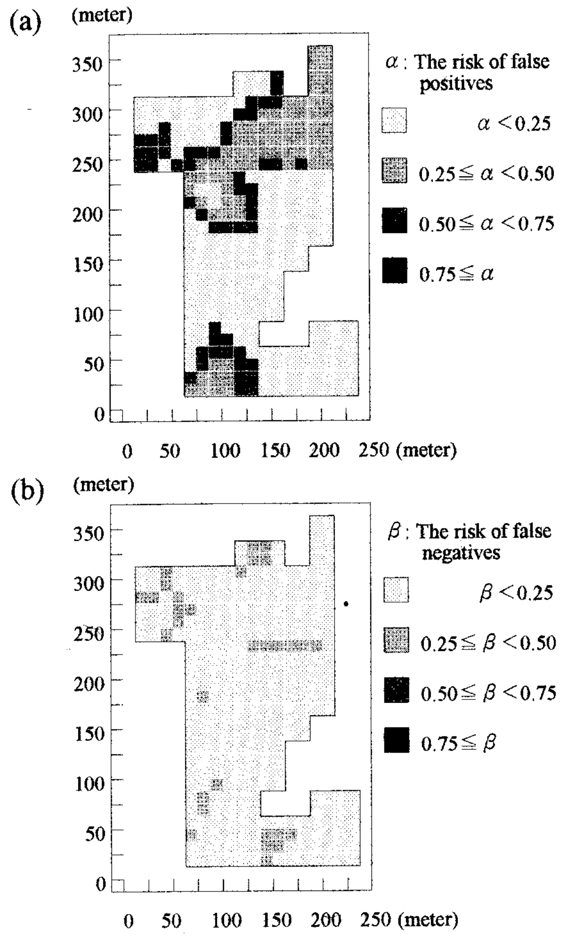

Fig.6. Choropleth maps of (a) the risk of false positives and (b) the risk of false negatives.

(1998).

6. Juang, K. W., D. Y. Lee., and C. K. Hsiao, "Kriging with Cumulative Function of Order Statistics for Delineation of Heavy-Metal Contaminated Soils," *Soil Sci.*, 163 (10), 797-804 (1998).
7. Juang, K. W. and D. Y. Lee, "Comparison of Three Nonparametric Kriging Methods for Delineating Heavy-Metal Contaminated Soils," *J. Environ. Qual.*, 29(1), 197-205 (2000).
8. Keck, T. J., W. F. Quimby, and G. A. Nielson, "Spatial Distribution of Soil Attributes on Reconstructed Minesoils," *Soil Sci. Soc. Am. J.*, 57, 782-786 (1993).
9. Litaor, M. I., "Spatial Analysis of Plutonium-239 + 240 and Americium-241 in Soils Around Rocky Flats, Colorado," *J. Environ. Qual.*, 24, 506-516 (1995).
10. Soutter, M. and Y. Pannatier, "Groundwater Vulnerability to Pesticide Contamination on a Regional Scale," *J. Environ. Qual.*, 25, 439-444 (1996).
11. Zhu, H. C., J. M. Charlet, and P. Doremus, "Kriging Radon Concentrations of Groundwaters in Western Ardennes," *Environmetrics*, 7, 513-523 (1996).
12. Dutter, R., H. Kurzl, and R. Rainer, "Geostatistical Analysis of a Mineral Deposit: a Case Study," *Environmetrics*, 8, 283-301 (1997).
13. Matheron, G., "A Simple Substitute for Conditional Expectation: The Disjunctive Kriging," Proccedings of the NATO Advanced Study Institute: Advanced Geostatistics in the Mining Industry, D. Reidel Publishing Co., Dordrecht, Netherlands, pp. 221-236 (1976).
14. Carr, J.R., E.D. Deng, and C.E. Glass, "An Application of Disjunctive Kriging for Earthquake Ground Motion Estimation," *Math. Geol.*, 18, 197-213 (1986).
15. Yates, S. R., A. W. Warrick, and D. E. Myers, "Disjunctive Kriging. I. Estimation and Conditional Probability," *Water Resour. Res.*, 22, 615-621 (1986).
16. Murthy, P.S.N., "Comparison of Ordinary and Disjunctive Kriging in an Indian Iron Ore Deposit," *Math. Geol.*, 21, 443-461 (1989).
17. Webster, R. and M. A. Oliver, "Optimal Interpolation and Isarithmic Mapping of Soil Properties. VI. Disjunctive Kriging and Mapping the Conditional Probability," *J. Soil Sci.*, 40, 497-512 (1989).
18. Yates, S.R. and M.V. Yates, "Disjunctive Kriging as an Approach to Management Decision Making," *Soil Sci. Soc. Am. J.*, 52, 1554-1558 (1988).
19. Oliver, M.A., R. Webster, and S.P. McGrath, "Disjunctive Kriging for Environmental Management," *Environmetrics*, 7, 333-358 (1996).
20. Journel, A. G., "Nonparametric Geostatistics for Risk and Additional Sampling Assessment," In L. H. Keith (ed.) Principles of Environmental Sampling, American Chemical Society, pp. 45-72 (1988).
21. Murray, M. R. and D. E. Baker, "Using Probability Spatial Analysis in Assessing Soil Cadmium Contamination," In I. K. Iskandar and H. M. Selim (ed.) Engineering Aspects of Metal-Waste Management, Lewis Publishers, pp. 25-47 (1992).
22. Juang, K.W. and D.Y. Lee, "Simple Indicator Kriging for Estimating the Probability of Incorrectly Delineating Hazardous Areas in a Contaminated Site," *Environ. Sci. Technol.*, 32(17), 2487-2493 (1998).
23. Chen, Z.S., "Cadmium and Lead Contami-

nation of Soils Near Plastic Stabilizing Materials Producing Plants in Northern Taiwan," *Water, Air, Soil Pollu.*, 57-58, 745-754 (1991).
24. Rivoirard, J., Introduction to Disjunctive Kriging and Non-Linear Geostatistics, Oxford University Press Inc., New York (1994).
25. Gamma Design, Geostatistics for the Environmental Sciences, Version 2.3, Gamma Design, Plainwell, MI (1994).
26. Yates, S. R. and M. V. Yates, "A User's Manual for the GEOPACK (Version 1.0e) Geostatistical Software System," U. S. EPA, USEPA Rep. 600/8-90/004 (1990).
27. Juang, K.W., D.Y. Lee, and Z.S. Chen, "Geostatistical Cross-Validation for the Design of Additional Sampling Regimes in Heavy-Metal Contaminated Soils," *J. Chinese Institute of Environ. Engineering*, 9 (2), 89-96 (1999).

Discussions of this paper may appear in the discussion section of a future issue. All discussions should be submitted to the Editor-in-chief within six months.

Manuscript Received: January 17, 2000
Revision Received: February 22, 2000
and Accepted: March 9, 2000

非連結克利金法在界定重金屬污染土壤的應用

莊愷瑋　李達源*

國立臺灣大學農業化學研究所

關鍵詞： 克利金法、非線性地理統計模式、重金屬污染、決策風險

摘　要

污染物的空間分佈是界定污染區清理整治範圍與面積的重要依據。近來，地理統計的克利金法常被應用在推估污染物的空間分佈，並可根據現有的監測或管制標準，在污染物的空間分佈圖上劃定出危害區域或需清理整治的範圍。由於污染物的空間分佈是依據若干採樣資料的分析與推估所得，採樣的誤差與資料的變異都可能造成空間分佈推估的不確定性，連帶產生危害區域或需清理整治範圍界定錯誤的風險；因此，劃定危害區域或需清理整治的範圍除了需要污染物的空間分佈外，仍需考慮界定錯誤造成的決策風險。非連結克利金法（disjunctive kriging）是一種非線性地理統計模式，可以用來推估污染物濃度的空間分佈，並估算出污染物含量大於監測或管制標準的機率，而污染物含量大於監測或管制標準的機率可配合污染物濃度的空間分佈算出決策錯誤的風險。本研究的主旨即在探討如何應用非連結克利金法於污染土壤的界定，並以臺灣省桃園縣的重金屬污染場址爲例做說明。經本研究結果發現，非連結克利金法能排除資料的變異性與高度偏歪對空間結構分析的干擾，推估出土壤中鎘濃度的空間分佈和鎘濃度大於管制標準（以10 mg/kg 爲例）的機率，並估算出鎘污染區界定錯誤的風險。因此，本研究所提出的方法將可提供做爲污染場址規劃與管理的重要工具。