中文摘要

在農地利用這個領域上，土壤圖(soil map)佔了很重要的角色。然而，內插一個組合資料(compositional data)需要滿足一些束制條件。由於這些數學上的束制(constraint)問題沒有解決，要內插(interpolate)一張土壤圖一向是很困難的事。首先，土壤是由各個組成份(component)依百分比組成，因此，各個組成份(component)的值必須大於0，而且小於100%。另外，由於傳統推估 (estimation)法是對各個組成份(component)獨立進行推估，因此，推估的結果可以想像，很難剛好使所有組成份(component)總和為100%。在解決這些數學問題之前有必要了解傳統推估法的特性，因此，在本文開頭將先探討幾種一般常見的推估法，並以實例來了解推估的效果。地理統計法中，克利金 (kriging)及補克利金 (cokriging)法是相當常見的推估法。在第二章中，本文用克利金 (kriging)及補克利金 (cokriging)來對宜蘭土壤資料做推估，而且暫時不考慮是否符合束制。此外，本文選定某些採樣密度來做內插，經由非線性迴歸分析，可以得到一個採樣密度(sample density)及推估精度(estimation precision)之間的經驗式。

如果有輔助變量 (secondary variable)出現，可進一步將克利金 (kriging)法擴展為補克利金 (cokriging)法，有效地利用輔助變量 (secondary variable)來改善推估精度(estimation precision)。克利金 (kriging)及補克利金 (cokriging)希望得到無偏推估。這項要求稱為無偏條件 (non-bias condition)。傳統的補克利金 (cokriging)法強制讓主變量 (primary variable)的權重 (weights)和為一，及，輔助變量(secondary variable)的權重(weights)和為零，可使補克利金 (cokriging)為無偏推估(unbiased estimation)。但事實上，要達到無偏推估(unbiased estimation)的無偏條件(non-bias condition)是可以有很多型式的。在第三章中，本文將探討各種不同

無偏條件 (non-bias condition)對補克利金(cokring)的影響。

傳統的無偏條件(non-bias condition)是強制使輔助變量 (secondary variable)的權重 (weight)和為零。這麼做將使得輔助變量(secondary variable)的貢獻變小。本文將提出一種新的無偏條件 (non-bias condition)。使得多變量 (multi-variate)的補克利金 (cokriging)法的公式和單變量 (uni-variate)的克利金 (kriging)公式完全相同。同時，本文導出平移補克利金 (rescaled cokriging)和正規化補克利金(correlogram cokriging)之間的關係式。當簡單補克利金 (simple cokriging)的平均值用推估的平均值 (estimated average)取代，本文導出簡單補克利金 (simple cokriging)和普通補克利金 (ordinary cokriging)之間並非一定相等。

在第四章中，本文將提出一個內插土壤粒徑組合(granulometric composition)的新方法。在已知文獻中，關於組合資料(compositional data) 的推估，大略上來講，可以說是將組合資料(compositional data) 的各個組成份(component)加以推估。不過，既然組合資料(compositional data) 是一個向量，用向量的方式考慮推估問題似乎更合理。本文所提出的方法，其中心概念為: 以向量的方式來考量的推估問題等同於一個最佳化規劃問題(optimization programming problem)。利用一個分解組合(decomposition) 的方法，這個最佳化規劃(optimization programming)問題可進一步轉變為一個在概念上很簡單，但強而有力的推估方法。本文提出的方法的獨特之處在於，先用克利金 (kriging)法對每一個組成份(component)做推估。推估問題即為以克利金(kriging)值為中心(center)的二次方規劃(quadratic programming)問題。這個方法的效能(performance)存在一個上限(就是克利金(kriging)法以及下限(平均值(sample mean)法)。本文將以一個含砂、砏、粘資料的實例，用本文提出的方法對土壤粒徑組成(granulometric composition)百分比做

推估，並和克利金(kriging)法及平均值(sample mean)法做比較，本文的方法其效能(performance)非常接近效能上限。

第五章中，本文把已知的數種推估法，加上本文提出的方法，以實例探討的方式討論它們的推估效能(performance)。區域化組合資料(regionalized composition)目前有三種方法可以用來做推估: 克利金 (kriging)法，比數對數法(ALR transform)，及基底法(basis method)。目前並無一個準則來評斷一方法的效能(performance)好壞。前人研究中對於這個課題通常僅著重在比較組合組成份(component)的統計數據上。然而，僅利用各組成份(component)的統計數據對於效能評斷上並沒有太大的用處。本文提出的效能(performance)量測準則(criterion)是用模數平均(mean norm)表現來量測。本文的方法和其他方法，即比數對數法(ALR transform)和基底法(basis method)，可據此準則(criterion)做一比較。一個實際的沉積質粒徑(sediment grain size)資料，可以用來比較這三種方法的效能(performance)。結果顯示當粒徑資料(grain size)含有零組成份(zero component)時，可以證明比數對數法(ALR transform)的效能(performance)會大幅下降。

有一類型的推估法，例如2-層最佳推估(SK2)、克利金(OK)、對數比數法(ALR)及基底法(BASIS)等，都是用克利金來做空間結構(spatial structure)的推估。然而，當一個區域化變數(regionalized variable)為對數常態分佈(lognormal distribution)的時候，對數常態克利金(lognormal kriging)的推估結果有時比普通克利金(ordinary kriging)更好。因此，在第6章中，利用對數常態克利金(lognormal kriging)的概念，本文將提出兩個新推估方法。第一個方法是偏推估式(biased estimator)。此方法利將推估問題解構(decompose)成兩層的規劃問題(optimization)，進而大幅降低了計算的複雜度。第二個方法則是無偏推估式(unbiased estimator)，不過計

算複雜度要比第一個方法大。由一個地下油庫資料的實例，本文將比較新方法和一些克利金類的方法(logOK，OK，SK2，ALR，BASIS)。而以均平方模數(mean-squared-norm)為準，可用來量測各方法的推估效能(performance)。交叉驗證(cross-validation)的結果顯示本文的新方法推估效能最好。本章最後將提出一個顏色對資料的編碼方法，可以將資料用彩色畫在2維地圖上。這個新的顏色編碼使得資料的變化及關係可以在彩色地圖上很容易看得出來。

最後，第7章為對於效能行為(performance behavior)的研究。本文將定義兩個效能測量器(performance measure)。效能曲線則是對鄰居搜尋半徑(search neighborhood radius)以及採樣密度做圖可以得到。結果顯示，不論鄰居搜尋半徑如何改變，本文的新方法及所有克利金類的方法，效能曲線的形狀都很接近。另一方面，增加採樣密度則可以增加推估精度。

關鍵詞： 克利金；粒徑；內插；地理統計法；矩陣式；無偏條件

Abstract

In the area of agricultural tillage usage, soil maps play an important role. However, it is usually difficult to get an interpolated soil map due to some problems generated by the constraints imposed on the value of grain size data. First, grain size data are consisted of percentage fractions, therefore it is required that every component must be larger than zero and less than 100%. Furthermore, it is difficult to obtain the sum of all components exactly 100% because traditional estimation methods perform estimation separately for distinguish component. Before solving these mathematical problems, we first performed estimation on a case study by several traditional estimation methods to understand the real effect of traditional estimation methods. In geostatistic methods, the ordinary kriging and ordinary cokriging methods are the most commonly used methods; hence, in Chapter 2 the ordinary kriging and ordinary cokriging were applied on the soils of Ilan without considering the aforementioned constraints temporarily. Moreover, to explorer the relationship between sample density and estimation error, some subsets of sample values were randomly chosen and then used to interpolate soil maps.

When secondary information is present, ordinary kriging method can be extended to cokriging to improve estimation precision. Kriging and cokriging require that the estimates are unbiased on average. This requirement generates a non-bias condition. The traditional non-bias condition of cokriging requires the primary data weights to sum to one whereas the weights of each secondary data are required to sum to zero. In fact, there exist various weight constraints which can ensure the non-bias. At Chapter 3, we discuss the influence of several non-bias conditions applied on cokriging methods.

Requiring that the weights of secondary information add up to zero will reduce

the influence of the secondary information. In this dissertation, we proposed a new non-bias condition. The programming of cokriging which is multi-variate, based on the new non-bias condition, has the same form as that of ordinary kriging which is uni-variate. In the meantime, we arrive at the relations between rescaled cokriging and correlogram cokriging. Further, when the stationary mean of simple cokriging is replaced by local mean, we find that there exists a necessary condition to making the estimation of simple cokriging being identical with that of ordinary cokriging.

At Chapter 4, we develop a new method for the interpolation of soil textural data. Previous work on the estimation of a soil composition has primarily involved the estimation of separate components of the composition. Since composition is a vector quantity, it is more appropriate to consider the estimation problem in terms of vectors. The estimation problem using vectors can be formulated in a single optimization programming and, by use of a decomposition method the optimization programming is transformed into a conceptually simple yet powerful estimation method. The significance of our method is that ordinary kriging is performed independently for each component. Given the krigings, the estimation of the target composition is equivalent to a quadratic optimization programming with center at these krigings. We find that there exists an upper bound (the ordinary kriging method) and lower bound (the sample mean method) for the performance of our method. By comparing our method with the ordinary kriging and sample mean method using data from a case study that reported the fractions of sand, silt and clay, we found that the performance of our method is very close to the upper performance bound.

At Chapter 5, we presented a performance comparison of the proposed method and several methods which have been used in various studies, i.e., log-ratio method and basis method. Currently there is no established criterion for the choice of these approaches. Previous works on this problem were focused on statistical properties of the individual component of the estimated composition, which were of little use for choosing the best estimation approach. A performance measure defined in the mean-norm sense is used in the comparison. Performance results obtained from one case study relating to grain size data were used to evaluate these estimation methods. For grain size data, we have shown here that the presence of zero components of compositions will result in performance degradation for the log-ratio method.

There exists a class of estimation methods of compositional data, in which ordinary kriging is used for the analysis and estimation of spatial structure, e.g., two-level optimal estimation (SK2), ordinary kriging (OK), log-ratio approach (ALR) and basis method (BASIS). When a regionalized variable is distributed lognormally, the lognormal kriging (logOK) may sometimes provide better performance than the tanditional ordinary kriging. Therefore, in Chapter 6 we derive two new methods based on lognormal kriging for the estimation of compositions. The first proposed method is a biased estimator. It is derived by decomposing the estimation problem into a two-level optimization problem and hence, achieves computational efficiency significantly. On the contrary, the second proposed method is an unbiased estimator but with a more complex computation than the first proposed method. By a case study relating to reservoir data, in this chapter we made a performance comparison of the proposed methods and several kriging-based methods (i.e., logOK, OK, SK2, ALR and BASIS). The mean-squared-norm performance measure is used to assess the performance of the different methods. Cross-validation results show that the proposed methods provide the best performance. A spectral encoding scheme was proposed in which the data was organized to corresponding organizations of three color constituents (red, green and blue) to represent tri-variate compositions on a two-dimensional map. The use of the proposed spectral encoding for data display allows interrelationships and patterns within data to be easily visualized.

Finally, the performance behavior of several estimation methods is presented in Chapter 7. The same analyses were repeated twice for each of two performance measures, i.e., the mean-squared-norm measure and simplex distance measure. The performance curves of these estimation methods were then obtained by varying the search neighborhood radius and sample size. The results show that the performance curves of the proposed lognormal methods and kriging-based methods remain similar to each other regardless of the change in search neighborhood radius whereas the accuracy of each of the estimations increases with increasing sample size.

***Keywords*:** Kriging; Grain-size; Interpolation; Geostatistic; Matrix form; Non-bias condition