

# 地理資訊系統的原則

*Principles of Geographical Information System*

編 著：馮豐隆

林鴻鵬

張小飛

張雪梅

曾家琳

楊大吉

詹雅文

薛怡珍

國立中興大學森林學系森林資源調查測計研究室印製

二〇〇〇年七月

# 地理資訊系統的原則

## *Principles of Geographical Information System*

*Burrough P. A. and R. A. McDonnell. 1998. Oxford University Press.*

### 內容(CONTENT)

- 第一章 地理資訊：社會、科學與系統（曾家琳 譯）  
Chap1、Geographical Information：Society, Science, and, Systems
- 第二章 資料模式與公理：實物的抽象表現（詹雅文 譯）  
Chap2、Data Models and Axioms：Formal Abstractions of Reality
- 第三章 電腦中的地理資料（張雪梅 譯）  
Chap3、Geographical Data in the Computer
- 第四章 資料的輸入、證實、貯存與輸出（楊大吉 譯）  
Chap4、Data Input, Verification, Storage, and Output
- 第五章 由點資料推導產出連續的面資訊（林鴻鵬 譯）  
Chap5、Creating Continuous Surfaces from Point Data
- 第六章 由地理統計學做適當的空間插入推估（張小飛 譯）  
Chap6、Optimal Interpolation using Geostatistics
- 第七章 空間中不連續實體的分析（曾家琳 譯）  
Chap7、The Analysis of Discrete Entities in Space
- 第八章 連續資料的空間分析（詹雅文 譯）  
Chap8、Spatial Analysis using Continuous Fields
- 第九章 資料誤差與品質控制（張雪梅 譯）  
Chap9、Errors and Quality Control
- 第十章 數值模式裏誤差的擴散（楊大吉 譯）  
Chap10、Error Propagation in Numerical Modelling
- 第十一章 模糊集合與模糊地理物件（林鴻鵬 譯）  
Chap11、Fuzzy Sets and Fuzzy Geographical Objects
- 第十二章 地理資訊系統目前狀況與未來趨勢（張小飛 譯）  
Chap12、Current Issues and Trends in GIS

附錄一 名詞解釋 (薛怡珍 整理)

APPENDIX1 Glossary of Terms

附錄二 地理及地理資訊系統有關之著名 WWW 網站

APPENDIX2 A Selection of World Wide Web Geographical and GIS Servers

附錄三 資料組例子

APPENDIX3 Example Data Sets

引用文獻

References

索引

Index

## 第一章 地理資訊：社會、科學與系統

### Chap1、Geographical Information：Society, Science, and, Systems

(一)地理資訊學會：簡史 .....	1-1
Geographical Information Society：a brief history	
(二)地理資訊系統 .....	1-3
Geographical Information Systems	
(三)本書的結構 .....	1-8
The structure of this book	
(四)GIS 的定義 .....	1-9
Definitions of GIS	
(五)地理資訊系統的組成 .....	1-10
The components of a geographical information system	
(六)問題 .....	1-14
Questions	
(七)未來閱讀的建議 .....	1-14
Suggestions for further reading	

## 第二章 資料模式與公理：實物的抽象表現

### Chap2、Data Models and Axioms：Formal Abstractions of Reality

(一)實際地理現象的觀念模式 .....	2-2
Conceptual models of real world geographical phenomena	
(二)地理資料模式與地理資料原件 .....	2-4
Geographical data models and geographical data primitives	
(三)地理原件以向量法和網格法表現 .....	2-9
The display of geographical primitives using vector and raster approaches	
(四)資料型態 .....	2-10
Data types	
(五)資訊系統中處理運用資料的定理與程序 .....	2-11
Axioms and procedures for handing data in information system	
(六)數據模擬與空間分析 .....	2-12
Data modelling and spatial analysis	
(七)使用數據模式的案例 .....	2-12
Examples of the use of data models	
(八)總結：實體或領域 .....	2-15
Summary：entities or fields	
(九)問題 .....	2-16
Questions	
(十)未來閱讀的建議 .....	2-16
Suggestions for further reading	

## 第三章 電腦中的地理資料

### Chap3、Geographical Data in the Computer

(一)電腦中的資料 .....	3-2
Data in the Computer	
(二)輸入電腦的基本資料模式的編碼 .....	3-4
Coding the basic data models for input to the computer	
(三)資料結構：電腦中的資料組織 .....	3-6
Database structures : data organization in the computer	
(四)檔案與資料的獲得擷取 .....	3-6
File and data access	
(五)資料庫結構與資料庫管理 .....	3-8
Database structures and database management	
(六)資料庫管理系統 .....	3-14
Database management system (DBMs)	
(七)選擇最適當的資料庫結構 .....	3-15
Choosing the most appropriate database structure	
(八)為展現空間現象幾何的資料庫結構 .....	3-15
Data structures for representing the geometry of spatial phenomena	
(九)空間資料的資料結構：網格式與向量式的選擇 .....	3-33
Data structures for spatial data : the choice between raster and vector	
(十)向量式與網格式資料結構的數據資料庫儲存量 .....	3-35
Database storage of vector and raster data structures	
(十一)物件導向的資料庫結構：結合屬性與幾何圖籍的貯存 .....	3-36
Object-oriented database structures : unifying attribute and geometric storage	
(十二)關連式資料庫與物件導向資料庫間的討論 .....	3-37
The debate on relational-hybrid GIS versus Object Orientation	
(十三)問題 .....	3-38
Questions	

#### **第四章 資料的輸入、證實、貯存與輸出**

#### **Chap4、Data Input, Verification, Storage, and Output**

(一)地理資料的來源 .....	4-1
Source of Geographical Data	
(二)地理資料的收集者與提供者 .....	4-3
Geographical Data collectors and providers	
(三)由資料供應者獲得數位資料 .....	4-4
Acquiring digital datasets from a data supplier	
(四)手工操作建立數位資料 .....	4-7
Creating digital data sets by manual input	
(五)資料呈現 .....	4-15
Data presentation	
(六)資料更新 .....	4-18
Data updating	
(七)資料貯存 .....	4-18

Data storage

(八)問題 ..... 4-20  
 Questions

(九)建議閱讀之文獻 ..... 4-20  
 Suggestions for further reading

**第五章 由點資料推導產出連續的面資訊**

**Chap5、Creating Continuous Surfaces from Point Data**

(一)空間插入推估：是什麼？何以需要？ ..... 5-1  
 Interpolation：what it is and why it is necessary

(二)連續的視覺化 ..... 5-2  
 The visualization of continuous surfaces

(三)空間內插推估的原理 ..... 5-2  
 The rationale behind interpolation

(四)空間插入推估的資料來源 ..... 5-2  
 Data sources for interpolation

(五)空間內插推估的方法 ..... 5-4  
 Methods for interpolation

(六)全球空間內插推估 ..... 5-5  
 Global interpolation

(七)使用分類模式做全球推測 ..... 5-5  
 Global prediction using classification models

(八)使用趨勢面進行全球空間插入推估 ..... 5-9  
 Global interpolation using trend surfaces

(九)使用便宜的全球迴歸進行空間推測，以獲可量測的屬性 ..... 5-12  
 Spatial prediction using global regression on cheap-to-measure attributes

(十)空間插入推估的地方性、確定性 ..... 5-13  
 Local, deterministic methods for interpolation

(十一)最鄰近資料推估法：徐昇多邊形法 ..... 5-14  
 Nearest neighbors：Thiessen (Dirichlet/Voronoi) polygons

(十二)線性空間插入推估：距離倒數空間推估法 ..... 5-17  
 Linear interpolators：inverse distance interpolation

(十三)曲線規法 ..... 5-18  
 Splines

(十四)簡單全球和區域方法的比較 ..... 5-20  
 A comparison of simple global and local methods

(十五)數位高程模型為由空間推估連續面的特例 ..... 5-21  
 Digital elevation models as a special case of continuous surfaces created by interpolation

**第六章 由地理統計學做適當的空間插入推估**

**Chap6、Optimal Interpolation using Geostatistics**

(一)地區化變數理論與克立金法的介紹 ..... 6-1

IA brief introduction to regionalized variable theory and Kriging

(二)配合變異圖模式 .....	6-3
Fitting variogram models	
(三)使用變異圖的空間分析 .....	6-5
Using the variogram for spatial analysis	
(四)等方向性與各向量性之變異 .....	6-6
Isotropic and anisotropic variation	
(五)不同尺度下以變異用表示空間變量 .....	6-6
Variograms showing spatial variation at several scales	
(六)不同覆蓋層級的空間變異 .....	6-7
Spatial variation within different cover classes	
(七)使用變異圖進行空間插入推估：次序性克立金法 .....	6-7
Using the variogram for interpolation：ordinary Kriging	
(八)使用克立金驗證變異圖模式 .....	6-11
Using Kriging to validate the variogram model	
(九)克立金的其他形式 .....	6-12
Other forms of Kriging	
(十)使用其他資訊的克立金法 .....	6-21
Kriging using extra information	
(十一)機率性的克立金法 .....	6-24
Probabilistic kriging	
(十二)模擬 .....	6-26
Simulation	
(十三)不同空間插入推估方法的相對性質 .....	6-29
The relative merits of different interpolation methods	
(十四)為取得理想的取樣結果，使用變異圖 .....	6-32
Using variogram to optimize sampling	
(十五)地理統計內插值軟體的來源 .....	6-36
Sources of software for geostatistical interpolation	

**第七章 空間中不連續實體的分析**

**Chap7、The Analysis of Discrete Entities in Space**

(一)空間分析操作的基本層級 .....	7-2
The basic classes of operations for spatial analysis	
(二)地理實體性質的屬性資料之操作 .....	7-2
Operations on the attributes of geographic entities	
(三)空間實體性質的屬性資料導出例子 .....	7-9
Examples of deriving new attributes for spatial entities	
(四)多項空間屬性資料在空間疊合時之操作 .....	7-15
Operations on attributes of multiple entities that overlap in space	
(五)由 A、B 間之距離進行緩衝區作業 .....	7-17
Operations that depend on a simple distance between A and B(buffering)	
(六)由連結進行操作 .....	7-17



Operations that depend on connectivity  
(七)資料擷取及實體性質模式的面面觀 ..... 7-19  
General aspects of data retrieval and modelling using entities

第八章 連續資料的空間分析

**Chap8、Spatial Analysis using Continuous Fields**

(一)由不關連的連續面進行空間分析的基本操作 ..... 8-1  
Basic operations for spatial analysis with discretized continuous fields  
(二)空間插入推估 ..... 8-3  
Interpolation  
(三)使用方格的空間推估 ..... 8-3  
Spatial analysis using square windows  
(四)由連續面導出一次和高次式面 ..... 8-7  
First and higher order derivatives of a continuous surface  
(五)導出面的位向關係與灌溉導路網 ..... 8-11  
Deriving surface topology and drainage networks  
(六)使用 LDD 路網進行空間分析 ..... 8-13  
Using the ldd network for spatial analysis  
(七)團塊化 ..... 8-16  
Clumping  
(八)具有或不具有不順資料的緩衝區化/擴展 ..... 8-16  
Dilation/spreading with or without friction  
(九)視覺圖、陰影凹凸起伏圖和照光圖 ..... 8-17  
Viewsheds, shaded relief, and irradiance  
(十)其他以網格為基礎的分析方法 ..... 8-21  
Other cell-based analysis operations  
(十一)使用於連續面和產出的作業方法、綜論 ..... 8-22  
Summary of operators that can be used on continuous fields and their products  
(十二)連續面的空間分析應用 ..... 8-22  
Partical applications of the spatial analysis of continuous surfaces  
(十三)結論 ..... 8-35  
Conclusions

第九章 資料誤差與品質控制

**Chap9、Errors and Quality Control**

(一)空間資料、成本和 GIS 產出的品質 ..... 9-1  
Spatial data, costs, and the quality of GIS output  
(二)空間資料的誤差來源 ..... 9-3  
Sources of errors in spatial data  
(三)影響空間資料可否信賴的因素 ..... 9-6  
Factors affecting the reliability of spatial data  
(四)有關空間實體真實性的前題假設所造成的錯誤 ..... 9-11  
Faults stemming from assumptions concerning the exactness of spatial entities

(五)結論：誤差和錯誤 ..... 9-20  
 Summary : errors and mistakes

**第十章 數值模式裏誤差的擴散**

**Chap10、Error Propagation in Numerical Modelling**

(一)數值化模型之誤差擴增之統計處理 ..... 10-1  
 Statistical approaches to error propagation in numerical madelling  
 (二)使用變異圖去進行適當取樣網路設計和減少空間解析度誤差所造成的成本 ... 10-18  
 Using the variogram to optimize sampling networks and reduce errors at the cost of  
 reduce spatial resolution  
 (三)智慧的地理資訊系統 ..... 10-20  
 Intelligent GIS

**第十一章 模糊集合與模糊地理物件**

**Chap11、Fuzzy Sets and Fuzzy Geographical Objects**

(一)不精確是一種思考方式 ..... 11-2  
 Imprecision as a way of thought  
 (二)模糊集合和模糊物件 ..... 11-3  
 Fuzzy sets and fuzzy objects  
 (三)選擇基於屬性的分類 ..... 11-5  
 Choosing the membership function :  
 1、顯示法 ..... 11-5  
 The semantic import approach  
 (1)多種模糊集合的作業 ..... 11-8  
 Operations on several fuzzy sets  
 (2)使用布林運算和模糊邏輯中選擇所造成的誤差分析 ..... 11-12  
 Error analysis of selections made using Boolean and fuzzy logic  
 (3)多邊形分界的標準方法 ..... 11-14  
 Applying the SI approach to polygon boundaries  
 (4)結合模糊界線和模糊屬性 ..... 11-17  
 Combining fuzzy boundaries and fuzzy attributes  
 (5)選擇隸屬函數 ..... 11-18  
 Choosing the membership function  
 2、模糊 K-means ..... 11-18  
 Fuzzy K-means  
 (1)疊合、融合和地理邊界 ..... 11-22  
 Class overlap, confusion, and geographical boundaries  
 (2)討論：應用模糊分類優點、缺點 ..... 11-23  
 Discussion : the advantages, disadvantages, and applications of fuzzy classification

**第十二章 地理資訊系統目前狀況與未來趨勢**

**Chap12、Current Issues and Trends in GIS**

(一)技術改變 .....	12-1
Change in technology	
(二)資料提供的改變 .....	12-3
Changes in data supply	
(三)使用者的改變 .....	12-5
Changes in the users	

附錄一 名詞解釋 (薛怡珍 整理)

**APPENDIX1 Glossary of Terms**

附錄二 地理及地理資訊系統有關之著名 WWW 網站

**APPENDIX2 A Selection of World Wide Web Geographical and GIS Servers**

附錄三 資料組例子

**APPENDIX3 Example Data Sets**

引用文獻

**References**

索引

**Index**

## 圖表目次

### 第一章

圖 1-1：空間現象的概念模式及展現 .....	1-2
圖 1-2：從 WWW 可取得的地圖例子，它是每星期更新資料一次 .....	1-4
圖 1-3：地理資訊系統主要的硬體組成 .....	1-11
圖 1-4：地理資訊系統主要軟體組成 .....	1-11
圖 1-5：資料蒐集與輸入 .....	1-12
圖 1-6：地理資訊數據資料庫的組成 .....	1-12
圖 1-7：資料的輸出 .....	1-13
圖 1-8：資料的轉換 .....	1-13
Box 1.1：使用電腦繪製地圖的理由 .....	1-6
Box 1.2：地理資訊系統現行領域 .....	1-7
Box 1.3：地理資訊系統的定義 .....	1-10
Box 1.4 .....	1-14

### 第二章

圖 2.1：處理地理資訊的所有層面包括人們的交互作用 .....	2-2
圖 2.2：不同的地理資料藉由人不同訓練不同目的而匯集 .....	2-4
圖 2.3：基本的地理原件點、線、面 .....	2-5
圖 2.4：確實的實體與連續性區域在不同資料模式的編碼 .....	2-6
圖 2.5：過程中的步驟：由真實世界現象觀察到標準化資料模式建立的步驟 ...	2-9
圖 2.6：由圖表展示資料的不同方式 .....	2-10
表 2.1：空間資料的各種資料模式 .....	2-5
表 2.2：資料型式 .....	2-10
表 2.3：歐洲土地利用不同估計方式表 .....	2-13
提綱 2.1：空間資料模式與資料結構 .....	2-2

### 第三章

圖 3.1： .....	3-10
圖 3.2： .....	3-11
圖 3.3： .....	3-12
圖 3.4： .....	3-13
圖 3.5： .....	3-16
圖 3.6： .....	3-17
圖 3.7： .....	3-19
圖 3.8： .....	3-19
圖 3.9： .....	3-20
圖 3.10： .....	3-21

圖 3.11 :	3-21
圖 3.12 :	3-22
圖 3.13 :	3-24
圖 3.14 :	3-24
圖 3.15 :	3-25
圖 3.16 :	3-25
圖 3.17 :	3-26
圖 3.18 :	3-27
圖 3.19 :	3-29
圖 3.20 :	3-30
圖 3.21 :	3-32
圖 3.22 :	3-32
圖 3.23 :	3-36
圖 3.24 :	3-30
表 3.1 : 在向量式和網格式中基本空間實體的特性	3-5
表 3.2 : 資料結構和電腦編碼	3-6
表 3.3 :	3-8
提綱 3.1 : 二進位數和算術	3-4
提綱 3.2 :	3-10
提綱 3.3 :	3-28
提綱 3.4 :	3-28

#### **第四章**

圖 4.1 : 使用全球定位系統可以得到世界上任何地點之經度、緯度及高度	4-3
圖 4.2 : 在不同系統間轉換數位資料	4-5
圖 4.3 : 擷取空間資料以建立網格資料庫之程序	4-8
圖 4.4 : 擷取空間資料以建立向量資料庫之程序	4-8
圖 4.5 : 手工輸入空間資料之數位板	4-9
圖 4.6 : 掃描網格文件之動作	4-10
圖 4.7 : 使線條變細	4-11
圖 4.8 : 建立正確向量多邊形地誌之步驟	4-13
圖 4.9 : 對 rubber sheeting 轉換向量	4-15
圖 4.10 : 大型地圖繪製機	4-17
提綱 4.1 :	4-1
提綱 4.2 :	4-2
提綱 4.3 : 網際網路及內部網路	4-6

#### **第五章**

圖 5.1：收集空間點位置的資料的不同種類的取樣網格 .....	5-3
圖 5.2： .....	5-7
圖 5.3： .....	5-9
圖 5.4：未經轉換的鋅的全球趨勢面的樣品 .....	5-11
圖 5.5：鋅至河流距離及氾濫區海拔高回歸的非獨立及獨立項數的分布直方圖 ...	5-13
圖 5.6：由氾濫區海拔級至河流距離圖籍由回歸計算鋅水準的結果 .....	5-13
圖 5.7：徐昇多邊形網及相等的 Delauney 三角形的例子 .....	5-14
圖 5.8：由徐昇多邊形的鋅 pycnophylactic 內插 .....	5-15
圖 5.9：人口數的曲線是由(a)相似的行政區(b) pycnophylacitc 內插展示 .....	5-16
圖 5.10： .....	5-18
圖 5.11：曲線規法的某些性質 .....	5-19
圖 5.12： .....	5-22
圖 5.13：在區塊直方圖上建立模糊廓的坡向資料 .....	5-22
圖 5.14：產生的數位高程模式 .....	5-24
圖 5.15：在 progressive sampling，網格密度自動調整區域複雜的地形 .....	5-25
圖 5.16：藉由具簡單的搜尋圓形的數值曲線可以建立內插表面嚴重的變形及誤差	5-26
圖 5.17： .....	5-28
表 5.1：n 個觀測值分成 k 個分類的單尾變異數分析 .....	5-6
表 5.2：氾濫分級的鋅變異數分析 .....	5-8
表 5.3：自然對數轉換的鋅的變異數分析 .....	5-8
表 5.4：線性回歸的變異分析項數 .....	5-11
表 5.5：決定內差的結果摘要 .....	5-21
Box 5.1：平均及變異數的基本原理 .....	5-6
Box 5.2：一般資料的轉換 .....	5-7
Box 5.3：在趨勢面使用高度多項式重要意義的評估 .....	5-12
Box 5.4：DEMs 的使用 .....	5-21

## 第六章

圖 6.1：區域變化原理衍生完整的空間變化 .....	6-2
圖 6.2：值距(range)、nugget 空間推估法及基質(sill)的簡單轉變變異圖的案例 ...	6-4
圖 6.3：一般變異圖模式案例 .....	6-5
圖 6.4：任意距離 $\beta$ 的半變異估計圓形插尋視窗 .....	6-6
圖 6.5：藉由次序性克立金法簡單舉例說明，預測 Z 在非取樣基地的值 .....	6-8
圖 6.6：使用次序性克立金法插入 Maas 資料組的結果 .....	6-9
圖 6.7：交叉確認是檢查變異圖預防次序性資料 .....	6-11
圖 6.8：以不同大小的區域敘述 .....	6-12
圖 6.9： .....	6-13
圖 6.10：比較次序性克立金法和使用對數轉換資料的交叉克立金法 .....	6-14

圖 6.11：在不同的方向下所計算之變異圖 .....	6-15
圖 6.12：使用各向量性變異圖之未轉換鋅資料的次序性克立金法的結果 .....	6-15
圖 6.13：個別計算氾濫頻率區域鋅質變異圖 .....	6-22
圖 6.14：使用分層變異圖預測鋅的等級 .....	6-22
圖 6.15：一般克利金法合併一個趨勢模式以估計鋅的程度 .....	6-24
圖 6.16：鋅質變異圖指示 .....	6-25
圖 6.17： .....	6-26
圖 6.18：Monte Carlo 模擬 .....	6-27
圖 6.19：藉由條件模擬的方式，zinc 層級 100 個處理的四種類型 .....	6-28
圖 6.20：條件模擬對全區域而言，能夠完成，或者對依據洪水頻率分級能夠分層 ...	6-29
圖 6.21：對不同樣本空間位置及模組大小的作用而言，等值預測變異值得等值線點的克立金變異值 .....	6-35
表 6.1：地理統計插入推估結果摘要 .....	6-30
表 6.2：不同方法標準偏差的比較 .....	6-31
表 6.3：比較內插法 .....	6-32
Box 6.1： .....	6-3
Box 6.2： .....	6-9

## 第七章

圖 7.1：資料擷取是預測資料的首要步驟 .....	7-1
圖 7.2：使用布林運算結果，其中陰影部分為“真實的 (true)” .....	7-5
圖 7.3：在屬性上單獨擷取實體時或從原有實體運算新屬性 .....	7-5
圖 7.4：擷取與重新編碼中，兩個鄰近的多邊形接受了同樣新的代碼，它們之間的邊界因而消失歸併成為一個綜合的圖 .....	7-6
圖 7.5：使用重新分類使地圖歸併 .....	7-6
圖 7.6：運算”top down”土地評估於界定玉米栽植土地適宜性的流程圖 .....	7-12
圖 7.7：包含跨時間沖蝕適宜性運算的沖蝕模式於土地評估估算之流程圖 .....	7-13
圖 7.8：為圖 7.7 流程圖評估結果，其中使用兩種不同的模式操作與使用不同的保護措施 ...	7-14
圖 7.9：多邊形疊圖使得資料庫內實體的數量遞增 .....	7-16
圖 7.10：進行多邊形疊圖時會產生大量無意義的小多邊形，必須除去它 .....	7-17
圖 7.11：產生緩衝區圍繞著預期的實體，像是點、線或多邊形，以產生出新的多邊形。這個新產生多邊形可以在原圖中進行多邊形疊圖來選擇界定的區域 ...	7-17
圖 7.12：從 A 到 B 遷徙次數的分析 .....	7-18
圖 7.13：透過運輸時間分析結果來看，市郊地區救護車可以在 9 分鐘從救護站抵達的區域。	7-19
Box 7.1：數學運算於改變屬性資料 .....	7-3
Box 7.2：運用土壤流失方程式 .....	7-14
Box 7.3：SLEMSA 沖蝕模式 .....	7-14

## 第八章

圖 8.1：網格的狀態是位置運算的功能與鄰近地區的側面相互作用 .....	8-2
圖 8.2：網格上點運算的案例 .....	8-3
圖 8.3：空間過濾的視窗運算 .....	8-5
圖 8.4：低解析過濾的平滑表面 .....	8-6
圖 8.5：在表面增加視窗的影響 .....	8-6
圖 8.6：平滑的複雜多邊形圖隨著中間平滑聚合區域，但不是減少分類級數 .....	8-6
圖 8.7：使用 edge filter 取摘取邊界 .....	8-7
圖 8.8：導出簡單的函數 .....	8-8
圖 8.9：DEM 的一次和二次導出方程式 .....	8-10
圖 8.10：區域排水坡向向量指示下坡路徑 .....	8-12
圖 8.11：由 DEM 網格導出集流網路 .....	8-14
圖 8.12：游擊流網路衍生產品 .....	8-15
圖 8.13：以 DEM 導出坡度長度 .....	8-16
圖 8.14：等方向性分散的網格 .....	8-17
圖 8.15：經過阻力函數而分散 .....	8-17
圖 8.16：從 DEM 上一個點看出的視覺圖 .....	8-18
圖 8.17：計太陽日照的照光度 .....	8-20
圖 8.18：直射的（一天當中各個時間）變化 .....	8-21
圖 8.19：暴雨在集水排放上影響視它區立所具有之集水量而定 .....	8-24
圖 8.20：運用排水系統網路來估計染物散佈 .....	8-24
圖 8.21： .....	8-26
圖 8.22：從森林中最佳木材萃取流程圖 .....	8-28
圖 8.23：木材萃取的成本 .....	8-28
圖 8.24： .....	8-29
圖 8.25：木材所處價值 > 運輸成本 .....	8-29
圖 8.26：當木材價格為 20000 個單位時，累積利潤及運輸路線達有利的狀況 ...	8-30
圖 8.27：輸入地圖把 <sup>137</sup> Cs 洪水時的重新分配模式化 .....	8-32
圖 8.28： .....	8-33
圖 8.29：散佈圖顯示了 <sup>137</sup> Cs 在河流 1km 內有高且更多變的程度 .....	8-34
圖 8.30：1993 年 7 月 Pripyat 河的氾濫 .....	8-35
表 8.1：由 DEM 所計算的屬概要及其應用 .....	8-21
表 8.2：統計上的相對於 1986 年同等的改變資料 .....	8-32
表 8.3：在每年洪氾區和 1km 區域土壤的比例 .....	8-34
表 8.4：在每年洪氾區和 1km 區域之土地使用分類比例 .....	8-34
表 8.5：GIS 分析實體與連續區域機能 .....	8-36
Box 8.1：使用 Zevenbergen 及 Thorne's 方法計算坡度 .....	8-8
Box 8.2：計算整體平均 .....	8-23



## 第九章

圖 9.1：整數算術、尺度或座標系統轉換可以使點接近邊界在內部或外部多邊形周圍	9-10
圖 9.2：當確實的投影是由網格近似值引起的誤差	9-12
圖 9.3：當網格太大而無法解決空間的細節，則混合像元問題引起	9-12
圖 9.4：網格化向量圖在 2 個網格尺度以推估網格誤差	9-14
圖 9.5：數化一條線是一個取樣的過程	9-16
圖 9.6：沿著被數化線的一個緩衝模糊帶的 Perkal's 概念	9-17
圖 9.7：在多邊形疊合後最初與衍生多邊形的測量	9-20
圖 9.8：當相同的線段被數化兩次，在圖上偽多邊形如何發生	9-20
表 9.1：在電指令長度，數化範圍與精確度之間的關係	9-11
表 9.2：在圖 9.3 使用 Switzer 方法於地圖上的結果	9-14
表 9.3：網格化方法、網格尺寸及網格化誤差的關係	9-15
表 9.4：Epsilon 誤差結果	9-17
Box 9.1：影響空間資料品質的因子	9-3
Box 9.2：由電腦指令覆蓋所產生的誤差	9-9
Box 9.3：Switzer 理論	9-14

## 第十章

圖 10.1：	10-3
圖 10.2：	10-4
圖 10.3：	10-5
圖 10.4：	10-12
圖 10.5：	10-14
圖 10.6：	10-14
圖 10.7：	10-15
圖 10.8：	10-17
圖 10.9：	10-17
圖 10.10：	10-17
圖 10.11：	10-19
圖 10.12：	10-20
圖 10.13：	10-21
Box 10.1：	10-7
Box 10.2：	10-15
Box 10.3：	10-16

## 第十一章

圖 11.1：自然語言藉由再定義及層級劃分解決模糊分類	11-3
-----------------------------	------

圖 11.2 :	11-5
圖 11.3 : 顯示方法的布林及模糊隸屬函數	11-6
圖 11.4 : 土壤結構的布林及模糊校正	11-8
圖 11.5 : 兩個重疊的模糊隸屬函數	11-9
圖 11.6 :	11-10
圖 11.7 : 對照使用常見及模糊當量的綜合土壤潛能分類的結果	11-11
圖 11.8 : 對於有興趣的資料層級使用模糊方法以連續性資料詳細保存及布林選擇損失的 連結	11-12
圖 11.9 : 量測誤差影響布林分類	11-14
圖 11.10 : 使用模糊隸屬函數描述邊界寬度	11-16
圖 11.11 : 使用模糊套疊計算屬性在模糊區域的平均權重	11-17
圖 11.12 : 傳統糊導出兩個具有逐漸、共享邊界的鄰接多邊形的土壤水分圖	11-17
圖 11.13 : 群集組成使用模糊 K-means 重複	11-19
圖 11.14 : 土壤污染資料的隸屬值的分散直方圖	11-20
圖 11.15 :	11-21
圖 11.16 : 模糊 k-means 於 Guyana 熱帶雨林樹種組成分類	11-22
圖 11.17 : 使用 SI 或模糊分類指導模糊分類的流程圖	11-25
表 11.1 : 圖籍單位的邊界寬度	11-17
表 11.2 : 模糊分級、氾濫污染的統計	11-20
表 11.3 : 模糊分級及熱帶雨林形式的統計	11-21
Box 11.1 : 計算布林及模糊圖籍的程序	11-8
Box 11.2 : 模糊集合的操作	11-9
Box 11.3 : 黏粒最大及最小的 MF 值的例子	11-9
Box 11.4 : 使用模糊 k-means 群集的演算法	11-19

## 第十二章

無

## *List of Plates*

(between pages 146 and 147)

### **Plate 1 數位化的照片地圖及 GIS(Digital Orthophoto Maps and GIS)**

- 1.1 比利時，安特衛普 (Antwerp) 的數位化照片。  
Digital orthophoto of the City of Antwerp, Belgium
- 1.2 城市的數位化照片，加入向量式線性資料顯示行政的邊界和資料，以繪製有用的地圖。  
Digital orthophoto of a city with added vector line data to indicate administrative boundaries and data for utility mapping
- 1.3 包括加入向量式線性資料的高速公路的半鄉村地區其數位化的照片。  
Digital orthophoto of semi-rural area with motorway including added vector line data
- 1.4 覆蓋一數位高層模型的數位化照片，提供地景的視覺呈現。  
Digital orthophoto draped over a digital elevation model to provide a perspective view of the landscape
- 1.5 都市景色的大尺度數位化照片，包括覆蓋向量資料在道路和建築的輪廓線作為有用的資料庫。  
Large-scale digital orthophoto of urban scene including overlaid vector data on road and building outlines for a utilities database
- 1.6 Plate 1.5 中道路和建築輪廓線的向量資料。  
Vector data for the road and building outlines for Plate 1.5

### **Plate 2 地方自治計畫，路線調查和製造銷售(Municipal planning, route finding, and marketing)**

- 2.1 Rotterdam 市大尺度網格影像的部分數位化資料庫顯示有用的、地籍的資訊，及產權的細項 (用顏色綠和深紅區分)  
Large-scale raster image of part of the digital database of the City of Rotterdam showing utilities, cadastral information, and ownership details (in colours green and magenta)
- 2.2 行進時可藉由烏鴉飛行距離及使用之道路做比較。  
A comparison of travel times via crow's flight distance and with the road
- 2.3 在荷蘭西部工作地區的可及性，以前進時的公共運輸工具來看。  
Accessibility of places of work in the west of the Netherlands in terms of travel times by public transport
- 2.4 在荷蘭西部工作地區的可及性，以前進時的私人車輛來看。  
Accessibility of places of work in the west of the Netherlands in terms of travel times by private car
- 2.5 銷售機會的空間敘述：社會團體 A 在利物浦的位置。(紅色)  
The spatial prediction of marketing opportunities: the locations of social group A in Liverpool (red)
- 2.6 銷售機會的空間敘述：在利物浦的威士忌消費 (紅色顯示威士忌消費量最大之地區)。

The spatial prediction of marketing opportunities: whisky consumption in Liverpool (red indicates areas consuming most whisky)

- 2.7 社會經濟團體“時髦的單身者”在英國的分佈。

The distribution of socio-economic group ‘stylish singles’ over the UK

- 2.8 在倫敦東北部販賣包裝的假期的機會分佈？

The distribution of the opportunities for selling packaged holidays in NE London

### Plate 3 地形和水文學(Geomorphology and Hydrology)

- 3.1 Rhine 河口水道深度之變化

Changes in channel depths in the mouth of the river Rhine

- 3.2 Rhine 河口的沈澱物及水道的地形分佈

The geomorphological distribution of sediments and channels in the mouth of the river Rhine

- 3.3 Rhine 河口的沈澱物橫切面

Cross-section through sediments of the river Rhine

- 3.4 柵欄圖 (fence-mapping) 的例子，顯示在三度空間中沈澱物組成的連續變化。

An example of fence mapping to display the continuous variation of sediment composition in three dimensions

- 3.5 空中紅外線雷射高度遙測掃描中取樣點的分佈。

The distribution of sampling points of an airborne infra-red laser altimeter scanner

- 3.6 紅外線雷射影像高度測量值在插入適當的格子後，紅色顯示高海拔，而藍色為低海拔。

Infra-red laser scan altimeter data after interpolation to a fine grid: red indicates high, and blue low elevations

- 3.7 運用 D8 演算法所產生之地區排水方向圖的例子。

Example of a local drain direction map created using the D8 algorithm

- 3.8 從它們提出之排水網路的現實呈現，累計上游的元素 (log 尺度) 覆蓋過 DEM。

Cumulative upstream elements (log scale) draped over the DEM from which they were derived to provide a perspective view of the drainage network

### Plate 4 遙測及誤差(Remote sensing and errors)

- 4.1 澳洲，水的綠色指標 (NDVI，標準的差異植栽指標)。

Greenness index (Normalized Difference Vegetation Index NDVI) for Australia, winter 1993

- 4.2 荷蘭 Dronten 有效土壤表面溫度、Minderhoudhoeve、實驗性農田的數位化紅外線影像。

Digital infra-red image of effective soil surface temperature, Minderhoudhoeve Experimental Farm, Dronten, the Netherlands

- 4.3 在加拿大亞伯達過去 periglacial 土地表面，其農田的空攝影顏色誤差。

Interpreted land cover data derived from Thematic Mapper satellite imagery draped over a DEM

- 4.4 模擬的、個別的連續表面，順時針方向從左上：沒有空間關連 (雜音)；右上：短距離的空間關連；右下：長距離之空間關連；左下：長距離的空間關連再加

上雜音。

False colour aerial photograph of farmland on former periglacial land surface, Alberta, Canada

- 4.5 當不適當的插入模式被用於從數位化的連續線中去產生 DEM，則會出現錯誤的大坡度角的虎紋。

Tiger stripes of erroneously large slope angles obtained when an inappropriate interpolation method is used to generate a DEM from digitized contour lines

- 4.6 從主題圖的衛星影像覆蓋 DEM 推論來說明土地覆蓋資料。

Simulated, discretized continuous surfaces. Clockwise from top left: no spatial correlation (noise); short distance spatial correlation; long distance spatial correlation; long distance spatial correlation with added noise

- 4.7 上：4 個模糊土壤層的連續地圖。下：地圖顯示全部 4 個層的最大組成值。

Top: Continuous mapping of four fuzzy soil classes. Bottom: Map showing the maximum membership value of all four classes

- 4.8 4 個層的最大組成值的地圖，重疊了最大不確定的（邊界）層的區域。

Map of the maximum membership values for all four classes with superposition of zones where the classes are maximally indeterminate (boundaries)

## 附錄一 名詞解釋

英文	中文	可查出處 (頁碼)			
		貓	小	數	統
1.Absolute georeference	絕對地理參數(值)				
●Geographical coordinates	地理座標			156	
2.Abstraction	抽象；抽取			2	
●Abstract space	抽象空間			2	
3.Acceptance test	接收測試	2			
4.Access time	存取時間	3			
5.Accuracy	準確度	4			
6.Addressability	可定址性	6			
7.Addressable point	可編址的點				
8.Algorithm	演算法	7		7	
9.Aliasing					
10.Altitude matrix	高度矩陣				
●Matrix	矩陣；二維陣列		136	229	
11.American National Standards Institute (ANSI)	美國國家標準協會	12	8		
12.American (National) Standard Code for Information Interchange (ASCII)	美國資訊交換標準碼	16	12		
13.Analogue/ Analog	類比	9	7		
14.Anisotropic	非等方性的				
15.Application	應用軟體				
16.Application program Application package Software package	套裝軟體	13	8, 9		
17.Arc	弧			15	
18.Archival storage					
●Archive/ File archive	歸檔；檔譜	14	9		
19.Area	區域			15	
20.Array	陣列	15	11	17	
21.Array process	陣列處理器	16			
22.Assenbler	組議程式	16			
23.Assembly language	組合語言	17	14		
24.Attribute	屬性				
25.Autocorrelation, auto covariance	自動關聯，自動共分散				
●correlation	相關；關係			49	

英文	中文	可查出處 (頁碼)			
		貓	小	數	統
●variable	變數			375	
●variance/ Var	變異數			376	
●variate	變數			376	
●variation	變化；變差；變分			376	
26.Automated cartography	機械化製圖				
27.Background processing/ mode	後部處理		17		
28.Backup	備份	22	17		
29.BASIC (Beginner's All-purpose Symbolic Instruction Code)	初學者通用符號指令碼(程式語言名稱)	24	18		
30.Baud rate	鮑率：一系列資料傳輸的速率	25			
31.Benchmark test	Benchmark test 績效測試				
32.Binary arithmetic	Binary arithmetic 二位元算數				
33.Binary coded decimal (BCD)	二進碼十進位	25	21		
34.Binary trees	二元樹	28	23, 220		
35.Bit	位元	28	24		
36.Bit map	位元圖				
●Bitmapping	位元映射	29			
37.Bit plane	Bit plane 位元面				
38.Bits per inch (BPI)	位元/吋	32	26, 58		
39.Block codes	模組代碼				
●Block	模組	30	24	32	
40.Block kriging	Block kriging				
41.Booleam operators	布林運算	32			
●Booleam algebra	布林代數		25		
●Booleam experssion	布林表示式		25		
●Booleam function	布林函數		26		
42.Break points	(中)斷點	33	26		
43.Buffering	Buffering 暫存器；緩衝區；緩衝器				
●Buffer	暫存器；緩衝區；緩衝器	34	27		
44.Bug	(程式)故障、錯誤	34	27		
45.Byte	位元組	35			
46.C <sup>++</sup>	C 語言(程式語言)	36	29		

英文	中文	可查出處 (頁碼)			
		貓	小	數	統
47.Cadastral map	地籍圖				
48.Cartography	地圖製作；製圖				
49.Cartidge tape	卡匣帶子				
●Cartidge	卡匣	39	31		
50.Cathode-ray tube (CRT)	陰極射線管		32		
51.CD-ROM	光碟唯讀記憶體	41			
52.Cell	儲存格、儲存單元；單元	41	33	46	
53.Central processing unit (CPU)	中央處理器	54	33, 48		
54.Chain	鏈結	41		49	
●Chaining	鏈接		33		
55.Chain codes	鏈結代碼				
56.Character	字元	42	34		
57.Chorochromatic maps	彩色地籍圖				
58.Choropleth map	地籍圖				
●Geological map	地質圖				環 172
●morphological map/ Geomorphological map	地形圖				環 172
59.Classification	分類				
60.Clump	叢集				
●Cluster	群集；叢集	46	36		
61.Code	代碼	46			
●coding	編碼				
62.Co-kriging	Co-kriging				
63.Colour display	色彩顯示	47			
64.Command	指令；命令	48	38		
65.Command language	命令語言	48	38		
66.Command Language Interpreter (CLI)	CLI 程式語言				
67.Compiler	編譯程式	50	40		
68.Composite map	合成圖				
69.Computer assisted/ aided cartography (CAC)	電腦輔助製圖				
●Computer-aided design (CAD)	電腦輔助設計	36	29		
●Computer aided instruction (CAI)	電腦輔助教學		30		
●Computer-aided learning (CAL)	電腦輔助學習	36	30		



英文	中文	可查出處 (頁碼)			
		貓	小	數	統
●Computer-aided management (CAM)	電腦輔助管理	37			
●Computer-aided manufacturing (CAM)	電腦輔助製造		30		
70.Computer graphics	電腦圖	50	43		
71.Computer	電腦	50	42		
72.Computing environment	電腦使用環境				
73.Conceptual model	概念模組				
74.Conditional simulation	條件模擬				
●Conditional	有條件的	51	44		
●Simulation	模擬	234	199		
75.Configuration	配置	51	44		
76.Confusion index	擾動索引				
77.Connectivity	連接(度)				
●Connector	連接器、連接符號	51			
78.Console ≡ terminal	控制台	51	44		
79.Contiguous	相鄰的				
●Contiguous graphics	相鄰圖形	52			
80.Contour	輪廓				
81.Convolution	迴旋				
82.Cross-hatching	交叉排線法				
83.Crossover point	交叉點				
84.Cross validation	交叉確認				
●validation	有效性檢查；確認、證實	269	226		
85.Cursor	游標	56	48		
86.Data analysis models	資料分析模組				
●Data	資料；數據	57	50		
87.Data link	資料連結				
●Link	連結；指標		120		
●Linkage	連結		120		
88.Data model	資料模組				
89.Data record	資料記錄				
90.Data structure	資料結構	59	55		
91.Data types	資料型態	59	56		
92.Database	資料庫	58	52		

英文	中文	可查出處 (頁碼)			
		貓	小	數	統
93.Database Management System (DBMS)	資料庫管理系統		52		
94.Debug	除錯				
●Debugging	除錯	60	56		
95.Debugger	調試器、調試程序				
96.Delaunary triangulation	Delaunary 三角測量				
97.DEM	數位高程模組				
98.Device	裝置	65			
99.Differentiable continuous surface	可微分連續表面				
100.Digital	數位的	66			
101.Digital Elevation Model (DEM)	數位高程模組				
102.Digitize	數位化	67			
103.Digitizer	數位化儀	67			
104.Discretization	Discretization				
105.Disjunctive kriging	Disjunctive kriging				
106.Disk/ disc	磁碟	70	62		
107.Diskette	磁片		63		
108.Dirichlet tessellation	笛里西列特棋盤鑲嵌				
●Dirichlet	笛里西列特			104	
109.Distributed processing	分散式處理	71	65		
110.Dot-matrix plotter	點陣式繪圖機	73	66		
111.Double precision	雙密度	73	66		
112.DPI (dots per inch)	點/吋				
113.Drift	Drift 漂流				
114.Drum scanner	滾筒式掃瞄機	75	68		
115.Edit	編輯	76			
●Editor	編輯程式	78	70		
116.Element	元素			116	
117.Ellipsoid	橢圓(球)面			118	
118.Entities	實體				
119.Exact interpolator	切確內插法				
●Interpolation	內插法；插值			193	109
120.Experimental variogram	Experimental variogram				
●variograph	變量計；變壓計				

英文	中文	可查出處 (頁碼)			
		貓	小	數	統
121.Extrpolation	外插法			132	79
122.Feature planes	特徵面				
123.Field	欄位	94	79	137	
124.File	檔案	95	79		
125.Filter	過濾程式、濾波器；濾套	95		138	
126.Finite difference modelling	有限差分模式			138	
127.Finite element modelling	有限元素模式				
128.Flatbed plotter	平板繪圖機	97	84		
129.Floating point	浮點	97	84, 85		
130.Floating point board	浮點板				
131.Floppy disk	軟碟；碟片	97	85		
132.Font	字型				
133.Format	格式 (化)	100	89		
134.FORTRAN (Formula Translation)	FORTRAN 程式語言	101	89		
135.Fourier analysis	富氏分析			143	
136.Fractal	碎維化；殘形			144	
●Fragmentation	碎片化		90		
137.Fuzzy set	模糊集合				
138.Gap	間隔	105			
139.Generalization	概化				
140.Geocoding	地理編碼				
141.Geodetical surveying	大地測量				
142.Geographical data	地理資料				
143.Geographical data model	地理資料模組				
144.Geographical information system	地理資訊系統				環 170
145.Geographical primitives	地理原始件				
146.GPS (Global Positioning Satellites)	全球衛星定位				
147.Graphics tablet	描圖板；數繪板	108	94		
148.Grey scales	灰階	109			
149.Grid	座標網格	108			
150.Grid map	座標網格圖				
151.Hardcopy	硬拷貝	112	96		
152.Hard data	硬數據				

英文	中文	可查出處 (頁碼)			
		貓	小	數	統
153.Hardware	硬體	113	96		
154.Hexadecimal system	十六進位系統				
●Hexadecimal	十六進位	114	97		
155.Hidden	隱藏的				
●Hidden line	隱藏線	116			
156.Hierarchical database structure	層級資料庫結構				
157.High-level language	高階語言	117			
158.Histogram	直方圖	118		170	
●Block diagram	區塊圖；方塊圖	30	25		
159.Hole effect	洞結果				
160.Host computer	主機		99		
161.Hypsometry	測高法				
162.Indexed files	索引檔		102		
163.Indicator kriging	指示者 kriging				
164.Inexact interpolator	不準確內插法				
165.Input	輸入	128	104		
166.Input device	輸入裝置	129	105		
167.Integer	整數	130	106	189	
168.Interactive	交互作用；互動式		108		109
169.Interface	介面	131	108		
170.Interpolation	內插法			193	19
171.Intersection	交集；交點			194	111
172.Intrinsic hypothesis	原始假設				
173.Isoline	等高線				
●Isobar	等壓線				環 209
●Isotherm	等溫線				環 209
174.Isopleth map	等值圖				
175.Isotropic	等向的				
176.Join	連結				200
177.Justification	辯護				
●Justify	對齊；校準	136	112		
178.Key file	關鍵檔				
179.Keyboard	鍵盤	138	113		
180.Kriging	Kriging				

英文	中文	可查出處 (頁碼)			
		貓	小	數	統
181.Kriging variance	Kriging variance				
182.LANDSAT	一種衛星系統				
183.Laser printer	雷射印表機	143			
184.Layer	圖層				
185.Legend	圖例				
186.Library	程式庫	147	117		
187.Light pen	光筆	147	117		
188.Line	行、線	148		216	
189.Line printer	行列式印表機	149	118		
190.Linear interpolator	線性插值			218	
191.Local drain direction (ldd)	區域排出方向				
192.Lookup table	查詢表	154	128		
193.Machine language	機器語言	158			
194.Macro	巨集	158	131		
195.Magnetic media	磁碟媒體	1311	131		
196.Mainframe	大型電腦(主機)	160	134		
197.Map	地圖；圖表	161			
●Mapping	製圖；繪圖；映射			148, 227	
198.MAP (Map Analysis Package)	製圖分析軟體				
199.Mapping unit	繪圖單元				
200.Map projection	地圖投影				
●Projection	投影			294	
201.Mass storage system	大量存儲系統		136		
202.Maximum likelihood	最大可能				
203.Mean	平均數；平均值			231	128
204.Menu	選單；表單	165	138		
205.Metadata	資料轉換				
206.Minicomputer	迷你電腦	169	141		
●Microcomputer	微電腦	165	139		
207.Model	模型；模組；模式			239	
●Mode	模式	171			
208.Modem	數據機	171	142		
209.Module	模數；模組、組件		143	239	

英文	中文	可查出處 (頁碼)			
		貓	小	數	統
210.Morton ordering	Morton ordering 莫頓序列				
211.Mouse	滑鼠	174	144		
212.Nested sampling	Nested sampling 淨抽樣；巢狀抽樣				
●Nested structure	巢狀結構		148		
213.Network	網絡；網路	177	148	160, 249	
214.Network database structure	網路資料庫結構				
215.Node	節點	178	149	251	
216.Noise	噪音；雜訊	178	151		
217.Non-differentiable continuous surface	非微分連續表面				
218.Non-linear kriging	Non-linear kriging				
219.Non-removable storage	不可移動儲存				
220.Non-transitive variogram	不可及 variogram				
221.Normalization	正規化；常態化	179		253	
222.Nugget	Nugget				
223.Numerical taxonomy	數字分類				
224.Object code	目的碼；目標碼	182	155		
225.Object-oriented database structure	(物件)目標導向資料庫結構				
226.ODYSSEY	電腦製圖程式				
227.Operating system (O/S)	作業系統	183	156		
●Operation code	操作碼		157		
●Operator	運算元；運算子	184	157		
228.Optimal estimator	Optimal estimator 理想化估計式				
229.Ordered sequential files	排序檔				
230.Ordinary kriging	Ordinary kriging				
231.Orthophotos	正確尺度像片圖				
232.Output	輸出	188	160		
233.Overlay	覆蓋；疊圖	188	160		
234.Package	套裝軟體	191	162		
235.Paint	繪塗				
●PAINT	繪圖指令	192			
236.PASCAL	PASCAL 程式語言	196	165	273	
237.ordering	Peano-Hilton 排序				

英文	中文	可查出處 (頁碼)			
		貓	小	數	統
238.Pen plotter	筆式繪圖機				
239.Performance	展現				
240.Peripheral	圓周的				
241.Photogrammetry	像片測量術				
242.Photomosaic	像片鑲嵌圖				
243.Pit	Pit				
244.Pixel	圖素	202	167		
245.Plotter	繪圖機	203	167		
246.Polygon	多邊形			283	
247.Polygon overlay and intersection	多邊形重疊部分 及交集部分				
248.Polynomial	多項式			285	
249.Precision	精密度；精確度	206	170	288	
250.Principal component analysis	主要成分分析				
251.Prism	棱柱(體)；角柱(體)			291	
252.Probability	或然率；機率			292	168
253.Probability distribution function	機率分佈函數				
254.Program	程式	211	174		
●Programming	程式設計			294	
●Procedure	程序	209	173		
●Process/ Processing	處理		173		
255.Proximity	接近				
●Propositional calculus	命題演算			295	
256.Puck	Puck				
257.Quadrant	象限			299	
258.Quadratic polynomial	二次多項式			299	
259.Quadtree	四元樹				
260.Range	值域；組距			305	
261.Raster	網格	218			
262.Raster data structure	網格資料結構				
263.Raster display	網格展示				
264.Raster map	網格圖				
265.Rasterization	網格化				
●Raster graphics	網格圖形；點掃圖形		182		

英文	中文	可查出處 (頁碼)			
		貓	小	數	統
266.Raster-to-vector conversion	網格轉化成向量				
267.Real data	真實資料				
268.Real time	真實時間				
●Real-time processing	即時處理		183		
269.Realization	真實化				
270.Record	記錄		184		
271.Redundancy	過多；冗餘		185		
272.Region	區域			309	
273.Regionalized variable	區域化變數				
274.Relational database structure	關聯式資料庫結構				
275.Relative georeferencing	相對地理參數值				
●Relation	關係(式)			310	
●Relative	相關			310	
276.Removable storage	可移動儲存				
277.Resampling	重複取樣				
278.Resolution	解析度	220	188	314	
279.Response time	回應時間	221	189		
280.R-trees	R-trees				
281.Run-length codes	運轉長度代碼				
●Run	執行；運轉；橫移距離	226	191	320	
282.Sampling	抽樣；抽樣	227	193	322	184
283.Scale	尺度				
284.Scanner	掃描器	227			
285.Scenario	情境模擬				
286.Semivariogram	Semivariogram				
287.Semivariogram model	Semivariogram 模型				
288.Sill	Sill 基石				
289.Simple kriging	Simple kriging				
290.Simulation	模擬	234	199		
291.Sink	槽	236			
292.Sliver	薄片				
293.Smoothing	平滑法			335	
294.Smoothing spline	平滑槽鍵法				
295.Spatial data model	空間資料模型				



英文	中文	可查出處 (頁碼)			
		貓	小	數	統
296.Spheroid	球狀體			342	
297.Soft data	軟體資料				
298.Software	軟體	237			
299.Source code	原始碼	239			
300.Spike	Spike				
301.Spline	Spline 槽鑑				
302.Stationarity	Stationarity 平穩				
●Stationarity point	平穩點；臨界點			344	
303.Statistical moments	Statistical moments 統計瞬間				
304.	Stereo plotter 立體繪圖機				
305.Stochastic imaging	推測想像				
306.Stochastic simulation	推測模擬				
307.Storage	儲存		204		
●Store	儲存(裝置)	246	205		
308.Stratified kriging	Stratified kriging				
309.Structure Query language (SQL)	資料查詢(程式)語言				
310.Support	支援				
311.SYMAP (Synagraphic mapping program)	SYMAP 製圖程式				
312.Syntax	語法		209	352	
313.Tablet	小型數化板				
314.Tape drive	磁帶驅動器		212		
315.Terminal	終端機	255	214		
316.Tessellation	Tessellation 棋盤鑲嵌				
317.Text editor	文字編輯程式	256	215		
318.Thematic map	主題圖				
319.Thiessen polygons	Thiessen polygons				
320.Tile	瓦格				
321.Tiling	瓦格				
322.Topographical map	地誌圖				
323.Topology	地誌學；拓樸學			359	
324.Transect	Transect 橫切				

英文	中文	可查出處 (頁碼)			
		貓	小	數	統
325.Transfer function	轉換函數				
326.Transform	轉換			360	
●Translator	翻譯程式			220	
327.Transitive variogram	Transitive variogram				
328.Trend surface	趨勢表面				
329.Triangular Irregular Network (TIN)	不規則三角網路				
330.Tuple	屬性值組；記錄				
331.Turnkey system	監管系統				
●Turnkey	轉鑰匙	262			
●Turnkey operation	監管操作		222		
332.Union	結合；合併；聯集			372	
333.Universal kriging	Universal kriging				
334.UNIX	Unix 作業系統		224		
335.Upstream element map	Upstream element map 向上元素圖				
336.Utility	公用程式	268			
●Utility program	功程式；功能程式		225		
337.Utility mapping	公共設施製圖				
338.Variogram	Variogram				
339.Vector	向量		227	377	
340.Vector data structure	向量資料結構				
341.Vectorization	向量化				
342.Vector to raster conversion	向量化轉換成網格化				
343.Viewshed	Viewshed				
344.Visual display unit (VDU)	視訊顯示器		226		
345.Voronoi polygon	Voronoi polygon				
346.Voxels	空間的三維立體單元				
347.Weighted moving average	加權移動平均數				
●Weighted mean	加權平均數			384	
348.Window	視窗	277	229		
349.Windows 95, Windows NT	視窗 95, 視窗 NT				
350.Word ≡ Byte	字	35, 74, 227			
351.Workstation	工作站	279	231		
352.Zero	零位	282		389	

英文	中文	可查出處 (頁碼)			
		貓	小	數	統
353.Zoom	縮放電子圖像				

Table 6. 3 A comparison of methods of interpolation (表 6.3 內插法比較一覽表) (可參閱講義 p.119)

說明 Method 方法	Deterministic/ stochastic 決定/ 推估 (確率)	Local/ Global 地區性/全球性	Transitions abrupt/ gradual 過渡突然/ 漸進	Exact Interpolator 內插法	Limitations of the procedure 限制	Best for 最佳	Computing load 電腦使用	Output data structure 輸出資料結構	Assumptions of interpolation model 內插值模式假設
Classification 分級; 分類	Deterministic "soft" information 確定性"軟體"資訊	Global 全球性	Abrupt if used alone 屬突然, 若單 獨使用	No 無	Delineation of areas and classes may be subjective. Error assessment limited to with-class standard deviations. 區域及層級的描繪可能主 觀。錯誤評估局限於有層級 標準偏差。	Quick assessments when data are sparse. Removing systematic differences before continuous interpolation from data points. 當資料是很少時, 快速評估。在連 續內插值前, 從資料點中移除系統 差別。	Small 小	Classified polygons 已分類的多邊形	Homogeneity within boundaries 邊界內均質
Trend surface 趨勢表面	Essentially deterministic (empirical) 本質上可決定的 (經 驗法則)	Global 全球性	Gradual 漸進式	No 無	Physical meaning of trend may be unclear. Outliers and edge effects may distort surface. Error assessment limited to goodness of fit. 趨勢實質意義可能不清楚。 外層和邊元效應可能使表面 變形。錯誤評估局限於適合 處。	Quick assessment and removal of spatial trends. 空間趨勢的快速評估及移除。	Small 小	Continuous, gridded surface 連續網格表面	Phenomenological explanation of trend, normally distribute data 正常趨勢分佈資料的 現象解釋。 (常態分佈資料)
Regression models 回歸模式	Essentially deterministic (empirical-statistical) 本質上可決定的 (經 驗統計)	Global with local refinements 視當地 (地區 性) 而修正	Gradual if inputs have gradual variation 屬漸進式, 若 輸入有漸進變 化	No 無	Result depends on the fit of the regression model and the quality and detail of the input data surfaces. Error assessment possible if input errors are known. 結果取決於回歸模型的適宜 性和輸入資料表面的質及細 節。錯誤評估可能發生, 若 可得知輸入錯誤處。	Simple numerical modeling of expensive data when better methods are not available or budgets are limited. 當更好的方法無法獲得或預算受限 時, 昂貴資料的簡單數字模式化。	Small 小	Polygons or continuous, gridded surface 多邊形或連續網格 表面	Phenomenological explanation of regression model 回歸模型的現象解 釋。
Thiessen polygons (proximal mapping) 徐昇多邊形 (近似製圖法)	Deterministic 可確定性	Local 地區性	Abrupt 突然式	Yes 有	No error assessment, only one data point per polygon. Tessellation pattern depends on distribution of data. 沒有錯誤評估, 僅是每一多 邊形的一個資料點。 棋盤 鑲嵌模式取決於資料的分 佈。	Nominal data from point observations. 從點觀測中的名義資料 (Nominal data)。	Small 小	Polygons or gridded surface 多邊形或網格表面	Best local predictor is nearest data point 最佳地區性預測器是 最接近的資料點。

(續下表)

Table 6. 3 A comparison of methods of interpolation (表 6.3 內插法比較一覽表) (可參閱講義 p.119) (續一)

說明 Method 方法	Deterministic/ stochastic 決定/ 推估 (確率)	Local/ Global 地區性/全球性	Transitions abrupt/ gradual 過渡突然/ 漸進	Exact Interpolator 內插法	Limitations of the procedure 限制	Best for 最佳	Computing load 電腦使用	Output data structure 輸出資料結構	Assumptions of interpolation model 內插值模式假設
Pycnophylactic interpolation Pycnophylactic 內插法	Deterministic 可確定性	Local 地區性	Gradual 漸進式	No, but conserves volumes 無，但有轉化量	Data inputs are counts or densities 資料輸入是計數 (用算的) 或密集(度)	Transforming step-wise patterns or populations counts to continuous surfaces. 轉變步驟-配對模式或者族群連續表 面計算。	Small - moderate 小-中等	Gridded surface or contour 網格表面或等高線	Continuous, smooth variation is better than ad hoc areas 連續平滑變化比特別的 (ad hoc) 區域好
Linear Interpolation 線性內插法	Deterministic 可確定性	Local 地區性	Gradual 漸進式	Yes 有	No error assessment 沒有錯誤評估	Interpolating from point data when data densities are high as in convert- ing gridded data from one projection to another. 當資料密度很高時，如同轉換網格 資料從一個投影到另一個投影，從 點資料內插資料密集(度)。	Small 小	Gridded surface 網格表面	Data densities are so large that linear approximation is no problem 資料密集(度)很大， 所以線性近似值 (inear approximation) 可 獲得。
Moving averages and inverse distance weighting 移動加權平均法	Deterministic 可確定性	Local 地區性	Gradual 漸進式	Not with regular smoothing window, but can be forced	No error assessment. Results depend on size of search window and choice of weighting parameter. Poor choice of window can give artifacts when used with high data densities such as digitized contours 沒有錯誤推估。結果取決於 權重參數及視窗搜尋的大 小。當使用如數化等高線的 高資料密度時，視窗的選擇 能夠給人為的製圖成品。	Quick interpolation from sparse data on regular grid or irregularly spaced samples. 在有規律的網格或不規則的空間樣 本上，從很少資料中快速內插值。	Small 小	Gridded surface, contour lines 網格表面，等高線	Underlying surface is smooth 下表面是平滑的
Thin plate splines Thin plate 曲線法	Deterministic with local stochastic component 隨地區性推估的確定 性 (機率性值的分佈 值)	Local 地區性	Gradual 漸進式	Yes, within smoothing limits 有，但受平滑法 限制	Goodness of fit possible, but within the assumptions that the fitted surface is perfectly smooth. 適合可能的優勢，但必須在 合適表面完美地平滑的假設 中。	Quick interpolation (univariate or multivariate) of digital elevation data and related attributes to create DEMs from moderately detailed data. 數化高程資料的快速內插法(單一變 化或多重變化)及從適當地細部資料 關聯屬性建立 DEMs。	Small 小	Gridded surface, contour lines 網格表面，等高線	Underlying surface is smooth everywhere 下表面全是平滑的。

(續下表)

Table 6. 3 A comparison of methods of interpolation (表 6.3 內插法比較一覽表) (可參閱講義 p.119) (續二)

Method 方法	說明 Deterministic/ stochastic 決定/ 推估 (確率)	Local/ Global 地區性/全球性	Transitions abrupt/ gradual 過渡突然/ 漸進	Exact Interpolator 內插法	Limitations of the procedure 限制	Best for 最佳	Computing load 電腦使用	Output data structure 輸出資料結構	Assumptions of interpolation model 內插值模式假設
Kriging 克立金法	Stochastic 推估性	Local with global variograms. Local with local variograms when stratified. Local with global trends. 屬地區性，但視全球變異圖而定。屬地區性，但當分層時，視全球變異圖而定。屬地區性，但視全球趨勢而定。	Gradual 漸進式	Yes 有	Error assessment depends on variogram and distribution of data points and size of interpolated blocks. Requires care when modeling spatial correlation structures. 錯誤推估取決於資料點的變異圖和分佈以及推估模組的大小。當模式化空間關聯結構時需要小心。	When data are sufficient to compute variograms, kriging provides a good interpolator for sparse data. Binary and nominal data can be interpolated with indicator kriging. Soft information can also be incorporated as trends or stratification. Multivariate data can be interpolated with co-kriging. 當資料足夠去計算變異圖，克立金為很少的資料提供一個好的內插值計算法。二位元及名義資料能夠以使用指示克立金的內插法被計算之。軟體資訊也能夠合併為趨勢或分層資料。多變量資料也能夠使用 co-kriging 內插法計算之。	Moderate 中等	Gridded surfaces 網格表面	Interpolated surface is smooth. Statistical stationary and the intrinsic hypothesis 內插值表面是平滑得。統計固定 (Statistical stationary) 和內在假設。
Conditional simulation 條件模擬法	Stochastic 推估性	Local with global variograms. Local with local variograms when stratified. Local with global trends. 屬地區性，但視全球變異圖而定。屬地區性，但當分層時，視全球變異圖而定。屬地區性，但視全球趨勢而定。	Irregular 無規律性；不規則	No 無	Understanding of underlying stochastic process and models is necessary 推估過程和模式的理解是必要。	Provides an excellent estimate of the range of possible values of an attribute at unsampled locations that are necessary for Monte Carlo analysis of numerical models, also for error assessments that do not depend on distribution of the data but on local values. 在對於數字模型的蒙特卡羅分析 (Monte Carlo analysis of numerical models：一種確率模擬的模式) 必要的 非取樣位置提供一個屬性可能值範圍的極好推估，也為錯誤評估不僅僅取決於資料的分佈也包括地區性價值 (local values)。	Moderate- Heavy 中等-重要的	Gridded surfaces 網格表面	Statistical stationary and the intrinsic hypothesis 統計固定 (Statistical stationary) 和內在假設。

# **Principles of Geographical Information System**

Peter A. Burrough and Rachael A. McDonnell 1998  
Oxford University Press 333pp.

## **CONTENT**

Chap1 、 Geographical Information : Society, Science, and, Systems

Chap2 、 Data Models and Axioms : Formal Abstractions of Reality

Chap3 、 Geographical Data in the Computer

Chap4 、 Data Input, Verification, Storage, and Output

Chap5 、 Creating Continuous Surfaces from Point Data

Chap6 、 Optimal Interpolation using Geostatistics

Chap7 、 The Analysis of Discrete Entities in Space

Chap8 、 Spatial Analysis using Continuous Fields

Chap9 、 Errors and Quality Control

Chap10 、 Error Propagation in Numerical Modelling

Chap11 、 Fuzzy Sets and Fuzzy Geographical Objects

Chap12 、 Current Issues and Trends in GIS

APPENDIX1 Glossary of Terms

APPENDIX2 A Selection of World Wide Web Geographical and GIS Servers

APPENDIX3 Example Data Sets

References

## Index



## **Chap1 、 Geographical Information : Society, Science, and, Systems**

Geographical Information Society : a brief history

Geographical Information Systems

The structure of this book

Definitions of GIS

The components of a geographical information system

## **Chap2 、 Data Models and Axioms : Formal Abstractions of Reality**

Conceptual models of real world geographical phenomena

Geographical data models and geographical data primitives

The display of geographical primitives using vector and raster approaches

Data types

Axioms and procedures for handling data in information system

Data modelling and spatial analysis

Examples of the use of data models

Summary : entities or fields

## **Chap3 、 Geographical Data in the Computer**

Data in the Computer

Coding the basic data models for input to the computer

Database structures : data organization in the computer

File and data access

Database structures and database management

Database management system

Choosing the most appropriate database structure

Database structures for representing the geometry of spatial phenomena

Data structures for spatial data : the choice between raster and vector

Database storage of vector and raster data structures

Object-oriented database structures : unifying attribute and geometric storage

The debate on relational-hybrid GIS versus Object Orientation

## **Chap4 · Data Input, Verification, Storage, and Output**

Source of Geographical Data

Geographical Data collectors and providers

Acquiring digital datasets from a data supplier

Creating digital data sets by manual input

Data presentation

Data updating

Data storage

## **Chap5 · Creating Continuous Surfaces from Point Data**

Interpolation : what it is and why it is necessary

The visualization of continuous surfaces

The rationale behind interpolation

Data sources for interpolation

Methods for interpolation

Global interpolation

Global prediction using classification models

Global interpolation using trend surfaces

Spatial prediction using global regression on cheap-to-measure attributes

Local, deterministic methods for interpolation

Nearest neighbors : Thiessen (Dirichlet/Voronoi) polygons

Linear interpolators : inverse distance interpolation

Splines

A comparison of simple global and local methods

Digital elevation models as a special case of continuous surfaces created by interpolation

## **Chap6 · Optimal Interpolation using Geostatistics**

A brief introduction to regionalized variable theory and kriging

Fitting variogram models

Using the variogram for spatial analysis

Isotropic and anisotropic variation

Variograms showing spatial variation at several scales

Spatial variation within different cover classes

Using the variogram for interpolation : ordinary kriging

Using kriging to validate the variogram model

Other forms of kriging

Kriging using extra information

Probabilistic kriging

Simulation

The relative merits of different interpolation methods

Using variograms to optimize sampling

Sources of software for geostatistical interpolation

## **Chap7 · The Analysis of Discrete Entities in Space**

The basic classes of operations for spatial analysis

Operations on the attributes of geographic entities

Examples of deriving new attributes for spatial entities

Operations on attributes of multiple entities that overlap in space

Operations that depend on a simple distance between A and B(buffering)

Operations that depend on connectivity

General aspects of data retrieval and modelling using entities

## **Chap8 · Spatial Analysis using Continuous Fields**

Basic operations for spatial analysis with discretized continuous fields

Interpolation

Spatial analysis using square windows

First and higher order derivatives of a continuous surface

Deriving surface topology and drainage networks

Using the D8 network for spatial analysis

Clumping

Dilation/spreading with or without friction

Viewsheds, shaded relief, and irradiance

Other cell-based analysis operations

Summary of operators that can be used on continuous fields and their products

Particular applications of the spatial analysis of continuous surfaces

Conclusions

## **Chap9 · Errors and Quality Control**

Spatial data, costs, and the quality of GIS output

Sources of errors in spatial data

Factors affecting the reliability of spatial data

Faults stemming from assumptions concerning the exactness of spatial entities

Summary : errors and mistakes

## **Chap10 · Error Propagation in Numerical Modelling**

Statistical approaches to error propagation in numerical modelling

Using the variogram to optimize sampling networks and reduce errors at the cost of reduce spatial resolution

Intelligent GIS

## **Chap11 、 Fuzzy Sets and Fuzzy Geographical Objects**

Imprecision as a way of thought

Fuzzy sets and fuzzy objects

Choosing the membership function :

1 、 The semantic import approach

Operations on several fuzzy sets

Error analysis of selections made using Boolean and fuzzy logic

Applying the SI approach to polygon boundaries

Combining fuzzy boundaries and fuzzy attributes

Choosing the membership function :

2 、 Fuzzy *k*-means

Class overlap, confusion, and geographical boundaries

Discussion : the advantages, disadvantages, and applications of fuzzy classification

## **Chap12 、 Current Issues and Trends in GIS**

Change in technology

Changes in data supply

Changes in the users

# 第一章 地理資訊：社會、科學與系統

## *Geographical Information: Society, Science, and Systems*

### 地理資訊科學：簡史

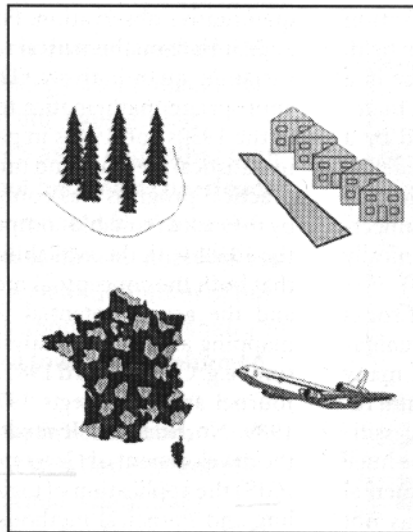
隨著世界人口的成長，對於地球上的土地、空氣、水、原料（raw materials）等資源之需求日益增加。以往過剩的人口可以透過遷徙至人跡稀少的殖民區域或是因瘟疫與戰爭而減少；像這樣的選擇是既無法長久，也不能在二十世紀的文明社會中被允許的。同時，人類社會變得更組織化，除了滿足使人們安定的基本需求之土地與自然資源外，還必須支持日益繁雜的社經活動行為模式。在面臨資源與土地需求與日俱增的壓力時，除了考量人類需求外，也應一併考量其它生物，有組織的協調資源分配。達成此一目的不僅是要瞭解資源的空間與時間形式，還需更進一步洞悉時間與空間的程序以管理它們效益。歷史多次證明了再生資源的減少與可承載的能力降低，意味著也許人口過剩與污染為生活帶來壓力，從此一觀點來看，文明生活是失敗的。

在地球土地上很多不同族群的人們跟自然資源每天都會發生激烈的衝突，人們並不總是訴諸於暴力解決土地的爭論，人們曾經建立基本的計算與數學能力去確保動物、農作物以及有價值的物件的交易與繼承的公平性。他們有能力對這片土地進行相似觀念的測量與分配。在國家法規中，規則與法規已被建立來管理與施行協議於區域和標誌有雨林或礦區的現代地籍系統之土地層級。

有關於土地使用與所有權屬的法規和規則的建立，必須制訂有別於個人或團體人類記憶的執行與協調記錄。從早期文明社會至近代期間，空間資料被地圖製造商（map-makers）和製圖者藉由航海家、地理學家與測量人員的蒐集，記錄於法典、畫刊等形式。在羅馬時代，agrimensores 或土地測量員在政府組織中佔了重要的部分，而且他們工作的成果也許在今日的歐洲地景中仍然可看見其遺跡形式。

測量技術與地圖製造隨著羅馬帝國衰退而沒落，直至文藝復興時代因地理學的探索而復甦。十七世紀許多著名製圖者，如麥卡托（Mercator），便已證明除了使用數理投影系統與精確的同位設定來增加土地區域定位與測量法的可靠性外，還需將自然事件與人為開拓地的空間分佈制訂成為標準一致的空間模型，那便是航海學、航線發現與軍事戰略價值所在。到了十八世紀，歐洲許多國家意識到有系統的繪製它們領土地圖之價值。地理資訊協會（Geographical Information Society）首先被建立於國家組織中，其主要的工作在於建立全國地籍測量與地形圖的繪製。這些高度訓練的學會至今仍然提供地表特徵或地形的空間分佈於地圖形式中。在這兩百年期間許多獨特風格的地圖已經被建立，但是那些過去傳統高製圖標準的方式仍沿用至今。

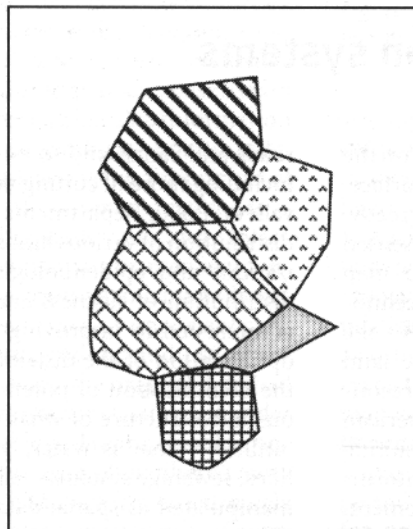
製圖科技—測地學量測、空照圖與製圖學，建立了有力的工具來精確的紀錄與描繪定義明確的自然與擬人論的現象之位置與特性。地理資訊基本單元很久以前就已經被確定，現代製圖者利用描點與線來描繪實際世界物體或行政上的單位以說明它們的重要屬性。區域屬性是由透過同型的色彩或濃淡表現，然而色彩濃淡的漸進變化有時應用於表示地表某些資產的改變；像是險峻的斜坡，深的湖泊或是從森林到沼澤的地表覆蓋變化。銅製刻蝕的印刷版技術自十七世紀前開始利用，依賴其精確與清晰輪廓的印刷特性來增強地理上的記號。今日，許多地理資訊與區位有關而且彼此相互影響，物體在空間中精確的位置如森林中的樹，街道中的房子，經由飛行至目的地或是行政上的單元如法國的 Dependances（如圖 1.1a）。



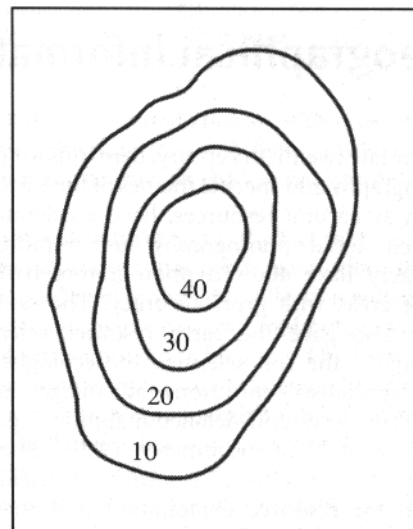
a) Objects in space



b) Continuous variation over space



c) Choropleth map



d) Isoline map

圖 1.1：空間現象的概念模式及展現（原文 p.3）

在科學研究地表之前，不同種類之屬性需要被圖形化。研究大地和它的自然資源—地理上的大地測量學、地質學、地形學、土壤學、生態學和土地等科學在十九世紀時便開始並延續至今。地形圖也許被視為一般用途是因為它未能提供滿足任何細項目的用途，礫石形式、土系和土地使用的分佈圖是使用於更狹隘的用途。這些特殊用途的地圖通常被稱為“主題的 (thematic)”地圖，因為它們所包含的資訊是有關於單一的目的或主題。然而，這種現象在各地是一樣的，製圖員無法紀錄和處理符合各種需求而產生的龐大資料圖籍。透過定性的描述岩石、土壤與典型的圖形變異，製圖員設定了“代表性的 (representative)”分類層級。把以點為基礎關連的分類層級歸類，如土地面積，適宜的製圖單元，用不同的顏色或濃淡繪製於地圖上。地圖依設定呈現出不同色彩與描述意義的區域而容易分辨，像一張複合色彩地圖 (chorochromatic)：它們是一種地誌圖的形式或是表現出等值的區域。製作單一主題資料的圖較描繪多項主題資料的地誌圖 (choropleth map) 容易。用這樣的圖可以表示出連續的變化，但是

複雜的主題事件迫使人必須刻意的從等值的區域劃分出所要辨識的對象，像是樹木、房舍或是設有圍欄的小牧場。

“主題圖 (thematic map)” 這個用詞是一個非常廣泛與鬆散的法，它除了表示出一般主題像是土壤、地質外，還表現出其單一屬性的分佈情形。單一屬性的價值在於可以於地圖中與空間實體的分佈做一完整連結，例如顯示人口統計資訊於國家的行政單位，行政區域或是地方管理機構層級。

可靠的屬性，像是海拔高度，實驗用田地上的土壤酸鹼值，都市中發生疾病區域的範圍變化或是用連續的等值線表示大氣壓力於氣象圖上。這些變化通常用慣用的等值線 (isoline) 或等高線表示之，這些線是透過等值的點連接而成或是透過屬性設定來應用單色灰階或色彩色階。

第一張岩石或土壤、植物群落或人的空間分佈圖是定性的。如同許多新的學科，許多研究首先致力於項目清單、分類與紀錄。因應許多環境研究的需要，定性模型的分類與製圖很難避免要加入龐大複雜的測量資料。量化的記述除了受制於資料量外，還缺少量化的評論結果。再者，還缺少適當的數理工具來描述空間量的變化。首先應用數值處理空間問題是 1930 年，接著 1940 年建立了相似的統計模型與時間序列分析。然而，缺乏適當的估算工具，有效的實際發展全然受到阻礙。直到 1960 年有數字顯示的電腦供使用；空間分析用的概念模型，數值主題圖與空間分析已發展到興盛期。今日，感謝地理資訊系統的建立，如本書所闡釋的適切的應用數值模型與統計理論於空間資料。

## 地理資訊系統

二十世紀末，地形學的資料和地球表面的各種主題（例如自然資源）兩方面的需求大大的增加，立體航空攝影 (stereo aerial photography) 和遙感影像 (remotely sensed imagery) 這兩種技術，可讓照相測量學家更精確地繪製大區域的地圖。這種相同的技術也帶給地球資源方面的科學家們——地質學家、土壤學家、生態學家、土地使用專家——便於勘查及繪製詳細地圖（例如 1960 年美國的照相測量法）。對於資源的開發與管理方面而言，有主題性的地圖是很有用的資料來源。為了配合土地能夠達到製造食物和支持人口的需求，氣候、土壤、水和應用科學這些可利用資源的土地評估研究因而興起。

對於空間方面的資料和分析之需求並非只有地球的科學家們在關心而已。都市規劃師和擁有地籍圖的機構，都需要城鎮之間的土地與資源分佈方面更詳盡的資料。土木工程師必須規劃路線、渠道以及評估工程費用，這還包括了剷除丘陵和填充凹地。警察機構需要知道各種犯罪發生的空間分佈；醫療機構和傳染病理學家會想瞭解各種疾病的分佈；而貿易則是透過開發銷路和鑑定潛在市場來提高其收益性。水、汽油、電、電話、排水系統是眾所皆知的人類需求的基本設施，都需要在地圖上紀錄、處理其空間資料。今天，可以運用政治及經濟的方法，像是地球資源的政治性機構，來促進人類的幸福。首先我們要瞭解如何處理這些資源——有多少肥沃的土地？有多少能源？有多少的人口？再來是瞭解這些土地、能源、人口是如何分佈的，分佈的地方，什麼人擁有它們，對於它們有何權力，以及我們要如何善用它們、管理它們。

對於空間的資料和分析方法的持續成長的需求可以用電腦來解決。在還無法使用電腦繪圖之前，把空間的電腦資料描繪在紙上或影片上是唯一的方法。這樣的資料利用符號的形式一點、線或區塊——用各種栩栩如生的技巧，製成密碼。以色彩或是原文碼做



成的記號是用圖例來解說，有些還附有研究報告的影印本。

這種地圖和它所附帶的研究報告是一種資料，因此在資料的收集、譯碼、和所使用的資料都是很重要的結果。第一，早期的資料為了容易瞭解和描繪，都會減少數量和分類；結果就會省略掉很多的細部。第二，需要精確地繪製地圖，而且要很清楚地表示出複雜的主題。第三，所謂完整的資料就是：在繪製大型區域方面要用大量的圖紙製作。一般的經驗就是，重要的區域經常會接近兩個區域的接合處，如果不是的話，地圖就要延展開來。第四，為了要和其他的空間資料結合，而去修正地圖上的舊資料，相當不划算也不容易。第五，印刷的地圖是靜態的，質化的文件。在主題地圖的製作上，沒有特別收集新的資料，而只以計量單位來描述數量方面的空間分析是很難的。

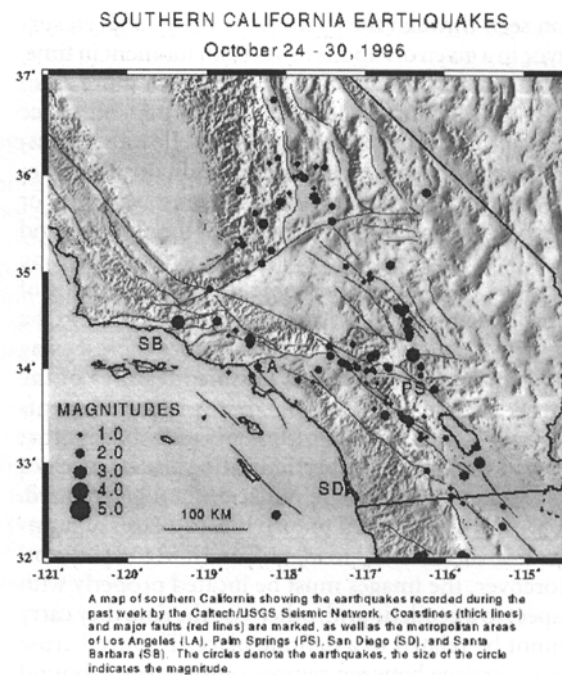


圖 1.2：從 WWW 可取得的地圖例子，它是每星期更新資料一次（原文 p.5）

相對於這些缺點，事實上紙製的地圖是一種便宜的東西，而且不需要用現代化科技來解讀，即使收集編輯資料、繪製、印刷，還有出版物是昂貴且相當費時。然而，如果還要把特殊主題的地圖以手工重新繪製，那麼花費就頗大。當一張地圖被認為是以 20 年以上為期的，這就不是個問題；但是現在對當前空間資訊的需求是有關地球表面如何改變的方法，而不能再用以前的繪圖技法。例如，一些地圖，像是氣象圖、地震資料，或是電信公司的分佈網，都是由空間電腦資料庫每天、每小時提供，不可能用手繪完成。

本來，手繪地圖或是編入資源的地圖，是由調查員在特定的時段內將位置快拍下來的資料製作出來的。近年來，空中攝影，甚至是人造衛星影像使我們能夠觀察景觀如何隨著時間而改變；注意到沙漠化、沖蝕、森林大火、洪水、蝗蟲害、天氣系統的進行。但是這些空中產物以及空間感應器都不是地圖，而是攝影影像和磁性影帶音訊代碼。這種數字資料不只是類似於點、線和區域所表示的地球表面的特徵，而且用影像元素譯成密碼—映像點—含有反射電磁波段輻射線的兩種尺寸的矩陣組織。新的工具能把發射出的數字轉換成影像，而且製定更有意義的圖形。此外，影像要用測量方格精確地定位，否則數據資料就和指定的區域沒有關連。因為遙感探測、地域測量和

製圖學彼此結合，使得這方面的需求提升；空間資訊處理的課程和地理資訊系統這方面的繪圖工具都有很大的可能發展性。

### 處理複雜資料的方案：概化或分析 (*Alternatives for Handling Complex Data: Generalization or analysis*)

在 1960 年代和 1970 年代，出現了新的方法，將土壤和景觀系統等自然資源的資料，拿來做資源評估、土地評估、以及規劃。人們瞭解到地表的各個部分並非獨立運作的，同時也缺乏足夠的方法來處理大量的各項獨立資料，因此嘗試著以一種整合、跨領域的方法，來評估這些因子。「完形」法 ("gastalt" method) (Hills, 1961; Hopkins, 1977; Vink, 1981) 根據所研究的不同屬性之間整體互動的情形，將地表分類為「自然發生」的環境單元 ("naturally occurring" environmental units)，或是稱之為可辨識、描述、定位的構成單元 (building blocks)。在這些「自然單元」之內，應當存在著地形、地理、土壤、植被、水體等環境特徵的組合，而這些組合應該是可辨識、獨特、且獨立的。同樣的基本概念也用在整合的資源調查之中——典型的範例是澳洲的整合資源或土地系統調查 (Christian and Stewart, 1968)，或前英國土地資源管理局 (例如 Brunt, 1967)，以及荷蘭 Enschede 的 ITC。此一方法將所有空間變數 (spatial variation) 減少為幾種有限的同質類型 (homogeneous classes)，藉此來處理資料量過多的問題。而這些同質類型應該要能畫在同一張地誌圖 (choropleth) 上面。

整合調查法 (the integrated survey approach) 存在著幾個問題，人們幾乎已經放棄不用了。在許多案例中，原本假定的同質單元，對許多應用模式來說卻顯得太概略 (too general)，若要修正關於某一景觀中某些特殊面相的特定資訊，不是非常困難，就是不可能。將景觀分割成空間單位的方法，也有極大部分必須依賴個人洞見；而調查的解析層級 (level of resolution)，絕大多數和地形基本圖的尺度較為相關，而不是根據繪製某特定屬性的分佈圖所需的調查層級 (level of survey needed to map specific attributes)。因而市場需要的，仍舊是更為傳統的「單一領域調查」 (monodisciplinary surveys)，例如地理學、地形、土壤、植被、土地使用等等的調查。土地 (使用) 上逐漸增加的壓力，使得以位置特性為主的資訊 (location-specific information) 的需求大增，調查也因而傾向於收集越來越多的資料。隨之而來的，資料量增加，以及分析空間模式 (spatial pattern) 與空間作用 (spatial processes) 的新洞見，已逐漸向前邁進，轉而尋求簡單、可再生產的方法，用以結合從各個不同來源和不同解析度所得到的資料。

早期，規劃師和景觀建築師 (尤其是在美國) 瞭解到，不同「單一領域」的資源調查可以結合、整合在一起——只要將各個資源分佈圖的透明膠片在墊光桌上疊在一起，並找尋各分佈圖重疊位置相符的界線即可。此一簡單技巧的最佳代言人便是有名的景觀建築師 (Ian MaHarg, 1969)。1963 年，另一個美國建築師兼都市計畫師，Howard T. Fisher 將 Edgar M. Horwood 的觀念改進，原本 Horwood 的方法是用電腦在方格紙上印出簡單的統計數值分佈圖 (Sheehab, 1979)。Fisher 的程式 SYMAP (Synagraphic MAPping system 的縮寫) (來自希臘文 synagein，意思是放在一起 [bring together]) 將一組分析資料用的模組結合在一起，並把它們拿來操作，產生分佈圖，或是等值線圖 (isoline interpolations)。其結果的表現方法，是用線印機 (lineprinter) 重複列印字母，以產生適當的灰階效果。

Fisher 後來擔任哈佛設計學院電腦圖形實驗室的主任，而 SYMAP 是第一個用來製

圖 (mapping) 的程式，由一群熱心、國際知名、能力過人的小組成員所完成。在這些製圖程式之中，格狀製圖程式 (grid cell mapping programs) (或稱之為 raster mapping programs) GRID 和 IMGRID 讓使用者可以在電腦上執行 McHarg 用透明膠片所做的工作。自然，哈佛小組在這個領域並非孤獨的，許多工作者發展了能力相當的程式 (隨手舉些例子，如 Duffield and Coppock 1975, Steiner and Matt 1972, Fabos and Caswell 1977) 起初，這些製圖程式都只能讓使用者做 McHarg 做過的事，它們只是加快其過程，並讓其能夠重複執行 (reproducible)。然而，使用者很快就瞭解到，只要在程式寫作上多花一點心力，他們就能針對資料作其他類型的空間分析和邏輯分析 (如 Steinitz and Brown 1981)，或是生態分析 (如 Luder 1980)，而對規劃研究有所助益。在以前，這些計算若以人工進行，是極為困難的。我們注意到，在這發展過程中，雖然資料來源和術語都很不一樣，但是在這些 raster map analysis 的程式中，許多最初用來進行遙測資料的圖像分析方法，如今還是沒有改變太多。

因為 SYMAP、GRID、IMGRID、GEOMAP、MAP 和許多其他較為簡單的程式，其設計的目的都是用來加速並簡化格狀資料的分析，其結果只能使用原始的線印機圖形來顯示，所以許多地圖繪製者都拒絕承認他們所產生的結果叫做地圖。地圖繪製者在 1960 年代開始採用電腦技術，但他們大部分都限制在自動繪圖和母版製作 (masters) 的方面，直到最近才有點改善。對傳統地圖繪製者而言，新的電腦技術並沒有改變製造地圖的基本態度——高品質的紙製地圖仍舊是主要的資料儲存處 (principal data store) 和最後結果 (end-product)。然而，到了 1977 年，使用電腦製作地圖的經驗有了長足的進步，使得 Rhind (1977) 能夠提出有力的理由來支持用電腦來製作地圖。(Box 1.1)

#### Box 1.1：使用電腦繪製地圖的理由 (原文 p.7)

1. 讓地圖的製作更快速。
2. 讓地圖更便宜。
3. 讓地圖能滿足特殊的需求。
4. 讓地圖在缺乏熟練工人的情況下還能製造。
5. 能夠針對相同的資料實驗不同的表現方法。
6. 當資料數位化後，讓地圖製作與更新的工作更為便利。
7. 讓需要統計分析與疊圖兩者互動的資料分析更為便利。
8. 將從現有地圖上取得資料的需要減至最低，進而將資料品質分類化 (classification) 與概化 (generalization) 的情形減至最低。
9. 能夠製造人工不易製造的地圖，如 3D 圖或實體鏡視圖 (stereoscopic map)。
10. 讓地圖生產過程中的選擇程序 (selection procedures) 和概化程序 (generalization procedure) 能夠顯而易見，並持續進行。
11. 自動化的引介讓整個地圖製造的過程易於回顧 (檢討)，進而也容易節省經費，並進行改進。

## 電腦製圖的投資

1970 年代末期，應用電腦輔助繪圖的投資已經開始，特別是美國北方的政府與民間機構。歐洲與澳大利西亞建立了尺度與作用相似且大於美國北方的電腦輔助繪圖設施，但是到了 1970 年代末期與 1980 年代初期，由幾個國家主要研究與使用電腦輔助

繪圖，特別是瑞典、挪威、丹麥、法國、荷蘭、西德與澳大利亞。數以百計的電腦程式與系統被建立，提供政府與大學應用於各類行的地圖。

電腦輔助繪圖的採用並未如預期的立即減少成本支出，要獲得或開發新的工具通常非常的昂貴；電腦硬體是極度昂貴的；缺乏受過訓練的工作人員與許多組織不願或無法採用新的工作業務。起初，電腦輔助繪圖市場被許多製造業者認為電腦輔助設計與電腦繪圖系統是不同的，其主要投資在於需要建立軟體，這是不可能從整體市場中獲得回饋的。因此，許多昂貴系統的買主被迫雇程式人員改編特殊的系統以滿足需求。

從使用電腦製圖與空間分析的歷史可以看出，它們相同的建立了自動化資料記錄、資料分析並且提出了幾個相關的範疇，例如地籍的、地形製圖、主題圖、民用工程、地質學、地理學、水文學、空間統計學、土壤科學、測量學與航照測量術、城鄉計畫、通用網路、遙測和影像分析。軍事的應用有疊圖，甚至支援數個單一規律的範疇。因此，其成果與專業術語有許多被引用於不同土地的應用。這些成果開始是區別的，但是其緊密相關的範疇最終在普遍目的的地理資訊系統中呈現。

在 1990 年代期間，數個重要的技術與組織的發展提供了廣泛的用途與地理資訊系統的應用。首先是知覺—更多的人現在知道地理資訊系統重要的是在於能夠有效率的處理大量的空間資料，儘管許多仍需要被確信的。許多知識已經結合於如何有效率的成立電腦製圖技術與地理資訊系統計畫。其次，到了 1995 年的電腦科技已經提供強大的資料數據處理能力，資料儲存價格的合理化使得地理資訊系統可以為個人或公司組織所使用。第三、許多電腦透過網路連線，允許貴重的資料與軟體彼此共享。第四、資料庫與其它電腦的介面標準化使得龐大數量資料的處理更容易。第五、基本處理空間資料的特性已被廣泛性的接受，少數商業性的系統支配著市場，使許多原本異質領域的階層得以統一。最後一點是現行的自動控制技術。然而，令人印象深刻的惡意技術發展，使得地理資訊系統空間交互作用理論的主體發展受到限制。

推測，地理資訊系統已經在全球安裝超過 39,000 個地點，主要使用地區為北美與歐洲。許多其它國家現在（指 1997）也開始利用地理資訊系統，這些系統被使用於許多不同的領域；自從 1986 年大約有 200 筆各式觀點的主題的書面著作被制訂成數以百計的共同協定，並且有數個重要的科技與協議紀錄致力於地理資訊系統的設計、技術、使用與管理。已經有國際與國家的組織聲明空間資訊在現代社會中對於規劃、行銷與資訊社會發展的重要性。

Box 1.2：地理資訊系統現行領域（原文 p.9）

地理資訊系統的生產者、種類與應用

主要的生產者與來源

地形上的製圖：國家製圖機構、私人製圖公司、土地登記、地籍

水文上的製圖

軍事機構

遙測公司與位興機構

自然資源調查：地質學家、水文學家、物理地理學家與土壤學家、土地評價者、生態學家與生物地理學家、氣象學家與氣候學家、海洋學家

主要可利用的地理資料類型

尺度幅度寬廣的地形圖

## 衛星與航空照片

管理上的分野：人口調查區域與人口調查資料、郵遞區號區域

人口、土地覆蓋與土地幅度層級的統計資料

從市場調查的資料

資料在公共事業（天然氣、水資源、電力線路、有線電視）與它們的地點

資料在岩石、水、土壤、大氣層、生物的活動、自然的危機與災害集中的大尺度空間與時間的層級的解析度。

## 一些現行的應用

農業：從農場到國家層級的監測與管理。

環境：監測、模式與管理、土地惡化、土地評價、農業規劃、山崩、沙漠化、水質與水量、瘟疫、空氣品質、氣象與氣候模式與預報。

流行病學與健康：地方性的疾病與環境因素的關連。

森林學：管理、規劃與適當伐材與栽植。

突發事件處理：有效的利用火災、警察與救護車路線記錄；增進對於犯罪與其地點的瞭解。

航空：空中、海洋與陸地。

市場：地點位置與目標群；有效的利用好的遞送。

不動產：在法律觀點下的地籍圖，地產價值與地點的關係，保險業。

區域／地方規劃：建立計畫、成本計算、維持、管理。

道路與鐵軌：規劃與管理。

場所評價與價值：削減與增加，估算材料的數量

社會研究：分析人口的移動與發展。

旅遊：設施與屬性的位置與管理。

公共事業：位置、管理雨水資源規劃、下水道、天然氣、電力、電話、有線電視服務。

## 本書的結構

這本書不只是為了描述地理資訊系統（GIS）在被討論聲中、實業、學術理論及資源管理上的優勢及成長，也不是為了解釋這些工具應用在現代的政府及商務上多麼的有效率。其他的書籍會完成這些工作（e.g. Huxhold and Levinsohn, 1995）這些工作也一直在進行。在第一次的版本中（Burrough, 1986），這本書主要闡述了地理資訊工作的科學及技術觀點，所以地理資訊系統的使用者能了解一般的原則、機會、及記錄上的陷阱、蒐集、儲存、修正、分析並且呈現場所資訊。這目標包含了不只是基本原則還有了解這項科技的使用限制，以及在資料蒐集、概念化、分析、成果的報告時因誤差所產生的作用。引用依據古老的電腦格言「進去的是垃圾，出來的也是垃圾」，在使用地理資訊系統時，記得要保有真實。這些系統最複雜的部份，使得正確的資料有可能因為人的從容或不經意而製造出垃圾，感謝種種的色彩，極高解析度的顯示，看起來像是高品質的產品。同時，透過聰明或高品質顯示把不好的資料值取消並非難事。問題的所在並非科技，而是場所資訊的複雜度（Monmonnier, 1993）。運用電腦時犯一個更大更好的錯比不犯錯簡單多了。

這本書提供了範圍廣泛，但是在需求範圍內的，概觀掌握 GIS 場所資訊的主要觀

點。在這一章中，我們將介紹發展的歷史及 GIS 的基本架構。在第二章中，我們探討搜集和描述場所資料的問題，以及處理在空間中不連續的實體的窮境，或是處理空間屬性的連續變化。當我們不知該使用哪種模型時，該如何著手呢？以哪些原則來著手處理會更好？該如和把觀察所得的及概念化的模型相結合？在第三章中，我們將探討在第二章所使用的範例技術性地在電腦上實行的方法，用以儲存，修正和顯示。

建立空間實體的地理數據資料庫的實質觀點，將在第四章被探討，涵蓋將資料數據化、掃描、更新的觀點，以及依據範例本質將地理資訊分類，即是一般的，特別是包含了地誌及地籍圖的申請。在第 5、6 章中描述某種屬性模型的基本數據資料，處理資料輸入及矯正遙測影像，從航照圖及其他資料來源調查數字立視圖模型起源，添補資料，從定量及定性屬性的連續範疇的觀察點。

資料分析及數值的模型將在第 7、8 章被探討。隨著兩個空間資訊的範例，第 7 章著重於基本數值本質的分析—在性質上、接近關係、及網狀系統的互動上，作數值化的及統計的操作。第 8 章探討許多從連續領域取得新資訊的方法。這兩章呈現實用的例子以圖解理論的部份。

許多資訊系統做出安全上錯誤的辨識。高品質的地理的介紹及正確地制定出質問的語言，通常從有關資料的性質真正的問題，還有發生錯誤的程度，以及相關於成果的自信的程度來圖解使用者。第 9 章及第 10 章探討發生在數據資料的錯誤，在分析上也許不會有什麼影響，統計處理的空間誤差可以被使用來改進資訊充足勝過貶低它。由這個觀點來看所有的討論，將就一個場所範例或其他來看，方法或許可以結合（舉例來看，從繪出一年的雨量變化（以連續資料之模型）透過一個自制的（清楚界限國家實體的設置模型）範例的選擇使之困難因為不同的科學家及開業者常常以不同的方法形容同樣的地區。甚至如果他們停止相同的訓練，以不同的方法意圖在地理的「物件」周圍去畫一個界限，這些分界線通常是漩渦式的並且極度的不規則。有個更糟的問題，當測量得來的數據資料卻產生不同程度的解析則必須相聯合。因此地理的現象被形容為此一世代自然及社會學，而非單一的被形容。

新的理論上及實務上的研究依循模糊的邏輯並持續分類建議並非總是需要得到空間的資料數據放到一個或其他的範例上一真實世界並非真的由界限或是連續的區域劃分—但是許多現象擁有可以被處理結為合兩者觀點的屬性。先前認為處理電腦中空間模糊不清的公式是不可能的，但是第 11 章表示要如此還有一段漫漫長路。事實上，從拒絕成為狹隘的，從被分類的範例，我們發現有關場所關係可以使之比使用陳舊的方法更明確表示。這個發現提供了許多新的機會在有用的研究以及將不須花費多年來工作的應用。

## GIS 的定義

GIS 的基本工具定義是從真實的世界中，針對特定的用途，透過工具的設定來蒐集、儲存、任意更改、轉化與展示空間資料。地理的（或空間的）資料從真實世界中描述事件可從下列三個方面看出：(a)知道它們位置的座標系統；(b)屬性與位置無關（像是色彩、成本、酸鹼值、疾病發生率等）；(c)它們在空間中的相互關係與如何連結。

GIS 其他的定義（Box1.3），其焦點不是在於空間資料庫就是在於組成型態。地理資料庫的定義在於強調不同的資料組成需要處理它們的位置、屬性與位向關係而其它的大部分資訊類型只需要處理它們的本體與屬性。組織的定義強調作用的設定與人

們管理空間資料所需的工具。從此一觀點來看，地理資訊系統已經服務社會很長一段時間了。

### Box 1.3：地理資訊系統的定義（原文 p.13）

#### (a)基本工具箱定義

- 透過對工具強大功能的設定，從真實世界中蒐集、儲存、任意更改、轉化與展示空間資料。
- 提供擷取、儲存、核對、處理、分析與展示資料的地表空間查詢系統。
- 一種包括空間與非空間資料的儲存、分析與展示之資訊科技。

#### (b)資料庫的定義

- 在資料庫系統中大部份的資料是空間索引方式，依據設定的運作程序從資料庫中回答有關空間主體的問題。
- 任何現有的或電腦基本設定的用來儲存與處理地理查詢資料的程序。

#### (c)基本組織定義

- 具自動化設定的功能以提供專業與高等的功用來儲存、更正、處理與展示地理位置資料。
- 一個制度上的實體，用來整合資料庫、專業技術與長期財務支援的技術之組成結構。
- 決策支援系統需要整合空間的參考資料以解決環境問題。

## 地理資訊系統的組成要素

地理資訊系統有三個主要的構成要素—電腦硬體設備，安置應用軟體模組與熟練的人員以便與功能完備的系統相互配合。

### 電腦硬體

一般地理資訊系統的硬體組成項目，如圖 1.3。電腦具有硬碟以儲存資料與程式，但是額外的資料儲存可以經由網路、卡式數位磁帶、光碟片與其它裝置完成。數化儀和掃描器是用來轉化地圖與記錄資料呈數位化形式，這些儀器可以透過電腦程式來應用。繪圖機、印表機或是其它輸出裝置是用來呈現資料經電腦處理過後的成果。電腦間的資訊交流可用數據機經由利用光纖或電話線的區域或全球網路達成。使用者可以藉由電腦螢幕、鍵盤、滑鼠以觸控設施來操作電腦及其周邊設備。

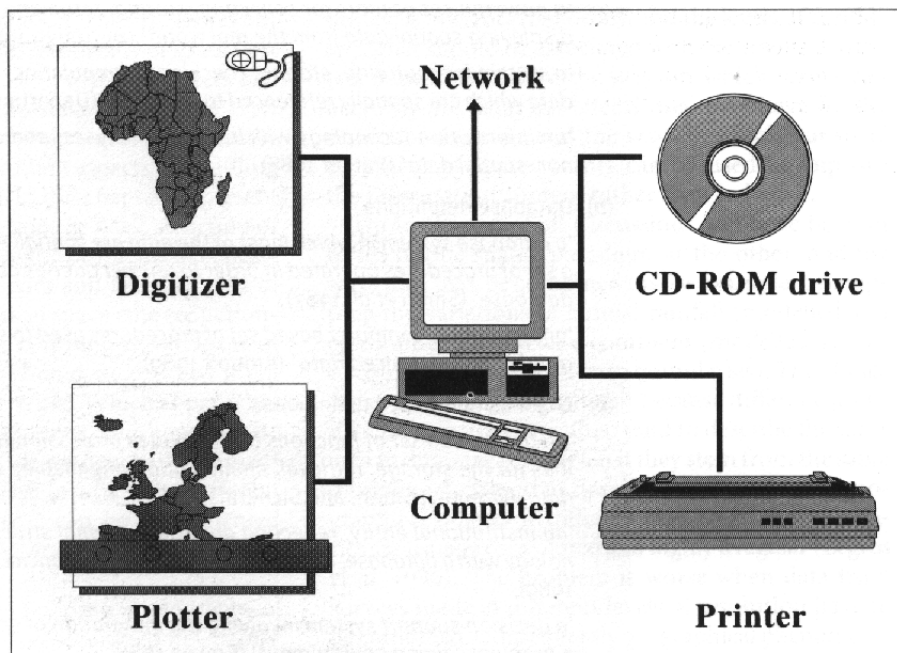


圖 1.3：地理資訊系統主要的硬體組成（原文 p.12）

### 地理資訊系統軟體

地理資訊系統軟體大致可以分為五種功能類型（圖 1.4）：

- (a) 資料輸入與確認
- (b) 資料儲存與資料庫管理
- (c) 資料輸出與呈現
- (d) 資料轉化
- (e) 與使用者互動

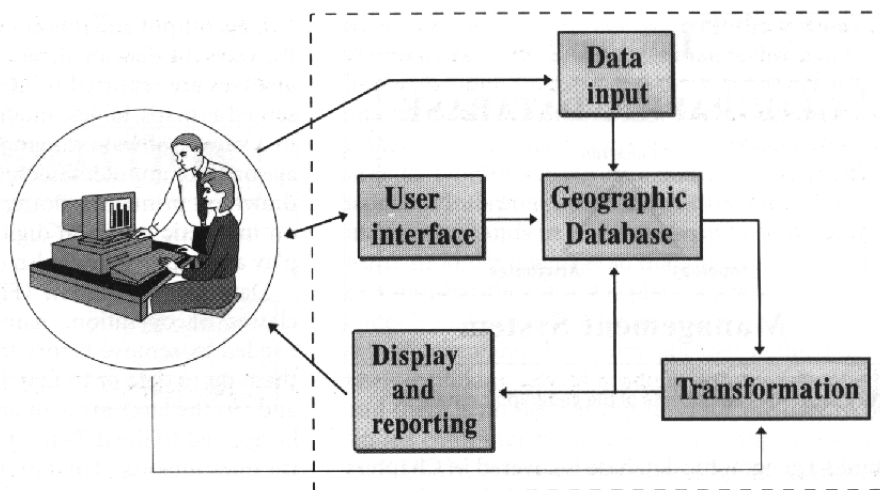


圖 1.4：地理資訊系統主要軟體組成（原文 p.13）



資料輸入可從現存地圖、田野調查、感應器（包括航空照片、人造衛星與精密記錄儀器）中獲得各方面的空間資料，並將其轉化為標準化的數位形式。其中有許多工具可供使用包括互動式螢幕與滑鼠、數化儀、word processors、試算表程式、掃描器等，並以檔案形式儲存於磁片或光碟片中。來建立一地理的資料庫在第 4、5 和 6 章中提及。

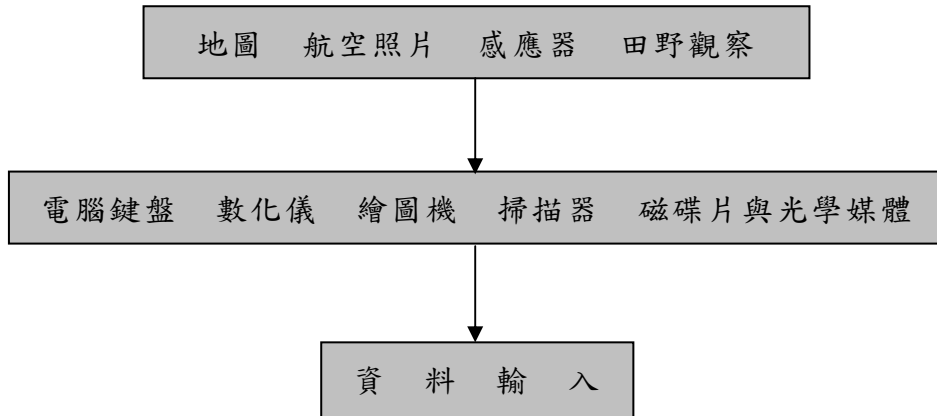


圖 1.5：資料蒐集與輸入（原文 p.13）

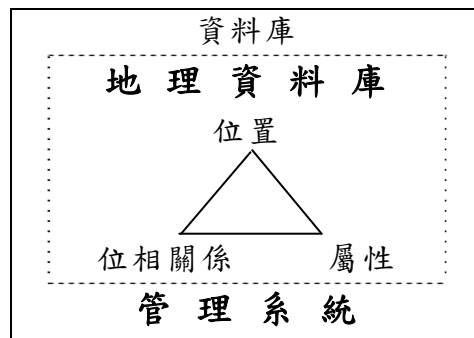


圖 1.6：地理資訊數據資料庫的組成（原文 p.14）

資料儲存和資料庫管理（表 1.6）關係到資料相關於位置、連鎖、以及地理的元素的屬性（點，線，面，以及在地球表面上更多複雜物件的屬性）被分解及組織，兩者著眼在他們必須掌握電腦及他們該如何被系統使用者意識到的方法。電腦程式被用來組織資料庫即是所謂的資料管理系統（DBMS）。資料模型（真實世界事件的形式化描述），資料庫結構，及資料庫組織的方法在第 2 及第 3 章中加以探討。

資料輸出及顯示（表 1.7）關係到資料表現的方法以及如何報告成果的分析給使用者。資料可能以地圖、表格及圖說（曲線圖 graphs 及 charts）的方式表現，以各種方法從電腦螢幕短暫的影像排列出，經由列表機或繪圖機輸出，使用吸引人的媒體以數字形式記錄資訊。表現及報告的方法，將在第 4 章加以探討。資料轉換（表取回、以及區域和周圍的計算，人們會期望以一種或其他形式的 GIS 取得各式資料。其他種類的操作也許是極端的具體應用，以及他們的組織編入特殊的 GIS 也許只能滿足特殊的那套系統的使用者。種化技巧的種類是可得，他們最理想的使用及誤用，設定簡單轉化的方法也許可以結合為了達到某種地理的及場所模型的類型，主要在第 7、8 和 11 章中完整的說明；某些場所轉化及轉化，必須確定資料的準確，我們在第 4 章中討論。1.8）包括兩種形式的作用，稱之為（a）轉化，必須從資料中轉化錯誤，或者提及他們時註明日期或是配和他們作其他日期的設定，以及（b）大量的分析方法也許為

了得到 GIS 問題的答案的資料是實用的。(Box1.4) 轉化可以操作場所的、地誌的、及非場所觀點的資訊，或分離，或結合。許多的轉化相關於單位尺寸的改變、合適於新計畫的資料、合理的資料圖 1.7 資料的輸出。

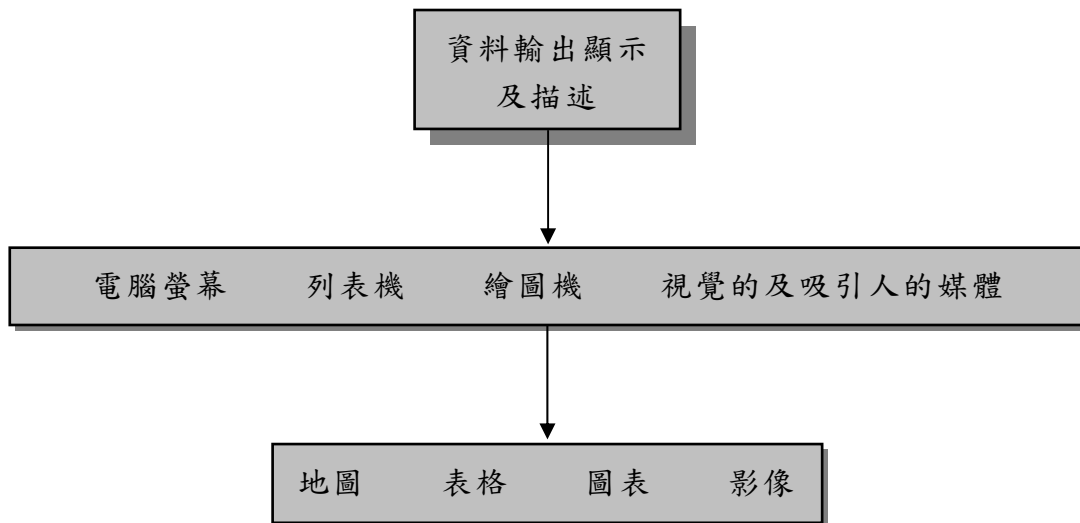


圖 1.7：資料的輸出 (原文 p.14)

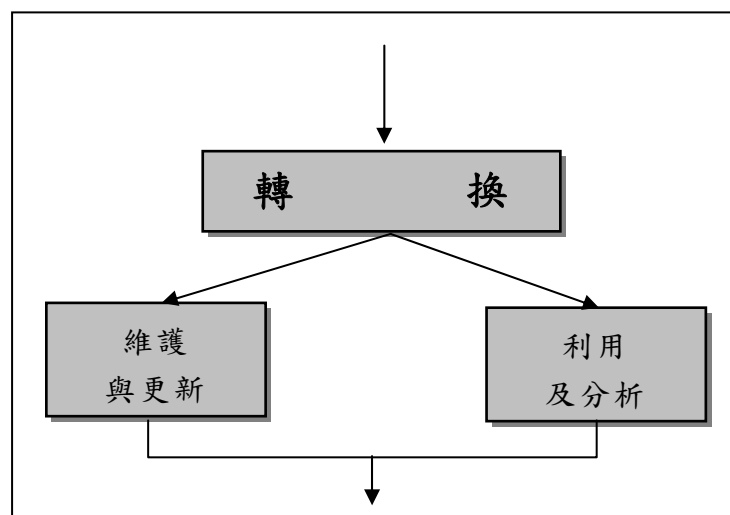


圖 1.8：資料的轉換 (原文 p.15 文)

GIS 的設計者明白使用者需要資料取回及轉化的需求受到了限制。然而大部分的系統提供了一個界面的範圍，使用者得以和系統產生互動。簡單的是 manu-driven 的命令只要用滑鼠點選即可，這個有效率的方法為一般的使用者提供複雜的機能。可選擇是為了使用者透過命令語言轉譯以簡單的命令輸入 (CLI)。以文字製造者，很多 GIS 的指令被寫成簡單明瞭的基準語言，所以使用者很直接明白他們在作什麼 (顯示土壤圖；從 X 郡標注郵遞區號)。使用者可以設計可複製的指令設定，使得經由連鎖數個基本指令實現良好定義的操作，不需要在這個層級上讓使用者而去成為電腦語言的程式設計師，就像 FORTRAN, Visule BASIC, 或是 C++。GIS 功能的主要類型，以及其簡單的表現方法 (representation)，在第 5-8 章都有所解釋，同時附上適當的例子。

## Box 1.4 (原文 p.15)

地理資訊系統的基本條件：

- (a) 顯示出 A 整個實體的位置。
- (b) 顯示出 A 整個實體的位置與 B 地方的關係。
- (c) 計算整個實體 A 與實體 B 之間存在事件的數量。
- (d) 求出函數  $f$  在  $x$  位置的值。
- (e) 估算 B 的大小 (區域, 周圍, 包含多少的計算)。
- (f) 判斷在交點的產生或是各種特殊不同的數據資料下的情況。
- (g) 在網狀或連續表面上, 沿著地面從 X 點到 Y 點, 判斷最少花費、妨害或距離的路徑。
- (h) 列出座落在  $X_1$  點、 $X_2$  點實體的屬性。
- (i) 判斷哪一個實體在有某種屬性結合的實體之後。
- (j) 將結合某種屬性的實體重新分類或重新定義顏色。
- (k) 已知在點  $X_1$ 、 $X_2$ ..... $X_n$  的 Z 值, 預算出在點  $Y_1$ 、 $Y_2$ ..... $Y_m$  的 Z 值。
- (l) 使用數值表示法從既有的屬性取得新的屬性, 或者從既有的實體得到新的實體。
- (m) 使用數據資料當作真實世界的模型, 模擬過程 P, 在已知的 S 情境下, 經過時間 T 所產生的影響。

當然, 並不是所有的操作都可以用基本的選單驅動式指令 (menu-driven command) 來執行, 使用者也可能會自己寫電腦程式來符合他們自己的需要。有些 GIS 系統提供所謂的巨集語言 (Macro Languages) — 簡化過的正式程式語言, 可以用來結合許多基本的應用程式。最新的發展是 (尤其是 GIS 的模擬模型 [simulation model]), 提供使用者高階的、相容的程式語言, 在其中能夠要有效率地撰寫模式 (參照 PCRaster—Wesseling et al. 1996, 或 MMS—Leavesley et al. 1996)。在其他情況中, GIS 被用來組合資料給某一複雜的模式, 而這模式是在系統之外用標準電腦語言寫作的 (Burrough 1996a.)。一旦執行該模式, 結果便會送回 GIS 顯示。

使用者和 GIS 之間的互動 (關於輸入查詢和資料), 以及模式的寫作 (關於資料分析), 是一直以來被忽視的部分, 直到最近才有所改善 (Hearnshaw and Unwin)。個人電腦、多重視窗系統、滑鼠或其他指示工具的出現, 讓許多人能夠使用電腦, 儘管打字仍然是許多工作必備的技巧。語音輸入、虛擬實境、多媒體, 以及聲音的輸入輸出等功能逐漸普及, 但仍然需要進一步的發展, 以完全配合 GIS。不論指令是打字輸入、按滑鼠、或是或是聲音辨識, 使用者仍然必須根據共同的規則來確保適當的系統化陳述 (proper formulation), 否則結果便是謬誤的。

### 問題：

- 1、說明現代的空間資料為何對下列活動是有價值的：市場、財產、銀行業務、旅遊、發電、森林學、農業。
- 2、為何國家組織應該免費提供空間資料於全球網際網路？審視其背後目的與動機。
- 3、探索數化製圖法、地理資訊與電腦輔助設計三者間有何差異，並說明為何此三者有不同市場功能。

## 第二章 資料模式與定理：真實的抽象表現

### *Data Models and Axioms : Formal Abstractions of Reality*

#### 前言：

當某人看見一個環境時，會藉由抽取關鍵性的答案來簡化原有的複雜相，以創造一個地區的模式，這種認知的訓練會因觀察者的背景文化及學習目的所影響。本章節測試好幾個模式的發展時期、地理資料圖或數字形式發生的過程，以及進行對地理資訊系統圖表使用在地理資料模式理論的測試是重要的事情。

兩個極端處理可察覺型式，可能是由所有物與座標系統製圖描述的一連串實體所佔據，亦或者是沒有直線邊界的連續性區域的多樣性。空間地理資料模式以概念想法為其屬性，也因此打破單位的記錄與製圖，主要的處理方式是利用好幾個點、線、面或網路來描繪景觀的多變性，採用特定模式影響了資料的類型，可能用來描述現象及空間分析，且用來解釋處理及修正空間資料的基本程序與定理。實際的例子在時常選擇與使用各種資料模式應用中便會出現。

想像當你在跟某人講電話，而他問你窗外的景致，你將會如何描述你所看到的變化？很有可能你會將景觀單位分類如建物、道路、田野、河谷或是小山丘，或亦參考地理位置如旁邊、左邊或是前面來形容事物。事實上，你已經有了景觀的概念模式，而你的經驗、文化背景及你對形容人的感覺都會影響到你對事物的描述或忽視。

當資訊需要轉換到用來描述一個較大範圍，將會變成一定的模式以確保資料的說明不會模擬兩可、及有效的傳達。本章節將會說明主要資料模式用來說明地理現象（見 Couclelis 1992, Frank *et al.* 1995 及 Campari 1993, Egenhofer & Herring 1995 及 Burrough & Frank 1996 formroe detailed discussion），本章節將做為接下來章節的基本背景，因為我們不能在電腦中儲存真實世界的現象，但卻能利用表現模式做為基礎，主要的步驟包括了人類觀察世界的行動，或是藉由工具的幫助來描述如空照圖、遙測或是當地樣本的統計，以呈現一個相似值或數據的描述如提綱 2.1、圖 2.1。最重要的第一步就是人們觀察並描述現象在空間上及時間上是固定不變亦或變化多端的，其知覺將會影響往後的分析；地理資訊系統的成功或失敗並不是決定於技術上的第一步，而是更多在固有的或其他方面的空間、空間互相作用的概念模式。

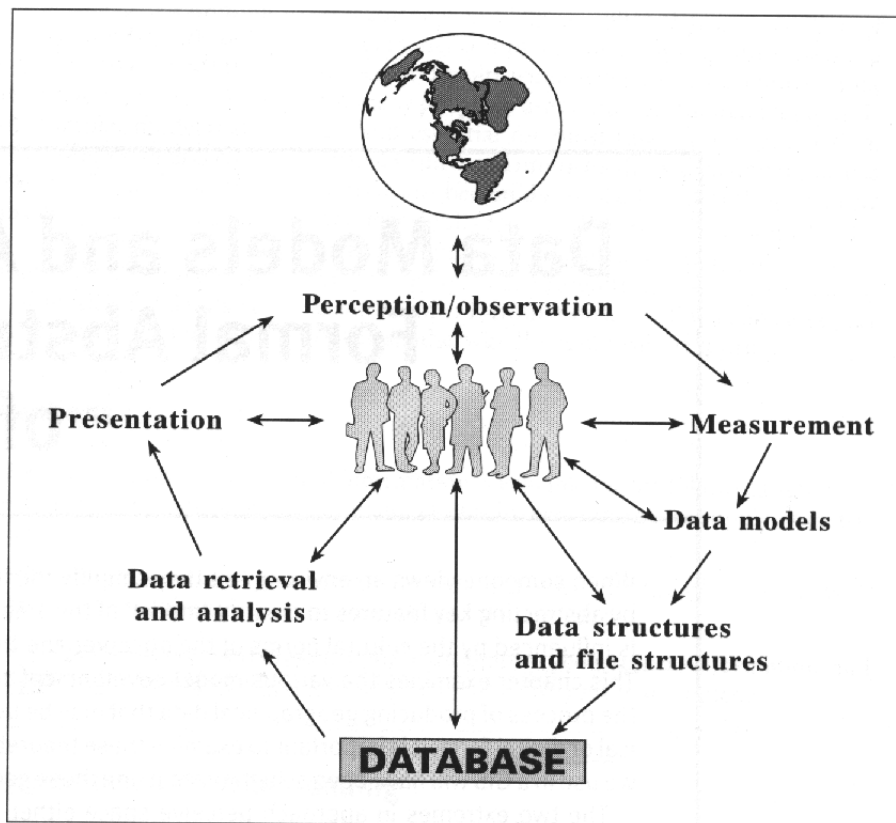


圖 2.1：處理地理資訊的所有層面包括人們的交互作用（原文 p.18）

提綱 2.1：空間資料模式與資料結構（原文 p.18）

空間資料模式與資料結構

空間資料組的近似值包括了七個模式的發展（比較 Peuquent 1984a, Rhind & Green 1988, Worboys 1995）：

- (a)實體（概念模式）
- (b)人類的概念引導一個模擬的抽象（相似值模式）
- (c)模擬抽象的形式沒有任何的代表與在執行上的限制（空間資料模式）
- (d)資料模式的描述反映了資料如何在電腦中記錄（數據資料庫模式）
- (e)一個區域結構是資料結構在電腦記憶體中特殊的描述（物理計算模式）
- (f)處理資料的定理與規則（資料操作模式）
- (g)藉以公認的規則與程序進行操作與呈現空間資料（製圖模式）

### 實際地理現象的觀念模式(Conceptual modles of real world geographical phenomena)

地理現象需要兩種電腦符號來呈現真實的世界、目前的狀況及目前的位置。在以前，現象學的概念如城鎮、河流、沖積扇、生態交錯帶及土地組成被用來建立街廓的基礎、綜合處理複雜的資訊，利用“物件”或是“實體”來描述現象較為符合正文中的標準（比較 Goudie *et al.* 1988, Johnston *et al.* 1988, Lapedes 1976, Lapidus 1987, Scott 1980, Stevens 1988, Whitten & Brooks 1972, Whittow 1984）。然而，要用這些詞彙來描述現象的許多方法是不足的，不同的詞彙會使用在不同層級的分析上，許多可察覺的地理現象藉由人類描述具體事物（如小山丘、城鎮或是湖泊）而不是明確的形式，其

區域可能會隨著時間而改變（例如參見 Burrough & Frank 1996）。

同時，建物街廓的類型用來敘述一個現象的尺度分析與其他形式是相當不同的，舉例來說，一條道路的印象傳輸來自於人造衛星的計測器，其形式可能像一條線的模式，但是建築基地的製圖可能是以區域模式來呈現其多方面的架構。根據層級界定的分類學，在其他層級的分析（如尺度），現象可能是成群的也可能是獨立分開的單位，舉例來說，國家—省州—城鎮—行政區，就是一個層級的政府管理單位，或則大部分的土壤、植生、甚至是動物分類的系統都是一個層級的分類學。

現象在空間上參考的詞彙可能就是運用到確切的幾何學或相關位置，前者使用當地或國際同等系統以定義橢圓體投影的規格系統，且其座標系與地球形式（橢圓體）在平面上是相類似的，這個座標系統可能是完全當地以 10 公尺為單位進行測量的、或則是國家網格或一個運用幾何學座標的經緯度且國際可接受的投影法。變更天然的地圖與其參考真實的地理性，不如一個絕對的空間幾何學如原始的岩石照片與 London Vnderaround 的倫廓；與其這些地圖中位置的界定參考其他空間中的特色、在實體中其他鄰近地區、方向角度，不如實際公制實際的距離。

## 空間的概念模式：實體或區域(Conceptual modes of space : entities or fields)

Is the geographic world a jig-saw puzzle of polygons, or a club-sandwich of data layers? 地理的世界是一個多邊形的智慧拼圖，或者是一個類似三明治疊圖方式的資料圖層？（Couclelis 1992）

從地理現象這些概念的想法當中，當考慮一個房間、地景甚至是大陸任何一個尺度的空間，地理現象可能是空間中所有事物呈現的形式；我們可能採用好幾種全然不同的方法來描述在地球表面上發生進行的各種事物，以下是兩種截然不同的空間處理方式：（1）觀察一個空間就像一個由其特質或所有物所描述的實體，依據且利用地理座標系統所繪製的位置；（2）或臆測一個在空間中有趣特質的變化，如某些連續的數學功能或區域。

實體，大部份的觀點是空間是由物件所佔據的，其第一個步驟是界定與承認一個實體（可能是房子、電纜、森林、河流或者是高山？）；第二個步驟是列出其特質、界定其邊界與位置。本書中我們使用“實體（entity）”這個字來代表這些事，雖然很多人使用“物件（object）”，但因我們在技術與製圖上的數據資料庫，物件導向有很特別的意義（見第三章）。在專業術語中，物件導向在電腦或電腦程式中用來參照結構資料的方法，並不完全代表完全的實體。

連續性的區域，在連續性的區域處理中用最簡單的概念模式呈現地理空間，用連續性的笛卡爾座標系統在 2D 或 3D 空間（或 4D 如果時間也包含其中的話），在空間上，其特質通常被假想為平穩且持續的變化。其特質（如大氣壓力、溫度、海洋水面海拔、土壤中泥土的成分）與其空間變化會優先被考慮；且只有當地理空間或時間的具有屬性特質值得注意時，如颶風或山峰或顯著的結果，才會被承認為“事件”（如 Laesar 颶風、Matterhon、Gulf stream 或泥土層含有豐富的 Indium 錒，Indium 錒原先被認為是小行星衝擊時期造成恐龍的滅亡）。

Objects in a vector GIS may be counted, moved about, stacked, rotated, colored, labeled, cut, sp; it, sliced, stuck together, viewed from different angles, shaded, inflated,

shrunk, stored and retrieved, and in general, handled like a variety of everyday solid objects that bear no particular relationship to geography 物件在向量地理資訊系統中可能被運算、移動、堆積、旋轉、著色、說明、刪減、分裂、切除、排列、不同角度的觀點、陰影、擴張、縮小、儲存及回復，而一般來說，處理如每天固體的物件，物件指的是敘述沒有特定關係的地理學。(Couclelis 1992)

實體在地理空間中也可能是一組最大特質等級時，一個實體模式或是一個連續區域處理的選擇是很困難的。在瑞士的例子，獨立的實體與海拔高度說明極端的差異。實際上，地理的解決方法基於數據資料庫使用者的目的而建立的，概念模式的選擇決定了資訊如何陳述。一個實體處理在山頂的選擇將是為記錄誰爬至山頂與時間提供了最佳狀態的系統，但卻不會為電腦提供邊坡坡度的資訊。坡度的計算是表面第一個導出來的地表起伏情形，如果第一個導出的坡度是 0，則視為平面，曲面都是向下的曲峰則稱為山峰。

...the phenomenon of interest is blithely bisected by the image frame ...for the mindless mechanical eye everything in the world is just another array of pixels.有趣的現象被影像傳輸（印象）給分解了....欠缺考慮的力學注視著世界上的每件事也只是像元的另一種排列。(Couclelis 1992)

由於總體過於簡單化，一個實體的選擇或區域的處理也是依賴觀察者科學及技術的訓練，訓練的焦點在於瞭解自然環境的空間過程，或許更像使用連續性區域的處理方式，當某人在一個管理情況下全然專心工作時，將會看見一個級數不同單位的區域。(圖 2.2)

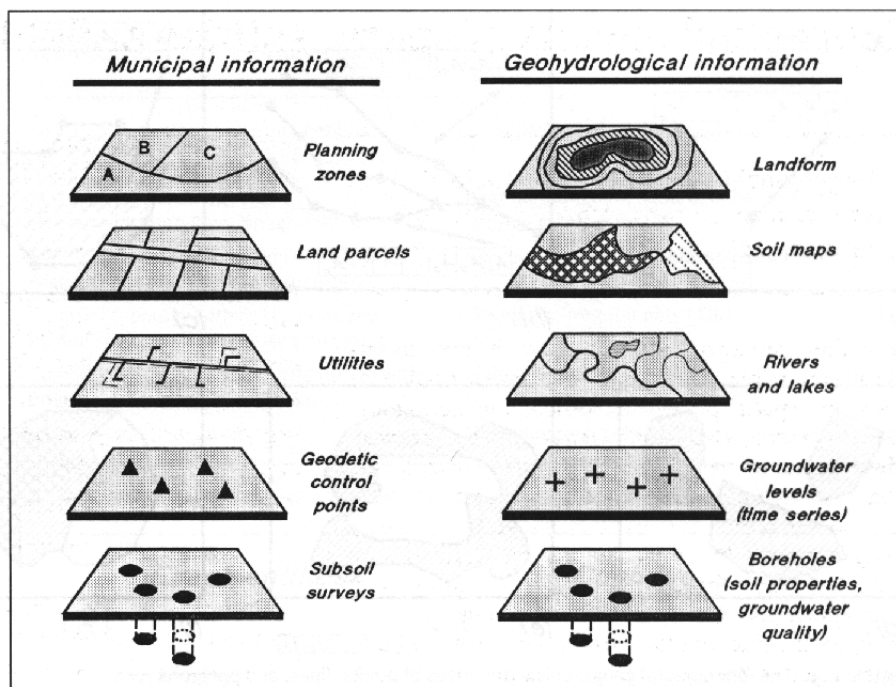


圖 2.2：不同的地理資料藉由人不同訓練不同目的而匯集 (原文 p.21)

## 地理資料模式與地理資料原件 (Geographical data modes and

### geographical data primitives)

地理資料模式是相當於概念模式的形式，被人類用來觀察地理現象（在本書中我們使用“資料類型”代表特質的數量）。他們形式化是為了分析與傳達如何將空間分成數個部份，以及假設此現象是唯一被認可的，其特質可被測量或具體說明，且其他地理座標系統是經正式鑑定的，就如資料可能經由不同的管道收集，理論上的資訊或觀察分析的層次或測量都可能是資料模式中重要的部份。

大部份人類起源發展的現象（房子、土地、政府管理單位、道路、電纜、油管在美國西部的農田耕作）可能處理在最佳實體應用的模式，最簡單與頻率最高的實際資料模式使用是基本的空間實體，可藉由特質及地理位置更進一步具體說明之。根據初期的基本地理資料之一，可再進一步地細分成“點”、“線”、及“面”（通常在地理資訊系統中使用“多邊形 (polygon)”）如圖 2.3。這些向量資料模式的基本單位在表 2.1 及圖 2.4 可見其多種的形式。方案意指規律性的多邊形鑲嵌在各組圖素當中。

隨著連續區域資料，儘管特質的變化如海拔、大氣壓力、溫度或土壤中泥土的成分，被假設為連續的 2D 或 3D 空間（也包括時間），其變化通常太複雜而無法由簡單的數學程式如多項式程式所解決。在某些樣區歸化方程式（平面傾向）可能使用在呈現大尺度的變化、不同的數值程式（見第五章），但通常區分地理空間為個別空間單位是必要的（見表 2.1 及圖 2.4）。棋盤式的結果與真實性經考慮過的分析層次其值接近，且假設運算其不同值進而可處理在連續的數學程式也可應用於各個近似值。實體與鑲嵌模式假設，現象可以用其特質與空間位置來具體說明其確實性，實際上，在某些情況下，資料模式可接受的真實描述，但某些情況下，其不確定會迫使我們選擇其中之一或其他處理方式（其空間分析的不確定性及錯誤的影響見第九、十章）。

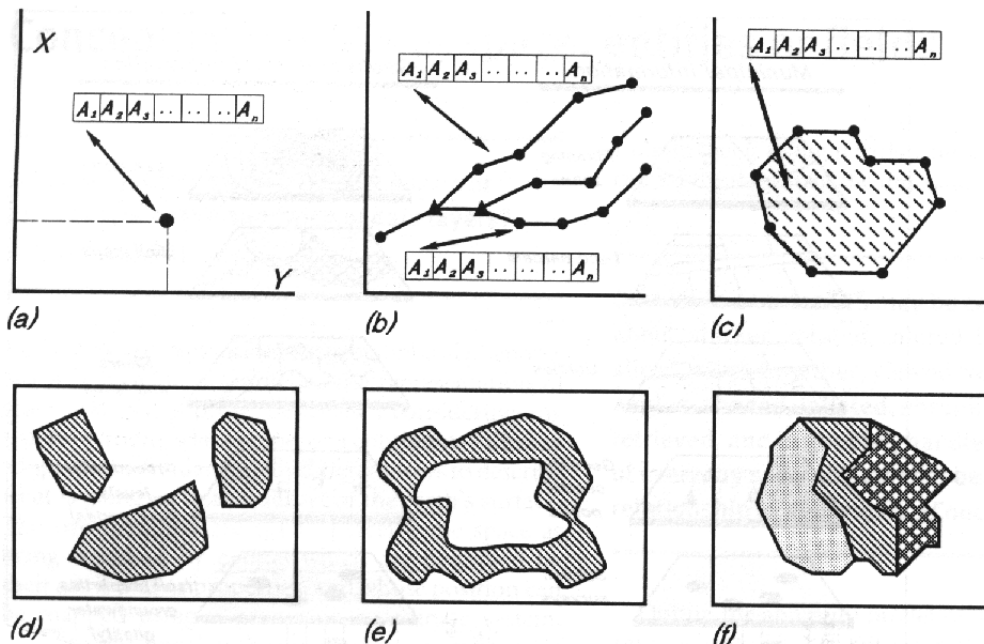


圖 2.3：基本的地理原件點、線、面（原文 p.22）

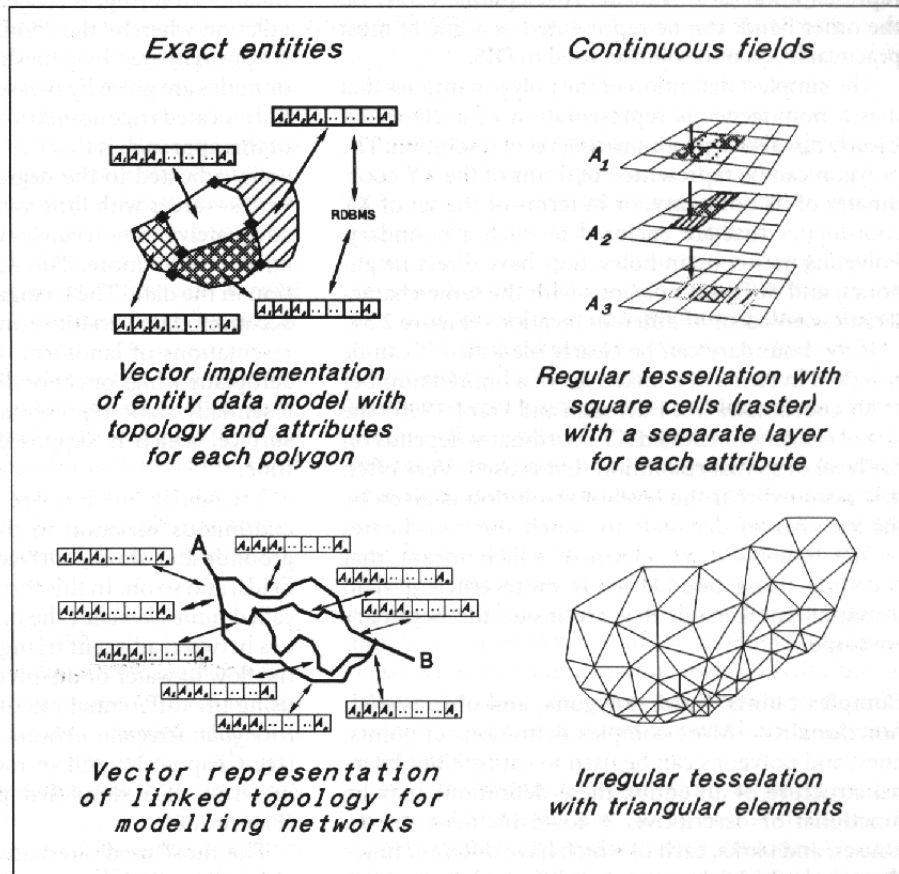
表 2.1：空間資料的各種資料模式（原文 p.23）

向量描述的確切實體	鑲嵌的連續區域
● 非拓樸學結構(沒有點與線"spaghetti")	● 規則的三角形、四方形或六邊形的格子
● 簡單拓樸學-連接線段：e.g.a 排水網路	(方格圖素=網格)



- 或地下水公共設備
- 複雜拓樸學-連接線段及一套物件結構：  
例如：連接的多邊形
- 物件方向複雜拓樸學-內部結構與關係

- 不規則的鑲嵌：thiessenlation 多邊形
- 不規則三角形網絡 (TIN)
- 限定的元素
- 一組規則的網格/一組不規則的四方格



(a) 上左：呈現清楚的面（多邊形）的向量  
 (b) 上右：連續性區域的網格模式  
 (c) 下左：呈現連接線段的向量  
 (d) 下右：連續性區域形成的不規則三角形

圖 2.4：確實的實體與連續性區域在不同資料模式的編碼 (p.23)

### 實體的向量資料模式(Vector data modes of entities)

實體向量模式呈現空間中連續存在的“點”、“線”及“面”，其圖解可參考圖 2.3 笛卡爾座標系。

#### ● 簡單的點、線、面

簡單的點、線、面是實際上呈現在 XY 軸座標的現象，假定不會改變也不含有時間或空間多變性的資訊，實質點是在地理上的延伸，這個實點運用 XY 座標軸呈現出抽象層級的分析，一個鎮可能是由一個點所呈現，但一個面就會如一個地區的層級。增加分析的層級顯示出現象的內部結構（舉例來說像一個城鎮、地方行政區、郊區、街道、房子、郵局、交通號誌），其結果對某些人來說是重要的，對某些則不是。

線表示物體的地理伸展，可被適當的以一對 XY 座標來表示，定義一個通過空間而連結的方法，但在說明其相關屬性時它是沒有真實寬度的，一條路在國家的層級去看，是被呈現為一條線；當在街道的層級下，則變為一條路徑的面，線的呈現是不實

際的。換句話說，電話線可被表達成在調現在使用 GIS 所決定的大部分現實層面中。面最簡單的定義，表示為一個均質的 2D 空間，清楚地可知，這也依賴於決定的層面。多邊形可被其邊緣 XY 座標所表示，或其區域內的數組 XY 座標所表示。多邊形包括全部他們有的正確鄰近關係，且帶有相同屬性地不同多邊形可能出現於不同的位置（圖 2.3）。

如果邊界可被清楚地定義成可被一連串的 XY 座標的最小值所描述，包括 XY 座標的這組的大小依賴被使用的空間其決定的層級，常常假定決定的層及是由十分位去給予描述的座標或邊界，那意謂著改用邊界密碼是較列出邊界中所有座標來的更有效果的。

### ● 複雜的點、線、面，和具有機能性的對象

點、線和面更複雜的定義可被用來獲取實物內在的結構；這些定義可被機能化或說明。一個鎮包括了街道、房屋、和公園，這每一個不同的機能且可不同地反應問題和工作在鎮這個層級上。位向關係連結（圖 2.4ac）可被用來指示線如何連結多邊形或連結網路。

在近幾年，壓縮（encapsulating）實體資料的較高建構方式已可以物件為導向的方式。這是預備網狀階層和在一個較高群集層次的一個單一單位，一起的關連性群狀實體間的機能性關係，其以物件為導向是被更完整的描述在第 3 章資料建構的關係中。

### ● 連續區域的鑲嵌

連續性地表面可被分離（discretized）到單一基本單元的組成中，例如：方形、三角形或六角形，或是不規則三角形或多邊形中，（Thiessen/Dirichlet/Voronoi 程序-見第五章）那是從地理的呈現所鑲嵌的。長期被用在地理測量上的不規則三角形的使用，是立基於三角測量的原則上，如何使土地的連續性表面接近三角形的網孔，其三角形的頂點或節點是藉測量其仔細地座落於三角學的點的“點高度”所給予的。這個方法的主要優點在於網孔的密度可被輕易地隨著表面的變化調整到那個等級，例如：帶有少量變化的面可藉少許的三角形去做表達，當面的變化較大時，則需要的三角形會更多。這支持一個在資料中變數的理由。三角形的表面也可輕易地在型式上的調整改變，如看地形的 3D 呈現或其它表面，如飛機機翼或車體。這些資料模式本質上是 hypsometric 多層表面的靜態呈現，那超過時間也不會改變的。

三角形網孔也被用來呈現連續性的變數在地表水流、地表水的移動、風場等的動力模式中。在這個型式下，他們提供一個稱為限定的元素（finite element）的結構，在屬性的不同點評價於鄰近三角形網格所產生水的流動或分解的材料被展示使用不同的微積分學。三角形不規則網路或 TIN 是一個資料建構（見第三章），被用來塑造藉簡單多邊形向量圖表所示之連續的表面。（圖 2.4d）

大部分使用的三角測量方案是規則的鑲嵌或規則的格子，2D 幾何學表面是分割成方形格狀（即像元），大小是由被決定要求去表現出的目的屬性的變數去做決定。空間的格狀呈現是像網格資料模式一樣（圖 2.4b）。當網格被用來表達連續改變之屬性的變數，每一個網格將有不同的屬性價值，網格間的變數將被假設為數學的連續，所以不同的微積分學可被用來計算區間的平均值、改變速率等。每一個網格可被當作是一個單獨的實體，和向量式的多邊形的不同在於向量式的規則行事和明確的圖型是較絕對的。

雖然規則網格是較常被用來描述靜態的現象，他可以輕易的被應用於處理動力學的改變，超過時間的改變可被紀錄在網格單獨的層中，所以改變從靜態的到動態的資料模式要求只在於基礎的結構是被重複於每個時間步驟，時間如同空間，被假竟去分離在這個模式下。規則的鑲嵌也可是網狀的準備更多空間的項目，經由線狀的區域四元樹和其它網狀的結構（見第三章）。

空間外旁生的改變也可被規則網格所輕易地應用，因為他接近到一個連續的、不同的數學表面材料的流動，經由空間可被計算使用限定的不同（finite difference）模式，因為網格不變的幾何學，意謂著第一和第二導函數可以簡單的減法和加法輕易地計算出來。

### ●2D 像元 (pixel) 和 3D 像元 (voxel) 如實體

在連續空間的規則鑲嵌分離的基本單位也可被用來準備一個幾何學的證明點、線、面的簡單資料單位。一個向量的“點”可被描述為一個單一的網格，向量的“線”為一組鄰接的網格，網格的寬度有相同的屬性，一個向量的多邊形為一組鄰接的網格，網格具有相同價值的屬性。向量描述常較受偏好，因為規則網格可簡化空間的項目，所以也減少了電腦和記憶體所需增加能量的問題。

空間的向量式和網格式是同值的，時常因表達自然的現象而搞混，例如：連續性區域也可以等值線或數組 XY 座標連結數組相等屬性價值的輪廓線來表現，輪廓線通常是在地圖和電腦螢幕中描述屬性價值的方法，但對於在空間交互作用的統計模式上處理連續空間的變數能力較弱。封閉的等值線可被像簡單多邊形或像實體步驟的角度來看的封閉線來討論，但他們僅只是描述的加工品，不是真實世界“物件”的外型。兩個方法的重疊是最大的，當我們必需去處理現象，如土壤圖單元或土地使用，或土地覆蓋單元。老式地圖的傳統方法是去定義土壤、土地使用、土地覆蓋等的類別，然後去確認符合這些類別的土地（實體）的區域。這些區域可藉由向量式邊界圍繞的多邊形或多組具相同價值相鄰的網格來描述。如此，描述就可以像是一張複合色彩地圖（chorochromatic map），因其每一個區域被表現出一個單一的色彩或陰影。

一個表達人造物如土地使用或土壤類別的方案方法是去假定土地使用或土壤是連續性的變數而非實體，但地理的表面是由區域所組成的，區域的屬性有相同的價值（多邊形）且區域的屬性價值突然地改變（邊界），這個方法先需要從基本網格資料，如間接地感知的形象去選取實體和均質的區域。注意的是，通過這些種類的屬性的變數是被認為像是連續的在空間中（因為每一個網格有關於土壤或土地使用的價值在，包括沒有土壤或濕地這些類別）。以不同的微積分學角度來看，表面不是連續的，這意謂著數學的工作，如坡度的計算必須不被適用在不能藉一個連續性數學機能所處理的資料上。

在 2D 和 3D 空間中，我們可以想到像元 (voxels) 或多邊形像是單元可被論述成一連串代有規則或不規則形式的開放系統。每一個網格（區間系統）的情況，由其屬性的價值所決定，屬性價值可被在問題中或從其他網格中使用的資訊的工作所改變。模式可見於向量式或網格式區域內，圖也做了概述和在向量的表達上製圖符號明確的作用。

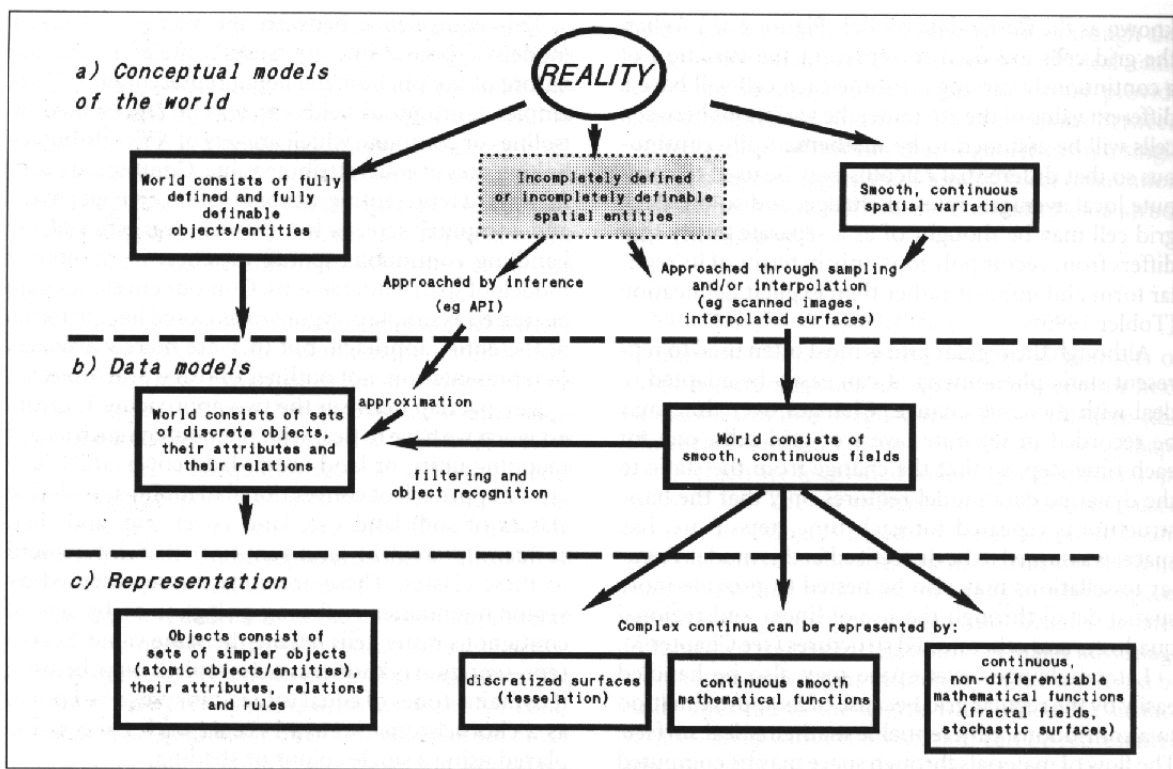


圖 2.5：過程中的步驟：由真實世界現象觀察到標準化資料模式建立的步驟

### 使用向量式和網格式方法展現地理原件(The display of geographical primitives using vector and raster approaches)

人們習慣看由線或點狀陰影所表達出的空間資訊，如果我們以慣例的方法去看所傳達的地理資料和紙上地圖，而這些都使用了向量和連續性區域的資料模式在面的屬性上，真實的世界不是以色彩化格式的點、線、面的符號來表現實體，就像一個在關於空間屬性上的價值中變動，如：像等高線高度的描述。

在電腦中，這些原本的外形或點狀的步驟，是進入表現向量和網格式的形式。圖 2.6 描述在這個簡單的地理資料模式的主要方法中，可被視為在向量或網格式的区域，圖中也概述並明確指出向量的表現方式，其對製圖符號學的影響。

VECTOR	Points	Lines	Areas	RASTER	Points	Lines	Areas
Feature data				Feature data			
Areal units				Areal units		-	
Networks				Networks	-	-	-
Sampling records				Sampling records		-	
Surface data				Surface data		-	
Label/text	Utrecht Arnhem	A12	Lake	Label/text	-	-	-
Symbols				Symbols			
Relations	attributes and pointers	attributes and pointers		Relations	attributes and relations	attributes and relations	

(a)左：向量式模式；(b)右：網格式模式  
圖 2.6：由圖表展示資料的不同方式 (原文 p.27)

### 資料型式(Data types)

每天的言談中，我們會從量的資料或數目中區別定性或名義上的屬性，並將認知中不同種類的工作，去做出適合不同種類的資料。用同一個形式化的資料模式和使用向量資料型式的資訊（這個資訊是可被寫下且儲存於電腦中的），這兩樣都是對的。（表 2.2）

表 2.2：資料型式 (原文 p.28)

資料型式	容許值	容許的運算
布林學	0 或 1	邏輯和標示運算：正確對抗謊言
名義上的	任何名字	邏輯運算，分類和證明
次序的	從 0 到 $\infty$ 的數字	邏輯和排列運算、大小比較
整數	所有的數字，從 $-\infty$ 到 $\infty$	邏輯運算、整數算術
真實的	實數（帶十分位），從 $-\infty$ 到 $\infty$	所有邏輯和數字表示的運算
位向關係	所有數字	實體間標準的鏈結

物件屬性可由布林函數名義上的、次序的、整數的或真實資料的型式來表現。真實資料型式包括十分位；整數和真實性共同被認為是數量（非向量）資料。地理的座標有時候被表達成整數，但大部分被表達為真實資料型式，且位向連結使用整數。不同連續性表象要求真實的資料型式，即使整數有時被當成近似值（接近值）使用，非不同的連續性表象和它的不連續形式（discretized）（網格式或像元）可使相同區域的資

料型式像實體。

邏輯運算可被實行所有資料型式，但算術工作被限制達到真實和整數的資料型式，因此，資料分析的種類是由使用在資料模式上的資料型式所支配。算術工作的準確性限制是由電腦訊息所記錄的數字程度去界定的。（見第八章）

## 資訊系統中處理運用資料的定理及程序(Axioms and proedures for handling data in information systems)

已探討過定義資料模式的問題，有如在真實世界中現象的表現，可從地理的原貌，如簡單或複雜的實體，或我們可以列舉中的邏輯面原則和定理被分離的連續表面中被建構，而那些原則和定理支配著資料模式可被討論的方式。以下說明為資料處理提供一個正式的基礎和提前瞭解在第七章與第八章敘述的材料。

- 1、確認某種分離（discretization），如帶有資料的實體是必要的。在地理資訊系統中，原始的實體是點、線、面和像元（格子元素）。複雜的實體具有一個被定義的內在結構可被從點、線和面的組成中建立。
- 2、所有基本的實體以它們的地理位置（空間座標或幾何學）屬性特質和關係（位向關係）去做定義。這些關係可被純粹的幾何化（考慮空間關係和鄰近物）或層級化（考慮屬性）或兼具兩者。
- 3、藉由各個實體的屬性、位置或內部或外在關聯與另一個實體進行區別，就單純的觀點而言，“實體”或“物件”通常被假定為內在均質的，除非他們是以數學方式呈現，或者除非他們建立的原貌是複雜的物件。在大部分 GIS 的案例中，通常只能用來區分內部均質及分界明顯的元素，但更完整的地理資訊系統可以幫助我們理解不精確的資訊，包括元素誤差或不準確的屬性、關係、位置及輪廓。
- 4、實體和屬性皆可被區分為有用的分類
- 5、這比例性的微積分（布林函數（Boolean algebra））可作為實體邏輯性的工具，包括其屬性、關係、及其群體的屬性。
- 6、在地理資訊系統中比例性的微積分可對廣大或長期的數字進行統計：
  - distance 距離
  - direction 方向
  - connectivity (topology) 連通性（位向）
  - adjacency 連接度
  - proximity 鄰接度
  - superposition 重合性
  - group membership 群體數
  - ownership of other entities 所有權及其他

地理資訊系統亦可利用微積分方式去計算不精確的數據（第十一章）

- 7、新的實體或設定實體可以幾何方法或現存的交集（線的交集、多邊行重合所組成）--見第六章。
- 8、利用點、線、面或顯示器辨識單元可建立新的合成實體。
- 9、利用現存實體的邏輯性或數學過程、模型可推演出新的屬性
  - New attribute  $U = f(A, B, C, \dots)$
  - 新屬性  $U = A、B、C\dots$ 的集合

這數學的計算依靠於各種的數據形式的演算（加法、減法、乘法、除法、三角形法、微積分等）見表 2.2。

新的屬性亦可利用幾何屬性（線性、面積、形狀、大小、周長）及插補方法推演出。

1. 實體正確的分別，可藉由個別子數據平面或重疊的屬性分析完成。
2. 在同一個 XYZt 座標中的數據可與整個數據位置連結。
3. 藉由 XYZt 座標表示一個單獨的數據可將其分類為一個單元或點位。
4. 新的 XYZt 座標屬性值可藉由與周圍環境的活動關係所衍生（傾斜度、外觀、連通度）。

## 數據模擬與空間分析(Data modelling and spatial analysis)

我們可藉由一些基礎的定理、模式數值、地理屬性的數據等的使用，以清楚瞭解其間連接關係，以下有幾個不同的例子可以說明：

1. 地理資訊系統最常使用的方式，是去標示一個不能改變卻需要瞭解的實體位置及形式，其可藉由新數據的輸入或數值模式的輸出，可轉變不同的屬性，以表示實體模式的向量關係。
2. 如果實體的屬性是固定的，除位置外，形狀及外型可以改變，如乾涸的湖，隨著湖水面積的改變，其向量模式也需要隨著時間被重新定義，無論如何，一個網格模式可以表示出這個湖表面的變化，操作這個過程則湖水的變化可以被連續的表現出來。
3. 假若一個實體的屬性可以改變，但其形式和實體間的連接關係不能改變，這關係可藉由一個實體導向的模式標示出其層級屬性及關係。
4. 假若一個實體並不能清楚的被辨別，則他通常被視為特殊、任意的連續空間。

## 使用數據模式的案例(Examples of the use of data models)

地理數據的使用者每天要決定使用的數據模式，以下有幾個重要的說明有助於我們瞭解這個概念。

### ●地籍圖 (Cadaste)

地籍圖及土地登記的目標是作為公家機關及所有權人的紀錄，他可以提供有關土地位置、面積、寬度等重要的內容和其屬性（例如地名、地址及所有人），而這些資料的匯集也是相當重要的，在這個案例中實體（向量）模式，可幫助我們對事物進行、概況及真正的數據進行瞭解。

在許多城市中土地登記是相當有計畫的，其邊界線的數值趨近於無限小，所以邊界的訂定相當精確。國家或地方邊界的座標必須相當準確，而其基礎數據的性質必須要進行結合，這將有助於同時表現基礎數據、邊界向量成為一多邊形的網路，以展示鄰接關係及其內容物。

### ●公共設施網 (Utility networks)

公共設施網是電路及管線的集合術語，他連結了各個住戶，藉著國家及地方的供給者供給其水、電、瓦斯、電話及有線電視，他也連結了污水處理系統，將廢水引入水溝及下水道，在許多地方網路是隱藏在地下，或是例如墊、電話及有線電視是利用

沿著道路設立的電線桿。

在記錄現象時，網路有三個重要的方向：（a）給予網路的屬性（他攜帶了什麼，用什麼連結，及其他附加資訊如使用者的姓名、年紀等）；（b）網路的位置（原本沿著公路所設置的位置，有助於需要修補時的速率）；（c）網路連結在一起有多少不同的部分。藉由其相關連的屬性連結可使我們清楚瞭解這些不同的數據形式（見表 2.2）。

### ●土地覆蓋物的數據資料庫（Land cover databases）

因為土地覆蓋物的不同，國家及國際行政機構可將其分類為都市土地、農耕地、草原、森林、海岸及山脈等。建立一個數據模式需要以下幾個步驟，首先，必須要去明確的定義這數據的種類；第二，需要決定如何去判斷他；第三，選擇一判斷原理（測點或衛星遙測）以解釋結果的產生。

最簡易的數據模式假設各種類均質且排他，而他們之間有明確的關係且皆位於地面上，假若這些是成立的，則每個種類的型態皆可藉由多邊形表示，這結果將有助於瞭解一地的事物，但這些問題的關鍵在於建立的數據的種類、圖面的判讀、種類附屬的屬性及其與同種地圖間的定位，這些數據必須由名詞上的意義（種類名稱記錄）轉變為數據（面積的計算）。

討論至今，在土地利用形式無法辨別時將如何處理，不同的個人及組織對於土地利用形式的判定有不同的方式，但他們的中心思想卻是相當類似的，在表 2.3 中將可展現其不同的結果，顯而易見的，從不同來源分析出的基礎數據會有不同的分析結果，他必須被嚴格的解釋及翻譯。

表 2.3：歐洲土地利用不同估計方式表—km<sup>2</sup>\*1000（原文 p.31）

土地使用分類	FAO-Agrostat	Pan-European Questionnaire by Eurostat	10 minutes Pan-European Land Use Database	Land Use Statistical Database	Land Use Vector Database
固定的農地					
德國	4.42	—	1.80	2.30	5.36
法國	12.11	—	12.07	12.18	31.45
荷蘭	0.29	—	0.22	0.34	1.07
英國	0.51	—	0.52	0.59	6.54
森林					
德國	103.84	103.84	98.56	—	100.46
法國	147.84	148.10	140.675	145.81	79.63
荷蘭	3.00	3.30	1.48	3.00	0.78
英國	23.64	24.00	18.96	14.29	10.03

現在我們討論測量方法對結果的影響，如果我們利用採樣的方式收集土地利用的資料，例如對設定點未曾被觀察的區域進行實際的調查和紀錄，這些未被取樣的基地會依據不同密度及品質的利用形式及調查技巧被加以區分，為了對我們所決定紀錄的區域進行詳細的了解，必須建構於原始地理面或形狀的基礎上。如果我們採用格子的方式對地表的使用情形進行進行紀錄，擇期屬性降因其所屬的利用情形而有所不同。

如果我們使用遙感探測去判斷土地利用形式，則是利用衛星掃描的方式，自動的對地表連續空間土地利用形式，以格狀的形式加以判讀，這圖形的辨識率受限於最小的掃描區域（Plate 1），由於資訊是因均質的網格所形成，不像實地調查的方式我



們必須對調查區域作完整的了解，所以這主要的問題在於衛星遙感探測的辨識率及轉換數據所得的網格尺寸，顯而易見的，這分類品質得成功是決定於整個過程。

這例子說明了兩者的不同，但完整的數據模式才可以顯示土地利用，這表示向量或網格模式的精確度，一部份必須依靠資料的收集，另一部份則是其被利用的形式。

### ● 土壤圖 (Soil maps)

在原始的地理上，大部分的土壤圖是以向量的多邊形表示實體的數據模式，多邊形可因土壤的均質的覆蓋單元形式被區分，其邊界線暗示土壤在極小的距離中快速的改變，並在圖上記錄土壤的外觀及內涵物。

這些數據模式來自實際的經驗，因為他標示了基地在圖面上的位置，並利用相關調查報告的土壤性質資訊以單位標示，但在科學上這樣的範例並不是精確的，因為其結果忽略了土壤結構在空間中的變化（見 Plates 4.4, 4.6）。

這個模式的概念與其他以多邊形為基礎的空間資訊結構相同，例如土地使用分區及土地所有權數據，在 1970 年代晚期及 1980 年代，這個土壤資訊系統發展的模式已經可以利用 GIS 進行假設，數據的形式被用作屬性的方式紀錄下來包括其名字、數值、真實狀態。

在地理空間上土壤分類的空間界線是相當被爭議的，藉著明顯、突然、顯而易見的地文特徵，例如岩石、水流、坡度的急速變化，土壤可以被明顯分界，我們稱其為主要分界（primary boundaries），可標示土壤結構的空間變化的我們稱其為二度分界（secondary boundaries）。

當前的使用者認為，在地文學上印刷的土壤圖並不能表示其真正的主要分界和二度分界，因此這兩種形式的邊界皆被相當微小的線所表示，在繪圖的過程中，不可避免的包含了土壤變化的概念及空間屬性的變化、地文繪圖的單位，根據 Boolean 布林模式，每一個基地可以屬於一個唯一的土壤單位，包括空間屬性及地質上土壤分類的屬性，可以 0（非成員或於面積外）或 1（為成員或於區域內）表示。

土壤、植被和其他類似自然狀態的主要的問題，在於其空間的變化，可以由細微的毫米至整個世界，即使土壤學家已經花很長的時間去辨識他，但地質圖的繪圖者人對不同程度的土壤圖採取地誌圖的模式，當一個空間產生連續而變化的土壤時，因為其假設內在單位的均質而產生嚴重的謬誤，此外，由於人為景觀造成的土地使用邊界，可以適當的被連續區分，真正的土壤邊界也因此更為清楚。

一個土壤的連續多邊形數據模式，是假設景觀中土壤的變化是逐漸的，其模式是藉由一系列的位置及數據所建立，最簡單的數據形式是指地理學上聯合屬性的分類，利用不同的連續空間可建立新的數據模式。

在原理上，土壤的數值模式非常相似於地表高程的測定，以輪廓線的方式連結地表上相同屬性的區域，或於網格中內插一個單元，皆可表示為一個網格空間模式，當土壤因為一些物理、化學、生物景觀的因素而發生不同時，其進行樣本的採樣必需依據一些採樣的原理，不論是面積或樣本數皆影響到數據的品質，這些改寫的方式及結果將在第 5、6 章進行說明。

### ● 水文學 (Hydrology)

水文上的應用模式受空間、時間及傳送材質的影響，會改變其位置、屬性及邊界形式，不僅是河流的層次亦包括湖泊在幾何及位置上的變化，水和土地間邊界的改變

會因為水體而有所不同，例如一個洪水可以打通一個新的河道，使舊河道荒廢，因此這改變同時出現於位置及地誌上。

由於幾何及地誌數據的改變包含相當多的數量，這簡單的實體向量點、線、面數據模式並不是極佳的水文現象表現方式，較好的方式是利用物體的方位物件導向原理去表示原始的實體與其機能（McDonnell 1996），這內部的數據結構組成允許進行轉換，所以這數據模式不僅包含了地理、位置、地誌和屬性資料，更包括期間如何變化的資訊。

連續而變動的數據模式可以容易的擷取傳送的資料，使用不同辨識率的幾何形狀及網路是最常見的水文模式，但這並不是一般的 GIS（例如 McDonald & Harbaugh 1988），連續而變動的網格數據模式不只傳輸了表面的資料，更可對數據資料進行增加或推演（見第八章）。

## 總結：實體或區域？(Summary：entities or fields？)

圖 2.5 說明了由知覺資料建立一個電腦化資訊系統的步驟，它可以幫助我們在實體或區域的基礎方法上進行清楚的選擇，在其他方面，他亦可以依據使用者的需求加以制定。

假若這最初的狀態經過清楚的鑑定，則實體向量與網格或連續表示的資料將可進行轉換。網格與向量間的關係在第四章將加以說明。無論如何，無可計數的科技可對狀態紀錄資料進行闡述，如果科學家 X 提出利用多邊形的方式去分析地景的機能，而科學家 Y 提出以利用連續的變化，則這二者皆可能曲解真正的事實，或任一者也不能完整的呈現。

圖 2.2 說明了在應用與訓練中，如何選定最佳的實體或連續區域的眾多方式，一般來說，這些原理已有詳細的項目與紀錄，在處理區塊的學習過程中，連續而不同的區域必須建立動態的數據模式。

在大部分的地理資訊系統中，其位置的數據與性質的值被視做相當精確的，所有的事情被當作已知，沒有空間是不確定的，由於我們所需求的每一個數據單位資料的收集過程及花費，使我們無法克服一些統計學上的失誤，但在原理上數據的品質不能因數據而增加這問題是沒有答案的。

事實上，簡單的空間狀態實體模式有一些不同於玩具的塑膠建物體，由於不同方式所組成的不同區塊，包括其邊長、種類，將遵守一些組成的基本定理，但在理論上，由地理資訊系統去建立一張新的圖，其組成的點、線、面並不會不同，若給與足夠的資源，我們將可建造房子，再造景觀，甚至與實際生活相同大小的動物模式，例如長頸鹿和大象。

就這例子而言，我們利用塑膠區塊去建構一個模擬的長頸鹿，包括大小、外形與顏色，但這模式並不具有長頸鹿的機能，他不能吃、睡、繁衍和呼吸，因為他的建構單元無法具備這些機能，較常見的，許多地質基礎數據可提供分析與處理相關的模式，除非選擇正確的基礎數據模式，數據才能真正提供地理上正確的機能。

## 在地理資訊系統進行空間分析時，所考慮九個因素

以下列舉 9 項非完全獨立，卻是我們在選擇數據形式及基礎數據過程中基本且重要的考慮問題：

1. 在概略或完整的學習中其形式及狀態是否真實

2. 種類實體形式與狀態的描繪是詳細或概略的
3. 數據形式是否記錄了屬性、名稱、次序、數值、真實或位向關係？
4. 這實體的基礎數據所表現的對象是否詳盡，或是元素描述完整，他的性質選定是有規則的或隨機的
5. 這基礎數據是否可以表示連續的實質的或區域屬性？
6. 這屬性的基礎數據是來自完整的細目或採樣！
7. 這些基礎數據將被用作描述、行政管理或分析等目的！
8. 這些使用者是依據邏輯、經驗或過程去推演出新的資料或情況！
9. 這考慮的狀態是靜態或動態！

## 問題(Questions)

1. 發展簡易資料模式使用在下列應用的項目：
  - 道路交通資訊系統
  - 速食店的位置
  - 高山地勢、山崩的發生
  - 地下水污染物質的擴散
  - 緊急事件單位（警察、醫院、消防隊）
  - 旅遊資訊系統
  - 高山地區植被改變的監測
  - 空氣污染移動的監測，如 1986 年車諾比事件（Chernobyl accident），由降雨造成銫  $^{137}$  ( $^{137}\text{Cs}$ ) 的沉澱，思考資料的來源、各種現象的限制、資料模式、資訊型式及使用者的主要需求。

## 第三章 電腦中的地理資訊

### *Geographical Data in the Computer*

#### 前言

提供電腦化空間模擬工具利用的可能性以了解資料模式如何用來表達在這些圖樣中被編碼的地理現象是必要的。本章敘述在哪些空間資料可被有效的編碼到電腦中以支持 GIS 運作的方法。電腦編碼資料和執行指令使用一連串存在兩個狀態“開”或“關”之一的開關旋鈕，這些狀態由數字 1 和 0 個別地編碼，所以對所有計算而言，基本的二進位系統是根本的編碼基礎。所以地理資料必需被轉換到電腦中個別的記錄，使用這些開關旋鈕去表達位置、存在或不存在、型式等的現象。被運用於資料的主要資料結構被稱為向量式或網格式，它們劃分空間為一連串的点、線、面，或是一個規則的鑲嵌。最近，物件關聯資料結構已被用來使用相同基本單位，如向量或網格系統，非建構實體至連接的或層級的型式。

當進入電腦中，資料必需被組織允許有效的獲得擷取、修正及操作運用，電腦中有組織的資料其系統範圍從簡單表列或索引檔，到依據層級、網路、關聯和物件關聯的圖表資料之高建構資料庫。這些結構決定資料如何被組織到資料記錄中，同樣地，控制它們保存在電腦中的方式。各種的組合利用資料呈現和處理運用的限制，影響了資訊的分析及模擬。電腦中編碼的不同方法和儲存各種的資料結構是被討論的。在近幾年研究已強調在為實體資料支持一個更實際的建構的 GIS 中使用物件取向資料庫結構的可行性。

地圖的電腦說明可能只有一個數字影像的紙張文件，儲存在一個電子儲存裝置中，如 CD-ROM，或可能是一個複雜 GIS 資料其相互關係的表現。地圖字母的型式可能已被制定使用程序，不同於傳統的位相關係圖或主題圖，可以使用其他各種符號系去傳達訊息，因此了解最終產物如何與在問題中現象的組織觀察作關聯。

電腦中提供一個較傳統地圖對記錄、儲存、修正、分析和表現空間資料更大範圍的可能性。然而，傳統的紙地圖是一個在終將給予的点的調查所產生普遍的目的資料庫。電腦可以為任何被賦予之目的選擇任何資料的呈現方式。電腦允許我們去加入資料、修改資料和從存在的資料中計算新資訊，所以一個電腦處理的空間資料庫不需依然為靜態的，但可能被用來模擬在空間現象其時間上的改變。這些新的可以性浮現了隨之而來有關空間、在空間資料及空間現象應如何以概念模式為觀點去敘述的本質的問題。

如果我們的目的是在於自動操作地圖製作的過程，去提供一個傳統地圖的類電子物（在空間現象的圖表表示是由標準符號及濃淡所編碼），然後概念模式要求為電腦化資料，與被使用在傳統地圖製作上有些小的不同點。主要的工作是去設計形式化地理資料模式編碼的方法，使電腦可以處理它們。形成概念—資料終止—資料修正—具像化間的連結是和被使用在傳統地圖製作上的那些是相似的，利用電腦從事修改困難的工作和繪製地圖。

然而，如果我們的目的是在於拓展我們的能力去處理空間資料，我們所看到的世界在那一個方法下，是不被揭束的印製地圖所限制。如果我們希望為特殊目的得到空間資料的次情況，或連結空間現象中的資訊至控制它們的外觀及屬性的實際過程，其次，電腦提供很多除自動操作的實工作以外的機會。電腦提供帶有資料的交互關係的意義，在那裡是不可能帶有被列印的資訊。資料可被改變、修正、重新計算，且表示的不只是已從真實世界收集而來，或以空間和短暫過程的數字模式為達案的新資訊的一個反映。此外，GIS 是較一個單一

自動操作者的實際工作更多的，它提供了包括在數字形式下空間資料的檔案記錄，及探討在空間和短暫現象中過程和型態間的交互作用的一個工具。

當地理資料被輸入電腦中，如果 GIS 認可的概念模式和使用者所習慣的概念模式是相同的，則資料將是非常方便的。（在第 2 章討論過）這將允許使用者和電腦間的交互作用成為可能且直觀的。無論如何，人類的空間認知並非經常是最有效率的方法去建構資料庫，且不能解釋物理的必要條件的儲存和重複使用數字資訊。雖然處理地理資料的電腦需要被計畫去表達現象生態學結構於適當的方法，同樣地允許以技巧之方式，將資料寫下或儲存於磁性的裝置中。不同地，很多其他種類的資料被近代的資料系統依慣例去處理，地理資料是複雜的，因為它們和位置、屬性及所記錄之現象其內在和外部位相關聯的連結，地理資料的位相關聯和空間觀點區別在於管理的記錄由被銀行、圖書館、飛機預約或醫療記錄上資料加工系統的處理。除了 GIS 外必須展示使用地理組成和形式的資料，例如：色彩、符號、濃淡和線型及寬度，那是被人們所說明的。雖然像本質的結果、量和地理資料的複雜性，一些邏輯的電腦概要資料已被發展為有效率的儲存、更新和修正去包裝資料進入裝置中。這些概要資料是一連續的模式/結構發展過程，描述於第 2 章和 Box2.1 中的 (d)、(e)。（為了方便，敘述於下）

過去的資料已被儲存於資料庫中，它們可被修正和轉換，接著根據形式化的邏輯和數學的原則。新的資料可被從舊資料中產生，資料可被一再使用在品質上不使變壞，雖然它們本質上的品質可能影響品質的結果—第 6、7、8、9 和 10 章，包括了這些地理資訊加工的方面，首先，考慮電腦如何儲存資料是重要的。

#### 空間資料模式及資料結構

（與 Box2.1 同）（p.51）

類似物的產生和數字的空間資料樣式包括七個模式發展及抽象概念的層次。

- (a) 實體的觀點（概念模式）
- (b) 人類所形成之概念引導到一個模擬的抽象概念。（相似模擬物模式）
- (c) 模擬抽象概念的形式化，不帶有任何慣例或限制在實行上。（空間資料模式）
- (d) 反映資料如何被記錄於電腦中的資料模式呈現。（資料庫模式）
- (e) 檔案結構，是在電腦記憶體中較特殊的資料結構呈現方式。（實質電腦模式）
- (f) 為處理資料而去了解之定理和原則。（資料操作模式）
- (g) 為人們展示和呈現空間資料而去了解的原則和程序。（製圖模式）

### 電腦中的資料(Data in the computer)

人們呈現數字和做算術的使用基礎為 10 的十進位系統，每一個在數目字被寫下的形式中的位置指出很多單位的 10 的能量如何符合那個位置。因此，第一段到十進位的左邊表示單一的倍數 ( $10^0$ )，第二個 10 的倍數 ( $10^1$ )，100 的第三個倍數 ( $10^2$ ) 等等，電腦使用不同的計算系統，它們儲存資料於只有兩個情況的開關旋鈕的陣列形式中，分別是“開”和“關”，且這些狀況是由數目字 1 和 0 個別地編碼，因此，電腦編碼資料和做電腦運算使用基礎為 2 的二進位算數。Box3.1 示範二進位系統和算術，在跟隨討論數 10 是用來指出在十進位系統之數目，和二進位系統中的數目。

### 二進位數及電腦顯現的數目和文字(Binary numbers and computer representation of numbers and text)

當帶有一般的數字時，在二進位系統的資料被表達使用一連串的欄、列，在電腦中，每

欄被用來表現一個開關旋鈕。開關旋鈕是在電腦磁片和記憶體中真正小的磁化的區域，其通常 8 個成一群，被稱為一個 byte，數個 bytes 可被連結在一起去形成電腦的字，字在資料記錄上是被群集的，且記錄在電腦檔案中是被群集的。數組電腦檔案可被層集的群集於指南或次指南中。

用 8 個 bits 為 1byte，我們可以表現 256 個數目從 0 到  $(2^8-1)$ ，即從 0 到 255。例如 8 個開關旋鈕的關聯  $11111111_2$  意指：

$$2^7 \times 1 + 2^6 \times 1 + 2^5 \times 1 + 2^4 \times 1 + 2^3 \times 1 + 2^2 \times 1 + 2^1 \times 1 + 2^0 \times 1 = 255_{10}$$

記得任何位置的數字的 0 次方相當於 1。

如果我們合併 2bytes 到一個 16-bits 的字，從 0 到 65535 編碼數目是可能的，無論如何，可以去編碼確實的和否定的數字也是方便的，所以只有第一個 15bits（從右邊算來）被為編碼那數字來用，第 16 個 bit ( $2^{15}$ ) 被用來決定符號，這意思是：用 16bits 我們可以從 -32767 到 +32767 編碼數字，一個缺少分數組成的數字被稱為整數，因此在 -32767 到 +32767 這範圍的數字可被稱為 16-bit 整數。

通常大於或小於 -32767 到 +32767 或有分數組成的數字需去編碼數字。因此帶有更多 bits 的電腦字是被需要的。帶有確實和否定的價值，以及小數的真實數字需要 32bit 或甚至 64-bit 的字（因此被稱為兩倍精密度），在某些為小數部份（小於 1）保留的 bits，剩下的給較大的數字使用。編碼數字的方法意指：帶有可編碼的任何數字的精確性是被使用之 bits 的數字的一個作用。這意指了數字不可被精確的編碼，當使用者企圖用那些較被允許之數字大的數字去運作時，可能導致循環的誤差。第 9 章討論在 GIS 更細的項目中其循環誤差的問題。

## 資料組織和編碼使用十六進位法 (Data organization and coding using hexadecimals)

我們已了解十進位數字的編碼進入機械使用的形式，但電腦也被用來儲存文字特性。為編碼最初數字的資料，最平常的使用系統之一被認為是（美國標準電腦資訊索引 ASCII：American Standard Computer Information Index），是基礎於 16 進位系統。

這個系統利用 4bits 的單位運作，4 位元 (bits) 是相當於  $16 (2^4)$  運算。16 進位數可藉單獨的特性（像在連續的 0-9）來呈現，以 A、B、C、D、E、F 分別取代數字 10、11、12、13、14、15。欄/地方系統被使用像二進位或十進位數字。例如：

十進位數字 17 和二進位 10001 是十六進位的 F1

十進位數字 32 和二進位 100000 是十六進位的 20

十六進位系統提供一個適宜的基礎，為編碼最初數字的資料，如：字句的字母和從 0-9 的數字，因為 8bits 的每個 byte 可以編碼為一個寫下的符號。ASCII 為數字特徵被像十六進位數字一樣從 0 到 FF（十進位數字 255 或  $2^8$ ）所寫下。例如：

- 字母 "A" 被編碼為十六進位的 41（十進位 065，二進位 1000001）
- 數字性質 "5" 被編碼為十六進位的 35（十進位 053，二進位 110101）
- 希臘字母 "μ" 是被編碼為十六進位的 82（十進位 230，二進位 11100110）

當為文件性質的數字和編碼兩者被數列被編成 0 或 1 的 bits 所呈現，加法指令需被加入電腦檔案去指出資訊被直接編碼或像一個被列印出的文件的呈現。

ASCII 系統是沒有意義的編碼文件的唯一方法，很多其他系統不只是可能的，且已為編碼測試性質、圖表、影像、聲音和其他資訊而裝置。Bits 的組織被認為是它的格式化，清楚地除非格式化被了解，否則資訊不可被從一系列的 bits 中抽出。

## 資料容量(Data volumes)

bits 的容量需要去改為密碼的資訊，是被表達在千 (kilo-)，百萬 (mega-) 或兆 (giga- 或 tera-) 的 bytes。一頁標準的 A4 被編碼在 ASCII 格式中附帶 64 線的測式，有 80 性質/線 (包括空間) 需要 40960bits (5120bytes 或 5.12k)。200 頁的書僅給性質需要 1.024Mbyte，在資訊中的頁數、文件字體、版面設計等需要增加空間，地圖和其他空間資料的形式需要大量的儲存空間。

從前面的段落，提到資料型式和 bits 數字間的關係需要去編碼資料，意思是：當涉及一種資料和另一種比較時，儲存電腦空間和過程時間是可能的。

### Box 3.1：二進位數和算術 (原文 p.38)

二進位系統是以 2 為基礎，在二進位系統中數字計算使用 1 和 0，當使用一般數字 (十進位系統) 時，資料是被呈現使用一連串的“欄”，第一個計算在  $2^0$  (即 0 或 1) 的單位，第二個計算在  $2^1$  (0 或 2) 的單位，第三個計算在  $2^2$  (0 或 4) 的單位上等等。

二進位連續的 0010 為： $2^3*0+2^2*0+2^1*1+2^0*0=2_{10}$

二進位連續的 1100 為： $2^3*1+2^2*1+2^1*0+2^0*0=12_{10}$

二進位數字與十進位數字相似，但有些不同點，

● 加法和減法是相似的： $14_{10}-3_{10}=11_{10}$   $1110_2-0011_2=1011_2$

● 乘法是總合數目  $23_{10}*4_{10}$  被分解為  $23+23+23+23=94$

二進位同樣的是  $1011_2+1011_2+1011_2+1011_2=1011100_2$

● 除法減去數字 (相反於上一個例子)  $1011100_2/1011_2=100_2$  等於  $4_{10}$

● 乘以 2 只要將欄往左移一個  $2*0101=1010$

● 除以 2 將欄往右一個  $1110/2=0110$

當然，如果最右邊 bit 是“1”，除法結果 1 會被移除，所以再以簡單地乘法乘以 2 並不會使數字回到原始值，除非加入其他的檢驗。

## 輸入電腦的基本資料模式的編碼(Coding the basic data models for input to the computer)

資料輸入的工作是資料的轉換從人們已看到，或可認可，或如被紀錄在自動裝置中的形式到適當的為電腦中數學加工的形式。在前一章，我們了解如何去定義立基於地理原件形式上的概念資料模式去表現資料是必要的。第四章涉及建立一個資料庫的實際外貌和產生有用的輸出，這章與在空間資料中可被有效編碼去實行任務和需要的工作的方法有關係。

將被從第二章清除，傾向於呈現實體和連續範圍的地理資料其個別的建物街廓是必要的，同樣地，是可能表現精確的或不精確的屬性、位置和關係。主要的空間單位使用在表現由向量 (點、線、面) 和網格 (像元) 原件的資料上。不同的資料型式和每一種也被強調的資料模式聯想在一起。為組織他們的儲存至電腦中，這些基本的建築街廓可被形式化到邏輯的圖表資料中 (被認為同資料結構)，那被使用去表現實體和連續資料模式兩者。表 3.1 摘要了不同資料結構的特性，同時重現實體位置資料、屬性儲存庫和位相關係。資訊有系統地陳述依據地理資料模式並不是必要去被組織到電腦中使用相各字的資料結構，網格資料結構可被使用去表達資料系統的陳述，依據向量資料模式，反之亦然。

表 3.1：在向量式和網格式中基本空間實體的特性（原文 p.40）

基本單元形式	位置資料	屬性保留方式	位向關係
向量單位			
點	X, Y, (Z)	記錄	隱藏的
節點	X, Y, (Z)	記錄	明確的
簡單線	N[X, Y, Z]	記錄	隱藏的
複雜線（弧）	N[X, Y, Z]	N[記錄]	明確的
多面體（立體）	M[lines]	記錄	明確的
網格單位			
網格單元（像元）	排、欄	單一價值	隱藏的
線	N[像元]	單一價值或指示物去記錄	隱藏的
面	M[像元]	單一價值或指示物去記錄	隱藏的
voxel	排、欄、層	單一價值或指示物去記錄	隱藏的

### 點、線、面：向量資料結構(Point, lines, and areas：vector data structures)

使用點、線、面或面去描述地理現象的資料結構被認為是一個向量資料結構，向量單位被賦予現實的特性，它們的地理位置可被獨立且精確地定義，像可能是它們的位相關係。它們通常不容許內部的變動（除非它們是較簡單單位的組成），所以它們的屬性提及全部的單位。

一個空間的現象被以它的地理位置及屬性所塑造。一個油井可被一個由單一對 XY 座標及“油井”層所組成的點單元所表達出。油管的部分可被一個由一個開始的 XY 座標和結束的 XY 座標所組成及“油-管線”層所組成的線單位所表達。煉油廠可被包括一組的 XY 座標加上“油-廠”層表達出。層可是像在這裡所給的實際名稱，或是他們可能是對照圖例或加入資訊的表格的數目，或者他們可能是空間符號。

簡單空間單元對有效率的資料處理是罕有充分的。當他們在資訊系統中是單一單位時可以去處理更複雜的現象是非常有用的，最簡單的複雜物件是“線”、“鏈”或“弧”（GIS 的專有名詞在同義字及近同義字上是豐富的），被用來表達一個彎曲或不直的線。當線連接，像在河流的接合點或道路交叉點，必須被稱為節點（node）的特殊一種點所連結。節點的機能是隨路線去帶資訊，線是被鏈結在一起且交通如何從這處流通到另一個地方。如果線不藉節點連結而橫越，電腦不可以區別橋、地下道或十字路口。

節點也被用來帶著有面的邊界弧線如何鏈結的資訊及提供有關左邊和右邊鄰近面和封閉島嶼的資訊。被用來連結弧到面的網路的模式稍後被更詳細的描述在本章。向量單位的屬性被儲存於電腦檔案中，如紀錄或 tuples 可藉點被連結它們。

在幾何學資料中典型的運算結合向量資訊的層去引出答案到布林型式問題或到在層之內且越對層各種單元間量的關係（幾何學的或位相關係的）屬性基礎的運算允許研究發生的一組特殊的價值或為新價值從存在的資料的中被計算。

### 網格單元：網格資料結構(Grid cell：raster data structures)

空間現象也可被數組規則形的鑲嵌單位（描述在第二章）所表達出。最簡單的形式是方形網格（像元）且鑲嵌的規則格子被認為是網格資料結構。實體的位置被定義直接關於格子網格，帶有與在地球表面一群方形的土地聯結在一起每個像元，網格資料的分解或大小甚至是像元大小和地表網格的像元大小間的關聯。但是在向量資料結構，不同單位間的位相關係



是被透過資料庫指示者明確地紀錄（稍後解釋於本章），在網格資料庫中，這只是被透過在像元中的屬性價值暗中地編碼。

現象的改變可帶有一個真實價值的像元數目，在格子陣列中表達。每一個屬性正被藉分別的疊圖所描述。在更複雜的網格結構中，網格單元價值可藉指示者去和一個帶有很多各別性的價值之紀錄做關聯。新的屬性可被從存在的屬性中產生（例：藉分類）或藉實行各種的空間研究或同等的運算。測量的資料很少是有效的去直接產生這些表面，所以內插法的技巧常被用來改變取樣點至連續的表面（見第五章）。

典型的運算包括布林修正和鄰近疑問。除了數學的模式與藉結合資料從各個層為相同或鄰近網格進行網格資料結構，藉結合資料從各個層為相同或鄰近網格，除了數學的模式與進行網格資料結構是相對簡單的。

### 向量式的電腦儲存及網格資料結構(Computer storage of vector and raster data structures)

這些不同的資料型式其儲存庫之需求在表 3.2 中摘要敘述。

表 3.2：資料結構和電腦編碼（原文 p.41）

資料結構		電腦編碼的資料型式
向量單位	幾何資料	X,Y,Z 座標通常是非向量的或使用 32 或 64bits 真實的資料型式有較高的可變性，可能是二進位、名義上的、次序的、整數的或實數，或定向的使用 16、32、64bits
	屬性資料	
	位相關係資料	
網格單位	格子位置及大小	32-bit 實數
	屬性	16-bit 整數；為數學模擬的 32-或 64-bit 實數

### 資料結構：電腦中的資料組織(Database structures：data organization in the computer)

了解細項之前，在空間實體的方法可被有效率地儲存於電腦中，我們必須先了解為最佳的儲存庫和儲存的組織資料其一般的結果。雖然為 GIS 使用者去詳細地了解資料如何被排序入電腦中式不必要的，較一輛車的駕駛更需要去知道有關內部引擎燃燒的運作，資料塑造的一點知識，而資料建構模式將幫助去更了解系統如何運作，以及他們的極限和平均值可能是多少。跟隨的層面只表示一簡單的導論。感興趣的讀者應查閱規範工作在資訊儲存庫和搜尋如 Date (1995)，Salton & McGill (1983)，Ullman (1980)，Whittington (1988)，Wirth (1976) 的研究。

### 檔案與資料的擷取(File and data access)

任何資料儲存系統的本質特徵是他們應該可以去允許資料去被儲存且快速對照。有數個達成這個的方法，某些是較其他方法更有效果的。不幸地，那似乎不是用在所有情況單一“最佳”模式，那解釋在部分大量投資在人力及金錢上，在有效的資料庫管理系統上，如電腦工作，被認為是控制從數字資料庫中資料的輸入、輸出、儲存和修改。

### 簡表(Simple lists)

資料庫最簡單的形式是所有項目的簡單表列，當每一個新項目被加入資料庫，它可被簡

單地置於檔案的尾端，檔案會越來越長，很容易去加入資料到這樣一個系統中，但修改資料上是沒有效率的。一個表列包含  $n$  各項目，需平均  $(n+1)/2$  個搜尋步驟去發現你所想要的項目，因此，對一個在卡中包含 10000 個敘述項目的資訊系統而言，給 1 秒去讀卡名和數目，將會需平均  $(10000+1)/2$  秒，或大約 1 個小時 30 分鐘去尋找你想要的卡。

大部分的人們知道透過非結構的表列有如“海底撈針”的搜尋是很沒有效率的。因此，把資料排序或組織化是一個必要的步驟，且提供關鍵字去建構以便快速修改資料。

## 排序檔 Ordered sequential files

字典中的字或電話簿中的名字是依字母排序來建構的，加入的新項目意謂額外的地方必須去插入而產生，但優點在於被儲存的項目可被很快速的搜尋。常常排序檔被二進位搜尋程序所儲存，代替在表列的開頭開始搜尋的方式，由中間的記錄開始檢查。如果關鍵值（即字中連續的字母）適合，則中間的記錄是正被搜尋的那個；如果值不適合，簡單表列被用來查看被要求的項目是出現在中間元素的前還是後。檔案適當的一半是被保留且重複搜尋，直到項目被歸位。二進位搜尋要求  $\log_2(n+1)$  個步驟，如果檔案有 10000 個項目那麼長，且搜尋項目/項目是 1 秒，比較於先前的 1.5 個小時尋找一個項目的平均時間是接近 14 秒。

## 索引檔 Indexed files

簡單連續的和排序檔要求資料根據關鍵屬性去做尋找，以字典的例子而言，關鍵屬性是拼字，但在很多應用上，特別是在地理資訊上，個別的项目（像元、點、線、或面）將不只有一個關鍵屬性，如：一個辨認數目或一個名稱，但也將帶有關聯屬性的資訊。常常資料帶有關聯屬性是被要求的，但只在於名稱。例如我們可能有一個已被土壤系名稱所建構的土壤斷面的排序表列，但我們會喜歡去尋找資訊，如：土壤深度、排水、酸鹼值、質地或沖刷。除非我們採用另一個資料庫方案，或我們的搜尋程序恢復到那些簡單連續表列。

運用索引檔儲存到原始資料檔可以在兩方面增快速度。如果資料項目在檔案中，它們自己提供檔案的主要排序，則檔案被認為是直接檔（direct files）（見表 3.3a）。在主要檔案中項目的位置也可被特別的依據被給予在第二個檔的項目，被認為是轉換檔（inverted file）（見表 3.3b）。如同書中所示，索引的檢查可決定項目去符合搜尋要求。

在直接檔中，為每個項目作記錄含有充足的資訊為搜尋去跳過不必要的項目，例如：了解一個資料檔案包括土壤系的名稱已被如字母般排序。每一個項目不只包含一連串的名稱和其他資訊，也包括了一個顯示資料庫位置帶有相同字母開頭的連串名稱的數目字。為特殊記錄的搜尋接著被更減化藉著建構一個簡單的索引檔，列出連串名稱的第一個字母和儲存庫位置間的相對關係。搜尋的程序藉索引的一個連續尋找，接著利用私人資料區的連續搜尋。搜尋步驟的平均數目是  $(n_1+1)/2 + (n_2+1)/2$ ，這裡的  $n_1$  是索引步驟的次數， $n_2$  是在資料區被索引所提及之項目的數目。

轉換檔的使用索引是第一個要求，可藉為每一個項目執行在資料中最初的連續搜尋所建構。結果是被聚於轉換檔或索引中，那提供更長遠資料儲存的關鍵。（見表 3.3c）

索引檔允許迅速地儲存到資料庫中，不幸地，它們有固有存在的問題，當使用運用檔案在檔案中，記錄會連續的被加入或刪除，例如：當發生在交互作用地圖系統時。直接檔案的記錄增加或刪除意謂檔案和它的索引都需被修改。當新的資料被寫入一個由轉換檔索引所記錄的檔案時，新的記錄不必是位於特殊位置上的，它可能被簡單地加入檔案的末端，但索引必須被更新。檔案修改可能是一個帶有大型資料檔昂貴的進行，特別是在一個交互作用的環境下。

索引檔更遠的不利點在於資料常常只有關鍵字被儲存在索引檔內，其他的資訊只能用排序查詢法獲取。

表 3.3 (原文 p.43)

## (a) Indexed files

Index		File item
Item key	Record Number	
A	1	A <sub>1</sub>
B	n <sub>a</sub> +1	A <sub>1</sub>
C	n <sub>a</sub> +n <sub>b</sub> +1	...
...	...	B <sub>1</sub>
...	...	...
...	...	C <sub>1</sub>

## (b) Inverted files

Soil profile number	Attributes					
	S	pH	De	Dr	T	E
1	A	4	deep	good	sandy	no
2	B	5	shallow	good	clay	yes
3	C	6	shallow	poor	sandy	no
4	D	7	deep	good	clay	yes
5	E	4	deep	poor	clay	no
6	F	5	shallow	poor	clay	no

S=soil series, De=depth, Dr=drainage, T=texture, E=erosion.

## (c) Index inverted file

Topic	Profiles (sequential numbers in original file)					
Deep	1			4	5	
Shallow		2	3			6
Good drainage	1	2		4		
Poor drainage			3		5	6
Sandy	1		3			
Clay		2		4	5	6
Eroded		2		4		

## 資料庫結構及資料庫管理 (Database structures and database management)

如上述例子所示，依賴真實世界實體的資料被寫入至電腦檔案中，這些檔案是簡單或有組織的表列。檔案的基本單元是資料記錄或 tuple<sup>1</sup>（可視為一種記錄），對每一個實體而言，是包括所有聯結的資訊。依賴資料種類的蒐集，資料記錄可能詮釋相同或多變的長度，且測量可被信賴在二種類型能夠有效地被編碼（Box 3.2）。然而，空間資料庫包括許多資料

<sup>1</sup> **Tuple**, a set of values of attributes pertaining to a given item in a database. Also known as a record. **Tuple**, 在資料庫中屬於一個給定的群項的一組屬性值。也如同大家所知的「記錄」。

的檔案，聯結相同實體層面或因為空間鄰近度或連接度必須連接或組成一起的實體資料。組織者些檔案被儲存及連接在電腦尚未模組實質世界現象及確認有效的儲存及資料修正是必須的。被設計用來儲存及管理大量資料的電腦被稱為資料庫管理系統 (DBMS)。

現代的資料庫管理系統唯有校地儲存及修正資料而使用很多種方法 (例如 Date 1995, Worboys 1995) 但全部已組織資訊的三個基本方式為基礎，也反映邏輯的模組使用在製作真實世界結構的模型：它們是層級、網絡及聯結圖解綱要，且全在地理資訊系統中被使用。近來一個被稱作物件導向的第四結構已於 CAD-CAM (電腦輔助設計 (computer-aided design) - 電腦輔助製造 (computer-aided manufacturing)) 及地理資訊系統所使用 (指的是“O-O” (物件導向))。“O-O”是為了作有效地模組介於空間實體，網絡模組的更深入發展，能夠相故連接及聯結。現在物件導向在某些商用地理資訊系統中被使用，雖然在發展概念的研究仍持續中。

做資料庫的模組是設計可呈現效率包括正確資訊的一個資料庫的工作。有一個邏輯的結構及易可能維護及延伸的... (Worboys et al. 1990)。為了解在 GIS 中如何被達成... 檢測不同組織結構的基本原則是首要的，且去看它們在空間資訊中如何是被使用，這些資料不是正確的實體，就是被紀錄使用的連續性資料模式 (查第二章 (第二章 Data Models and Axioms: Formal Abstractions of Reality 資料模式與通則：真實的抽象表現))。

### 層級資料庫結構(The hierarchical database structure)

當資料有一個「成人-兒童」或一對多的關係，正如行政土壤土層分級或多邊形不同層級區域的像素單元，層級法提供資料存取快且便捷的方式。資料組織的層級系統正如所知是為東植物分類'土壤分級等等的環境科學法。它們假設層級的每一部份能夠利用一個可完全描述資料結構的關鍵點而到達介於關鍵屬性 (不同準則) 級項目可能控制的組成屬性，良性的聯結性假設也可被成立。

圖 3.1a 說明二個多邊形圖的簡例，圖 3.1b 說明一句層級模式資料如何被編碼，模組主要的優點是簡易級存取定義層級的關鍵容易。它們是易於瞭解、資料更新、延伸以及在大量儲存系統中組織資料是有用的 (見 Kleiner and Brassel (1986))。在關鍵屬性中資料存取是容易且快速的，然不性的是對組織屬性是相當困難的。因此，假如所有可能假如所有可能查詢的結構已事先可知，層級系統對資料修正是好的。目錄學、銀行或航道修正系統的案例是常見的。然而，為了環境資料，許多修正要求的探究本質不能藉由嚴謹的層級及使用者拒絕系統，如同不可能堅固的，是不能供給的。例如 Beyer (拜耳) (1984) 報導英國邱園 (英國皇家植物園) 起初建立為超過一百萬份依據植物分類原則而製作的草本植物標本層級的資料庫，直到邱園主任欲到墨西哥 (Mexico) 採集新的草本植物材料且查詢這些已從墨西哥採獲的邱園植物資料庫，這個問題才被發覺到 (自註：當初建立資料庫時，未考量的問題稍後被發現 (如上述問題))。不幸地是為了資料庫的供給者及使用者，邱園主任得不到任何 (上述問題) 的答覆，這是因為標本採集地並不是植物分類關鍵的部分。

層級資料庫結構更多的缺點是大量索引檔案必須被維護且某些性質可能必須被重複很多次，導致儲存量增加及存取成本的資料冗長。圖 3.1b 說明每一座標組的冗長可能必須被重複二次，且座標 3 級 4 可能被重複四次，因為 c 線必須被重複二次。層級結構也可能是空間的過度浪費；例如，若操作被製成多邊形 I 及 II 相同名稱，沒有容易的方式去將可能不需要展現的 c 線展現穩定。

Box 3.2 (原文 p.45)

### 網絡資料庫結構(The network database structure)

在層級結構上，資料庫的行逕路徑被限制在有分類路徑的前後（上下）。在很多情況中，很多更快速的連結被需求，特別是在資料結構中，為了儘管有關座標的正確資料在資料庫非常不同的部分被寫入，在地圖中鄰接項目需要被連結在一起的製圖特質。

在冗長及連結的問題，藉由如在圖 3.1c 所示的簡潔網絡結構可被避免掉，每一條線組及每一組座標僅需要顯示（出現）一次。使用者個結構無論何時藉由相同名字的多邊形參數之去穩定 c 線組的模型。因此製圖一般化更容易。在製圖學中，網絡結構中的環形指示（ring-pointer）結構常被使用。環形指示（圖 3.1d）是如多邊形網絡似的，對具複雜位相（拓撲）結構的航海環境而言，是非常有用的方式。

在複雜系統可能變成資料庫完全實質的部分，每次指示必須被維護稍作一個改變，為了資料庫及指示結構的建立及維護對資料庫系統可能完然過肩可考慮的。

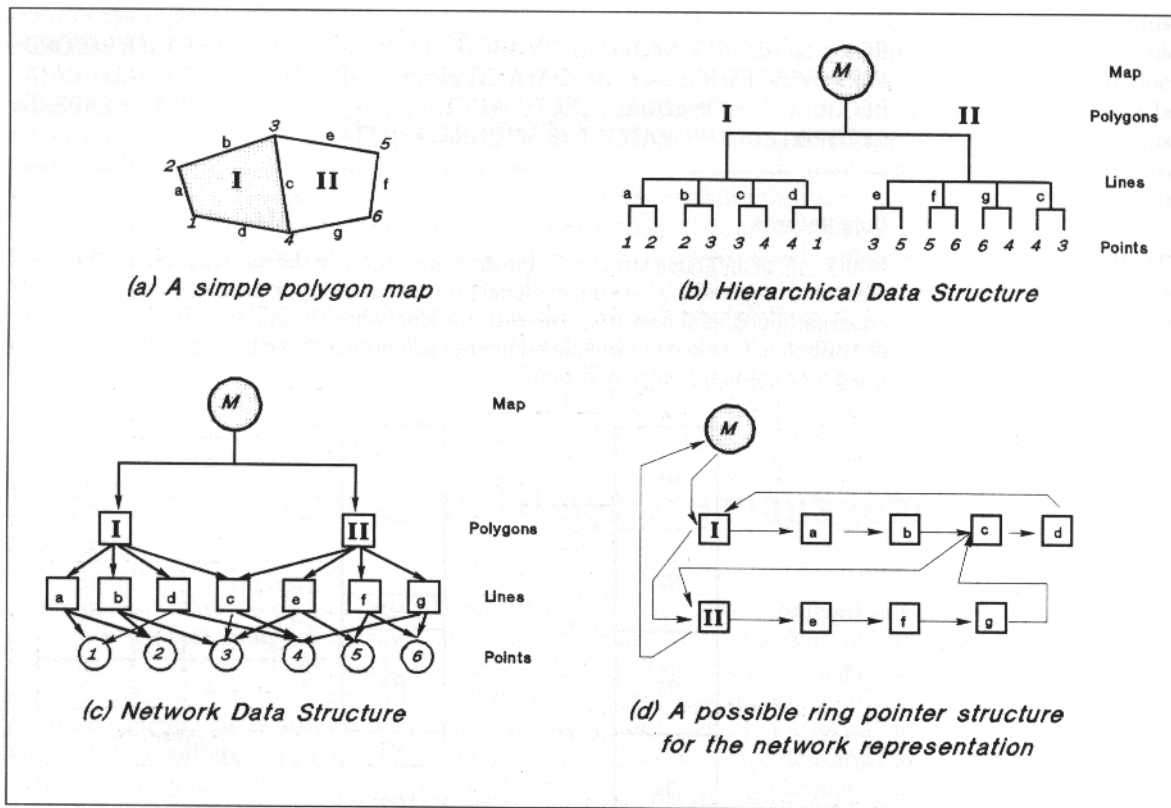


圖 3.1 (原文 p.46)

### 關聯性資料庫結構(The relational database structure)

在最簡化的形式中，關聯的資料庫結構儲存沒有指示及層級。取而代之的是在簡單記錄中，資料被儲存，如同所知的“tuples”，是每一含括一個屬性的領域組群；tuples 在二維表格中是被組合起來，如同所知的關聯性，試算表計畫的方式。每一個表格或關聯性是一個很有用地個別檔案。在網絡結構中的指示結構及層級結構的關鍵藉由辨識形式的資料冗長是被替代的，（而這個辨識形式）在每一個個別的檔案中，編碼備用來為辨識記錄而視為獨一無二的關鍵。

透過使用者定義問題適宜性的關聯性，資料從關聯性的資料庫被粹選出來。關聯性在存在的檔案中並不需要呈現，所以控制計畫使用連結代數的方法去建構新的表格。這些規則長

在稱為結構查詢語言 (SQL: Structured Query Language) 所編碼一見 Date (1995) 及第六章 (第六章: Optimal Interpolation using Geostatistics 由地理統計學做適當的空間修正)。圖 3.2a 顯示圖 3.1a 檢試圖的關聯性結構包括許多我們熟知的冗長及方法如“標準化”常被用於建立更多有效的編碼。例如寫入位址線組當作直線部分的組群簡化維沒有資訊流失的關聯性結構 (圖 3.2b)。

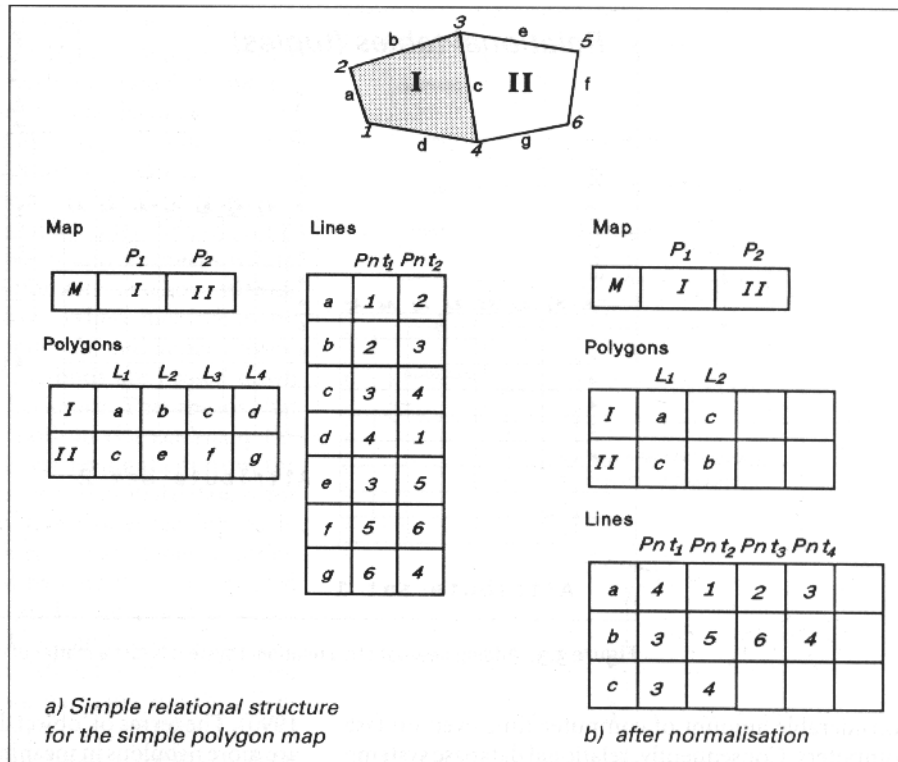


圖 3.2 (原文 p.47)

聯結資料庫有極大的優點，它的結構是非常易操縱請可能滿足所有查詢的需求，就是能夠使用數學操作及布林邏輯的規則而被公式化 (形式化)。它們允許資料以不同類型被搜尋、結合及比較。資料的增加或移除也是容易的，因為這正涉及加或減一個 tuple，或甚至是一個全表格 (如圖 3.3)。查詢存取不同聯結的表格藉由透過常用的區域加入而製成。對那些全部記錄有相同數字屬性及沒有自然地層級的情況而言，是好的。然而，介於標隔間的聯結處是複雜且加入的數量是被需要的，操作可考慮大量的電腦時間，甚至在速度快的電腦群中。因此，聯結資料庫系統必須非常有技巧地被設計，為了提供合理速度的搜尋容量 (這就是為什麼它們價格昂貴的理由)，在 1980 年初期，電腦首次用於地理資訊系統上 (Abel 1983, Lorie and Meier 1984, van Roessel and Fosnight 1984)，且現在它們在許多知名商用地理資訊系統上，為處理空間實質屬性已建立成為主要的工具。

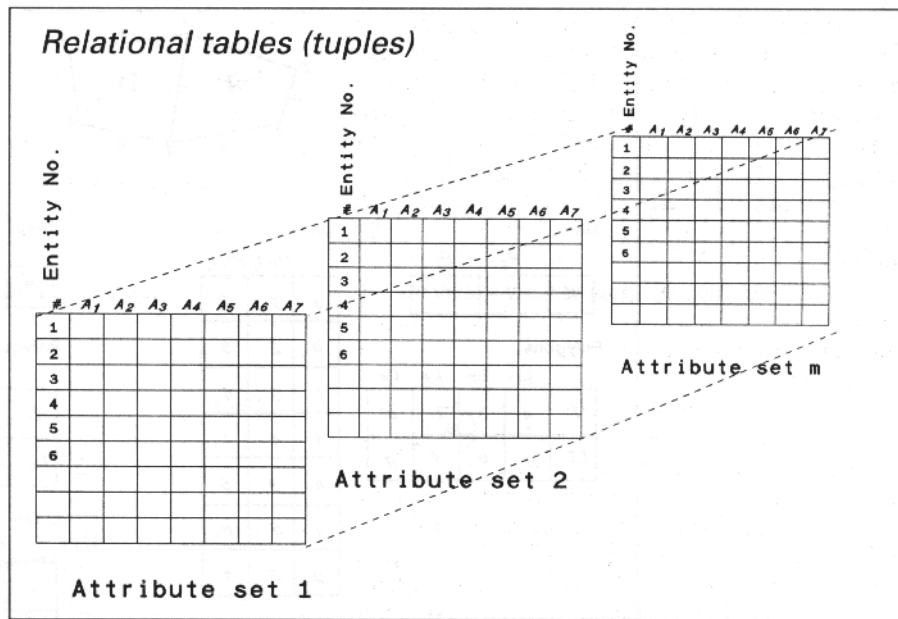


圖 3.3 (原文 p.48)

### 物件導向的資料庫結構(The object-oriented database structure)

物件導向的概念始於程式語言如 Simula (Dahl and Nygaard 1966) 及 Smalltalk (Goldberg and Robson 1983)，且意念應用到資料庫藉由在聯結結構中，冗長及依次搜尋的問題而被模擬。在地理資訊系統中，它們的使用藉由去處理複雜空間實體比一個簡單的點、線、原始的多邊形的需求更智慧化已被模擬，當分析運作多邊形疊圖一樣且也藉資料庫修正的問題被展現出來 (見第七章：The Analysis Using Continuous Fields 使用空間中連續區域分析)。雖然有許多在定義使用電腦群 (computing community) 沒有明顯的一致性對概念的許多不同形式化而言 (Worboys *et al.* 1990)。物件基礎項目或物件中心在意義上 (方法上) 比物件導向程式語言可定義的特質更為模糊不清 (Jacobsen *et al.* 1992)。

被發展使用物件導向的程式語言的物件導向資料庫結構，組合介採用組織周圍正確實物當作對抗正被程序化方程式的資料聯結可塑層級及網絡方式的速度 (Chance *et al.* 1995, Kim and Lochovsky 1989)。在聯結結構上，邏輯聯結的實體資料記錄群項中，資料在一句任何本質結構被組織併入相似現象的組群中 (正如所知的目標 (物件) 階級) 的一連串獨特單一物件的群項中被定義。透過明確清楚的連結，介於不同物件及不同階級聯結被建立起來。

一個物件的特質可能在屬性群項 (稱為狀態) 及一組描述行為的 (稱為運作或方法) 程序的資料庫中被描述。這些資料藉由資料庫中獨特的定義者被定義的物件包裹著。這都保留相同的位描述它們特性數值的任何改變。例如一個建立的物件可能在超過結構的群組中的時間而改變，但它獨特的鑑識者將完全保留。腳踏車的聯結模組僅是部分的表列；物件導向模組連結所有部分，如此它們的功能在相互聯結及物件的行為是很明確地被表達出來。

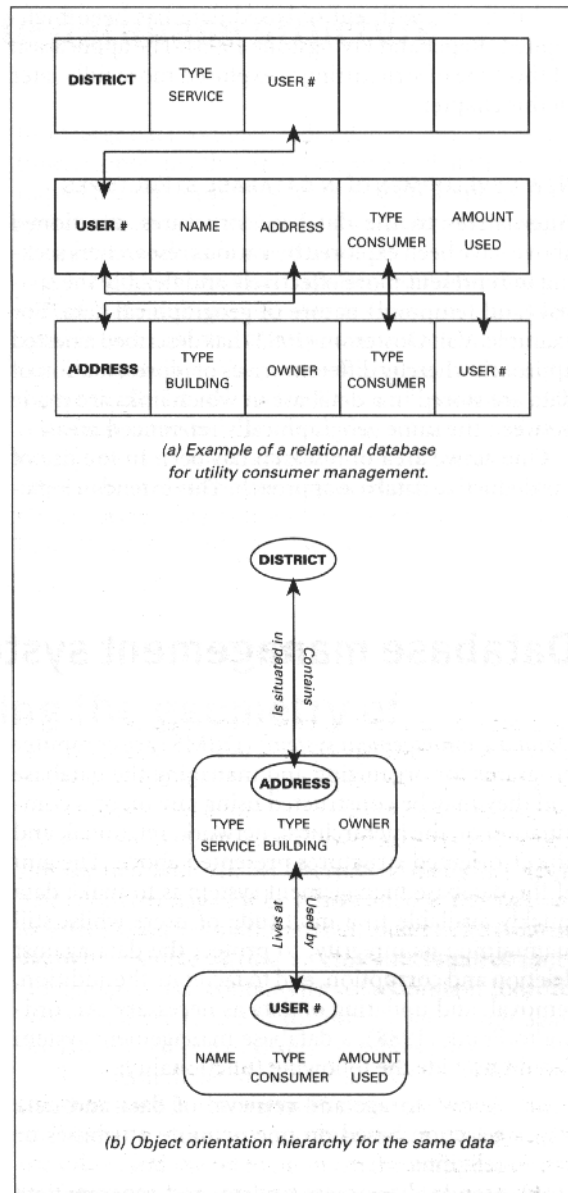


圖 3.4 (原文 p.49)

在資料庫中的物件結構，運用直接涉及獨特鑑識者的指示而被建立。在它們之中，等級及實例藉由指示去呈現不同的聯結及層級而被連結起來。依循前則，向上及向下的層級被建立，不同定義狀態或方式的層級透過承接的系統被延續下來。這意味效率可能在賦予物件屬性特質及從資料庫修正中被建立起來。圖 3.4 比較為了採取一個物件導向方式通用群件組 (utility company) 的聯結資料庫。在每一個主要表格聯結方法 (區域、使用者 #、位址) 藉著從一個表格至下一個表格被重複的資料所連結。在介於它們被定義聯結的物件導向方法中，區域、位址及使用者 # 被設計成物件，正如含括、位處於~、藉~被使用，處居於~等等。所有資料僅使用一次，且直接指示不僅提供資料快速遞修正，也滿足來回傳遞指令資料庫層級。例如在已被給予的區域未判斷它們列入的安排，使用物件導向方法區發掘被給予的類型的所有使用者是更為容易的。

在資料庫中，資料已被包裹起來視為一物件，正如去改變它們或查詢它們傳遞送達一個需求，如訊息去實現一個運作 (Chance et al. 1995)。質疑問題的類型視為已經被使用定義物件的運作而定。當介於不同物件接收或在不同的涵構，物件回應到一個訊息將視可能攜帶



不同反應的相同訊息或狀況而定；這是被群組成多形式的。

在物件導向資料庫中，被使用的資料需要如同獨特實體般的清楚定義之。提供對組織層級，交互聯結資料非常有效的結構資料庫（如網路及層級相對應物）。建立資料庫顯而易見的是時間耗率，如物件被定義為清楚明確且多樣的連結需要被建立起來。資料庫提供非常有效率的結構被完成，特別是對於採取參數至明確的物件的查詢。聯結資料庫的資料在已屬性值為基礎所呈現的查詢上比較好，雖然專門的索引方法及簇群技巧對某些資料庫正已被開發（發展）（Arctur and Woodsford 1996）。

為了地理資訊系統採用物件導向的資料庫結構的可能性被提供在近幾年藉許多研發者而被研發（例如 Worboys 1994）且它們採物件基礎資料的使用也已經很顯著（Raper and Livingstone 1995）。結構類型的應用在本章節稍後會有更完善的探討。

### 在資料庫結構上的新研發(New development in database structures)

藉不同研發者探尋呈現更有效率且堅穩的地理資料空間（或時間）的本質，導入以資料結構上述的替選方案已被探究。例如 Van Oosterom（1993）藉在介於已被連結組成相同地理參數區域，已被儲存的網式（nested approach）資料代表的不同尺度資料庫已被描述中。

研究熱門的領域在推演式的資料庫方法的使用已存在。這延伸一個以邏輯為基礎的方法至資訊儲存，查詢及使用程式語言如 Prolog（電腦人工智慧語言中的一種）的過程步驟。推演式資料庫儲存資料及定義一個事實或表達聯結性的邏輯；資料記錄以延展在聯結資料的形式是相類似的及強烈表達邏輯運用其它聯結性的實質聯結的部分所組成（Yearsley and Worboys 1995, Worboys 1995, Quiroga *et al.* 1996）推演式資料法的優點是允許可能比承擔使用聯結的、網絡或層級結構更複雜的物件及模組去建構且對空間資料處理，提供更多可能性。

以這些資料庫結構為基礎的地理資訊系統發展，目前在理論階段仍是必要的。

### 資料庫管理系統(Database management systems (DBMs))

資料庫管理系統（DBMS）對組成及管理資料庫而言，是電腦程式，且它們可能被件夠用在層級、網絡、聯結及物件導向結構的組成或任何一個地方展現。資料庫管理系統的目標在於製作可快速獲得的資料，對大多數使用者而言，當維護它的公正性去保護資料對抗刪除及損毀，且使得添加、移除及資料的更新更為容易。依據 Frank（1988）所提出的資料庫管理系統，應具下列功能：

1. 允許資料的儲存及修正，及已一或更多屬性或聯結為基礎的資料選擇。
2. 標準化至應用程式以維護在程式中的獨立性從資料的使用資料，及分離資料儲存及存取。
3. 提供一個介面（界面）在資料庫及以沒有實質儲存需求細項的邏輯描述，為基礎的應用程式。
4. 在實質儲存結構的應用獨立性上，製作一個存取方程式（功能），為了界在儲存媒介改變而程式不受其影響。
5. 允許好幾個使用者可同時存取資料。
6. 從未經選擇或非法改變過程中，保護資料庫。
7. 提供審慎的法則對自動執行資料的一致性。這些法則從資料庫中移除錯誤及不協調性是一個很棒的方式。

大部分的資料庫管理系統允許透過高層級程式語言及親使用者結構查詢語言（SQL：一

種應用在電腦資料庫查詢上的程式語言)來存取資料是最為常見的。當高階程式介面允許資料庫如地理資訊系統，被直接連結到應用程式之中，使用者介面允許資料庫的非正式查詢。

好的資料庫管理系統也將確認為降低資料存取及電腦轉化過程的時間在儲存媒介的實質鄰接區域，空間相偕的資料將被儲存。快速地到達大量資料之中，回應使用者需求面上操作順利，已成為設計地理資訊系統的重要(設計)點。許多地理資訊系統使用資料庫管理系統當作系統去利用資料處理容量的一部份。

### **選擇最適的資料庫結構(Choosing the most appropriate database structure)**

這四個基本資料庫結構應該顯而易見—層級、網絡、物件導向、聯結性—為空間資訊系統提供一些東西。層級系統允許大的資料庫被容易分割位可管理的質、量，但為建立新的搜尋路徑是鑑定的，且它們可能包括許多冗長的資料。網絡系統包括少許冗長的資料，若任何快速且直接提供，以及兼顧的聯結介於聯結的實物之中。物件導向系統允許聯結性、功能性、持續性及相依性，為的是建立一個系統在程序工具的花費、電腦電源上更複雜及更多需求。聯結系統式開放的、易操縱的即可適應的，但許多遭大量資料量、冗長及長時間的搜尋。因此在空間資訊系統(SIS)技術比被安排到可管理的其中一項工作，(前者)常被為相互補強而一起使用。

一個層級式的方式對分割空間資料到可管理的主題或區域常是有用的，如此連續、無瑕疵的製圖方式就變得可能。網絡式對位相連結向量線組及多邊形而言是理想的。聯結方式對修正物件在屬性的基礎、或建立新的屬性，或者已存在資料的屬性值是好的。當物件在特殊方法中分配屬性或交互感應作用，這種方式對物件導向是有用的。

### **為展現空間現象幾何學的資料結構(Data structures for representing the geometry of spatial phenomena)**

這些主要的資料庫結構影響在電腦系統中地理資料如何被個別分及儲存，如地理資訊系統。許多系統不是善用向量式或網格式的疊圖，就是建造空間資料特徵面的方法到有用的群組。

### **網格式資料結構的資料組織(Data organization in raster data structures)**

網格式資料庫從使用者察覺許多的笛卡爾疊圖中被建立起來；給大量座標組織略圖企圖更理想化資料存取及儲存過程化的需求至最小化(見圖 3.5a)。

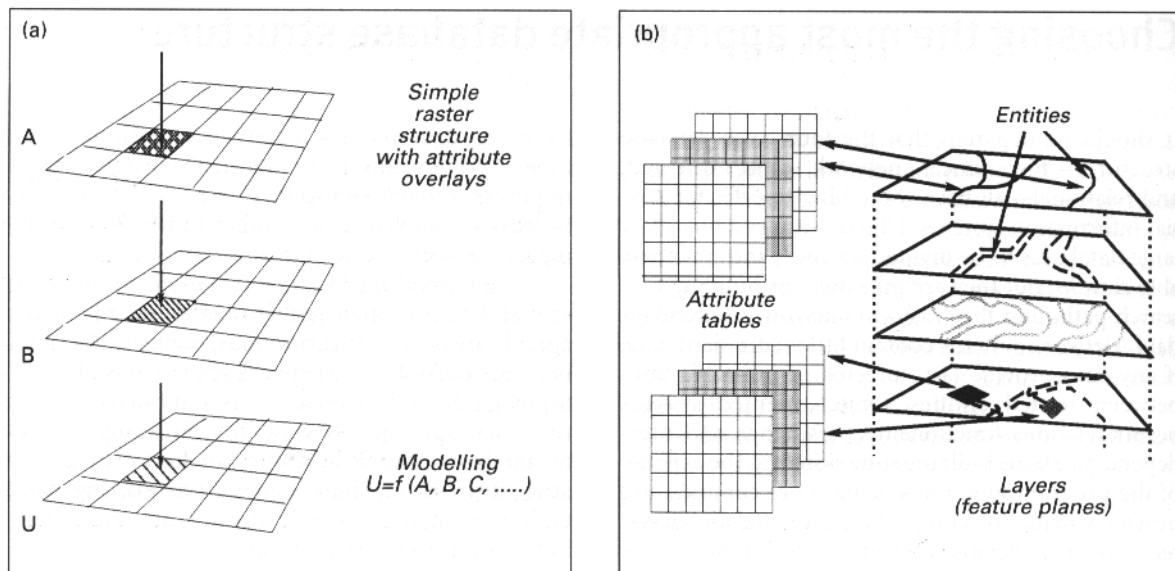


圖 3.5 (原文 p.52)

在每一個圖層的每一個別單元簡單的網格式結構被假設位在資料庫中是獨立不相關的單元（介於資料值、像元（pixel）及每一個圖層如圖 3.6a 的屬性值被定義。為降低儲存需求，在單元大小或展現的符號上沒有壓縮技巧被使用。在一個替選方案中（圖 3.6b），資料庫中每一圖層可能被展現當作單一屬性值的點的二維矩陣。這仍要求許多儲存空間當作它含括為每一圖層重複且單元大小或展示符號再一次沒有資料被掌控的冗長座標的表列。

圖 3.6c（在 Tomlin（1983）的地圖分析中已被使用）顯示層級結構在屬性值及製圖單元的點群建立一個多對一關係，所以製式的區域（多邊形）可能很容易地被標示。如展示符號、編碼或改變可變的事物、量等變得容易，當對立前二各結構的每一小單元值時，它要求再寫入對每一圖層每一製圖單元只有一個數字。資料壓縮技巧（在本章稍後討論）備用來降低儲存需求，且在處理結構資料上是有效率的。SPANS 地理資訊系統以上述這個層級結構的四分樹（quadtree）形成為基礎。這個結構主要的缺點是對連續區域的資料而言，是不靈活的（不好用的）。

每一圖層第四結構被儲存當作依各用一般資料起始點（header：包含發送及接收地址及其它錯誤檢查、認證等等信息）含括資訊的分割檔案，如圖的投影、小單元大小、行列數及資料類型；藉由依據行列排序被命令的簡單表列值所緊接著（圖 3.6d）。當座標值對每一個別小單元及一般幾何學而言沒有被儲存時，這顯得更有效率，且展示值在圖層的資料起始點已被寫入。PCRaster（Wesseling *et al.* 1996）就是使用這個結構。

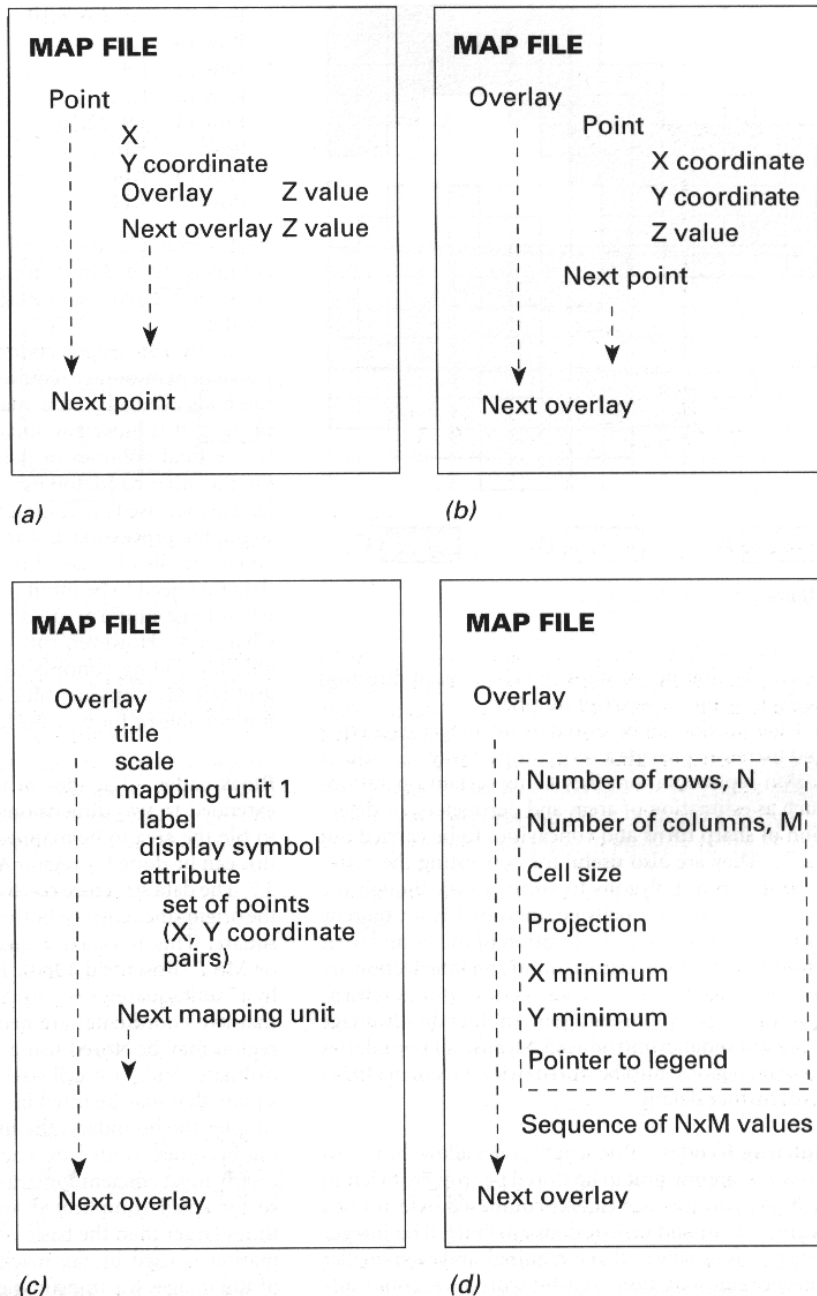


圖 3.6 (原文 p.53)

### 為儲存網格資料的壓縮方法(Compact methods for storing raster data)

當網格資料結構被用來呈現一個連續性的表面，每一單元有一獨特的值且它攜帶所有 n 個行及列去對每個圖層編碼，在圖的投影上、網格的原始點、網格大小以及資料類型。儲存最貧乏的資料圖層是無向量關係切每一單元包括一個真實數字的資料類型，正如對數化層級而言，以高度為基礎 (altitude matrices) 的案例中和從修改其他連續表面中獲得。因為在甚少數的位元中 (位元 (bits) : 電腦計算資料儲存的小單元) 它們能夠編碼 (如稍早前的討論)，另一資料類型將需求較少的空間。

當網格式結構備用來呈現具相同值像元的線組或區域，為假若以資料結構適當被設計為條件的網格式資料在儲存需求上影響可考慮的儲存方式是可能的。在圖 3.6a 及 b 所描述的結構可能使用列陣座標去降低數字實質正確總量的儲存，伴隨著所有空間運作必須在行列列陣數量被實現的限制上。這些系統在介於製圖單元值及小單元座標一對多聯結的形式上不去

作編碼的資料，所以對編碼而言，壓縮的方法是不能被使用的。

在圖 3.6c 參數每個地區（或製圖單元）的點組被給予的第三結構及允許壓縮儲存不同多樣的方法被使用。在為制圖單元或多邊形可能更有效地被儲存的空間資料上，有四個主要的方式：分別是鏈編碼（chain codes）、連續（運轉）長度編碼（run-length codes）、區塊編碼（block codes）及四分樹（quadrees）。

### 鍊編碼(Chain codes)

見圖 3.7 中，A 區域的邊界可利用專業術語說明其起源，及其組成單元的方向向量，其方向可以被數字化（東=0、北=1、西=2、南=3），例如我們要描述這個區域的邊界，可起始於第十行第一列，其編碼如下

$$0, 1, 0^2, 3, 0^2, 1, 0, 3, 0, 1, 0^3, 3^2, 2, 3^3, 0^2, 1, 0^5, 3^2, \\ 2^2, 3, 2^3, 3, 2^3, 1, 2^2, 1, 2^2, 1, 2^2, 1^3$$

數字代表了像元的數目，其方向由上標數字代表。

鍊編碼可以被儲存為完整的數據形式，所以可以以一個壓縮的方式去表現一個區域，它可以提供一個容易且正確的操作方式，例如去估計面積及周長或是去測量物體形狀的轉變及其中的凹窪。它也可以將多邊形的網格表示轉變為向量形式，不過由於其鋸齒狀的格子將被平滑化，所以在邊界位置的確定將出現誤差。在疊圖的操作上，若沒有完整的格子表現，在連結或交集上將出現困難，另一個問題是所有區域間的邊界在輸入時，皆被重複儲存兩次（Freeman, 1974）。

### 連續長度編碼(Run length codes)

連續長度編碼在每一個製圖單元，是以每一行由左至右的方式，儲存不同類型的資料，包括起始點、結束點及屬性，這完整的數據形式通常是全數需要的，甚至是一個 8 位元的字元層級，以滿足一些不同的需求。

在圖 3.7 的 A 區域其編碼如下：

第 9 行	2,3 6,6 8,10
第 10 行	1,10
第 11 行	1,9
第 12 行	1,9
第 13 行	3,9 12,16
第 14 行	5,16
第 15 行	7,14
第 16 行	9,11

在這個例子中，這 69 個網格在 A 區域中被 22 個數字完整的編碼，其結果將簡化這些被儲存數據所佔的空間。

清楚的連續長度編碼在傳統的資料儲存方式中，被認為是相當重要的多對一關係呈現方式，特別是在數據大小必需被有限度的維護時，換句話說，太多的數據壓縮將導致製圖過程及處理的增加，當大而同質的區域需要被數字化時，連續長度編碼（Run-length codes）亦可被用來減少鍵入網格數據的體積（見第四章），無論如何，連續長度編碼（Run-length codes）並不適用於連續不同的區域，因為每一個網格皆具有獨特的性質，且其數據並不能進行壓縮。

### 區塊編碼(Block codes)

連續長度編碼 (Run-length codes) 的原理可藉著四方形的製圖方式擴展為兩個不同的尺度，圖 3.8 表示在圖 3.7 的 A 網格區域，將如何進行製作，其數據形式包含 3 種不同數值包括起源 (中心或左下角) 及每個四邊形的半徑，這被稱為中間座標體系轉換 (medial axis transformation)，A 區域可被 17 個 1 單位的四邊形、9 個 4 單位四邊形、1 個 16 單位四邊形加以儲存，所以提供兩種數值形式是必要的，在 A 區域的儲存過程中，共有 57 個數字 (54 個座標數值、3 個網格大小)，顯而易見的，這大的四邊形可被區塊編碼轉變為區域或簡單的邊界。連續長度編碼及區塊編碼皆適宜表示大而同質的區域，不適宜表示小而變化的區域，中間座標轉換 MAT 可減少傳真圖像的影像大小，這個方法同時具有表示連結及交集性質的優點，例如 elongated-ness (Rosenfeld 1980)。

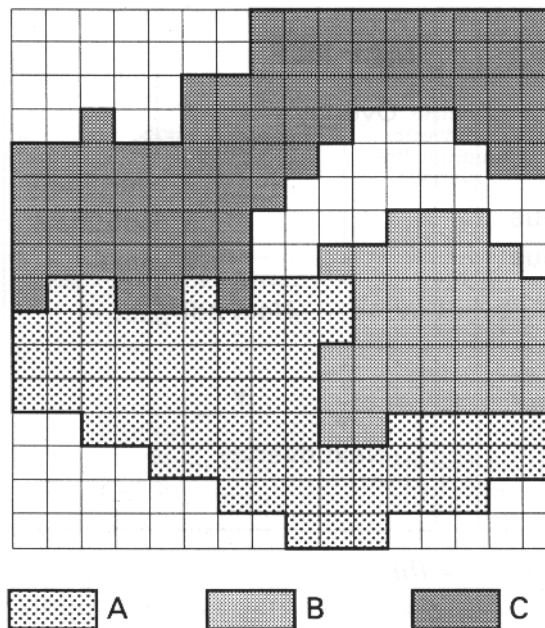


圖 3.7 (原文 p.54)

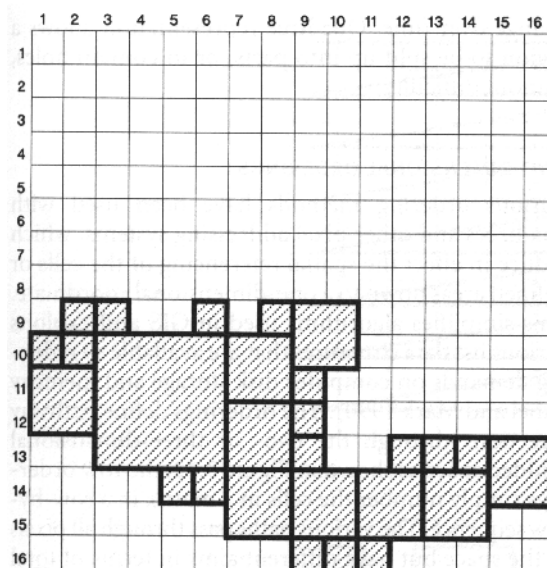


圖 3.8 (原文 p.55)

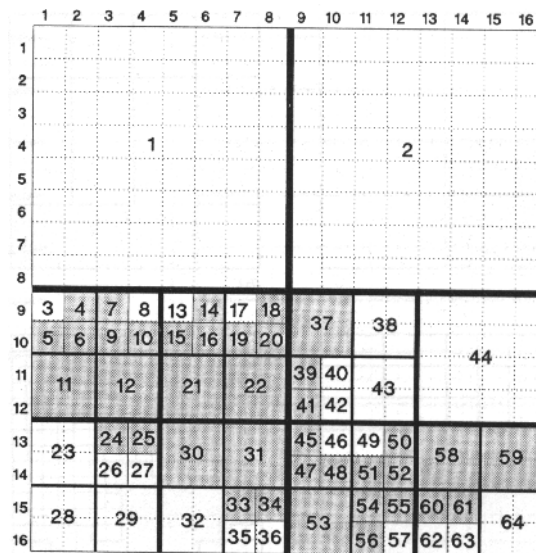


圖 3.9 (原文 p.55)

### 四分樹及二分樹(Quadtrees and binary trees)

利用這些規則的網格有一個問題，其數據的辨識率受限於基礎網格的大小，四分樹及二分樹對於細微層次的細節，提供了一些原理及無限的確立方式 (Samet 1990a, 1990b)。

這最壓縮的空間表現方法是利用連續性的基礎，在  $2^n \times 2^n$  的陣列中加以區分，如果這個每次將面積區分為一半的方法被稱做二分樹 (binary tree) (之後將詳細討論)，如果這個區域是由小區塊組成，且可一步步被四分的單元所包括，則此種分區被認為是四分樹，這是較常被使用的一種方式，這兩個例子中，其區分的限制皆為圖形的最小像元。

圖 3.9 可以利用四分單元表示 A 區域連續的部分，這個單元的結構被描述為階層 4 的樹，即為被認知的四分樹，在圖 3.10 中，陣列的  $2^n \times 2^n$  個點起始於樹的根部，而樹高指的就是數目最多的 n 層，每一個點有 4 個分支，各為 NW、NE、SW 和 SE，由於符合本區的需求，故不再需要更小的分區。

在 1980 年代，在 GIS 上有很多有趣的四分樹使用方式 (Martin 1982, Mark and Lauzon 1984)，它是一個清楚的表示工具，Samet (1990a, 1990b) 使用了具權威的運算原理去計算週長及面積，以轉換網格成為四分樹。對於其他的網格的表示方法而言，四分樹有一些有趣的優點，標準的區域性質可以被容易而有效的計算。四分樹是變數轉換 (variable resolution)，可用於陣列資料不足或並不是特別需要去紀錄的時候，在 3D 系統中，其四分樹的分析方法與 2D 情況。

如同其他方法，四分樹法所遭遇的最大問題，是無法轉化兩個形狀大小不同的四分樹區域，因此進行形狀的分析與組成的辨識時，並不是十分明確，這問題出在物體隨著時間移動或改變，但四分樹法可允許對物體進行分割或包括一些瑕疵，這是沒有問題的。

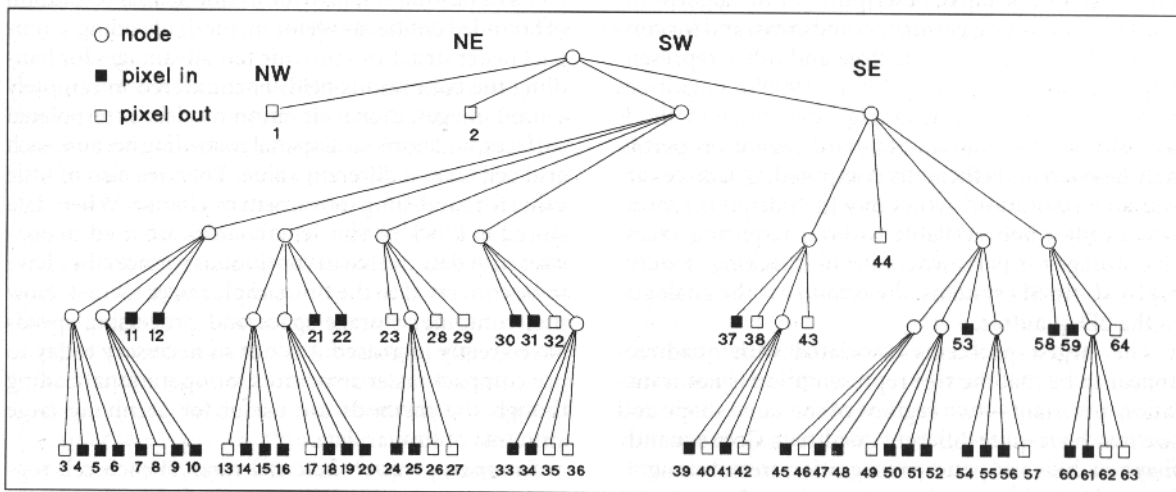


圖 3.10 (原文 p.55)

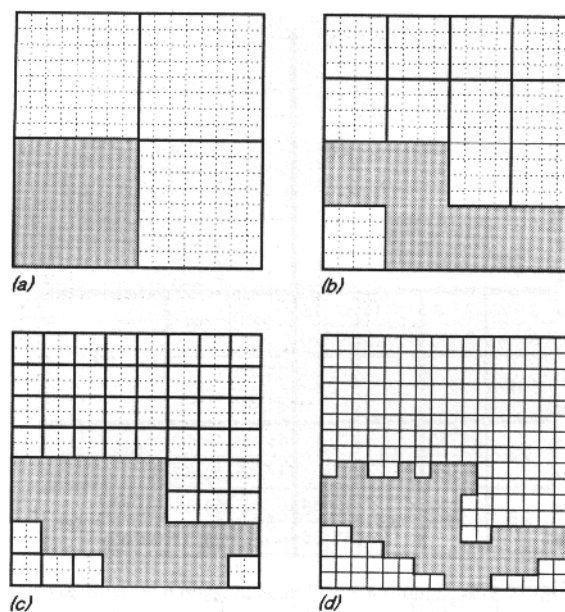


圖 3.11 (原文 p.56)

### 二次元的排序(Two-dimensional orderings)

四分樹及其他像元定位系統的排序方式，是為了減少網格或被定義範圍成為同質區域的空間影響，GIS 使用單一化的運算方式是為了減少電腦磁碟的儲存及記憶體 (Abel and Mark 1990)，利用一行接著一行的順序。圖 3.12 顯示出排序在空間路徑與方向上的定義，不同的路徑穿過空間中所有的網格，有不同的目標、術語、單位長度或連接方法，其中最常被使用的是圖 3.12c,d 的 Morton 和 Peano-Hilbert 方法，其在亞四分樹 (sub-quadrants) 中根據像元空間的回歸序列為基礎，在數據存疑及網格單位不同時，提供一個定位的技術。



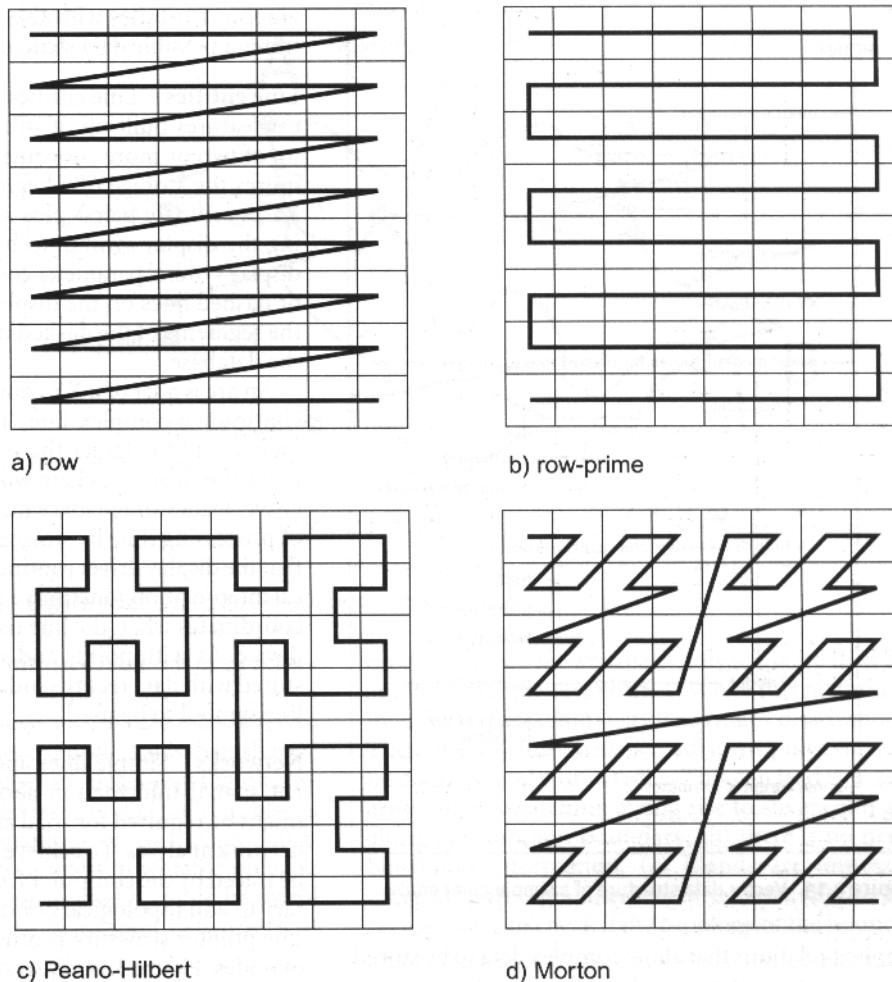


圖 3.12 (原文 p.57)

### 壓縮的網格形式與數據分析(Compact raster structures and data analysis)

當空間單位（像元、線或多邊形）正確或為靜態實體時，壓縮的網格形式可以有效的對數據形式進行分析，四分樹有足夠的層次，對於以向量編碼的固定實體，提供精確的空間分析，由於不同的網格具有不同的價值，壓縮的網格結構在處理遙感探測影像（remotely sensed images）、海拔高度數值模型（digital elevation models）、內插值表面（interpolated surfaces）、數值空間模型（numerical spatial modelling）等連續空間影像時提供少數優點，它們同時具有模擬運動及改變的小型運算。當數據被儲存為區塊編碼或連續長度編碼的形式以進行分析時，它被視為一個完整而壓縮的網格系統。現在以電腦儲存及其處理速度已相當進步，不在需要在處理大量相似數據時，利用這些方式進行網格結構編碼。

摘要—網格數據結構。如果每一個網格皆出現不同的結構，則這個  $N * M$  的簡單陣列在利用上會出現困難，它的限制與數據量及記憶體大小有很大的關係。在許多主題圖中，區域展示的儲存需求，受鍊編碼、連續長度編碼、區塊編碼及四分樹的影響，其中連續長度編碼對於處理需要考慮區域面積的展現及儲存的大型像元相當有利，當圖形的辨識率增加時，每個區域像元的數目也會增加，因此這時使用區塊碼和四分樹較為有利。四分樹方式對於變動的資料結構，可減少多重的處理過程。

### 向量數據結構的數據組織(Data organization in vector data structures)

一個向量資料庫的建立，來自於使用者藉著觀察區分不同種類的數字（Figure 3.5b），

這同等的空間將被假定為連續，這和量化的網格結構不同，可將所有的位置、長度和尺度被精確的定義。但這也不是完全精確的，因為其受限於長度如何用電腦文字確切的表示，及所有的顯示器皆具有最基礎的顯像單位。此外這個假設數學上正確的座標系，在數據儲存於最小的空間時，其間被暗示具有關係，因此在 GIS 向量結構的範圍上並沒有絕對適用的方法去儲存點、線、面資料。

### 點實體(Point entities)

點實體被認為包含了所有利用 XY 數對所表示的地理和圖形上的實體，根據 XY 座標，數據資料將被儲存且表示出點的種類及其附加資訊，數據並必需被具有意義的符號所表示，如果這個點是一個具有意義的實體，則數據的紀錄必需包括以下的資訊，包括：特性、形式、定位、尺度、方位及其他非圖面的屬性，圖 3.13 可說明一個點實體數據結構的形式。

### 線實體(Line entities)

線實體被定義為所有由兩個或更多的座標構成的直線段，展現的線性特徵。簡單的線是起點與端點（兩個 XY 座標數對）利用符號所表示，例如展示的符號參數可利用實線或虛線加以表示，甚至所有的虛線線段在數據資料庫中皆不被儲存。

一個鏈形的設定是利用連續而完整的 n 個 XY 座標數對所形成，利用最短的線段及最多的 XY 座標數對，一個完整的鏈形將被最近似的連結，數據儲存的空間及處理的時間可利用顯示器的數學功能，在線數據傳送入展示器時加以節省（例如 B-splines），當簡單的點、線與鏈的數據資料被紀錄時，可利用線的符號形式加以指示。

### 網路(Networks)

簡單的線和鏈並不能攜帶有關聯接度的空間資訊，譬如公路、傳輸或排水網路的分析，為了利用電腦以完成線性網路，這些數據模式需要增加位相的指標，這些指示的結構是為了幫助節點，3.14 圖示的數據結構排序必須先要再每一個網路分支中建立連通度，此外鏈形的指標傳輸其角度鏈結的節點，亦會帶來數據記錄的指標，可完整的分析整個網路的位相意義。這簡單的連結結構由於記錄了每一個成為一個總數（共 n 個連結點+1），其線段附屬的屬性將可被用來選擇最佳的路徑（見 3.14 圖上的黑點）。

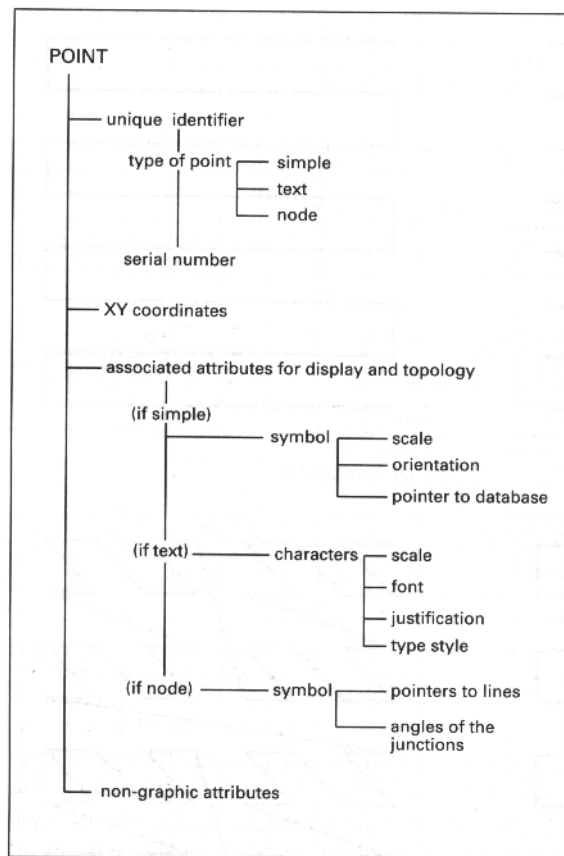


圖 3.13 (原文 p.58)

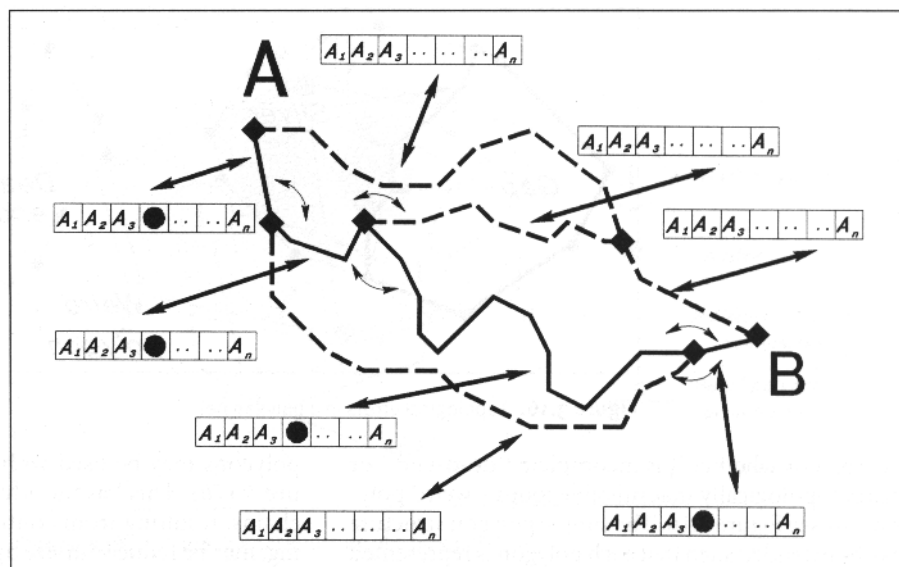


圖 3.14 (原文 p.59)

### 多邊形(Polygons)

多邊形可以被不同的向量數據形式所表示，也就是許多種空間數據形式被連結為多邊形，這方法值得注意的地方在於其可能被如何表現與利用，以下的討論著重於 Peuker 與 Chrisman (1975), Cook (1978), Weber (1978)的工作，並包括幾個已知或常被使用的建構多邊形數據方法。

多邊形數據結構的目的是描寫此面積的位相特性（它們的形狀、鄰近者、階層），如基礎空間建構單位的特性，也許就如同一個主題圖數據般被展現及利用，在利用多邊形數據結

構方法表示前，必需先界定多邊形網路地理數據的利用的需求。

首先，所有組成圖面的多邊形將具有獨特的形狀、週長與面積，這裡沒有如同網格系統般標準的基礎組成單元，甚至在最秩序或依秩序安排的美國街道計畫中，網格系統將不明智的把所有或一部分的區塊，假定為形狀或大小相同，但對土壤與地理圖而言，是很難清楚的劃分空間的不同。第二，在地理分析中，數據形式的分析需要紀錄每一個多邊形的鄰接關係，就如同流域的流通性一般。第三，在主題圖上的多邊形並非全部皆在同一個階層島嶼 (level-islands)，如發生在一個大型島嶼上的湖泊等。

### 簡單多邊形(Simple polygons)

這有一個簡單的方法去表現一個多邊形是簡單鏈結的延伸，例如每一個多邊形皆是由 XY 座標系設定邊界 (圖 3.15)，名字及符號的使用是告訴使用者每一個多邊形是由簡單的主要實體所設定。這個方法優點在於單純化，單其卻具有很多缺點，包括 a.在多邊形間的線段必需數化且儲存兩次，這將導致沿著一般邊界上不重合的裂縫或間隙出現嚴重的誤差；b.缺乏鄰接區域的資訊；c.除非是單純的地理結構否則島嶼不能存在；d.當發生位相上邊界的錯誤或不確定點時，沒有容易的方法去核對其中錯誤 (見圖 3.16)。簡單多邊形結構可能會藉由其鏈結數目擴大，但其必免不了一些基礎問題。

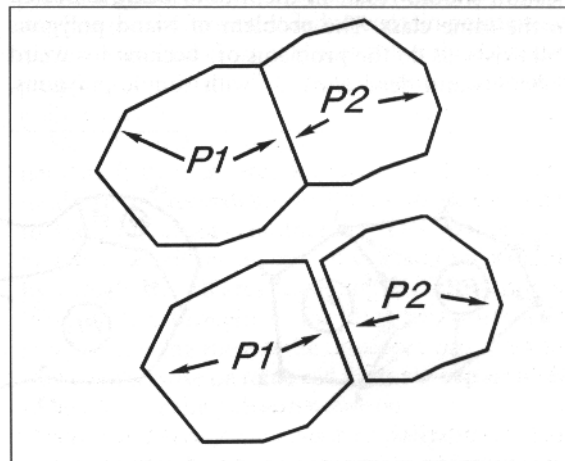


圖 3.15 (原文 p.59)

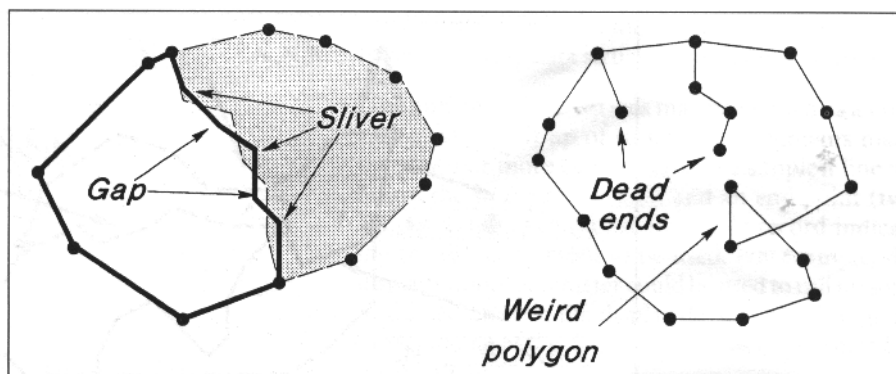


圖 3.16 (原文 p.60)

### 點目錄多邊形(Polygons with point dictionaries)

這個表示法中，所有連續的數對被有秩序的編碼，且每一個多邊形被以目錄的模式加以聯合紀錄 (圖 3.17a)，點目錄多邊形的數據形式優點在於鄰接多邊形的邊界是獨立的，但功能仍然存在，也因此如果這個結構在進行重新編碼或分類時，不能輕易允許介於兩個鄰接

多邊形的邊界被禁止或取消，但問題仍然存在於島嶼多邊形及不容許的環形相位 weird polygons 或不完全連結 dead-end 中。有一些簡單多邊形是被鏈目錄所使用（圖 3.17b），在沒有對目錄進行修正時，這有助於利用運算簡化其大小（見第四章）。

多邊形的性質利用指標對數據表加以連結（圖 3.17c）。

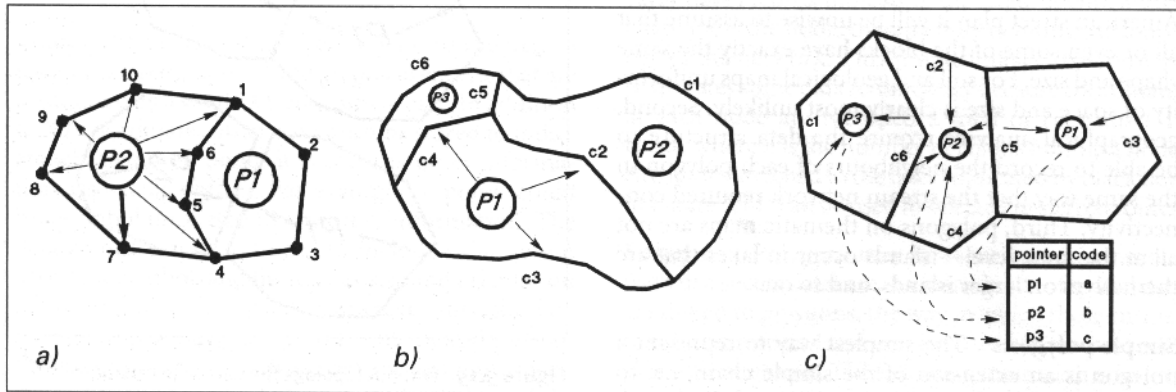


圖 3.17 (原文 p.60)

### 外顯多邊形的相向結構(Polygon systems with explicit topological structures)

孤島及鄰近者只能在複合的外顯位向資料結構中被正確掌握，這位向結構可以被一或二種方式建立，一是在數據輸入時進行位向連結，或是利用軟體從一組連環建立位向，在第一個案例中，其負擔在於操作者所插入的位向，後者則關係到電腦能力的強度，為了要描述完整的多邊形結構，這兩種方式都會使得被儲存的數據增加。

第一個在地理數據結構上建立清楚位向關係的是美國普查局 Dual Independent Map Encoding 的 DIME 系統，DIME 資料檔基礎的組成元素是被定義的簡單線段，其合成的線段是由一系列的線段所展現，多邊形任一邊的線段其節點與編碼有兩種指標，當節點不能點出線段或線段不能點出相鄰線段時，對多邊形的概述會相當困難，此外簡單的線段由於數據資料過多，對於線段壓縮相當麻煩。

### 多邊形網路結構完全地位向關係(A fully topological polygon network structure)

在圖 3.18 的結構顯示，這結構的邊界或字串的設定，可以依據任一命令或方式被數化，這系統允許孤島和湖泊在任一層級中被嵌入，對於曲折多邊形 (weird polygons) 或死巷 (dead-end) 進行自動的檢查，並對於結果的多邊形的非空間屬性進行自動或半自動的連結，也支援鄰近物的尋找，儘管具有不同的詳細資料，在哈佛大學中這個系統對於描述使用過程仍相當簡易 (Peucker and Chrisman 1975)。

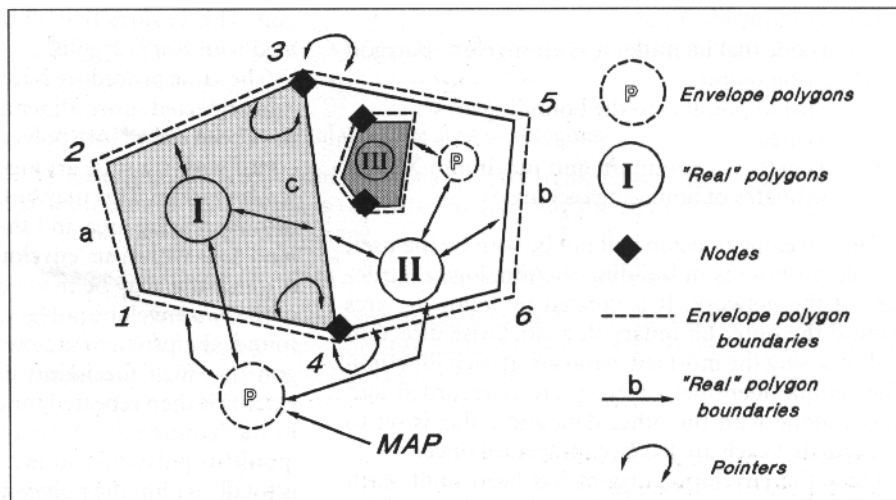


圖 3.18 (原文 p.61)

數化多邊形邊界（見第 4 章）與建立多邊形位向最好被分別處理，位向多邊形邊界的建立只需要兩個輸入數據的假設，即為多邊形邊界必須以鏈形方式進行編碼，且在一些多邊形的邊界，其名字或被儲存的圖形連結屬性資料會被數化為可辨認的點實體資料。

**第一階段**，連結弧為網路邊界。這些鏈首先必須根據其限度（最小和最大的座標）進行儲存，所以鏈的位向會趨近於另一個，也會聚集於一個數據檔中，可減少搜尋鄰接鏈的時間，並進行相交程度的檢驗，接合點由所有鏈的結束點形成，其數據紀錄擴大包括指標及角度。穿過其他空間的鏈將形成新的鏈並產生指標。

**第二階段**，檢驗多邊形的閉合度。結果的網路需利用掃描檢視封閉度，觀察其是否具有指標以建構另一個，其餘的則被視做單一獨立例如案例中的孤島，所有的鏈若不能傳遞，其原因可能是它們被特殊的方法展現，或由多邊形網路的鏈結集中被移動。

**第三階段**，連結線為多邊形。這第一個連結線成為多邊形的步驟，是為了在圖面的外在邊界建立一個新的封套多邊形（envelope polygon）（圖 3.18），這個封套實體必需包括以下的紀錄：

- (a) a unique identifier 單一確認者
- (b) a code that identifies it as an envelope polygon 確認外框多邊形編碼
- (c) a ring pointer 一個環形指標
- (d) a list of pointers to the bounding arcs 一張指標對於邊緣列的清單表
- (e) its area 它的面積
- (f) its extents (minimum and maximum XY coordinates of bounding rectangle) . 它的範圍（最小及最大XY座標的圍塑矩形）

使用者將不會發現封套多邊形的目的是建立一個網路的位向結構，其利用沿著外部空間順時鐘方向的鏈選擇左邊之交點進行建構，每一個獨一無二的標示將會隨著其他數據被記錄和儲存。

從前封套多邊形（envelope polygon）可以被建立，現在每一個單獨的多邊形也可以被建立，這項工作的完成起因於與先前的位置相同，但這時的研究必需著重於順時中方向焦點的選擇，在穿過的計數上需加以注意，一但它被穿越了兩次，它將落於研究的外圍，回到起始點，它被定義為可組成所有線段，在這同時，亦必須核對這些累計的多邊形（圖 3.18），如果它不是 360° 那就是有了數化上的失語，或出現不連續的多邊形（不連續多邊形的出現，

必須回溯步驟 1，但如果這個鍊連接的是主要節點，則核對的工作便是必須的），就如同封套多邊形，每一個多邊形實體皆有設定的資訊：

- (a) unique identifier 一個唯一的確認者
- (b) an ordinary polygon code 普通的多邊形編碼
- (c) a ring pointer from the envelope polygon. At the same time the identifier of this polygon is written in the ring pointer of the envelope polygon 來自外框多邊形的環形指標
- (d) a list of all bounding arcs. At the same time, the polygon's unique identifier is written into the record of the line 所有邊界鏈的列表
- (e) a ring pointer to the adjacent polygon in the network 網路內鄰近多邊形的環形指標
- (f) minimum and maximum XY coordinates ( extents ) of the bounding rectangle. 長方形最小與最大的XY座標系範圍。

在同一個網路層級中，尋找下一個多邊形的過程，就如同建立所有個別的多邊形，當最後一個多邊形嵌入指標（e），這個指標將回歸為封套多邊形。這有助於確立所有的邊界線皆連結二個多邊形。

這同樣的過程出現在所有的島嶼和不連接的子大陸。鍊可以連結多邊形，但島嶼與子大陸在位向階級上必須被正確的排列，首先，儲存它們並觀察是否島嶼會落入一個封閉的大尺度中。壓縮兩個封閉多邊形可以很快的進行測試。如果出現相對的問題，便是有孤島的出現，如果出現了一次重合，在多邊形內的點的運作，可用來觀察島嶼是否在多邊形內，如果出現了重合，便表示一個單獨的數據或分析過程有了問題。如果沒有重合，指示物將從網路的多邊形中被寫到封閉的多邊形中，如果沒有重合與平衡的話，代表著兩個多邊形網路是互依的。環狀指示物結構的封套多邊形—網路多邊形—島狀封套多邊形—島狀網路多邊形允許一個無限大的範圍，常常，範圍只需要合計一次，然後網路可由以下的指示物輕易橫跨。

**第四階段**，計算多邊形面積：下一個步驟包括利用梯形定律計算各別多邊形的面積（如 Box 3.3），先計算全部再扣除表面的孤島，然後當作屬性存起來，因此在電腦區域中通常較有效率。

Box 3.3 (原文 p.63)

**第五階段**，結合非地理屬性的多邊形：建立數據資料庫的最後一個步驟是連結多邊形以結合描述呈現的屬性。這可能由好幾個方法進行之，首先，計算一個獨特測試每一個多邊形的實體，包括資料實體的部分，或多邊形形式化後的內部活動。這測試指示物對結合可能或不可能在圖籍資料中存取的屬性。這個測試對視覺顯示器可能是有用的，利用點與面的搜尋去聯結多邊形（如 Box 3.4）。其次指示電腦為每個位在多邊形中心的每一個面下一個獨特的定義，同一時間印表機列印所有的多邊形定義。這表列可被包含其他經由對照獨特多邊形定義的多邊形非地理屬性的檔案。

Box 3.4 (原文 p.64)

藉由合成的軟體建構相位上的多邊形進行描述，圖 3.19 可見其數據結構有以下的優點：

- (a) 多邊形網路完整的結合間隙、裂縫和一需多餘的座標
- (b) 所有的多邊形、鏈和聯合的屬性皆為相互關聯的部分，所以可以進行所有鄰近物的

- 分析，在紀錄這個系統時需要盡量的描述這些鏈所具有的非空間屬性。
- (c)大陸島嶼的鑲嵌數目位受限制。
  - (d)地區資料庫的精確度只受限於數字轉化與電腦文字的長度。

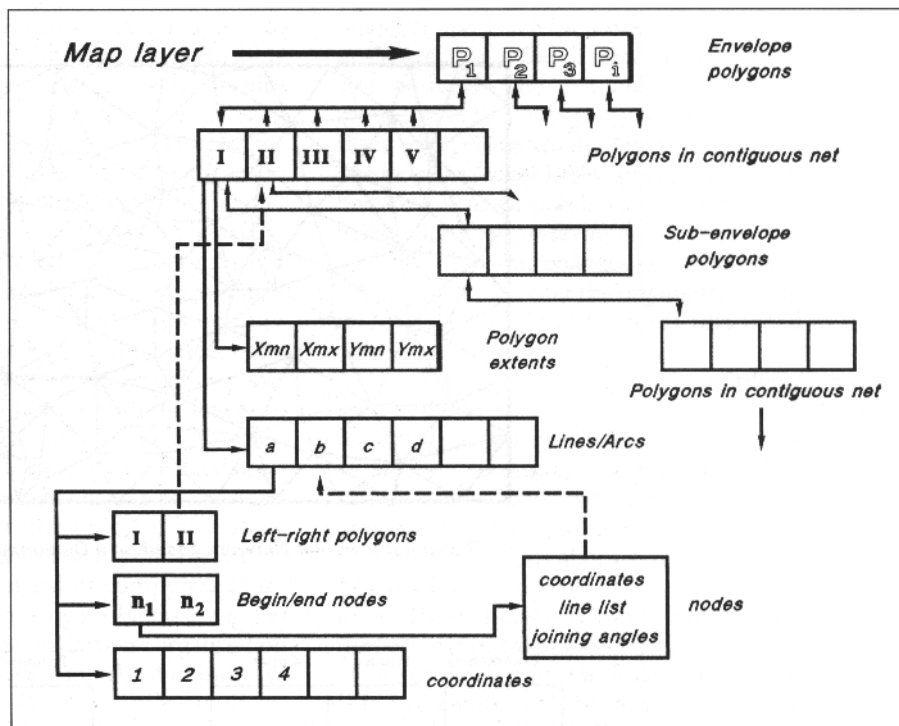


圖 3.19 (原文 p.65)

### 編輯與更新多邊形網(Editing and updating the polygon net)

向量式多邊形網路可能是由個別的点與節點組成移動式座標系統、多邊形屬性的改變、刪除或增加線，甚至是整個面的部份所編輯而成。改變座標系統或關聯屬性是簡單且要求沒有變形的位向關係。藉由刪除或增加線與面的變形，網路需要當地重新估算的位向關係與資料庫的重建。因而這些各式的資料結構不因經常改變的空間格局而有效益。

### 特殊用途的向量式資料結構—不規則三角形網路(Special purpose vector data structures—the Triangular Irregular Network (TIN))

一個重要且使用最多的向量式多邊形結構就是 TIN。它是由已知的點，基於 Delauney 三角形形成一連串的三角形所建立的。如圖 3.20a，在三角形中允許不定的密度與點的干擾被使用在一個面中影響屬性值的變化。結構模式關係諸如主要單位的網路節點。位向關係的關聯是由每一個節點與其鄰近的節點所建構的指示物，而建立成的數據資料庫，鄰近的邊界以順時針方向分類環繞在北邊開始的每個節點。由 TIN 模式化的區域，由名義上的節點在由 TIN 所設的位向關係球體狀的顛倒面所呈現描述的。這名義上的節點促進了鑲邊、點的位向關係描述與簡化了其過程。

圖 3.20b 表示網路的一部分，資料結構（三個點與二個三角形）被用來界定 TIN。這數據資料庫包括三組記錄稱之為節點表列、指示物表列、以及由三角形所組成的表列。在指示物表列中，這節點表列包括記錄、定義每一個節點及含有其座標系統、鄰近節點同樣的起始位置，區域邊緣上的節點有名義上的指示物，準備去說明與外界的接鑲。



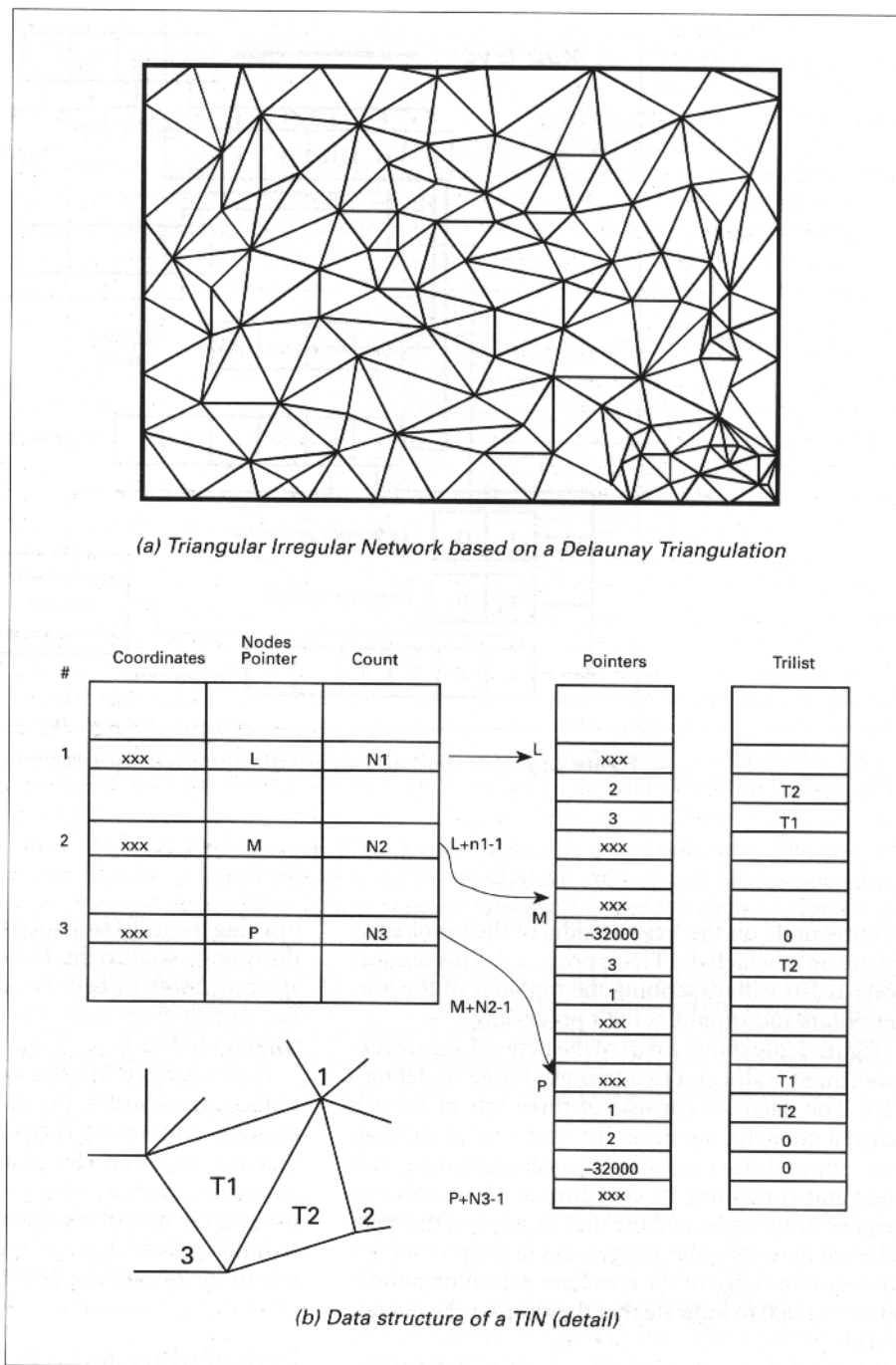


圖 3.20 (原文 p.66)

這節點表列與指示物表列包含所有基本的海拔資訊與關聯，因此，它們能滿足許多的應用。其他的應用能直接參考三角形是必要的，如坡度圖、丘陵微細變化、或與三角形建立其屬性。這是由使用三角形表列去建構每一個三角形右面的直邊而成的。在圖 3.20b 中，三角形 T2 是由在指示物列表中三個直邊所建構的，即節點 1 到 2、節點 2 到 3、以及節點 3 到 1。

節點座落在最大變化的區域幫助減少在衍生形式中的錯誤。在 DEMs 的例子中，山脊與山谷的值是幫助確定地面海拔的結果沒有異樣，如河流上游的流動。使用這些不規則立體的點，TINs 避免重複規則的格子與為電腦獲得資料，如坡度，提供有效的平均值。

## 在向量式資料結構中的發展以提升資料存取時間與有效的儲存庫 (Developments in vector data structures to improve data access times and efficient storage)

向量式資料模式是關係著儲存只有相關座標系統有記錄的地理資料在幾何學資訊中有效的平均值，主要的問題是資料存取的建構，特別是位向關係與屬性資訊。舉例來說，基本向量的點、線或面之間的相互關係是藉由指示物或說明特殊的認定，位向關係的結構只有由這些特殊說明的方法可連貫。在早期 GIS 這些說明某些時候稱為主要指示物索引，通常容納一個存取資料庫依賴關鍵的連續表列。這是表有兩個主要的問題：第一，難得有完全符合邏輯次序的，在編輯中，孔隙的出現或長度的增加都意謂著搜尋幾何學不變的是連續的。其次，表格長度準確的搜尋時間增加意謂著地圖處理時間增長非線性和資料庫的大小。

在 1980 年，許多 GIS 資料庫技術的發展集中在空間資料較少依賴資料庫的大小處理的製作進行的問題。今日電腦的速度代表了其儲存量與存取問題不似以前那麼顯而易見。資料庫的組織與內部參考的改變也同樣幫助改善效率。

### 成群資料的儲存媒介(Data clustering on storage media)

第一個嘗試改進資料庫存取時間使用”brute-force”計算理論以快速地審視點陣列，或集中在小型、鄰近磁碟或核心儲存量上區域的主要索引陣列。這無疑帶來某些改善，亦是最佳治標物與可解決下面的問題。其他處理複雜成群主要指數點不只是根據實體類型，也根據空間位置，在磁碟資料中根據地理位置成群。

### 使用 B-樹狀與 R-樹狀的資料庫索引(Database indexing using B-trees and R-trees)

加速懷疑指示資料庫的重要性在本章節先前已進行討論。然而，在 GIS 應用方面，大型或複雜的資料庫的索引可能自行變成長期與費事的疑問。結構因此發展索引中的索引層級，使搜尋更有效益與管理性。他們知道如多層次的索引以及實際使用向量式資料結構，許多位向關係與屬性資料容納在索引檔。

B-樹狀結構提供結構用“內部節點”與“leaf-節點”的多層次索引，其機能類似個別的樹枝條。圖 3.21 的例子顯示出街道檔案命名的資料是以字母順序排列的索引。第一個索引將資料分成字母群 a 至 f、g 至-m、n 至 s、t 至 z，列出各個節點資料的指示物，索引的下一個層級將字母群分成 finer (鱗) 部份。當搜尋啟動時，也限制每一層的儲存時間。B-樹狀結構解決了資料庫的力學改變經由插入或刪除記錄之後變更組織的算法規則系統 (Worboys 在 1995 更詳細地討論)。

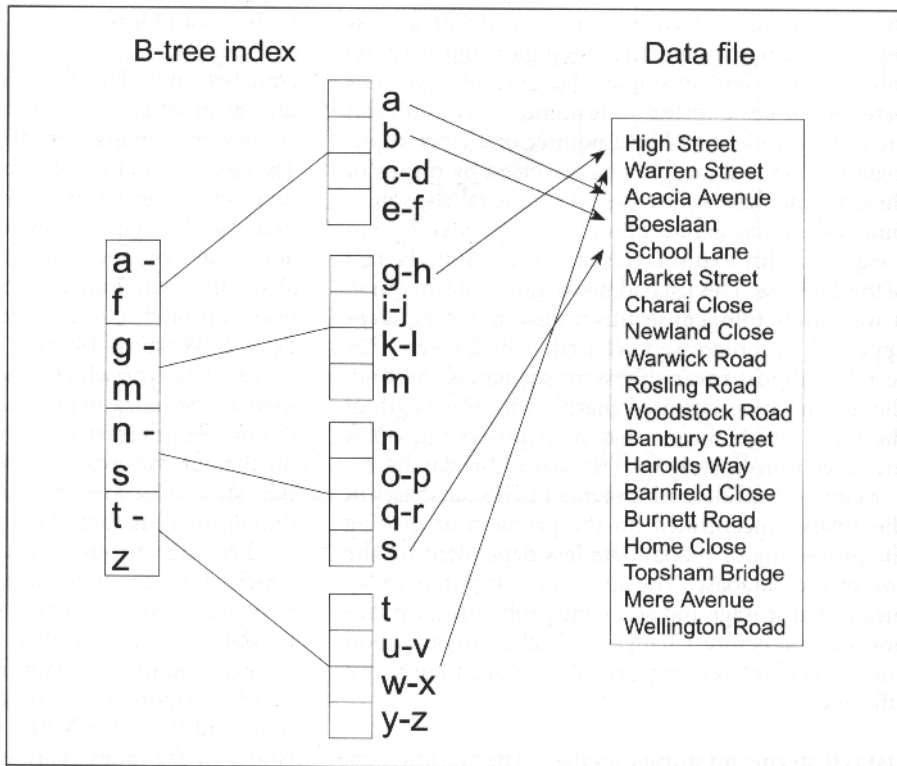


圖 3.21 (原文 p.68)

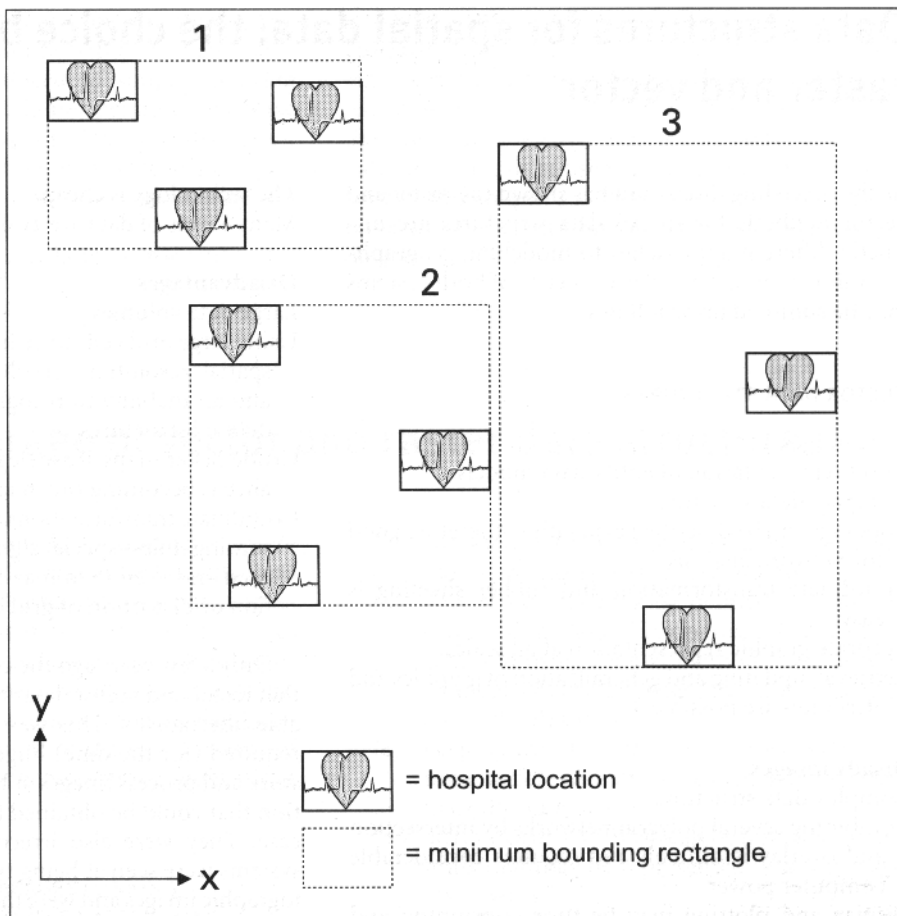


圖 3.22 (原文 p.68)

在 B-樹狀任何有直線次序的資料型式，可能如索引檔般地使用。在 GIS 中，索引檔的

數量或測試一連串的價值對搜尋屬性的記錄都是有用的 (van Oosterom 1993)。然而，此結構並沒有注意到 GIS 懷疑的許多空間 (2D) 本質。

有多種允許資料庫幾何的所有物的 B-樹狀模式方案，包括所引的結構。舉例來說，R-樹狀 (Guttman 1984) 將空間分成一連串的含量 (boxes)，如一般所知的最小周長矩形 (MBRs)。如圖 3.22 所示，一連串醫院的位置與三個 MBRs 用來分解空間。索引中的節點呈現這些矩形與搜尋實際醫院的位置，將會由當中之之一的算法第一個指導。MBRs 不同矩形大小的層級結構為索引，可能使用樹形結構建立之。搜尋算法檢查含有實體的大矩形，然後接著樹狀下行各種層級直到遭遇到疑問的狀況。

#### 四分樹(Quadtrees)

四分樹的基本概念本章節之前已有說明，其用來描述減少磁碟空間需要呈現網格資料結構。這個資料結構技術也可能用在向量式基本資料、屬性存在或不存在制定了一個如同限定義的編碼 (Laurini & Thompson 1992)。在某些複雜的四分樹系統中，網格也可被參考了幾何學單位呈現如邊界或頂點或點的數值。

四分樹結構進行有幾何變化的空間資料的編碼。在許多層級中，空間的分割已經如不規則的方格一樣地被描述。然而，隨著某些結構資料本身制定了再細分部份的位置與形式。點、線與向量的分佈或位置，在每一個細分部份將空間分成不規則的形式，以幫助描述在各種大小資料值的密度與分佈的變化 (Laurini & Thompson 1992)。不規則四分樹的結構形式，是高度依賴各種點嵌入時間的部份 (Worboys)。在向量式資料使用四分樹的主要問題是缺乏明顯位向關係的參考，以及點與線值的準確性被限制。

#### 連續向量範圍：覆瓦(Continuous vector coverages：tiling)

真實世界是連續的，而且不會停止在地圖與電腦檔案的邊界。一個無縫的資料庫是由根據系統所知的如覆瓦為鄰近區域的連結的檔案所模擬的連續性。理論上，世界是由無限延伸的方向所陳述的。每一個覆瓦可能參考確實的資訊數值，尤其是細項 (整體中的一小部份) 可由每一個區分主要的覆瓦，用在四分樹結構中相似的種類，調解成亞覆瓦。覆瓦輸入一個特別的複雜度，在其中所有弧在覆瓦邊界與位向關係指示物會自動終止，而不只參考其他實體。在較大的資料庫耗費大，但一個在搜尋時間可能因有關位置的層級的效能而保持非常低的成本。

在理論上，覆瓦允許限定性的圖籍去創造與儲存如從少數覆瓦需要參考任何時間靜止儲存在磁碟的唯一資料。實際上，資料的絕對容積將超出財政上可允許的磁碟空間，甚至允許在硬體經營花費的減少，因此，對廣闊的國家版圖來說，好的依賴必須放置在網路資料庫或儲藏庫如 CD-ROMS、光學磁碟或磁片，以儲存今日大量的數據資料庫。

#### 向量式資料結構的新發展 New developments in vector data structures

常見 GIS 結構眾多問題之一是遇到了細項層次或儲存與展示安裝與制定在輸入層的資料分析。當從事不同地圖的尺度時，資料的細項保持相同，且使用者的觀點被製圖展示的分析與急速上升的程度所管制。學者之一 Van Oosterom (1993) 發展在細項層次與使用者互相作用的幾何描述，反映資料結構的想法是由展示的尺度所制定的。急速上升的大尺度地圖比小尺度陳述更多的細項。線性的實體展示如前者的多邊形與後者的線形。

#### **空間資料的資料結構：在網格式與向量式的選擇(Data structures for**

### spatial data : the Choice between raster and vector)

根據之前的討論，已經顯示空間資料結構的網格式與向量式理論應用在模式化的地理資訊系統是截然不同的。兩個系統的相關功能可做出以下的結論：

	向量式資料結構	網格式資料結構
優點	1) 實體資料模式描述佳 2) 壓縮資料結構 3) 位相關係可以明確地描述—所以有好的網路分析 4) 座標系的轉換與 rubber sheeting 是容易的 5) 所有尺度的準確繪圖 6) 繪圖與屬性可恢復、更新與普遍化	1) 資料結構簡單 2) 屬性資料具體位置操作容易 3) 使用各種空間分析及過濾 4) 數學模式簡單因所有空間實體具有簡單規則的型態（輪廓） 5) 技術便宜 6) 可獲得許多資料的形式
缺點	1) 複雜的資料結構 2) 由交叉點與疊圖結合多邊網路困難且要求考慮電腦動力 3) 展示與繪製平面圖可能耗時且昂貴,尤其是高品質的繪圖、色彩與描影 4) 基本單位（如多邊形）的空間分析沒有額外的資料，因考慮其內部的相似性 5) 因每個空間實體都有不同的輪廓與形式，所以由明確的拓樸學界定空間相互作用的過程的模擬模式是困難的	1) 資料體積大 2) 加大網格式、縮減資料體積、減少空間分析，導致失去資訊及無法定義現象學的結構 3) 精緻繪圖使粗糙的天然網格式問題較少 4) 除非使用特別的算法（規則系統）與硬體，則座標系統轉換不易又耗時，甚至導致失去資訊與圖形失真

幾年前，網格式與向量式資料結構是不相容的方案，這是普遍所知的事情。這是因為網格式的理論在空間分析的層級上，要求極大的電腦記憶體來儲存與處理影像，而在向量式結構中可被輕易建立。同樣不相容的是，因在常見的製圖影像、基本的高品質繪圖與位相關係的準確性上，向量式系統似乎是正確的。早先技術發展嘗試將向量式簡單化，因對製圖者來說，這向量式的結構是較熟悉的。而網格式系統只有在疊圖分析是恰當的，而較不要求繪圖的精緻。

在之前，網格式對向量式的討論沒有長久的相關，也沒有顯示互相的特別性。很清楚的第一就是重要概念問題，也就是實際上的技術。製圖、硬體拷貝描述、及網格式系統資料儲存量的問題，在高解析的電腦螢幕與印表機有很大的影響，以及資料儲存與壓縮技術相關易得的媒介。在 1970 年末期，好幾個工作者，包括 Peuquet (1977、1979)、Nagy & Wagle (1979) 都表示，在向量式結構發展的多邊形資料的許多演譯不只有網格式方案才有，但在某些案例中是較有效的。

舉例來說，多邊形的周長、區域、總和、平均值、及其他點、半徑操作的計算減少，在網格式裡簡易的計算操作。某些過程的操作仍然保持更多組成資料結構之一，網路分析比位相關係資料結構來的簡單，有時候在網格式系統中地圖的交叉點與疊圖是不重要的。

今天，許多 GIS 提供向量式與網格式結構且也提供轉化問題。然而，這些通常要求資料在相關結構的分析形式橫過圖層。近幾年來在數據資料庫組織的方案、物件導向理論、允許向量式與網格式資料結構在同一時間被使用（稍後會討論）如它們處理不同空間單位、點、線、面或像元為獨特物件一般。

## 向量式與網格式資料結構的數據資料庫儲存量(Database storage of vector and raster data structures)

本章先前已描述多種的數據資料庫結構。其提出存取網格式與向量式結構實際的優缺點。這說明了,使用在大部份 GIS 系統的資料庫結構有助於有力的商用的關聯資料庫管理系統 (RDBMS)。更多近來物件導向數據資料庫已經使用在商用的 GIS 新形式數量中,如他們所提供應用在區域總數的優點。

## 混合關係資料庫：連接幾何學描述屬性(Hybrid relational databases : linking geometric representation to attributes)

商用的 RDBMS、INFO、ORACLE、INGRES、INFORMIX 及相似產品容易用於 GIS 系統設計者,如他們可以為其處理需要用已經發展與測試的系統。這些數據資料庫允許設計者將空間資料管理問題分成二個部份。第一個部份是如何描述空間物件的幾何學與位相關係—應使用向量式資料結構或網格式資料結構?第二個部份是如何處理可能使用商用的 RDBMS 的空間物件的屬性。

這結果混合結構(某些時候涉及到地表關係模式)有不同的優點:

- a) 屬性資料不需要隨著空間資料庫而儲存,但可能保持在任何地方的系統或甚至經由網路的線段上。
- b) 屬性資料可被擴充、增加、刪除、更新而不需去修改空間資料庫
- c) 商用的 RDBMS 確保新發展可以如標準(規格)般具體化
- d) 資料結構可能定義在資料字典中的標準方式:資料可恢復使用一般的理論,如依賴 RDBMS 的 SQL(標準質問語言)
- e) 保持在 RDBMS 中的屬性資料不會妨礙在 GIS 圖層的基本原則
- f) 在 RDBMS 的屬性可連接至以多變化方式描述的空間單位。

使用商用的 RDBMS, GIS 設計者創造混合結構的變化,包括如下:

### 1. 弧-節點-關聯資料庫管理系統(ARC-NODE-RDBMS)

這可能是大部份使用充滿向量式弧-節點位相關係的資料庫系統被用來描述如上解釋的線與面邊界的網路。每一個空間單位用獨特的數碼或編碼所定義,且可能放置在一個精選的圖層或疊圖。空間單位的屬性如記錄在可能由 RDBMS 處理的相關表格所儲存。某些位相關係的資訊與屬性如座標系統、邊界、區域、最小周長矩形的座標系及簡單的資料也可能儲存在 RDBMS 的表格中。圖 3-23a 表示這類資料結構的典型例子。

當圖層是相交的時,此類系統的複雜性是新空間實體與其組成屬性的產品,因為這意謂著在 RDBMS 中建立記錄與連接的新組合。

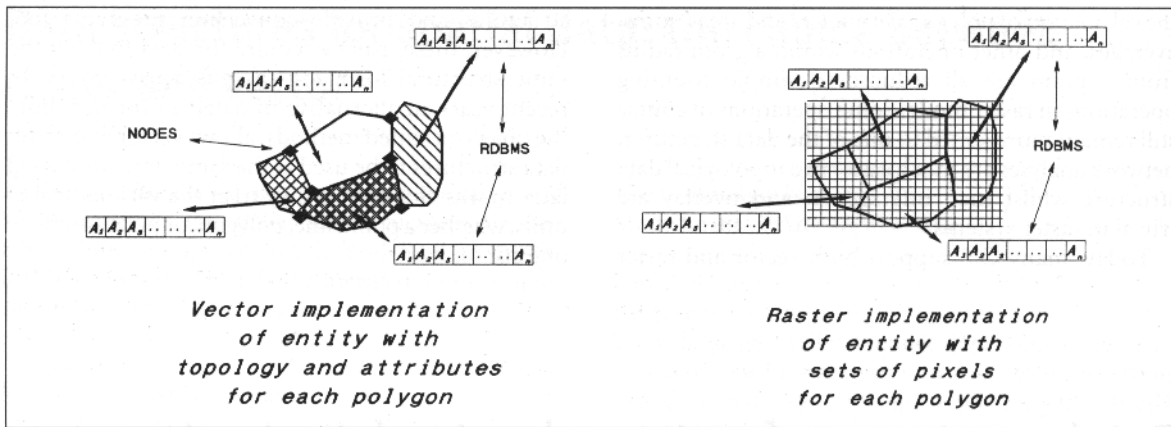


圖 3.23 (原文 p.72)

### 2.由網格組成-關聯資料庫管理系統(COMPACT RASTER-RDBMS)

如果空間物件是由幾組像元代替位相關係連接線段所描述的，則網格等於如上述是可被創造的。當大部份的空間資料涉及內部均質的主題單位，如地籍圖單位 (choropleth map units)，則 run-length 編碼的網格式壓縮理論可用來儲存空間。圖 3.23b 就是講一個簡單網格混合應用均質面的例子。

### 3.四分樹-關聯資料庫管理系統(QUADTREE-RDBMS)

四分樹可能也用在資料壓縮，但因為允許資料被描述在空間集合體的各種層級，它們允許不同空間分析的層級以使用在不同的圖層。當交叉資料在向量式的狹窄的河流山谷與寬闊平原的單位，四分樹是極有用處的，但在允許只有單一網格大小或分析的層級的網格式系統是困難的。

### 4.物件-關聯資料庫管理系統(OBJECT-RDBMS)

在近幾年來，物件導向處理 (早先討論過) 已經採用在同 GIS 中向量式與網格式資料結構的組織。在這些系統中，多種幾何學與屬性資料儲存在相關表格 (Gahegan & Roberts 1988)，且物件導向編制語言、提供像製圖資料基本物件的分界面一樣的分析功能。這系統允許物件導向的處理或地理資料的組織以去開發相關資料庫的環境。

## 物件導向的資料庫結構：結合屬性與幾何圖籍的儲存(Object-oriented database structures：unifying attribute and geometric storage)

物件導向資料庫要求地理資料如一連串微粒單位般地定義。這顯著地幫助資料界定使用實體概念模式。地理資料由一連串的屬性與定義其空間、輪廓、時間的數字大小的功效值為其特性、描寫 (Worboys 1994)。屬性界定的部份將描述物件的幾何的種類 (點、線、面或網格)，多於一個幾何類型可能反映出在不同空間尺度發現種類的不同點。

屬性與行為變數是自己為其所有物與理論定義物件分類。層級關係用各種分類所建立的，舉例來說，“弧”物件可能是“面”物件的亞網。位相關係連接各種物件例子與分類之間，經由物件指示物與操作者如“方向”、“交叉點”、“鄰近的”、“疊圖”、“左邊”、或“右邊”所明確建立的。

在物件導向資料庫中的空間資料組織，提供引人注意的對某些 GIS 使用者就如一個模式方法真實世界的語義與過程，比可能的關聯系統提供更多綜合的、直覺的方式 (Kidner & Jones 1994)。人們製作人造物件如公共設施部門已經發現這些系統提供適宜資料形式一個處理方式，他們使用與懷疑容量的需要。在物件分類間的層級結構與關聯複雜關係的描述可被直接地控制，因此，在資料庫的更新與改變容易被曲解。

資料庫的結構在一連串的自制、基本的單位帶來問題與可能性。很困難地將連續空間區域分類成單獨的單位。在分析與模式化中你會如何將小丘陵分散成一連串明顯的物件呢？界線的選擇通常是主觀的 (見第 11 章, Burrough & Frank 1996)。然而，在何處可見資料支持這個觀點與客觀的分別出現，資料改變的可能性增加。這些物件使用在新的應用方面甚至要求不同的結構。當可以重複使用資料程式庫包括相當大金錢與時間的投資，他們已經顯示出一個相當大的報告書 (Worboys 1995)。空間資料物件程式庫現在進入網際網路。

對資料來說，在 GIS 中物件導向資料庫的執行已經被限制。這問題是有一些一般性的物件導向資料庫可利用的產品充當支援 GIS 機能的發電機。

## 關聯式的資料庫與以物件導向的資料庫之討論(The debate on relational-hybrid GIS versus Object Orientation)

物件導向資料庫近期的發展助長重視存在關聯式系統優點的討論。主要的結果可能做以下的結論：

### 關聯式地理資訊系統(Relational-hybrid GIS)

- 輸入系統後空間資料的可變性
- 由 DBMS 所提供的資料恢復與圖解功能
- 從其他系統簡單的資料完成，尤其是屬性資料
- 所有資料方面儲存在特殊的檔案結構
- 容易使用
- 因為關聯式資料庫而成正確的理論機能

### 時間性資料的拙劣處理(Poor handling of temporal data)

- 座標系統資料並不有助於準確數據資料庫管理的主题，當可能被應用在屬性資料，因此可信度的問題與完整的存在。
- 為疑問或模式化依賴的空間位置或屬性
- 疑問的慢速處理尤其當處理複雜的物件時
- 疑問與模式化限制了由 GIS 提供的功能 (或資料必須輸出)

### 物件導向的資料庫(Object-oriented database GIS)

- 孔隙在真實世界物件與概念與資料庫的描述是少於關聯式資料庫
- 狀態與理論的儲存庫確信數據資料庫維修降至最低
- 網格式與向量式資料結構可能完全在相同的資料庫
- 資料物件的交流是被支持的
- 快速的資料庫的疑問是必要的，尤其當複雜的物件與關係必須處理時 (如少數接合操作)
- 要求比關係實體少的磁碟空間，此磁碟空間需要儲存更多索引檔
- 使用者界定的可使用功能



## 缺點(Disadvantages)

- 並沒有普遍公認的物件導向模式，因此，不同的資料庫產品有不同的標準與參照特定的 0-0 語言（程式語言）
- 確認物件通常是困難的，特別是在連續空間表面
- 要求功能的確認與物件一樣的位相關係
- 限制索引的應用，因為壓縮與物件本身的觀念不能相容
- 系統中物件的複雜度沒有建立標準如 SQL 以及提供普遍的疑問語言或疑問，充分運用是困難的
- 程式語言的應用比關聯式的理論有較少的理論與實際的經驗
- （after Arctur & Woodsford 1996, Graham 1994, Herring 1992, Milne et al. 1993, Worboys 1994, Worboys 1995）

近來 GIS 發展已經計畫給使用者更多適應性，以定義空間單位與提供更多分析的技巧。應該允許集中更多在空間中多變的關係不如在電腦中呈現與被資料庫結構所限制。內部操作的新發展與開放 GIS（第 12 章討論）將有幫助。

## 問題(Question)

1. 討論電腦在空間資料結構編碼的限制。
2. 為什麼資料庫管理系統如此重要？什麼是其主要功能？
3. 思考日復一日生活中無意發現的主要數字地理資料。何種資料模式被用來呈現資訊，為什麼？
4. 為什麼使用者瞭解資料庫結構何時使用 GIS 是重要的？
5. 重新檢視不同理論使用在加速資料的存取與壓縮。思考不同 GIS 應用的範圍且這些技術是重要的。
6. 為一個或更多的處理方式設計一個以物件為導向與關聯式的數據資料庫：
  - 油管透露的環境議題
  - 史前定居點的考古學調查
  - 警察部門
  - 公共設施部門（水、瓦斯、電力、電信）
  - 旅遊查詢
  - 土壤監測組織

哪一種資料你會使用？你如何在資料庫當中定義它們？思考每一種處理方式的限制與益處。

## 第四章 資料輸入、驗證、儲存及輸出

### *Data input, Verification, Storage, and Output*

#### 前言

建設精確的空間實體的地理資訊系統資料庫是一個艱難任務。地理學原始資料，諸如許多不同類比或者數位的形式，例如地圖、空中照片、衛星圖像或者表單。建構數位地理學資料庫有三種不同方法：(一)從資料供給者以數位形式獲得資料，(二)使現(存)類比資料數字化，和(三)人們自己的數位調查。通常必須有系統的紀錄這些資料並且這些資料必須是共同型式。不管是類比或數位形式的資料，這些資料需要轉變成地理資訊系統可使用的內部資料庫架構。這些現存的數位資料傳輸它經常涉及在供給者和當事人的系統之間使用標準的轉換方式。原始資料通常以類比形式顯示，而這些資料之數位化必須要經由一些儀器轉換，例如一個數化儀、掃描器或者立體繪圖機(stereoplotter)的使用裝置。當這些資料轉換後，我們必須檢查這些資料是否有錯失位置和誤差。大部分地理資訊系統提供這個工作相關之編輯工具。必須把屬性與實體資料庫連接起來，包含相關表格或空白表格之聯結。原始資料和其他的重要特性可以儲存在單一電腦上，也可以儲存在網路上，便利使用者於在網路上分享資料。經由許多人的努力而創造 GIS 之資料庫。磁和光學的媒介提供資料儲存和資料的分散。部分或者完整的空間資料庫可以在電腦台螢幕上或精裝的報告或影片短暫的顯示。

#### 地理資料的來源(Sources of geographical data)

GIS 資料庫之建造包含資料的擷取、檢驗之複雜操作和架構的過程。因為原始地理學資料可由許多不同類比或者數位的形式例如地圖、空中照片、衛星圖像或者表格取得，空間資料庫能夠以多種互容的方法建構。這些方法分別是：

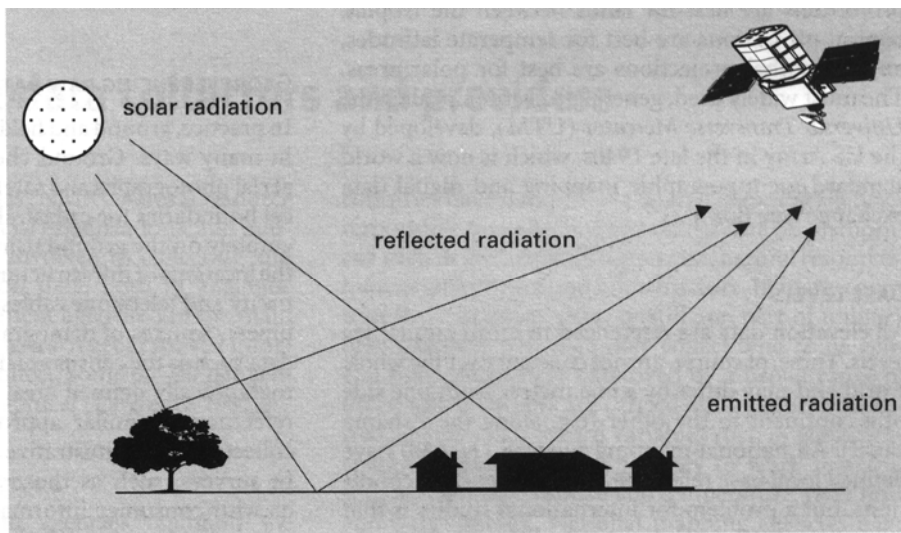
- 從資料供給者以數位形式獲得資料，
- 以現存類比資料數位化，和
- 人們自己的數位調查
- 從點的觀察內插推估至連續面的資料。

在一般情況下，通常必須以幾何型式登記這些資料和恰當地定義座標系統和編碼，他們才能夠在已使用 GIS 的內部資料庫架構中(被)存儲。所希望得到的結果應該是一個現存的完整的資料庫，能夠支援爾後資料分析和模擬。

就大部分的人來說，地理學資料的來源通常來自文章或者數位的地形學上的或者主題圖，這是空間現象分布的寫實陳述。這些地圖是用某種尺度以不同符號展示實體的屬性或者顏色。借助於一個相同的座標系統展示了地球表面上實體的位置。

從安裝在飛機及衛星的光學和數位的遙感系統得到圖像提供不同尺度時間和空間的分辨率(參看章節 5)。立體空中照片是人類解釋地質學、土壤、植物，或者土地的覆蓋物地圖的主要資料來源，並且他們也是有價值背景文件，用於把其他空間資料放在恰當地地理的文獻。數位攝影製圖法和數位的 orthophoto 地圖直接提供地形高度及地表覆蓋物資料而不需要從類比式文件轉變為數位形式文件。(插圖 1.1)

Box 4.1 ( 原文 p.76 )



各式各樣的掃描器安裝在衛星或者飛機上直接提供數位資料。這是以被動的接收器接收反射的輻射的系統，並且為記錄較大範圍的反射波長能量的種類，包括的熱的和微波的電磁光譜區域。被動系統對地球的表面成分的吸收和反射起不同層次回應，從而提供不同的空間模式資訊。積極系統包括側面帶狀雷達掃描器，雷射高度掃描器(插圖 3.5、3.6)和聲納掃描器(在安裝運用在氣墊船或者船水面下)提供表面海拔高度和物質密度的資訊。由傳感器的類型及時間和空間決定傳送資料的用處，這取決於這個支援站臺的路徑和高度。收集地理資料的傳統方法是用已知的位置調查記錄地面或者田地的樣品數值，使用設備從 questionnaires 和土壤的螺旋鑽孔機到自動化操作的化學探針(Stienstra 和 van Deen 1994)。這些一系列點位置調查的結果及屬性值通常記錄在一個表單中；這些可以使用章節 5 和 6 的方法。今天，透過數位資料記錄器使記錄數值更加容易，基本上迷你電腦與測量設備或者手的手持型元件相連了，更容易使使用者控制裝置紀錄時間間隔。那時可以直接或者透過 GIS 電腦資料庫裡的轉換型式閱讀。

Box 4.2 (原文 p.78)

### 地理參考文獻(Georeferencing)

所有 GIS 空間資料擁有共有架構的參考文獻的是最重要的。大多數 GIS，除了地區性研究外，由大地測量學的座標系統提供參考文獻的共同架構。

最通常和方便使用於 GIS 的座標系統是平面直角笛卡兒座標，通常使用北方-南方和東方-西方。因為地球為非平坦的，而是近乎球形，經度或者東西方位與格林威治子午線有關，緯度或者北方南方位置與赤道有關。因為地球不是真實球體而是在磁極兩邊較平坦，大地測量者在飛機上面將地球的真实曲面分割成幾個橢圓以紀錄國際間及地區性之資料以製成地圖。在自然世界地理學位置的調查中通常按照幅度/經度或者依國家柵格而記錄。

### GPS 之地理參考文獻(Georeferencing with GPS)

透過廣泛地使用許多類型的資料收集器的全球定位系統，現今已能很容易定義和記錄一個資料點的位置。這些設備能定義這個地理位置及高度，使精確度的程度改變，並且任何地球表面上使用以三角測量幾何學為基礎的，由 NAVSTAR GPS 衛星發射訊號。一個手持型的地面接收器一定能從這三個衛星的至少其中一個(圖 4.1)得到信號，提供他們的軌道和原子(時)鐘校正。每一個衛星到接收器距離是整個波長和階段轉變(Fix 及 Burt 1995)數字(目)

的一個功能並且這個資訊，與位置資訊結合，足以計算出接收器的位置，幅度/經度或者格線的參數和高度。在 GPS 螢幕上顯示出這些結果，也可以下載到電腦系統裡。在限制地圖覆蓋的地區(範圍)，GPS 特別是得到位置資訊的重要來源。GPS 使用上的主要限制是無法取得固定精確的地理位置，這是由於美國國防部的精確時間代碼干擾，接收之衛星數目不同和 GPS 接收器之品質。透過一個較好的定位點，例如平坦的地區或者燈塔，將其當作地區基地站，差分 GPS 研究可以改良至 1 公尺精確度(Kennedy 1996)。

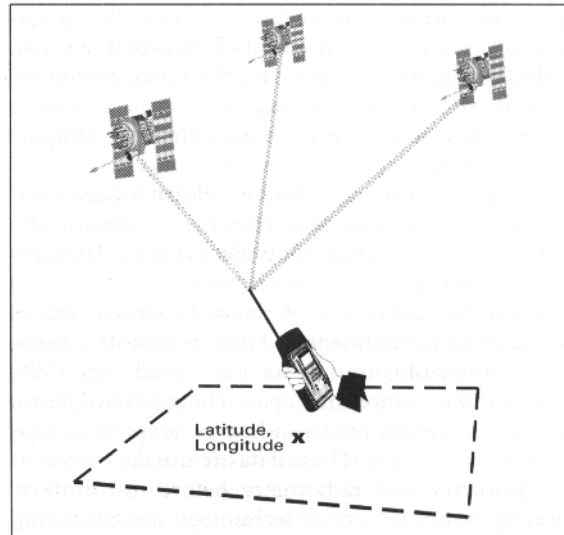


圖 4.1：使用全球定位系統可以得到世界上任何地點之經度、緯度及高度（原文 p.79）

## 地理資料之收集者及提供者 (Geographical data collectors and providers)

國家地圖機構(NMA)、自然資源調查機構、商業團體和個體的研究者均為類比和數位形式地理資料收集者和散布者。政府機構是地理資訊的主要收集者、供應者和用戶；關於他們收集的所有資料大約有 60%~80%為地理資料(Lawrence 1997)。

### 有系統的資料收集者(Systematic data collectors)

在傳統上，政府機構持續進行著地理資料的系統收集。許多國家有負責地形、自然資源、人類居地等基礎自然性質之收集系統之機構。在許多國家中這些機構仍然是軍事機構的一部分。他們所收集的資料傾向於一般使用目的和有著較廣泛使用者。土地所有權、職業和旅行、土壤和地質學、降雨量和溫度、水流和水的品質等，則由較專業機構調查記錄。國家地圖機構負責定義和保持地圖扁球體、規劃、格線起源、方向和間距、參考海平面、資料測量和收集的標準，商業地圖公司同樣的負責提供標準的地形圖覆蓋。

在許多國家，特別在北美和歐洲，地籍的調查和數位資料庫實用的登記使用是在一個有系統的基礎下進行。

有系統的遙感探測資料收集是由加拿大、美國、俄國、印度、法蘭西和日本政府和私人的遙感探測公司提。收到的資料儲存在磁媒介上，並能夠在卡帶上或者在網際網路上提供。在分發前通常經過幾何學和 radiometry 校正這些資料。傳感器技術的發展增加多-光譜的資料的數量，用以改進的光譜放射、空間和時間分辨率，並提供更多詳細的資訊。最近發射的 ERS2 和即將發射的加拿大 Radarsat 的以及法蘭西、日本和美國即將發射的地球觀察衛星將提供這些資料來源。有系統的遙感探測資料收集，主要有下列問題(一)雲的遮蓋限制了他們的效能；(二)他們的費用(不同政府和商業供給機構之間的不同)；和(三)處理系統的對特別圖

像的需求，才能夠從這些圖像得到可以認知的實體資訊。

### 特別的資料收集者(AD HOC data collectors)

由商業機構收集資料如私人調查者、工程師、市場研究者、政府機構或者學術的機構經常是較專門。他們收集一個專門目的資料，例如市場研究練習、礦場評價、環境影響估價或者綜合性項目。雖然實體及記錄屬性的種類可能覆蓋無數主題，但是這些考察可以使用由系統調查提供的基本座標系統和基本地圖資訊。研究和觀察方法的尺度及資料分類和解釋經常不是唯一的，因此限制他們的應用。經常由於商業利益而限制到資料的通路。

### 資料提供者(Data providers)

資料供應者提供或者出售許多詳細，不同版式、尺度和架構的地理學資料。傳統資料收集者和供應者在相同的機構，但是前不久，這些角色被分別歸於不同的機構中，包括商業公司，獲得自己沒收集的資料權利以轉賣。例如，在許多國家之代理機構，批發由美國陸地衛星和法的 SPOT 系統的資料。

地理學資訊可以由資料供應者的分發的受各種法律的限制，特別是在軍事機構所收集的資訊尤其嚴格。政策對資料所有權有影響力，政府和商業收集的資料其版權有嚴格的限制。更進一步的限制是由於政策限制了資料收集和分發的一些費用(Burrough and Masser 1997)。

直到前不久，大部分地理學資料以數位形式提供，並且這經常是使用 GIS 用戶的主要限制，那樣經常需要很大精力將地圖轉變成數位文件，並且必須使用現代設備重新部分調查。在最近的十年，政府和商業的組織機構感受到數位地理資料市場潛力，雖然受到價格以及版權限制，至少在一些國家中容易獲得這樣的資料。現今已經可以取得或者購買數位地理資料以應用於自然和社會經濟。在國家之間獲得這些資料的費用差異相當大，例如在美國收集資料的費用由國家基金負責而其他國家的 NMAs 是需要付費的。

由政府提供許多自然海拔高度、城鎮、道路、河流等用戶廣泛使用的數位資料，這些大多由市場負責分發。美國、歐洲和澳洲等國家發展一些不同尺度或者主題數位資料庫之綜合地圖資料。這些提供一個有用資料產品，這產品能滿足各種用戶需求。例如 Lytle 等人(1996)描述土壤有組織碳、可用的水容量、土壤深(度)和其他的性質的數位土壤資料庫(STATSGO)，其中並含有此地圖的發展。Kineman(1993)說明一個由美國的 NOAA 和環境保護局(EPA)發展之全球生態系統資料庫之五年計畫，以應用於全球變遷之研究，此資料包括植物覆蓋邊界、表面氣候學、土壤、礦脈和地形。Langas(1997)描述由多國共同努力完成之巴爾幹集水之 GIS。這些資料庫主要用於一系列資訊應用。在最近幾年許多國家討論整合各個國家、大陸和全球地理資料庫。這些執行政程序在 12 章有更詳細之討論。

與政府資助之企業之發展相應，商業機構開始了發展數位地理學資料以滿足增加的資訊需求。供給者傾向於發展適當主題的資料產品例如 郵政區號、社會經濟資料、運輸路線或者地形海拔高度之陳述。其中某些公司收集使用其他機構的現存資料，整合而成為'新'資料以增加其價值(及其智慧資產)；然後可以賣這些沒有原始資料的版權限制的一個商業產品(Lawrence 1997)。例如，在許多國家可以從商業機構買十年人口普查詳細資料，拿這些資訊結合發展功能性詢問和地圖以創造新商業產品。其他公司也已經意識到將自己收集的資料版權賣給廣大的讀者群以回收投資。

### 由資料供應者獲得數位資料(Acquiring digital datasets from a data supplier)

資料供給者提供的數位資料從小規模地區一直到寬廣的覆蓋地，從網路上免費提供至非常昂貴之資料(Burrough 和 Masser 1997)。許多本地政府和商業的用戶提供不可或缺的數位資料來源，以使這些用戶不必須應付數位化的費用以收集自己的資料。

雖然使用現存的資料是具有吸引力的，此外必須注意不同供給者的資料結合在一個項目時，資料間之兼容性。也許在設計規劃、尺度、基本尺度和屬性的描述中有差別。例如，在沙漠地形的考察中，遊牧民這個地區之分類結果很有可能不同於地質學工程師。由遊牧民解釋之地景特徵可能為水軸承的能力，遮蔽，或放牧之生產力，工程師所關心之事為此地區之經濟潛力或者地理危害之物。因此用戶當使用現存資料時需要確保它的涵義與他們的一致。這總是不容易弄清楚，許多資料並不包括任何支援資訊，一般所知之 meta 資料，包含使用方法及收集資訊方法。

用戶必須注意下列資料特性以確保他們應用時之兼容性：資料的流通、記錄的長度、資料的尺度、使用的 georeferencing 系統、收集資料之技術及抽樣資料使用的策略、資料的品質、資料的分類和修改的方法、個別地圖單位的尺寸和形狀。

從一些來源使用資料，特別是行政邊界的交界區域研究，有著不同地理參考系統、資料分類和個體調查的抽樣方法，因此造成此地區較難以整合。用戶需要意識到國際間的資料傳輸時特別有這些問題(Mounsey 1991, Lytle *et al.* 1996, Burrough and Masser 1997)。

一旦這些相容性問題擴展到由來源到 GIS 資料的轉換。使用數位式錄音帶 卡帶分散數位資料是相當好的使用媒介，其他如 CD ROMs、軟碟和網際網路變得越來越重要。必須將來源資料編碼及轉換其資料架構以能使用於 GIS。從不同供給者供應數位地理學資料處理系統的地理資料轉變成另一個資料是困難的，並且有著各種不同形式之轉換資料。多數 GIS 供應商用自己的辦法處理許多商業資料庫管理系統來解決資料的幾何學和拓撲學的方面的屬性資料(如同在第 3 章討論)，過去經常提供最簡單的資料轉換的界面。這些辦法意味著將所有空間實體資料轉換到另外的系統，必須對地誌和資料的架構重新組合以使新 GIS 系統可以讀入此資料(如同在插圖 4.2 顯示那樣)。

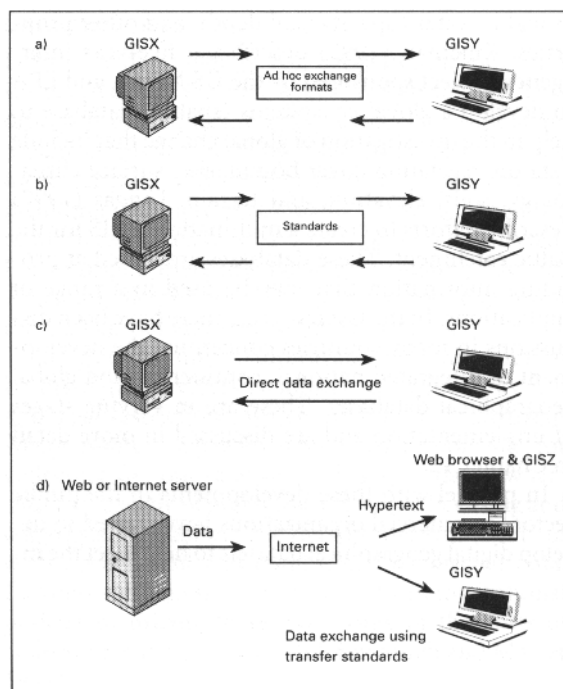


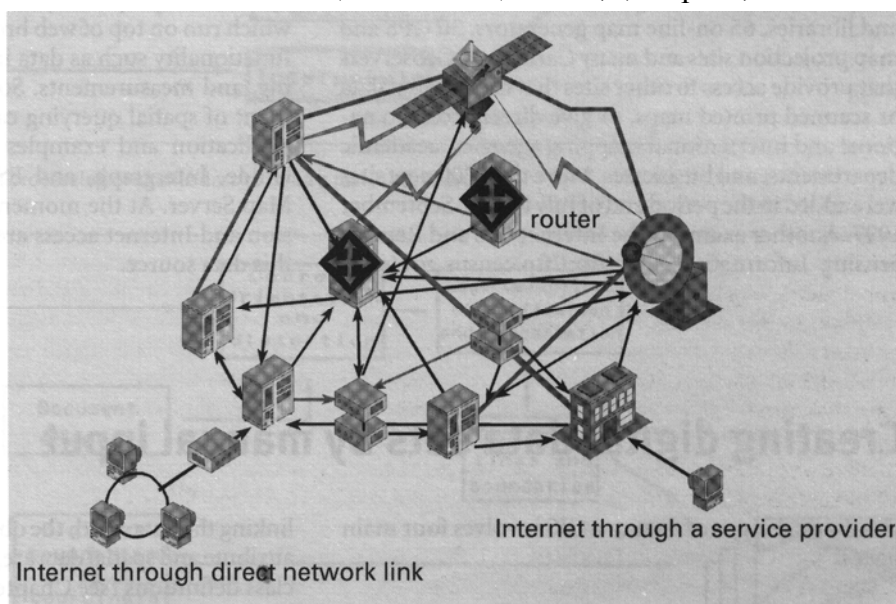
圖 4.2：在不同系統間轉換數位資料 (原文 p.82)

近年來轉換資料的問題已經開始透過標準轉換技術之協議而漸解決。近來有來自於政府

和 GIS 供應商的資料轉換之國際標準定義。例如 DXF 和 E00 的商業轉換型式，由資料供應者和 GIS 的供應商輸入與輸出。可以從一個本地資料庫架構系統輸出資料至另一型式的系統，並能閱讀文件及使它轉變成自己的系統，如同在插圖 4.2b 方面顯示那樣。例如許多 GPS 資料由 IUS NSDI (National Spatial Data Infrastructure-National Research Council 國家空間資料基礎研究委員會 1994)和 CEN(Commission European Normalisation 歐洲標準化委員會-David *et al.* 1996)的政府和國際性組織。許多國家已經建立如英國 NTF (National Transfer Format 國家轉換格式)之轉換型式，此型式不但可轉換格式亦可以記錄部分基本資訊(Guptill 1991)。

在進一步的改進方面，許多商業 GIS 能直接接受其他系統資料而不需要轉換，如圖 4.2 所示。這很有可能在'開放式 GIS' 的概念下討論和研究改進方面，更容易在系統之間轉換資料(在第 12 章節有進一步討論)。它將解決資料輸入的瓶頸並鼓勵更充分利用現存之 GIS 資料。

Box 4.3：網際網路及內部網路（原文 p.83）



數位遙感探測資料或者掃描空中照片資訊已經保留在像素形式(表格)裡，但是，它也許與光柵或者向量 GOIS 的型式不相容。需要各種的預先加工，如調整分辨率、像素形式(方向和變形)和製圖的規劃，以確保與資料庫地誌之兼容性(以章節 5 詳細討論)。其他預先加工活動包括把像素分類為例如城市、玉米田、橄欖林等等的土地覆蓋型，經由網格向量轉變法(在以後的章節中更詳細描述)來產生線性和多邊形的實體。產生的地圖圖像可以轉換為 GIS 可使用之型式。

網際網路，為一個公開企業，是由極多部電腦連結之數位網。由不同資料傳輸媒介如衛星和無線電的聯結線和光學纖維、unshielded-twisted pair、同軸和電話的通信網之陣列組成。將不同裝置的排列成陣式透過簡單信號擴大器連結至數據機在經由路由器發送程式和資料。使用標準網際網路通訊協定 TCP/IP(傳輸控制協議/ 網際網路協定)傳送所有資料。Auser 直接透過網路連接至另一部電腦，或者從商業網供應者(使用電話通信線的形式)撥號而連結。由下圖顯示：

透過在其他位置存取電腦台網之間完成一些角色：

1. 作透過電子郵件和討論組之通信機制
2. 資訊途徑工具

3. 作一個傳送文件工具(使用檔案傳輸協定-文件傳送協定)
4. 由另一個位置(使用 TELNET)執行電腦程式之終端機。

全球資訊網是網際網路使用術語。事實上這是網路之間的主要資訊工具，透過它，使用超連結媒介製作"網頁"資料。利用網路瀏覽器網路'web browser' 連結伺服器'web server'可以連接之全世界。全球資訊網為 GIS 用戶提供資料，甚至提供就業機會的資訊來源。

內部網路是提供本地所有網際網路設備直接連接的網路。它為組織之內散布資訊和通信之網路，它與公眾網路資料分開並可確保連接之資料安全性。

利用網際網路及內部網路為尋找 GIS 資料帶來了新方法(見圖 4.3)。如同往常由不同資料庫得到之向量、網格和物體的資料，網路上亦能取得。舉例如 1997 年 9 月 Roelof Oddens Bookmarks(see <http://kartoserver.frw.ruu.nl/html/staff/oddens/oddens.htm>)列出超過 2,350 個新的數位地圖，比荷蘭的 50 個更多，20 個道路導航查詢系統(見第 7 章)，124 個電子地圖和查詢系統，65 個線上地圖，30 個 GPS 和地圖投影圖和許多的 Carto- 和 Geoservers 提供接近其他地區數位資料給使用者，掃描印製地圖或國家和國際性的繪圖代理機構、學術機構和公司等。自 1997 年 7 月到 9 月中旬增加 200 個新位址，另一個例子，網際網路 GIS 和遙感探測資訊位址(<ftp://ftp.census.gov/pub/geo/gissites.txt>)，提供線上 GIS 和遙感探測資料之位址，在阿爾法 betical 或者訂單。許多資料是免費的(例如 1 公里解析度之數位世界地圖)儘管越來越多的人取得同意書已獲取資料而造成資料供應成本增加，並非所有資料都是最新的，某些具有歷史價值，某些提供每日或者每週的資訊。

網路瀏覽軟體用來存取特定供應地點(網路伺服器)資料或者使用"線上"GIS 軟體(如圖 4.2d)或者使用標準的轉換型式下載至使用者電腦(叫作 client)。專有軟體由在網路瀏覽器的上層執行程式並提供如資料整合、空間瀏覽和測量的基本 GIS 功能。某種支援包括自動書桌地圖指南、圖表和 ESRI MapObject 網際網路地圖伺服器之空間查詢能力，此刻資料傳輸和網路是使用這個資料來源的主要限制。

## 手工操作建立數位資料(Creating digital data sets by manual input)

手工輸入資料到 GIS 包括下列四個步驟：

1. 輸入空間資料，
2. 輸入屬性資料，
3. 檢驗和編輯空間和屬性資料，和
4. 必要時候把空間與屬性資料連接起來。

使用於 GIS 的各種資料庫架構也有不同的輸入方式。主要差別在於第二和第三步驟。空間和屬性的資料個別儲存在 GIS 相關資料庫內(參考第 3 章)並需要先於任何分析連接。作為一個物件型式有關之資料庫，不同資料庫之空間及屬性資料需要類似儲存步驟以便連接這個資料的物體之屬性資料。透過物件導向之資料庫把屬性和空間的資料連接起來了(參考第 3 章)。沒有必要分別連接。

## 輸入空間資料(Entering the spatial data)

以實體模型、地理資料以點、線或者使用一系列座標的地區/像素形式。這些是由地圖或者空中的照片的地理參考系統，或者由覆蓋於地圖上的一個標線或者格線獲得。

輸入資料最簡單的方法是鍵入座標到 GIS 的檔案或者輸入至 GIS 的程式裡。寫入座標



的龐大的勞動可以以打字機替代輸入他們到電腦文件裡，可以使用例如數位轉換器(數化儀)、掃描器或者立體繪圖機的硬體裝置以給 X 和 Y 座標，並大大地減少點編碼。

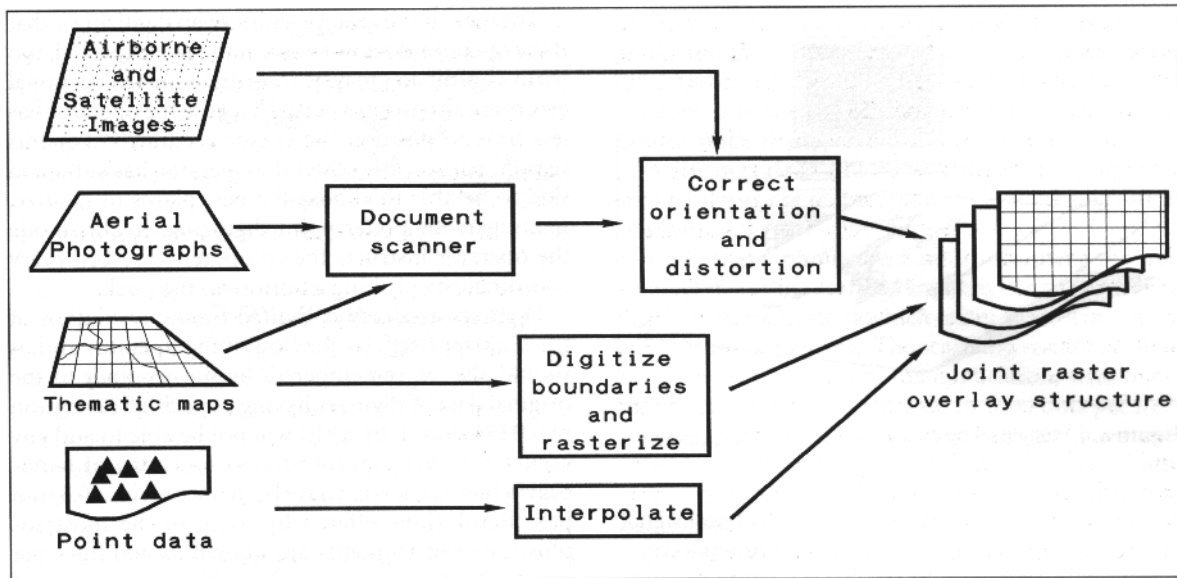


圖 4.3：擷取空間資料以建立網格資料庫之程序 (原文 p.85)

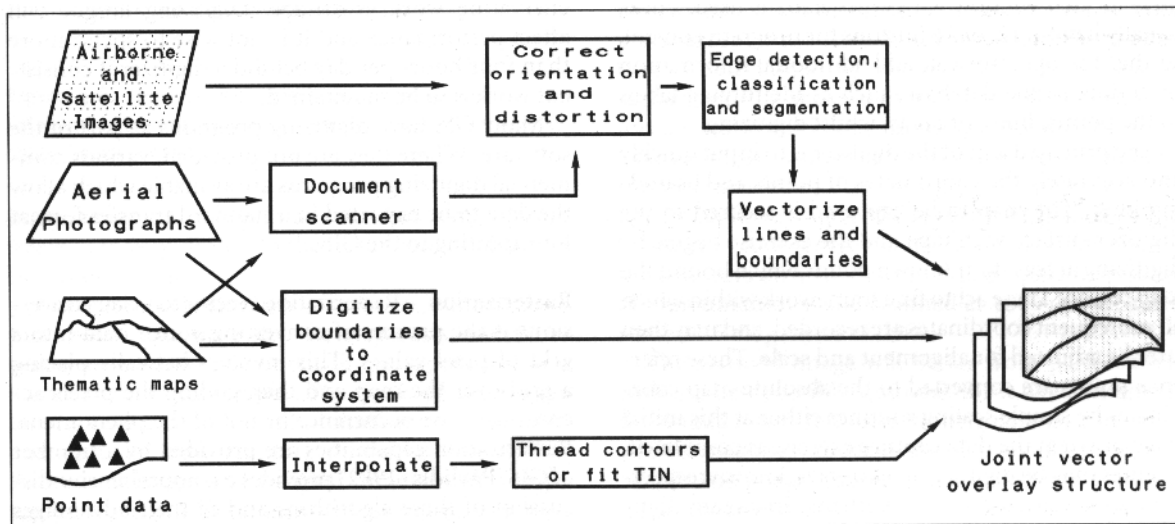


圖 4.4：擷取空間資料以建立向量資料庫之程序 (原文 p.85)

### 數化儀(Digitizers)

一個數化儀是放置地圖或者文件的一個電子或者電磁的版子(圖 4.5 顯示)，將地圖直接放入或直接置於其上，用來紀錄地圖資料點的裝置。目前最常使用的類型是電直角精密電線柵格(electrical orthogonal fine wire grid)或者排列接近尺寸 30\*30 公分(12\*12 英吋)至 1.1\*1.5 公尺(40\*60 英吋)之電磁的面積。可以這個指示器保留在例如滑鼠，或者 puck 或者類似鋼筆之設備中;這些可能有軟線或者沒有軟線。這個游標或者鋼筆放在地圖上的一點上並壓下按鈕，游標的座標信號直接傳送數位化的參考架構資料到電腦。這需要相當大的精確度，由嵌入塑膠中，具有 cross-hairs 的精確線圈組成的 puck 定位。通常 puck 有 4、12 或 16 個控制按鈕，所以操作者可以增加額外的訊息或編碼至資料庫中，例如數位化點、線、或區域時之確認標籤。

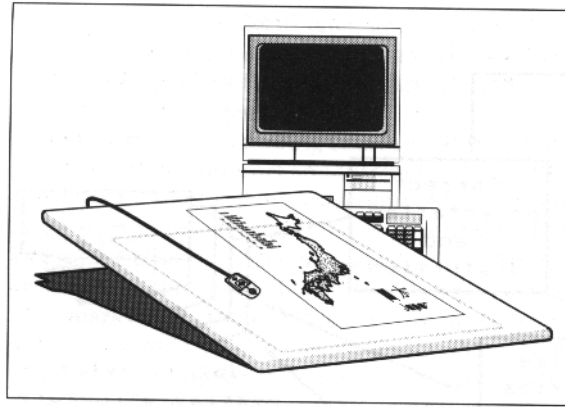


圖 4.5：手工輸入空間資料之數位板（原文 p.86）

數化儀的主要目標是要迅速和精確輸入點和線的座標。以磁帶擷取數字轉換器表面的地圖由測試至少四個已知點開始。固定框架後便可以記錄之後的座標，然後可以調整校正和尺度調整。使這些參考點用例行程序轉變成絕對地圖座標，其時間可以在初期或者完成資料擷取之後。

可以利用兩種方法使線數位化，河流或點分別數位化。在河流數位化方面，這個游標放置在線的開始處，把命令送到電腦台以開始在同等的时间間隔或在 X 或 Y 方向同等的間隔記錄座標和操作員之游標移動盡可能與線的所有彎曲和起伏緊密結合。在線的末端或者在接合處，電腦接受指示停止紀錄座標。數位化的速度取決於這個操作員能夠追蹤線的速度，並且通常平直線之記錄比複雜部分快速。

河流數位化的主要缺點是如果這個操作員不期望較快之工作速度，可能會紀錄太多座標。這個操作員沒有較多時間把這個游標精確定位而使更大的位置錯誤可能會發生。由於這一原因，當這個操作員有足夠技術能選擇最好的位置數位化時，許多人更喜歡點數位化。在這個模式中這個操作員透過一個按鈕指示這個電腦記錄每個座標。

數化儀之精確度受數化儀本身的解析度限制、操作員的技術，也應該注意原始資料的精確度。數化儀有著 0.0254 毫米(0.01 英吋)的解析度，在擷取過程中用 0.4 毫米鋼筆手繪邊界是無法增加其精確度。由一些老舊的設備提供的解析度為每毫米有 400 條線(每英吋有 10160 條線)為  $\pm 0.05$  毫米(Calcomp 全球資訊網)，儘管有現代桌上型數化儀，數位化還是非常耗費時間的工作。如果持續每天花費多於四小時之勞累的數位化工作將影響其性能表現。許多 GIS 軟體的一部分有數位化程式。若是沒有提供，各種商業數位化程式可用的允許以標準的轉換格式輸出這個資料到 GIS 內。

### 網格化(Rasterization)

Rasterization(網格化)是使向量資料轉變成方格像素值的方法。這包含在地圖上放置柵格然後根據現象給像素編碼。Rasterization 能夠提供 GIS 數值，Pavlidis(1982)提供這些算法優點和缺點之全面討論。

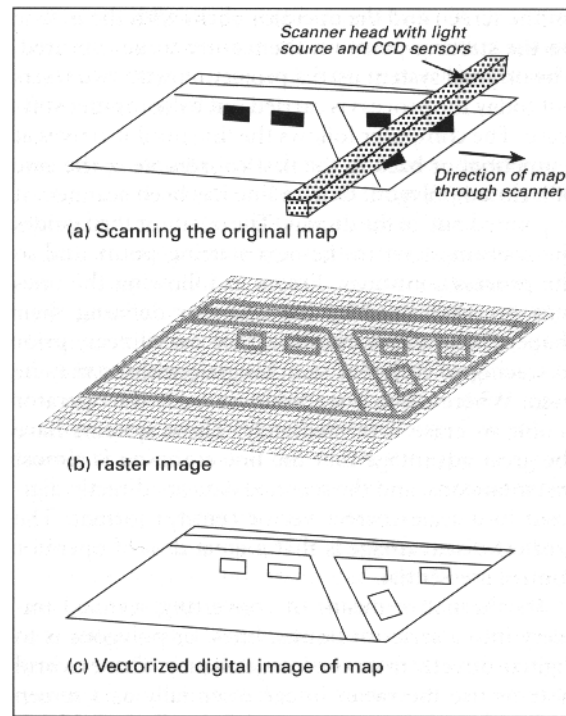


圖 4.6：掃描網格文件之動作(a)掃描原始地圖(b)網格圖形(c)向量數位地圖（原文 p.87）

### 文件掃描器(Document scanners)

掃描器是將類比資料轉變成數位方格基礎的圖像的裝置。他們用於擷取地理學資料以使地圖轉變成高解析度光柵圖像，可以直接處理成向量圖像。其原理是根據將一個地圖上作標記並以反射光束的原理工作；其差別是數位記錄強度(使用 32bits/空格之高品質數位 orthophotos)以柵格反射值編寫的數位圖像。這樣的圖像與由遙感獲得的網格圖很類似。即記錄資料為一個步驟接著一個步驟，與電子照相掃描整個文件之操作模式類似。第一種掃描器是以一個移動光源臂(通常為發光二極體或者穩定螢光燈)和一個高解析度鏡頭的數位照相機。此種照相機有著特別的 CCDs 排列其上，這些是記錄落在他們的表面上光子數量轉換為數位值的半導體裝置(Almelio 1974)。

由移動掃描器或者地圖建立數位二維的圖像(如同在插圖 4.6 顯示)。掃描的地圖安裝在一個平台或轉動圓輪上，平台掃描機之光源為有次序的在掃描文件上上下下移動著。對於寬度超過 1.5 公尺的地圖(60 英吋)，使用掃描器，是以安裝在直立照相機上之光源照明。這時地圖經由輸送機制移動，與傳真機類似之作用類似。滾筒式掃描機則是在一個轉動鼓的外部放置地圖，感應器之移動與地圖方向成直角穿過地圖。

對於這些光學機制的掃描器由移位距離控制掃描的圖像像素的大小，並且由文件或者光源移動速率控制。掃描器之解析度可以任由使用者改變，最高品質之解析度到達 1800 點/英吋(dpi)。此類掃描器如雷射掃描器 VTRACK 系統可以清楚準確的發射雷射光束。(下一章節)

現代文件掃描器使用類似於電子照相法的方法雷射印表機相反之方法，因為掃描表面有大量之光感應點，能夠直接用軟體決定其解析度。沒有移動部分，期望活動光源和解析度由傳感器表面的幾何學和記憶的數量決定而非由機械臂決定。顏色掃描透過每一個像素上一系列光感應的材料作用。機械和電子照相的方法產生大量資料文件：為了縮小給讀者的資料檔案體積，僅以 1600 個 dpi 解析度掃描 30x50 公分之地形圖，大約為 11.5 兆位元資料。

擁有最好掃描器並不一定有完美的掃描圖像，掃描愈清楚，原始地圖上之污跡或缺點亦一一呈現。此時數位圖像可能需要利用交互式美化處理，以操作或者用 binarization 使額外之資料過濾去除(所有像素分類為固定範圍之灰色階以成二進制圖像)和過濾單獨之像素。如果原始地圖品質差，則必須用低灰階尺度對比，可變的背景強度，和干擾合適的 binarization 方法，可以改進掃描圖像之結果(Trier 和 Taxt 1994)。特性和符號的識別程式可以使用於圖像中的標籤符號和

掃描圖像之結果也許是向量化(在下面描述)或轉變成另一個種可以直接輸入 GIS 之網格架構。相稱、尺度校正和調正，可以透過已知點位置來控制，例如可以從掃描的圖像直接定義和自動地處理紀錄基準點。一旦完成這這個數位地圖資料，我們也可利用轉換格式輸入 GIS 中。

### 向量化(Vectorization)

向量化(網格轉變為向量)通常使用特別軟體提供像素轉變沿資料排列之陣式之算法。透過穿越像素線的掃描圖像過程完成。這些把彎曲線還原成僅僅一個像素寬(插圖 4.7 a-c 顯示)。然後自動計算或者透過用戶控制計算相同數值之線性或者面積像素掃描結果。隨之而來的資料需要編輯成代碼，以使這些個體單位和線之連通性正確，尤其如果使用全自動的向量化方法。例如雷射掃描器 Laser-Scans VTRACK 和 Hitachi 的電腦輔助設計 CAD-Core 追蹤者系統的各種系統已經發展，它給這個操作員給得更多對單元擷取過程的控制(Jackson 和 Woodsford 1991)。在螢幕上顯示出這個掃描的圖像結果，並且這個操作員從滑鼠按入擷取之開始點以擷取資料。電腦跟隨著線直到接合處，或者回到第一座標成為封閉多邊形。一旦已經掃描了線，便可以顯示'繪出"結果。此時操作員導引這個雷射對下一個點開始掃描，此過程持續進行。操作員可以由先前追蹤不同搜查算法定義他們的形狀，例如直線或曲線，使跟隨線的錯誤減到最少，一但追蹤線的地方發生錯誤，這個操作員能立即抹掉它。

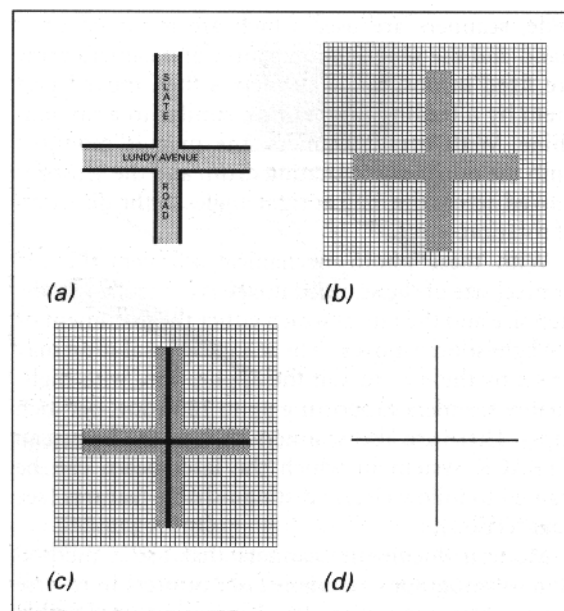


圖 4.7：使線條變細(a)原始地圖片段(b)掃描圖形  
(c)擷取中央像素使其變細(d)中央線以向量呈現 (原文 p.88)

這些系統有很大優點，就是掃描線幾乎是即時的，並且直接使這個掃描的資料以正

確尺度的向量(實體)形式直接數位化。最大的缺點為大量操作控制步驟是不可避免的。另一替代轉換資料方法是使掃描的影像成一系列點、線或者多邊形直接從圖像顯示數位化。這些混合系統基本上把這個網格圖像用作螢幕 **backdrop** 並且由滑鼠控制游標定義各種實體的數位值。這個方法已經成功使用於掃描的地圖，數位照片和衛星的圖像(參考插圖 4.3)。

### 分析立體繪圖機(Analytical stereoplotters)

擷取數位地理資料的第三種技術是立體繪圖機。這 photogrammetric 設備是用成對立體的空中照片直接記錄層次和實體的(從一個些微不同的位置觀察相同的地區)。在最近發展方面，已經用衛星傳感器紀錄視頻和數位的照相機的數位立體圖像，並利用處理圖像之 photogrammetric 系統算法來產生海拔高度資料。此資料之高度和空間的解析度是可變的，但是新衛星之發射給為將來之數位海拔高度地圖提供更詳細之資料。新的 photogrammetric 方法之發展可以產生數位三維的地圖。Beers(1995)描述一個攝影魚眼記錄系統，允許記錄一個大視野之圖像，這可以用來產生三維的視野和座標。

立體繪圖機可以利用於擷取連續性的海拔高度資料之數位海拔高度模型或 orthophoto 地圖，並且在第 5 章有更詳細之討論。

### 輸入屬性資料(Entering the attribute data)

屬性資料(有時叫特性編碼)是在 GIS 內需要處理的空間實體性質，但是不是自己的空間。例如，擷取道路資料可能將鄰接像素一起擷取，由某種顏色，符號，或資料位置顯示實體線和資料於 GIS 上。描述道路類型的資訊(高速公路或者塵土軌跡)可能利用常用之製圖標誌。至於寬度、表面的類型、任何專門交通規則、每小時車輛流量之道路其他資料等也需要記錄，並需要包括在 GIS 內。雖然或許這些屬性值可以直接輸入時與圖表實體相連，但是輸入大量複雜交互屬性時是沒有效的。因此這些資料分別儲存在 GIS 之空間資料內，或者連同空間性質一起輸入物件導向之資料庫。

屬性資料可能來自如紀錄紙，現存資料庫，試算表 (spreadsheets) 等不同來源，他們也許透手工或者透過使用例如 TXT、CSV 或者美國資訊交換標準碼 ASCII 等標準的轉換格式，將資料輸入 GIS 資料庫裡。相關之資料庫利用標識符(在下一部分中更詳細討論)將這個空間屬性資料記錄的和連接一起。

### 資料檢驗和編輯(Data verification and editing)

一旦已經輸入了這個資料，其可能的錯誤，疏忽，及先前之問題被連結至空間及屬性資料。圖 4.8 概括敘述建立 GIS 資料庫之完整程序。

檢查空間資料中的錯誤之最好方法是產生在原件相同規模的電腦計畫或者薄的紙。然後將此兩圖放於透光桌子上以肉眼觀看比較兩個地圖，有系統地從地圖之左至右和上至下操作。遺失的資料、錯誤位置和其他的錯誤應該以不同顏色清楚地標誌在印出之資料。某種 GIS 直接顯示一些地誌或者直接以顏色標識螢幕圖像編碼的錯誤。如果地圖是唯一的，僅僅在地圖邊界之內顯示錯誤位置；如果地圖是覆蓋更大區域的連續地區，或者數位化的資料必須與電腦中已經存在之地圖資料連接起來，也必須檢查接近地圖邊界之空間資料。某些操作，例如多邊形之形成也可以顯示空間資料中的錯誤。

可以由印刷出文件和用眼睛檢查行列之屬性資料。一個更好的方法是經由電腦程式掃描資料檔以檢查所有之錯誤，如同文字一樣而不是數字，超過一定範圍之數字等等。

(第 8 章節更詳細討論)把在空間的和屬性的資料的擷取期間可能出現更多錯誤。如下歸類：

(一) 空間資料是不完全或者雙份的。

- 於手工輸入點、線或者細格的空間資料時，經常省略部分資料而使得資料不完全。在掃描向量方格圖的資料轉變過程中，通常省略線之間縫隙，沒能線的所有部分連接起來。同樣地，掃描的資料轉變為向量方格時導致不需要的”spikes”出現。數位化的線也許不只一次輸入線和節點或許不會接合成交叉。

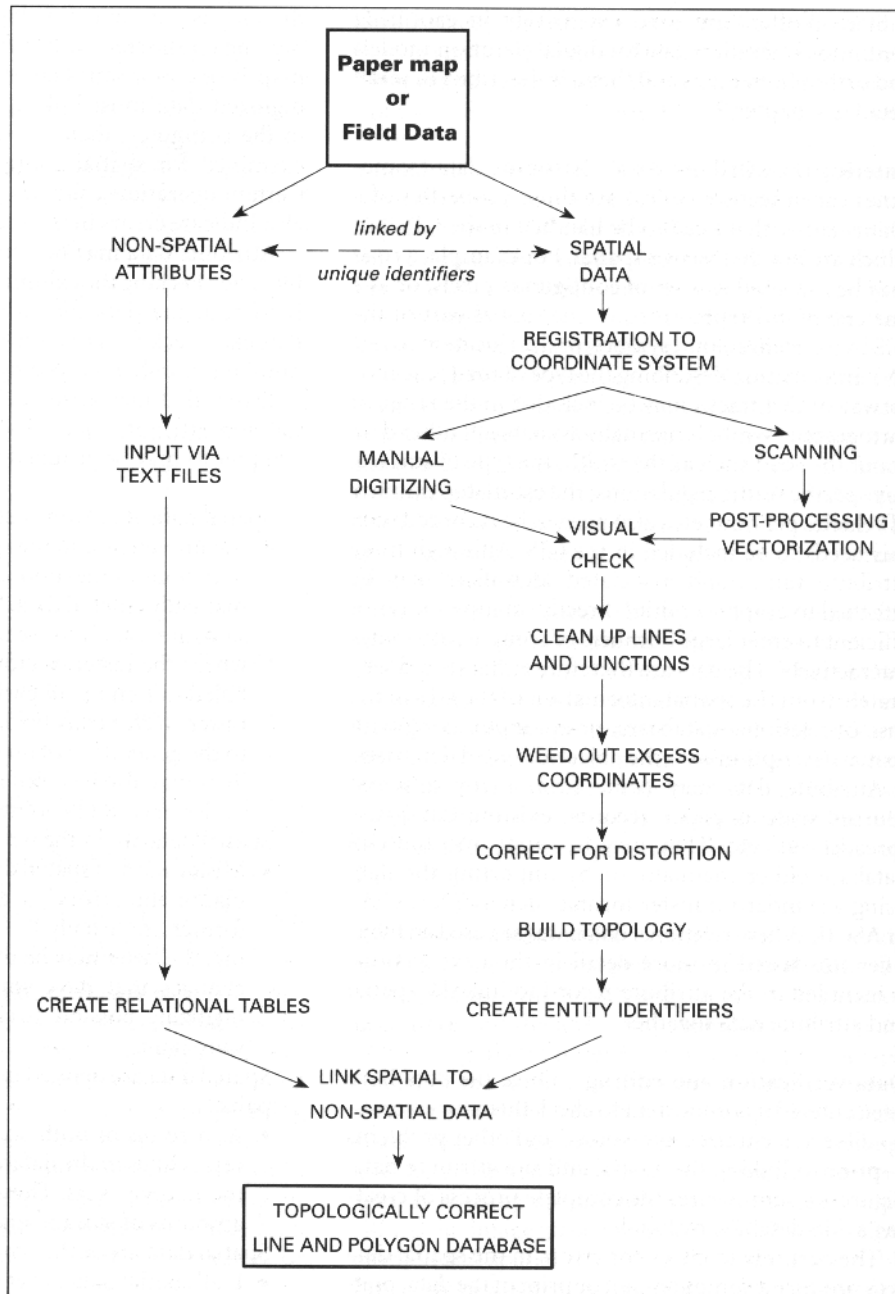


圖 4.8：建立正確向量多邊形地誌之步驟（原文 p.90）

(二) 錯誤的空間資料

- 位置錯誤可能從小錯誤延伸為總體空間錯誤。前者通常是草率數位化的結果；後者也許是因為在數位化期間，初期出現不明原因錯誤或者尺度變化的結果，也有可能是硬體或者軟體的錯誤。

### (三) 使用太多種座標定義空間資料

- 由於掃描或者數位化過程，使用太多點資料定義資料庫中的線。這些可能佔去太多的儲存空間。

### (四) 錯誤的空間資料尺度

- 如果所有空間資料出現錯誤尺度，則通常是數位化之尺度出現錯誤。在 georeferencing 過程中使用了不正確之數值，掃描資料的錯誤通常會增加。

### (五) 空間資料變形

- 空間資料可能因為用不正確尺度的基本地圖來數位化以致變形。大多數空中照片由於飛機的傾斜，地勢起伏和物體與鏡頭距離不同而無正確尺度。所有紙地圖都由於展開時在一個方向比另一個方向更大。此外，紙地圖與田地文件可能由於暴露在雨、陽光、咖啡、啤酒和時常摺疊導致變形。從一個座標系統轉換至另一座標系統，例如 Universal Transverse Mercator(UTM)可能需要不同於輸入文件或者圖像的資料庫座標系統。

這些錯誤大部分需要透過 GIS 直接提供各種的編輯和更新的支援功能。交互式的過程可能比輸入資料本身消耗更多時間。資料的編輯校訂通常由查看電腦台螢幕上所顯示之地圖錯誤部分，經由軟體，透過使用鍵盤、滑鼠控制游標，或小型數位板之控制來校正。

可以透過螢幕游標使這個空間實體移動或者指出他們在數位板上的位置以改正向量資料庫中的小的位置錯誤。某些 GIS 系統中，需要時可以直接利用電腦指令來使圖形實體移動、轉動、去除、插入、延伸或者截斷。可以透過數位儀或者鍵盤增加新資料。某種實體編輯之操作無法單獨操作，但是，必須由檢驗或者運算以確保資料庫的一致性。例如，在實用地圖的網路中(例如電話線)編輯單一線到兩條分支線需要重新組合信號連貫性，在多邊形網路中，其中如果有一條線線或者一部分線移動或者改變，這個多邊形地區必須重算。

資料按比例排列之問題可以由把簡單數字因子應用於資料中而克服。更複雜之轉動和轉換操作需要將一個較精確之基本地圖和一個變形的地圖放置一起。應該把有缺失之地圖與基本圖比較，並且與基本圖上一些點連接起來，以校正缺失圖上位置至正確的位置(圖 4.9)。數學轉換延伸和壓縮原始地圖，直到這些連接向量長度調到零位，並且彼此顯示出這些 tie points。然後原始地圖上的所有其他點正確地重新安置。這個過程叫作 robber sheeting 或者 warping，缺失之地圖像彈性片一樣向四面八方伸開以適合另一個地圖。網格資料由於有固定的方格和資料的架構所以不能直接使用；在第 5 章描述歪曲網格架構的合適技術。

過量之座標定義線可以移走，使用為人所知 Douglas 和 Peucker(1973)，Reuman 和 Witkam(1974)之鋤草算法(weeding algorithms)移除。也許透過使用 B-splines 可以改良平直線片斷之視覺效果。

網格資料中的屬性值和空間的錯誤必須藉由改變失誤細格值來改正。它是由簡單命令輸入正確的屬性值至數位化細格中的。如果大量錯誤細格，新的資訊將會覆蓋現存數值。

一旦已經改正了空間錯誤，向量線之地誌和多邊形網路能夠如第 3 章節所解釋的那樣產生。

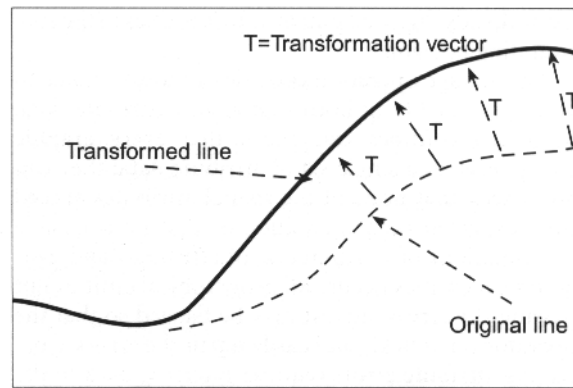


圖 4.9：對 rubber sheeting 轉換向量（原文 p.91）

### 連接空間及屬性資料(Linking spatial and attribute data)

與 GIS 資料有關之資料庫架構，手工擷取資料之最後過程包含涉及將屬性和空間的資料庫透過共同識別符號相連接。道路可以把它確認為 A1。這些聯結在首次建立資料及資訊處理期間核實，但是不是不可變更的，並且在用戶叫特別資料部分時重新建立。

點和線資料的校正能夠自動地完成或者需要需要在數位化或者掃描/向量化過程間以手工完成。多邊形的校正需要特別的方法校正，可以透過交互式數位化(interactive digitizing)，或者透過”點在多邊形中(point-in-polygon)”算法，從已經數位化的點或者文字實體至多邊形周圍來轉換校正代碼。

連接操作提供一個合適機會以證實空間及屬性資料之特性。例行審查手續能夠檢查每個圖表實體都收到一組屬性資料；也能夠確保沒有空間屬性值超過他們的期望的範圍，或者出現非常理的屬性及地理實體結合。所有無法通過測驗的地理實體可能被打上記號，操作者可以快速且容易的修補錯誤。改正錯誤的屬性需要查看原始資料，以檢查任何錯誤並輸入所有的失誤及錯誤數值。

連接空間和屬性資料之程式可以來檢查所有聯結是否洽當。所寫的程式應該可以將錯誤標示出來，空間和屬性資料之間的不正確聯結通常起因於數位化或掃描期間，或者建立多邊形地誌時，於空間資料中輸入不正確識別代碼的結果，編輯空間資料包含改正多邊形地誌及資料的標識代碼中的任何錯誤。當這數值於與法正確但是實際上是錯誤時，點出屬性資料的錯誤是非常困難的。

### 建構資料(Data structuring)

在資料擷取過程中，可利用時下兩種方法之一種校正資料，一是幾何學方法校正網格資料，或者用拓撲學和幾何學方法校正向量資料。新型 GIS 系統接受此兩類資料，但必須於自一個格式轉換至另一種格式時重新建構空間資料，這也許要取決如何應用(參考第 7 及 8 章節)。向量至網格及網格至向量之轉換在這個章節前面已經討論過。

建構全面性的 GIS 資料庫需要有效率的儲存能力及查詢系統，透過如章節 3 描述那樣的合理指令和資料壓縮技術。例如至少在一個商業 GIS 系統中支援資料的 quadtree 編碼，並且完成基本選單命令。藉由屬性和空間的資料的索引或者透過 RDBMS 的正常化提昇資料庫之效率(章節 3 再次描述)。

### 資料的呈現(Data presentation)

一旦資料庫完成，同時亦能進行查詢和分析的操作，在第 7 及 8 章節有詳盡描述。資料



檢索和分析的結果必須以用戶可理解的形式製作或者允許資料轉送至另一台電腦系統。大多數 GIS 系統包含有支援資料輸出的軟體。

數位或者"computer-compatible"輸出，能夠以光磁碟文件形式或者其他的存儲裝置使另一系統能再利用或者讀取。另一方法則是可以以光纖或電話線例如網際網路來輸出資料。

類比或者'people-compatible'形式輸出為地圖、圖形和表格之產品輸出裝置，這些能夠在例如電子螢幕短暫的顯示或在紙或者聚酯薄膜的穩定基本材料上產生永久的圖像的。最近幾年，許多小公司發展出整合於 GIS 資料庫中之特別製圖工具以輸出美觀的圖形。這些例如地圖圖例、標題、方向指示(北方箭頭或者座標記號)和一個比例尺或者比例尺圖，以及允許加入易於表達資料的顏色，符號和文字。顯示時亦可以選擇使用平滑等值線或顏色或者色彩較暗之 choropleth 地圖顯示。三種輸出高品質地圖製品的裝置為：直接顯示之物件 visual display units(VDUs)，繪圖機或印表機。

### 直接顯示之物件(visual display units (VDUs))

電腦螢幕或者視覺的顯示單位直接對人們展示出 GIS 分析結果之主要工具。尤其對於網際網路使用者及線上工作者更為重要。各種尺寸之電腦螢幕使用常規陰極射線管(CRT)或 light emitting diodes(LEDs)陣列技術，使影像顯示在電腦螢幕上。陰極射線管之內分別有發射紅色、綠色和藍色光之電子槍組成這個螢幕的光感應粒子以匯集成螢幕影像。當光發射至螢幕粒子上，此粒子便被激發並發射出一種顏色的光，而當成千上萬個發光點集成之光束發射則形成影像。陰極射線管光束有系統地在水平和垂直的方向，依次輪流激起這些粒子穿過螢幕，並且一秒更新掃描這個整個螢幕 60~80 次。

大型電子顯示(LED)則有三層，每一種顏色有一層。由 GIS 系統軟體控制陰極射線管發射光到螢幕的特定部分，並且由電腦控制這個圖形轉接器(graphics adapter)。軟體把信號依次送到圖形接合器(graphics adapter)以輸出圖形。這個後面是分離的卡片或者控制送到陰極射線管信號的電腦內的處理機板，對於控制圖形有小記憶容量。

這些輸出裝置的差別取決於硬體本身(圖像顯示部件和圖形學的接合器)(the VDU and graphics adapter)以及 GIS 顯示軟體。顯示之詳細程度取決於螢幕解析度，顏色的數字，及顯示資料的尺寸和尺度。複雜資料庫並無法完整顯示，所以多數系統有變焦放大"zoom in"功能以放大顯示某部分資料庫。

大多數 GIS 用戶所看到分析的結果顯示在螢幕上的一個靜態地圖，圖形或表格。並可以使用特別裝置得到如同攝影般清晰的複製影像或者儲存在電腦記憶裡成為暫存檔，隨後便可列印出。最近電腦技術上的進步，允許增加存儲和觀察的資料的數量，允許顯示出更複雜連續動態影像，不管是觀眾或者環境，特別是當研究範圍改變時可以應用於 GIS 模型上。預先顯示是一項的很大的改進，每一個連續時間步驟的數值可以以單一靜態圖像展示來解釋結果。極少的商業 GIS 系統支援動態顯像能力和產生這種類型的顯示需要特別的軟體。

透視圖法，三維的顯示或者區塊的圖表(有別於 drqping)是是顯示地形主要資訊之方法。能夠提供二維的地圖不可能完成的視覺印象。在圖 1.4 的例子中顯示出，(自地形的數位海拔高度模型取得 orthophoto-第 5 章節)，章節 3.8、3.7 和插圖 5.13，8.13 和 8.16。通常以顏色做皺折圖像顯示，因為人類眼睛不能區分單色灰色色階以完全明白主題資料之涵義。透視畫法及動態顯像之結合，像電動遊戲一樣使用，是在 GIS 系統顯示展出空間時間模型結果的強大工具(例如 Van Deursen and Burrough 1998)。能夠由除了 fenne 地圖以外的幾個專門供應商銷售的一系列方法，顯示出土地三維區塊內的變化和 cut-away 普通圖表(如章節 3.3 和 3.4)。

## 繪圖機(PLOTTERS)

繪圖機是一個自動化製圖器，他們是複製文件或者檔案之地理學資料，而將其輸出之裝置。他們擁有供紙或者保持機制，通常是一個滾桶或者平的表面和放置顏色的鋼筆繪圖臂(圖 4.10)。由水平及/或垂直地移動使這個圖畫裝置和紙成為二維的線圖(X 和 Y)。所有資訊由軟體之指令控制繪製一系列線：移動繪圖裝置點出 XY，設定筆往下，把筆移動到 X'Y'，把筆移動到 X''Y''...提起筆。移動 XY 座標是使用絕對或者漸增的座標。

繪圖機的靈活性和速度取決於 preprogrammed 資訊的數量，是由繪製複雜字母和符號得來。'精明的'繪圖機含有 preprogrammed, 'hard-wired'特性和符號的設置僅僅需要由簡單幾個電腦指令完成。繪製圖像的品質主要取決於控制這個繪圖裝置的馬達的移動距離大小。

最近在繪圖技術中有主要變化，是以噴墨裝置的或者氣泡噴射裝置取代顏色筆，僅僅從前 physically 印刷機(參見下一部分)。這些有更快和更安靜優點得和他們可以像繪製線圖一樣迅速和均勻地將實體面積填滿。他們列印在各種材質上，包括平紋布和有光澤的紙、軟顎 velum 和聚酯軟片。為單色之解析度為 720 dpi，彩色為 360 dpi。於攝影材料上使用之雷射層繪製地圖，其精確度更高。

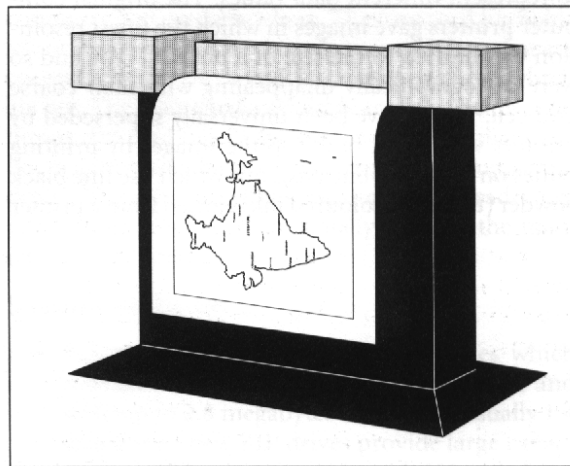


圖 4.10：大型地圖繪製機（原文 p.93）

## 印表機(Printers)

早期印表機為類似線圖印表機或者列印的終端機簡單型式的網格輸出裝置，必須用不同鍵盤特性的結合來印製不同資料數值。這些原始電腦印表機所印之圖，其最好的解析度受限於圖像尺寸大小及物理限制，而繪製出無吸引力之粗糙的網格 cartographicall 圖。雖然現今印表機已經取代以往之印表機，但是仍然是以印刷數值為基礎，使用黑粉(碳粉)一行一行的列印圖像或者從印刷機頭印出顏色以建立圖像。現今已能列印出更高解析度的圖像，最高到 A0 大小。解析度是由移動速度和噴墨頭的尺寸決定，某些印表機可以有 720 dpi 之高品質輸出。

以碳粉為基礎的印表機如雷射印表機，就像們所熟悉的影印機(零過程)一樣。他們透過電腦處理程序輸出雷射信號('on' or 'off')對一個帶電的滾筒掃描使它轉變成烙印。此時滾筒經過碳粉開使其帶電，滾筒上的那些點黏住黑粉信號以形成圖像。紙透過這個滾筒並將是量的碳粉轉移到紙上。再經過兩個加熱的盤子之間而使它碳粉永久黏附於紙上。彩色雷射印表機根據相同的原理，以三個不同雷射光束射向滾筒以完成全彩需求。

噴墨印表機是根據 GIS 軟體之電子信號，利用微細噴頭噴出不同顏色之墨水印出圖像，

控制噴射的兩個主要機制已經發展完成。噴墨印表機用電子方法將墨水自墨水匣轉移到紙上。氣泡式印表機有含有一系列厚度大約 1mm 之噴管。使用一個簡單的機制使噴管加熱產生氣泡並從那裡釋放墨水。每一個噴管都有發熱器，當電流產生脈波使每秒增加幾千個溫度。每一個創造出之微小氣泡使噴管之內的壓力增加，迫使單一，微小的墨水滴排出。此時壓力降低，真空驅使新墨水開始同一過程。

## 更新資料(Data updating)

許多地理資料並非一成不變，而是會隨著時間變動。僅有少數的變動是固定性的，因此他們可以自動的呈現。例如，政治界線會隨著議會的突發奇想而改變，土地的利用與田地的限界會隨著重新劃分的結果而改變，土地增長與剝蝕的結果會改變土壤的線界，若這些地景的改變沒有列入資料庫中，則資料庫的可信度及完整性將會降低，因此資料的更新是需要的。前面所提到的基本編輯方法即適用於輸入新的資料、變換現存資料位置及引入新的或經過轉換的屬性參數。當空間架構需要改變時，資料集的架構即需要重新更改。

資料的更新不僅止於舊資料庫的修改，它還意味著重新的調查與新資料的收集。一些地球表面的特徵改變的很緩慢(如岩石的種類)且重要的改變很少，因此資料的更新問題不大。然而，有些其他種類的地理資料則可能每幾年全部重新調查比更新舊資料更符合成本效益。相較於更新屬性資料，提供屬性資料記錄與空間實物一對一連結較為重要。

## 儲存資料(Data storage)

建立一數位資料庫極耗費與耗時，其中將數位地圖資訊由電腦區域性記憶體中轉換至一能提供較長時間且安全的保存媒介是必要的步驟。有些地誌、地籍與環境製圖的資料庫在適當的保存狀態下通常適用期可達 1-25 年。

數位資料儲存於磁性或光學媒體中，其形式多樣且依形式不同而價格各異，另依資料使用的時機決定適用的形式。以前資料由電腦的硬碟(較貴)傳送至九軌 0.5 吋的磁帶上，此種形式較安全、便宜且可靠，因為磁帶上的磁性分子不會隨著時間而訊號流失，它們被儲存在安全的區域，之後經由磁帶驅動機於每次使用時讀取，此種可攜式媒體稱為”可攜帶式儲存媒體(removables)”。今日技術的快速改變使得硬碟更完備且便宜。越來越多的資料被儲存在”固定式儲存媒體(non-removables)”即使資料並不常使用。

## 可攜帶式儲存媒體(Removable Storage)

可攜帶式儲存媒體為可以將資料帶離原電腦至別處使用的磁性或光學儲存媒體，它們被用來暫存與備份(為安全理由)經常使用的資料以及儲存不常使用的資料，主要可獲得的媒體有列於下，其中桌上型電腦中最常見的讀取機器為軟碟機與 CD ROM，其他的產品則需要特殊的讀寫機且通常用於主要資料儲存電腦。

(一) 磁性媒體有：

- 軟碟 3.5 吋塑膠外殼中有一圓形的磁帶片，通常儲存 1.4 megabytes，最多可儲存 2.8 megabytes，新的壓縮程式可將容量提升為 100 megabytes 至 1 gigabyte。
- DATs 全名為 Digital Audio Tapes，通常為卡帶或筒狀容器其中內含 4 mm 與 8 mm 磁帶，可儲存數 gigabytes 的資料。
- 0.5 吋大線軸儲存於金屬盒的磁帶，10 年前最為通行，現在某些機構還有，儲存資料量約為 520 megabytes。

(二) 光學媒體有

- 塑膠樹脂碟片可用雷射紀錄數位資料，其最常見的形式為錄製音樂。
- 光學碟片(同意詞為 CDs – compact disks)通常直徑約為 13 公分(5.25 吋)，可儲存 675 megabytes 且可依可否寫入資料至碟片上分為：
  - 2CD ROM(Read Only Memory)為唯讀媒體，許多資料與軟體通常以此儲存。
  - 2WORM Optical Disks(Write Once Read Many)僅能寫入一次資料，之後使用者僅能讀取資料。
  - 2Rewritable Optical Disks 通常保存於塑膠盒中，可讀寫多次，為儲存量且有效率的碟片。
  - 2laser disks(are large 12 in disks ??? 不知在說啥)通常為唯讀

通常兩個基本步驟將資料儲存於這些媒體，Archiving 為將資料由電腦藉由專業的軟體或硬體將資料壓縮傳至媒體以達最大的儲存量，此類的資料若要用於 GIS 必須再”儲存 restored”一次。第二個方法為大量儲存系統(Mass Storage Systems MSS)現在正大量的被應用，有別於 archiving 系統，此類系統的資料以原來的形式儲存，因此可以直接用於 GIS，可複寫式光碟已漸廣用於此系統，有些電腦具有可以裝許多片光碟的裝置，稱為光碟抽換盒(optical juke box)，因此電腦即可隨時讀取任何一片光碟。

### 不可攜帶式儲存媒體及網路(Non-Removable Storage and Networks)

每 megabyte 的硬碟價格下降使得可攜帶式儲存媒體越來越不經濟，gigabyte 的資料現在已可儲存於個人電腦，此轉變主要的問題為資料的分享，使用者必須使用他們原來的電腦或是有灌入他們所需資料的電腦，如此非常的不方便且會導致資料的重複儲存，因而難以知道哪個版本是最新的。在許多使用者都要取用相同的資料或軟體的地方，有些機構利用電腦網路以減少這方面的問題。

有兩種主要網路，區域網路(Local Area Networks LANs)與廣域網路(Wide Area Networks WANs)。區域網路為一地域中的電腦連結而成，廣域網路則為遙遠地理區的連結網，電腦是由纜線(銅線或光纖電纜)或非纜線(無線電、微波、雷射與紅外線)連結。非纜線的連結多為點對點的連接且通常用於廣域網路，電纜的連結則多用於區域網路，亦可用於廣域網路。

網路中的電腦架構也許可以描述為下列幾種形式：

- Peer-to-peer，兩部電腦連結在一起分享檔案，此類的連結方式較少用於 GIS，因為大檔案的傳輸會影響電腦的表現。
- Client-server，有一或多台電腦(伺服器 server)專職於儲存資料及軟體，用戶端電腦(clients)具有自身的處理能力但可與伺服器連結取得資料與軟體。伺服器管理檔案的傳輸與儲存、中心列印，在某些的系統中亦可支援資料庫的諮詢與操縱作業。
- Central processing systems，主要是由主架構與小型網路構成，具有一極強的中央處理電腦用來儲存所有的資料與軟體，另經由一系列的使用者終端機讓使用者進入，所有的過程皆由主電腦完成。

GIS 的網路系統通常屬於第 2 及第 3 類型，且漸往 client-server 系統走，此系統提供較高效率的資料傳輸與列印功能，因容量的大小使得大多數的 GIS 軟體都裝置於客戶端，網路式的儲存方式並非沒有其問題，其問題包括複雜的資料取得與確保能提供最新的版本，一個以上的使用者同時更新或使用同一個資料庫，允許大量的使用者進入與更改資料源的能力。因此讀取與使用的限制須加強。

### 問題(Questions)

1. 解釋如何收集地理資料建立下列結果

速食店之位置

山區之山崩地區

地下水污染之擴散

緊急事件單元(警察、火警、救護車)

旅遊資訊系統

高山地區植被相改變之監測

空氣污染傳播之監測，例如 1986 年車諾比爾事件 137 Cs 隨雨水沉降之追蹤

2. 進行下列工作時錯誤之防範步驟(a)數位化及(b)掃描空間資料時
3. 有哪些裝置或媒介可以儲存或備份資料？你如何改進其安全係數？
4. 什麼是地圖所傳達最重要的訊息？製作好地圖之條件？你會採用彩色或灰階展示資料？

### 建議閱讀之文獻(Suggestions for further reading)

- Burrough, P. A., and Masser, H. J. 1997. *European Geographic Information Infrastructures : Opportunities and Pitfalls*. Taylor & Francis, London.
- Guptill, S. C. 1991. Spatial data exchange and standardization. In D. J. Maguire, M. F. Goodchild, and D. W. Rhind(eds.), *Geographical Information Systems, I : principles*. Longman Scientific and Technical, Harlow, Essex, pp. 515-30.
- Kennedy, M. 1996. *The Global Positioning System and GIS*. Ann Arbor Press, Inc. Ann Arblr, 268 pp.
- Maling, D. H. 1973. *Coodinate Systems and Map Projections*. Georage Phillip, London.
- Rhind, D. H. 1992. data access, charging, and copyright and their implications for grographical information systems. *International Journal of Geographical Information System* 6 : 13-30.

## 第五章 由點資料建立連續面資料

### *Creating Continuous Surfaces from Point Data*

#### 前言

本章是解釋由空間上屬性變化的製圖，以建立分離及連續的表面的方法。資料來源一般是觀察稀疏分布的樣點如地下水來源，土壤剖面，氣象站或是的地覆或動物數量存在及缺乏資料或基本空間單位如網格是行政區的人民或是市場出口(outlets)。結果通常由固定的網格內插並由彩色或灰階的地圖或輪廓線展示。文章描述空間取樣的決策及空間預測的方法，包含全球分類方法、回歸及區域決定的內插方法如徐昇多邊形，反轉距離權重及 thin-plate 曲線尺。每種方法藉由做過的例子說明，以提供數學及證明應用方法的基本介紹。數位高程模式及數位正射圖則由特殊的連續表面的例子說明。

#### 空間內插推估：是什麼？為什麼需要？

內插是，由在相同的地區(範圍)或面積內測量獲得的點的位置預測於未取樣地位屬性值的程序。由存在的觀測值，預測所覆蓋地區(範圍)外的地位的属性值叫作外差法。內插是用來將資料從點觀測值轉變成連續性的區域，因此能夠把這些測量的空間格局取樣的模式與其他空間實體的空間格局作比較。內插是必須的：

- (a)當離散的表面有不同水準的解析、格子大小或是方位的需求。或是
- (b)當一個連續面是由一個不同的於需求資料模式代表時。或是
- (c)當我們的資料不完全包括主要的興趣（例如，他們是樣區）。

(a)的例子是掃描的圖像由一個固定大小的棋盤形網格或其他方式轉變（文件，空中照片，或遙測影像）。這個程序一般所知如迴旋(convolution)。

(b)的例子從某種棋盤形到另一種棋盤形的連續面的轉換（例如：TIN 轉成網格、網格轉 TIN，或是向量多邊形轉網格）。

(c)的例子是資料將設置的樣點成離散、連續面。我們必須區分用稠密集取樣的路網到稀少取樣路網(network)的情況，或者沿著寬廣空間的橫切面收集的資料。

密集抽樣網路常由便宜並且能夠直接觀察屬性的空中照片和衛星影像的資料來源建立高程表面，以代表土地表面海拔高的變化。（稱為數值高程模式DEM，也叫作數值地形模式DTM）。稀少取樣網路經常利用borings費用，實驗室分析，和最常見的田野調查，感興趣的屬性的空間變化無法直接看到，需間接地取得。許多使用連續面的空間模式會在第八章解釋。

從內插獲得的連續面能夠用在地理資訊系統的地圖套疊，或由本身展示。表面可藉由輪廓線(等值線)、分離的固定網格、不規則方塊的數值模式表示。使用的資料結構是固定的網格（網格），等值線，或是三角形不規則網路（方塊）。因為內插表面在空間上非常連續，固定網格的內插需要藉由在每個具有不同值的網格單元的資料結構表示：因此空間儲存網格資料結構如run-length codes（長的編碼）及quadtrees（四分樹）不容易被使用。原始資料可能由規則或不規則的空間或時間樣點收集而來，或由網格化的地表如遙測影像或經由文獻掃描所獲得的影像來取得。經內插的屬性預測通常由量測的相同資料形式表示，但一些內差的

方法提供指標函數評估方法顯示所得到的值是可接受或分類發生的機率。

## 連續地表的視覺化

連續地表通常由影像或線條表示。影像的方法包括以規則或不規則的網格及tessellations由彩色或灰階值的等級表示地圖屬性值的變化。'Two and a halfD' (2又1/2立體) 表示在連續的表面上褶皺'draping'屬性的表面的完成，代表土地的地理位相（如插圖1.4，4.3或圖5.13）。

線性表示包含等值線（具均等值的線），垂直的切割(slices)（剖面圖），關鍵(critical)線，如嶺線，河流，海岸線，和坡度折線。線及影像的表示方法可結合增加圖形表示。

雖然內差的表面顯示沿著三個資料軸的變化，即x、y座標軸及作為內插屬性的座標軸，它們通常不被認為是3D表示。'三次元'名詞通常（適當地）保留在由3D空間架構提及的連續性的屬性變化的情形。實際三度及四度表現及視覺化需藉由不常在標準的GIS工具箱發現的特殊軟體(Nichols et al. 1992)。連續的值通常由fence diagrams(柵欄圖示)表示（由3D影像切取，插圖3.3及3.4）或藉由3D展示波狀集中地區의影像，如ore body或污染的煙雲。

## 空間內差的原理

空間內插及外差的背後的原理非常常見的觀察資料，平均上，空間上非常接近的點的值比分離的點特徵更為相似。一般而言，兩分離幾公尺的觀測點比距離數公里的小山具有更相同的高度。然而，分類在推估地圖單位平均預測未取樣點位置或分類等級的中心概念也是一個普遍的方法，不顧任何分級內量測值的空間關係。當平均屬性值是由同質的分級或土地單位計算，如地質、土壤、土地覆蓋、或植被圖的地圖單位，所有小範圍的資訊將會流失。多數的內插法器嘗試使用區域資訊以提供更完整描述地區內屬性變化的方法。假如這些變數可以成功計算，我們或許可以期待我們評估的看不見地位的屬性值將會比單獨由分類平均獲得的值要好。如此結構的圖籍在後來的GIS套疊分析或是量化模式。可以產生較小的誤差。

## 空間插入推估的資料來源

連續地表的資料來源包括：

- 立體的航空照片或套疊衛星攝影影像。
- 由衛星或飛機掃描或文獻掃描。
- 直接量測的樣點屬性或間接的逢機、結構的、或是直線取樣格局，如固定切面或數值化曲線。
- 數值化多邊形/地誌圖。

許多內插的資料來自變動複雜的格局較少數點的取樣。這些量測法通常是所知道的'硬體資料'。當資料稀少時，它在自然的過程及現象所具有的資訊是非常有用的，可能由'軟體資訊'製造我們所知的格局，並可以支持內插。許多的例子，無論如何，自然的過程是未知的，我們必須滿足在自然空間屬性變化的問題的不同假設。這些包括假定取樣變異的平滑性及關於機率分布的及統計的固定性的統計假設(見下章)。

## 空間取樣

取樣點的位置對後來的分析是重要的關鍵。理想上的製圖，樣點在全區須均勻的分布。

完整的規則的路網取樣是具偏差的，然而，當它在地景的固定格局頻率一致，如規律間隔的水流及林木，基於這理由統計學家喜愛某些計算無偏平均及變方的逢機取樣。完全逢機的樣點位置也具有許多缺點。首先，每個點的位置是分離的，反之規律的網格僅需要起源的位置，方位，及間距以固定每點的位置。這些在茂盛森林及艱難的地形是非常簡單的，甚至是GPS。第二，完全的隨機選擇可以導致點不平均的分布除非許多點已經量測過，但這個經常受限於成本。

圖5.1代表主要有價值的選擇。逢機及固定取樣的好的折衷方法是藉由分層逢機取樣，即單獨的點是由固定排列的區塊及層級逢機選取。群叢(或nested)取樣可以使用在數種不同尺度的空間變化的測驗。固定的剖面取樣通常使用在量測河流、海灘、山坡、數值輪廓線的剖面，在由取樣印刷地圖製作數值高程模式上是一個常見的方法。

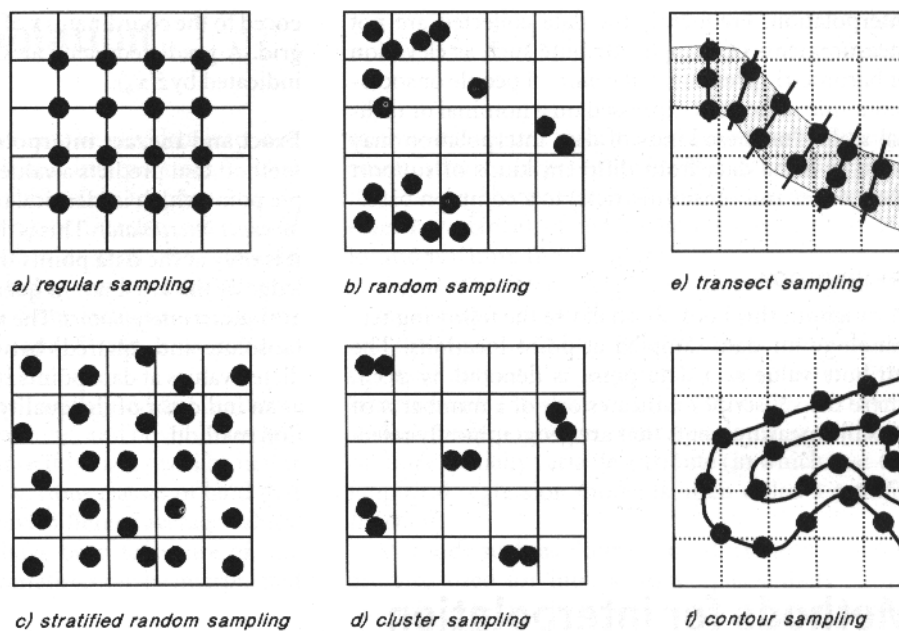


圖 5.1：收集空間點位置的資料的不同種類的取樣網格（原文 p.101）

## THE SUPPORT

**SUPPORT** 是使用在地理統計上的一個技術性名字，由自然取樣所量測的面積或體積獲得。假如樣點是由土壤剖面凹處取得一公斤重的土壤，則 **SUPPORT** 大體上將是面積為 10\*10 公分及厚 5 公分。假如取樣是由取樣試管取得的一公升土壤水，則 **SUPPORT** 具有體積的特點，因為實驗室的分析方法通常將複雜及費事的取樣均質化，所有內部結構或變數會遺失。因此測量法指出取樣的材料，面積或體積不要過大或過小。

在特定的 **SUPPORT** 收集的資料是使用在預測未取樣位置的相同屬性的值，則預測指出位置同樣去有那種 **SUPPORT**，除非程序如區塊或空間平均是使用在關聯大面積或體積的觀測值。這些全部已知的程序稱為升級(upscaling)的程序。最簡單升級程序是收集許多由圍繞在輪廓分明區域的樣點的幾何位置所獲得幾個小樣品組成大量的樣品，及在研究室分析前將大量的樣品均質化。例如收集 10 個次級樣品並混合在 10\*10 公尺面積，**SUPPORT** 將成為一具有相同尺度的正方形。

當小範圍的空間屬性變動的問題很大到大範圍變數無法清楚辨別時，由大量的樣品擴大 **SUPPORT** 是合理的。當由不同方法收集的資料如由遙測感測器所收集的土壤或是水的資訊需要被合併時，它也是有用的。多數遙感探測掃描器所收集的資料是記錄成單一網格的值



(像元的值)，像元大小可由幾公分到幾公里。記錄的數值是由面積權重平均接收的輻射值，所以像元面積定義 SUPPORT 的大小。然而地面的量測，通常由大量的小位置獲得，並可察覺像元內的變化，除非是適當的體積。假如觀測值的設定與 SUPPORT 大小沒配合好，由模式或空間分析去結合設定的資料將很困難(Burrough 1991a)。

由人口普查區域或郵政區域的人口統計學的研究，測定的 SUPPORT 通常是不規則形狀的面積。這些大小及形狀(有時為時間上)的變化，造成內差的困難。通常，資料的收集並不是量測單一的屬性如海拔或是大氣壓力，但人口數量或是社經指標是由類別或是等級尺度表示。這些種類的資料的內插可能適合由不同種類 SUPPORT(郵政號碼或是人口普查區域)使資料變為普通的基礎。

## 專用術語

我們常使用以下遍布本書的點位置取樣的資料的專業術語。點資料的屬性值由  $z(x_i)$  表示，下標  $i$  是指可能量測的數量( $n$ )之一，是在地理上提及的  $x$  座標上任何方便的笛卡兒網格。在未取樣地區的預測值是由  $\hat{z}(x_0)$  表示。

### 準確或不準確的內插

由取樣點預測值的屬性和量測的屬性值相同的內差方法為準確的內插。這是一個完美的情形，因為在討論時，它僅是我們所直接知道的屬性的點的資料。所有其他方法為不準確的內插。統計差異(絕對及正確)在量測及預測值的點資料值的  $\hat{z}(x_0) - z(x_0)$  中通常使用來作為量化不準確的外插方法的指標。

## 空間內插推估的方法

本章所討論的空間內插方法包括：

- **全球推估方法**
  - 使用外部資訊分類
  - 地理座標軸上的趨近面
  - 替代屬性的回歸模式
  - 在空間分析的方法
- **區域量測方法**
  - 徐昇多邊形及 pyenophylactic 方法
  - 線性及倒轉距離權重
  - thin plate 曲線規法

這些方法比較簡單，僅需瞭解測量或取樣統計方法。這些方法通常包含在商業上的 GIS，雖然操作方法很難取得。

地理統計學的方法使用空間自動相關的方法，所知如'kriging'，需瞭解統計上空間自動相關的原理。這些方法使用在不規則的屬性變化，及上述的取樣密度，簡單的內插方法可能造成不可靠的預測。地理統計的方法提供內插品質評估的可能。同樣地，它可能獲得大於 SUPPORT 的土地區塊的預測。另外，地理統計學的方法允許指標函數的內插和可以合併軟體(soft)資料引導內插，因此增強預測的結果。

一些 GIS 包括簡單的地理統計方法但是通常有範圍限制，它最好釋輸出資料到專門的套

裝地理統計學軟體。因為理論背景，內插的地理統計方法分別由第六章描述。

## 資料設定的例子

為提供本章及第六章所有不同內插方法使用在由稀少族群樣品內插的清楚對照，我們使用單一設定的資料(見圖 5.1a 及附錄 3)。在本章及第六章，我們使用由位於接近荷蘭南方 Stein 村莊的 River Meuse 河流氾濫區較大的研究區域中分離出 5\*2 公里、具 0-20 公分表土的沖積土壤，大於 155 個土壤取樣的觀測值所取得的 98 個觀測值。所有的點資料參考 10\*10 公尺的 support，地區內大量的取樣是由分層逢機取樣獲得。取樣地位的海拔及到河流的距離也被記錄。樣品是進行重金屬濃度的化學分析(鎘、鋅、lead 及汞)，在這裡我們以 ppm 描述鋅含量以比較解釋所有的內插方法。

## 全球空間內插推估

內插的方法可分為兩群，為全球空間內插及區域空間內插。全球空間內插使用所有可利用的資料提供感興趣全部區域的預測，區域性內插於環繞樣點的小區域進行內插，以確保所得的評估值僅由當地最鄰近地區資料得到，並和區域一樣適合。

全球內插最不是最常使用在直接內插而是在檢查、可能的改變，如由主要趨勢造成全球變化的影響或目前土地分類變化比率，可能指出具有不同得平均值的地區。一但全球影響被注意到，則剩餘的全球變數可以進行局部(區域)的內插。

全球方法通常是簡單的計算及基於標準的統計變異數分析及回歸的想法。分類方法可使用間單有價值的軟體資訊(如土壤型或行政區)劃分面積為可由在這些區域位置內屬性量測的統計數(平均、變方)特徵化。

回歸方法是探討可簡單及可預測屬性間的相關函數。這些方法可僅基於樣點的地理座標(這些例子的方法被稱為趨勢面的分析)或在一個或多個空間變化的屬性的關係(這些例子中，經驗回歸模式常稱為轉換函數)。

如傅立葉轉換及小波方法同樣使用內插(特別是遙測影像分析 Lillesand and Kiefer 1987, Buiten and Clevers 1993)，但是它們需要許多不同階析度的大量資料，在這裡未曾被考慮。

## 使用分類模式做全球預測

當空間資料稀少時，有時方便假設觀測值是由統計上固定的族群取得(如的資料的平均數或變異數在 support 的位置及大小是獨立的)。兩者之一，我們可以決定這些觀測樣品一些空間座標的改變。假如我們選擇前者，我們自動選擇一個類別的方法進行空間預測，意謂變數的空間結構是由外部定義空間單位。假如我們選擇分類，我們可以使用已知的變異數分析方法計算我們的預測值。

由均值多邊形的分類假設單位內較單位間的變數較小，最重要的改變是發生在邊線上。一個概念模型通常使用土壤及地景圖來定義均值的土壤單位、地景單位，生態類型等等，在對象上如土壤製圖單位、河道、集水區、山坡、坡度等，已被認定是可取得其他地景方面資料有用特徵。

最簡單的統計模式是變異數分析模式： $z(x_0) = \mu + \alpha_k + \varepsilon$  5.1

$z$  是屬性位於  $x_0$  位置的值， $\mu$  是  $z$  在主要興趣的平均值， $\alpha_k$  是在  $\mu$  及單位  $k$  的平值間的偏差， $\varepsilon$  是剩餘(pooled within-unit)誤差，有時如所知的干擾。

模式假定每個分級 k 的屬性值是常態分布，理想上，每個 k 具不同的形式。每個分級 k 的屬性平均值是  $\mu + \alpha_k$ ，是由設定的獨立樣品評估，是空間地的獨立的假設。分級內變異數的平均由  $\varepsilon$  取得，並假設  $\varepsilon$  在所有的分級中都是相同的。

相關變異 ( $\sigma^2_w / \sigma^2_t$ )， $\sigma^2_w$  是分類間共變異數， $\sigma^2_t$  是樣品的總變異，是量測分類的真諦。當所有地圖單位包含一個以上的樣點時，兩者都可被量測。較低的相關變異代表較佳的分類。變異數分析設定(如表 5.1)。連續分類的重要性可由常用的統計的 F 測驗進行變動比率的自由度 m，n-k 來測驗。

Box 5.1：平均及變異數的基本原理（原文 p.104-105）

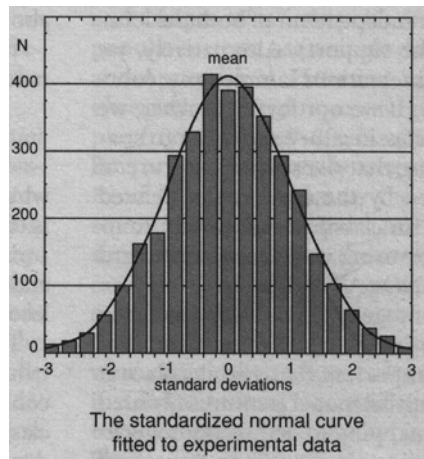


表 5.1：n 個觀測值分成 k 個分類的單尾變異數分析（原文 p.105）

來源	平方和	自由度	平均平方	變動比例
種間	$SS_b$	k-1	$MS_b$	$MS_b/MS_w$
種內	$SS_w$	n-k	$MS_w$	
總數	$SS_t$	n-1	$MS_t$	

### 假設

這種方法是由下列關於空間變數的假設獲得：

- 地圖單位上，z 值的變異值是逢機且沒有空間接觸。
- 所有繪圖單位在多邊形內具有相同的內部分級變異(干擾)。
- 所有屬性為常態分布。
- 所有在邊緣發生的空間變化是急劇的而非逐步的。

注意這些假設不需要一定存在。有些地圖單位或單獨出現在地圖單位可能內部比其他圖具更多變數。假設內部的分級變數是逢機的，暗示差異不可以安排在更大製圖尺度，那通常是不真實的：在大範圍 nested 尺度獲得的土壤圖，在所有尺度上的差異是可見的。資料並不一定需常態分布，或許是近常態、矩形、曲線分布等其他的分布方式。在一些例子，計算自然對數的正規化，logit 轉變或其他適合的轉變是必須的(Box 5.2)。當轉變是必須的，假如它具充足的資料，它將更合理去處理每個地圖單位或是每個地圖單位真實的輪廓，如一個分離的實體並計算個體的平均及標準偏差。

以變異數分析方法所得的重金屬鋅水準的空間預測將使用氾濫頻率圖的軟體 t 資訊解釋

成三種分類。一、經常氾濫(連年)，二、每兩至五年氾濫及三、每次氾濫時間少於五年。當鋅是由污染河流沉積所獲得，氾濫頻繁地區具較大濃度的鋅是合理的假設。

圖 5.3 顯示鋅的量測值於單一模式並不接近常態分布。他們可在不同的平均及變異數呈對數常態分布或是多模式 (multimodal) 分布。可以看見影響非常態分布的變異數分析是由非轉換資料及改變成自然對數的資料獲得。

未經轉換的鋅水準

分類的平均及標準偏差是(鋅的單位 ppm)：

	平均	標準偏差
1	769.76	423.17
2	264.97	176.62
3	205.77	105.32

圖 5.2b、c 是設計由每分級中位於氾濫頻率分級的點資料計算平均值顯示氾濫頻率圖及土壤中鋅含量推導圖。因為圖僅顯示每分級中的平均值，位對數轉換及轉換的鋅含量圖是相同的。

由此分析得來的結論是僅有連年氾濫的這些地區土壤中具高水準的鋅。氾濫頻率圖是由土壤中鋅含量區別最常到最不常氾濫的地區，但圖 5.3 顯示，第二及第三種分類具相當大的重疊。每個分級的標準偏差是很大的，因為是非常態分布及 95%的信賴區間(-76.6-1616.0ppm 鋅)包括零，這是不合理的。

Box 5.2：一般資料的轉換 (原文 p.106)

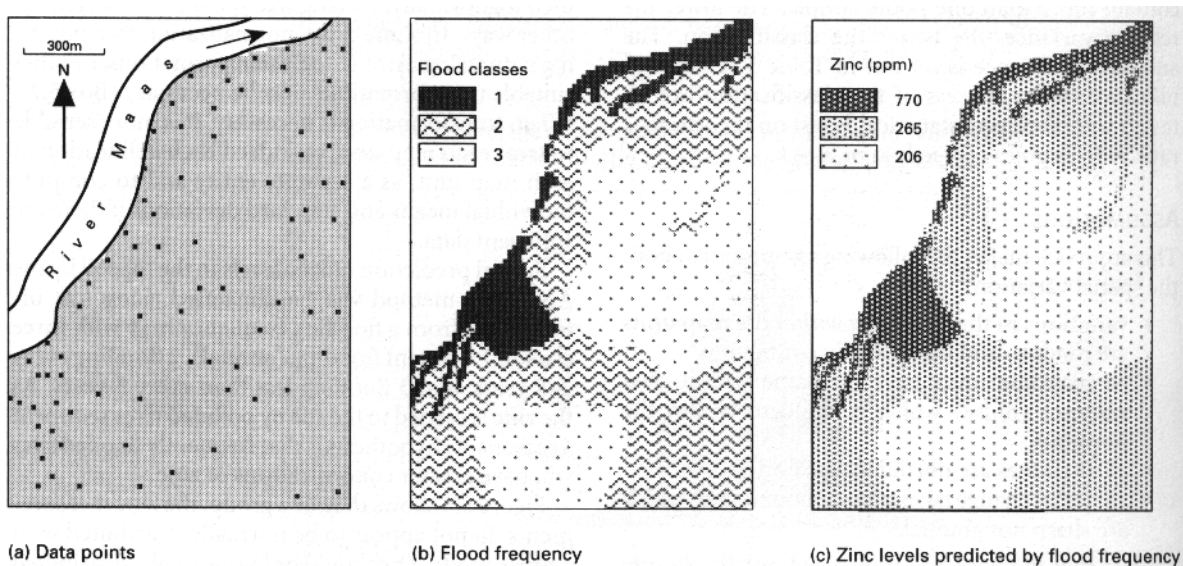


圖 5.2：(a)測驗資料設定的地位位置(b)洪水頻率的分級(c)每個洪水頻率分級的鋅含量 (原文 p.106)

自然對數轉換

分類的平均及標準偏差及經轉換後的平均值：

顯示變數分析

	平均值(取對數)	標準偏差(取對數)	平均(指數)
1	6.484	0.609	654.58
2	5.421	0.531	226.11
3	5.239	0.415	188.48

這些結果顯示對數轉換具極大的變數增加率及增顯分類的品質。資料經取對數後的 95% 的信賴區間不再包括零，因此他們是合理的，但所得標準偏差很大(193.6-2212.8ppm)。在 T 測驗的對數單位平均圖證明沒有統計涵義可以描述分類二及分類三的差異。這些資料一起由直方圖表示(圖 5.3)，結果證實第一種分類與其他兩種具相當重疊的分類是具較大差異。結果建議我們可以使用由第二種分類及第三種分類的氾濫分類較少的資料來修改軟體資料，使兩種分類具不同平均及變方(氾濫頻繁及不頻繁者)。簡單的分類顯示氾濫(每年)頻率對污染的沉積物是具很大的重要性(cf. Middelkoop 1997)。

表 5.2：氾濫分級的鋅變異數分析 (原文 p.108)

來源	自由度	Mean square	變異比率	相關變異
種間	2	3206979	33.8	0.60
種內	95	94880		
總體	97	159048		

表 5.3：自然對數轉換的鋅的變異數分析 (原文 p.108)

來源	自由度	Mean square	變異比率	相關變異
種間	2	14.55	46.6	0.52
種內	95	0.312		
總體	97	0.6056		

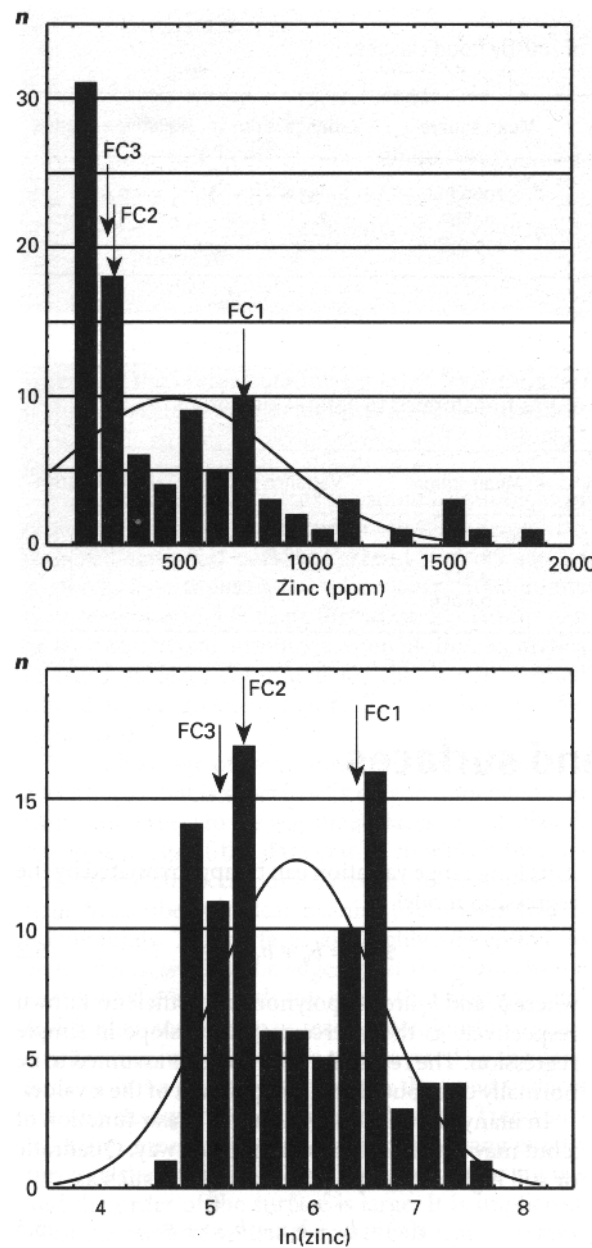


圖 5.3：(上)鋅水準的直方圖(下)鋅(對數轉換後)的直方圖（原文 p.107）

### 使用趨勢面作全球推估

當屬性的變異發生於連續的地景上，它或許可由平滑的數學表面模擬。有許多的方法可以做到：所有屬性適合某些多項式的形式來觀察點資料，因此未取樣位置的值可由座標值計算。

最簡單的方法是由複回歸模擬大範圍空間變異相對於地理位置的屬性值。這想法是推估多項式的線或是面的資料，根據一個或是兩個方向的資料，由點資料的最小平方由此  $\hat{z}(x_i) - z(x_i)$  得到最小平方和。它假設空間座標(x,y)是獨立的變數，z 是感興趣的屬性，獨立變數為常態分布。同樣地，它假設位置的回歸誤差是獨立的，雖然通常這不是事實。

如一簡單例子，假設環境屬性 z 所得的值是由  $x_1, x_2$  到  $x_n$  的點量測而來。假如部分主要變數，z 值是隨位置呈線性增加，x 的大範圍變數可由下列回歸模式推估近似值：

$$z(x) = b_0 + b_1x + \epsilon \tag{5.2}$$

$b_0$  及  $b_1$  是簡單回歸中已知的多項式係數，代表截距及形狀。殘差(干擾)是假設為常態分布及  $x$  值是獨立的。

許多情形下  $z$  並非為  $x$  的線性函數，但是在更複雜的方法則可能會是。二次或更高階的多項式的模式如：

$$z(x) = b_0 + b_1x + b_2x^2 + \varepsilon \tag{5.3}$$

可以被使用。由項數數量的增加，它可能由精確地複雜曲線推估任何點的設定，因此而減少  $\varepsilon$  至 0。由複回歸在  $x, y$  座標取得的二次多項式是表面的形式

$$f\{(x,y)\} = \sum_{r+s \leq p} (b_{rs}, x^r, y^s) \tag{5.4}$$

最先的三種形式是

$$b_0 \qquad \text{單一值} \qquad 5.5$$

$$b_0 + b_1x + b_2y \qquad \text{線性} \qquad 5.6$$

$$b_0 + b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 \qquad \text{二次式} \qquad 5.7$$

整數  $p$  是趨勢面的維度。通常  $p = (p+1)(p+2)/2$  係數，通常接近至最小。

$$\sum_{i=1}^n \{z(x_i) - f(x_i)\}^2 \tag{5.8}$$

$x$  是  $(x,y)$  的向量記錄。所以水平面為零次式，斜面為一次式，二度空間表面為二次式，立體表面為具 10 個參數的三次式。

可發現  $b_i$  係數是複回歸的標準問題，可簡單由標準統計軟體計算。趨勢面可由量測固定網格上所有點的  $z(x)$  值展示。輪廓線的演算法可以用來和繪圖員繪製圖比較。如圖 5.4 是未經自然對數轉換計算的鋅資料所繪製的趨勢面(見 box5.2 的解釋)。

趨勢面分析的優點在它是一個很淺顯易懂的技術，至少包含關於表面計算的方法。廣泛的資料特徵可由低階趨勢面來模擬，但是這會增加在複雜、更高的多項式在描述自然含義的困難。面是具高敏感的邊緣影響，波動邊緣去推估面積中心的點，產生二階和更高表面可能剛好達到在包含資料的面積之外的大或小的荒謬值。

因為這是一般的內插，趨勢表面對 outliers 的資料非常敏感。趨勢面是一個平滑的函數，很少正確通過原始的資料點，除非它們是少數或表面的階數是很大的。毫無疑問的在複回歸下，回歸線或面的殘差是常態分布的獨立誤差；事實上一些趨勢面分析使用最有效之一是已經顯露部分的研究領域，顯示一般趨勢的最大誤差(Burrough *et al.* 1977; Davis 1986)。後來趨勢面最主要的使用，並不是在區域中的內插，而是像是先移動資料概略的特徵來使用一些其他區域的內插。

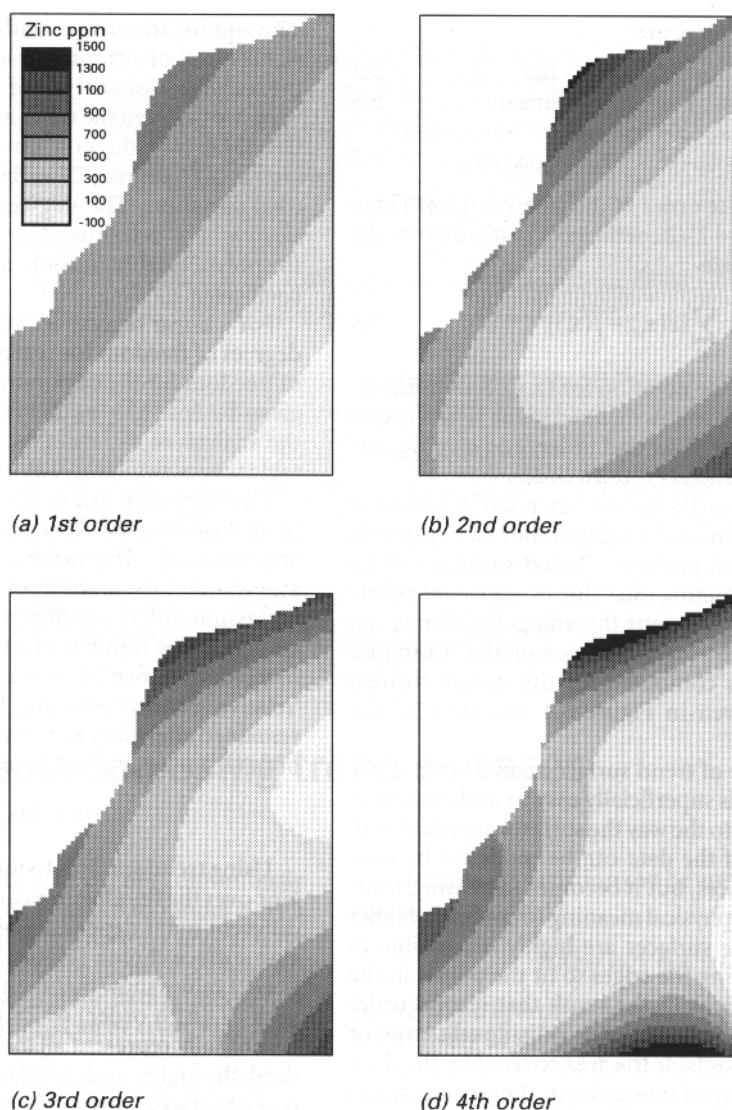


圖 5.4：未經轉換的鋅的全球趨勢面的樣品（原文 p.110）

趨勢面的重要性

統計學對趨勢面的重要性可以由變異數分析的技術測驗，以劃分割趨勢及趨勢誤差的變異。當觀察數為  $n$  時，總變異數的聯合變異的自由度為  $(n-1)$ 。回歸的自由度  $m$  是由多項式回歸式項數的數目測定，除係數  $b_0$  之外。

由一個線性回歸， $z(x,y)=b_0+b_1x+b_2y$ ，回歸的自由度  $m=2$ 。回歸的誤差自由度為  $(n-1)-m$ 。表 5.4 代表變方分析的關係，讀者需注意它和表 5.1 的分類相似。

表 5.4：線性回歸的變異數分析項數（原文 p.110）

來源	平方和	自由度	平均平方	變動比例
種間	$SS_r$	$m$	$MS_r$	$MS_r/MS_d$
種內	$SS_d$	$n-m-1$	$MS_d$	
總數	$SS_t$	$n-1$	$MS_t$	

回歸線或是面在一般的變方分析方法可以假設為類比的分級。變動的比率或是 F 測驗，評估不管變異總數是由從隨機的族群中相同數量具相同自由度的地位導出明顯差異的回歸。



如”相關變異”估計多少變數在分類之後還遺留下來，所以回歸的適合度(R<sup>2</sup>)可以計算如：

$$R^2 = MS_d / MS_t \quad 5.9$$

使用趨勢面模擬在洪水區鋅的濃度如連續面產生不同圖是依據回歸面選擇的階數。適合度 R<sup>2</sup> 值顯示甚至是高階面包括係數 21(五階)或 28(六階)，並不能完全顯示資料的變異性。當然，高階面(第四、第五、第六-最後兩個沒顯示) 在東南方角落地區預測鋅的負值，很清楚的嚴重扭曲事實。

階數	1	2	3	4	5	6
適合度	.183	.475	.560	.687	.767	.802

雖然適合度值隨回歸階數改善，首先發現的是我們對於是否增加多項式的階數就可以明顯增加資料的適合度沒有判斷的方法。明顯的改善同樣可以使用變異數分析估計，如 Box 5.3。雖然明顯較佳的推估可以包含在更高階的多項式，它並不是合理自然的去選擇具自然解釋的趨勢面。

### 使用便宜的全球回歸進行空間推測，以獲可量測的屬性

圖 5.2 及 5.4 的考慮顯示資料設定的例子，土壤的鋅含量及到河流的距離之間有清楚地地理關係。由其他研究(e.g. Leenaers et al. 1989a, 1989b)了解氾濫區土壤的重金屬污染物是由幾個因素造成，最重要的是與來源的距離(如河流)及氾濫區的海拔高。在這區域，粗的污染沉積物是和河堤的沙一起沉澱，細緻的沉積物則沉澱在氾濫頻繁和長期氾濫的低窪地區。因為至河流的距離及氾濫區的海拔高屬性的製圖便宜，因此我們可能可以發展鋅含量的空間預測，假如我們可以取得鋅及這些獨立變數的經驗模式。

Box 5.3：在趨勢面使用高度多項式重要意義的評估 (原文 p.111)

這種回歸模式的形式

$$z(x) = b_0 + b_1P_1 + b_2P_2 + \epsilon \quad 5.10$$

b<sub>0</sub>, . . . b<sub>n</sub> 為回歸係數及 P<sub>1</sub>, . . . P<sub>n</sub> 為獨立變數。在這例子，P<sub>1</sub> 是至河流的距離，而 P<sub>2</sub> 是海拔。

圖 5.5 表示獨立變數至河流的距離之直方圖(河流距離單位為公尺)，氾濫區海拔(海拔單位為公尺)，及非獨立鋅的濃度(ppm)：河流距離及鋅的對數直方圖也被包括。檢視圖 5.5 指出在複回歸下最好使用對數轉換後的變數。在多數的標準統計軟體中，複回歸是的方法最有價值的。

由 98 個鋅的資料點的例子，ln(鋅)的複回歸對照海拔及 ln(至河流距離)具下列關係：

$$\ln(\text{鋅}) = 10.000 - 0.292(\text{海拔}) - 0.333 * \ln(\text{至河流距離}) + 0.394(\text{resid 誤差}) \quad 5.11$$

$$R^2 = 0.749$$

像這種模式通常稱為轉換方程式(Bouma and Bregt 1989)；它可由多數的 GIS 簡單進行計算(見第七章)。製圖來源包含數種方法或許為向量或網格形式(見第七章、第八章)。圖 5.6 顯示結果。

相同的程序可能使用許多其他獨立及非獨立變數的設定，如溫度及海拔，至海邊的距離的雨量，植被組成如過剩濕氣函數，顧客數量及收入水準等。地理座標及有關聯的屬性可以

包含在一個回歸中，使用許多這些可能由資料得來的資訊。最重要的一點是認為回歸模式建立自然的感覺。注意所有回歸的轉換模式是不準確的內插。

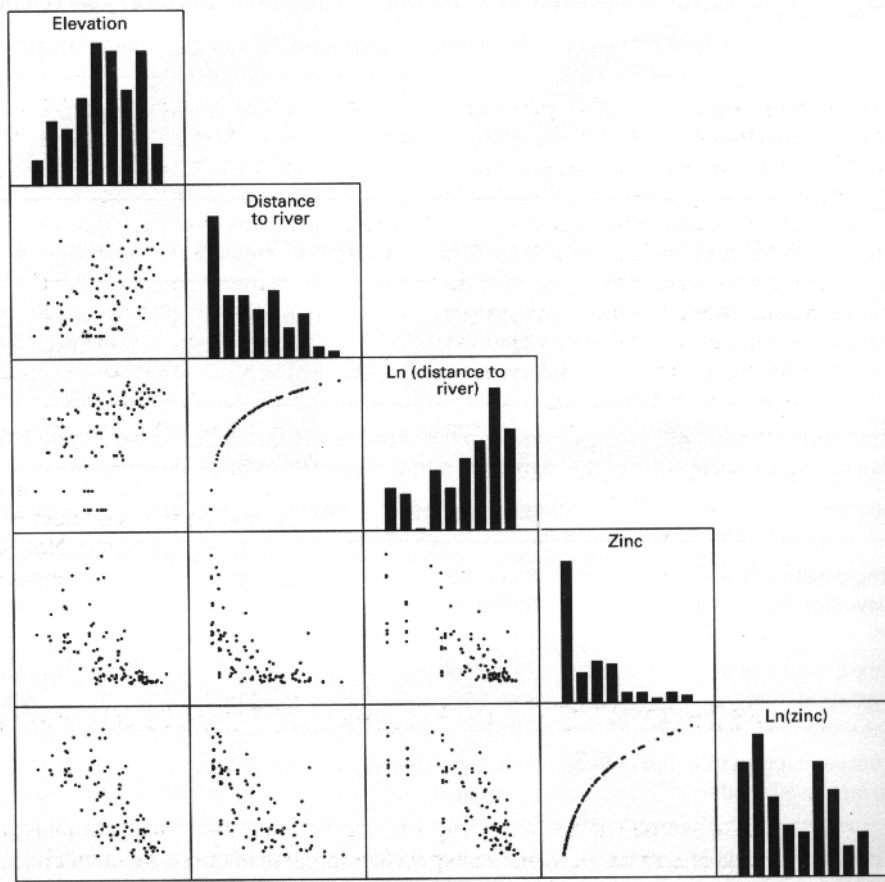
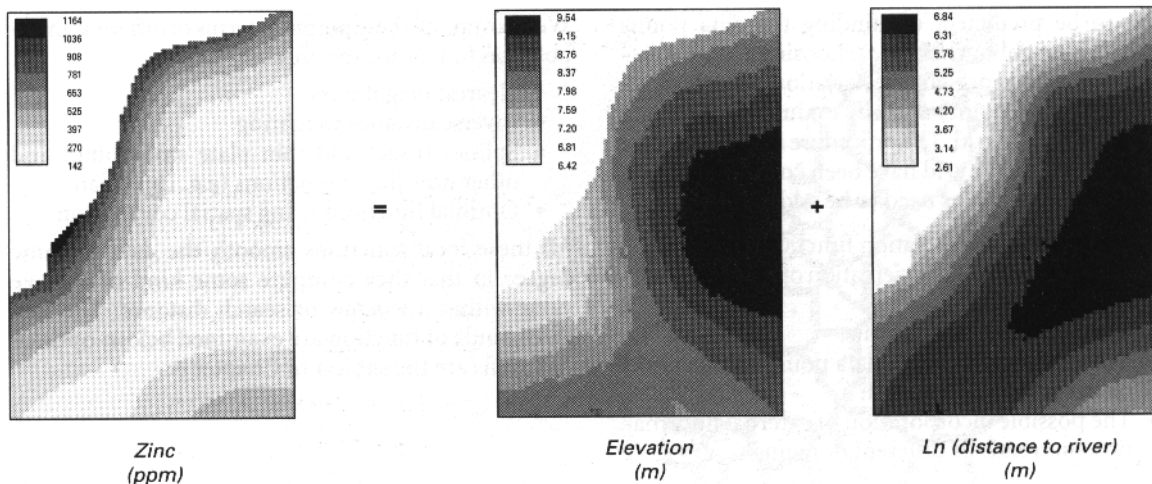


圖 5.5：鋅至河流距離及氾濫區海拔高回歸的非獨立及獨立項數的分布直方圖（原文 p.112）



$$Zinc = \exp(10.000 - 0.292 \text{ elevation} - 0.333 \ln(\text{distance to river}))$$

圖 5.6：由氾濫區海拔級至河流距離圖籍由回歸計算鋅水準的結果（原文 p.113）

### 空間插入推估的地方性、確定性

目前為止所有的表示方法已經再內插時加上外部、地球空間結構。在所有短期例子，區域變數已經不考慮逢機、非結構性干擾。直覺上，這對於預測不可見的點的值變成與一旁量測的值相似是不合理。因此人們尋找區域性內插的方法，使用最鄰近點的資料直接的資訊。這種方法，內插會關聯到(a)定義一個研究區或是在預測點周圍的鄰近地區。(b)找出鄰近地

區內點的資料(c)在有限數量的點之中選擇數學函數表示在變數上，點數量的限制。(d)評估這些在規則網格上的點。重複這程序直至所有在格子上的點都被計算過。

以下的問題需要被滿足：

- 使用的內插函數種類
- 鄰近地區的大小、形狀及方位
- 點資料的數量
- 點的資料分布：規則網格或是不規則分布
- 轉折或不同區域的外形資訊合併的可能

我們以不同項數的內插函數檢查這些點，特別是：

- 最鄰近地區
- 距離的倒數權重
- 曲線規法(精確及薄的金屬(thin plate)平滑)及其他非線性函數(如 Laplacian)
- 使用空間 covariation 的最適函數。

所有這些區域函數會平滑這些在電腦視窗或調查距離計算某些平均值資料到某種程度。最前面的三種函數在下面解釋，最適函數是第六章的主題。

### 最鄰近資料推估法：徐昇多邊形法

徐昇(又稱為 Dirichlet 或 Voronoi)多邊形以空間預測的分類模式預測由最接近的單一點資料提供非取樣地區的屬性到極端值。徐昇多邊形法劃分地區推估，是由每個單元觀測的點的資料構成完整測量的方法。假如資料是位於規律的正方格，則徐昇多邊形相同的邊到網格空間是全部相等，規則的單元和；假如資料是不規則間隔，則產生不規則的多邊形格子 (Figure 5.7)。線結合點的資料顯示為 Delaunay 三角，是一個相同位相如 TIN(第三章)。

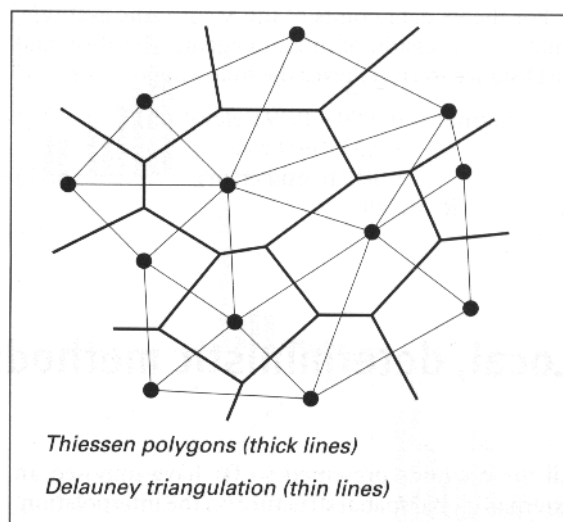


圖 5.7：徐昇多邊形網及相等的 Delaunay 三角形的例子 (原文 p.114)

徐昇多邊形通常使用在 GIS 及地理分析上，如由點資料關聯到空間的快速方法；一個常見但是固定的徐昇多邊形的使用是假設任何地點的氣象資料可由最鄰近的氣象觀測站獲得。除非有很多觀測值(通常沒有)，這種假設並非真的很適合逐步變化的現象如降水及溫度和氣壓，因為(a)最終地圖的形式是由觀測值的分布測定，及(b)維持地誌圖的方法是想像邊緣內

及所有在邊緣的改變是均質的。每個方格僅有一個觀測值，則沒有方法估計方格內變化，除取得重複的觀測值外。

徐昇多邊形的優點是它可以簡單的使用定性的資料，如植被分類或土地利用型，假如你所需的都是地誌圖及沒有注意邊界的幾何格局。因為在點的資料內所有預測值是相等的，徐昇多邊形是可以精確預測。

圖 5.8a-c 說明徐昇多邊形使用在氾濫區土壤中鋅含量的圖籍。由外部定義的單元取代的計算的平均值，鋅在氾濫區的空間變化簡單的由量測值到它最接近鄰近單元指定。(圖 5.8b, c)。徐昇多邊形圖建議鋅圖籍包含由氾濫頻率的多邊形掩飾想像中鋅水準的空間變異，特別是在氾濫頻率為 2 及 3 等級之內(比較圖 5.2c)。

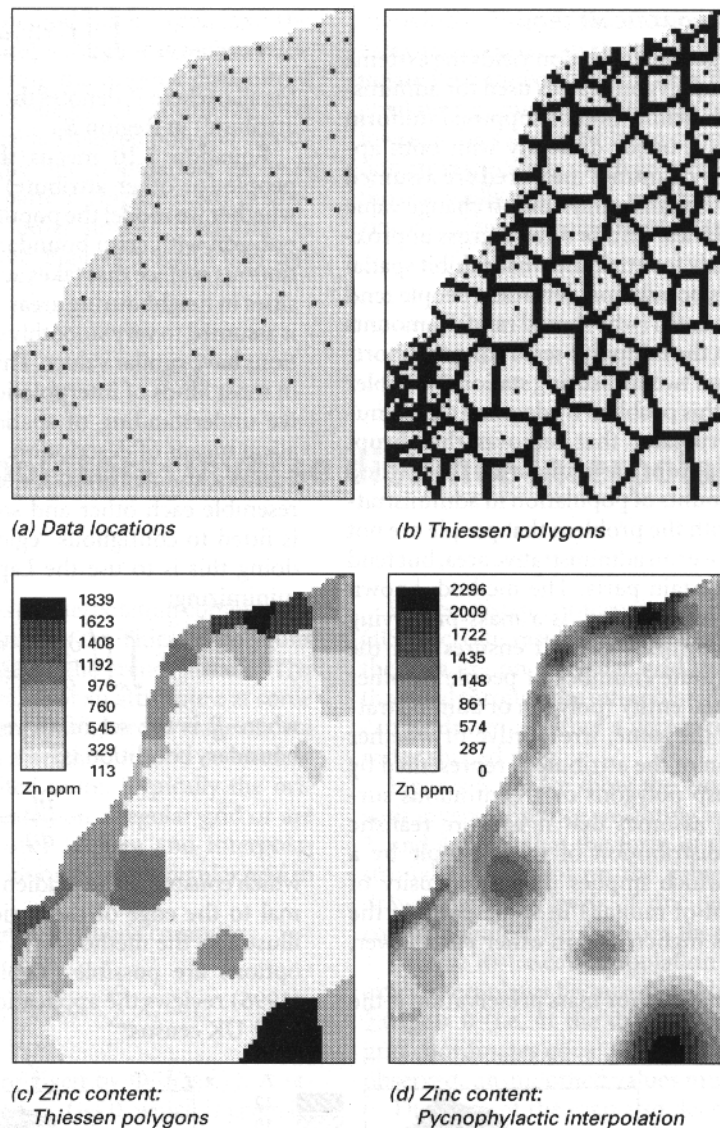


圖 5.8：(a)資料位置(b)徐昇多邊形(c)鋅含量：徐昇多邊形(d)鋅含量：由徐昇多邊形的鋅 pycnophylactic 內插（原文 p.115）

## TOBLER'S PYCNOHYLACTIC METHOD

徐昇多邊形內插產生建立分離的多邊形圖最大框(case)如使用管理單位或是土地利用型相同面積輪廓線值。兩種方法主要困難是定量的值是假定單位內是均質的，僅邊界值會改變。它通常是一個概略的近似值，特別是顯示空間連續事物的屬性值，如人口密度(人口聚

集趨勢)或降水(為什麼降雨量改變在兩測站最短距離區分的中點是突然的?)。Tobler (1997) 由設計連續及平滑內插，移動突然變化的不相似邊界滿足這個問題。它的方法是設計在行政區的人口數量以應付人口不能均勻散佈至行政區但趨於而集中在某一部分。這種方法，如同已知的 pycnophylactic 內插，是主要資料真正位置的數量保存。它保證在空間實體(多邊形及行政區)大量的屬性(人口數量及其他屬性)保持相同，不論是地球屬性變化是由均質、波動的多邊形或是連續表面所代表。這方法假設它更實際表現平滑面的人口分布，暗示著人口密度(或降雨量)就平均值而言，某些地區被高估、某些地區被低估。

整體保留的主要狀況是微分的情形：

$$\int_R \int f(x, y) dx dy = V_i \tag{5.10}$$

所有的  $i$ ，由  $V_i$  表示在  $R_i$  地區的值(人口、數量、屬性)。

5.10 方程式是指每個多邊形人口的總量(或其他屬性)是不變的，不管是在任何我們是由具波動或平滑邊界的單一多邊形模擬的人口數量；連續面在根據鄰近區域取得的人口密度。它強迫面是被假設為非常的平滑，因此鄰近位置具相同的值。這種假設常見於多種的內插形式中，因為它結合直覺上可了解的空間變數藉由數學方法描述。除非是自然的邊界，否則鄰近地區的密度趨於彼此類似，因此一個節點或平滑面是由連續區域推估。最簡單的方法是使用 Laplacian condition，如藉由最小值。

$$\iint_R \left( \frac{\partial f^2}{\partial x} + \frac{\partial f^2}{\partial y} \right) dx dy \tag{5.13}$$

$R$  是所有地區的常數。最普遍的邊界狀況是

$$\frac{\partial f}{\partial \eta} = 0 \tag{5.14}$$

以限制推估表面類型的坡度使地區邊界更平緩。

圖 5.9 以一個簡單例子解釋這種方法。其他選擇是可能的(Tobler 1979, 1995)；Martin (1996)回顧此種方法應用在 1991 年的香港的人口普查。

Pycnophylactic 方法可以應用點屬性的鋅或其他性質取樣資料，如降水量，由原始資料轉換至密度函數。這是藉由計算鋅在每個徐昇多邊形的總含量以獲得的每個多邊形鋅族群'population'近似的總量。產生的格局(圖 5.8d)和由其他平滑的內插(見圖 5.2、圖 5.3 及圖 5.9 及 6.6)所獲得的是相似的，雖然是連續的，但是明顯的並不是一個正確的內插，感覺上由單一網格內插的值和最初在相同單位所量測值不同。由 pycnophylactic 內插所獲得的最高及最低值和實際量測的值相比，各自是差別是很大及很小，這些取樣的資料如鋅可以帶來一些內差的問題。

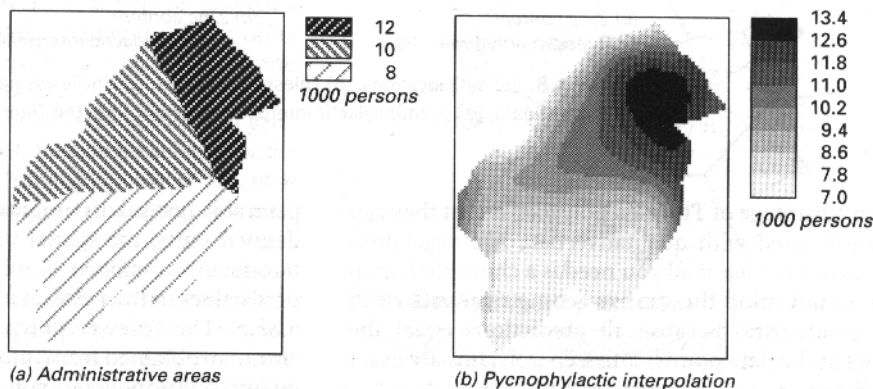


圖 5.9：人口數的曲線是由(a)相似的行政區(b) pycnophylactic 內插展示（原文 p.116）

### 線性插入推估：距離倒數插入推估

距離倒數的插入推估方法結合接近徐昇多邊形學說及趨勢面的逐漸變化的想法。這假設是在某些未看見的點的屬性 z 值，是一個發生在未見點周圍鄰近地區或視窗的點的資料的距離權重平均。典型的原始點的資料為了製圖是位於規則的網格或是分布在不規則的地區上，為了製圖內插值是由密集的規則網格位置獲得。加權移動平均方法計算

$$\hat{z}(x_0) = \sum_{i=1}^n \lambda_i \cdot z(x_i) \sum_{i=1}^n \lambda_i = 1 \tag{5.15}$$

權重  $\lambda_i$  是由  $\Phi(d(x, x_i))$  獲得。有一個需要是  $\Phi(d)$  的 d 趨近於 0，是由倒數或是負的指數函數  $d^{-r}$ ,  $e^{-d}$ ,  $e^{-d^2}$  所得到。最常見的  $\Phi(d)$  形式是倒數距離權重預測，其形式如下：

$$\hat{z}(x_0) = \sum_{i=1}^n z(x_i) \cdot \frac{d_{ij}^{-r}}{\sum_{i=1}^n d_{ij}^{-r}} \tag{5.16}$$

$X_j$  是表面要進行內差的點， $x_i$  是一個資料點。最簡單的形式稱為線性內插，權重是由設定的點資料及預測的點距離的線性函數計算。

距離倒數內插通常使用在 GIS 中以建立覆蓋在點資料上的網格。一但資料在規則網格上，輪廓線可以通過內插值及圖可以被描繪成網格或是向量圖。

圖 5.10a, b, c 顯示鋅內差的資料由距離倒數權重所得到的值為 r 或 1, 2, 4。注意這些產生圖籍的形式是根據 r 值而來。結論推斷鋅污染的水準，在 r=1 及完全不同於 r=4 時。距離倒數內插強迫為精確的內插因為當  $\sum(d_{ij})=0$  (如點資料) 它會無限擴張，因此當輸出的網格座標等於這些觀測的取樣點時，非平滑的值必須重複複製。

地圖的形式同樣根據聚集資料或是存在的 outliers。當值和其他環境非常不同時，距離倒數內插通常具有一個‘鴨蛋’格局圍繞在單獨資料點，雖然可由更改資料點的搜尋標準修改精確至某程度，這方法並沒有內建的測驗方法來預測品質，所以圖籍的品質僅可由取得精確的觀測值評估。注意這些必須如原始觀測值性質 (support)，儘管由倒數距離內插來平滑鋅的比例到 r 值，這是真實的內插解答及沒有取樣的性質 (support)。

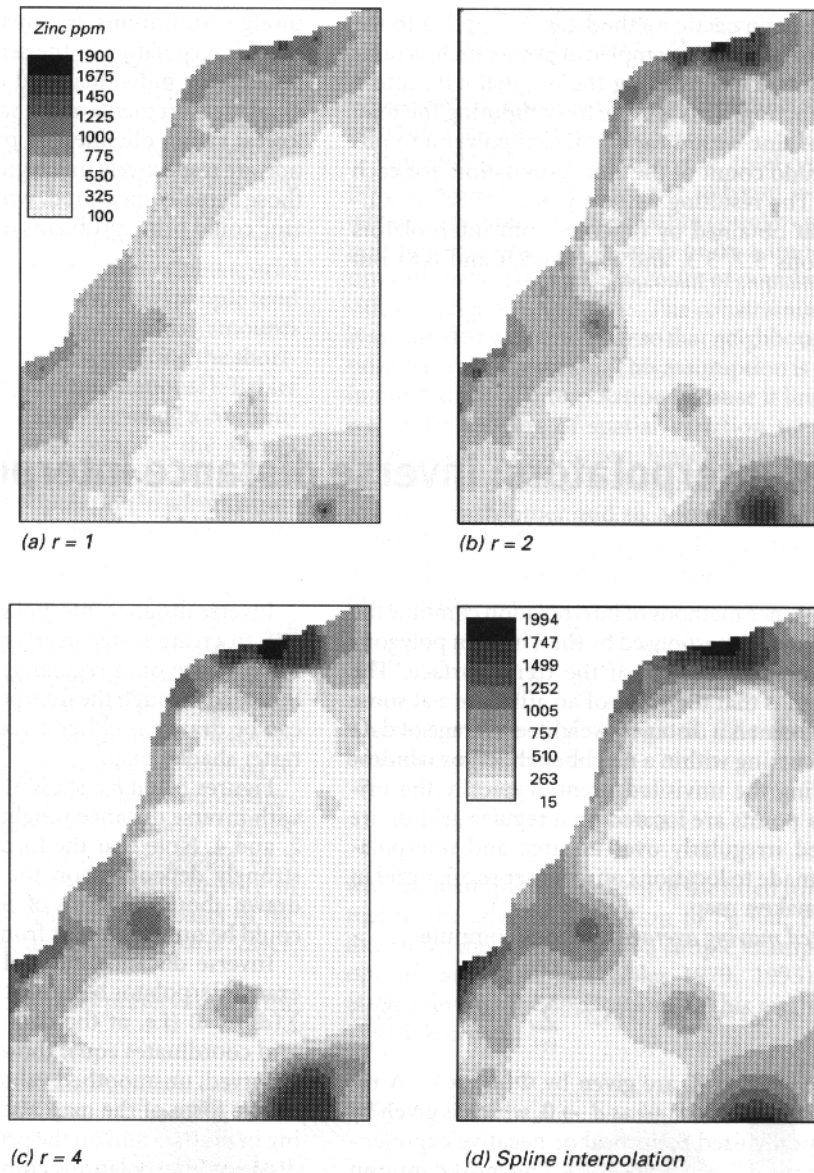


圖 5.10：(a-c)距離反比內插顯示參數權重影響的結果(d)曲線規法的鋅製圖（原文 p.118）

### 曲線規法

在電腦使用在推估的曲線資料點以前，製圖者使用可彎曲的尺由眼睛來達到最佳的區域推估的平滑曲線。可彎曲的尺稱為曲線尺。當在描繪線段時，可彎曲的尺由重量適當的保持固定資料點的位置(Pavlidis 1982)。近代相同產品是塑膠外膜的可彎曲的尺，在辦公用品商店販賣。它可以顯示線段沿曲線尺描繪是接近三次多項式，是一個連續且有連續第一及第二導出。

易彎曲的尺函數曲線尺的數學式。它們是 piece-wise 函數，是說它們推估較小數量的精確點資料推估，當在相同時間可以確定接合點在曲線部分或其他部分是連續的。這表示曲線尺，可能修改部分曲線而沒有全部再計算，它不可能具有趨勢面(Figure 5.10)。一般定義 piece-wise polynomial 函數  $p(x)$  是：

$$\begin{aligned}
 p(x) &= p_i x & x_i < x < x_{i+1} \\
 p^j(x_i) &= p_{i+1}^j(x_i) \quad j=0, 1, 2, \dots, r-1; \\
 i &= 1, 2, \dots, k-1
 \end{aligned}$$

點  $x_1, \dots, x_{k-1}$  區分  $x_0$  到間隔到  $k$  次間距稱為斷點曲線點在這些  $x$  值通常稱為結點 (Pavlidis 1982)。函數  $p_i(x)$  是  $m$  或更小次方的多項式。項數  $r$  是使用曲線尺限制的節點。當  $r=0$ ，函數就沒有限制，當  $r=1$  時，函數是連續沒有任何導函數的限制，假如  $r=m+1$ ， $x_0$  到間隔可由單一的多項式表示。所以  $r=m$  是強制的最大數量的解釋導致一個 piece-wise 結果。 $m=1, 2, 3$ ，表示曲線為線性、二次式或立方。導函數是  $1, 2, m-1$  階，因此二次方程式曲線規法於每各節點上具一個連續性導函數。立方曲線規法在每個節點上必須具有兩個連續性導函數。如簡單的曲線規法，當  $r=m$  時，僅有  $k+m$  個自由度。這個例子的  $r=m=3$  具特殊重要性，因為曲線規法的項數首先使用在 piece-wise 多項式函數。雙立方曲線規法是用在當表面由需內插的線取代時三度空間的情形。

因為簡單的曲線規法計算在一個大範圍的分離的次級間距(sub-intervals)是困難的，像是可能具有數位化線段的例子，多數實際上的應用是使用特別的曲線尺，稱為 B-曲線尺。B-曲線尺是值定義在我們有興趣的間距為 0 的其他曲線尺的總和(Pavlidis 1982)。因此 B-曲線尺容許低階多項式的簡單方法的區域推估。

B-曲線尺通常使用在展示平滑的數位線，如土壤的邊界或製圖協會預期平滑、流動線的地理圖。Pavlidis 注意到複雜的形狀如發生在文字字體上，B-曲線尺所定義的項數可能比點資料的設定更具經濟性。B-曲線尺使用在平滑多邊形邊界，會導致某些混亂，特別是計算週長或是面積時。假如多邊形面積是由使用梯形數位化的點資料計算而來(Box 3.3)，它將不同於 B-曲線尺平滑邊緣所產生的面積。當高階的 B-曲線尺使用於平滑的彎曲邊緣包括急遽的、矩形的轉角，許多問題可能會出現(Figure 5.11b)。

使用曲線尺內插的問題是不管應該選擇斷點來符合點的資料或是它們的內插。曲線規法由內插產生的差異或許是由兩種方法來產生(Figure 5.11c)。必須注意的是最大及最小的曲線規法並不是必需發生在點資料上。

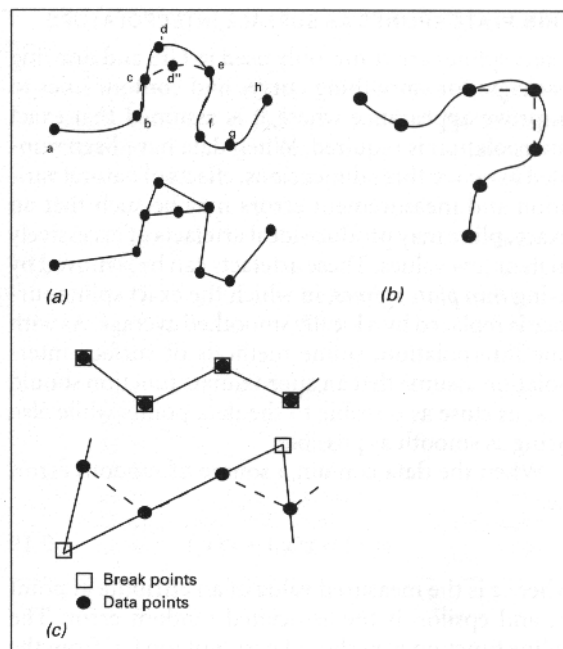


圖 5.11：曲線規法的某些性質(a)位置校正顯示位置的改變(b)精確的曲線圍繞於彎曲處(c)選擇斷點位置在或是導致曲線具有大的影響間的資料點 (原文 p.119)



### THIN PLATE 曲線如表面內插

精確的曲線規法通常使用在 GIS 及繪圖軟體於平滑曲線及輪廓線以改善假設必須精確內插的外貌。當資料是由二維或三維取樣，自然變數及量測誤差的影響在精確的曲線規法可能產生過高或過低的區域製品(artefacts)。這些製品可以使用 THIN PLATE 曲線移動，在精確的曲線尺面是以局部平滑平均替代。如線性內插，曲線規法內差的面假設一個近似函數應該儘可能同時及平滑的穿過點資料。當資料包含逢機誤差來源，如：

$$y(x_i)=z(x_i)+ \epsilon (x_i) \tag{5.19}$$

z 是 x 點屬性的量測值，ε 是聯合逢機誤差。

$$A(f) + \sum_{i=1}^n w_i^2 [f(x_i) - y(x_i)]^2 \tag{5.20}$$

A(f)項代表函數 f 的平滑度，第二項代表資料的'接近'或'精確度'。權重  $w_i^2$  是選擇至誤差變異的倒轉比例；

$$w_i^2 = p / \text{Var}[\epsilon(x_i)] = p / s_i^2 \tag{5.21}$$

P 值反映由使用者對不同平滑曲線尺特性給予的相對重要性。平滑，THIN PLATE 曲線尺常使用在高程的內插以建立數位高程模式當它必須迅速且有效率的內插大面積時(Hutchinson 1995, Mitasova et al. 1995)。正文證明它們在多變數內插的屬性的使用如經度、緯度及海拔高三個變數曲線尺函數內插年平均降水量。

#### 使用鋅資料的曲線規法

圖 5.10d 顯示曲線尺鋅的變化的製圖。注意曲線尺趨勢面描繪取樣地位附近具大量鋅值面積。

#### 曲線規法的優點及缺點

因為曲線規法是使用一個時間上少量點的 piece-wise 函數，內插值可以快速的計算。平滑表面的測驗資料顯示預測值相當接近內插值，假如量測誤差結合資料是小的(Mitasova et al. 1995)。對照趨勢面及平均權重，曲線規法保持小尺度特徵。線性及地面曲線規法指出美觀令人滿意的資料能及迅速產生清楚資料概況。曲線規法的平滑度可以由直接分析幾何表面及位相簡單計算數學導函數(見第八章及圖 5.14a)。線性參數次模式(回歸模式)的 INCORPORATIONC 獲得點簡單性質(SUPPORT)的獨立變數的內插。

某些曲線規法優點已被提出。其他是曲線尺內差的聯合誤差沒有直接量測，因此可能由遞歸的技術如已知的'jack-knifing' (Dubrule 1984)獲得。最重要的優點是 THIN PLATE 曲線可能提供不真實平滑的真實看法；在某些情形，如屬性值評估數值模式，這種特性可以產生令人誤解的結果。

### 簡單全球及區域方法之比較

我們說明數種不同方法的內插技術，但是它們如何比較呢？哪一個可產生最相符的結果，當這方法比其他優先？首先，表 5.5 就所有最大及最小的內插值比較，及量測面積比例在 500, 1000 及 1500ppm 鋅的評估。

表 5.5：決定內差的結果摘要（原文 p.121）

方法	最小值(ppm)	最大值(ppm)	單位面積 >500ppm	單位面積 >1000ppm	單位面積 >1500ppm
氾濫頻率(未轉換)	206	770	11.96	0.00	0.00
趨勢面-第一階	-98	791	38.53	0.00	0.00
趨勢面-第一階	109	1350	31.88	3.37	0.00
趨勢面-第一階	-13	1256	33.26	3.66	0.00
趨勢面-第一階	-100	1500	28.3	5.88	0.69
距離及海拔回歸	142	1164	17.82	0.64	0.00
徐昇多邊形	113	1839	28.11	9.65	5.31
Pynophylactic 方法	0	2296	28.35	10.15	3.87
倒轉距離 r=1	136	1541	28.92	1.16	0.03
倒轉距離 r=2	114	1827	28.82	3.99	0.61
倒轉距離 r=4	113	1839	28.47	7.05	2.49
Thin plate 曲線規法	15	1994	30.44	6.74	1.87

比較的結果我們可以看到高於三階的全球趨勢面是不合理的。因此真實的趨勢面僅對寬廣地理趨勢描述有用(比較 Burrough *et al.* 1977)。Pynophylacticz 方法及 THIN PLATE 曲線延伸比記錄高或比記錄低的最大及最小值，但是在倒轉距離  $d=-2$  得到接近原始值的結果。面積預測的範圍超過三個 arbitrary 限制是由這些方法產生最大超過 1000ppm 的值；其他資料設定沒有獨立資料及實行回顧時，它們通常不能說是最好的方法。倒轉距離及薄的金屬曲線尺似乎可以產生看起來最自然的表面，但是這種方法僅可反映我們計算結果，這些表面總是比實際情形為平滑。

## 數位高程模型為由空間推估連續面的特例

本節測驗數位高程模型(DEM)由一個在 GIS 內許多使用內插、連續面的特例。原來的 DEM 是用來計算地誌圖，但今天則具有需多其他的應用(Box 5.4)，包含‘眼睛抓取(eye-catching)’的能力在土地利用型構成基礎的空間資料展示(如插圖 1.4)。這部分說明 DEM 的建立及如何由電腦模擬、主要獲得的產品如區塊圖示及數位地誌圖。其他產品，如由數學導函數建立的 DEM 是在第八章連續面的空間分析部分說明。

Box 5.4：DEM 的使用（原文 p.122）

## DEM 表現的方法

地區上面的高程變化可以由許多方法模擬。DEM 可以由數學定義的面或是點或線的影像表示。線資料可以使用輪廓線或是剖面來表示及重要特徵如河流、山脊、海岸線及坡度斷點。在 GIS，DEM 可以由固定的網格 (*altitude matrices*)及規則的三角網格(TINs)模擬。兩種形式為內部轉換(inter-convertible)且優先權是根據需取得的分析資料。

高程網格是處理高程面最常使用的形式。原本是由立體航空照片在立體分析如

GESTALT GPM-II (Kelly *et al.* 1977)的定量測量獲得。DEM 附產品是需要建立尺度校正的地誌圖(如下)。高程網格可由規則或不規則空間點資料其他定量資料的相同方法的內插產生。

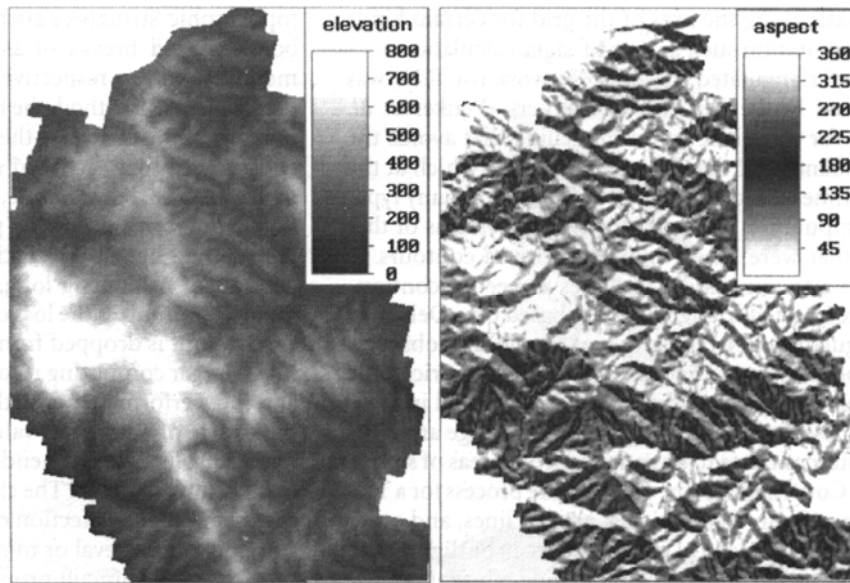


圖 5.12：左：灰階尺度網格(像元大小 30m\*30m)；右：遮蔽坡度圖增加坡度隱蔽輪廓影響  
(原文 p.123)

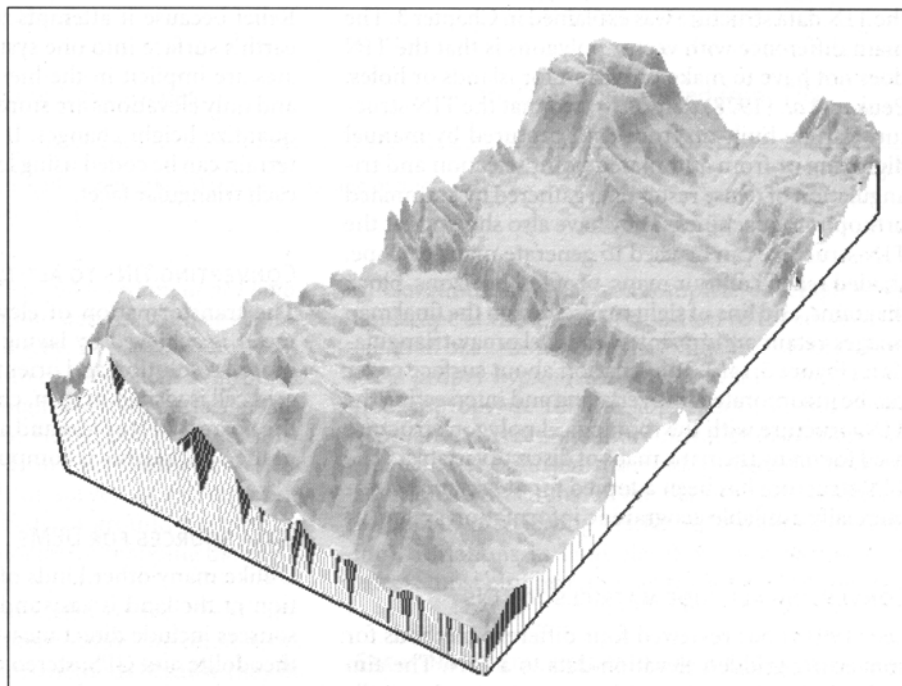


圖 5.13：在區塊直方圖上建立模糊輪廓的坡向資料 (原文 p.123)

因為電腦內可以簡單移動網格，特別是在 GIS 網格中，高程網格變成很有用的 DEM 形式。英國、澳洲、美國及世界上其他地方，完全覆蓋在 1：250000 比例尺的地形圖(網格單元為邊緣 1-5 公里的大小，在網路上是有用)。高解析度的網格是基於 1：50000 或是 1：25000 地圖或航空照片圖，在這些國家或其他國家更有價值。圖 5.12 及圖 5.13 是 DEMs 海拔高網格的例子，圖 5.12 由 30\*30 公尺的網格單元顯示 DEM。海拔高資料或許可以尤須多不同方法展示，圖 5.12 顯示簡單海拔高的灰階尺度影像及坡面表面的灰階尺度影像(第八章顯示如何取得)。由選擇適合的灰階尺度，影像來自北方(上方)的光顯露。因此可獲得立體空

間輪廓的感覺。圖 5.13 顯示其他地區的一個相似影像，但現在坡度的資訊是夾雜在 DEM 中，產生區塊圖示表現自然的輪廓。海拔高網格由許多獲得的土地利用型：如坡度、剖面、光輻射、可見的線及位相面等(第八章的例子)的資訊為起始點。

### 不規則三角形路網

儘管在計算輪廓線、坡度角度及坡向、山的遮蔽、習慣的盆地描述(見第八章)，altitude matrices 是很有用的，規則網格系統並不是沒有缺點。這些缺點包括(a)大量過多剩餘的資料在相同地區面積中(b)再沒有建立網格大小時，不能去適應(adapt)具有不同複雜輪廓的面積。過於強調沿著網格中心的某些計算性質如觀測線(line-of-sight)的建立。

不規則三角形路網(TIN)是由 Peucker and co-workers (Peucker et al. 1978)設計，為了數位高程模式以避免過多的高程網格，在相同時間同樣也可以由許多形式計算(如坡度)，比在那時間上僅基於數位輪廓線的系統更有效率。TIN 是一個地形模式，使用連續、Delaunay 三角形的關聯的三角面在不規則空間節點或觀測點的表單(圖 5.7)。不像規則網格，TIN 允許在輪廓複雜地區加入外部的資訊，將沒有需要在龐大數量的多餘資料由簡單的輪廓聚集。因此，TIN 資料攫取分析可以明確隨著山脊、河流線或其他可以數位化重要位相特徵達到所需的正確性。TINs 提更有效、正確以海拔高資料為主的資料儲存，介紹可能妨礙一些空間分析種類的三角形的處理，如幾何及位相面的導函數。TINs 是由相似的地理向量結構模擬這些在多邊形網的使用，TIN 資料結構再第三章已經說明。和向量多邊形主要的差異是 TIN 並不需要產生分離區塊或是缺陷。Peucker *et al.* (1978)證明 TIN 資料結構可由手動數位化或是自動選點或由自動製圖器獲得的三角密集網格資料等攫取資料來建立。它們同時顯示 TIN 結構可以產生坡度圖、隱蔽輪廓、輪廓線圖、剖面、地平線、區塊圖示及 line of sight 圖，儘管最後圖影像會保留 Delaunay 三角形(圖 5.14b)。關於覆蓋面資訊可由套疊及交叉具位相多邊形結構的 TIN 結構結合，使用在分離變數的主題圖。TIN 結構至少在地理資訊系統具有商業價值。

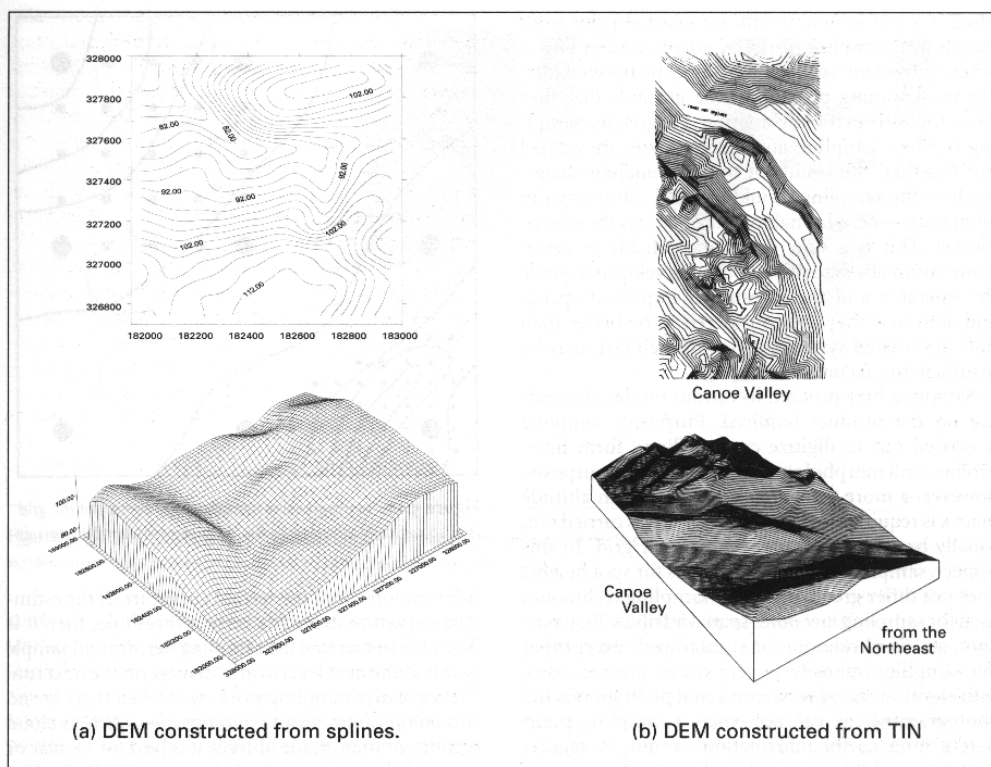


圖 5.14：產生的數位高程模式(a)由曲線規法(b)由三角網格(data courtesy of A.de Roc and T.Poiker) (原文 p.125)

## 轉換 ALTITUDE MATRICES TO TINS

Lee (1991a)檢查四種轉換海拔高資料網格方法至 TIN，針對所見的轉換，可取得不規則空間的高度點最小可能的位置以提供關於地形結構最大資訊如山頂、山脊、山谷底部及坡度斷點。它用三種方法計算(分別為最重要點方法(VIP)、等級改變方法(HT)、啟發式方法(DH))，在長、薄的三角形產生的支出上，HT 方法具非常有效率的資料結構方法。DH 方法是每次點資料繪至網格時，遺失最小資訊，但它需要大量的計算資訊。Lee 發現三種方法確實的功能根據選擇的參數值，所以不會具有應該被使用的最優良的演算法。使用者是根據針對發現的極端值、有效的資料儲存及回覆或遺失最小資訊選擇演匯法；它們的特性很難彼此比較。

近來 Dutton (1996)建議壓縮方法模擬 planetary 輪廓基於球面遞歸的多面體格子至等邊三角形面。這種方法所知如全球輪廓的大地海拔模式(Geodetic Elevation Model)因為它嘗試取得全部地球表面至一個體系。水平座標是隱藏三角固定等級僅有海拔高被儲存。使用單一區塊(flags)到量化高的改變。它要求全部地形可以由每個三角面資料使用小於一位元的編碼。

## 轉換(TINS TO ALTITUDE MATRICES)

海拔高資料由 TIN 轉換至網格是藉由 TIN 上在放置需要分析的網格及定位來完成。每個網格轉換是可見的，檢查最鄰近的 TIN 點及線或面的平均的 TIN 高度已經計算好。

DEMs 資料來源不像其他量化的資料，土地高度是由簡單及便宜來量測。資料來源包含野外經緯儀及 GPS 的直接量測，立體的航空照片、飛機及衛星掃描系統，及紙張圖籍數位輪廓線。由系統的製圖，高度資料由攝影方法(ASPRS 1980- Manual of Photogrammetry)由套疊立體航空照片及衛星影像而來。因為特別的目的，也使用其他種類的掃描器，如航空雷射計 interferometry 量測高精確面(Plates 3.5, 3.6; Hazelhoff, pers. comm.)。

船、潛艇的聲納或氣墊船同樣被使用來調查海床及河床的高度格局，或是由地面穿透雷達及地震儀來繪製次級表面層的高度圖。

假如常有豐富資料，區域、簡單(線性)內插方法通常比複雜的外插方法好，因為它們不需要假設關於空間影響，它的計算快速。當立體航空照片及衛星影像是高度資料來源，我們有影像解析度水準的完整土地利用型覆蓋。DEM 的建立那時是選取資料建立適於應用的合適高度表面，包含幾何修正及移動變形，及定位。資料失真由傾斜及晃動的載台(飛機及衛星)造成。土地高度的變異，及大氣影響同樣會移動良好的結果。

## 由數位航空照片及衛星影像建立數位高程模式

Makarovic (1976)區別幾個 DEMs 照片取樣方法，取樣的選擇是當樣點是之前選擇或在取樣分析期間，適當的取樣是多餘的樣點在地面取樣過程期間可以被拒絕而可以攜帶最少的額外資訊。進步的取樣是取樣及資料分析可以一起完成，資料分析的結果會支配取樣應如何被分析。取樣是可由手動的，例如，由人的操作來領導立體繪圖機。這是緩慢過程並容易產生誤差。半自動系統已被發展來引導操作，這些是改進速度及正確性的結果；它們被認為將

會比雖然快但是不夠精確的全自動系統更好。

取樣根據結果需要可能由不同方法分析。有目的的取樣由數值化輪廓線，形成線、剖面及形態線段。在許多的目的，基於高程基礎的更普遍 DEM 是必須的，所以基於規則或不規則的網格的 areal 取樣是可以實現的。在這方面，由準確的航空照片取樣和由其他空間屬性使用的取樣技術沒有太大的不同(如逢機、分層逢機或固定網格)除了取樣強度非常大時。同樣的，由航空照片表示的海拔高攝影家可以看見樣點是恢復有效或沒笑的資訊。固定的取樣網格具較低的面變化尺度適應性；在變化較少的區域太多點可能會被取樣，面積變化較大的區域，樣點數可能較少。假如操作者是隨心所欲自由的選取觀測值，取樣可以具有很高的主觀性。Makarovic (1973)提議一個稱為進步取樣'progressive sampling'的方法，提供非常複雜的地形主觀及全自動的取樣方法為了產生網格。進步取樣牽連到一系列的連續趨勢，首先由粗略的網格開始，然後繼續為密度更高的網格(圖 5.15)，在不同的連續取樣分析網格密度為兩倍，被取樣點由電腦分析包含先前分析獲得的資料。

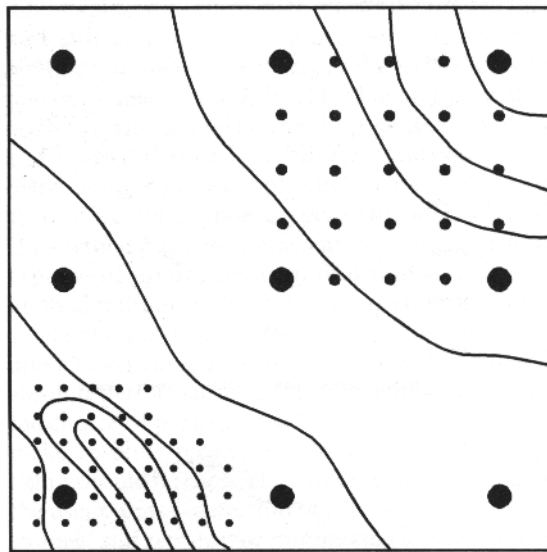


圖 5.15：在 progressive sampling，網格密度自動調整區域複雜的地形（原文 p.126）

電腦分析過程如下：選擇一個具九個點粗略的方形的碎片，每個相鄰區塊點的海拔高的差異順著行及列被計算。第二個差別在那時被計算出來。這些攜帶關於地形曲率的資訊。假如量測曲率超過某些門檻，則它需增加取樣密度及在下一個分析時另一級水準下的網格密度的樣點。

良好的進度分析取樣工作是當照片上沒有不規則面積，或人造物體；它最好由固定網格或包含水平、輕微隆起或平滑波面的半規則地形。具不同形態特徵區別的適當的粗糙地形及某些不規則區域可以由修改進步的取樣稱為混合取樣獲得更佳的处理(Makarovic 1977)。在混合取樣突然干預自然或不規則物體的地型或邊界，首先在這些區域內取樣之前先由手描繪輪廓。許多突然改變的粗糙地形表面或許不可能由半自動進步或混合取樣方法有效覆蓋，所有資料可能由選擇性取樣獲得。

最後，由進步或混合取樣的資料收集需自動化覆蓋填滿全部一致的海拔基值。

### 由數值化輪廓內插至網格

最常使用的地形的線模型是由印製好的地圖及數化的輪廓線提供一個準備好的數位高程

模式的資料。國際製圖機構已經花費很大心力來使用掃描器來自動製圖(見第四章)。不管數化已存在的輪廓線比直接由照片所產生的 DEMs 具較差品質的爭論(Yoeli 1982)，不幸的，數值輪廓並不特別適合計算坡度或製造陰影緩和模式(shaded relief models)，因此它們需轉換成高程網格。

當人們企圖由數值輪廓線建立所有的 DEMs，及使用區域內插的方法如倒轉距離權重及克立金法(見下章)將輪廓線內插至固定網格常獲得不滿意的結果。假設由內插權重計算後的假設的問題不是很大，如幾何學的搜尋演算法。假設的個別問題是平滑的輪廓線是地形及位置誤差的真實代表，是僅由紙的伸張及數化的配置、內插數值輪廓限制固定網格所產生的面是劇烈的人工產物的結果。很奇怪的，更多的關心是喜歡由許多樣點數化輪廓線，這是更大的問題。誤差的理由可以看圖 5.16。精準的數值輪廓線反映所有具相同 z 值的資料點。由網格上未取樣位置來評估 z 值關係到發現研究半徑內點資料的某些數量及當時計算的權重平均。假如所有點資料具相同的 z 值，結果將具相同的 z 值。網格(淨)結果是所有內插的窄的區域每個輪廓區域是具相同 z 值，因此每個輪廓線是被轉換階梯形。問題通常是較大面積的浮雕品趨勢線是分離的特徵及偶然搜尋演算法僅抓取最大的趨勢線資料。

當然，由階梯狀'padi' DEMs 產生的其他誤差(第八章)；計算坡度通常產生具不自然線條的圖(附圖 5.4)。通常忽略這些誤差，因為誤差的來源無法了解。例如，圖 5.8、圖 5.9c、圖 5.10c 證明，由 Isaaks and Srivastava (1989)使用的 Walker 資料設定好像因為內插的數值輪廓線的誤差結果變的較為退化。

輪廓線問題較好的解答是使用其他設計來處理由數化輪廓線產生的資料種類的內插方法。假如這是無效的，最好的策略是使數化輪廓線的點變得極細，增加額外的點於資料的設定來指出山峰、河流、溪谷坡度轉折點及以一個大的搜尋視窗使用許多資料點去內插。它的計算要求將更高但是它比取樣內插的方法提供更佳的結果。

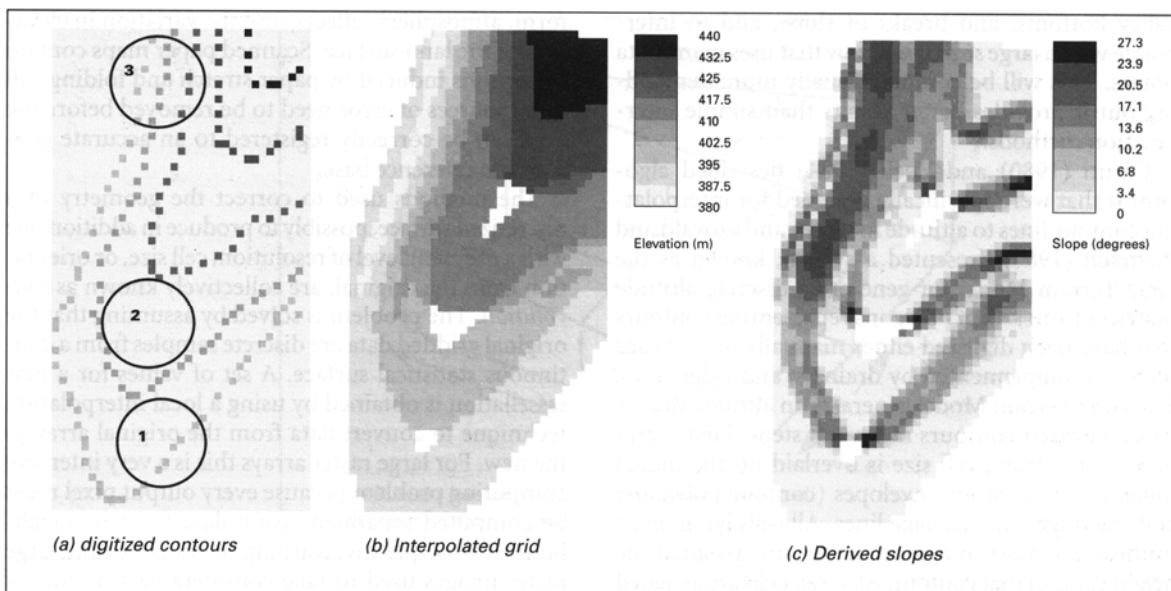


圖 5.16：藉由具簡單的搜尋圓形的數值曲線可以建立內插表面嚴重的變形及誤差 (原文 p.127)

Ceruti (1980) and Yoeli (1984)描述特別設計來內插輪廓線到高程網格的演算法並由 Oswald and Raetzsch(1984)展示這個系統，如所知的 Graz 高程模式是用來產生分離的多邊形

高程網格表現已經以手或掃描器數值化的流域或山脈線條的輪廓線。Graz 高程模式由許多等級的數值輪廓線產生高程網格。首先將一個適宜單位大小的網格套疊在數位影像作為輪廓線的外膜(輪廓線多邊形)及山脈、流域線條。所有單位網格貼覆或直接緊鄰輪廓線以分配輪廓線的高程值。所有其他單位網格分配一個為-1 的值。這些其他單位網格在下一步驟分配的高程值是在網格資料庫的矩形或視窗一個線性內插程序工作。內插通常沿著四個搜尋線段方向如 N-S, E-W, NE-SW, NW-SE。經計算視窗內區域陡峭坡度內插過程如一個在已分配好不同高度的高程單位的差異的簡單函數。每個視窗的坡度被分為四個等級。由最陡峭的坡度開始，為分配的視窗內網格單元被分配一個高度；這程序是由其他坡度(除在所有以計算後的 DEM 陡峭部分的被分離計算的平坦區域)等級重複。Oswald and Raetzsch (1984)聲稱這個內插方法(連續陡峭坡度演算法)是一個健全及有用的技術。注意商業的 GIS 可能包含輪廓線但是並不顯示演算法，因此使用者無法得知結果的品質。Carrara et al. (1997)報導在商業上成功執行由數位化輪廓線獲取 DEMs 被廣泛利用的方法比較對照。

### 由 ALTITUDE MATRICES 及其他網格資料的幾何修正

所有由航空及太空載具收集的掃描影像會包含地理失真是因為地球曲率、載具傾斜及晃動大器影響及土地表面高程的變動。掃描的紙的地圖包含的失真可以由紙的伸張及摺疊減少。所有這些誤差來源需要在資料可以正確地表示到精確的地理參考基礎之前被分離。

這些使用在正確的幾何分離表面的分法，可能生產另外一個具有不同水準的解析、單元大小或由最初的方位，如所知的迴旋。這問題由假設最初網格資料是由連續統計面的分離取樣解答。一個新的網格的設定值是包含使用一個區域內插技術去轉換由原始陣列到新的陣列的資料。由於大的網格陣列是一個密集的計算問題因為每個輸出的像元使用由鄰近區域所得資料需被個別的計算。因此彎曲或轉換大的網格影像使用來花費相當多數量的計算時間但計算處理器的技術發展及記憶體以景減少時間需求來轉動一個大的網格矩陣由數小時至轉瞬間。

變形關於兩個分離的程序。首先是輸出影像的位置計算結果和數入影像的矩陣相比較。一般第一階的方程式變形關於兩個分離的程序。首先是輸出影像的位置計算結果和數入影像的矩陣相比較。Adams 等(1984)提出一般性的第一階變形(轉換、旋轉、尺度)的方程式

$$\begin{aligned} u &= a_0 + a_1a_2x + a_1a_3y \\ v &= b_0 + b_1b_2x + b_1b_3y \end{aligned} \quad 5.22$$

$x, y$  是原始像元座標， $u, v$  是新的座標， $a_0, b_0$  是轉換值， $a_1, b_1$  分別是  $x, y$  的尺度值。 $a_2, b_2$  及  $a_3, b_3$  是根據旋轉的角度  $\theta$ ，如下

$$\begin{aligned} a_2 &= \cos \theta & b_2 &= -\sin \theta \\ a_3 &= \sin \theta & b_3 &= \cos \theta \end{aligned} \quad 5.23$$

變形並不是共面的，如推估衛星影像至地球曲線表面以抓取常見的地圖投影如橫麥卡托投影座標，需使用高階變形。有幾個方法被使用在區域內插上。最簡單的及最多限制的是由原始面最接近的鄰近區域內插變形表面網格單元值。較佳的選擇是一個雙線性的內插如新值是由四個輸入像元圍繞輸出像元(圖 5.17)。最佳的內插可能是使用鄰近區域的 16 個像元及權重和的方法基於兩方向( $\text{sincx}(\text{sincx}/x)$ )函數描述的立方迴旋(圖 5.17b)。



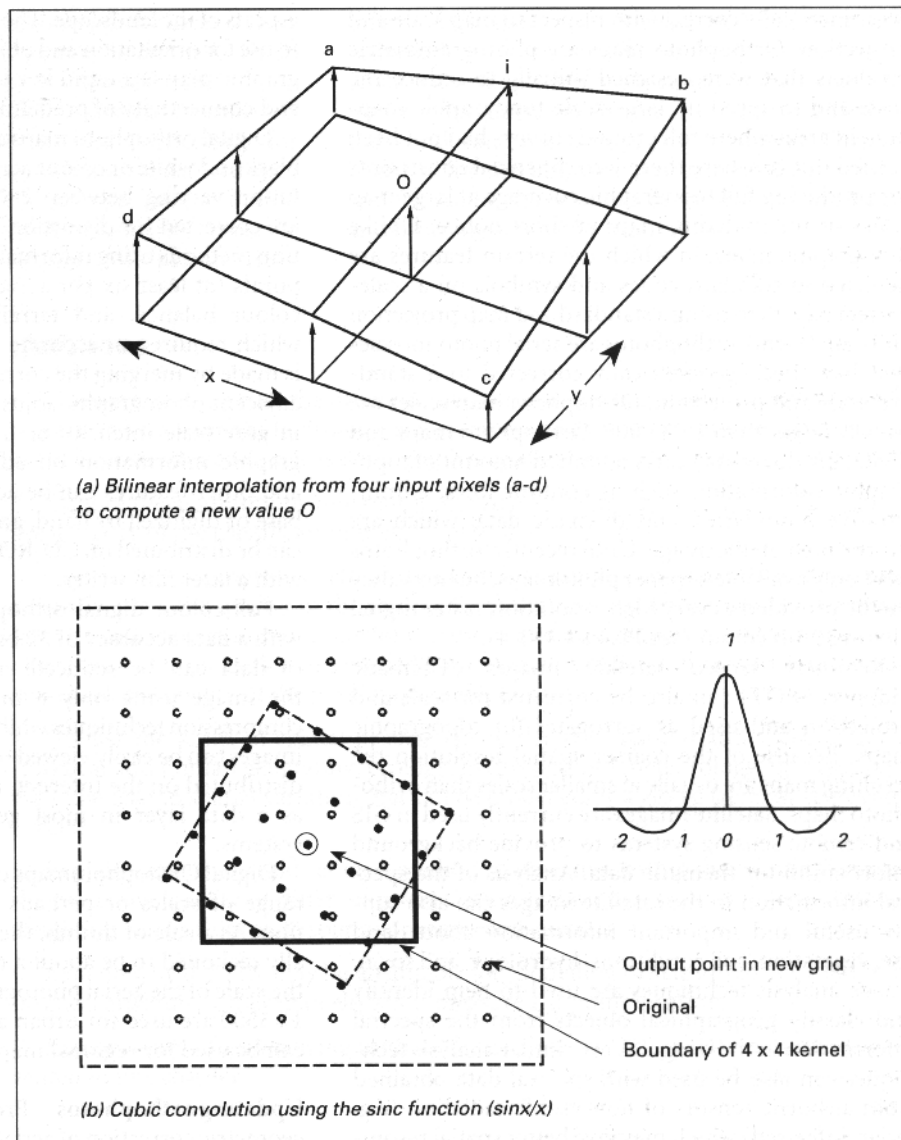


圖 5.17：(a)由雙線性內插四個輸入像元在網格單元 o 計算新值(b)使用  $\text{sinc}(x)$  函數的立方迴旋（原文 p.129）

### 數位化 ORTHOPHOTOS

第八章解釋許多可以由高程網格分離的有用結果，但或許最有用的可以獲得 DEM 的數值製圖結果是近代(modern)數值製圖。越來越多的數值製圖圖籍使用在提供幾何校正、在土地使用及土地覆蓋所取得資料的高精密攝影影像及如有用的背景資料在公用事業、市政管理、及環境研究上(附圖 1.1-1.4)。

Orthophoto 是一個相同幾何校正方法的攝影圖是一個以尺度及投影的幾何校正的地形圖。Orthophoto 圖是最初為減少花費及區域內大尺度地形製圖的攝影產品，是地面調查不可獲得或是生產大圖籍比例尺上的全部地形覆蓋或可以即時更新圖籍的產品。不像地形圖，它的地形特徵是由標準的規定所描繪，及在一個尺度校正的基本圖的象徵是使用一個標準化地圖投影及圖例，orthophotos 是已經幾何校正為標準尺度及投影的鑲嵌的航空照片。通常 Orthophoto 圖的比例尺大於 1：25000。Orthophoto 圖及數值化結果套映在影像上可能獲得很小數量的地形資訊如趨勢線、行政區邊界及主題資料。直到最近，orthophotos 過去僅和紙張製圖具相同價值，但是現在它們可以提供非常高解析度的 32 位元全彩的網格數位檔案(見附

圖 1.1-1.4)。

遙測衛星影像(主題圖, SPOT)同樣可以進行比例尺及投影校正, 並如地形圖般使用。因為出略的空間解析所產生的圖通常比 orthophoto 圖具更小比例尺。衛星影像現在使用在 GIS 及遙測系統以提供主題資料的地面資訊。衛星影像的空間資料分析同樣可以提供土地使用、植被、土壤組成、水文及諸如此類的有用及重要的資訊。影像分析技術是用來幫助地易及分類影像光譜資訊的地形主題。非常相似的分析技術同樣可以使用在由航空非可見光感測氣的光譜資料(如雷達、遠紅外線)可能比衛星影像得到更好的空間解析。遙測如一個可負擔的工具通常使用在獲得地球表面在時間改變的資訊。遙測影像及分類結果通常主要利用如數位化資料庫; 同樣的可以使用高品質底片輸出在紙上或底片。

Orthophoto 圖(或衛星影像圖)及地形圖的主要差異, 當地形圖表現結構及地景方面選擇的分類資訊, 前者包含所有地景方面的攝影資訊在使用的解析水準是可見到。Orthophoto 圖很容易使用在定位及目標判釋上; 地形圖的關鍵則是結構位置及連結先前定義的物體。

由黑白或彩色照片掃描所獲得的數值 orthophoto 圖, 解析度變化在  $200 \mu\text{m}$  至  $10 \mu\text{m}$  之間。失真影像的校正是由數學校正方法使用資訊(如地面控制點(每張航照圖至少六個點)、照相機光學、色彩平衡、地形高度的差異)來要求一個正確的 DEM。鑲嵌照相是由鄰接的照片得到的校正後的數位影像合併所成, 藉由不同灰階值強度或色彩平衡調整。行政邊界的地形資訊及其他特徵附加在 GIS 的資料庫或由手數化, 數化的 orthophoto 可以分散於 CD-ROM 如由雷射檔案寫入的硬碟拷貝。

全彩數位 orthophoto 影像由準確的每像元 32 位元儲存一個資料。資料的容量可由近使用每像元 8 位元紀錄影像或影像壓縮技術大量減少, 使 orthophoto 影像可以在個人電腦輕易查看或散布於網路上, 並可以在多數的地理資訊系統合併資料層級。

數位的 Orthophotomaps 可以在一個更廣的尺度範圍或是或許更佳的解析水準生產。如一個實用方法, 地圖比例尺通常由最大三個時間的航空照片尺度計算。比例尺 1:1000 至 1:3500 是使用在都市區域, 而比例尺 1:50000 可以使用在區域製圖。

#### orthophotos 更新

由航空照片幾何校正的 DEM 結果是在製作數值 orthophotos 最昂貴的部分。然而, 從很多的地景外觀並不隨時間改變的例子, DEM 不需由每次飛行的航空照片重新計算, 已存在的面可以用來支持校正新的照片。當數位資料是在 GIS 內, 它可以更簡單原始情形(如土地利用)來和新的情形比較並可以簡單獲得量化的變異評估。在某些情形, 如戶外的採礦業、海岸的沙丘侵蝕、土木工程新建移除操作, DEM 的更新允許改變的植被快速、正確的計算。有用的影像對於數年間全部過程的改變是可以研究的, 如海岸線或河流的評估。太多地形結構資料可以由航空照片蒐集, 特別是在大及中比例尺的應用上。地景方面不可見到的資料結構, 如埋藏的輸油管及部分邊界調查, 需由野外調查蒐集。近來 GIS 允許向量資料轉換成可看及分可分析的網格結構, 如由衛星影像的掃描資料或數值 orthophotos。因此地形資料的結構可以顯示全球概況, 可以在適合的背景觀看。對於檢查資料品質、正確性及為了看到資料結構和其他方位位置的關聯性相當有用。

## 第六章 由地理統計學作適當的插入推估

### *Optimal Interpolation Using Geostatistics*

#### 前言

當數據十分大量時，大部分的插入推估技術可以給予相似的結果。當數據較為稀落時，無論如何，有關於下述變動的假設，必需進行取樣及方式的選擇，且其參數必須要能夠迴避一個會令人誤解的結果。地理統計學的插入推估方法最為人所熟知的便是克里金法，其企圖將變化的空間理想化的分為以下三個元件：(a)自然的變化（不同的層級或趨勢）可以被視為有益的或軟性的資訊、(b)空間上的自動相關性，但物理上的變化較難去解釋，最後可視作(c)沒有關係的雜訊。這些空間上不同關係的特性被包括於自動變異或半變動的函數中，且可以提供理想化的插入推估影響及搜尋半徑。根據一、二或三個空間尺度的變動樣本數據實驗結果，可以對其進行計算，這些實驗的數據適用於一個有限的變動模式，以供作插入推估的計算影響。

地理統計方法為插入推估提供了極佳的彈性，在推估大於資料提供者的面積或體積時可利用克里金空間推估法，此方法可以進行二元數據（指示克里金 indicator kriging）的內插，並合併有關趨勢為通用克里金（universal kriging）及階層克里金（stratified kriging）的資訊。所有插入推估的方法將產生一個平滑變化的表面以表示現存的變化。對照這些平滑的內插值，這是一個有條件的模擬，利用變量及現有的觀察，推估未知點最佳的數據值。這結果的表面不需要平滑，但將比克里金法的預測值有較小的誤差。結合相關資訊及有條件的模擬，將有利於計算以網格數據為基礎的環境模式。

最後，這些變量的資訊可以被用來幫助點數據製圖的合理化。

#### 地區變化數理論與克里金法的介紹 (A brief introduction to regionalized variable theory and kriging)

當比較過所有第五章的氾濫平原土壤鋅值的圖，我們可以發現使用不同的方法在綜合取樣、地方詳細資料統計及最大最小值預測中帶來相當大的不同。這些不同的統計（見表 5.5），按照不同的統計方法，推估面積比率上各地方鋅值在 500、1000 或 1500ppm 以上，顯而易見的，不論我們面對的是所有不同的結果，或是對於一個特定含鋅量的等級決定一個清潔土壤值，都是相當困難去選擇的。沒有一個插入推估的方法可以按照估計的不同，直接討論預測值的品質。在所有的案例中，只有一個方法去決定估計值的正確性，那便是對一個未經過推估的額外點位進行計算。但除非對判斷技術的品質研究練習，很少進行判斷優劣的研究，因為這將需要花費相當多的金錢。

此外，所有的研究都有缺陷，沒有一個先決的方法知道如何選擇最適當的權衡參數值，或者這個選出的尺寸是否合用。就我們所觀察，層級表面的控制參數及相反的距離權衡，由於不同的推估方式能夠產生不同的圖像。再者，沒有方法可以察覺以下資訊：

- 點的數據必須計算局部的平均
- 經由點指出鄰近的尺寸、方向及形狀
- 是否有更好的插入推估方法取代現行簡單的距離函數
- 因為插入推估值所合併而來的錯誤（未知之處）

為了採礦工業，法國的地理數學家 Georges Matheron 和南非的採礦工程師 D. G. Krige 針對這些問題，發展出插入推估理想化的方法，現在這個方法已經被增加作為地面水模式、土壤圖及相關區域的使用，而其也使得地理統計上的插入推估方法在商業用的 GIS 方面更為重要。

地理統計學的插入推估方法起源於承認每一個連續屬性的空間變化，這通常過於不整齊而無法用簡單、平滑的數學函數進行描述，取而代之的是在一個隨意的平面上，能夠對變化進行描述的方法。這個特性被用作表示區域性的不同，例如大氣壓力、海平面以上的高度或是連續區域人口統計分佈指標。在 D. G. Krige 後，地理統計學上的插入推估方法即被視為克利金法。

地理統計學方法解決了上述插入推估方法的限制，並保證了這個推估的屬性結果是按照假設條件下的理想數值。為了選擇變動值，最佳的策略是利用動態程序原理將函數理想化 (Bullock and Stallybrass 1977)。這個插入推估方法是由 Matheron、Krige 及其同業所發展，目的是為了選擇插入推估的影響值，使其功能函數更為完美，例如 BLUE (最佳線性合理推估 (Best Linear Unbiased Estimate))，這個方法也可作為網路取樣的理想化。

地區性不同原理假設任一個無偏估值，可以被三個主要的元件所概括(圖 6.1)。這裡有(a)一個結構的組成原件，具有一定的意義或傾向，(b)一個任意數，但空間的關聯組成原件被認為是其變量，和(c)一個空間上非關聯性的任意雜數或剩餘的錯誤項目。x 被認為是 1、2 或 3 尺度中的所在位置，而其任意變量的值 Z 與 x 關係如：

$$Z(x) = m(x) + \varepsilon'(x) + \varepsilon'' \tag{6.1}$$

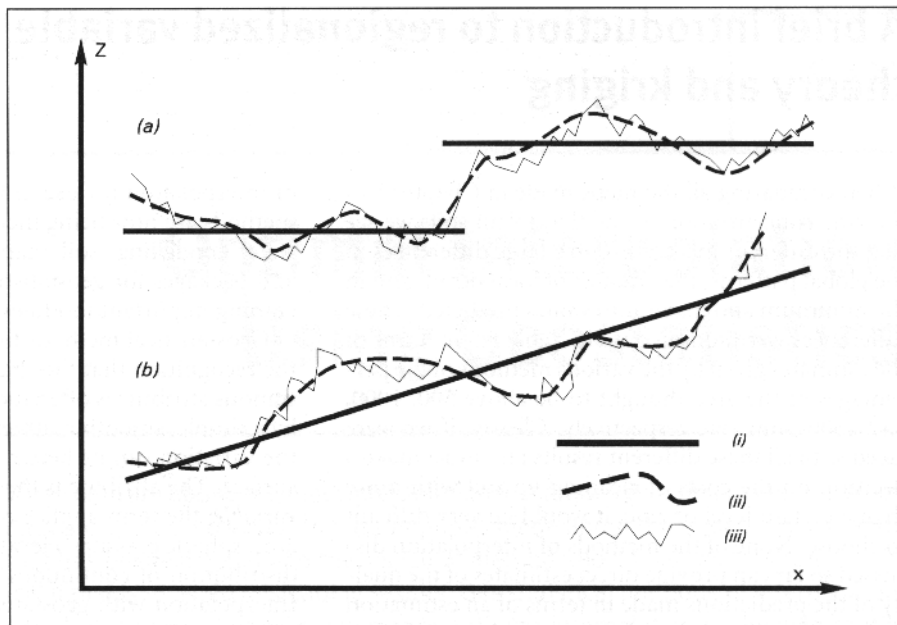


圖 6.1：區域變化原理衍生完整的空間變化為(i)不同上下變化的平均特質，(ii)空間上連續但隨意變化，及(iii)由定義錯誤或小的值距變化所造成的隨機、不連續區域 (原文 p.134)

在這裡  $m(x)$  是一個不等的函數，用以描述  $Z$  在  $x$  情況下的結構組成， $\varepsilon'(x)$  則是指示一個隨機的變化，但空間上區域的變化卻隨著  $m(x)$  即區域變量而異，而  $\varepsilon''$  是一個餘數，表示空間上的獨立，Gaussian noise 則有  $0$  及  $\sigma^2$  變量的意思。記錄這些字母所代表的意義，表示  $Z$  是一個隨意的函數且沒有一個標準的屬性  $z$ 。

第一步是決定  $m(x)$  合適的函數，在沒有傾向或水流的案例中  $m(x)$  相等於面積或平均數取樣的值，或是表示利用距離向量  $h$  分離出兩個空間  $x$  及  $x+h$  的不同，這將被表視為：

$$E[Z(x) - Z(x+h)] = 0 \tag{6.2}$$

這兒  $Z(x)$  及  $Z(x+h)$  是任意變量  $Z$  在  $x$  及  $x+h$  時的數值，也因為如此其影響變量的原因便是兩個位置的距離  $h$ ，所以：

$$E\{[Z(x) - Z(x+h)]^2\} = E\{[\varepsilon'(x) - \varepsilon(x+h)]^2\} = 2\gamma(h) \tag{6.3}$$

這兒  $\gamma(h)$  被認為是半變動的，其具備兩個條件，一是固定的變動，一是有差別的變動，可藉由區域性變量原理所假設的需求加以定義。這個意思指的是即使一個結構影響也會被記錄，其剩餘變量將是均質的變化，所以兩個不同的位點僅被一個介於兩者間的功能函數所表示，我們可修正函數 6.1 為：

$$Z(x) = m(x) + \gamma(h) - \varepsilon'' \tag{6.4}$$

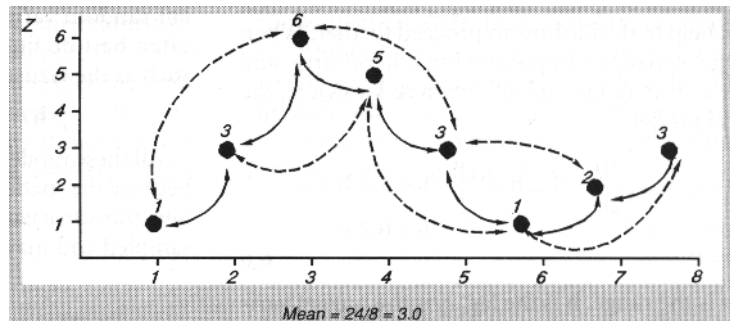
以表示  $\varepsilon'(x)$  與  $\gamma(h)$  之間是等值的。

假如這個條件被原本的假設所滿足，則這個半變動的值將可被現有的樣本數據加以估計：

$$6.5$$

這兒  $n$  代表的是屬性值  $z$  的觀察取樣點數對，其因為距離  $h$  而分離（見表 6.1），一部分  $\gamma(h)$  值相對於  $h$  被視為實驗值變量。這值驗值的變量是描寫區域變化的第一步驟，其變異圖可為插入推估工作、理想化的取樣及決定空間形式提供有用的資訊，無論如何，這樣的步驟對於實驗變異圖必需先適用於一個理論模式。

Box 6.1 (原文 p.135)



### 配合變異圖模式 Fitting variogram models

圖 6.2 表示出一個典型的非平滑變化屬性數據變異圖，例如土壤性質，這個曲線被安排通過實驗所導出的點，以表現出其重要的特徵。首先，位於後方最大值  $h$  已經脫離了此層次，其水平的部分被認知為基質，其表示介於點數據間部分沒有空間依賴，因為所有不同的空間變量估計將被不變的數值所表示；第二，由  $\gamma(h)$  所表示的向上曲線延伸至  $h$  所表示的範圍被認知為值距，這是變異圖中精密的部分，可表示內部空間的不同，在值距內具有很多相同處，且可以提供我們在進行插入推估時加權平均值的變動。清楚的，假若這個點位距離值距過遠，將無助於進行插入推估的工作。

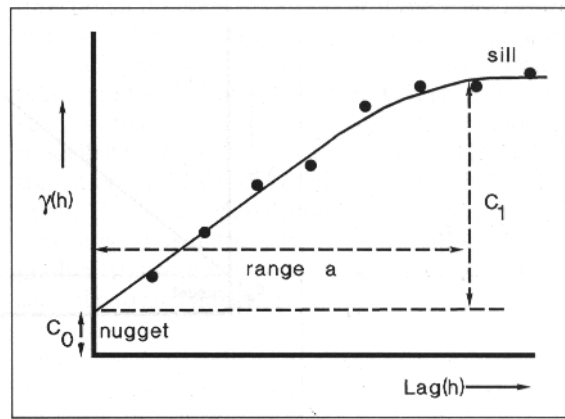


圖 6.2：值距(range)、nugget 空間推估法及基質(sill)的簡單轉變變異圖的案例 (原文 p.135)

圖 6.2 顯示的第三點，表示這個合適的模式並不會通過原點，但會明確的切 y 軸於  $\gamma(h)$  上一個定值，根據等式 6.22，當  $h=0$  時這個半變動的等式亦為 0，因為其中的不同被定義為 0。當  $h \rightarrow 0$  時  $\gamma(h)$  值即被認為是一個  $\epsilon$  的估計值，其剩餘空間上非關聯性的  $\epsilon$  值，即為所認知的 nugget 空間推估模式，這個空間變量的尺度誤差組合，相較於無法斷定的間隔取樣點是相當短的。

空間變異圖的形式能夠表示一個面積內空間變化的種類，並有助於未來的發展，當 nugget 空間推估模式的變動不大且重要時，將會有清楚的值距和基質，其曲線將被認知為球的模式：

6.6

其中 a 值便是值距，h 即為 lag， $C_0$  為 nugget 空間推估模式的變量，而  $C_0+C_1$  具有相等的基質，且通常適合做良好的變異觀察。

假若有清楚的 nugget、基質及逐漸靠近值距，其函數模式 (6.7) 通常是很好的選擇。

$$\gamma(h) = C_0 + C_1 \{1 - \exp(-h/a)\} \tag{6.7}$$

假若變量非常的平滑且其 nugget 變量  $\epsilon$  相對於空間任意變量  $\epsilon(x)$  相當小，則其變異相當適用於一個變化的曲率，例如高斯模式 (Gaussian model)。

$$\gamma(h) = C_0 + C_1 \{1 - \exp(-h/a)^2\} \tag{6.8}$$

所有的模式因為其空間關聯結構隨著 h 值變化，被認知為為一個轉換的變異，而非轉換的變異在面積中沒有基質，故可被塑造成線性模式：

$$\gamma(h) = C_0 + bh \tag{6.9}$$

在這兒 b 指的是線的斜率。一個線性變異圖屬性形式改變在所有的尺度中，例如簡單的 Brownian motion 動作。圖 6.3 顯示出一些案例，一個變異圖會伴隨著 h 逐漸上升，可指示出數據應被個別的塑造。

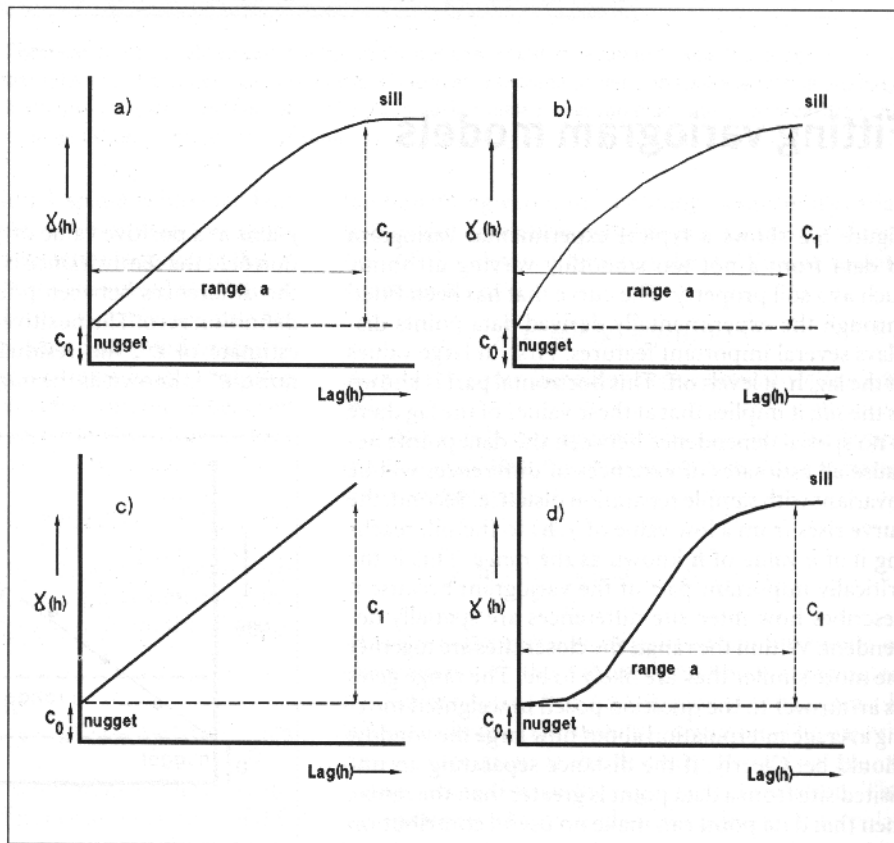


圖 6.3：一般變異圖模式案例(a)球體(b)指數(c)線性(d)Gaussian (原文 p.136)

變異圖的估計與塑造對結構分析與插入推估是相當重要的。變異圖選擇任何偶發的函數，也就是說他必須跟隨著相當正確的實驗系統規定參數 (Isaaks、Srivastava, 1989)，在沒有取得每一個 lag 的點數對值時，模式將不能適用於實驗的半變動估計，因為這些數據必需重複的於變異圖的估計中使用，且其自主性的影響程度對於最短的 lag 相當重要 (比較 Taylor & Burrough 1986)，因而其影響值被利用數對的數值進行計算，所以變異圖的適用度是一個需要判斷力與技巧的過程，而目前的地理統計學軟體通常為了半意圖的適應性，必需包括一些相互常規的影響。

### 使用變異圖的空間分析(Using the variogram for spatial analysis)

在進行插入推估的工作時，為了決定最理想的影響值，變異圖示方法上相當重要的步驟之一。當 nugget 變量  $\epsilon$  被區域變量支配時，實驗的變異圖將無法展現減少的趨勢如  $h \rightarrow 0$ ，原因是其數據過於繁雜，使得插入推估工作無法察覺而處理。在沒有空間記錄支援的情況下，最好的  $z(x)$  估計方法將是對區域內所有取樣點進行計算，這將意味著需要花費許多的時間與金錢。

在實驗中所衍生出的半變動變異圖要是相當混亂，變暗示出於計算  $\hat{\gamma}(h)$  時取樣過少，法則中建議，在完成一個穩定的變異圖時，取 50-100 個數據點是必要的，儘管平滑的表面需要較少的點勝過一些不合規則的變化，但仍需考慮到空間中會遭遇的變化，且平滑的變異也可能經由擴大研究的範圍而產生。

變異圖的值距清楚的提供了應該使用的研究範圍資訊，假若一個數據點超出了值距的範圍，則其將太遠而無法進行任何貢獻，假若所有的估計點皆遠離了值距，則最佳的估計為一

般值。這些距離可藉由一些方法進行修正，其中圓形內鄰近點的修正將使其變為橢圓形。

在實驗的變異圖（半變動的距離大於範圍的 dip）中穿洞效果 *hole effect* 的存在，顯示出一個假的循環是因為過大的變動範圍，使的研究面積太小不足以包含所有的變化，真正的週期將提供變異圖在基質中一個與區塊波長相合的週期的變量，並提供最初的採樣點一個週期的軌跡。

假若值距過大，則其中最長的值距變量將支配變異圖，假若其過小，則主要的變量將存在於最短的距離之上。在變異圖中方向的不同會提供區域內相當的影響，在取樣不充分足以確立估計方向時，則方向將會產生誤差。

一個變異圖能適用於高斯（Gaussian）變異圖模式，則表示其區塊內具有平滑的變化，例如一般的海拔數據。利用球形變異模式塑造變異圖有一個支配轉變的轉變點。實驗變異模式的選擇可提供區塊內變量有一個逐漸擴散尺度及區域間影響的展現。

### 等方向性與各向量之變異(Isotropic and anisotropic variation)

在前面的敘述中，我們可以隱約的了解變異圖的數據來源，可藉由有規則的空間線條或網格加以蒐集，在許多案例中，我們只有不規則的資料，所以這些數據通常需要經過實驗變異圖的計算。為了進行這樣的工作，我們使用圓形的研究範圍，如腳踏車輪胎一般，去定義一個從中央點距離  $h$  的區域。這個輪子被安放於數據點上，而落如其中的點將可用來估計所有數據  $(z_i - z_j)^2$  的變化（圖 6.4）。一般影響的法則中，通常無法察覺並計算在許多 lag 的變異。假若我們忽略了方向的影響，則這個變異圖的結果被認知為等方向性的，其平均了變異在所有方向中。無論如何，如圖 6.4 所示，在特定的距離  $\beta$  內變異圖的計算是容易的。這些被認知為非等方向性的變異圖，且其不同  $\beta$  值的值距、基質皆可藉由方向指示其空間變化，例如一些垂直沉澱或平行的河流。

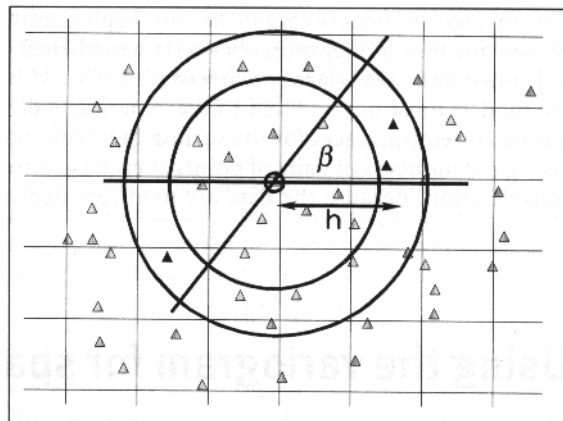


圖 6.4：任意距離  $\beta$  的半變異估計圓形搜尋視窗（原文 p.138）

藉由計算所有可能方向  $(z_i - z_j)^2$  可建立一個以 0,0 為中心點的橢圓形變異圖表面，這個變異圖的座標軸有更長的橢圓形座標，且其方位可藉由與正北方的角度訂定。

### 不同尺度下以變異表示空間變量 (Variograms showing spatial variation at several scales)

這原先的共變異模式簡單的將屬性的空間變量區分為三個元件：(i) 一個一般的意義



或傾向，(ii) 一個變異圖，及 (iii) 剩餘的 nugget。在某些情況下，特別是有充分的數據時，也許可以區別超過一個的變異組成，例如有兩個任意的區塊且其具有個別的值距相互影響，在這裡完整的變異圖將根據 coregionalization 線性模式被區分為幾個元件 (Isaaks & Srivastava 1989)：

$$\gamma_T(h) = \gamma_1(h) + \gamma_2(h) + \gamma_3(h) + \dots \tag{6.10}$$

所以可將 6.4 的程式替換為：

$$Z(x) = m(x) + \gamma_1(h) + \gamma_2(h) + \varepsilon \tag{6.11}$$

每一個亞變異圖將因為其具備的參數形式而被定義。

### 不同覆蓋層級的空間變異 (Spatial variation within different cover classes)

直線趨近的去計算變異圖，前提是假設所有的數據皆位於同一個覆蓋層級或範圍，這樣的計算變異圖方式為一個通用模式，它可以決定整個區域內插入推估的影響值，區域內將不需要在土地覆蓋、硬體形式、軟體形式及氾濫頻率上有重大的不同，每個區塊將擁有其各自的空間關係結構或圖案，例如河流中重金屬的分佈情況，氾濫平原與半山腰上將有顯著的不同的，這便是因為重金屬的來源及傳輸過程將相當不同。

當面對不同範圍的重要決定因素時，必需觀察研究區域是否能夠用單獨的通用模式，更由於表示變異圖的空間特性，其問題通常是因為數據的不足，造成通用變異圖不當的使用及有限的區域模式，比較實際的做法，是企圖以現有的區域模式的空間變異圖單位，求出變化，成為新的多邊形屬性變異圖，且其參數將可被相關數據資料庫儲存 (Pebesma 1996, Laurini and Pariente 1996, in Burrough and Frank 1996. Voltz and Webster 1990)。

### 使用變異圖進行空間插入推估：次序性克立金法 (Using the variogram for interpolation：ordinary kriging)

被賦予空間地相依靠隨意的變量不被無關的雜訊所淹沒，合適的變異圖可被用來判定需為區間差入之  $\lambda_i$  的權重，和被用於權重移動平均值差入是相似的。除非現在的

$$\hat{Z}(X_0) = \sum_{i=1}^n \lambda_i * Z(X_i) \tag{6.12}$$

權重是由地理統計上分析的資料推論而來，而不是從一個一般的、可能不適當的模式所推論得到的。我們可得：6.12，當  $\sum_{i=1}^n \lambda_i = 1$  時， $\lambda_i$  的權重被選出，因此推估  $\hat{Z}(X_0)$  是公平的，且推估變異數  $\sigma_e^2$  是較任何其它線性組合的觀察值來得少。

$\hat{Z}(X_0)$  最小的變異數是由

$$\hat{\sigma}_e^2 = \sum_{i=1}^n \lambda_i * r(X_i, X_0) + \phi \tag{6.13}$$

所給予的，而且當  $\sum_{i=1}^n \lambda_i * r(X_i, X_j) + \phi = r(X_j, X_0)$  對所有 j 則可得到。

數對  $r(X_i, X_j)$  是 Z 在取樣點  $X_i$  和  $X_j$  間的二次變異， $r(X_i, X_0)$  是取樣點  $X_i$  和非調查點  $X_0$  間的二次變異，這些數皆是從配合的變異圖中獲得，數  $\phi$  是一個所要求最低限度的 Lag 值距乘數，這個方法即是次序性克立金法。它是一個精確的插入，就那方面而言，當方

程式如上被運用時，插入值或最佳區間的平均數將與資料點的值同時產生。在地圖中，值將被插入點在規則的格子中，如果那是自己的一個格子，較被用來取樣的地區更佳。接著插入值可被改變成為一個運用完全描述的手法的輪廓圖。相同地，判斷誤差  $\sigma_e^2$ ，即克立金變異，也可被製圖給予的有關插入值的確實性的有價值資訊，超過區域的重要性。常常克立金變異可如克立金標準偏差（或克立金誤差）被制成地圖，因為這有如預測相同的單位，且這個慣例在本章接著會敘述，Box 6.2 和圖 6.5 顯示次序性克立金法的方程式如何建立與解釋。

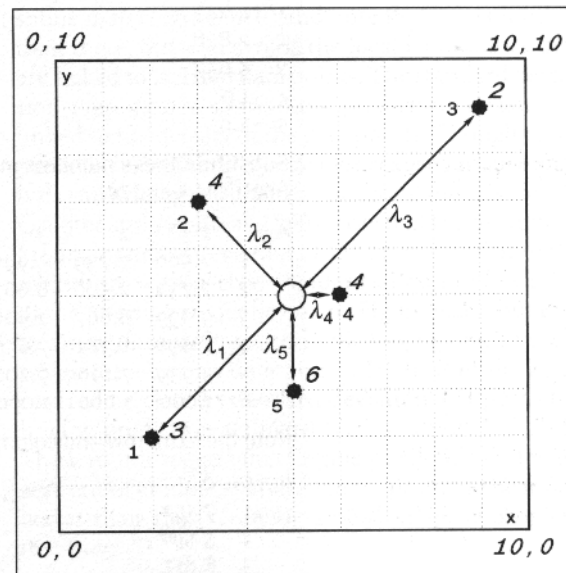


圖 6.5：藉由次序性克立金法簡單舉例說明，預測 Z 在非取樣基地的值（原文 p.139）

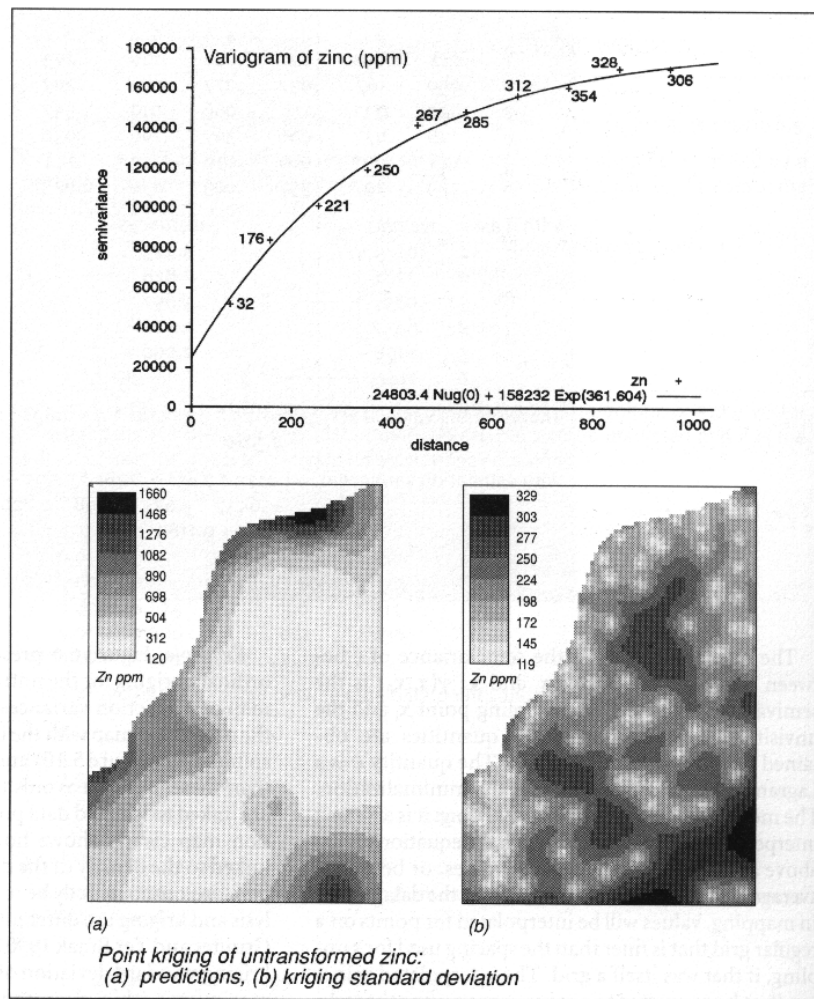


圖 6.6：使用次序性克立金法插入 Maas 資料組的結果。上：變異圖使用適合的指數模式；(a) 鋅層級的克立金敘述，(b)克立金標準偏差（原文 p.142）

例子圖 6.6 呈現使用次序性克立金法於未轉換之鋅資料的結果，預測變異的地圖也被顯示。比較預測地圖，運用反距離地圖和雲形規 spline 圖（圖 5.10），注意是較這些更平滑的，並避免區間值與隔離資料點的鏈結。雖然比較不可能被嚴密的製作，因為變異分析和克立金法使用不同之統計模式（見 de Gruijter and Ter Break 1990），注意最大的克立金標準偏差 458ppm 和內在類別標準偏差的類別 1 的氾濫頻率圖（圖 5.2c）相似，但這個最大值只發生在沒有資料點的區域之克立金地圖。在大部分被類別 1（標準偏差 423ppm）所覆蓋的區域中，克立金預測標準誤差大約為 140ppm，這些結果暗示克立金預測漸進的空間改變呈現加以考慮是較全球的、全區的、地誌圖所給予的更佳。

Box 6.2：計算克立金法權重值的應用範例（原文 p.140-141）

在圖 6.5 中，計算克立金法權重用於非取樣點 Z ( $X_i=0$ ) ( $X=5, Y=5$ )，使取樣的 5 個點其空間變量的屬性，由母數  $C_0=2.5, C_1=7.5$  和範圍  $a=10.0$  的球形變異圖所塑造，5 個取樣點的資料為：

$i$	X	Y	Z
1	2	2	3

2	3	7	4
3	9	9	2
4	6	5	4
5	5	3	6

以最大值的觀點，我們必須去解釋下述：

$$A^{-1} \times b = \begin{bmatrix} \lambda \\ \phi \end{bmatrix}$$

A 為數對資料點間二次變異的母體，b 是每個資料點和被預測的點間二次變異的向量，λ 為向量的權重，φ 是解釋方程式的一個 Lagrangian，開始藉由產生資料點間的一個差異母體：

<i>i</i>	1	2	3	4	5
1	0.0	5.099	9.899	5.00	3.162
2	5.099	0.0	6.325	3.606	4.472
3	9.899	6.325	0.0	5.0	7.211
4	5.0	3.606	5.0	0.0	2.236
5	3.162	4.472	7.211	2.236	0.0

資料點和未知基地間的向量差異：

<i>i</i>	0
1	4.243
2	2.828
3	5.657
4	1.0
5	2.0

替代這些數目字進入變異圖中，去得到相同的二次變異數（母體 A 和 b）

<i>A=i</i>	1	2	3	4	5	6
1	2.500	7.739	9.999	7.656	5.939	1.000
2	7.739	2.500	8.667	6.381	7.196	1.000
3	9.999	8.667	2.500	7.656	9.206	1.000
4	7.656	6.381	7.656	2.500	4.936	1.000
5	5.939	7.196	9.206	4.936	2.500	1.000
6	1.000	1.000	1.000	1.000	1.000	0.000

注意其他列和欄（i=6）以保證權重總和為 1

<i>b=i</i>	0
1	7.151
2	5.597
3	8.815
4	3.621
5	4.720
6	1.000

<i>A<sup>-1</sup>=i</i>	1	2	3	4	5	6
1	-0.172	0.050	0.022	-0.026	0.126	0.273
2	0.050	-0.167	0.032	0.077	0.007	0.207
3	0.022	0.032	-0.111	0.066	-0.010	0.357

4	-0.026	0.077	0.066	-0.307	0.190	0.030
5	0.126	0.007	-0.010	0.190	-0.313	0.134
6	0.273	0.207	0.357	0.003	0.134	-6.873

帶入 λ 如：

i	權重	差異
1	0.0175	4.423
2	0.2281	2.828
3	-0.0891	5.657
4	0.6437	1.000
5	0.1998	2.000
6	0.1182	φ

Σ=1

因此  $Z(X_i=0) = 0.0175*3 + 0.2281*4 - 0.0891*2 + 0.6437*4 + 0.1998*6 = 4.560$

代入判斷變量

$$\sigma_e^2 = [0.0175*7.151 + 0.2281*5.597 - 0.0891*8.815 + 0.6437*3.621 + 0.1998*4.720] + \phi$$

$$= 3.890 + 0.1182 = 4.008$$

### 利用克立金法查核變異圖模式 (Using kriging to validate the variogram model)

交差變異數是實行使用克立金方程式，回頭去查核變異圖模式。它包含計算每個資料點瞬間的  $(\hat{Z}(X_i) - Z(X_j))$  的分佈，當每個資料點被連續地省略，且從剩餘的資料(圖 6.7)所預測，程序被設計成只測試變異圖於後一致和缺少的傾向，藉由一個意謂接近 0 的不同點和一個 Z 值 (不和 Z 資料混淆) 來指出。

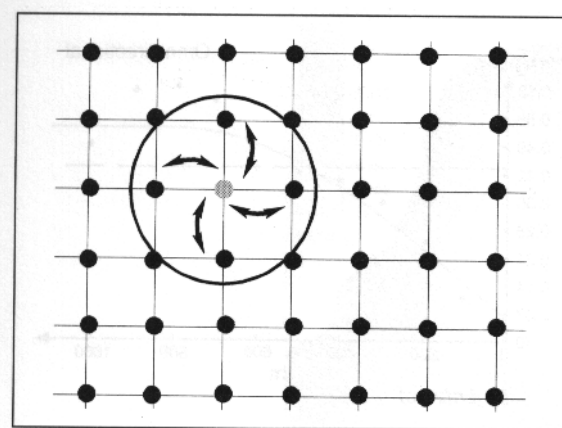


圖 6.7：交叉確認是檢查變異圖預防次序性資料 (原文 p.143)

### 區域克立金(Block kriging)

無疑地，克立金實現了發現更好的方式去估計插入之權重的目的和提供有關誤差的資訊之目的。插入值的結果呈現圖不可能是精確的完全被敘述。然而，因為點克立金或簡單克立

金，方程式 (6.13 和 6.14) 指出所有插入值和支持 (support) 相關，support 是一個次序性取樣的面積或體積。常常，如同在土壤或水質的取樣，這個取樣只是一些公分穿過 (centimetres across)，被給予的常是很多自然現象的大振幅、小範圍變異，如土壤或水質 (Burrough 1993)，一般克立金法可以成為有很多銳利的釘子和洞在資料點中的地圖，這可被修改克立金方程式去估計超過土地 B 之地區其可變 Z 的平均值  $Z(B)$  來克服，當一個預期去估計 Z 的平均值於實驗計畫給予之區域，或插入值於量的模式的特殊尺度網格 (圖 6.8) 時，這是有效的。

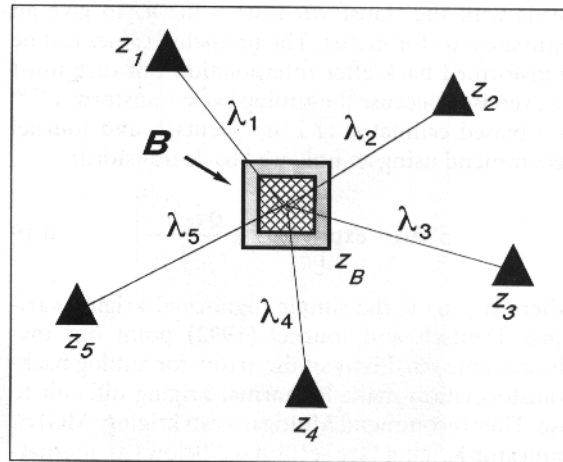


圖 6.8：以不同大小的區域敘述 (原文 p.143)

超過區域 B 的 Z 的平均值為

$$Z(B) = \int_B \frac{Z(X)dx}{\text{area}B} \quad 6.15$$

由來估計

$$\hat{Z}(B) = \sum_{i=1}^n \lambda_i * Z(X_i) \quad 6.16$$

如同之前代入  $\sum_{i=1}^n \lambda_i = 1$ ，最小的變異為

$$\hat{\sigma}^2(B) = \sum_{i=1}^n \lambda_i * r(X_i, B) + \phi - r(B, B) \quad 6.17$$

當  $\sum_{i=1}^n \lambda_i * r(X_i, X_j) + \phi = \bar{r}(X_j, B)$  對所有的 j 是可得的。

從區域克立金中得到的估計變異通常大體上是較從克立金中得到的少，當這些方程式被運用時，結果平滑的插入表面是免於洞 (pits) 和釘 (spikes) 的去產生點克立金。

## 克立金法的其他形式(Other forms of kriging)

### 簡單克立金法(Simple kriging)

簡單克立金法是由綜合的二次線性回歸固定點有一已知意義的假設下所敘述的 (Olea 1991)。因為第二次序固定點的假設也常限定於從物理環境得到的大部分資料。次序性克立金法 (沒有一個先決的意義) 是大部份常被使用的，雖然 Voltz 和 Webster (1990) 和 Webster 及 McBratney (1987) 提出實例。

### 非線性克立金法(Non-linear kriging)

Lognormal 克立金法是對數分佈的插入，而非正常分佈資料的插入 (Deutsch and Journel 1992)。資料首先轉換成自然對數或 10 的對數，因此，變異圖模式和插入過程帶有轉換  $y(u) = \ln Z(u)$  去給  $\ln Z(u)$  一個估計值  $y^*(u)$ 。被敘述的值可在插入後被轉換回去，但注意須被運用，因為逆對數 (antilog) 回歸轉換後為  $e^{y^*(u)}$  是一個偏向  $Z(u)$  的估計值。Deutsch 和 Journel 建議使用一個無偏向的回歸轉換 (back-transform)  $Z^*(u) = \exp[y^*(u) + (\sigma_{sk}^2(u))/2]$  在式中  $\sigma_{sk}^2$  是簡單 lognormal 克立金變異。Deutsch 和 Journel (1992) 指出對於逆對數回歸轉換的極敏感誤差使 lognormal 克立金在使用上不同。他們建議多高斯克立金法 (Multigaussian Kriging (MG)) 或標誌克立金法 (Indicator Kriging) (見 section 6.5 上) 中任選其一。儘管如此，lognormal 轉換對於很多帶有確定歪斜的分佈的物理資料，如土壤化學物和物理屬性 (Webster and Oliver 1990, Burrough et al 1992)。

圖 6.9a 說明實驗使用 log 轉換的資料之鋅資料的變異圖，圖 6.10a、b 顯示地圖的敘述和 log-轉換的標準偏差，普遍的格局和圖 6.6 的那些是相似的。

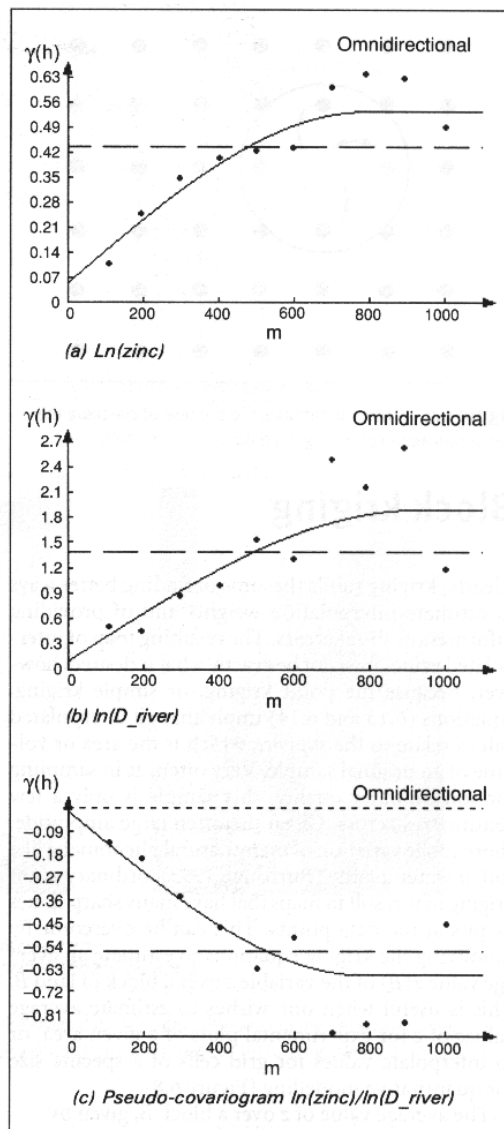


圖 6.9：變異圖(a) $\ln(Zn)$ ，(b) $\ln(\text{到河流之距離})$ ，(c)偽 (pseudo) 交叉變異圖 (原文 p.144)

### 非等方向性或巢狀變異圖運用於一般克立金法(Ordinary kriging with anisotropy or nested variograms)

合併非等方向性到一般克立金法程序對於修改距離、基質的改變到二次變異 A (見 Box 6.2) 的基質，加以考慮二次變異的方向的變異，是簡單的一個問題。圖 6.11 顯示變異圖計算於 NW-SE 向 (垂直於河流) 及 NE-SW 向 (平行於河流)，注意兩倍的變異圖適合的主要軸 (NW-SE)，是一個計畫引出變異圖有相同"塊" (nugget)，圖 6.12 顯示合併 anisotropy 到插入、產生長、線性條紋平行於河流的影響。儘管明確的變異圖它們無法被設計去與其他地圖作比較，但結果是合理的。

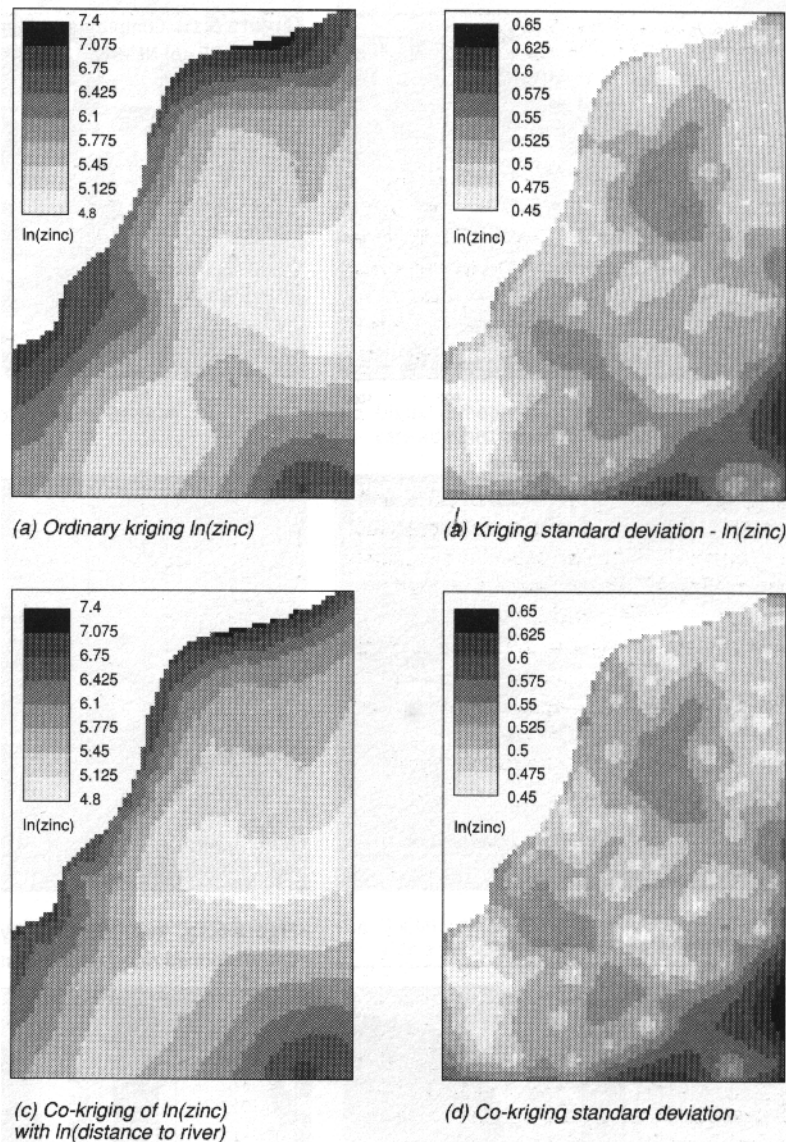


圖 6.10：比較次序性克立金法和使用對數轉換資料的交叉克立金法：(a)次序性克立金法  $\ln(\text{Zn})$ ；(b)克立金標準偏差— $\ln(\text{Zn})$ ；(c) $\ln(\text{Zn})$  的交叉克立金法，帶有  $\ln(\text{到河流之距離})$ ；(d)交叉克立金標準偏差 (原文 p.145)



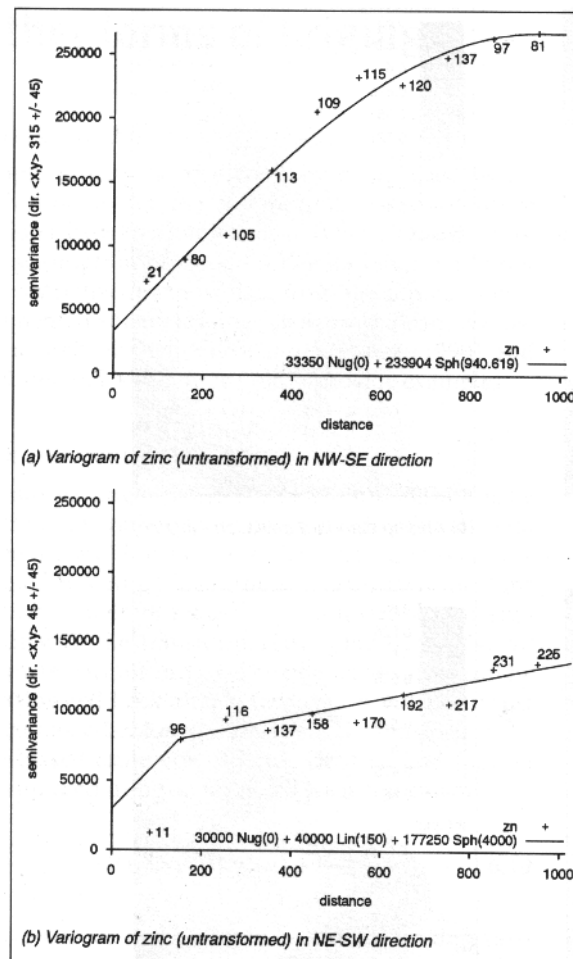
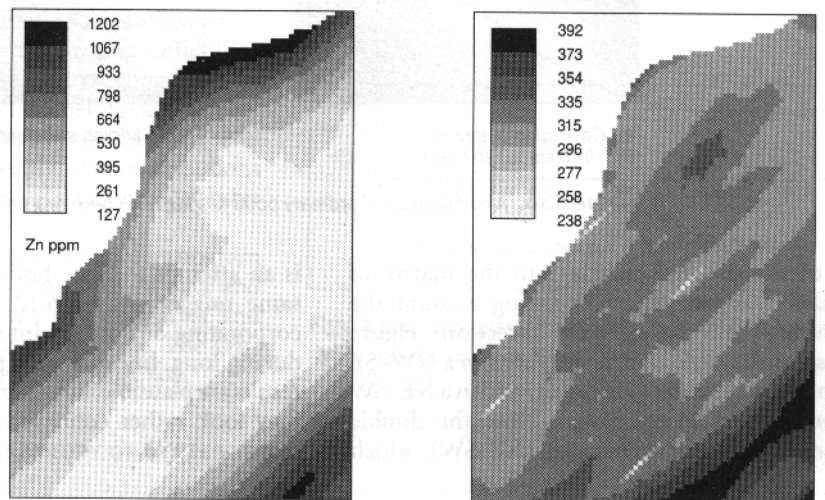
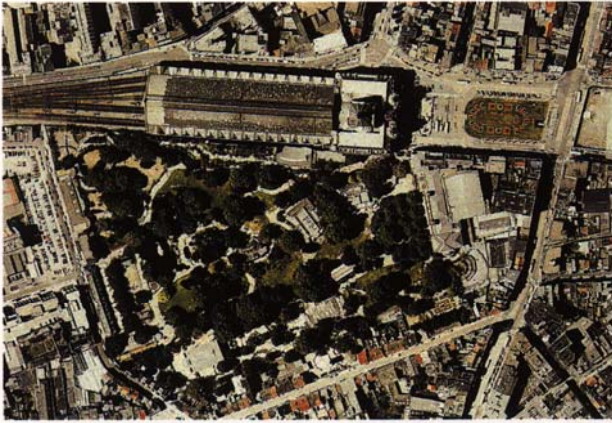


圖 6.11：在不同的方向下所計算之變異圖：(a)鋅（未轉換的）在 NW-SE 方向上之變異圖；(b)鋅（未轉換的）在 NE-SW 方向上之變異圖（原文 p.146）



(a) Zinc levels from anisotropic variogram (b) Anisotropic kriging standard deviation

圖 6.12：使用各向量性變異圖之未轉換鋅資料的次序性克立金法的結果：(a)從各向量性變異圖中的鋅層；(b)各向量性克立金標準偏差（原文 p.146）



1.1 Digital orthophoto of the city of Antwerp, Belgium



1.2 Digital orthophoto of a city with added vector line data to indicate administrative boundaries and data for utility mapping



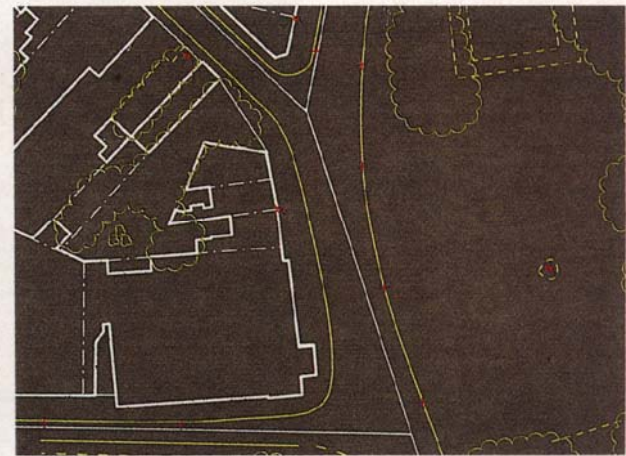
1.3 Digital orthophoto of semi-rural area with motorway including added vector line data



1.4 Digital orthophoto draped over a digital elevation model to provide a perspective view of the landscape



1.5 Large-scale digital orthophoto of urban scene including overlaid vector data on roads and building outlines for a utilities database



1.6 Vector data for the road and building outlines for Plate 1.5

Plate 1：數位化的照片地圖及 GIS（原文 p.146-1）

- 1.1 比利時，安特衛普（Antwerp）的數位化照片。
- 1.2 城市的數位化照片，加入向量式線性資料顯示行政的邊界和資料，以繪製有用的地圖。
- 1.3 包括加入向量式線性資料的高速公路的半鄉村地區其數位化的照片。
- 1.4 覆蓋一數位高層模型的數位化照片，提供地景的視覺呈現。
- 1.5 都市景色的大尺度數位化照片，包括覆蓋向量資料在道路和建築的輪廓線作為有用的

資料庫。

1.6 Plate 1.5 中道路和建築輪廓線的向量資料。

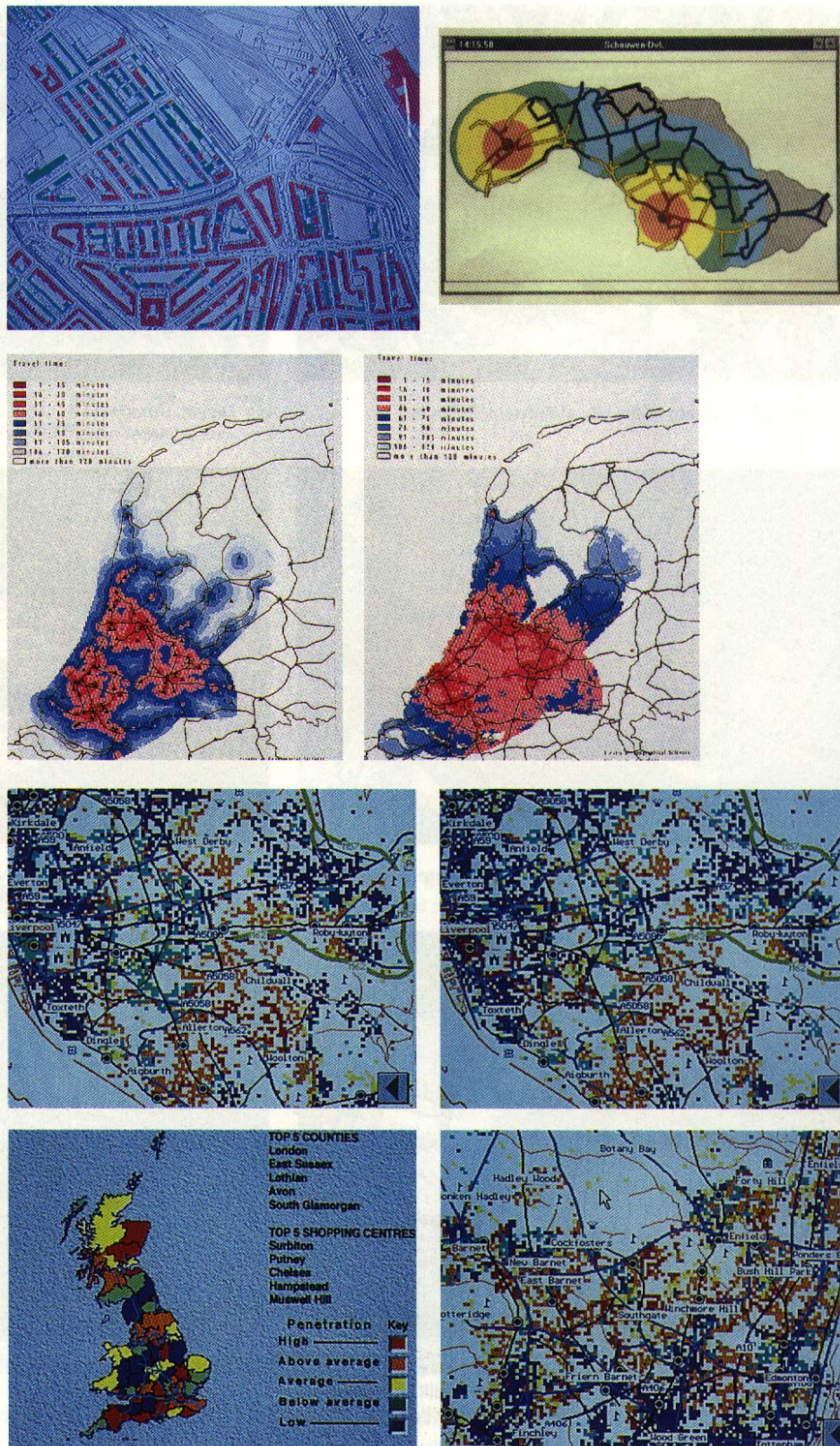
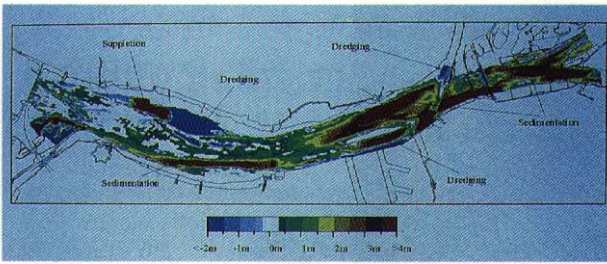


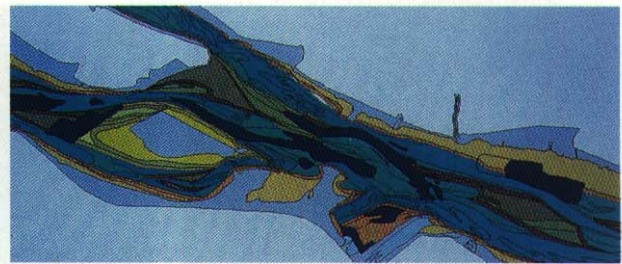
Plate 2：地方自治計畫，路線調查和製造銷售（原文 p.146-2）

- 2.1 Rotterdam 市大尺度網格影像的部分數位化資料庫顯示有用的、地籍的資訊，及產權的細項（用顏色綠和深紅區分）
- 2.2 行進時可藉由烏鴉飛行距離及使用之道路做比較。

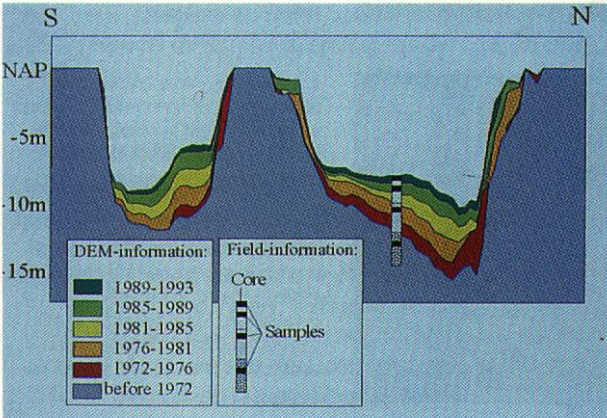
- 2.3 在荷蘭西部工作地區的可及性，以前進時的公共運輸工具來看。
- 2.4 在荷蘭西部工作地區的可及性，以前進時的私人車輛來看。
- 2.5 銷售機會的空間敘述：社會團體 A 在利物浦的位置。（紅色）
- 2.6 銷售機會的空間敘述：在利物浦的威士忌消費（紅色顯示威士忌消費量最大之地區）。
- 2.7 社會經濟團體“時髦的單身者”在英國的分佈。
- 2.8 在倫敦東北部販賣包裝的假期的機會分佈？



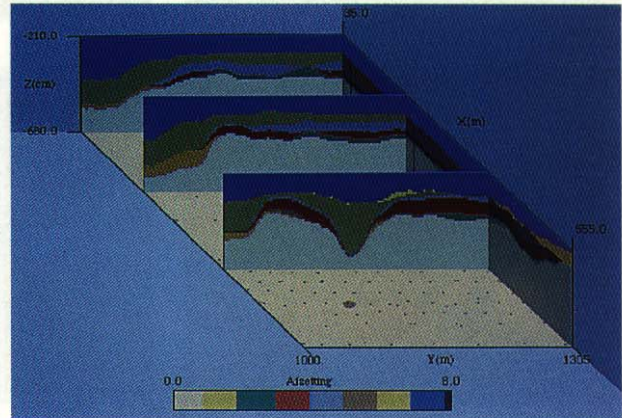
3.1 Changes in channel depths in the mouth of the river Rhine



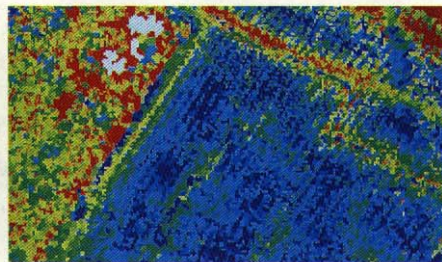
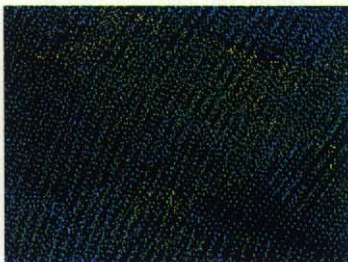
3.2 The Geomorphological distribution of sediments and channels in the mouth of the river Rhine



3.3. Cross-section through sediments of the river Rhine



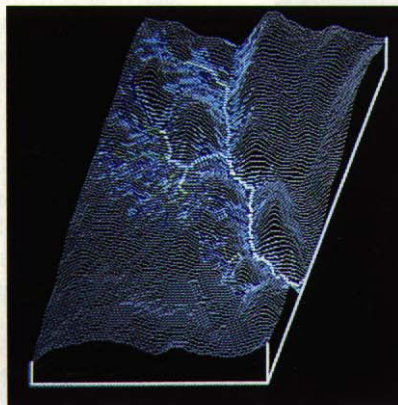
3.4. An example of fence-mapping to show continuous variation of sediment composition in three dimensions



3.5 (far left) The distribution of sampling points of an airborne infra-red laser altimeter scanner

3.6 (left) Infra-red laser scan altimeter data after interpolation to a fine grid. Red indicates high, and blue low elevations

Plates 3.1-3.6 courtesy Lodewijk Hazelhoff and the Ministry of Public Works, The Netherlands



3.7 (far left) Example of local drain direction map created using the D8 algorithm

3.8 (left) Cumulative upstream elements (log scale) draped over the DEM from which they were derived to provide a perspective view of the drainage network

Plates 3.7-3.8 courtesy P. A. Burrough

Plate 3：地形（Geomorphology）和水文學（Hydrology）（原文 p.146-3）

- 3.1 Rhine 河口水道深度之變化
- 3.2 Rhine 河口的沈澱物及水道的地形分佈
- 3.3 Rhine 河口的沈澱物橫切面
- 3.4 柵欄圖（fence-mapping）的例子，顯示在三度空間中沈澱物組成的連續變化。
- 3.5 空中紅外線的雷射高度遙測掃描中取樣點的分佈。

- 3.6 紅外線雷射影像高度測量值在插入適當的格子後，紅色顯示高海拔，而藍色為低海拔。
- 3.7 運用 D8 演算法所產生之地區排水方向圖的例子。
- 3.8 從它們提出之排水網路的現實呈現，累計上游的元素 (log 尺度) 覆蓋過 DEM。

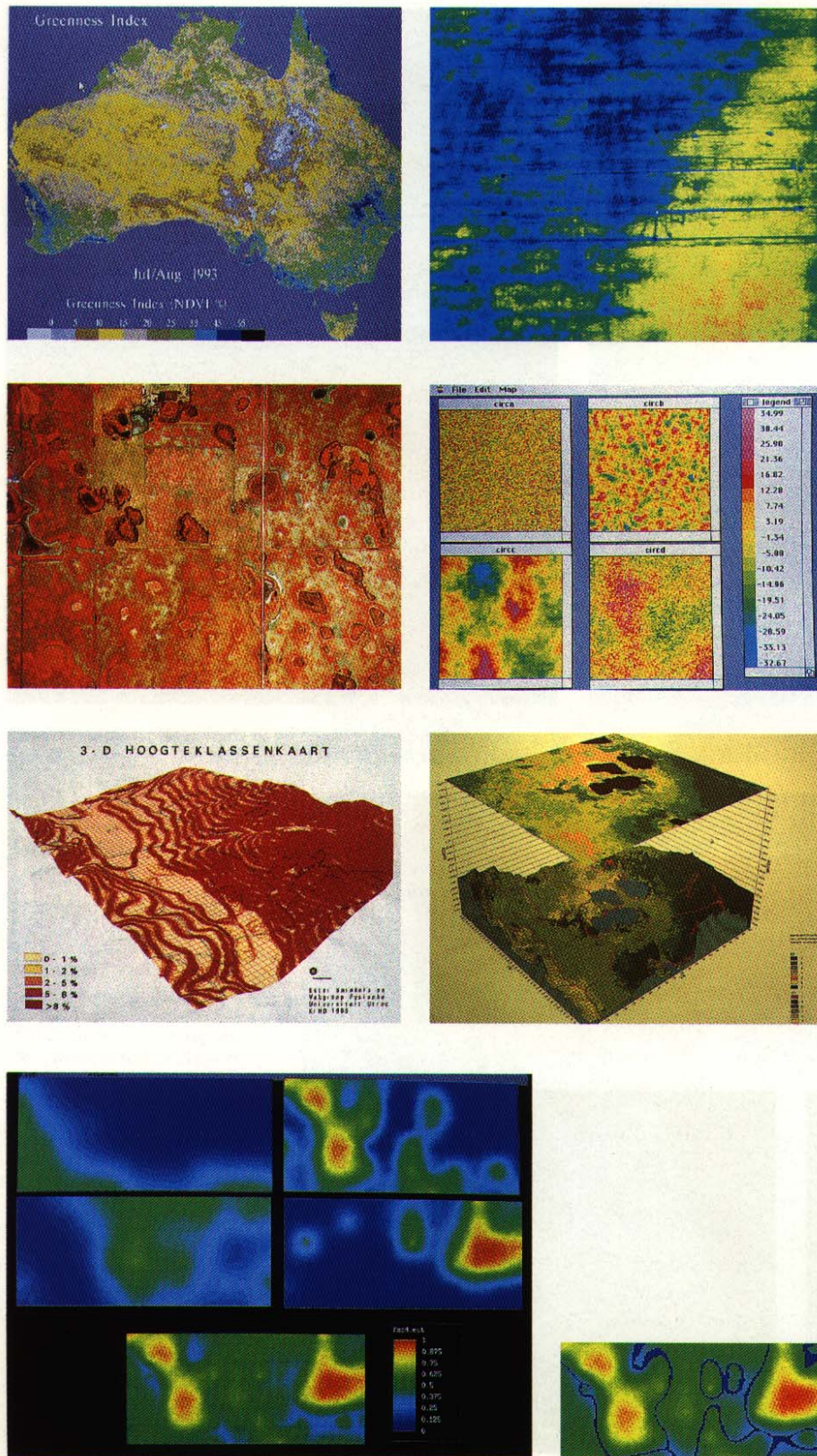


Plate 4：遙測及誤差 (原文 p.146-4)

- 4.1 澳洲，水的綠色指標 (NDVI，標準的差異植栽指標)。

- 4.2 荷蘭 Dronten 有效土壤表面溫度、Minderhoudhoeve、實驗性農田的數位化紅外線影像。
- 4.3 在加拿大亞伯達過去 periglacial 土地表面，其農田的空攝影顏色誤差。
- 4.4 模擬的、個別的連續表面，順時針方向從左上：沒有空間關連（雜音）；右上：短距離的空間關連；右下：長距離之空間關連；左下：長距離的空間關連再加上雜音。
- 4.5 當不適當的插入模式被用於從數位化的連續線中去產生 DEM，則會出現錯誤的大坡度角的虎紋。
- 4.6 從主題圖的衛星影像覆蓋 DEM 推論來說明土地覆蓋資料。
- 4.7 上：4 個模糊土壤層的連續地圖。下：地圖顯示全部 4 個層的最大組成值。
- 4.8 4 個層的最大組成值的地圖，重疊了最大不確定的（邊界）層的區域。

### 使用其他資訊的克立金法(Kriging using extra information)

通常資料點不止是關於鋅分佈的資訊來源，而且是可以運用幫助內插的其他知識。主要的來源有：(a)在清楚而不同的版圖中特有的階層形；(b)來自於品質低劣的量度，更多資料點樣本的同等變化；(c)轉換過程的物理或觀察的空間模式。

### 階層式克立金法(Stratified kriging)

當有足夠的軟體將一個區域分類成有意義的亞區域以及足夠的資料為每個版圖計算變異圖時，則內插會為每個獨立的區域帶來點或面，調整克立金法以避免在分類邊界上不連續的點。根據洪泛頻率分類與鋅濃度不同的分析（見第 5 章，圖 5.2）提出兩個分類的重要性：洪泛分類 1（每年的洪泛）與靜止。圖 6.13 呈現變異圖為此兩個分類的計算。它顯示了兩個單位間空間變化的強烈關聯，但是在洪泛分類 1 的變化幾乎是與少頻率洪泛區域一樣的。圖 6.14 顯示內插將會顯示鋅濃度在分類中的變化。比較層級式克立金法在圖 6.6a、b 的誤差，證明層級化如何導致內插的錯誤，以及許多區域導致依賴資料點的分佈。在這個例子中，層級化也維持了狹路的空間結構，在資料點間平均空間較小的洪泛水道，也因此困難插入點資料而消失。

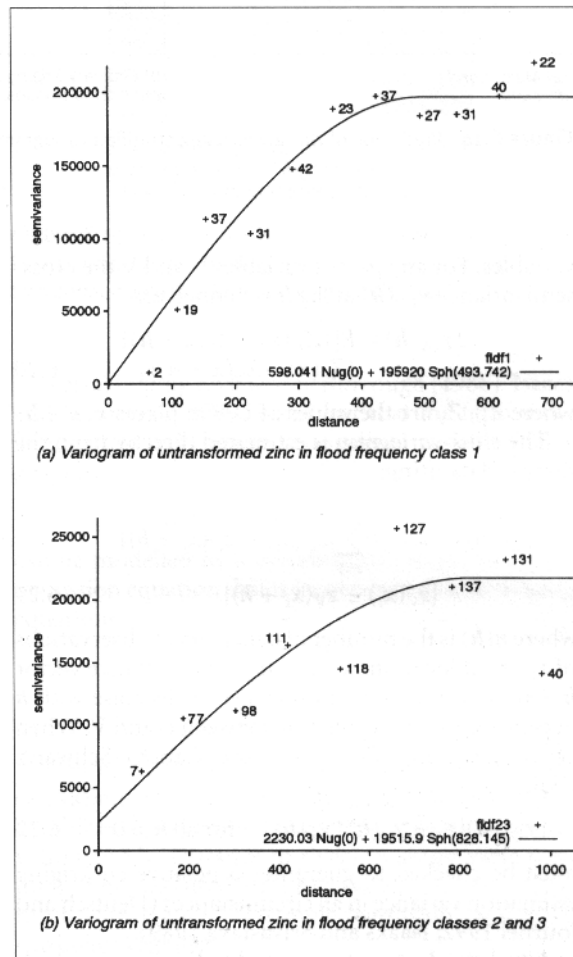


圖 6.13：個別計算氾濫頻率區域鋅質變異圖：(a)氾濫頻率變異圖分類 1；(b)氾濫頻率變異圖分類 2 和 3，但分類 1 基質較分類 2、3 大 10 倍（原文 p.147）

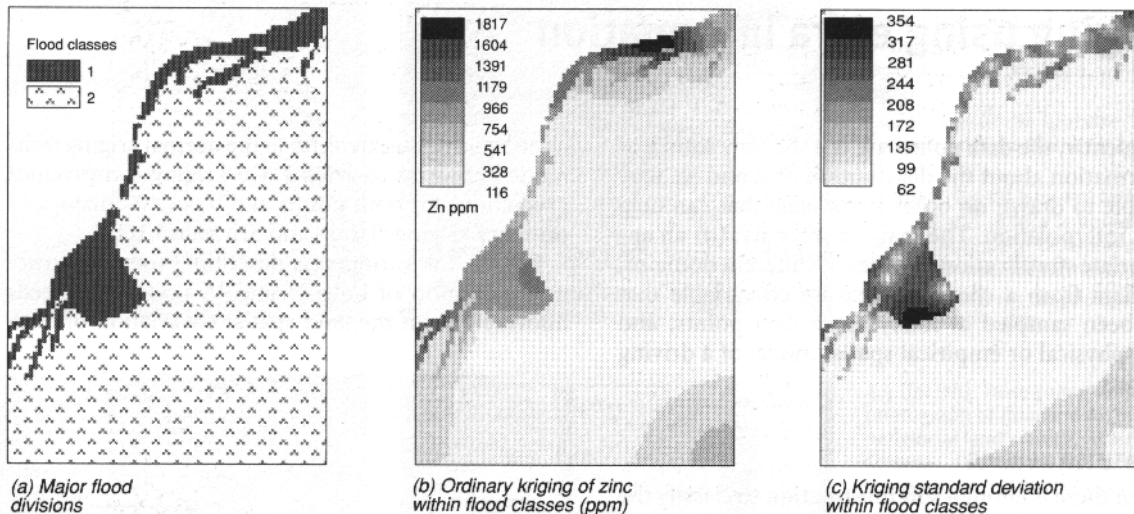


圖 6.14：使用分層變異圖預測鋅的等級：(a)主要氾濫分佈；(b)鋅質氾濫類型的依般克利金法；(c)氾濫種類的標準差克利金法（原文 p.148）

### 整合克利金法(Co-kriging)

通常資料可因每一樣區多於一個的屬性所獲得。當另一組 V 可能是粗劣的量度及更多的觀察時，一組 (U) 可能是費事的量度，也因此是罕見的樣本。如果 U 和 V 是空間關連的話，則使用的 V 的空間變化的資訊來建立圖籍 U 則是可能的。這可由使用標準的克立金法



的延伸技巧如整合克利金法所成。整合克利金法用類比的一般克立金法預估產生點及面 (Isaaks & Srivastava 1989)。

除了變異圖描述U與V無結構的變化外，co-kriging需要在連接空間變化的資訊。任何一組常變的U和V，其cross一半變數  $\gamma_{UV}(\mathbf{h})$  在間隔時間lag  $\mathbf{h}$ 被定義為：

$$2\gamma_{UV}(\mathbf{h}) = E \{ \{Z_U(x) - Z_U(x+\mathbf{h})\} \{Z_V(x) - Z_V(x+\mathbf{h})\} \} \tag{6.20}$$

6.21

6.22

6.23

6.24

6.25

6.26

整合克利金法有個優點，就是如果拙劣量度屬性與費事屬性是資料點的不同數字被測量。如果兩者（或更多）屬性是在同一點上測量，則整合克利金法不會給予一般克立金法不同的判斷。

如果 cross-變異圖是由同組觀察點測量變化之間計算的，便稱之為真 cross-變異圖。如果是在不同位置觀察測量屬性所計算的，就如假 cross-變異圖，因為一個確實的假設的形成，樣本的命名去比較畫分變數的關係。

舉例來說，在第 5 章，方程式 5.9 及圖 5.5 標註了在 ln（鋅）層級與 ln（河流長度）之間強烈與線性的關係。舉例說明，假設有 98 個”河流長度”的測站，但只有 49 個有鋅。圖 6.9 顯示 ln（鋅）及 ln（河流長度）的變異圖以及 cross-變異圖之間的變化。圖 6.20c、d 顯示出 ln（鋅）的 co-kriging 點基於 98 個 ln（河距）的幫助觀察 49 個鋅。比較圖 6.10d 與圖 6.10b 顯示的克立金法圖的內插錯誤與單個 ln（鋅）點克立金法是幾乎相同的。如果隨拙劣量度的變化有強烈的空間關連，這顯示 co-kriging 的可能性是存在費事的實驗分析中。

許多 co-kriging 在土壤圖的例子在 Burrough 1993 列出；Solow and Gorelick（1986）建立潮流。Co-kriging 在變化的格局使用由一般物理過程連接是相當成功的。

### 一般的克立金法—內插的建構趨勢(Universal kriging–interpolation with a built-in trend)

Co-kriging 是不容易執行的，要求考慮地理統計理論的洞察力以及某些專家相信適合、服從假設本身限制的 cross-變異圖的成果與 co-劃分區域的線性模式。一個運用在許多環境科學的方案，是使用觀察（歸化）的知識傳遞模式如克立金法步驟的一部分。有時候，如一般的或結 KT 克立金法般所知 (Deutsch & Journel 1992) 或克立金法在外部趨勢的呈現。現代理論完全取代 Burrough 在 1986 一般克立金法觀點的第一個想法。Isaacs & Srivastava 在 1989 年指出  $m(x)$  在方程式 6.1 中  $z$  在  $x$  中的變化：

6.27

由定數論現代化，如在克立金法中合併的歸化方程式：

6.28

圖 6.15 顯示由一般克立金法程序計算鋅濃度的結果，其程序是結合  $\ln$  鋅在海拔與河距的歸化方程式。鋅單一的變異圖被用在所有的區域。比較圖 5.6（簡易歸化）與圖 6.1（一般克立金法）的結果。

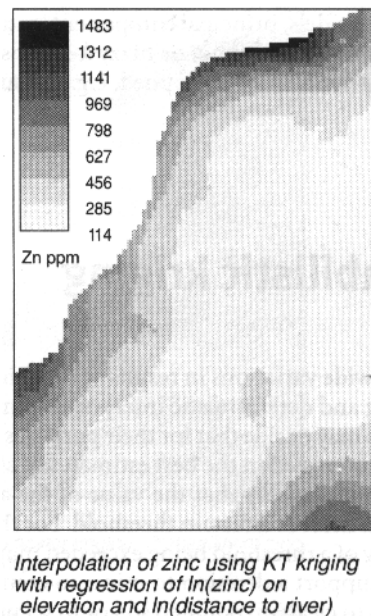


圖 6.15：一般克利金法合併一個趨勢模式以估計鋅的程度（原文 p.149）

### 多變的克立金法(Multivariate kriging)

多變的克立金法是第理統計運用在多變的轉變，如歸化模式的結果，主要的成分變化，平均數的互補或模糊 k-means。某些時間會限制其運用，如所有模糊 k-means 值的總合是等於 1，因此，克立金法方程式需要被修飾。見第 11 章及 Gruijter (1997)。

### 機率性的克立金法(Probabilistic kriging)

根據不同的克利金方法及決定的插入推估方法在結果上會有所不同，一些使用者基於他們的目的而非真正對  $Z(X_0)$  最佳的判斷有興趣，但在機率中，屬性值得問題會提供一個相當重要的開端，為此，其可能性或許足夠支持開礦的決定，而去委託操作者清除受污的土壤或建立行銷基礎。指示克立金法是在一般資料自連續尺度至二進位尺度的形式來自一般克立金法的非線性形式（1 假設  $z$  小於等於  $T_j$ ，和 0 相反）。變異圖是用來計算二進位的資料，隨著轉變資料一般克立金法的程序，這結果對應關係展示範圍 0-1 表示  $T_j$  超出或未超出的可能性的連續資料。計算變異圖和其他開端的內插，提供了洞察開端超出如何變化開端層級的可能性。

圖 6.16 呈現運河中鋅層級在 98 個點測得在 500ppm 與 1000ppm 的指示變異圖。這結果的對應關係（圖 6.17a、b）顯示了鋅層級超出的可能性。這結果清楚且容易了解，也容易與 GIS 的數據資料庫合併，對決定製造經營者也不須了解克立金法的複雜。

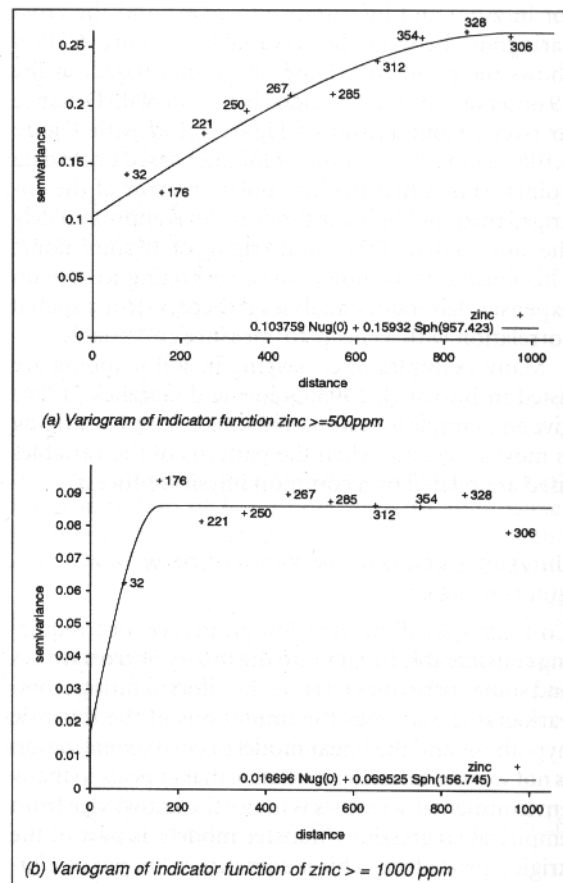


圖 6.16：鋅質變異圖指示：(a)  $> 500$ ppm；(b)  $> 1000$ ppm (原文 p.150)

指示克立金法可與軟體合併，如洪泛頻率的分級：圖 6.17c、d 顯示了鋅由洪泛頻率分級界定區域中超出 500ppm 的機率，以及數字提供一些在洪泛過程及小泛濫周遭受污染有趣的觀察。

有好幾個指示克立金法與指示整合克利金法在土壤科學、土壤受污染與地質學發表的例子。如果原來的資料沒有隨著任何簡易的分佈（高斯 Gaussian 模式或輸入指令標準）分立的克立金法提供一個非線性，分佈依據預估者。當資料不是標準的分佈時，這個理論是嚴格計算，但對一般分佈屬性可能是有用的，如此可以在污染的研究中發現（如 Burrough 1993）。

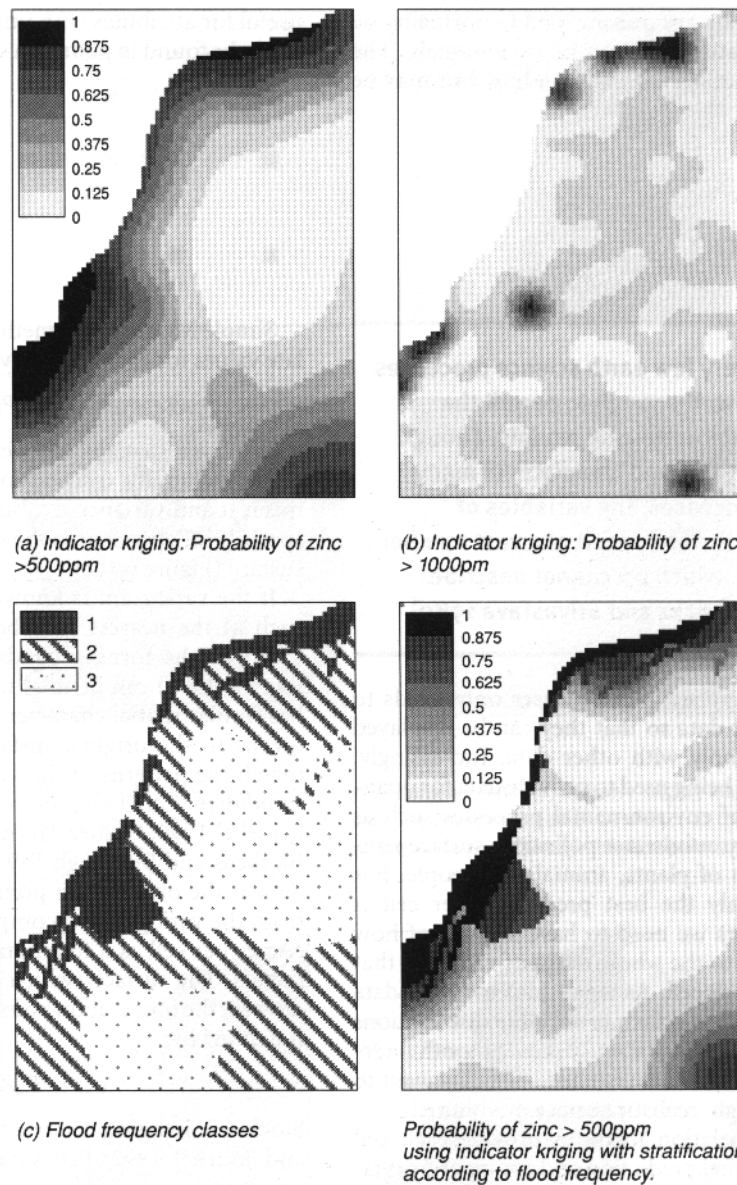


圖 6.17：(a)500ppm 程度的克利金指示；(b)同上於 1000ppm；(c)氾濫頻率分類；(d)氾濫頻率分類 1 和 2+3 的（500ppm）成層克利金指示（原文 p.151）

## 模擬(Simulation)

不幸地，少之又少的地球科學進展被完全充分了解，以至於發展出來的理論無法變成確定值模式（deterministic models）去應用。雖然我們知道須多基本過程的物理、化學概念，主張的多變...是一個大量過程的最終結果...那是我們所無法去定量描述的（Lsaaks and Srivastava 1989）。

對於許多地理資訊系統應用而言，使用者僅需插入點資料（point data）的動作，以便於插入的資料能夠被展現或結合。然而，地理資訊系統漸漸地被使用在為環境處理過程定量模式而提供資料，如氣候改變、地表水污染、地表逕流，動、植物或人類的傳播上。對某些模式而言，對每一網格單元（per cell）作最佳的預測是必要的；我們須有有一個意念，就是在單元中模式如何反應變異物全部範圍是可能的。當對每一單元及其可能的分佈無法預測資料值時，我們必須再分類推估模擬（機率模式 stochastic simulation）。平滑內插法不可能成為

我們想要的：我們可能想要修改一個粗糙的實際表面或量體。

克立金內插法提供每一單元使用一個地區性、理想化的預測，且依賴變異圖及資料結構的標準偏差。為了測定一個提供的數字模式可能如何反應輸入值比地區性的數值大或小，我們可能從原始的內插法及它的標準偏差，計算較高級較低的邊緣表面，而這些表面只提供可能值的平滑推估，及為空間鄰近性重要模式的敏感度評估（地下水模式、沖蝕、逕流水、散佈模式），最好是有在有一獨特及機率相同值的每一單元錯誤值。

簡單 Monte Carlo 法（Simple Monte Carlo methods）對每一個個別的單元能夠被用來模擬數值：

$$z(x) = Pr(Z) \tag{6.29}$$

Z 為一個隨  $\mu$  值及  $\sigma^2$  變數，一個在空間上獨立、正常分佈的機率分佈函數（PDF），但威脅每一單元如同空間獨立性，所以結果是一個固定雜訊表面（stationary noise surface）（圖 6.18c）。

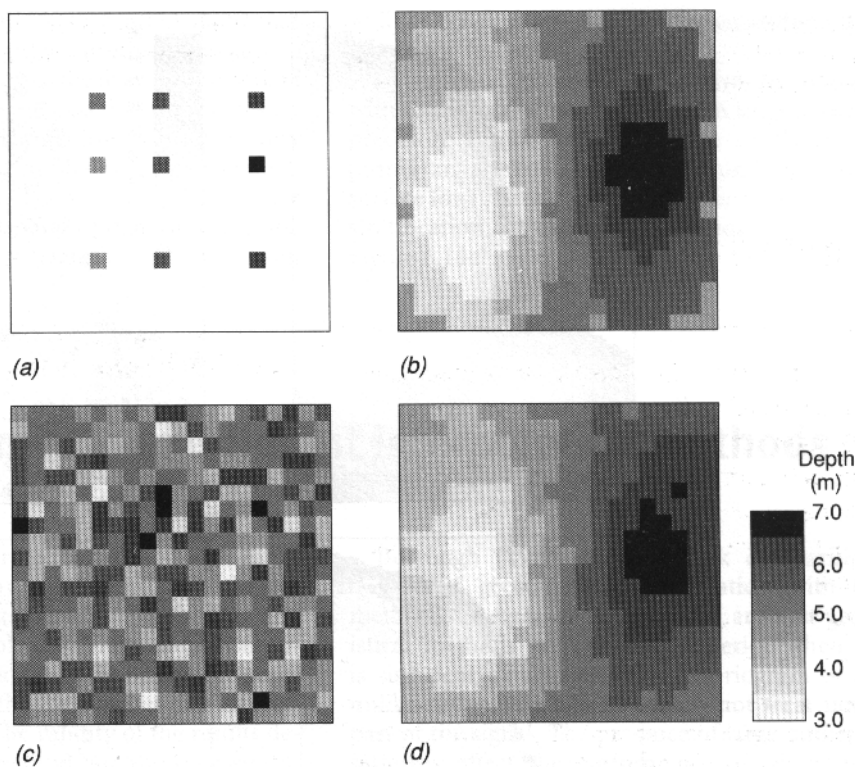


圖 6.18：Monte Carlo 模擬：(a)原始資料；(b)一般克立金圖；(c)擁有相同平均數及變異數的固定雜訊範圍 100 個處理平均值；(d)透過資料點及相同變異圖如原始資料，模擬 100 個處理平均值（原文 p.153）

若可得知變異圖，幾個方法如最鄰近法（the nearest neighbour method）（Heuvelink 1993）或旋結法（the turning bands method）（Deutsch and Journel 1992）能被用來模擬一個具有相似空間特質（nugget 空間推估模式、基質基質、範圍、平均值）到原始面（original surface），但它不是對抗已取樣區域的模型（patterns）。條件模擬相對地正如我們所知的推估想像（stochastic imaging）（Journel 1996, Gomez-Hernandez and Journel 1992, Bierkens and Burrough 1993a, 1993b）結合從變異圖求得的單元最可能結果，如變異圖參數方程式。表面被定義成：

$$z(x) = Pr(Z), \gamma(h) \tag{6.30}$$

推估想像說明如下（Deutsch and Journel 1992; Pebesma and Wesseling（forthcoming））：

1. 建立一個簡單克立金常用的方程式。
2. 隨機選擇一個非樣本的資料。計算克立金的預測及從資料組使用資料的單元位置之標準偏差。
3. 藉由預測及定義標準偏差的機率分佈中，描述一個隨機值。增加這個值到資料點的表列。
4. 重複 2-3 步驟，直到所有單元都已被採用，且完成一個模擬處理（the simulation of one realization）。
5. 重複 1-4 步驟，直到足夠的處理已經被創造出來。
6. 必要時，從所有的處理中計算表面的平均值及標準差。
7. 使用每一個別處理執行環境模式，看看不同的輸入，產生什麼樣的變化結果。

條件模擬比克立金（kriging）或整合克立金（co-kriging）花更多時間計算。例如一般點克立金（ordinary point kriging）（圖 6.6）在 16RAM 及 100Mhz Pentium 的電腦環境中，約需花 1 分鐘處理，但在相同的機器上，對每一個處理而言，條件模擬需花 2.2 分鐘，或者每一百個處理，需花 3.5 小時（即平均 2.1 分/處理）。當變異圖可得知且大部分可能推估時，比起平滑平均值（smoothed averages），填滿資料單元模式化是必須的。

圖 6.18 可看出一般克立金、簡單模擬及條件模模的差異性。對模式的 Monte Carlo 分析而言，通常在一個模擬中至少隨機表面的處理有 500 個，能夠被當作不同的輸入使用（見第 10 章）。圖 6.19 從鍍鋅物（zinc content）不同的 100 種模擬中，說明 4 個處理且對所有 100 個處理而言，圖 6.20 顯示平均值及表準差表面。值得注意的是，依據條件模擬被製成的標準差表面沒有被結合至資料點的分佈，正如同採用克立金。也留意條件模擬的平均表面應該與一般克立金內插法相似。（若足夠的折疊（replicates）被使用，那麼它應該是相同的）。採用比 500 還少的處理，從一般克立金標準差大小持續不同是可能的，且這可能成為計算更少處理的附屬品。

條件模擬也能多為了得到更好精密度而結合軟體資訊。圖 6.20c、d 顯示計算一個平均表面及為 100 個使用處理圖 6.14a 二個洪水頻率年分級（flood frequency classes）的變異圖及層級標準值表面的結果。



圖 6.19：藉由條件模擬的方式，zinc 層級 100 個處理的四種類型（原文 p.153）

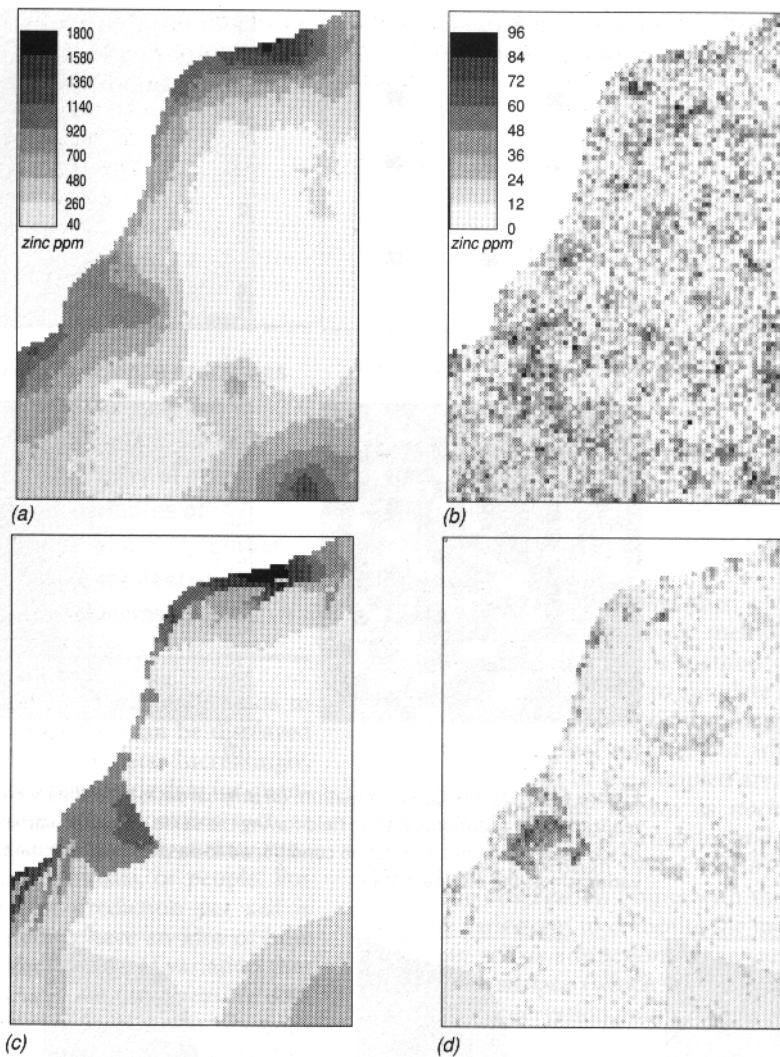


圖 6.20：條件模擬對全區域而言能夠完成，或者對依據洪水頻率分級能夠分層(原文 p.154)

## 不同空間內插推估方法的相對性質(The relative merits of different interpolation methods)

地理統計理論不是一個單位，但一個依據資料基本空間關連結構的了解的技巧的廣泛範圍以及控制內插的使用。理論假設包括不變的概念，本質的假說與普遍性，以及有時難得碰到真實的資料。結果的確實性不只依據理論也假設年域的同值。地理統計理論有提供建立點或面的優點；在空間軸向的資訊，一組變化的尺度，以及外部資訊的組合得到最精確的資料。

Burrough (1993) 把地理統計內插與其他理論作一個比較。一般的結論是地理統計理論在成功的資料去評估變異圖是比較好的，不像曲線，因為這樣的理論不會處理如信號的部分。大型離層的呈現對所有理論的預估的執行具有相當的影響以及其他作者對由勘查資料分析的理論的離層移動的推薦，或健全理論的使用。一般克立金法在陡峭邊界的呈現時，是最小成功的，經由階層化在不同區域可能帶來相當大的改進。普通的地籍圖在預估確實基地質是貧乏的，當空間變化幅度不大以及曲線不明顯的情況下。當資料是貧乏的時候，分類法的應用通常會有最佳的預估（見第 10 章）。

## Box 6.3：克立金法的步驟（原文 p.155）

1. 如果使用指示性克立金法轉換二進位的值（0/1）的話，檢查資料的標準、空間趨勢以及執行特有的轉換。
2. 計算測試性的變異圖以及安裝合宜的模式，如果空間變化比純礦更有關連的話。如果資料不是關連性的，則內插會是不合理的。
3. 檢查橫過確實性的模式。
4. 選擇克立金法或條件式的模擬。
5. 如果克立金法使用變異圖模式去內插位置在不規則網格時，則位置就次會同等於一般的樣本（點克立金）或成一個較大的面（面克立金）。如果條件式的模擬，計算至少 100 的關係在不規則網格，從這些計算的平均值與標準誤差生到表面。
6. 利用網格圖或線輪廓（非隨條件式模擬）展示結果，亦不個別也不懸垂在其他資料圖層（如 DEM）。
7. 入結果於 GIS 中以及使用其結果與其他資料結合。

### 比較第 5 章與第 6 章中不同內插理論的結果(Comparing the result of the different interpolation methods in Ch5 and 6)

在表 5.5，我們比較了內插的定數理論的結果，以及表 6.1 也同樣對地理統計理論進行比較。儘管沒有絕對的比較標準，如資料的效力、確實性，我們註解了定數理論與地理統計理論的結論。通常所有只有從點觀察判斷區域的比例資料的理論（定數論與地理統計）超過臨界假設的值，如洪泛分級與距離/海拔歸化模式。

在克立金法標準誤差容易計算中的理論，似乎顯示成層克立金法由最佳結果所產生（表 6.1、表 6.2）。在一般，來自資料與外部資源洪泛頻率分級較佳的資訊，較低預估的標準誤差。地理統計預估在空間連續變化中減少內插錯誤是有效的；成層的組合，甚至預估錯誤。不論這些結果是確實要求測試資料組，以及在第 10 章之前。這些證據強烈激發使用地理統計與軟體以改進 GIS 預估的動力。

表 6.1：地理統計插入推估結果摘要（原文 p.156）

方法	最小值 (ppm)	最大值 (ppm)	百分比 >500ppm	百分比 >1000ppm	百分比 >1500ppm
Ordinary point kriging with isotropic variogram 等方向性變異的一般點克利金法	119	1661	29.28	4.91	0.36
The same, but with transformation to natural logarithms 同 上，但伴隨自然對數轉變	122	1348	20.24	1.45	0.00
Ordinary point kriging with anisotropic variogram 非等方向性變異的一般點克利金法	127	1202	31.43	3.92	0.00
Ordinary point kriging within major flood categories 內部 具有主要氾濫分類的一般克利金法	116	1817	15.56	3.53	0.85
Co-kriging on ln (distance to river) 整合克利金法(與河流的距離)	144	1300	16.20	0.97	0.00
Universal point kriging with a regression model on ln (distance) and elevation 具有距離或高層回歸模式的通用克利金法	114	1483	18.89	2.33	0.00



Conditional simulation with one general variogram 具有一般變異的有條件模擬	140	1606	28.89	3.96	0.21
Conditional simulation with stratification according to flood frequency and 2 variograms 根據氾濫頻率及 2 變異值而成層的有條件模擬	105	1800	15.11	3.53	0.86

表 6.2：不同方法標準偏差的比較

方法	最小標準差	最大標準差
分類 1 的氾濫頻率圖	423	423
分類 2 的氾濫頻率圖	177	177
分類 3 的氾濫頻率圖	105	105
等方向性變異的一般點克利金法	119	329
非等方向性變異的一般點克利金法	238	392
具有主要氾濫分類的一般點克利金法	62	354

表 6.3：比較內插法 A comparison of methods of interpolation (原文 p.158-159) (可參閱講義 p.119)

說明 Method 方法	Deterministic /stochastic 決定/推估 (確率)	Local/ Global 地區性/全球性	Transitions abrupt/gradual 過渡突然/漸進	Exact Interpolator 內插法	Limitations of the procedure 限制	Best for 最佳	Computing load 電腦使用	Output data structure 輸出資料結構	Assumptions of interpolation model 內插值模式假設
Classification 分級；分類	Deterministic "soft" information 確定性"軟體"資訊	Global 全球性	Abrupt if used alone 屬突然，若單獨使用	No 無	Delineation of areas and classes may be subjective. Error assessment limited to with-class standard deviations. 區域及層級的描繪可能主觀。 錯誤評估局限於有層級標準偏差。	Quick assessments when data are sparse. Removing systematic differences before continuous interpolation from data points. 當資料是很少時，快速評估。在連續內插值前，從資料點中移除系統差別。	Small 小	Classified polygons 已分類的多邊形	Homogeneity within boundaries 邊界內均質
Trend surface 趨勢表面	Essentially deterministic (empirical) 本質上可決定的 (經驗法則)	Global 全球性	Gradual 漸進式	No 無	Physical meaning of trend may be unclear. Outliers and edge effects may distort surface. Error assessment limited to goodness of fit. 趨勢實質意義可能不清楚。 外層和邊元效應可能使表面變形。 錯誤評估局限於適合處。	Quick assessment and removal of spatial trends. 空間趨勢的快速評估及移除。	Small 小	Continuous, gridded surface 連續網格表面	Phenomenological explanation of trend, normally distribute data 正常趨勢分佈資料的現象解釋。 (常態分佈資料)
Regression models 回歸模式	Essentially deterministic (empirical-statistical) 本質上可決定的 經驗統計)	Global with local refinements 視當地 (地區性) 而修正	Gradual if inputs have gradual variation 屬漸進式，若輸入有漸進變化	No 無	Result depends on the fit of the regression model and the quality and detail of the input data surfaces. Error assessment possible if input errors are known. 結果取決於回歸模型的適宜性和輸入資料表面的質及細節。 錯誤評估可能發生，若可得知輸入錯誤處。	Simple numerical modeling of expensive data when better methods are not available or budgets are limited. 當更好的方法無法獲得或預算受限時，昂貴資料的簡單數字模式化。	Small 小	Polygons or continuous, gridded surface 多邊形或連續網格表面	Phenomenological explanation of regression model 回歸模型的現象解釋。
Thiessen polygons (proximal mapping) 徐昇多邊形 (近似製圖法)	Deterministic 可確定性	Local 地區性	Abrupt 突然式	Yes 有	No error assessment, only one data point per polygon. Tessellation pattern depends on distribution of data. 沒有錯誤評估，僅是每一多邊形的一個資料點。棋盤鑲嵌模式取決於資料的分佈。	Nominal data from point observations. 從點觀測中的名義資料 (Nominal data)。	Small 小	Polygons or gridded surface 多邊形或網格表面	Best local predictor is nearest data point 最佳地區性預測器 是最接近的資料點。

(續下表)

表 6.3：比較內插法 A comparison of methods of interpolation (原文 p.158-159) (可參閱講義 p.119) (續一)

說明 Method 方法	Deterministic/ stochastic 決定/推估 (確率)	Local/ Global 地區性/全球性	Transitions abrupt/ gradual 過渡突然/漸進	Exact Interpolator 內插法	Limitations of the procedure 限制	Best for 最佳	Computing load 電腦使用	Output data structure 輸出資料結構	Assumptions of interpolation model 內插值模式假設
Pycnophylactic interpolation Pycnophylactic 內插法	Deterministic 可確定性	Local 地區性	Gradual 漸進式	No, but conserves volumes 無, 但有轉化量	Data inputs are counts or densities 資料輸入是計數 (用算的) 或密集(度)	Transforming step-wise patterns or populations counts to continuous surfaces. 轉變步驟-配對模式或者族群連續 表面計算。	Samll – moderate 小-中等	Gridded surface or contour 網格表面或等高線	Continuous, smooth variation is better than ad hoc areas 連續平滑變化比特別的 (ad hoc) 區域好
Linear Interpolation 線性內插法	Deterministic 可確定性	Local 地區性	Gradual 漸進式	Yes 有	No error assessment 沒有錯誤評估	Interpolating from point data when data densities are high as in convert- ing gridded data from one projection to another. 當資料密度很高時, 如同轉換網格 資料從一個投影到另一個投影, 從 點資料內插資料密集(度)。	Samll 小	Gridded surface 網格表面	Data densities are so large that linear approximation is no problem 資料密集(度)很大, 所以線性近似值 (inear approximation) 可獲 得。
Moving averages and inverse distance weighting 移動加權平均法	Deterministic 可確定性	Local 地區性	Gradual 漸進式	Not with regular smoothing window, but can be forced	No error assessment. Results depend on size of search window and choice of weighting parameter. Poor choice of window can give artifacts when used with high data densities such as digitized contours 沒有錯誤推估。結果取決於 權重參數及視窗搜尋的大 小。當使用如數化等高線 的高資料密度時, 視窗的選擇 能夠給人為的製圖成品。	Quick interpolation from sparse data on regular grid or irregularly spaced samples. 在有規律的網格或不規則的空間樣 本上, 從很少資料中快速內插值。	Samll 小	Gridded surface, contour lines 網格表面, 等高線	Underlying surface is smooth 下表面是平滑的
Thin plate splines Thin plate 曲線法	Deterministic with local stochastic component 隨地區性推估的確定 性 (機率性值的分佈 值)	Local 地區性	Gradual 漸進式	Yes, within smoothing limits 有, 但受平滑法 限制	Goodness of fit possible, but within the assumptions that the fitted surface is perfectly smooth. 適合可能的優勢, 但必須在 合適表面完美地平滑的假設 中。	Quick interpolation (univariate or multivariate) of digital elevation data and related attributes to create DEMs from moderately detailed data. 數化高程資料的快速內插法(單一 變化或多重變化)及從適當地細部 資料關聯屬性建立 DEMs。	Samll 小	Gridded surface, contour lines 網格表面, 等高線	Underlying surface is smooth everywhere 下表面全是平滑的。

(續下表)

表 6.3：比較內插法 A comparison of methods of interpolation (原文 p.158-159) (可參閱講義 p.119) (續二)

說明 Method 方法	Deterministic/ stochastic 決定/推估 (確率)	Local/Global 地區性/全球性	Transitions abrupt/ gradual 過渡突然/漸進	Exact Interpolator 內插法	Limitations of the procedure 限制	Best for 最佳	Computing load 電腦使用	Output data structure 輸出資料結構	Assumptions of interpolation model 內插值模式假設
Kriging 克立金法	Stochastic 推估性	Local with global variograms. Local with local variograms when stratified. Local with global trends. 屬地區性，但視全球變異圖而定。屬地區性，但當分層時，視全球變異圖而定。屬地區性，但視全球趨勢而定	Gradual 漸進式	Yes 有	Error assessment depends on variogram and distribution of data points and size of interpolated blocks. Requires care when modeling spatial correlation structures. 錯誤推估取決於資料點的變異圖和分佈以及推估模組的大小。當模式化空間關聯結構時需要小心。	When data are sufficient to compute variograms, kriging provides a good interpolator for sparse data. Binary and nominal data can be interpolated with indicator kriging. Soft information can also be incorporated as trends or stratification. Multivariate data can be interpolated with co-kriging. 當資料足夠去計算變異圖，克立金為很少的資料提供一個好的內插值計算法。二位元及名義資料能夠以使用指示克立金的內插法被計算之。軟體資訊也能夠合併為趨勢或分層資料。多變量資料也能夠使用 co-kriging 內插法計算之。	Moderate 中等	Gridded surfaces 網格表面	Interpolated surface is smooth. Statistical stationary and the intrinsic hypothesis 內插值表面是平滑得。統計固定 (Statistical stationary) 和內在假設。
Conditional simulation 條件模擬法	Stochastic 推估性	Local with global variograms. Local with local variograms when stratified. Local with global trends. 屬地區性，但視全球變異圖而定。屬地區性，但當分層時，視全球變異圖而定。屬地區性，但視全球趨勢而定	Irregular 無規律性；不規則	No 無	Understanding of underlying stochastic process and models is necessary. 推估過程和模式的理解是必要。	Provides an excellent estimate of the range of possible values of an attribute at unsampled locations that are necessary for Monte Carlo analysis of numerical models, also for error assessments that do not depend on distribution of the data but on local values. 在對於數字模型的蒙特卡羅分析 (Monte Carlo analysis of numerical models：一種確率模擬的模式) 必要的非取樣位置提供一個屬性可能值範圍的極好推估，也為錯誤評估不僅僅取決於資料的分佈也包括地區性價值 (local values)。	Moderate- Heavy 中等-重要的	Gridded surfaces 網格表面	Statistical stationary and the intrinsic hypothesis 統計固定 (Statistical stationary) 和內在假設。

## 為求得理想的取樣結果，使用變異圖(Using variogram to optimize sampling)

本章已針對內插法及其優缺點作一描述。無疑地，就算是最佳的方法也受限於資料，且雖然許多地理資訊使用者都不可能親自蒐集它們所需的資料，若資料不同數量或資料點安排不可能得到很好的資訊，對使用者而言，這樣的方式是值得考慮的。有關這個概念在第 10 章會再討論；而在本章，我們僅是探討如何將變異圖相關樣本空間的位置安排（及配置上）至克立金空間推估法（block kriging）的單元大小。

在一般模組克立金（方程式 6.17 及 6.18）中， $Z_B$  的預測錯誤（the prediction error）受限於變異圖及取樣的結構（sampling configuration）。因此，一單變異圖可得知，那麼導致在任一必要最小錯誤內插值的取樣步驟就可能設計出來。特別是土地 B 模組所依賴的「預測  $\sigma_B$ 」：

- (i) 變異圖的形式（線性、球面、或其他函數），非等方性或非標準（non-normality）的存在以及 nugget 空間推估模式變異數或雜訊的量。
- (ii) 被用來計算點或模組推估的鄰近資料點的數量。
- (iii) 取樣結構—最有效率的一不規則取樣，或一個規律性方格網（regular square grid）或三角形網格（triangular grid）。
- (iv) 推估已做好的土壤模組大小—是一個區域等值到一個支持點或較大的土地模組？
- (v) 如何將取樣點安排至模組。

依據 Burrough (1991a)，這些點已得知，是以 Webster 及其研究群（已於 Webster 及 Oliver 1990 年報導）最初的工作為基礎；本章中，我們僅考慮第 (ii) 點及第 (v) 點，即是在一個規律性網格上鄰近資料點的數量及取樣空間位置需要對提供的模組取得一個最大  $\sigma$  值。

一旦變異圖已知，便計算模組大小的結合及在規律性三角形的取樣空間位置上或正三角形網格上到使用以提供的預測錯誤，取得模組平均的預測。一個模組 B 的預測變異數介於模組平均數的克立金預測及真實質（the true value）是可預期的平方根差分。模組平均值的標準錯誤是模組 B 預測變異數的平方根（square root）。變異數的預測可由下列方程式取得：

$$\sigma_B^2 = E \left[ \left( Z_B - \hat{Z}_B \right)^2 \right] = 2 \sum_{i=1}^n \lambda_i \gamma(X_i, B) - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(X_i, X_j) - \gamma(B, B) \quad 6.31$$

$\gamma(x_1, x_2)$  是介於  $x_i, x_j$  點的半變異值，介於其中（及在不對等的角度）考慮  $x_i$  至  $x_j$  的距離， $\gamma(x_i, B)$  介於  $x_i$  及模組所有點的平均半變異值，且  $\gamma(B, B)$  是在提供（如在模組變異數內）的平均半變異值。

從 6.31 方程式中，可得到變異數的預測而不是 constant，但依賴模組 B 的大小，變異圖的形式，及介於資料點之間的距離（如在關鍵模組被推估的取樣點結構）。值得注意的是這些變異數不是依賴已觀測數值本身（透過變異圖除外）。對不同模組大小及平方網格（square grid）空間位置等，我們能夠計算  $\sigma_B$  的數值（見圖 6.21）。這些曲線能幫我們決定取樣的強度（intensity of sampling）及樣本空間位置安排。嘗試從點觀測到收集如網格平均的資料連結可能有用（如遙測資料）或真實資料被用來驅動使用已決定空間解析度的一個質量處理模式（quantitative process model）。

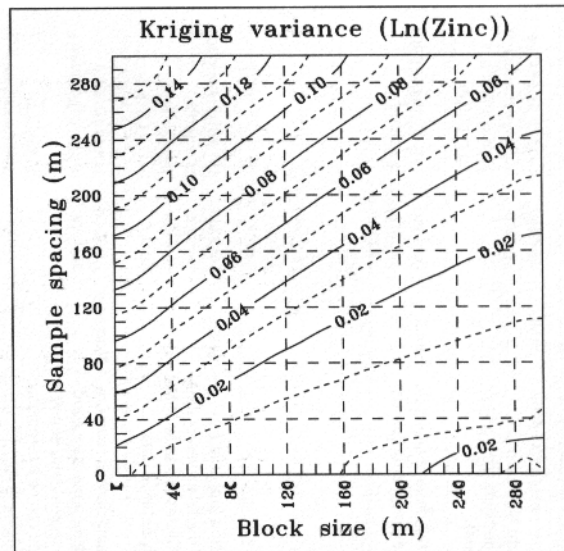


圖 6.21：對不同樣本空間位置及模組大小的作用而言，等值預測變異值得等值線點的克立金變異值（原文 p.160）

### 當變異圖未知時，採用理想化取樣(Optimizing sampling when the variogram is unknown)

若變異圖無法知道時，藉由巢狀取樣法，變異圖可能可以被推估的很近似（Webster and Oliver 1990）。簇群取樣被安排在樣本彼此間配對位於短距離的方法；這些配對在一個較大且已知的距離部分等等。變異數的巢狀分析被用來推估平均樣本空間位置的漸增變異數（cumulative variance）變化，藉由 Miesch（1975）推估的變異圖已得知。巢狀取樣對內插法而言，不足以提供變異圖的推估，但可顯示巢狀取樣的結果對規律性製圖而言，藉由內插法能夠指示理想化近似樣本空間位置（near-optimal sample spacing）。因此，若變異圖不得而知，理想化取樣（optimizing sampling）可以在規律性取樣前，更有效決定最佳的樣本空間位置。藉由刪除一些相近樣本空間位置重複處的不平均法（unbalanced approach）可以使得巢狀取樣更有效率。

### 地理統計內插值軟體的來源(Sources of software for geostatistical interpolation)

本章中介紹的所有內插值及模擬，可於 Gstat 軟體（Pebesma 1996）及 PCRaster 軟體（Wesselin *et al.* 1996）中使用得到。其他便宜或免費地理統計軟體，可從如 Gstat、GSLIB（Deutsch and Journel 1992）以及於 Varekamp（1996）等可以獲得。某些統計軟體（如 SPLUS, GENSTAT 等）也包含地理統方法。Pannatier（1996）出版一組以視窗為基礎的頂級交互式程式.....。而在一般地理資訊系統軟體的使用中，我們也建議讀者小心不完全文件說明的軟體（incompletely documented software）！

### 問題(Questions)

1. 說明為何在克利金（kriging）中變異圖（variogram）有一個好的模式？
2. 比較在推估（interpolating）海拔高度資料製造一個 DEM 的方法中，一般點克利金

(ordinary point kriging) 及薄曲線 (thin plate splines)。

3. 檢視
4. 討論使用指標克立金，從社會調查法或植被研究中去推估現有或本來不存在的資料的方法。
5. 說明如何使用模組克立金來確認點資料已推估至有相同空間位置及空間概化層級的網格，如 Thematic Mapper 遙測影像。
6. 說明決策者為何使用一個不確定的已知層級 (a known level of uncertainty) 具內插退估資料的價值。

## 第七章 空間中不連續實體的分析

### *The Analysis of Discrete Entities in Space*

#### 前言

地理資訊系統的目的除了建立數位化描述的地理事件之資料庫外，還需要提供選擇、擷取與分析它們的工具。本章節闡述了一些方法有效的來關連空間中鬆散的單元實體——就這些實體的屬性而言，它們如何在資料庫中被選取；如何使用布林邏輯與數學原則運算新的屬性來產生有用的群體或分類，或是用來推及複和的圖像。許多像這樣的方法並非真實空間性質的，因為它們只影響了屬性，而不包含大小、形狀或空間實體的形式，像是可以被地理的基本單元如點、線、面或是像素。空間分析起始於空間包含與排除部分的界定，同時藉由不同種類線與面的交互關係來產生新的實體。空間交互作用並不僅侷限於現存實體的邊界，還包括了鄰近地區的功能，像是烏鴉飛行的路程、地形接近度與網絡上的距離，例如道路或河流距離。這些程序是被附上插圖的，例如透過氣象學、考古學、地理人口統計學、土地評估與規劃。

在討論過蒐集資料與建立空間資料庫的問題後，下一個議題是如何使用這些資料提供資訊來解答真實世界的問題。這涉及了從簡單的資料擷取與展示的資料操作理論的廣幅來建立與應用複雜的模式於不同的規劃情形的分析與比較。

一些轉換的功能，像是這些必須的資料清除與更新於改變尺度、預測或內插推估，在第4-6章已被討論。地理資訊系統在許多用途方面提供了大幅度的分析能力。這些分析能力通常被組織成模組化的指令，因此每一種分析可以個別的執行，或是與其它的分析共同建立資料分析模式。事實上，使用者介面可以被提供於指令類型中，視窗上預設程式模組的按鈕，或是使用者自行描述性的寫入高階的程式語言。本章與接下來的章節主旨不在於告知讀者如何撰寫程式模組，但是對於可利用的資料分析功能與種類提供適當的瞭解以便完成欲進行的工作。

資料分析的一般性問題見圖7.1。資料庫包含的資訊可以被使用於回答使用者的問題，同時將提供答案於地圖、表格與附圖上。欲回答這些問題，需要正式的設定資料擷取與分析方式來召回資料，估算新資訊與展示結果。本節接下來的部分是有關於這些正式的連結如何被界定與使用。

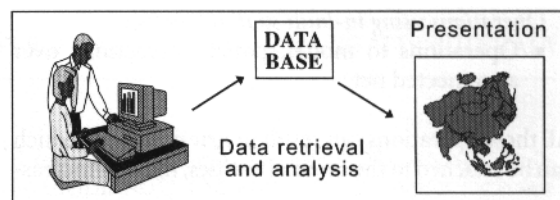


圖7.1：資料擷取是預測資料的首要步驟（原文p.163）

#### **Spatial analysis is more than asking question**

可以用於空間資料的分析方法的種類，主要取決於已經被使用的資料模式與描述。有一點很重要的是，不同的空間資料模式與描述可以制定不同的空間問題。其基本問題是，不管基本資料模式是歸因於空間內的實體或是超越空間的屬性連續變化。在案例的實體中，資料



擷取與分析涉及屬性、位置與實體的連續性與計算它們的空間分佈；在案例中連續的田野，資料分析涉及了田野的空間特性。但是所產生的問題是，連續的田野被設定的三角網或四方網格所分割，同時單獨的三角或正方網格 (regular grid) 也可以被當作是獨立的網格 (grid cell)。

本章節主要集中在資料分析的理論，它們大多是有用的且與空間實體連結，不管是與物件導向的理論或是空間中連續實體分析 (見第八章) 的關連。資料模式的基本原則與分析呈現於第二章。本章與接下來的章節證實了這些原則的應用與它們如何被轉化為電腦指令來解釋實際的問題。

## 空間分析操作的基本層級(The basic classes of operations for spatial analysis)

物件的實體模式在空間中有三個重要的資訊，便是它是什麼？它在哪裡與它跟其它實體產生何種關係？實體的本質是藉由屬性而定，包括地理位置、座標與不同實體間的接近度與連接度 (connectivity) 之空間 (topological) 關係。區位觀點、接近度與從多種類形資料中透過地形識別的地理資料。這些已是資訊系統中的例行操作程序。

我們整理出下列實體資料分析操作的基本層級：

### 屬性操作

- 實體的單一或多個屬性操作。
- 多實體空間疊圖的單一或多個屬性操作。
- 操作單一或多個屬性的連結。

### 距離／位置的操作

- 找出實體間的歐幾理德距離與區位規範。
- 建立實體的緩衝區域。
- 使用內建的空間位向來操作。
- 模式空間與網路交互作用的操作。

## 地理實體性質的屬性資料之操作(Operations on the attributes of geographic entities)

如同第二章的說明，屬性是實體的特性，它們說明了這個實體是什麼。這些屬性可以被分成三個類型—歸類為區位的 (緯度、經度與海拔的地理屬性)；被簡單附加定性與定量的非空間特性之符號；與從實體自身的空間特性所得來的。例如地塊的編號、所有權人的名字與一塊土地的覆蓋形狀之非空間特性。圍設的柵欄與道路的距離、面積、形狀連接度等屬性是由一塊土地的形式所取得的。

如同協定的資訊系統，新的屬性可以被附加於實體上如同資料庫操作的結果。例如一個新的屬性 (或是一個屬性值) 可以被估算大於原本給定的面積大小的土地片段，或是這些土地擁有人居住以外的區域。在資訊的展示方面，新的屬性可以透過上色與加上符號來表現地圖上實體的種類。新屬性的取得可以透過合理的邏輯模式與數學分析，包括操作實體的接近度與位向關係特性。當屬性的設定附加於某個符合搜尋原則的實體，簡單資料擷取可以被視為建立一個臨時的屬性“選擇”。

選擇的程序或新屬性的建立可以被形式化如下：x是任何給定的位置，屬性質 $U_i$ 由下列是子取得：

$$U_i = f(A, B, C, D, \dots) \quad 7.1$$

式中A、B、C.....用來估算 $U_i$ 的屬性質。函數 $f()$ 可以被下列單一或組合的因素帶入：

- (a) 邏輯（布林）運算。
- (b) 簡單與複雜的算數運算與數值模式。
- (c) 單變量統計分析
- (d) 多變量統計理論與貝氏（Bayesian）統計於分類和區別
- (e) 多準則理論（Multicriteria methods）、基於人工智慧（AI-based）、（類）神經網絡（neural networks）。

### 單一或多個實體之屬性的邏輯（布林）運算

實體的屬性可以透過附加於資料庫（如SQL）中的程式語言，進行布林代數而加以更正與重新分類。布林代數使用了邏輯運算—交集（AND）、聯集（OR）、XOR與差集來決定特定的狀態是真實的或錯誤的（見Box 7.1）。運算元AND（符號 $\wedge$ ）指包含所有實體的交叉區域；OR（符號 $\vee$ ）指包含所有實體的組合區域；NOT（符號 $\neg$ ）只包含某個指定實體而不包含其它實體的區域；XOR（符號 $\underline{\vee}$ ）指所有指定實體的聯集來差集所有實體的交集部分（圖7.2）。值得注意的是，邏輯運算可以被應用於所有類型的資料，不論是布林（Boolean）、名目上的（nominal）、順序上的（ordinal）、數值（scaler）或是方向上的（directional）資料。

Box 7.1：數學運算於改變屬性資料（原文p.165）

- (a) 邏輯運算  
兩個以上的單元使用聯集、交集、差集、XOR獲得真實或錯誤（0或1）的結果。
- (b) 算數運算  
透過加、減、乘、除、升冪、降冪、取對數（以10為底的自然對數）、刪減、平方根等方法求出新屬性的結果。
- (c) 三角函數運算  
新屬性是sin、cos、tan等三角函數，或是它們的反函數（arcsin、arccos、arctan），或是從弧度等級來改變或grad representation。
- (d) 資料類型操作  
新屬性是原有屬性的呈現，像是不同的資料類型（布林的、名目的、順序的、方向的、整數的、實數的或是位向資料類型）。
- (e) 統計運算  
新屬性是被給定實體描述的屬性的平均、眾數、中位數、標準差、便異數、極小值、極大值、幅度、歪度（skewness）、峰度（kurtosis）等。
- (f) 多變量運算  
新屬性透過多變量迴歸運算。  
新屬性透過實質程序的數值模式運算。

新屬性透過主成分分析、因素分析來運算。

相似性分析於多元資料的轉換。

透過多元數值分類 (multivariate numerical taxonomy) 理論把實體歸類成既定的分類 (新屬性等於分類名稱)。

透過判別分析、最大概似法或貝氏技巧 (Bayesian techniques)，把實體指定成以統計機會為基礎的或然率。

實體在給定的條件下，被指定成模糊集合的值。

使用神經網絡理論將實體指定成為一個分類。

兩個簡單的例子說明了這個理論。首先考量空間資料庫被使用於真實財產，代表A典型擷取問題於買家a，其可能強況如下：請顯示出所有房屋的位置，價值在200000-300000之間，需要四個房間同時土地面積不得小於300平方公尺。如果資料設定包含屬性“價格”、“房間數”、“土地面積”與位置，則期望前提下的地圖，可以容易的透過多元交集被詳載的屬性而獲得最吻合的土地：

如果價格200,000以上 **AND** 價格300000以下 **AND** 房間數為四 **AND** 土地面積大於300平方公尺 則ITEM=1 否則ITEM=0

被選擇的實體如果符合屬性清單，其布林值為1；反之為0。結果的展示透過著上新的顏色於ITEM=1的實體。

新考量一個疑問於土地適宜性分類。在資料庫的土壤圖單元中，每一個單元也許有屬性來描述它的質地與表土的酸鹼值。如果設定地圖單元為A，叫做奧勒岡壤土（名目資料類型）；設定B為表土酸鹼值大於7的製圖單元（數值資料類型），則資料更正敘述如下：

$X=A \text{ AND } B$  可以找出所有奧勒岡壤土且酸鹼值大於7的區域。

$X=A \text{ OR } B$  可以找出所有奧勒岡壤土與表土酸鹼值大於7的區域。

$X=A \text{ XOR } B$  可以找出所有奧勒岡壤土與表土酸鹼值大於7的區域，但是不包括同時具有奧勒岡壤土與表圖酸鹼值大於七的區域。

$X=A \text{ NOT } B$  可以找出所有奧勒岡壤土區域，但是不包含表土酸鹼值大於7的區域。

被選擇的實體也可以重新命名並（或）給定新的展示符號（如圖7.3a）。這是一個明顯的邏輯敘述例證“如果條件為（C），則實現詳述的工作”。

值得注意的是，布林運算不同於算數運算的是，條件順序是不可以互換的。 $A \text{ AND } B \text{ OR } C$  的結果取決於順序上先運算AND其次運算OR。附帶說明的，通常考量超過兩個實體時，同常會先說明其運算順序。依上例，設定C為貧乏缺水土壤，則 $X=(A \text{ AND } B) \text{ OR } C$  為（a）奧勒岡壤土且酸鹼值大於7；聯集（OR）；（b）土壤貧乏缺水的地圖單元。 $X=A \text{ OR } (B \text{ AND } C)$  為（a）所有奧勒岡沃土地圖單元；交集（AND）；（b）包含所有表土酸鹼值大於7與土壤貧乏缺水區域。

此外也值得注意的是，布林運算也許需要一個符合於屬性的性質於資料中，同時它並未考量了誤差與不確定性，除非這些是特別被歸併到串組（sets）的定義中。

如果“海拔高”這個屬性質值被設定高於2000公尺者為山脈（mountain），則海拔1999.99999999.....的山丘（hills）被否定了。這不是順序或名目資料類型的問題，但是當評估有關會表現出“量”的尺度資料時便會產生問題，諸如海拔高、pH值、泥土含量、土壤深度、大氣壓力、鹽分、人口等問題。而這個問題在於各式各樣的測量誤差與不確定性。

如果誤差結合於資料中而擴及了設定的邊界值，則布林原則的精確應用也許產生違反直覺的結果（見第11章）。

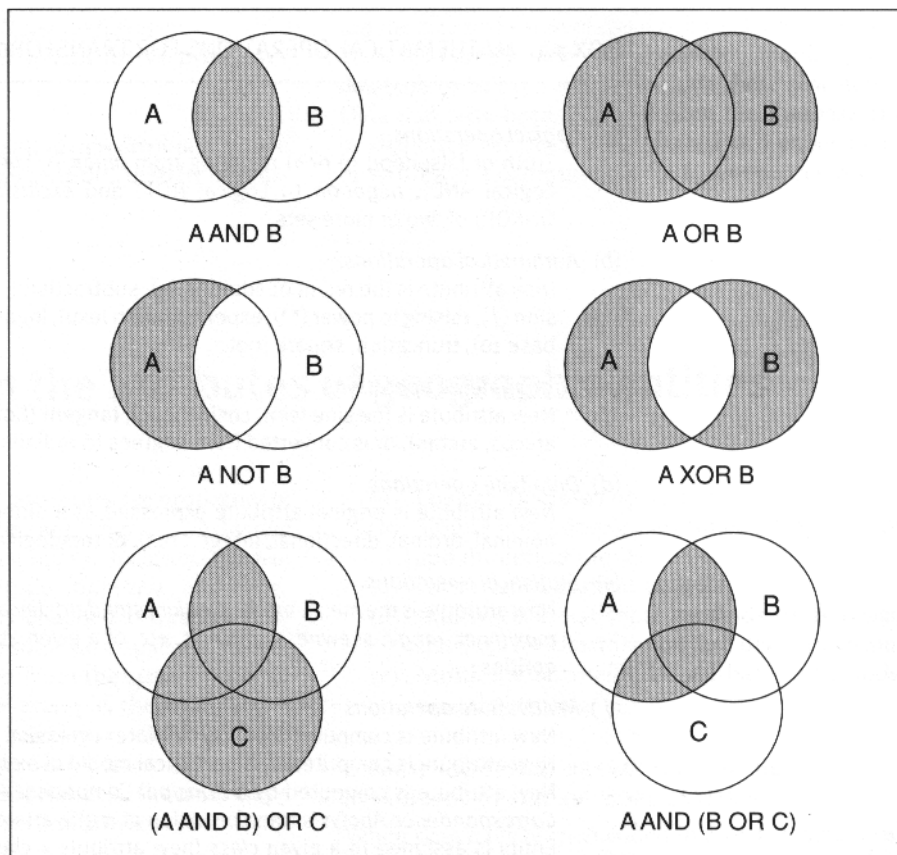


圖7.2：使用布林運算結果，其中陰影部分為“真實的 (true)”（原文p.166）

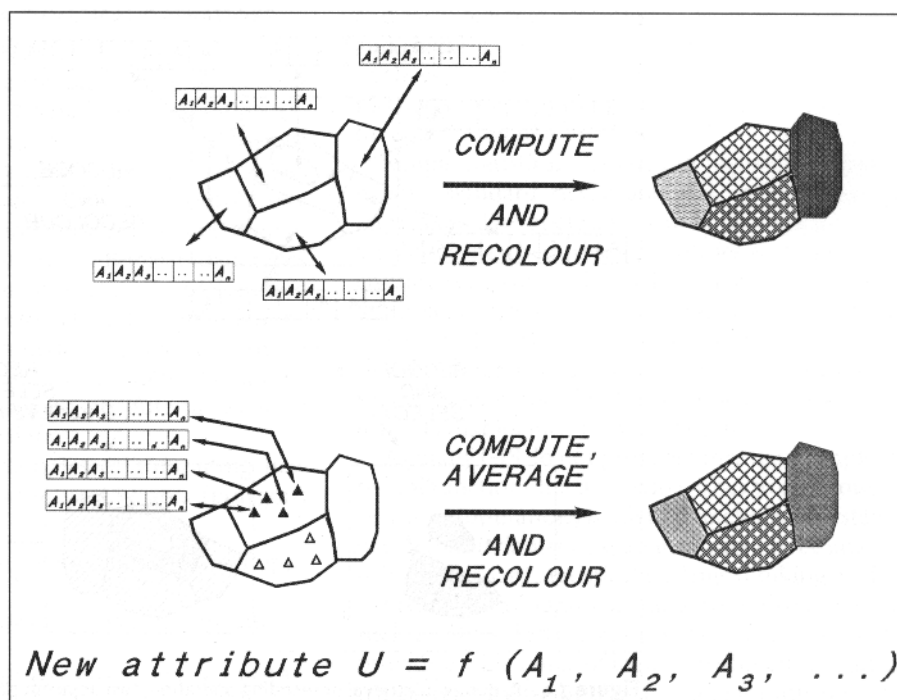


圖7.3：在屬性上單獨擷取實體時或從原有實體運算新屬性，地圖中的空間單元並未改變形狀，只有顏色與陰影部分改變（上圖）。同樣的，當透過多邊形中的點搜尋包圍點的物件時，它們是運算平均面積（下圖）。（原文p.167）

### 單一實體多屬性布林運算的空間觀點

除了符號與邊界的移動，實施邏輯擷取與重新分類於空間中實體的非空間屬性會對地圖的影像產生影響。估算一個新的屬性或是條件需要安排圖例與被選取實體的顏色或陰影（見圖7.3a與插圖2.1）。當選擇使得鄰近的多邊形收到一樣的符號，這可能使它們之間原本明顯的邊界消失，因而完成一種形式的綜合圖（圖7.4）。圖7.5說明了選擇這樣的方式來單純化複雜的土壤圖。

布林運算不僅適合於地理元素的非空間屬性，因為它還可以應用於地理位置與源於空間的屬性或地理實體的地形特性。例如，某人希望找出面積大於五公頃且具有酸鹼值大於7.0的黏土區域。更複雜的搜尋也許包含了面積的形狀，區域的邊界特性或是鄰近區域的特徵，像是森林地區與都市地區的疆界界定。在這些例子中，搜尋的結果將會影響空間的紋理。

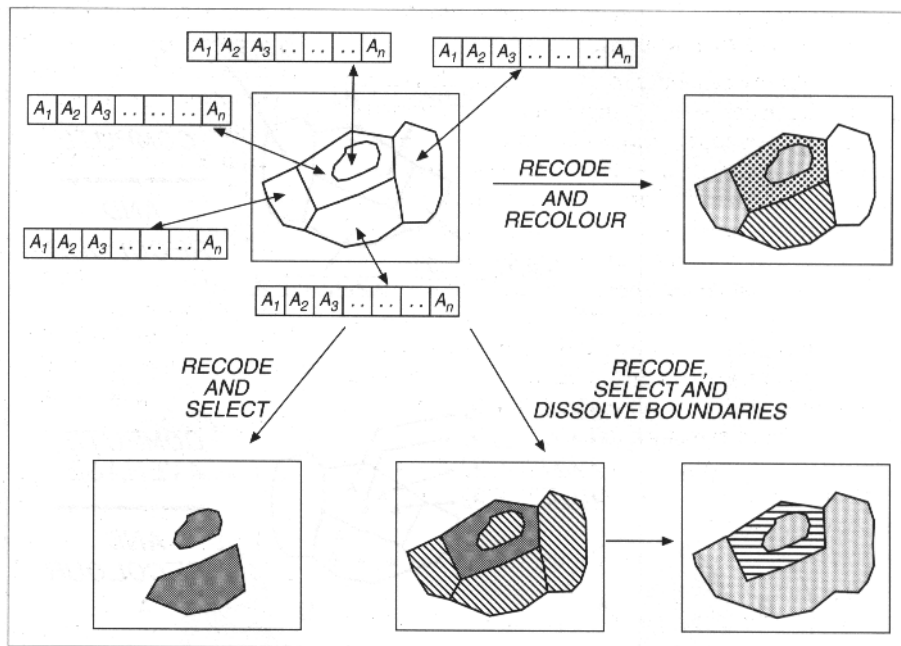


圖7.4：擷取與重新編碼中，兩個鄰近的多邊形接受了同樣新的代碼（code），它們之間的邊界因而消失歸併成為一個綜合的圖（原文p.168）

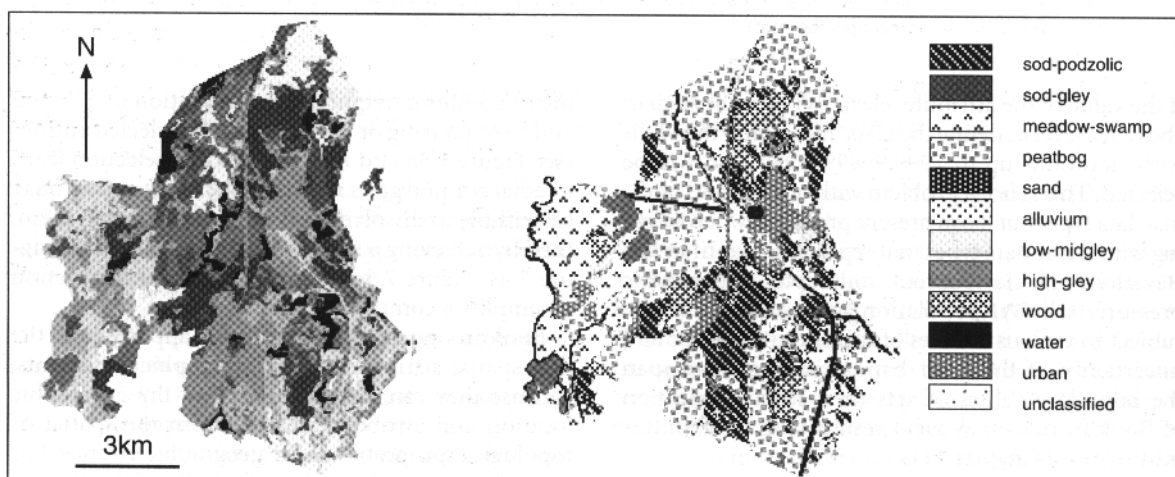


圖7.5：使用重新分類使地圖歸併（綜合）。左圖：原本的土壤圖包含95個不同的土壤單元。右圖：將土壤圖重新分類成12個單元。Note：重新分類後仍保有原本的幾何特徵，同時圖例與陰影符號只應用到單純化後的圖（原文p.168）

### 簡單與複雜的算數運算於單一實體的屬性

新的屬性可以被所有普通的算數方式估算（加、減、對數、三角函數、指數與所有組合包括複雜的數學模式）。算數的與三角的運算方式可以明確的只被使用於數值資料與可靠種類的順序資料。算數運算在布林資料類型與順序資料類型是無意義的所以不被允許（算式如  $X = \text{sqrt}(\text{London})$  是無意義的）。算數的運算可以非常簡單或非常複雜，但是在所有例子中運算是一樣的——一個新的屬性（結果）是從現存的資料被運算的。

運算新的屬性於給定的管理區域（多邊形）或點的位置假設例子如下：

- 人口數增加 = 人口數1990 - 人口數1980。
- 總消費能力 = 平均收益 × 人數
- 每個農場平均小麥產量 = 所有產量 / 農場數
- 預測小麥產量 =  $f(\text{收成})$   $f(\text{收成})$  是一個複雜的運算小麥產量之數學模式，如同土壤、濕度、土地肥份特性（點的實體）。
- 等級分配 = 結果（多變量分類） 多變量分類也許是任何實體中的數值屬性的統計或多準則分析（multicriteria analysis）。

對一條河川集水區而言，經由排水口排出雨水的比例，可藉由從排水口累積的年河流排水量區分每一集水區年雨量來估算之。

算數運算與邏輯運算可以被容易的被結合：

IF (A+B) / (C) ≥ TEST VALUE THEN CLASS = GOOD

## 屬性的統計分析

本書中沒有足夠的空間來解釋單一變量與多變量的統計分析，如果讀者需要這些理論的細節應該查閱標準的教科書。例如 Davis (1986) for geology and earth, Jongman *et al.* (1995) for ecology, Haining (1990) for the environmental and social science。在本章節中，我們認為統計模式只是一組估算新屬性的步驟。

簡單的統計分析可以被用來估算中位數、標準差、相關性與回歸。這些估算方式可以應用於一組附加於單一實體的屬性或任何透過邏輯搜尋（logical search）被更正的實體。例如，估算所有土地的地表水硝酸鹽等級的中位數與變異數或估算所有郵遞區域包括火車站的快餐銷路平均數。

大部分的地理資訊系統並不提供簡單統計以外的統計功能，使用者欲進行統計運算前，需要選擇一個可以使用統計套裝軟體來分析空間中實體屬性的系統。連結地理資訊系統至試算表程式或是標準統計套裝軟體如 SPSS、Statistica 或 S-Plus 以提供寬廣幅的邏輯、算數與統計運算來轉換實體的屬性資料於地理資訊系統中。所有的試算表與統計套裝軟體均包含有用的圖解程序於平面圖表、直方圖與其它類型的統計圖表。使用超連結的功能來連結資料庫、圖表與地圖展示，使用者能透過功能強大的資料分析工具來檢視其可能的分佈格局與屬性資料的相似性和實體在空間分佈上的關連性（Gunnink and Burroughs 1997, Haslett *et al.* 1990）。

大部分的統計套裝軟體提供下列功能於統計資料分析：

- 基本的中數統計（statistics — means），標準差（standard deviations），變異數（variance），歪值（skewness），峰度（kurtosis），極大與極小值等。
- 無參數的（non-parametric），統計中位數—眾數，upper 與四分位。

- 直方圖，2D與3D的分佈圖，Box and whisker plots, 樹狀圖。
- 變異數的單變量與多變量分析。
- 線性迴歸與關連性分析。
- 主成分原理與因素分析。
- 群團分析。
- 典型分析（canonical analysis）。
- 判別分析。

## 數值模式（numerical models）

算數運算應用於數值屬性是沒有限制的，它們常常用來運算屬性的值。而這些屬性的量測是非常困難的或是可以輕易的從有用的原始資料中導出，例如從統計量或是自然資源調查中取得。數學運算的標準已經被迴歸模型導出，在專業領域中有時稱之為轉換函數（transfer functions）（Bouna and Bregt 1989）像土壤科學與土地評價。更複雜的數學函數式子被用來代表實質作用如作物生長、空氣品質、地表逕流、殺蟲劑的溶解、傳染病源、與日具增的污染壓力等通常求助於數學模式。大部分的地理資訊系統中並不提供複雜的函數模式之程式，反而它們常被組合成資料形式在網路上其它電腦中輸出成模式。經過運算成為模式的成果再被傳回地理資訊系統中成為新的屬性用於展示或評價。然而新建立的程式語言與特殊工具的介面，提供了地理資訊系統空間模式的製作。（e.g. ARC/INFO—Batty and Xil 1994a,1994b; PCRaster—Wesseling et al. 1996 or GRASS—Mitasova et al. 1996）。

## 神經網絡、多準則評估與模糊邏輯(Neural networks, multicriteria evaluation, and fuzzy logic)

到目前為止所有導出新屬性的理論都是參數的（parametric），用來說明一個新屬性的定義可以透過邏輯的或是數值的方程式來表達，其中權重與變數是可以被客觀分配的。迴歸分析與the calibration of numerical models兩個例子其“最佳”變數值被選擇用於分類與預測。附加說明的是根據許多案例顯示數值模式多是線性的。例如此處是針對變數值與其影響的結果之間的關係。

這些假定源自於典型的機械科學。參數理論是難以使用的，然而在複雜非線性的統計中屬性值是非常態分佈的，而且分佈原因甚至在統計上的關連是輕微的。這些困境圍困著空間資料的相互關係也許違反了變數理論的精神，而且其問題與如何對複雜的空間主題分類一樣棘手。這些問題在用於土地覆蓋分級的遙測資料分類中已經被特別注意（Lees 1996a）。

神經網絡提供了新的地理實體（實體或像元）分類方法於明顯的群體中（Fitzgerald and Lees 1996, Lees 1996b）。在許多例子中，分析的並非是真實空間的訊息，但是需要選定一個非統計理論運算的實體作為分類的基準。神經網絡是一個程序的手段，提供像是規則系統（algorithm）或是電腦程式來模仿人類的智能。神經網絡是由許多簡單的程序元素（processing element）所組成，這些程序元素透過單向的通訊管道之連結來傳遞數字資料。這些元素的運算僅止於局部的資料，但是整體來看，神經網絡依據本身對格局（pattern）的理解，對輸入的資料加以組織。這些格局可以透過“自行學習（self-learning）”來建立—神經網絡系統從資料中決定最佳分類方式或提供一個分類需求的模板（template）來“監督分類（supervised classification）”。

應用神經網絡於地理資訊系統仍然處於研究階段 (Lees 1996b)，但是暗示出這些工具將提供複雜資料關連之有用的與簡單的方法，此外也許非常需要建立複雜的模式。

神經網絡不是連結複雜與非線性資料的唯一方法。多標準評估理論已被建立來提供使用者決定新的屬性，但這反應出涉及複合的與相互矛盾標準的問題。在11章我們探索在地理資訊系統中特殊的模糊邏輯與連續分類於空間資料分析。

## 空間實體性質的屬性資料導出例子 (Examples of deriving new attributes for spatial entities)

新運算的屬性與重新分類的結果通常透過重上陰影與著色於實體上的方式來呈現 (圖 7.2)。如同布林選擇，除了鄰近的實體因相似而被分成同一類之外，空間實體的特性 (位置、性狀、類型、地形) 並未改變。值得注意的是，除非資料結構是近似於網格資料符號的基本單元圖，若資料是網格 (raster) 的形式，則可透過單一像元 (pixel) 個別運算之 (見第三章)。接下來的文章提供一個針對新屬性產生的基本空間分析的例子。

### 使用實驗 (empirical) 迴歸模式

在此利用一個評估瑞士高山的溫度函數之線性迴歸作為統計模式的例子。瑞士高山的海拔高與溫度之線性關係如下：

$$T = 5.697 - 0.00443 \times E \quad 7.2$$

式中T是攝氏溫度，E是以公尺為單位的海拔高。在平面的基質中，每一個網格單元被視為個別的實體，因此應用方程式 (7.2) 於所有單元，從數值高程模型 (DEM) 中求出網格溫度圖。類似的例子在第五章與第十章已探討過有關重金屬污染與獨立屬性的關係，像是距河流的距離與澇原評估。

### 使用多變量簇群 (Using multivariate clustering)

地理人口統計學的分割 (geodemographic segmentation) (Webber, 1997) 是被跨國際市場的公司，利用住在當地的消費者之統計資訊，透過區別鄰近的基本類型來對西部鄉村住宅區分類。這些空間的實體透過行政區域人口普查、郵政編號區域或郵件訂單位址連結至全球的跨國際房屋地址所提供。這提供了一個優秀的空間的答案，即每個空間單元內大約有十七個住址。每一個空間單元由四個標準來表現其特徵：年齡 (年輕的、中年的與老的)；收入 (高收入、中等收入、低收入)；都市化程度 (首都、都市、都市外、農村) 與家庭型態 (有小孩的已婚夫婦、獨居的、無子女的夫婦、領撫卹津貼者)。這些屬性產生108種可能的分類組合，經過重新編碼後成為10個核心的分類層級用來識別社會經濟 (socio-economic) 類型的特徵。層級分配的多變量理論是被用來確認基本空間單元在10種分類中的類型，同時也與經驗尺度資料與透過問卷調查所獲得的消費者偏好屬性作連結。層級類別內的空間單元之邏輯更正與屬性提供了地圖於區域的或全國性層級，揭示出了市場機會與品牌偏好的空間分佈。

### 使用簡單布林邏輯

在世界許多地方，使用不足的資料如濕度、能量與雨量等所建立的數值模式來估算作物



產量，所以用簡單標準為基礎的定性預報也許是唯一的評價土地農業發展的方法。這是一種哲理支持著現在著名的聯合國糧食與農業組織（FAO, Food and Agriculture Organization）的土地評估程序（FAO 1976, Beek 1978）。慣例的（prescriptive）土地評估（Rossiter 1996, FAO 1976）是一個基本的簡單概念，即地景可以被明顯的邊界分割成“製圖單元（mapping units）”。土壤調查通常是地景分割類型的基礎，但是除此之外的理論有使用地貌、生態的分帶或植物群落。就典型的屬性值而言，昔日的基本空間單元已經被繪製成圖與定義之，它們對給定的用途之適宜性也許被重新分類或基於目前資訊運算新屬性所限制。一般的程序被稱為“由上而下（top-down）”邏輯，因為它起使於預料全球性的標準或存在於轉化主要的特定用途單元資訊於單元中的物理模式。不論是資料被儲存成向量式多邊形或網格單元，這個程序剛好相同；唯一不同的是空間的描述於概念上的實體所運送的資訊種類。

下面是一個基本原則（rule-based）理論的例子。為了使作物生長良好，需要潮濕的、排水良好的與肥沃的土壤。種植單一作物使土壤被裸露或是一年之中部分時間接近裸露，此時土壤應該可以抵抗雨水沖蝕的影響。這四個屬性，可利用水分、可利用的氧氣、養份與沖蝕敏感度（erosion susceptibility）是大家已知的“土地品質（land qualities, LQi）”，這些資料可以透過農學家與土壤專家，使用簡單邏輯轉換功能從主要的土壤與土地資料中取得。整體而言，適合的地點被土地特徵界限所限制—（FAO 1976, McRae and Burnham 1981）。

圖7.6說明了完整的程序。使用從肯亞共和國的Kisii部分地區的土地系圖、描述性的與數值高程模型所獲取的資料（Wielemarker and Boxem 1981）。研究區域面積1,046（3,750×3,750）公頃，使用的地圖為Marongo地區的1：12,500詳細土壤調查圖（Boerma *et al.* 1974）。選擇因素憑藉parent material（七種重要地質類型）、地勢起伏（海拔幅度從4,700-5,300英尺，約1,420-1,600公尺）與土壤（12種地圖單元）。詳細的土壤調查資料被描述成parent material、土系、風化岩床的土壤厚度、卵石性質（stoniness）、岩石性質（rockiness）；同時表格資訊顯示土系與土地品質的關係。每一個屬性被數化成為分離的多邊形疊圖，並改變成為每方格62.5公尺的60×60方格陣列。數值高程模型的獲得透過從數化的輪廓線、標高與改變的區域地勢起伏（最低40公尺、最高560公尺）等資料內插推估而來。關於氣候、土壤化性、土地使用與文化常規的資料也是有用的（Wielemaker and Boxem 1982）。斜坡距離（需要被估計耗損）可透過航空立體照片與個別的疊圖來說明。

在這個例子中，我們藉由土地品質—養分供給、氧氣供給、水分供給與沖蝕敏感度來考量種植玉米的適宜區域。這些土地品質可以被個別的列成指定的值1、2、3於下列的分類：

無限制	指定為1
中度限制	指定為2
嚴重限制	指定為3

導出土地品質的原則是：水分的可利用性是土壤深度與土系的布林交集（AND）；氧氣與養分可利用性可以從土系資訊中透過查表獲得；沖蝕敏感性或風險可以藉由坡度等級與土系之布林交集（AND）導出。原則性的例子如下：

如果土系是 $S_1$ 則透過查表 $L_w$  指定養分品質為 $W_3$

如果土系是 $S_2$ 同時坡度等級是平緩的，指定沖蝕敏感度為1

如果土系是 $S_2$ 同時坡度等級是陡峭的，指定沖蝕風險為3

過去個別的土地品質已經被指定，所有多邊形或是像素的適宜性之限定值已經被確立了：

適宜性 = 最大值 (LQ<sub>water</sub>、LQ<sub>oxygen</sub>、LQ<sub>nutrients</sub>、LQ<sub>erosion</sub>)

這些提供了下列三個等級的適宜性：貧乏的—最低等級嚴重限制的；不嚴重，中度限制；無限制。

這是一個簡單的方法，使用其它的因素值重複分析相同區域土壤與土地品質關係。以便瞭解如何透過灌溉、覆蓋掩蔽物與改變成梯田來增加土地適宜地區的面積。這個情境模擬可以透過取代真實土地特徵資料來實現，表現出一個更低等級或是改善的情境。在這個方法中，可以使用布林模式依不同的因素而定來探索地區的適宜性。然而，有一點是不可以忽視的是結果和資料與洞察力一樣在土地評估中提供。

### 空間實體屬性的(Temporal modeling)

一個根本的問題是簡單的基本原則模式並未能描述出土壤與地景間資源的改變的時間歷程。我們也許預料到栽植玉米的田地會使得土壤養分耗損與降級，即使當初是適宜的土地，也許在重新分類後會成為低度適宜等級。所以，應該要可以有用的估算研究區域玉米田今後40年的土地適宜性。過份的農人已經種植玉米超越了整個區域來供養牲畜，根據1985年調查，玉米田呈現每年4%成長。根據研究，農夫可以經由簡單的保育理論自行實施簡單措施來保護土壤，在部分區域這些適當的保育理論需要一些附加的準備使其更成功。

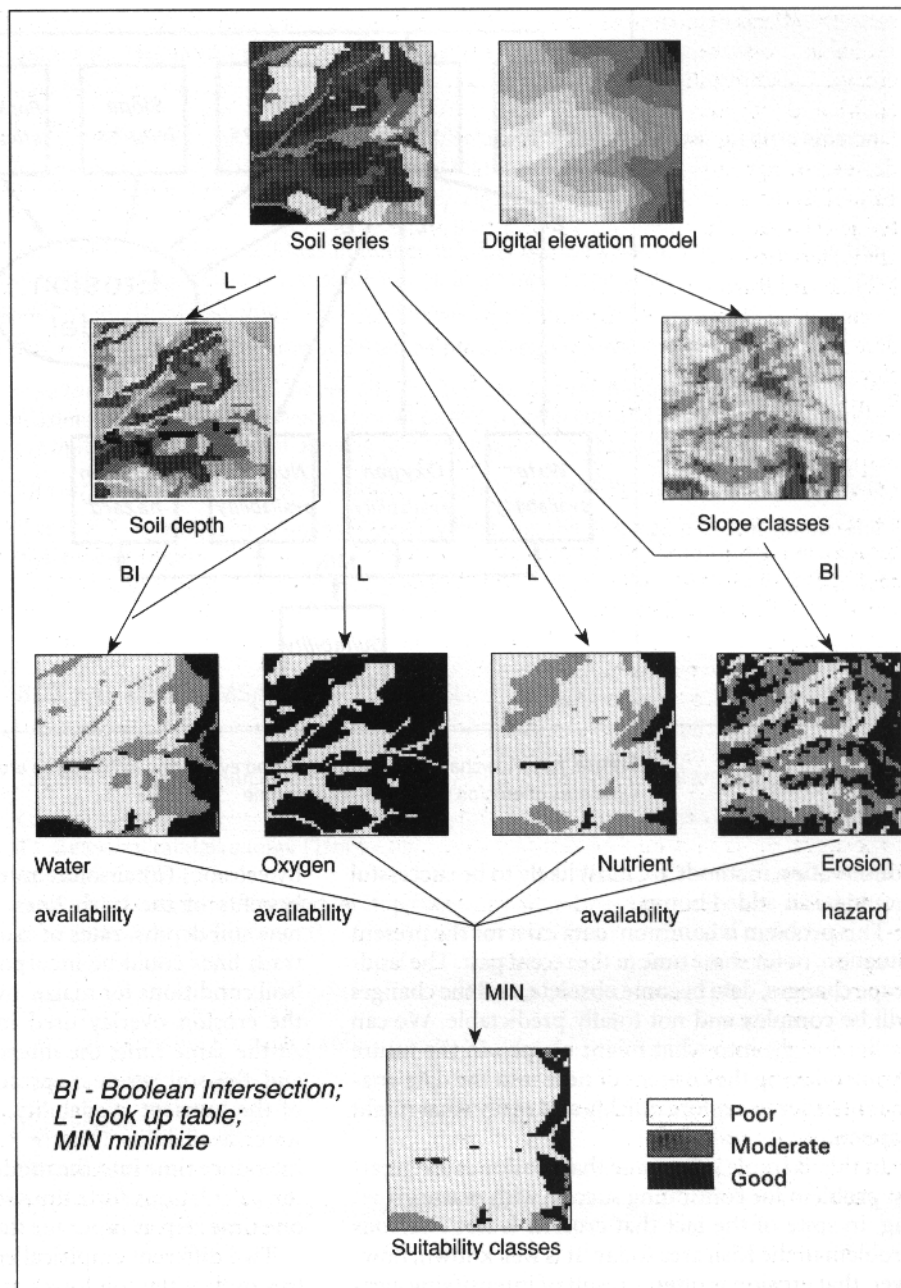


圖7.6：運算“top down”土地評估於界定玉米栽植土地適宜性的流程圖（原文p.173）

這個問題是共同的：資料呈現的是現在的情況或是過去的。土地變遷，則資料變成無用的，而且改變是複雜與無法預測的。為了更容易與迅速的探索未來可能發生的事情，我們可以納入時間因素於資料分析中。

在這個例子中，我們假定土壤沖蝕是影響玉米田連續栽植最大的問題。事實上，今日在Kisii地區地區土壤沖蝕並非是一個嚴重問題。然而，沖蝕通常是農業增強的結果，同時Kisii地區農夫已經開始被鼓勵使用玉米植物體的莖、葉，沿著等高線構築成“資源堤（trash lines）”以降低土壤流失（Wielemaker and Boxem 1982）。所以，主要的問題是如何估算土壤流失。這個問題可以透過包含“實驗沖蝕模式（empirical erosion model）”於土地評估程序中，用來估算（a）土壤深度如何因沖蝕而減少。（b）純粹的沖蝕速率與（c）trash lines的利益。有關於新土壤深度、沖蝕速率與資源堤的利益等資訊可以被合併於新的、未來的“玉米土壤狀態”疊圖，用來取代本例中原本的沖蝕疊圖。同時，新深度與土系圖交互關連

可以提供新的估算於養分適宜性、氧氣適宜性與水分適宜性（圖7.7）。使用多組時間速率結果，重複運算時間速率是最簡單的方法來將時間因素加入模式程序中。

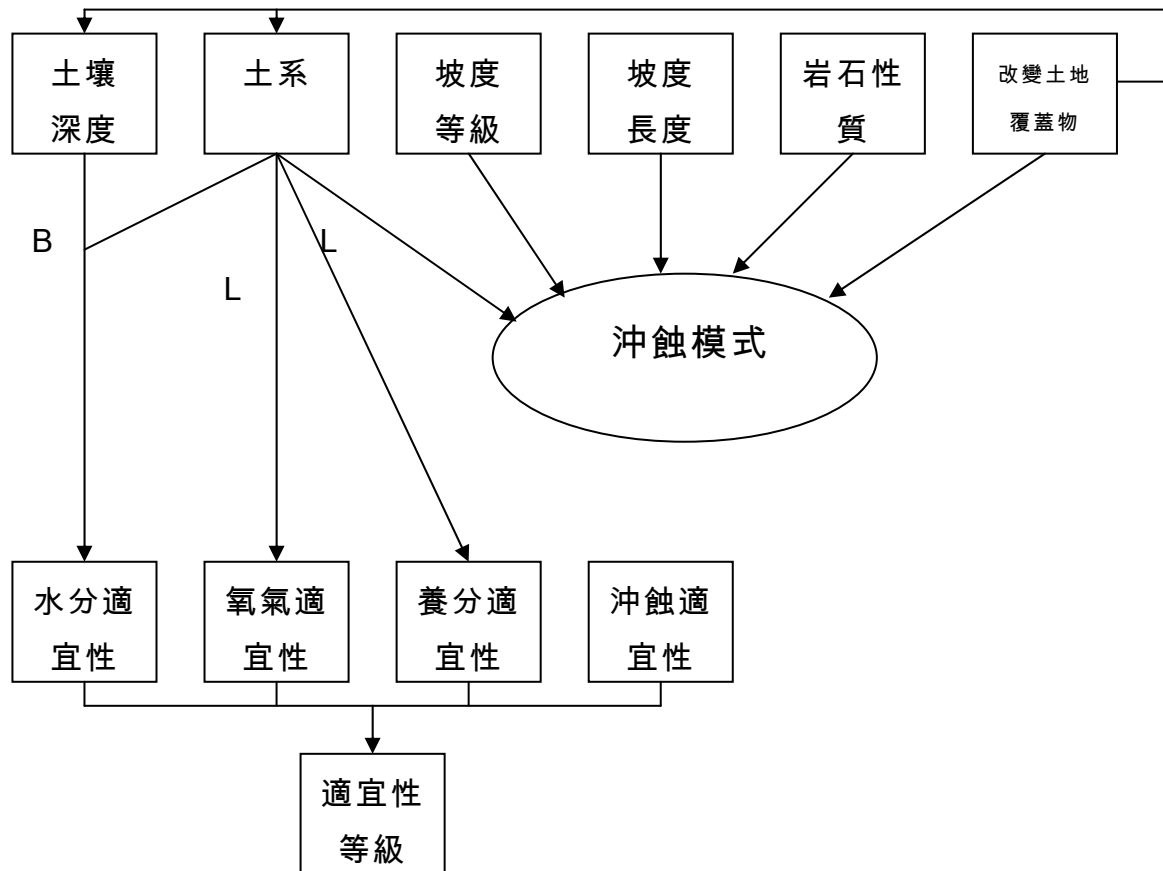


圖7.7：包含跨時間沖蝕適宜性運算的沖蝕模式於土地評估估算之流程圖（原文p.174）

兩個不同的實驗沖蝕模式被用來模擬土壤流失，換句話說，通用土壤流失方程式（Universal Soil Loss Equation, USLE）被Wischmeier與Smith所建立（1978）。隨後Stocking（1981）建立了南非土壤流失評價模式（Soil Loss Estimation Model for Southern Africa, SLEMSA）。這些模式之所以被使用是因為它們容易計算，而所需的資料也已經被建立於資料庫中（參閱Box 7.2與Box 7.3）。不利於實驗模式的是它們有關於真實沖蝕問題的運算過於草率，但是專家模式（specialist models）中，通常區位特性與所需資料無法在一般目的的資料庫中尋獲。

沖蝕模式是點的運算；它們運算每個實體的土壤流失（本例是使用網格形式），而未考慮到土壤的流失與增加超越了鄰近的網格。資源堤對沖蝕的影響透過重新分類斜坡長度疊圖與重複模擬來運算。這個模擬重複了新土壤深度、岩石性質、所有沖蝕速度模式與不包含資源堤影響等總共12次的疊圖。在給定的模式中，疊圖運算去掉了沖蝕速率與資源堤的疊圖，使用沒有梯田利益的無資源堤的圖（圖7.8）。兩個模式同樣考量去除沖蝕速度疊圖的運算，產生不同的疊圖於兩個模式運算條件下。

這個跨時間模擬的結果顯示，地景將呈現緩慢的退化—40年間最佳區域從14.1%下降至11.6%；最低區域從51.3%下降至50.6%。這個模擬建議了沖蝕將集中在幾個區域，同時在一些沖蝕層級中，資源堤會對其產生少許改變。

乍看之下，這個結果似乎是有說服力與逼真的，但是我們沒有辦法知道如果它們什麼地

方是接近正確的。可以明瞭的是實驗方程式難以轉換於不同的地區，儘管許多努力是在於找尋普遍性的模式。更深一層的問題是在疊圖運算中，已經被假定的資料是完全的與不變的。這明顯的不是事實，即使重新更正Kisii的詳細土壤調查。特別是許多地理資訊系統的使用者，在地理模型模式中使用地誌主題圖時，誤差程度與誤差擴及的問題在地理資訊程序中很少被討論到。很明顯的，進行地理分析時，誤差問題具有決定性的影響。全部的問題將在第9、10章完整的被討論。

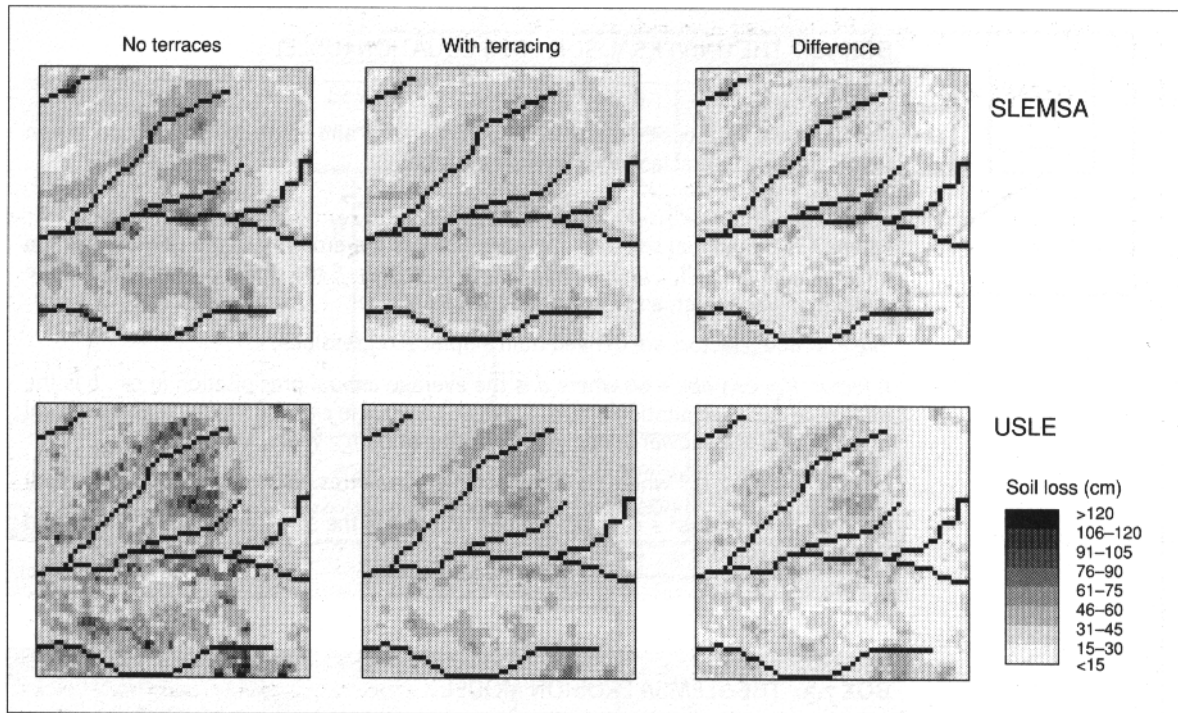


圖7.8：為圖7.7流程圖評估結果，其中使用兩種不同的模式操作與使用不同的保護措施（原文p.176）

Box 7.2：通用土壤流失方程式（Universal Soil Loss Equation，USLE）（原文p.175）

通用土壤流失方程式（Wischmeier and Smith 1978）透過實驗關連預報農業的土壤流失

$$A=R \times K \times L \times S \times C \times P$$

式中A是一年土壤流失量（公噸）；R降雨的沖蝕；K是土壤的沖蝕性質；L是斜坡長度（公尺）；S坡度（%）；C是耕種指數；P是防護指數

因素R、L與S是從實驗迴歸中導出

R因素， $R = 0.11abc + 66$  式中a是年平均雨量（cm）；b兩年間最大降雨日之降雨量（cm）；c兩年間所有陣雨降雨量（cm）。

L因素， $L = (I/22.1)^{1/2}$ 。式中I是斜坡長度（Wischmeier與Smith，1978）

S因素， $S = 0.00655s^2 + 0.0454s + 0.065$ 。式中s是坡度（%）

Box 7.3：SLEMSA沖蝕模式（原文p.175）

南非土壤流失評價模式（Soil Loss Estimation Model for Southern Africa, SLEMSA）

控制變項

E 季節降雨量（ $J/m^2$ ）

F 土壤沖蝕性指數

i 被作物阻斷雨量（%）

S 坡度險峻 (%)

L 斜坡長度 (公尺)

次模式 (submodels)

土壤裸露情況  $K = \exp [ (0.4681 + 0.7663F) \cdot \ln E \div 2.884 - 8.1209F ]$

作物覆蓋  $C = \exp [ -0.06i ]$

地形  $X = L^{0.5} (0.76 + 0.53S + 0.076S^2) / 25.65$

輸出

預測每年土壤流失 (公噸/公頃)  $Z = KCX$

## 多項空間屬性資料在空間疊合時之操作(Operations on attributes of multiple entities that overlap in space)

這裡我們將屬性操作的討論延伸到包含2個或多個完整或部分覆蓋在同一空間的實體屬性，換句話說我們考慮下面的問題：

A包含B或

A包含於B

重疊與交叉的問題

A相交B

A與B重疊

式中A與B是兩個不同的空間實體

### 包含

第一個例子A包含B，且A包含於B，從實體屬性的延伸布林運算規則得到解答來測量實體如何佔用空間。哪家餐館位於蘇活區？哪條地下水監測井鑽孔在X區？一但這些實體被選定加上，就能運用屬性分析的程序，不論是一個或全部實體。舉例來說，一個地下水井的最低最高水位能在幾年內被抽出，或是全部地下水井平均水位也能被估算出來。這些計算的結果能用其他顏色、陰影或是標籤來標出一個封閉的多邊型。

其他可運用此分析的例子是：考古學方面，「決定A區後鐵器時期埋葬地點數目」或檢索葡萄通道 (passage graves) 與決定土壤種類與造園發生的位置；或是土壤學方面，找出SI單位土壤的所有的剖面，與計算表土裡黏土層的中數與標準差。

### 重疊與交叉實體

在剛才的例子A包含B且A包含於B，及A交叉B且A重疊B，其中的A與B是線或是不同型式的多邊型，邏輯檢索的第一步驟是定義出新的區域或是線段，多邊形時的程序如同重疊或是交叉多邊形，而導出新的空間實體。圖7.9依照是否為空間性的布林結合 (圖7.9a)、覆蓋 (圖7.9b)、交叉 (圖7.9c)，顯示出的3個不同的結果。多邊形疊合用來解釋「找出Westminster城市被公園覆蓋的區域」這樣的問題，第一個疊合顯示城市行政區域的邊界，第二個疊合顯示不同的土地覆蓋區域 (圖7.9c)。實際的運用上，假如地圖A是一張土壤圖，地圖B是流域圖，結果就是流域圖的土壤圖。

有一些情形是多邊形疊合導出所謂的假多邊型（圖7.10）。這是因為一些邊界的數位化過程被認為是在相同的位置。這樣的錯誤來自放錯數位資料，也可能是不同原始文件所造成，或是測量上的誤差。令人詫異的，試圖將線段用更多的點做更精確數位化，並不能解決這樣的問題反而使問題更糟糕，這有一些解決方式，第一，指定一張主要層的邊界作為主邊界而其它的必須順著這個邊界。第二，檢查假多邊形並且消除所有小於這區的關鍵值；必須做一個決定就是被假多邊形覆蓋的較大多邊形必須加回去。第三，略過假多邊形的所有調節沿著結合邊界區域的平滑視窗，計算一個新的平均邊界。這常過度定義（over-defined）而且使用Douglas-Peucker演算法或其他中數時比較單純，而且使計算與顯示新多邊形實體邊界時較平順（圖7.10）。

### 一或多個直接點對點連結的操作（物件導向）

利用空間資料庫大量的計算代價是，在線與多邊形上導致實體產生裂縫或從資料庫移除的邏輯操作，因為可能依靠這個操作實體的數量被執行完成。在實際的情況下，如果兩個簡單的多邊形相交創造第三個多邊形，第三個多邊形與所繼承的屬性都會被加入資料庫中，如果重新分類兩個毗連的多邊形導致不必要的邊界，那麼兩個多邊形會被移除並加入一個新的到資料庫中。在實行上，多邊形加入或移除的數量可能很大或是很混雜，所以很難說這個數量有多巨大。儲存修改原始資料庫，這改變常只是計算原始資料的次要部分，且結果儲存在分開的檔案或資料夾。

在hybrid-relational GIS中，加入與移除多邊形表示同時更改空間資料與屬性資料，修改的空間資料比加入或刪除一個表格的進入點還多，因為所有的拓樸連接都需要被計算，大多數商用GIS有一個妥協的解決方式，用他們自己的軟體與商用資料庫相關程式來處理修改的空間的資料來儲存相關的屬性。

關連網絡的混合資料庫的優點是原則上分析的數量與種類沒有限制，物件導向的GIS試圖用合併大數量的資訊建構資料來解決這些計算上的問題，這樣的方法是以不大量改變資料卷數來完成。這種最通用的資料取得與分析法選擇的方法需要事先思考，此乃為何物件導向資料庫的建構會花很多時間。

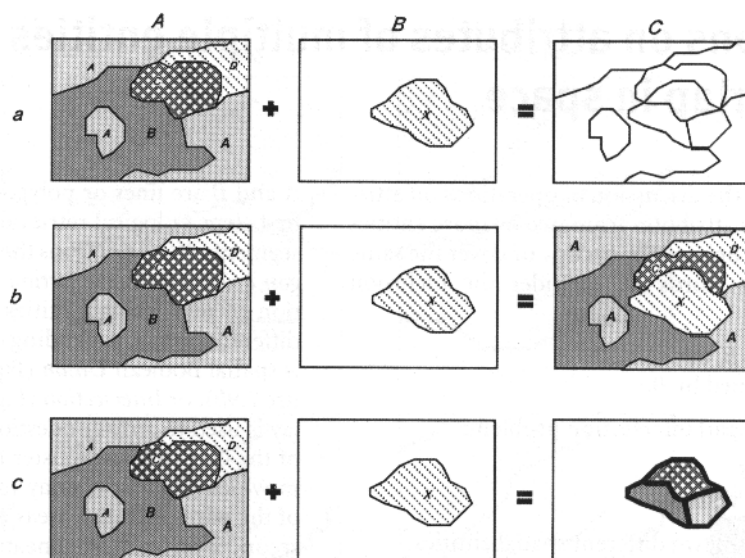


圖7.9：多邊形疊圖使得資料庫內實體的數量遞增(a)簡單疊圖—所有邊界保留(b)第二張圖覆蓋了第一張圖，使的地圖發生小部分改變(c)從第一張圖“切出（cut out）”一小部分（原文

p.178)

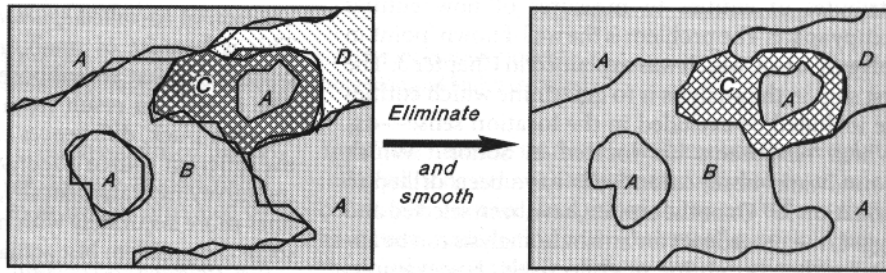


圖7.10：進行多邊形疊圖時會產生大量無意義的小多邊形，必須除去它（原文p.178）

### 由A、B間之距離進行緩衝區作業(Operations that depend on a simple distance between A and B (buffering))

操作類型“A到B的距離低於/超過距離D”，其中D是透過緩衝指令協助的烏鴉飛行距離。這是畫出一個區域包圍原本的實體。該分割區或緩衝帶的邊界是所有的D從原本實體的座標構成。如果那是一個點的實體，則分割區是一個圓圈；如果是一個筆直的線，則分割區呈完整邊界的矩形；如果是一個不規則的線或多邊形，則以同樣的方式向外擴張（見圖7.11）。緩衝帶實際上是一個新的多邊形，被用來臨時的協助解決空間問題或自行附加於資料庫中。決定一個實體是否在內部、於外部或重疊於緩衝區域只是使用描述的操作方式（包括多邊形疊圖的問題—見圖7.7），同時在這些實體中，邏輯的與數值的操作進行與前面的一樣。

接下來是使用分割/緩衝指令的典型例子於其它分析的操作：

- 決定距離五公里內同時為白色房子的快餐店之數量。
- 調查自然河道潛在性水污染於鄰近的廢土場中。
- 運算房子橫置200公尺內所有總值來建議新路之路徑。
- 估算全世界所有居住距離海洋100公里以內所有人數。
- 估算水源第五公里內的牧草數量。
- 定出Neolithic village一小時步行可及距離內潛在耕地面積總量。

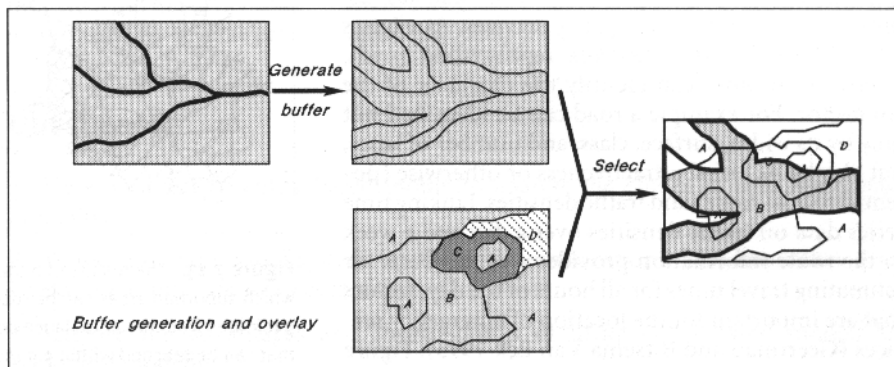


圖7.11：產生緩衝區圍繞著預期的實體，像是點、線或多邊形，以產生出新的多邊形。這個新產生的多邊形可以在原圖中進行多邊形疊圖來選擇界定的區域（原文p.179）

### 由連結進行操作(Operations that depend on connectivity)

實體中有一些操作方式是直接的連結在資料庫中，這些連結是空間性的就像剛才的例子



A是直接鄰近B，或是A經由拓樸學的那些道路或其他流通過線模型的網絡連結B。實體也能經由內部拓樸被連接，所以複雜的空間實體能夠被由一系列的次實體組成，就像例子裡用物件定址一樣。

操作式A是直接鄰近B與A是經由拓樸網絡連結到B是同一件事約兩個版本。拓樸的連接線和多邊型使用從空間資料庫(見第3章)詳盡的資訊來決定兩個實體或位置是如何連結。內部實體距離在網絡或其他的連結的計量如旅遊時間、路線吸引力等之上，能被用來決定交互影響的指數。這些操作式常用於決定緊急遞送或是最有效的傳送路徑。

舉例來說。當一個邊界在兩個土地覆蓋的多邊形間也定義為一條道路，那是簡單的事去選擇出那些在兩邊有不同種類的土地覆蓋的路線。這樣的分析將容易從城市的道路(兩邊建成區)從海岸的道路(中識別鄉村的路(兩邊是農業區))。

在這拓樸學直接的網之上連線分析常用於自動的路線尋找(汽車卡車的指標系統)與最適宜的緊急服務位置。屬性附屬於代表道路、河流與鐵路等能標示出連結物的特性的線狀元素。舉例來說一條路能被標釋出的不只是因他的寬度、表面、等級或是車道線數。還有他的視覺吸引力或其他的東西(潛在的旅遊路線)與交通密度對路線資訊來說，一天或一週的交通密度的連續時間資料提供一個估計全天時間旅遊的基礎因素對緊急服務路線位置是重要的。圖7.12顯示出如何在網絡上從A到B路線可能靠道路屬性取出，而且可能相常的不同於從計算基於單一的緩衝區的烏鴉飛行航線。插圖2.2比較經由緩衝區從單一烏鴉飛行距離值計算旅遊時間與連結道路上的累積的時間值:這旅遊時間的比較依靠的是一天之中不同時間的通行方式，這是建成區緊急服務分析的一部份。插圖2.3與2.4顯示分析經由計算荷蘭西部使用私人轎車或公共交通工具旅遊時間忍受度的結果。

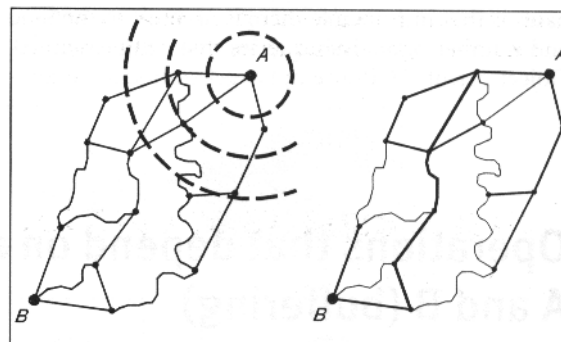


圖7.12：從A到B遷徙次數的分析於 (a) 烏鴉飛行距離與 (b) 網絡中沿不同路線飛行次數來界定在不同的路線中預期的旅行次數 (原文p.180)

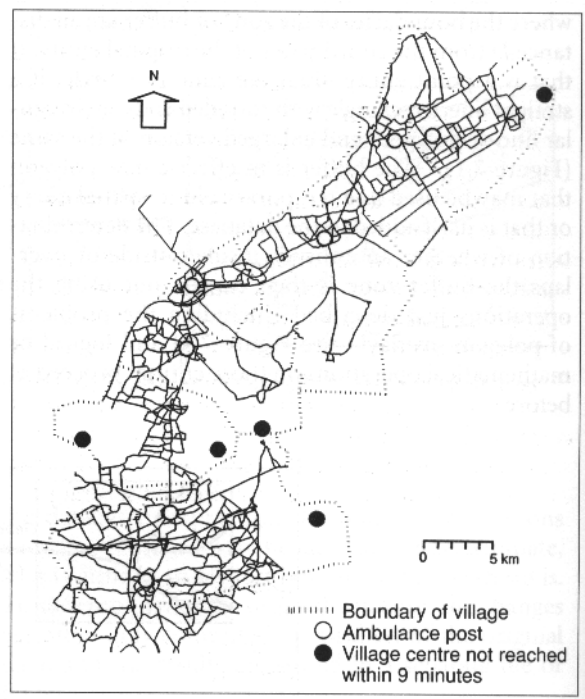


圖7.13：透過運輸時間分析結果來看市郊地區救護車可以在9分鐘從救護站抵達的區域。黑色線的部分表示救護車可以從救護站9分鐘內到達的地區。黑色圓圈表示無法及時到達區域。（原文p.180）

## 資料擷取及實體性質模式的面面觀(General aspects of data retrieval and modeling using entities)

空間實體能被取得而且新的屬性能經由廣義的邏輯與數值方法被計算。數值的程序也能被應用在拓樸學連接上包含、交叉問題、相似性分析與關連分析。此方法能被連到去創造定位許多不同種類空間問題的複雜模型，當它能被非常資訊的光坐在電腦前面瀏覽資料庫裡，看有什麼或不同的程序在作用時，注意那些分析的操作並不是交替的，所以指令中順序的執行是非常重要的。非正規的程序對單一資料復原與轉換是非常好的無論如何。當一個複雜系列指令必須被頻繁的使用去擷取與轉換資料，那麼去創造一個結構的指令檔是非常合情理的。可讓不同的使用者回顧、修改與使用。這樣一組指令組成一個模型或說是程序，能經由指定一個圖像或一個名稱被儲存在GIS中，並且能用在其他的資料庫完成同樣的操作。

沒有任何的分析方法出現在這一章，將注意力放在資料的品質與錯誤上，有一個未明訂的假設是所有的資料與關係都是精確而已知的。忽略掉這些錯誤，用GIS的空間模型是探索不同方案非常好的值。比起讓土地使用計畫者迷失在他們沒有或只有一點點經驗的情況中來說，去訓練他們運用地景數值模型還比較好一些。希望與駕駛在模擬器中飛行、建築師和道路設計者做模型來擴展他們的經驗與發展創意一樣，土地使用計畫者在鑄成地景上無可彌補的錯誤前，也能先從數值模型中學習到可能發生的錯誤。

## 第八章 連續資料之空間分析

### *Spatial analysis using continuous fields*

#### 前言

連續區域的範例為空間模式提供豐富的基礎，特別是當包括在不規則方形網格的資料中。單單只是信號數字疊圖一般容易以及促進數字化模式的形成，代數對應關係理論可得到數學運算可導出在所有網格的疊圖。在連續區域的數學運算可分為點及空間的運算。點的運算與第七章屬性的討論是一樣的；而空間的運算包括空間的過濾、表面衍生的計算（坡度、地形、凸面）、地面位向關切如觀測站、陰影凹凸起伏圖、照光圖來計算。本章節將會解釋這些運算式；理論藉由圖面的應用於水文沖蝕、以及地面逕流，充分運用出自森林的木材及在有機土中銻的连接如主要河流氾濫鄰近地區的機能。

如同在第二章的解釋，呈現連續區域的有兩種方法。第一種是Delaunay三角形（數化海拔模式的TIN）；第二種是更多同基質的高度或網格用來分析GIS網格及影像。Delaunay網路通常用在GIS的邊界以支持地下水呈現流動力學的有限元素模式（MODFLOW-McDonald and Harbaugh 1988），沖積扇（如Gee等 1990）或空氣品質（Fedra 1996）。有限元素模式（FEM）通常不是大部分GIS標準元件中的一部分，僅管Kuniansky與Lowther（1993）的報告中指出有限元素網格在向量式GIS中成立最大程式。使用FEM的Numerical模式與GIS使用來聚合資料與經由分界面的模式來說，對連結GIS通常有缺失。從模式而來的結果回歸到GIS，改變方形網格或線而容易掌握，然後整合數化圖進而展示之。

有限元素是本書範圍外之探討。這裡我們描述連續區域空間分析由規則方形網格所描述的運算，每一屬性是個別的圖層所呈現，以及每一網格是不一樣的數量值（第三章）。附註一點的是，僅管在許多案例中，我們通常參照海拔基質（altitude matrix）（如數化海拔模式），Z的屬性可描述任何連續變化屬性或地區的變化，如土壤污染物的層級、大氣壓力、每年的沉澱物、潛在市場的指數、生物密度，或地區的附加價值。

### 由不關聯的連續面進行空間分析的基本操作(Basic operations for spatial analysis with discretized continuous fields)

#### 地圖代數和製圖法模式(Map algebra and cartographic modeling)

在之前的章節中，Kisii土地推估案例證明了當一個使用來自好幾個地理圖層實體的資料時，如果所有資料參考相同的幾何學、規則的方格時，則對於維持數據資料庫和計算新屬性值是較為容易的。資訊的缺失導致網格平滑多邊形的邊界有更多比不去創造新多邊形的分支。而且藉由選擇網格尺寸去配合由遙感影像的使用可增加衛星資料的優點，也可以如資料分析與模式的輸入一般使用。如電腦在動力與資料的來源上的增加提供由網格組成的空間格局的削減的分析（比較Plate 1）變成比位向觀光大數字創造連接線與面更為容易接受。

更進一步來說，網格呈現在每個屬性的主要優點是記錄各個分立的圖層，是任何數學操作為同樣的網格施行在一個或更多屬性上容易應用在疊圖上。這意謂著可以使用精確的代數符號去操作網格資料。這理論稱之為地圖代數（Map Algebra）（Tomlin 1983,1990）以及使用的過程代數技術去建立空間分析的模式稱之為Cartographic Modeling (sic)。

地圖代數的理論意謂著使用者只需要具體說明空間運算的使用及資源圖層的命名及電腦程式的結果，然後應用運算在所有圖層的網格上。這使得寫電腦模式如連續計算變得更為容易，也以使第一點延伸至2次空間變得容易。舉例來說：

$$NEWMAP = MAP1 + MAP2 + MAP3 \quad 8.1$$

是所有需要計算在三個圖層MAP1、MAP2、MAP3的屬性值的總合，得到下列式子：

$$NEWMAP = (MAP1 + MAP2 + MAP3) / 3 \quad 8.2$$

計算其平均值後，再推至下列式子：

$$NEWMAP = (SQRT(MAP1) + SQRT(MAP2) + SQRT(MAP3)) \times 2 \quad 8.1$$

計算總合的平方根。

這些例子計算由小格組成基礎的新值：他們了解點的運算及相當於同樣數學運算運用於點、線、面的屬性（如Box 7.1）。當我們發現使用表面概念的差異性及平滑是可能計算某些地區的空間機能或鄰近地區網格的屬性時，特別是網格處理的優點變的清楚。最後因網格基礎處理是類似連續表面能決定表面上新的屬性，或選錄陡峭下坡路徑的位向關係。這些運算提供假說相當多的工具如橫過地區清晰度的變化及水文機制或充分運用通路於特殊地帶的種類。

從前面的討論，我們可以了解如果共同語言界面（CLI）允許使用者去表示在數學語言中基本空間機能然後去寫數學模式計算網格資料就容易多了。許多GIS提供簡單綱要的語言稱之為巨集語言（macro language）。在本書中我們假設所有的運算可能以簡單的、一般的、數學的表示之因此力學模式編碼用來解釋在變異系統例子中可容易成為工具。不是所有的地圖代數作者都熱於使用數學公式，然而，Tomlin（1990）為這些不能用數學方式編碼提供了一個英文語言方案。我們相信，僅管使用這些數學術語可以更為完美。

在本章我們可以檢視使用於網格資料的主要數學運算的類型，舉例說明運算的範圍-簡單及更複雜實際的例子。

### 點的運算(Point operations)

所有的邏輯及數字運算描述之前章節中的點、線、面連結成關聯式資料庫管理系統（RDBMS）可用來描述連續區域中個別的網格。同樣的網格在不同的圖層其值必然符合邏輯取樣、增加、減去，或任何允許資料型式的方程式的數學理論運算（圖8.1；圖8.2）。因此，可得到一個結論：增加或減去真實的數字，但卻不是如布林（Boolean）或名義上資料型式編碼的數字。布林（Boolean）或名義上的資料可由邏輯概念計算得之。

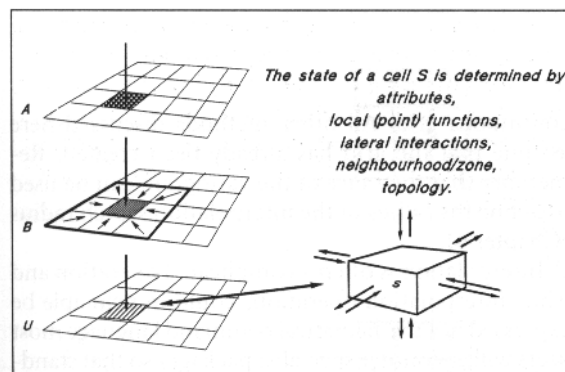


圖8.1：網格的狀態是位置運算的功能與鄰近地區的側面相互作用（原文p.185）

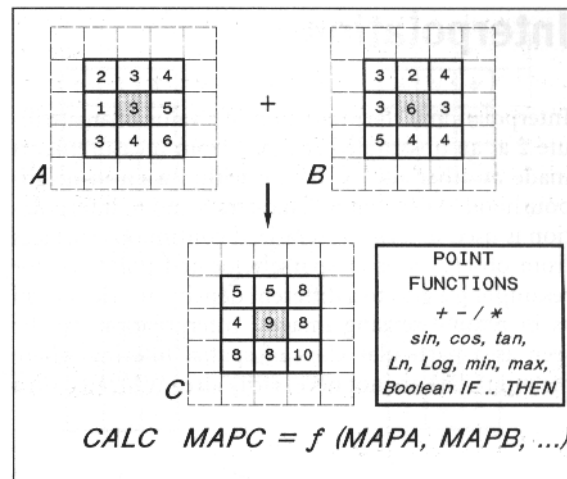


圖8.2：網格上點運算的案例（原文p.185）

## 空間運算(Spatial operations)

與位向關係連接向量式數據資料庫的定義相比較使用網格資料有其優點及缺點。缺點包括了沒有確切實體只有網格形成的大概及無法從地面的所有物去推論出位向關係的問題。優點則是連續區域模式提供一組有實際使用的空間分析運算。

接下來的操作計算網格新的屬性如某些確實空間鄰近地區網格屬性的機能。鄰近地區不一定是同質性的（如方形或圓形）。GIS中大部分的例子網格的尺寸是固定的且相同的。變化密度網格的選擇需修飾算法但不是本質上改變可導出的運算特性。這些空間運算包括：

- 內插
- 空間過濾
- 一次與高次導出
- 地表位向關係的導出：排水網及流域
- 擴展估計（成塊）
- 非線性擴張（衝突的展開）
- 視覺圖、陰影凹凸起伏圖、照光圖

## 空間插入推估(Interpolation)

內插是屬性 $Z$ 在不同地點 $(x_0)$ 來自測量在其它位置 $x_1$ 座落於鄰近地區值的預估。如同第五章與第六張，內插是用來創造描述來自於少數點或不同密度的取樣或遙感影像中連續的區域。內插如同空間過濾（spatial filtering）特殊的分類數入的資料通常不是已經座落在連續的網格中。這裡所有理論的討論假設網格是已經創造出來的。記得變異圖的範圍是可以用來定義內插搜尋半徑的範圍（第六章）。

內插通常是複雜的運算且內插運算在原則上可表示在數學語言中大部分的使用者會遇上特殊的組件因此標準的術語便不能使用。

## 使用方格的空間分析(Spatial analysis using square windows)

### 空間過濾(Spatial filtering)

最簡單以及可能最廣泛使用空間過濾理論的無關聯的、連續地面包括傳遞方形視窗（另

有核心或過濾)及計算視窗 $C_{i,j}$ 主要網格的新值如視窗覆蓋網格值的機能。此類計算通常如捲積、迴旋 (convolution) 所知。視窗通常是3x3的網格，但也有其他類形的網格尺寸 (5x5, 7x7或測量距離)。一般方程式如下：

8.4

當f標準的視窗計算則視窗邊為 $2m+1, 2n+1$ 。

$C_{i-1,j-1}$	$C_{i,j-1}$	$C_{i+1,j-1}$
$C_{i-1,j}$	$C_{i,j}$	$C_{i+1,j}$
$C_{i-1,j+1}$	$C_{i,j+1}$	$C_{i+1,j+1}$

通常使用視窗運算 ( $f$ ) 是低解晰度及高解晰度的區別。

### 平滑低解晰度過濾(Smoothing (low-pass) filter)

在視窗主要核心網格的值就如同其它網格值的簡單算法平均數的計算 (圖8.3)。在系統中使用真實數字的能力代表著能如十進位組成的真實數字計算之但在許多遙感系統中輸入及輸出被編碼成為整數，因此，在十進位的部分則小數點會被刪除掉。

在遙感系統及影像中分析代表值可由視窗的 $n \times m$ 值中網格值加減計算得之。舉例來說，3x3的過濾，代表視窗的主要核心可由增加每一個1/9寬的網格計算得到結果。在5x5的視窗每一個網格寬是1/25。特別權重可由輸入非線性權重值得其主要網格，如這些接近主要核心的網格有更大的權重值，像距離權至重在普通內插的概念。舉例來說，3x3的視窗：

1	1/15	2/15	1/15
2	2/15	3/15	2/15
3	1/15	2/15	1/15
	1	2	3

或5x5的視窗

1	1/65	2/65	3/65	2/65	1/65
2	2/65	3/65	4/65	3/65	2/65
3	3/65	4/65	5/65	4/65	3/65
4	2/65	3/65	4/65	3/65	2/65
5	1/65	2/65	3/65	2/65	1/65
	1	2	3	4	5

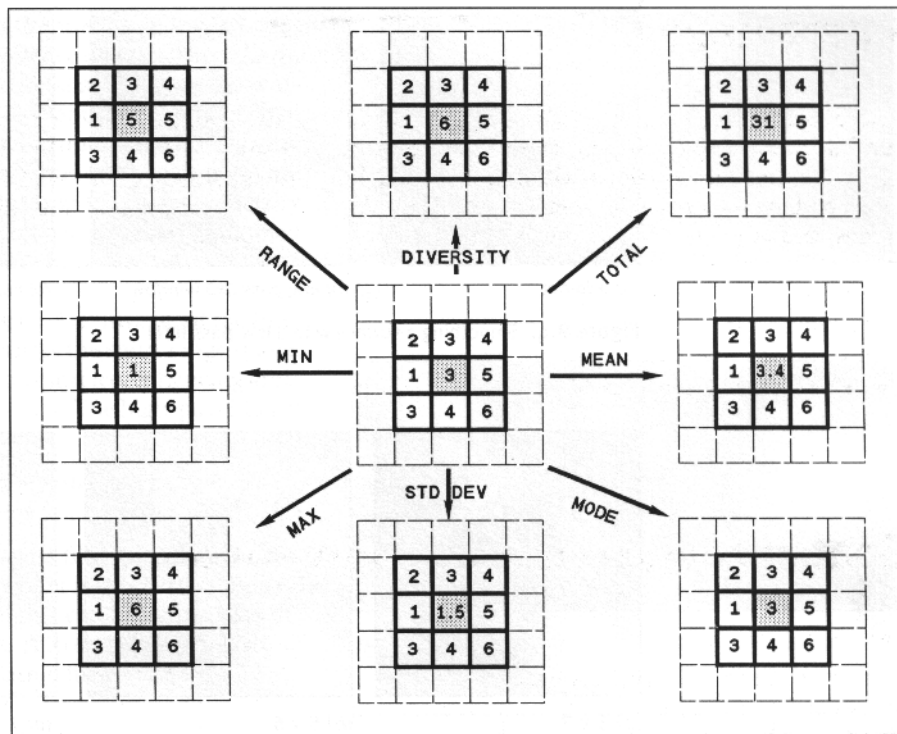


圖8.3：空間過濾的視窗運算（原文p.187）

這低解晰度有消去來自資料上極值的影響以產生平滑的影像（圖8.4；8.5）。對名義上與次序上的資料（整數及比率資料）來說這意謂可以由共同值的模式代替（多數的）。使用形式過濾在名義上的資料（如土壤單元）可以是簡化複雜地圖有用的方法（圖8.6）。但附註的是平滑一個形式過濾的網格影像比地六章重新分類屬性與合併土壤多邊形的綜合歸納步驟是不同的運算方式。

過濾的一般式子：

計算低解晰與高解晰過濾：

低 = 視窗平均數（連續表面-n）

高 = 連續表面 - 低解晰

當n是方形視窗在網格或距離單位的邊時，去計算形式過濾：

形式地圖 = 視窗多數（連續表面n）

在類似的方法，位置最大與最小值其不同在於其範圍可以容易的被計算出來。其不同點（在視窗中不同值的數字）或任何一個在視窗中四個軸線上兩個網格間的不同是具選擇性的。對少數名義上及次序上的資料來說以及不同點（在視窗中不同值的數字）是表示位置空間格局的複雜度有用的運算。每一步驟可由適合的操作者所運用（如text Box）。

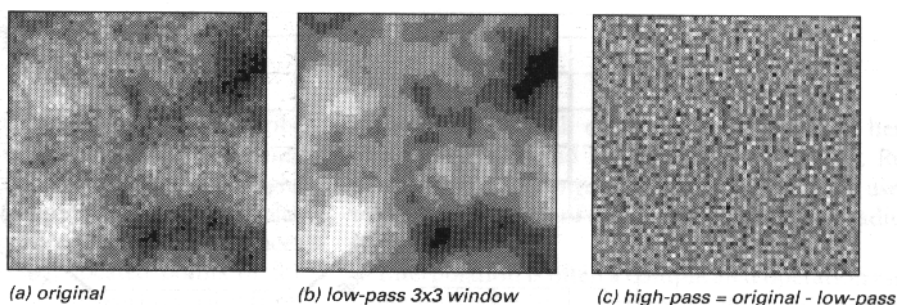


圖8.4：低解晰過濾的平滑表面（原文p.188）

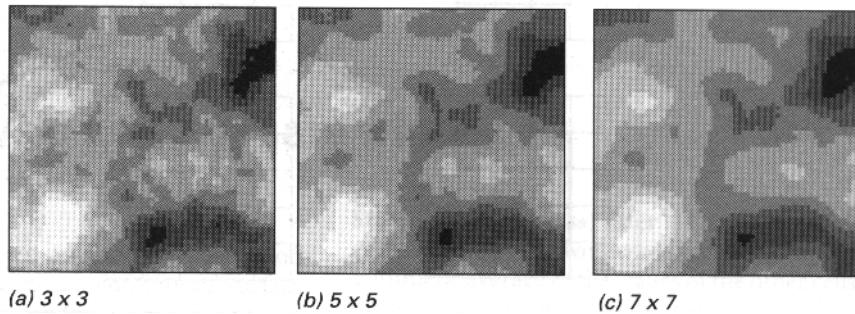


圖8.5：在表面增加視窗的影響（原文p.188）

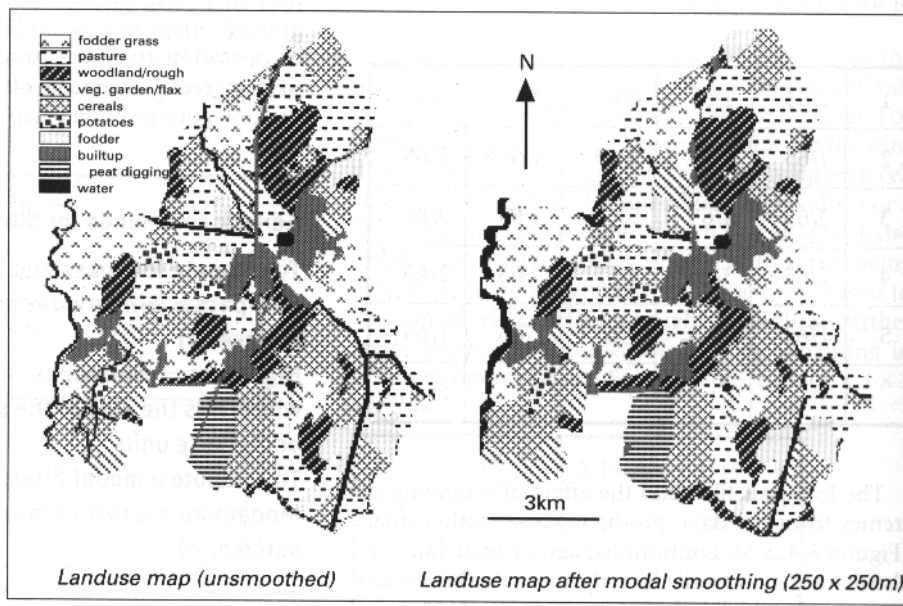


圖8.6：平滑的複雜多邊形圖隨著中間平滑聚合區域但不是減少分類級數（原文p.188）

### 高解晰邊緣及過濾(High-pass and edge filter)

低解晰過濾的相反就是提高連續表面所有物的短範圍空間，增加快速改變或複雜的地區。高解晰過濾被定義為：

$$\text{原來的表面} - \text{低解晰影像} = \text{高解晰影像} \quad 8.5$$

高解晰過濾的品質可以依賴低解晰過濾的定義。方案選擇性的，一組權重值可以定義視窗（Pavlidis 1982）。通常使用視窗組稱之為Laplacian filter：

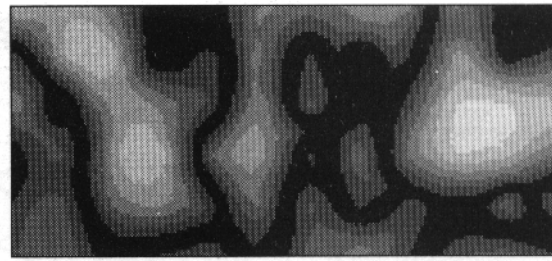
0	1	0
0	-4	1
0	1	0

圖8.7顯示應用邊緣去決定在連續表面最大位置改變的速率。邊緣過濾通常也是用來增加由遙感影像提供在連續統區相對相同的地區。尖銳邊的由來及邊界像素組通常用來推論可以符合確實性及向量要求影像中分立空間實體的存在。

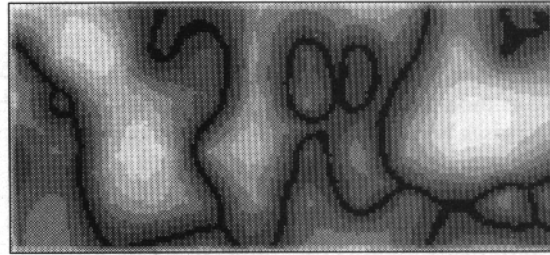
建立“邊界”邊緣過濾的影響：

- (a) 次序的表面
- (b) 高解晰過濾
- (c) 改變的最大速率產生的邊界

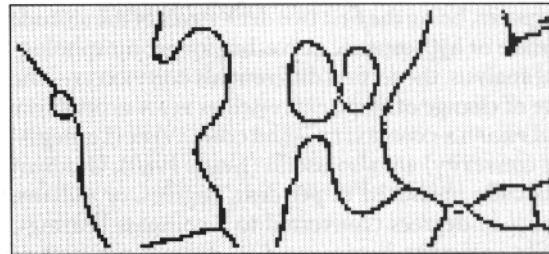




(a)



(b)



(c)

圖8.7：使用edge filter取摘取邊界（原文p.189）

## 由連續面導出一次和高次式面(First and higher order derivatives of a continuous surface)

因為網格表面在原理上於任一位置皆可導出數學函數，故被當作是數學上的連續，在實際操作中，由於其表面設定自由，故其導出的函數必需近似於不同的二次方程式計算，或適用於一個多項式函數。

兩個一次的函數是坡度和坡向，而兩個二次函數則是曲線的描繪及計劃（Evans 1980）。其中坡度被定義為一個水平接觸表面的切線，即利用DEM塑造的特定點且包括兩個元件，換句話說，也就是坡度，最大的屬性改變值；坡向指的就是最大改變比率的坡向。這些術語僅被以下的Evans（1980）專業上，許多作者以坡度表示傾斜度以作為定義（例如Peucker *et al.* 1978, Marks *et al.* 1984）及說明坡向（例如Marks *et al.* 1984）。傾斜度及坡向可以表現出許多意義，如屬性表面或高度曲線函數，但就地形學上的分析，以凸起、凹陷的描繪及計劃表現坡度變化也是相當有用的（Evans 1980）。傾坡度通常以百分率、度及弧度表示，坡向也轉換為指示性的度數，而凸起則被視為每單位的距離，例如度/每100公尺。

## 使用坡向函數估計坡度及坡向(Using directional filters to estimate slope and aspect)

高度曲線函數通常是計算每一個區域的高度矩陣，以3x3的小區或視窗在圖面上移動（圖8.8）。在點*i, j*最簡單的X軸傾斜度差別估計即為最大的向下斜度。

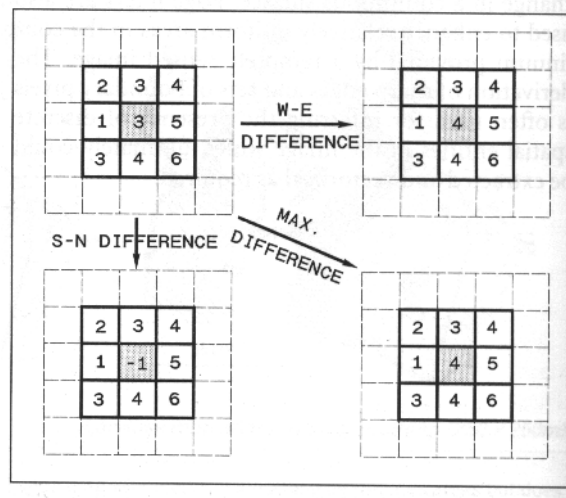


圖8.8：導出簡單的函數（原文p.190）

$$[\delta z / \delta x]_{ij} = \max (z_{i+1,j} - z_{i-1,j}) / 2 \delta x \quad 8.6$$

其中  $\delta x$  為離網格中心的距離（注意沿著對角線  $\sqrt{2}$  修正  $\delta x$ ）。這個估計方法的不當之處在於地帶性錯誤將對坡度造成嚴重的影響，因此許多二次函數的限定方法出現（Fleming and Hoffer 1979, Ritter 1987, Zevenbergen and Thorne 1987），以有限制的運算法則，適用視窗內鄰近的四個單位，其指出坡度為：

$$\tan S = [(\delta z / \delta x)^2 + (\delta z / \delta y)^2]^{0.5} \quad 8.7$$

其中  $z$  是屬性， $x$  和  $y$  為等位的軸線。

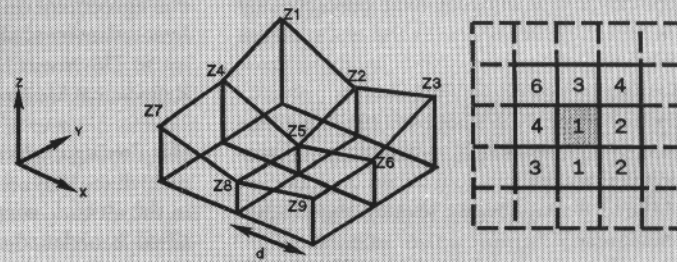
這個坡向為：

$$\tan A = -(\delta z / \delta y) / (\delta z / \delta x) \quad (-p < A < p) \quad 8.8$$

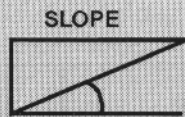
Zevenbergen and Thorne (1987) 表示如何以一個6個參數的二次方程式表示屬性、凸起及凹窪，其適用數據見Box8.1。

Box 8.1：使用Zevenbergen及Thorne's方法計算坡度（原文p.191）

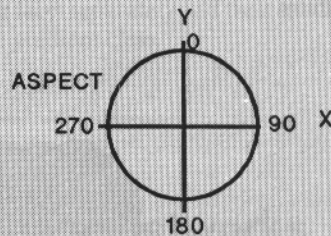
**BOX 8.1. COMPUTING SLOPES USING ZEVENBERGEN AND THORNE'S METHOD**



$$\begin{aligned}
 A &= [(Z1+Z3+Z7+Z9)/4 - (Z2+Z4+Z6+Z8)/2 + Z5]/d^4 \\
 B &= [(Z1+Z3-Z7-Z9)/4 - (Z2-Z8)/2]/d^3 \\
 C &= [(-Z1+Z3-Z7+Z9)/4 + (Z4-Z6)/2]/d^3 \\
 D &= [(Z4+Z6)/2 - Z5]/d^2 \\
 E &= [(Z2+Z8)/2 - Z5]/d^2 \\
 F &= (-Z1+Z3+Z7-Z9)/4d^2 \\
 G &= (-Z4+Z6)/2d \\
 H &= (Z2-Z8)/2d \\
 I &= Z5
 \end{aligned}$$



$$SLOPE = \text{SQRT}(G^2 + H^2)$$



$$ASPECT = \arctan(-H/-G)$$

Profile curvature



$$PrC = 2(DG^2 + EH^2 + FGH)/(G^2 + H^2)$$

Plan curvature



$$PIC = -2(DH^2 + EG^2 - FGH)/(G^2 + H^2)$$

concave = positive  
convex = negative

Horn (1981) 利用一個三次方程式估計是窗外部8個點為：

其東、西傾斜度為

$$[\delta z / \delta x] = [(z_{i+1,j+1} + 2z_{i+1,j} + z_{i+1,j-1}) - (z_{i-1,j+1} + 2z_{i-1,j} + z_{i-1,j-1})] / 8 \delta x \quad 8.9$$

而其南北傾斜度為

$$[\delta z / \delta y] = [(z_{i+1,j+1} + 2z_{i,j+1} + z_{i-1,j+1}) - (z_{i+1,j-1} + 2z_{i,j-1} + z_{i-1,j-1})] / 8 \delta y \quad 8.10$$

在3x3視窗中9個高程的點可選擇二者之一的方程式導出斜度及坡向。

使用不同的通用方法去計算坡度與坡向有助於了解何者是最佳的。Skidmore (1989) 回顧了6種估計坡度與坡向的方法，包括以上敘述的，他認為第二和第三個方法優於8.6的簡單運算，但Horn's方法與多項是方法再結果上會有少許不同。Hodgson (1995) 使用 Monison's (1971) 的綜合測試，以研究運算法則應用在4或8個鄰近位置，結果顯示數據在4個鄰近位置時坡度與坡向的數據較8個鄰近單位精準，如同Horn的敘述。近年來Jones (1997) 替8個計算坡度與坡向的真實物件和DEM綜合表面的運算法則帶出其他分析，當以RMS剩餘的錯

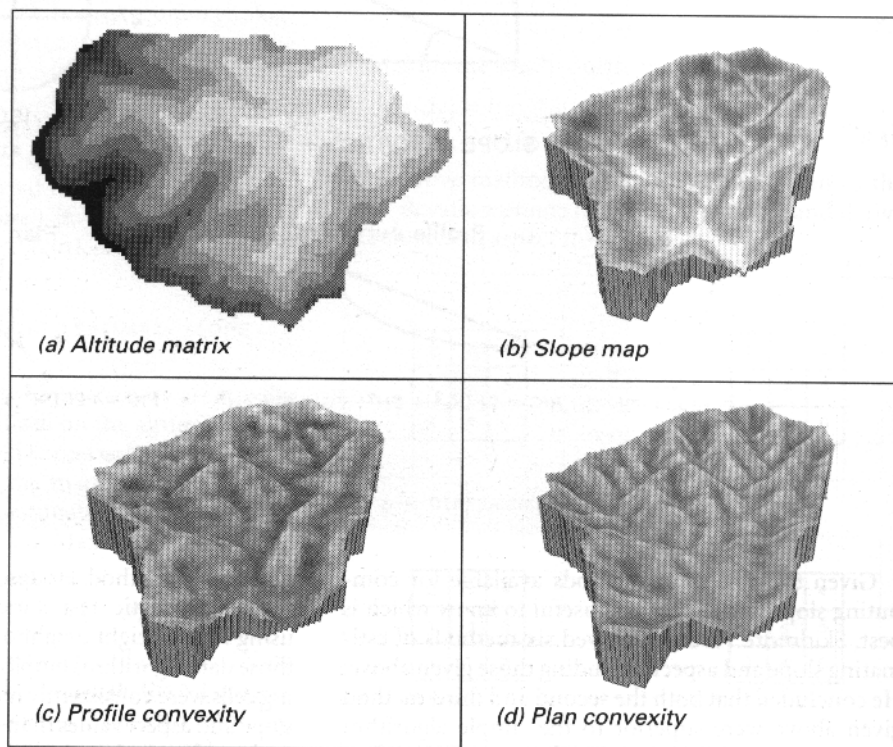
誤值導出不同時，以運算法則與真實表面測試坡度與坡向的結果，其先後順序如下：

1. Fleming and Hoffer (1979) /Ritter (1987) /Zevenbergen and Thorne (1987) 認為以四個鄰近單位可達最佳的平滑表面。
2. Horn (1981) 提出8個鄰近單位的方式優於一個粗糙的表面。
3. 'One over distance weighting'距離權重植是利用8個鄰近單位。
4. Sharpnack and Atkins (1969) 使用8個鄰近單位方法。
5. Ritter1987的對角線方法是採用四個鄰近單位。
6. 簡單方法'Simple method'有3個單元。
7. Travis等1975年最大向下坡度方法是利用中心點及8個鄰近單位。
8. 最小回歸二次方程式亦是利用中心點及8個鄰近單位。

可靠消息指出，Horn方法與Zevenbergen and Thorne運算法則被使用在幾個為大眾所熟知的地理資訊系統上，所以這較佳的運算法則具有一般的一致性。根據Jones指出，被認為是最不好的運算法則為最大向下坡度與複合線性回歸方法。

### 坡度與坡向的展示圖 Displaying maps of slopes aspect

在適當的導出最適合的海拔高度矩陣小區後，結果需要加以分類以達到清楚的圖面展示的目的，其查詢表格的定義通常以適當的分類和顏色或是灰值尺度加以判定，每一個查詢表格中的小區可進行比較，適當的灰度與顏色可輸送致展示設備。在今日具有高解析度的展示工具可以簡單的獲得肉眼可辨識的圖像，為了視覺辨識，展示主題數據（坡度、坡向等）以高度數值模式是相當有效的。圖8.9提供了坡度、凸面描述及小型及水區計劃的案例，可在30x30公尺的區域中有穩定的辨識率。



(a)屬性矩陣(b)坡度圖(c)側面起伏(d)計劃起伏  
圖8.9：DEM的一次和二次導出方程式（原文p.192）

坡向圖可被9種種類加以展示，其主要坡向有北、北東、東、南東、南、南西、西、北西，另一個則為平坦的地帶，另一個方法是利用循環的灰度值，以北東為最亮的，這可行成一個3D的結果。

坡度通常在不同的區域中有相當大的不同，即使在一個既有的分類系統中，通常也須符合統一的分類定律，最佳的刻度圖像製造必需掌握分類限制的意義與標準的分佈頻率。若有6個分類，則表示有 $\pm 0.6$ 標準差，則 $\pm 1.2$ 通常是相當令人滿意的結果（Evans 1980; see also Mitasova et al. (1995) 以原始的方法展示坡度資訊）。

通常由海拔矩陣導出的圖像特性其圖形是相當複雜的，在標繪前可藉由laser plotter或導出平滑的函數，利用光影或色彩圖形加以改良，平滑的DEM具有a low-pass filter，在計算函數時可以減少雜訊，但其結果將會移動極端的數據且低估了坡度角度，如果需要，其結果可利用網格的內插推估方法更加平滑的展現。

在系統中使用整數計算，因為數據量化的雜訊問題會使得海拔矩陣的插入推估網格過小，形成不明智的結果而導致坡度估計的數字問題（Horn 1981），利用輪廓線進行數化的插入推估可在導出的函數中嚴格的展現誤差（Chapter 5, Plate 4.5）。

坡度與坡向圖能以DEM的TIN法計算在每一個三角形的平面中，並可以根據傾斜度的分類描繪出來。

## 導出面的位相關係與灌溉導路網(Deriving surface topology and drainage networks)

在進行灌溉區域與灌溉導路網的分析前，必須先辛苦的拷貝航空照片或印刷的地理圖，此外這樣的工作不可避免會造成誤差的增加，在一般的區域中，並非一直十分容易的以肉眼去辨識一個流域的邊界，尤其是在多物的森林區中，即使是非常詳細的位相圖與灌溉導路網以藍色的線條加以顯示，亦有可能錯估了整個區塊中的水系路線。這時利用DEM的遙感探測獲得一年中不同時間的水流資訊，是相當有助益的。

灌溉導路網和水流、流域、分水線或山脊在真實景觀中視相當重要的，他有助於了解物質的流動，其可藉由TIN或海拔矩陣直接數化建立，但也可以自動的由海拔矩陣中衍生出（Band 1986, Hutchinson 1989, Jenson and Domingue 1988, Marks et al. 1984, McCormack et al. 1993, Morris and Heerdegen 1988）。水文學者提供新的工具使得灌溉導路網自動化估計景觀中的水流及沉澱，和鏈結地理資訊系統中水文學的動態模式。

當自動衍生灌溉導路網的海拔矩陣時，以下的步驟是必要的。

### 1. 決定路線(Determine routing)

網格表面上物質的流動是決定於最陡坡的坡向，這有幾個運算法則去計算他，包括D8或8點流動8-point pour運算法則、坡度加權slope-weighted運算法則、水流路徑stream-tube運算法則（Moore 1996）。

D8 (deterministic) 運算法則近似於以3x3個單元的視窗表示最陡坡度的坡向，這會自動的導致任意的坡向為 $45^\circ$ 的單位，這被一些人士看做相當嚴重的問題。以D8運算法則計算水流坡向，能夠形成8個不同的坡向值，其值也可以以度數或數值碼表示。最通常的實作方式是利用電腦鍵盤輸入數值，對於凹槽或坑洞以5為表示，其他的則以常規表示（Moore 1996）。

其結果新的網格疊圖被稱作區域排水坡向設定或ldd (Figure 8.10)，每一個往格皆包含

一個坡向的數值整數形式FD (flow direction水流坡向) 為：

$$FD = d \text{ where } d = f \text{ for } \max_{(f=1,8)} [W_f | z_{ij} - z_{i+1,j}, z_{i-1,j}, z_{i,j+1} | ] \quad 8.11$$

其中距離的權重 $W_f$ 對北、南、東、西鄰近位置是1而 $1/\sqrt{2}$ 是對角線。

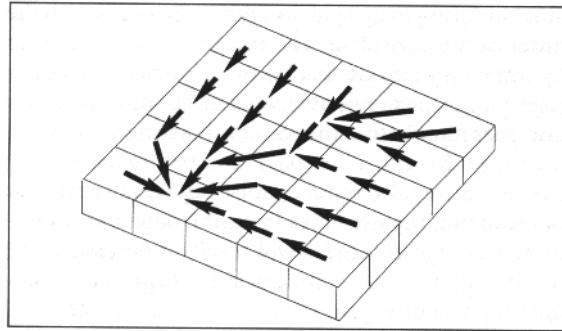


圖8.10：區域排水坡向向量指示下坡路徑 (原文p.194)

圖8.11b是ldd圖的案例可展示在DEM的背景上，因為簡單的D8運算法則結合了幾個商業用地理資訊系統，在一致的坡度表面它可以提供長而直的流動線和一致的流動坡向，且對於並非罕見的平行流動線不會令其交叉。他不能塑造流動的分佈。

Rho8 (隨機 (random)) 運算法則是統計學方法D8的一種形式，它可以較佳的展示地形上隨機的變化。其將 $w_f$ 中的對角線 $1/\sqrt{2}$ 以 $1/(2-r)$ 取代，其中 $r$ 為介於0至1間隨機均勻分佈的隨機變量。即使例如D8運算法則模仿散佈狀態 (見Desmet and Covers 1966)，Moore (1996) 仍要求這樣的模擬必需更符合真實的流動網路。在Monte Carlo的模擬中，也許D-8或Rho8皆能較為真實的呈現。利用DEM和D8運算法則以電腦儲存網路會增加RMS的誤差，他會在最可能的產生網路中重複100次 (見第十章)，Monte Carlo模擬優點在於DEM的誤差可以調整的更接近真實情形而隨機的流動路線也可以產生。

FD8和Frho8運算法則是為了修正原來運算法則塑造的流動及流域分佈。流動的分部可以複合的出現在最近區域 (nearest-neighbour) 的節點中，其情形包括在路地上流動或甚至集中於河道中，可使用D8/Rho8運算法則，其下游節點複雜的流動比率，可以坡度銓重為基礎進行運算 (Freeman 1991, Quinn *et al.* 1991)。

河流流動的方法。Costa-Cabral和Burgess (1993) 決定以河流來源像素面積的一小部份總合，加上每個像素流向，即線性指示所決定的灌溉坡向和像素的邊界進行計算見Mitasova and Hofierka 1993 (cited in Moore 1996)。

## 2. 窪地的移動(Removal of pits)

當一個平滑而連續的表面被趨近於一個正方形的格子時，不可避免的一些小區域將被周圍海拔較高的區域所圍繞，這些窪地可藉由真實的閉合而減少或僅利用人工化的網格過程。人工化的窪地通常會出現在一個狹窄的山谷，谷底的寬度小於網格的尺度，而他便能發生於所有程度辨識率的圖形中，他也能發生在插入推估面積時減緩誤差的過程中 (如圖8.11a)

人工化窪地的問題分散了灌溉位向 (drainage topology) 並必需移動以獲的ldd的連續網路，其移動方式有二，分別為切斷 (cutting through) 及填上 (filling up)。切斷一個或多個邊界網格層以尋找下游的網格，需要擴大尋找的視窗尺寸，才有助於找到一個或一系列同樣高低的窪地網格。一但這個路徑沿線的網格被發現具有適當的位像連接，其真實的高度便不

具關係了。而填上是增加核心網格的高度，使其鄰近的一個或多個網格相等，並測試是否鄰近的排水坡有其目的地，假若這沒有發生，則高度將持續的增加至鏈結被發現。

窪地的移動被視為是一些必需問題交互影響的過程，Hutchinson (1989) 發展了一種 spline 雲形規插入推估方法以使得窪地不會發生，但他並非總是可以察覺所有窪地自動的移動，因為關閉和半關閉是一些景觀中真實的特徵。MacMillan *et al.* (1993) 在加拿大亞柏達省有一個詳細應用說明，這研究是關於春雪融化時表面水的儲存，並依靠許多相互關聯的抑制條件。Van Deursen (1995) 提供一種結束窪地自動移動的方法，便是給這些受壓抑的體積、面積或深度甚至降水一個值（例如一些被壓抑的窪地或是水道的某些部分）。McCormack *et al.* (1993) 提供一個以特徵為基礎的趨近方法以賦予高原或受壓抑地區一個排水方向。

一旦窪地和高原被判識，DEM 便能夠進行調整。對於一個需要大量、複雜的數據設定大型面積網格，較實際的方法便是對其整個河流流域進行數化，並轉化其向量為網格，然後 burn in 輸入河流網格於較低階的灌溉網路中。

---

例如一般從網格 DEM 提取位向表面

$lddmap = (dem.map, a, p1, p2, p3, p4, \dots)$

其中 lddmap 是由位向衍生，a 是運算法則的使用，而 p1, p2, p3, ... 是移動窪地根據流動深度、核心體積、核心面積等的參數。

---

## 使用 ldd 路網進行空間分析 (Using the ldd network for spatial analysis)

運算法則的使用與計算與流動坡向是沒有關係的，其結果是在原本的網格上，重疊一個清楚的位向表面（如圖 8.11b），ldd 網路對於計算其他數值高程模式是相當有用的，因為它可以明確的包含不同網格連接度的資訊。這些資訊使其能夠指出液體或物質流動的方向，在沒有 ad hoc 尋找視窗中。

## 網路上物質累積的流動 (Accumulating fluxes of material over a net)

因為在位向矯正網路中，每一個網格會連接下游的鄰近區域，縮以計算屬性會變的相當容易，例如每一個網格物質劉經的累積總數。網格新的累積運算估計情形，為原網格的總合，加上上游元素流動（upstream elements draining）的網格

$$S(C_i) = S(C_i) + \quad 8.12$$

假如每一個網格值為 1，其結果會產生上游的元素圖，或換句話說，累積的上游網格數在目前的網格中不會改變。這上游的元素圖通常展現於數對的尺度中（圖 8.11c——也可見 Plates 3.7 及 3.8）。

假如這物質的值被其他的疊圖所提供，例如有效降水，其總合的計算將會累積於一個理想表面，如他可以很容易的計算大規模的每一個均衡網格以：

$$S = P - I - F - E \quad 8.13$$

其中 S 表示的是剩餘的水量，P 是總降水量，I 是流出量，F 是入滲量而 E 是蒸發量。這網路上累計的流量可藉由鏈結網格上的 S 的值計算出。位向網路亦是 GIS 動態模擬工具的寬度範圍基礎，van Deursen and Burrough (1998) 對其有更詳細的解釋。

上游元素圖本身可以被用來作其他地形指標的計算，例如降雨指標圖可以被定義為：

$$wetnessindexmap = \ln(A_s / \tan \beta) \quad 8.14$$

其中A是以 $m^2$ 為單位的流域面積（上游元素數目 $\times$ 每個網格面積），而 $\beta$ 是以度為單位的坡度值（Beven and Kirkby 1979），圖8.12a是一個由DEM衍生出的降雨圖。

水流量指數（The Stream Power Index）被定義為：

$$\omega = A_s \times \tan \beta \tag{8.15}$$

這個直接的水流能量比例 $P = \rho g q \tan \beta$ ，其中 $\rho$ 為水的密度， $g$ 為重力加速度， $q$ 每單位寬度的水流變化（Moore *et al.* 1993），這個時間上能量的消耗率被視為陸地上水流的充實能力。

沉積物運送指數被定義為：

$$\tau = [A_s / 22.13]^{0.6} \times [\sin \beta / 0.0896]^{1.3} \tag{8.16}$$

這個指數的特性在沖蝕與沉積的過程，特別是對土壤流失的影響，他類似於長度與坡度的因子在一般的土壤流失關係式（Universal Soil Loss Equation）中（Wischmeier and Smith 1978），但可適用於三度空間的表面。圖8.12b表示沉積物運送指數，能夠沿著一個水流長度變化。

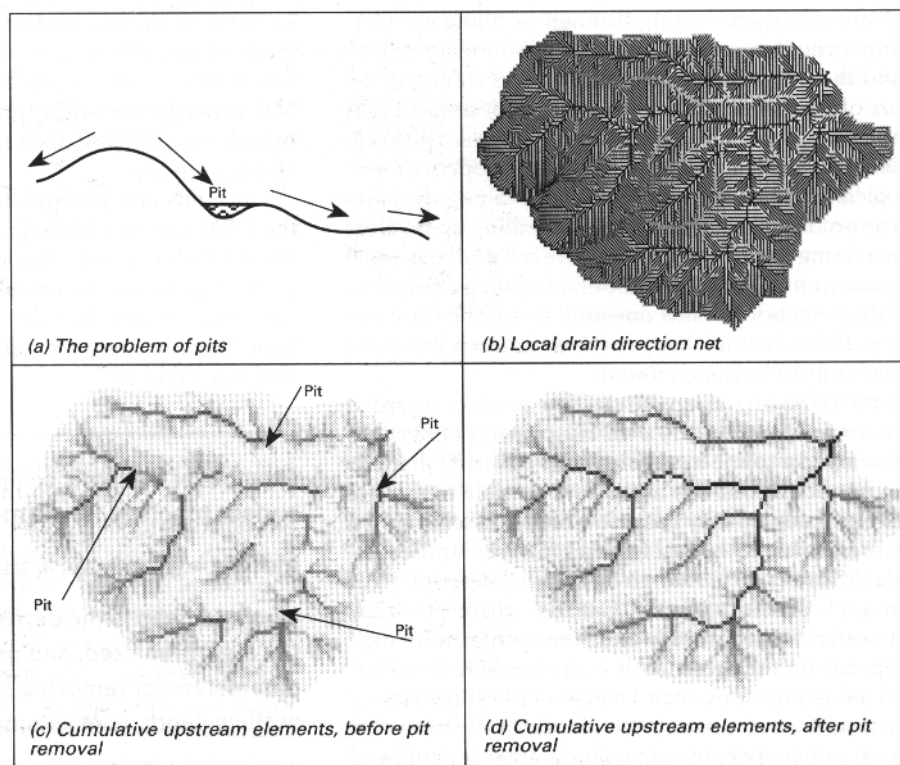


圖8.11：由DEM網格導出集流網路(a)坑洞問題(b)區域排水方向網路(c)在坑洞移動前累積的上游upstream元素(d)在坑洞移動後累積的上游upstream元素（原文p.196）



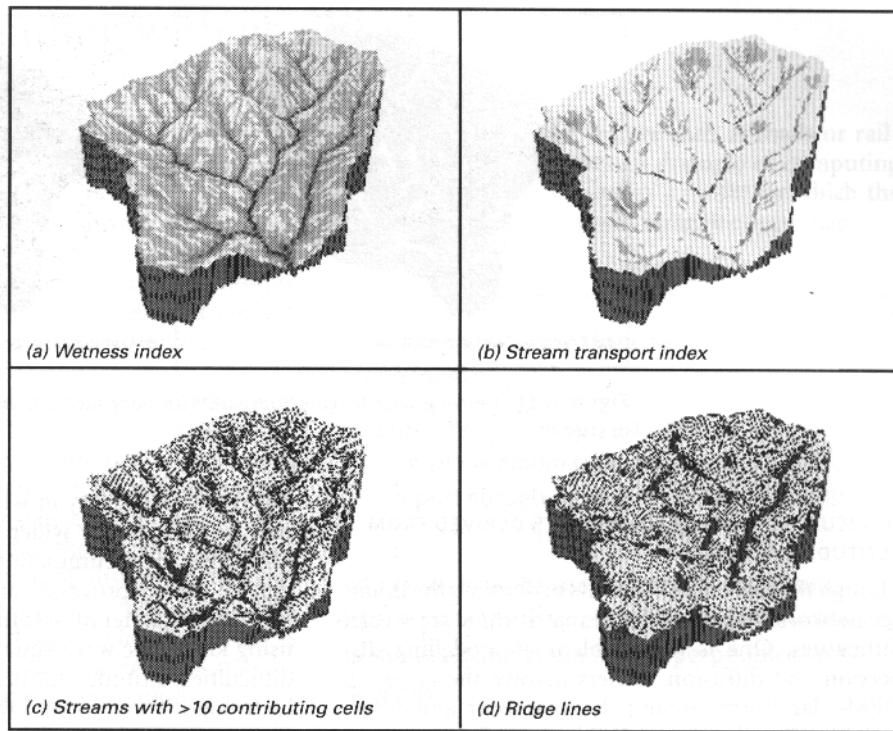


圖8.12：游擊流網路衍生產品(a)降雨指標(b)河流轉變指標(c)以>10 contributing輔助網格表示河流(d)山脊線 (Ridge lines) (原文p.197)

**從區域灌溉方向圖中衍生的其他產物(Other products that can be derived from the local drain direction)**

河流水道能夠被定義為擁有超過N contributing上游元素的網格。河流水道亦可被布林運算定義為：

$$\text{Streams} = \text{if} (\text{upstream elements} \geq 50 \text{ then } 1 \text{ else } 0) \tag{8.18}$$

這建立了一個具有50或更多上游元素的二元圖，並被定義為屬於這條河流 (Figure 8.12c)。

**山脊(Ridges)**

在定義上，山脊並沒有上游元素，所以選擇所有UPL值為1的網格，以提供最初山脊的估計 (圖8.12d)。

**集水區(Catchments)**

因為所有被流經的網格被視為流域網格的一部份，所以在Idd自動計算中會計算其面積並定義其網格的流域。在流域中可以將所有流經網格定義為1值，其他外部者為0，這可以作為cookie cutter (餅乾切割器) 以由遠遠的感覺出流域特性的數值，或作為其他相同圖形辨識率的來源。使用高的傳遞high-pass或邊緣過濾可提供一個線性的流域邊界，其會轉換網格表示法為一個向量形式的鏈碼 (第三章)。

坡度長度運算近似於累積運算，但他基於網路中距離移動的原有網格值總合及上游網格的累計，計算出一個網格新的屬性值。

$$S(c_i) = S(c_i) + \sum_u^n (c_u \times d_u) \tag{8.17}$$

圖8.13a表示出這種坡度長度計算的方法。其距離可以依據網格的尺寸以歐基里德的距

離表示（長寬為1單位的網格，其對角線為1.414），或包括一個相衝突的術語在網路的網格中。

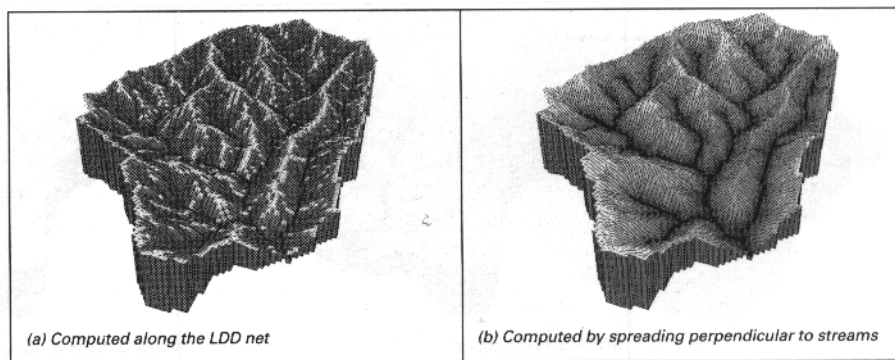


圖8.13：以DEM導出坡度長度(a)根據LDD(b)以垂直於河流分散程度（原文p.198）

### 由高度矩陣衍生出灌溉網路的困難(Difficulties with drainage nets derived from altitude matrices)

儘管以海拔矩陣推演出灌溉導路網路有許多利益，但也有一些困難，一是在散佈與蔓延方面，另外是DEM所造成的錯誤，因為地形並非永遠平滑且可以微分的。網格的尺寸選擇也許會影響到坡度、坡向及水流連通度的估計。在海拔矩陣中，河流總是一個網格寬，但真實的河流其長寬變化會比網格尺度更寬或更窄，塑造流動假設的累積趨近於重力所衍生的過程，且忽略快速流動的水體（近似於運動學上的波動等式）所產生的慣性。忽略這些困難，DEM的海拔矩陣被發現是極佳的水文學工具，並在應用過程中導出有用的結果，其在本章的第二部分有舉例說明。

### 團塊化(Clumping)

許多布林選擇或網格屬性分類的結果，在網格設定上表示連續，但這並不能被定義為空間中的實體。接下來每一個網格的計算檢驗是為了看在3x3的視窗中，使否具有相同的分類，如果如此，網格將被視為相同的團塊，並被賦予一個區別其不同於其他的特性值。這結果使得每一個連續的網格團聚集成更大的空間單位，對許多目的來說更有用處。例如在利用上游元素圖辨識所有的山脊時，其也許會創造幾個屬於不同山脊的網格集合，利用團塊特性可對團塊進行計算。

### 具有或不具有不順資料的緩衝區化/擴展(Dilation/spreading with or without friction)

這不是一個視窗的運作，但是一個實體上擴展或緩衝的連續分析。然而，確切實體的擴展或緩衝，通常受限於等方向性或同形的延伸（一個緩衝區，較原本的範圍大-見圖8.14），一個連續表面的擴展卻不能展現對於擴展過程中的阻力有其不同類的變更（圖8.15）。

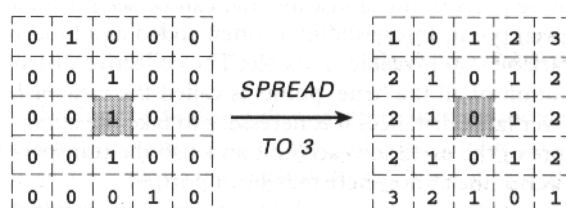


圖8.14：等方向性分散的網格（原文p.199）

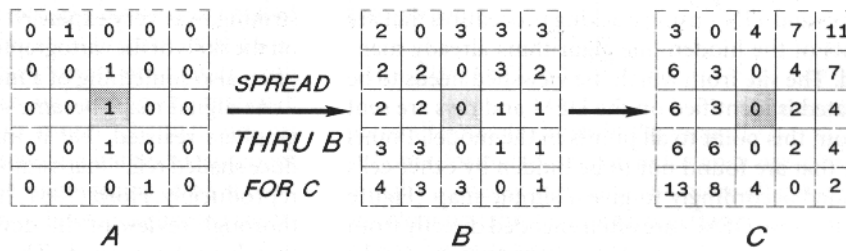


圖8.15：經過阻力函數而分散（原文p.199）

在非等方向性的擴展中，有兩個元件的值可由起點中累積。第一是距離，可藉由小區不數或真實單位計算。第二個依靠單元屬性的是累積的距離。當通過一個小區，緩衝區屬性最及距離累積大值，其結果的影響力較其阻力為大，所以幾何學上較長的路徑也許會被較簡易的方法，達到一個約定值。

緩衝的區化計算

Spread map 延伸圖 = spread 延伸 (starting points 起始點, v. friction 摩擦力)

其中起始點給予開始延伸或緩衝一個起始的定位，v 是一個最初的值，而摩擦力是一個個以網格為基礎其內部的阻力。

簡單或摩差力型式的延伸可被用來估計坡度長度，其垂直於河流網路，可由上游的元件圖中衍生出（或任何其他線性的特性如公路或鐵路），圖8.13b 給予一個計算 'distance from the stream 與河流的距離' 的案例，是使用坡度圖計算緩衝區。

Commands to create slope length and inverse 建造坡度長度及其倒數的指令

slope length 坡度長度

slope length = spread (strm, O, sip. map);

slmx = map maximum (slope length);

report slope length = ((1 - (slope length / slmx)) \* slmx);

其中 strm 是圖形中河流的定位元件，sip. map 是由垂直於河流的坡度長度所導出的坡度及坡度長度圖。

非等方向性的延伸會在連續表面上產生一個自由洞 pit-free，排水網路即為一個表面每一個網格被設定至起始點為最適路徑，計算及水區及上游元件在這樣的表面，可沿著最佳的路徑指示得出。

取代所有網格的延伸，非等方向性的延伸可被沿著 ldd 圖或其子集的路徑加以定義，當抗力多重複合於其上時，順著問題的累積，面積可在網路上被顯示。

視覺圖、陰影凹凸起伏圖和照光圖 (Viewsheds, shaded relief, and irradiance)

三個即為相關的方法涉及光線的光源和 DEM 上的路徑計算，並影響其他的定位元素，然而，所有的空間分析方法皆討論網格的屬性值，或在水平網格的疊圖中不同的屬性值，下列的計算可建立新的屬性，為三種連續表面上的空間形式。這些方法目的是計算視線（決定

視覺圖)，或藉由陰影展現一個類似3D的情形，及計算每日或每年進入的太陽能量。

### 視線圖(Line of sight maps)

這是一個最簡單的運算，目的在確定由約定點所看見的景觀範圍，可見度通常被編碼為一個二元的變量—0為看不見的，1為可見的。所有真實點的集體的散佈被稱作視線。在相當大的距離中，他必須要計算地球的曲率，而大氣的透明度也是相當重要的。

藉由傳統的輪廓圖去決定可見度intervisibility並不容易，因為大量的數據資料必需進行摘錄及比較，可見度圖可被屬性矩陣及使用TIN的追蹤過程製成，其引長線運算法則hidden line algorithms以在先前提及。這視線的基地需要在DEM上被有計劃的定義，而光線將由一點向模型上其他點送出。點（網格）被找出而不是被其他點隱藏並進行編碼成為一個簡單的圖（圖8.16）。因為DEM's通常由航空照片值間編碼，高度的記錄也許不被計入特徵，例如在真實地形中的森林或建築物，其結果必須要小心的解釋。在一些案例中，景觀元素的高度被建入DEM中以了解他們在景觀中對於視線上的影響，而是線也可由有計劃的TINs計算出（De Floriani and Magillo 1994; Lee 1991b）。

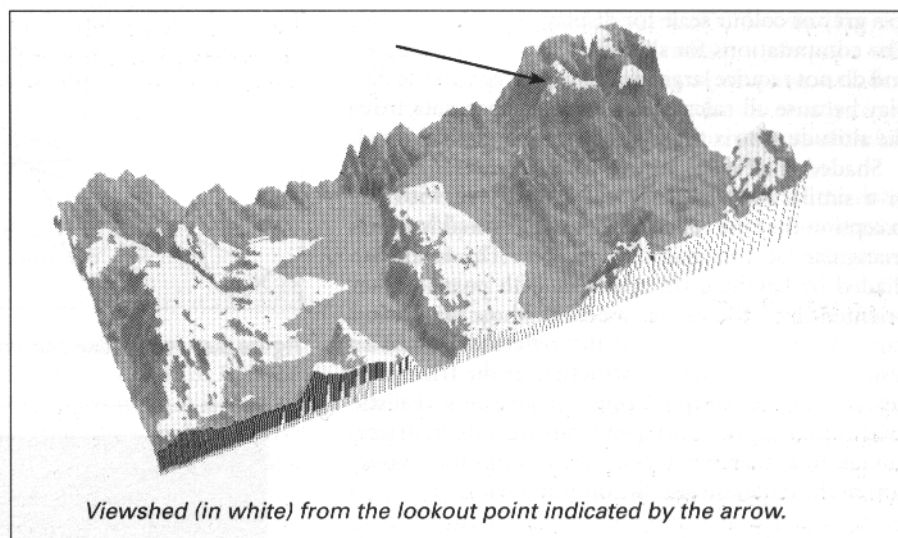


圖8.16：從DEM上一個點看出的視線圖（原文p.201）

在模擬訓練飛行、微波傳送站位置、景物辨識、森林火災警戒站等方面，估計基地的可見度是GIS一個重要的用途，DEM計算視線的誤差影響已被Fisher（1995）加以研究。

### 陰影凹凸起伏圖(Shaded relief maps)

製圖者已經發展許多技術以改進圖像的品質，特別是一些多丘陵與多山地區的描繪。凹凸起伏的描繪是其中一種相當成功的方法，他大量發展於奧地利及瑞士的製圖學校，並有其明暗配置的根源，而文藝復興時期的藝術家也使用光線及陰影去表現三度空間的物件。這些手繪的技巧有賴於繪圖方法與噴霧技術以符合需求，酸然其成果在外觀上相當驚人，但也需耗費龐大的經費並仰賴製圖者，令人質疑的是，有些人通常是登山者。

當數化圖形變的可能，許多製圖方法能夠自動精確的產生並複寫陰影凹凸起伏圖，Horn（1981）給予被試驗過的發展與方法一個廣泛而詳盡的回顧，自動陰影凹凸起伏圖的原理基礎是一個將地帶看似什麼理想化的模式，照光圖便是一個約定的位置。其最後的結果（如第

五章, 圖5.12b) 會因為灰色的尺度及連續區域的展現技巧而看似一張航空相片, 但陰影的凹凸起伏圖可利用航空照片屬性矩陣不同, 以許多方式加以計算。第一, 陰影的凹凸起伏圖度能展使區域的覆蓋, 只能數化土表面面; 第二, 光源通常必需選擇地平面上45度角的西北方位, 這樣較有利於人類察覺; 第三, 這個區域模式通常是平滑且一般化的, 因為航空照片中數據集合的過程將不會展示之前的細微數據。

陰影凹凸起伏圖能夠簡單的被製造, 且所有的圖形皆被要求估計以知表面元素的方向(例如坡度單位), 及當光源由西北方45度角高照射而來時, 其其表面元素如何反映光線。表面元素光影的展現有賴於重視光源的方向及其元素的材質, 有光澤的表面會比有孔隙或無光澤的表面反射較多的光線。在陰影凹凸起伏圖的發展中, 最常被討論的是如何計算反射比的問題(Horn 1981)。

根據(Horn 1981)所延續的方法是充分的生產合理且據品質的陰影凹凸起伏圖, 首先, 第一個步驟是計算坡度 $p$ ,  $q$ 為每一個網格 $x$ (東-西)及 $y$ (南-北)的方位, 如算式8.9及8.10。這些值將在反射比圖中reflectance map被換算成反射比的值。這是一個以圖表方式表現坡度 $p$ 的方式, 而 $q$ 則作為反射比模式之用。Horn建議採用以下公式以獲得較好的結果:

$$(i) R(p, q) = 1/2 + 1/2 (p' + a) / b \quad \text{當 } p' = (p_0 p + q_0 q) / \sqrt{(p_0^2 + q_0^2)} \quad 8.19$$

此為位於光源遠方的坡度。為標準製圖位置(西北方距海平面45度角)  $p_0 = 1/\sqrt{2}$ 及 $q_0 = -1/\sqrt{2}$ 的光源。

參數 $a$ 、 $b$ 允許水平面灰值選擇, 且關係著灰值改變的變率, 如 $a=0$ , 則推演 $b=1/\sqrt{2}$ 。

$$(ii) R(p, q) = 1/2 + 1/2 (p' + a) / \sqrt{(b^2 + (p' + a)^2)} \quad 8.20$$

圖上所有可能坡度皆在0-1之間。

一些反射比的計算是相當複雜的, 對於轉換坡度成為反射比, 較有效的方法是建立一個查詢表格, 並將每一個網格的反射比值轉換成灰色或彩色尺度加以展現(如圖5.13)。陰影凹凸起伏圖的計算並不複雜或需要大量的記憶, 除展示外, 因為所有的計算不需從屬性矩陣中取太多不必要的數據去填滿核心。

陰影凹凸起伏圖是TIN DTMs以一個簡單的方法描述一些例外, 每一個三角形表面取代網格決定反射比, 其表面的陰影以平行表面坡度的線表示, 並以分離的距離表示光驗強度, 其結果強烈的保持三角形網格結構, 而在作者的觀點中, 雖然大部份是依靠圖形辨識率或輸出工具, 它比屬性矩陣有較少的真實影像。

## 陰影凹凸起伏圖的應用(Applications of shaded relief maps)

陰影凹凸起伏圖可以確實有用地用在三維方面正確地描述地帶符號影像的存在。不只是在日光系統是有用的, 也可以在地形數值分析上做應用。當使用合成主題資訊時陰影凹凸起伏圖可以提高地圖的寫實性。

## 照光圖(Irradiance mapping)

這是陰影凹凸起伏圖原則的延伸去計算日光直接落在表面上的總數。太陽在天空中不是固定在任何一個位置, 但可以根據緯度、一天的時數、一年的天數得到位置。這需去合併能量實際收到的總數中大氣吸水性的影響, 以及形式化地點陰影影響(天空視察因素), 其影響是考慮在冬天或一天的開始、結束丘陵景觀的重要性(Dubayah & Rich 1995)。散開的照光如計算地方局限的確實影響一樣的困難。

在位向關係詳述資訊上照光圖模式可以在Dozier (1980) , Dozier & Frew (1990) , Dubayah (1992) , Kumur (1997) 中發現及參考引用。在任何點的直接照光吸收是日光頂點角度、大氣頂端的日光流動 (exoatmospheric flux)、大氣透射比、照光角度、及天空覆蓋度的機能 (圖8.17)。

每天與每年時間中頂點角度和大氣流的變化，以及大氣透射比變化就像大氣吸收者與消散物 (雲、塵) 的複雜機能一樣。大氣透射比也會因吸收者及消散物數量的減少隨著高度而增加。

大氣透射比 $T_o$ 與在坡上直接照光度 $I$ 的式子表示：

8.21

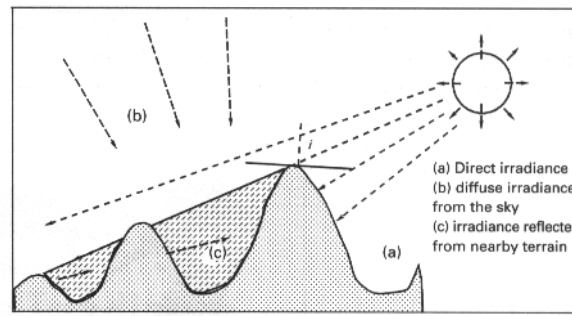


圖8.17：計算太陽日照的照光度 (原文p.202)

$\cos_i$  是坡上照光度的cosine， $S_o$  是大氣照光流動， $\theta_o$  照光頂點角度， $\phi_o$  是照光方位角， $A$  是坡的方位角， $\beta$  是坡度。因 $B$ 與 $A$ 可以由海拔資料獲得，方程式8.19描述了照光度優勢成分的空間變化。如一天一年的變化，去計算照光時間及空間變化是可能的 (圖8.18)。附註的是天空視察要素限制了當太陽落在地平線上的直接照光以及計算 (如直接照光只有在直接光照地帶或太陽的視覺圖 (viewshed) 是可計算的)。

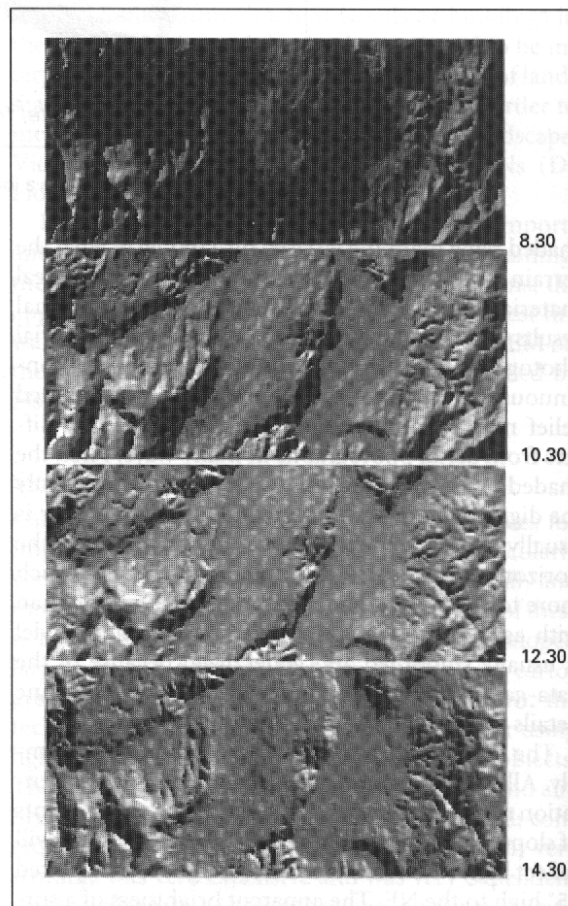


圖8.18：直射（一天當中各個時間）的變化（原文p.202）

在附註一點是去綜合歸納季節或年每月每年照光推估的事情是容易的，因此去創造一個區別植物成長能量輸入位置圖、家用暖氣、或岩石風蝕是容易的。綜合歸納一個暖流分級圖與雨天雲的位置圖可以由上遊元素獲得（如方程式8.14）暖濕、暖乾、冷濕、冷乾連接的圖。因此土壤資訊的結合與分類的植物指標遙感影像可以快速地提供詳細搜查假設產生與野外實地調查計畫的景觀模式。

### 其他以網格為基礎的分析方法(Other cell-based analysis operations)

以網格為基礎的連續區域運算在GIS與影像分析有大部分的交集。其他類似型式的運算可在數學形態學區域（Mathematical morphology）（Gutowitz 1991,Takeyama & Couclelis 1997）發展（Serra 1968）及細胞化學（Cellular Automata）。

時間的改變。隨著網格資料同樣的方法在本章及第六章的屬性計算是容易連接運算的，因此提供了一個動力機制的模式（Van Deursen & Burrough 1998）。大部分的細胞化學（Cellular Automata）運算使時間改變更為複雜。

表8.1：由DEM所計算的屬性概要及其應用（原文p.204）

屬性	定義	應用
高層	海平面或地平面上的高度	潛在能量的應用：氣候變化-壓力、溫度、植生、土壤、物質量的加、減計算。
坡度	海拔變化的速度	地帶的陡峭、陸路及地下涇流、土地能量分級、

		植生、上坡阻力、遙感影像的校正。
地形	坡的凹凸方向	照光、蒸散作用、植生、遙感影像的校正。
曲面	坡改變的速度	流動加速度、腐蝕/沉積物增加的地區、植生、土壤及土地海拔指數。
平面	地形改變的速度	會聚/分歧溪流、土壤水
排水網Ldd坡向	陡峭下坡溪流	如溪流位向關係機能的排水屬性計算，評估當物質轉動的網路。
上游元素/地區/ 特定流域	上游每個網格的輪廓寬度	上游流域的位置、流域排水物質容量
河流長度	上游路網最長的道路	流動加速度、會聚速度、沉積物的產生
河道	流動水與上游元素的數值	溪流飽和度、溪流位置、會聚/沉澱作用
山脊	非上游組成的網格	排水分界、植生、土壤、會聚、地理分析連接
雨量指數	Ln (特定流域/坡度)	濕氣滯留指數
河流動力指數	確切流域區域*坡度tan	會聚測量陸路溪流的動力
沉澱物輸送指數(LS Factor)	$(n+1)\left(\frac{A_s}{22.13}\right)^n\left(\frac{\sin \beta}{0.0896}\right)^m$	會聚與分歧機制的特性 (如USLE)
視覺圖	內部清晰度的地區	微波傳播塔的觀測站、觀火站、旅館、軍事應用。
照光	日光能量投射每一單位地區的總數	植生、土壤、蒸散量、儲存能量建物的位置。

## 使用連續面和產出的作業方法、綜論(Summary of operators that can be used on continuous fields and their products)

本章已說明了產生的一個大範圍可由個別規則格子的連續面導出 (表8.1)，某些導出的資料，如：坡度、方位、和視線、丘陵陰影和光亮皆可輕易地從不規則三角形網路去獲得，但尋找系統使用TINs所提供的功能去導出表面的位向關係及計算物質流動通過這些被導出的面，則較不常用。常常，TINs的使用者會將例如位向關係相關聯的線或物件等，明確地輸入他們的網路。

大部分一般所碰到的連續面是數化的高層模式，且大部分導出所提及的，在地形高層的使用和說明上有著直接的關係，所呈現的作業方法可被用在任何的連續面上，因此，如遙測影像或內插的結果，或是空間模式化，將於接下來的例子中說明。

## 連續面的空間分析應用(Practical applications of the spatial analysis of continuous surfaces)

本章剩餘部分將呈現幾個例子，說明在GIS系統中包括空間操作於連續面的額外值。作業方法的準備可被用於分析空間的交互作用，及為表面位向關係的創造打開了許多新的可能性，尤其那些牽涉物質或流體穿越空間的輸送。水文學首先要思考水經過排水池和集水流向為何處，這是一個步驟的重要部分，但有很多其他的地區，加入空間的交互作用，只在點或一個實體基本原理 (成塊模式lumped models) 將先前的操作模式化，提出新的敘述性力量和



洞察力。本章的重點，除了估計的重要外，我們限制討論在空間交互作用的靜態模式，連結時間序列資料和空間模式，打開很多更刺激的可能性，但超出這本書的範圍。在GIS環境中的動態模式被解釋在相關的書籍當中。(Burrough and van Deursen 1998)

地區調查包括 (a) 集水的簡單地水文模式，(b) 沖蝕災害的模式，並外加遙測的輔助，(c) 木材從森林中萃取的最大效用，及 (d) 夏季洪氾及在Ukraine北部河岸地區及植物，<sup>137</sup>Cs程度增加的關係的車諾比研究調查。這例子的目的在於證明空間作業方法如何能被使用，而不在於呈現空間模式的完整發展。

### 空間分析運用在表面水文學上(Spatial analysis in surface water hydrology)

在本章從材料的呈現上，它清楚說明連續面的分析提供更多資訊於水文的模式上，特別在坡度及方位資訊的準備、排水網的誘導水文指標的計算、及集水量的設計上。附帶潛在及實際的蒸發、土地使用及土壤型態的其他資料，GIS可以提供更多輸入資料給流動且存在的表面水水文模式，如TOPMODEL (Beven *et al* 1984) 及ANSWERS (De Roo *et al.* 1992)。Beven及Moore (1994) 及Maidment (1993, 1996) 提出更多細節部分。

甚至帶有一個簡單網格的GIS及一個有限的指令語言結構，它可能帶有水文表面過程的簡單分析，了解位向關係上網格間的關聯，不只允許物質從格子到關聯格子的傳送可被估算，也轉換成為被格子的某些其他特性所調整。例如：為每一個格子通過網路從一個質量平衡去積聚過量的水，允許去模式化水的流動，出奇的從格子到格子，像一個流動過程 (Box8.2)。

Box 8.2：計算整體平均 (原文p.206)

相似編碼用魚計算整個網路上水分的整體平均，對每一個網格而言：

逕流 = (降雨量 - 截流 - 蒸發 - 蓄存) + 連接值。

當水的總和是從上游網格所接受而來，再加上點，水平均值超過0，則每個網格的逕流產生，逕流會被加到接下來下游的網格中。

圖8.19a、b顯示過量的降雨在圖8.11的集水量，從模擬的降雨結果高位或低位將如何加速反應排水系統下方到集水出口。當結合資料於土地使用、滲透、及蒸發，一個系統對調查說明在不同土地使用下，降雨的最後量是有用的。

如果污染物會排放到河流，一個相似的方法給予污染物從污染源點擴散的簡單評論，程序以流程圖及例子顯示在圖8.20。

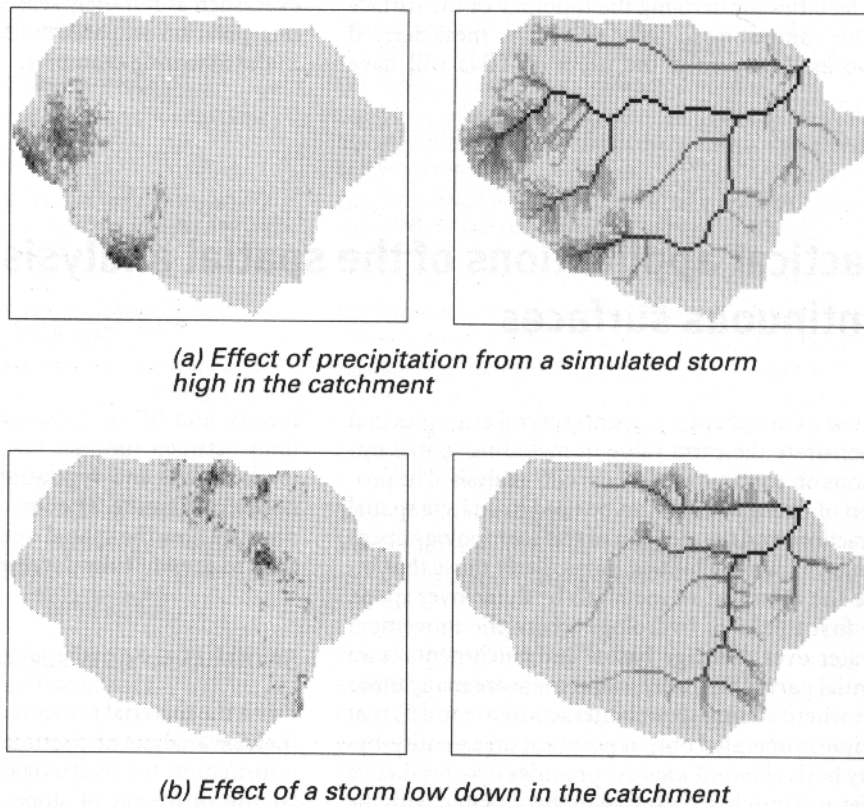


圖8.19：暴雨在集水排放上影響視它區位所具有之集水量而定 (a) 降雨量的影響，從一個模擬的高暴雨所得的集水量 (b) 暴雨的影響，低下處所得集水量 (原文p.206)

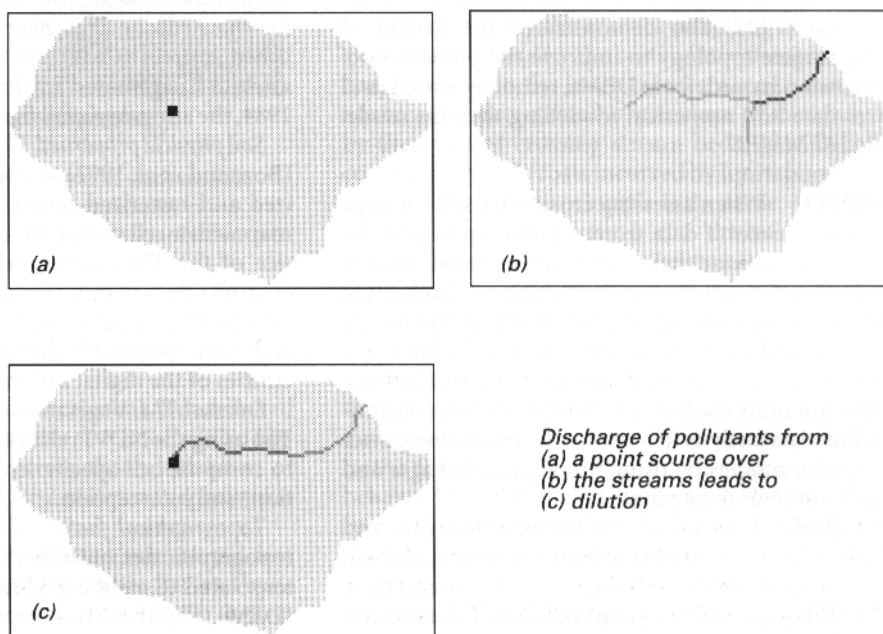


圖8.20：運用排水系統網路來估計污染物散佈。污染物排放從 (a) 一個點的來源 (b) 溪流引導 (c) 稀釋、沖淡 (原文p.207)

### 模式化土壤沖蝕災害 (Modelling soil erosion hazards)

直到最近，土壤沖蝕以被很多基地經營者，像一個靜態過程來討論在每個基地已被個別地評估的觀念下，一些注意已被用於沉澱物在陸上的輸送。就像說明在先前的章節的土壤沖蝕模式，例如USLE (Wischmeier and Smith 1978)，SLEMSA (Stocking 1981)，或the

Morgan, Morgan, and Finney Method (Morgan *et al.* 1984) 是基地特殊的點模式，他們的估計計算忽略了土壤沖蝕的輸送及沉澱。加入沉澱物的表面模式已被發展，但專家可能須要去執行它們，尤其如果它們被寫成自我控制的基本單位，在FORTRAN或相似的計畫語言中。而且，使用基地特殊模式，如USLE（見第七章），去計算土壤沖蝕越過被模式化為多組的相依實體（多邊形或像元）地景，忽略多邊形或格子所有空間的交互作用，無疑地，這個方法在物理學上是不實際的。這個運用在土壤沖蝕aspatial方法，已被很多使用Wischmeier計畫於土地計畫，研究土地的沖蝕被四周地景的磚或瓦牆以人為方式切斷的田野實驗所強化。

說沖蝕是一個過程，不如說是一個描述，較當作一個捲曲實體的屬性，更能適當地描述使用一個連續變數的資料模式。導入沉澱物輸送到沖蝕模式的最簡單方法就是，去視受沖蝕的土壤為一個類似於水可以流動過坡向網路的物質。每個網格其可能的沖蝕（potential erosion）可被計算使用點模式。接著每個網格的輸送容量（transport capacity）決定有多少可能受沖蝕的土壤可被搬移到附近的網格。如果網格有減緩表面流動的屬性，則小部份物質將被輸送。如果網格的土地已裸露，不能提供多的阻力於旁邊的流動，則有多數或全部可能受沖蝕的土壤將被輸送。當每個網格是地形上的連結到上游臨近地區，它也將接收到沉澱物，如果接收沉澱物的總量是較排出的為大，則沉澱就會產生。如果沒有，則有網的沖蝕。網路的輸送容量視地形的幾何外貌而定，例如：坡度、坡度的距離、及粗糙，也視水落於其上的潛在和動能能量而定。

加入輸送成分到USLE是非常容易的，只要坡向網路被確定，模式立刻說明在山谷窪地的沉澱現象，點模式可能不曾處理。藉強調陡峭坡地、沙州、未受保護的土壤及侵略的地表逕流的結合，修改的模式在指出沖蝕可能所在大區域的位置較好（比較Desmet and Govers 1996；Mitasova *et al.* 1996）。De Jong（1994）也證明了使用表面位向關係去改善土壤沖蝕災害的益處，從DEMs、遙測土地所包含的資料及數字的模式，這個被稱為SEMMED的模式來做說明。

SEMMED當假設大區域的沖蝕災害是有用的，是否遙測所得之資料至少可被用做一個主要輸入資料到模式中，因為在空間和時間中，它的連續變化和決定較測量在少數連結重要土壤單元（多邊形）的點座標，能更適當的反應出自然的空間和時間關係的變化。SEMMED使用連續方程式分析去改善敘述使用位向關係連結及遙測資料的沖蝕模式的能力。

SEMMED是以the Morgan, Morgan, and Finney Method為基礎去說明每年的土壤流失（Morgan *et al.* 1984），修正提出地表逕流的報告。SEMMED畫分了土壤沖蝕在水方面及沉澱物方面的點變遷。模式考慮土壤沖蝕起因於土壤粒子因雨點影響而分開，及因地表逕流這些粒子輸送，但不考慮波、濺的輸送或最後的分開。

水方面運用每年降雨量去決定降雨對潑濺分離和最後量的能量。地表逕流量被計算使用“平均雨量/降雨天數”和“土壤溫度儲存量”。水方面不包括水在網格間的交換。

模式的沉澱物方面，考慮土壤粒子由雨水潑濺和土壤粒子在最後的輸送分開，分離被模式化為一個降雨能量的方程式，運用Meyer（1981）實驗上研究的結果，修正因植被遮攔降雨的報告（Lafren and colvin 1981）。在沉澱物方面，輸送量是被計算為一個地表逕流量、坡度險峻及農作物覆蓋影響的方程式。USLE c-factor value（Wischmeier and Smith 1978）被用在農作物覆蓋影響的報告上。

## 資料來源(Data sources)

**氣候學資料 (climatological data)**：模式沖蝕的分配方法須要分配降雨量資料，但這些並不是總是可以獲得的，雖然降雨量的雷達製圖正大量改進當中，Ardeche省綜合的氣候學資料從公開來源取得 (CMD 1984)，當使用在這個研究中，眾多短暫時間的衛星影像在1984年獲得，那年的降雨量圖被使用。

**土壤物理性質 (Soil physical properties)**：法國領域，比例尺1/100000的土壤圖 (Bornand *et al.* 1977) 被數位化和網格化為一張與TM-image像元大小相符的，格子30×30m大小的空間連續網格地圖，土壤性質需要去計算，由文獻資料補充的面資料所獲得且被認為數化土壤圖，其土壤多邊形實體屬性的土壤溫度儲存量。

**大地衛星系統TM (Landsat TM)**：由大地衛星系統TM運用NDVI關係 (Tucker 1979) 所得到的植物資料被用來計算格子一個別的農作物覆蓋、蒸散及遮攔。

**地形因子 (Topographical factor)**：s=1/25000地形圖 (IGN 1985) 的等值線被數化且插入去提供格狀數化高層模式 (DEM) (圖8.21a)。坡度險峻由這個使用視窗程序—方程式8.6的DEM計算得到。可能的排水線地形網路被計算使用8-點流出演算法 (the 8-point pour algorithm)。

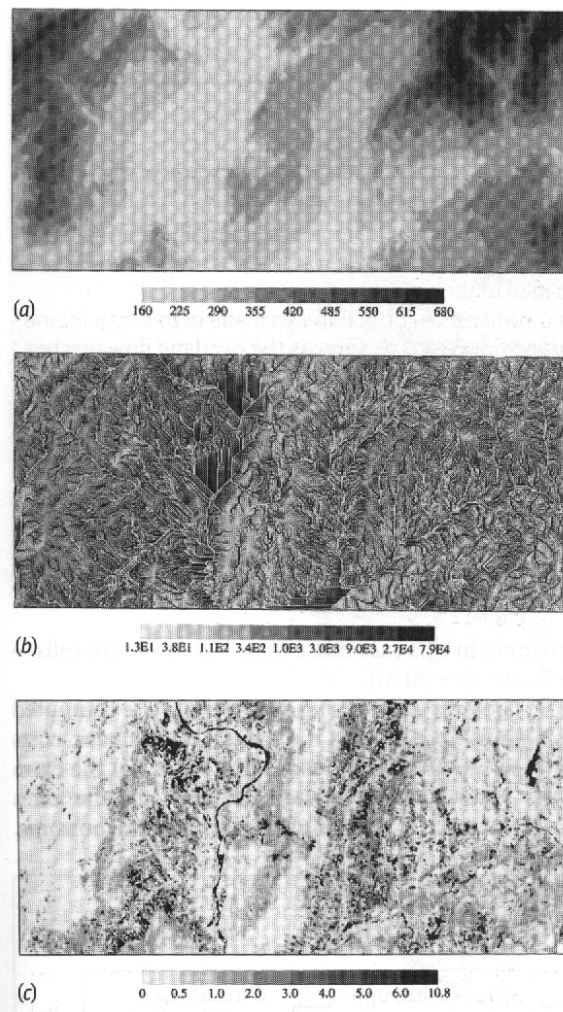


圖8.21：(a) 研究區數化的海拔高模式用於SEMMED (解析度30M) (b) 運送能力可計算使用排水系統網路 (c) 沖蝕災害 (預測土壤流失 $\text{kg m}^{-2} \text{y}^{-1}$ ) (原文p.209)

DEM也被用來替SEMMED計算分開地圖的輸送量 (圖8.21b)，一旦上層土壤濕度/網格

是過量的，這個地圖控制了地表逕流的計算。

**SEMMED 成果：**土壤流失的敘述比例範圍從0-10.16kg/m<sup>2</sup>（圖8.21c）。使用較少插入值的地區顯示最大值，例如：石灰泥地區四周Lussas的村莊，和在石灰石台地的邊緣。在Coiron較高部分的Massif Central地區低價值被發現，而且石灰石台地由於植被使得高降雨量遮攔價值。SEMMED斷言沖蝕主要發生於台地（Coiron和石灰石台地）的四周，符合田野觀測值及接近顯示在研究區域的地形地圖上的溪谷和小溪。超過90%的年土壤流失描述值是低於1.5 kg/m<sup>2</sup>。

SEMMED說明了製造沖蝕災害的區域地圖是可能的。沖蝕災害的區域地圖是較簡單外推法的計畫實驗為佳，自從模式被標定刻度，可被用來探討測量的有效值，設計去減少在景觀上最危急地區的沖蝕災害。

### 從自然林中最佳的木材萃取(Optimal extraction of timber form a natural forest)

這個例子說明了如何從自然林中萃取木材是最適宜的，為通過森林的成本、有價值樹木的位置及它們的市場價格的公式，所需之基本資料包括必要的資訊去判定基地運輸的成本、有價值樹木的位置及它們目前的市場價格。

在這個例子當中，我們假設運輸的成本是被沿單一通道橫越的成本及橫越一個沒有道路的土地的成本所判定的。在之後的案例中，我們假設運輸的成本是由土地的坡度、及穿越溪床的成本所判定的。如果如我們所想，我們也可以包括由限制土壤運輸的成本負擔，但原則上相同。

分析分為兩部分，第一為判定每個基地的可行性，接著判定要被萃取的有價值樹木確定的市場情況。

想法之後的分析是所有森林的點皆可到達，而不在於價格。這價值在每一個位置都是不同的，但從進入森林中後，可被模式化為一個連續、累積的方程式。這個連續累積的方程式運轉如DEM，但它的Z值是累積成本，運輸表面的最短通道，是從任何網格到森林入口最佳的運輸路線。

圖8.22是最佳使用的流程圖，圖8.23顯示最基本的地圖資料、數化高層模式、運輸道路的位置及東南方進入點、以極有價值樹的位置。樹的分佈可由田野測量所判定或由航空照片說明，或遙測影像的分類來判定。

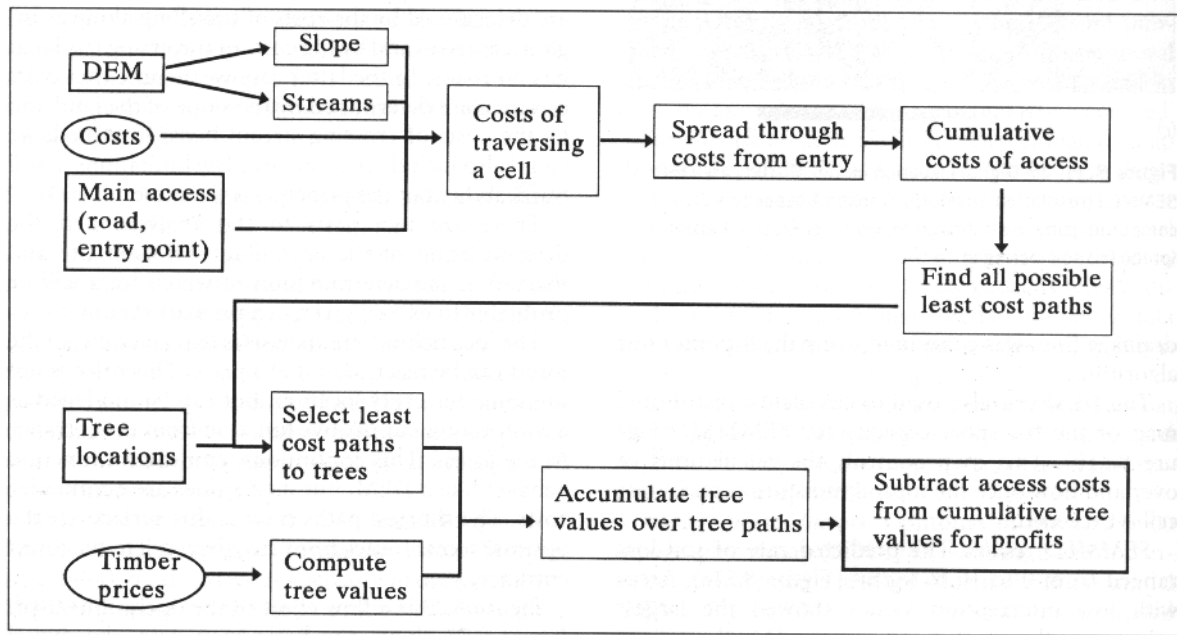


圖8.22：從森林中最佳的木材萃取流程圖（原文p.210）

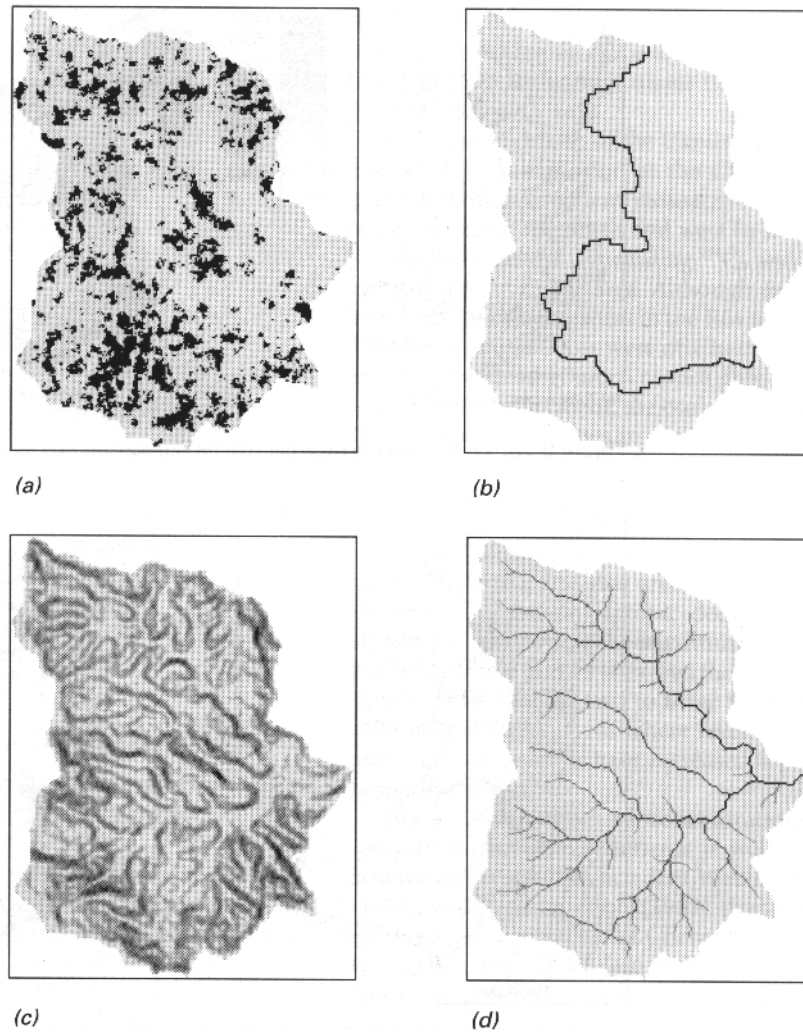


圖8.23：木材萃取的成本其成分(a)樹的位置(b)運輸之道路(c)坡度(d)河流（原文p.211）

### 分析步驟

1. 一般的坡度和排水分析所製的坡度圖和主要溪流的地圖（在這個例子中，超過100個

上游格子—格子大小為30×30M)

2. 橫越任一個格子的成本被計算為一個點運作，即（I）清除通道或沿道路駕駛的基本成本，（II）因陡峭地形所額外負擔之成本，及（III）因必需越過溪流窪地所額外負擔之成本，（I）（II）（III）的總和。圖8.24a顯示每各地區的總成本/網格。
3. 運輸的累積成本由開始於進入森林的擴張運作所計算，使用網格橫越的成本。運輸產生連續面最佳通道的全部組合接著導出如溪流網一般的相同方式，由位向關係的DEM所導出。圖8.24b顯示沿最佳通道運輸而到達所有網格的累積成本。
4. 到達樹的最佳通道次組合是由堆積的樹流“tree flow”從樹的所在位置，越過運輸的累積成本所建立的。
5. 木材價值運輸這些最佳通道的累積增加量，可由估計目前木材的市場價格、增加樹木的位置及目前的價格和擴展上游的次佳通道的價值來計算出來。原因的表面（圖8.25）給了木材價值的累積增加量一個從入口點算起，最適成本的方程式，減去累積樹的價值地圖的通道成本，使地圖顯示出可以有利萃取的地區（圖8.26）。分析可被重複用於不同的市場價格上，去瞭解木材萃取的有利性，如何在空間上和全部的金錢價值上做變動。

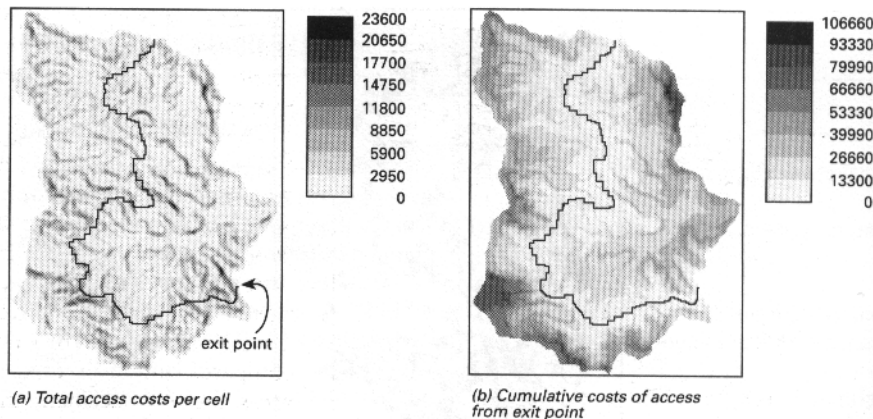


圖8.24：（a）所有通道成本/格子（b）累積成本從出口到達每一個格子（原文p.212）

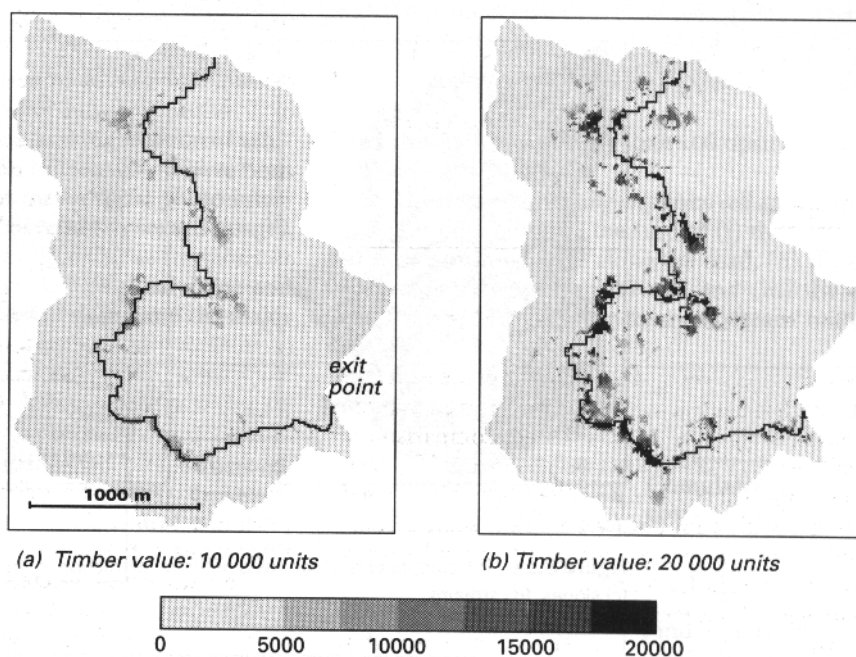


圖8.25：木材所處價值>運輸成本（a）木材價值：10000個單位

(b) 木材價值20000個單位 (原文p.212)

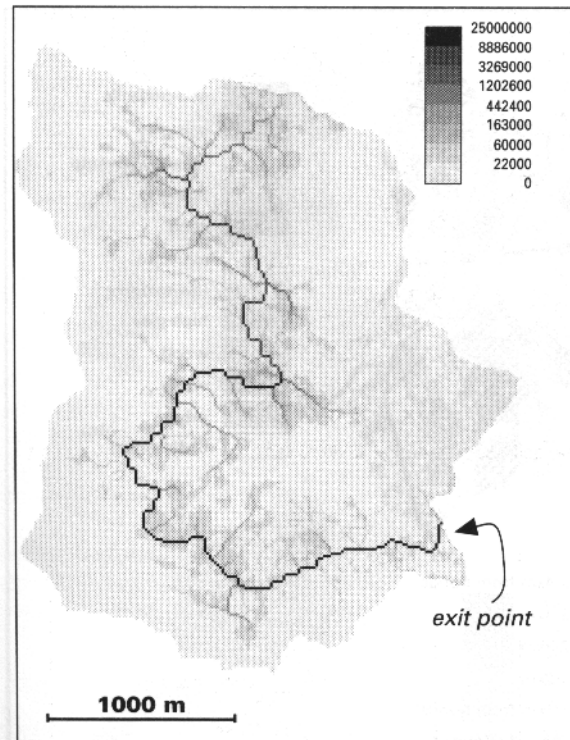


圖8.26：當木材價格為20000個單位時，累積利潤及運輸路線達有利的狀況 (原文p.213)

重點在於最適成本的通道如何能被模式化，對其它成本因素作考慮，也在於模式化這個過程不只去顯現出樹的位置，更顯現出由於不同樹種和不同成熟期而產生區域中木材價值的變化。分析中也可與樹木的生長模式相關聯，去顯示森林的開發如何能被引導到一個持續的方式上。

### 洪氾時車諾比的銫<sup>137</sup>的重新分配(The redistribution of Chernobyl <sup>137</sup>Cs By flooding)

這個例子顯現出一個有趣的問題：在相同屬性的資料，在這個案例中，在土壤裡的輻射銫程度，收集了數組在1988年和1993年單點的位置 (Burrough *et al.* 1996)。很顯然地，在1993年的程度很多在1988年的測量值來的高，這違反了銫的半衰期30.2年的這個預期，直接的比較是困難的，因為不同位置的每年使用，所以簡單地價值的減法不能被執行。對於相同來源的網格年與年間的內差法是可行的。但內差法的面間簡單的不同點是不必指出在<sup>137</sup>Cs程度的真實改變。因為在年間的不同點可因取樣一個高的變數型態，或物理過程而引起，例如：氾濫。藉著使用GIS結合地理統計的模擬，那假設在1993年測試了<sup>137</sup>Cs的增加量，那是由於有一個主要的洪氾發生在1993年7月。

**問題(The problem)**。Ukraine大部分因1986年4月車諾比核子爆炸而受影響，植物焊接下來的放射性雲柱，被盛行風帶到北部和西部。放射性同為素：<sup>134</sup>Cs和<sup>137</sup>Cs被沈澱在廣大的地區。雖然型態上輻射銫在車諾比大約350km西方Ukraine的Rovon地區，其沈澱程度和其它大部分鄰近意外事件更高污染的地區為低。在區域內顯著的土壤型態的生態情況，承認比較而言高的植物輻射銫的了解，導向持續長期各種實務生產，尤其牛奶、肉和葷污染。



在1988、1993和1994年的土壤取樣證明在某些地區<sup>137</sup>Cs程度已隨同位素的蛻變下降，但有重要的地區在1993和1994年輻射銫的最大程度超過在1988年2-3次的測量最大程度。程度下降的地區和程度增加的地區間主要的不同在於緊鄰溪和運河。

在河川的沈澱物輻射銫的重新分配和集中是充分的被瞭解的。在1986年意外事件後不久，Walling和Bradley（1988）和Walling *et al.*（1989）證明車諾比radioisotopes在Wales的River Severn水道和洪氾平原的沈澱物增加，所以相似的過程可能起因於在Ukraine的Rovon地區當地<sup>137</sup>Cs程度的增加。

資料從一個典型集體的農場地區，在Rovon地區內（the Chapayev and Kolos集體的農場）蒐集而來，Rovon地區已在1988年被Ukraine當局確認為一個有強大空間和短暫改變輻射銫大背景程度的基地。地形是非常平的，而且主要的排水是藉蜿蜒的河道，河道排水接近6500km<sup>2</sup>朝北向的一個地區。直到Ukraine-Belarus北部邊緣流向東River Pripyat，在車諾比連接了Dnieper河，所有的Belarus-Ukraine邊緣地區藉由soddy gley和泥炭土（peaty soils）的廣大地區顯示特性。那些地區被繪製了目前和過去1：3000000和1：500000的地圖。實際的研究基地測量約10.5×12.5km，和包括Rovon當地的Dubrovitsa Pistrict的76.5km<sup>2</sup>。

取樣座標的資料庫，自然和人為的排水渠道，受每年春季洪水影響的地區、田野數目、土壤系，和土地使用圖由數位化1/10000比例尺的紙製地圖所產生（圖8.27a、b）。一個為了定量的空間分析和內插所做的50\*50M解析的格子狀資料庫，由基本地圖資料所產生，受污染的土壤取樣在1988、1993和1994年，對72、87和47個基地個別地蒐集，且分析了<sup>134</sup>Cs和<sup>137</sup>Cs。每個曲樣的位置皆被數位化（圖8.28a、b）。

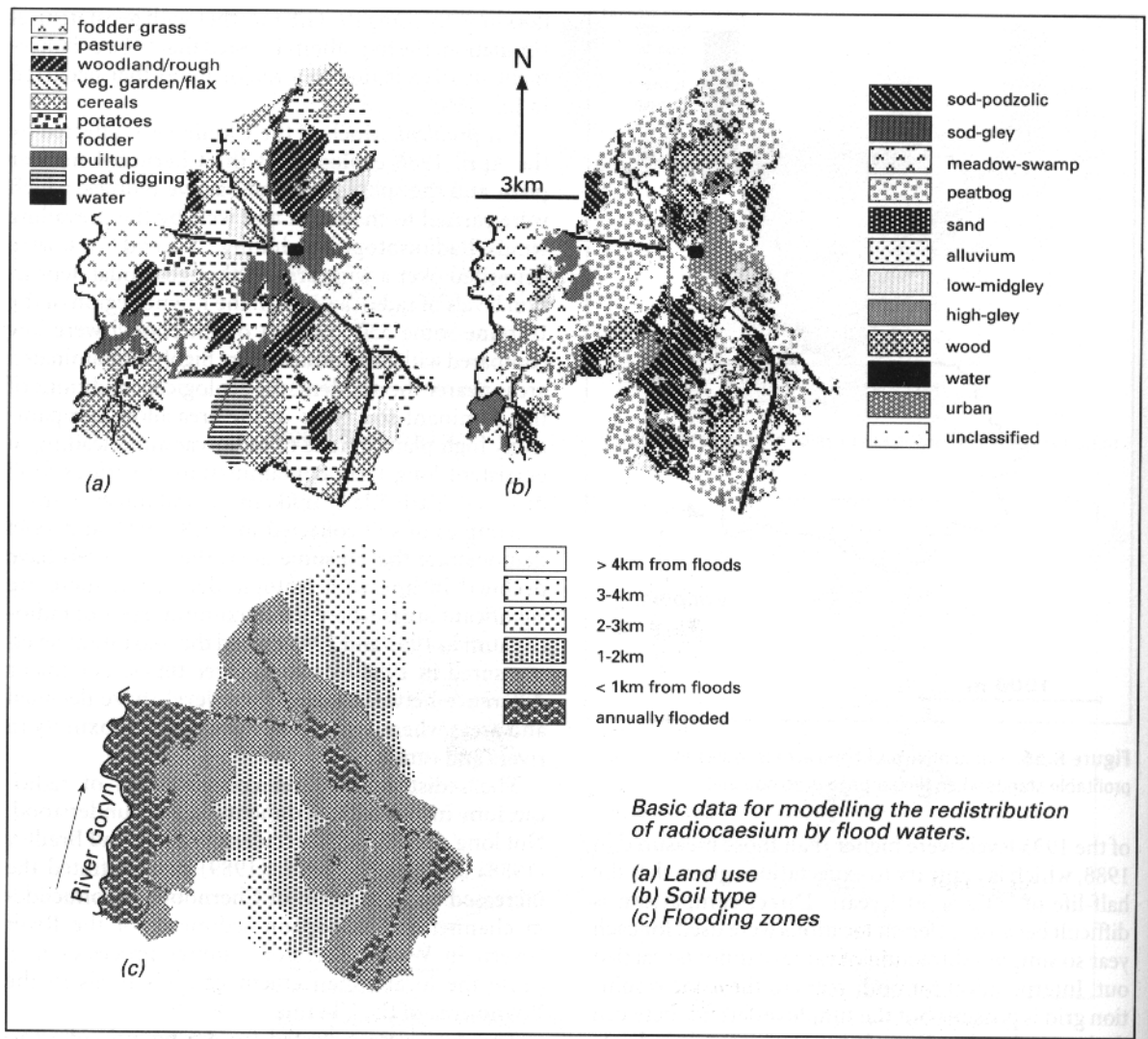


圖8.27：輸入地圖把<sup>137</sup>Cs洪水時的重新分配模式化（模式化輻射性銫金屬於洪水時的重新分配（a）土地使用（b）土壤型式（c）洪水區）（原文p.214）

<sup>134</sup>Cs和蛻變的校正。研究<sup>137</sup>Cs的程度，其短暫的改變可能已藉著水文學研究的原始資料所產生，首先校正兩個使用1986年的同位素輻射<sup>137</sup>Cs/<sup>134</sup>Cs=2:1的蛻變比率。所有的分析引用了輻射銫在1986年被轉換成的相等值的資料（表8.2），資料被強硬的實際的偏斜，且被換算為自然對數而標準化。所有空間分析使用轉換資料，但結果已被重新轉換來展現和輕易的解釋。

表8.2：統計上的相對於1986年同等的改變資料（KBqm-2）（原文p.213）

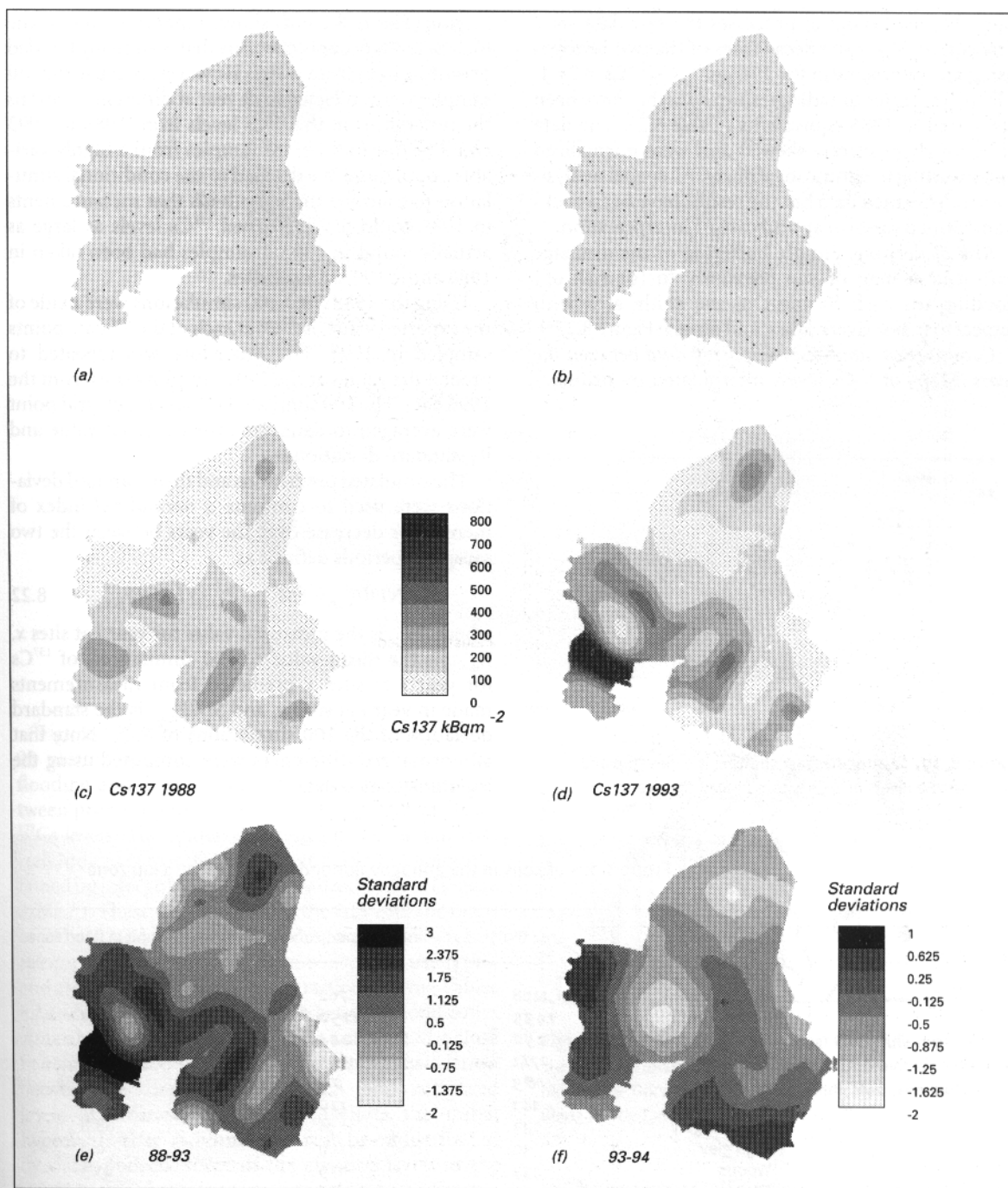


圖8.28：(a)(b)資料點在1988年及1993年的位置(c)(d)1988年和1993年 $^{137}\text{Cs}$ 程度的內插法  
(e)(f)1988-93年及1993-94年個別地標準化的不同計算（原文p.215）

**洪水氾濫的風險。**排水渠道圖的簡單緩衝給一張接近洪水（或有風險）的基地地圖，為在研究地區每一個位置，附帶關注的主要河流、湖和運河（圖8.27c）。

**地理統計學分析：年之間的比較資料。** $^{137}\text{Cs}$ 程度由次序性克利金法內插所得的地圖（圖8.28c、d）顯示了每年最高的成度出現在每年洪氾或從水1km內的區域。所以，因為土壤取樣蒐集每年在不同位置， $^{137}\text{Cs}$ 的程度從1988到1993年不同，可能是由於從一個高變化群體的不同樣本。我們決定使用有條件的沈澱物去調查假設，如果1993年的取樣樣本即為1988年索取樣的基地，在1988年測量已經可以轉換的 $^{137}\text{Cs}$ 程度和實際在1993年發現的一樣大。

使用1988年的資料，100份模擬為被製成 $^{137}\text{Cs}$ 的期待值在87個1993年被取樣的資料點的

每一個點。過程被重新去從1993年資料來預測在1994年取樣基地的位置。在每個網格點上100個模擬值，被求平均數去測定期待值和它的標準差。

模擬的敘述和它們的標準差被用來計算兩個取樣時期之間的增加或減少的標準化指標，定義為：

$$NDIF_{ij \cdot x} = (M_{i \cdot x} - P_{ij \cdot x \cdot y}) / SDIF_{ij \cdot x \cdot y} \quad 8.22$$

當 $M_{i \cdot x}$ 是於年*i*在基地*x*的測量值， $P_{ij \cdot x \cdot y}$ 是<sup>137</sup>Cs的100種模擬的平均值，對年*i*在從年*j*在基地*y*的測量所預測的基地*x*，而 $SDIF_{ij \cdot x \cdot y}$ 則是100個模擬的 $P_{ij \cdot x \cdot y}$ 其標準差。要注意的是所有標準化的不同被計算使用對數轉換的資料。

圖8.29顯示1988-89年標準化的指標計畫違反了距離從水或每年的洪水，且證明了接近於洪水和大的（和多變的）<sup>137</sup>Cs程度兩者之間緊密結合。1988-93和1993-4兩者標準化指標的空間分佈由內插標準化指標而得到（圖8.28e、f）。這些在GIS終將土壤和土地使用圖疊圖起來，這些圖證明了洪氾平原土壤、土地使用型態、和輻射銫高的地區間有緊密的關聯（表8.3，8.4）。雖然洪氾區域和高標準化程度間關聯不是那麼精確。洪氾、洪氾的可能、和<sup>137</sup>Cs程度的增加間強烈的關聯性這個假設，更清楚了支持最初輻射銫可藉水來重新分佈的假設：也的確，某些西南高程度地區可能是由於沈澱物是由上游運送而來（比較Walling and Bradley 1988, Walling *et al.* 1989）。

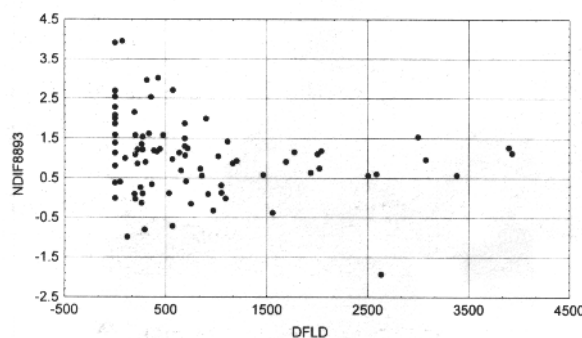


圖8.29：散佈圖顯示了<sup>137</sup>Cs在河流1km內有高且更多變的程度（原文p.216）

表8.3：在每年洪氾區和1km區域土壤的比例（原文p.216）

表8.4：在每年洪氾區和1km區域之土地使用分類比例（原文p.217）

進入險峭的型態指示洪氾地區較整個區域有較高比例。

結合了GIS和地理統計分析，清楚的顯示洪氾頻率、土壤型態、土地使用、和土壤樣本的資料，可被結合用來估計在1988年後由於洪氾和附著的污染物其<sup>137</sup>Cs增加程度的測量假設。在1993年7月前後主要食物在牛奶方面<sup>137</sup>Cs程度的增加，其更進一步的資料證實了食物在食物鏈中有直接地影響，特別在泥炭土這些未經改善的牧場。在研究接近完成後，我們發現一組在網路上的NOAA影像，證實了1993年7月的洪水，和提供更多有關<sup>137</sup>Cs藉Ukraine洪水變動的支持（圖8.30）。

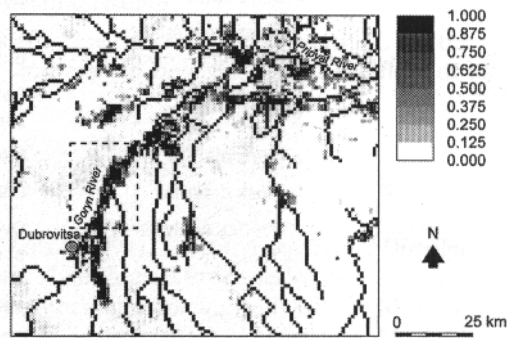


圖8.30：1993年7月Pripyat河的氾濫。研究區域是位於Goryn和東岸的Dubrovitsa東北方。灰色範圍指示出將發生氾濫的潛在性。（原文p.217）

### Box 8.3：使用SEMMED的程序（原文p.210）

SEMMED的步驟為：

1. 計算的降雨量/降雨天數 = 年降雨總量/降雨總天數
2. 計算每一個格子中，土壤水分儲存容量。
3. 由結合每年的降雨量去計算（地面的流量/網格），即雨量/天數\*土壤水分儲存容量。
4. 導出當地的區域排水方向圖。
5. 累積地表流量超過坡向網，修正於河流水道的地圖上：地表水流一到達河道，則地表水流的沖蝕就會停止。
6. 結合地表水流的網格圖（步驟3）和潛在地表水流圖（步驟5），修正被用於滲透入（飽和的）頂端土壤的水量，扣除帶有飽和滲透力的格子地圖（滲透圖是由重新分類數化的土壤圖所產生的）
7. 結合步驟6的結果和坡度險峻及植被因子圖產生散佈的輸送能力圖。

## 結論(Conclusions)

本章節廣泛地顯示了在推論來自於如連續地表屬性模式的新資料。在表8.5下了一個結論就是GIS在處理確切實體與連續區域的機能。通常連續表面並不總是描述地形。這個理論應用在許多區域，不僅僅是水文學還有沖蝕、森林經營以及土壤、水污染。儘管所有的例子關係著地景的自然問題，這些分析也可以應用在研究一個或更多連續表面的屬性，如潛在市場的表面、暴露的弊病或良好的經濟。

表8.5：GIS分析實體與連續區域的機能（原文p.218）

幾何學Geometric	地理座標系統的轉換
相互作用Interactive	地理編輯與屬性資料更新與編輯的相互作用
儲存Sorting	屬性或地理資料的分類
位置Location	位置實體定義的屬性組
總結Summarize	總結每一個地理實體的屬性（點/線/面/網絡）
統計Compute statistics	點/線/面/網絡的統計（定義、區域、附件）
鄰近Proximity	處理鄰近地區與搜尋-創造緩衝區及導出廊道分析
內插Interpolate	來自於點資料的內插的規則的網格或同線性
圖表Block diagram	計算三次方資料的圖表
疊合分析Overlay analysis	使用布林運算算術機能/過濾，不論是向量式或網格式中的疊合與結合好幾個地圖去處理地理屬性資料
多邊形網格Polygon to raster	轉換多邊形網格的輪廓
邊緣探測Edge detection	半自動化的影像邊緣探測
網絡分析Network analysis	交通量密度的路徑變數網絡中發現最短路徑。連接具體實體至其他網絡
數化地帶分析 Digital terrain analysis	利用3D表面呈現地形。計算坡度、坡向、陰影凹凸起伏圖、氾濫、決定集水區
模式Models	模糊模式的分界面

## 第九章：資料誤差和品質控制

### *Chapter Nine : Errors And Quality Control*

#### 前言

本章調查誤差如何發生在空間中的資料，和它們在資料分析與模式中可能有的影響。誤差包括粗心的錯誤（人為誤差）和不講理的對話（gaffs）（系統誤差），但它們也是選擇的資料模式和計算模式原有的一部份。統計的不確定和空間變化是在空間資料中任何誤差的評論觀點。描述估計誤差的方法在實體範圍中，分成三種情況：向量—網格轉換、數位化和多邊形疊圖。

#### 空間資料、成本和 GIS 輸出的品質 (Spatial, costs, and the quality of GIS output)

GIS 產品的品質通常被最後產生在電腦螢幕、圖表、或影像的圖案的視覺外觀來判定。電腦製圖者正設計更令人心動的技巧為傳播視覺給人們，由視覺外觀所控制的品質是不適當的，然而，如果呈現的資訊是錯誤的，或是因誤差而訛誤。不確定和誤差對於空間資料是固有的，需要被適當的提出來，而不是被掃除到想像圖示的表現下，可以有一個錯誤吸引力，有關誘人的、製圖者高品質製圖的產品，和現在電腦圖示的專家，提供給 GIS 的使用者。在 1980 年 Chrisman(1984a)指出“我們已發展可能性，例如平滑輪廓線並不總是由適當的證據來支持。” Blakemore(1984)推論注意到 naive 要求的某一電腦製圖的支持者電腦—輔助製圖產品是準確的在用來製造它們的金屬器具的解答。他注意到只有一些評論作家，如 Boyle (1982), Goodchild(1978), Jenks(1981) and Poiker(1982)提出在地理資訊過程中對誤差問題的注意，但在 1996 年甚至是發展後的 25 年，仍然沒有充分的去注意到誤差如何產生，和如何被傳送。雖然空間資料品質其有系統的研究現在正被發表(Guptill and Morrison 1995)，但大部分在誤差上的研究仍然處於調查層面。(Fisher 1995, Goodchild and Gopal 1989, Heuvelink 1993, Lodwick *et al.*1990)

在最近的研究，Wellar and Wilson(1995)提出：雖然 GIS 已對空間理論化的定性、定量、and/or 意象程序上有影響，但它也在空間理論化的過程上有一些影響，也因此自然的變化和誤差有一個更佳的了解。資料獲得和投資所給予的成本和 GIS 的使用有所關連，這是意外的。在地理統計和空間的統計範疇中，無論如何，已有很多理論上的和實際上的研究在於如何去處理在可被討論為連續面的屬性和其空間變化中的疑慮 (e.g Isaaks and Srivastava 1989, Journel 1996, Deutsch and Journel 1992, Cressie 1991)。而且同時去連續發展在這些地區的構想，去提出一個聲音基礎，瞭解在空間資料及空間資料分析中不確定的角色。

根據主題的正確度 (thematic accuracy)，位置的正確度 (positional accuracy)，及時間的正確度 (temporal accuracy)，資料的正確度常集結 (Aalders 1996)，但在空間資料誤差可能發生於從觀察到呈現的過程中的各種階段裡，在認知的誤差 (不適當的確認) 可能發生於概念階段。誤差和判定地理位置的近似值是依據測量技術，準備的硬體 (GPS 衛星、雷射經緯儀等)，以及地圖投影 (projection) 和橢圓體 (spheroids) 的選擇。在屬性測量的誤差，由於在問題現象的變化，測量裝置的正確度，或觀察者偏見可能發生在原始資料的紀錄間。對於被討論如：連續面樣本密度、它們的提供 (support) (樣本的物理大小)，及取樣

的完善的現象，是不確定的所有來源。

當資料被儲存於電腦中，誤差可以緩慢進入，因為太少電腦空間被分配去儲存高精密度（high-precision）數目，需要記錄資料到所給的正確度的水準，某些資料可能是如此簡單或困難去蒐集，必須附隨著一些樣本，且信賴不正確的相互關係連帶其他、便宜於測量屬性，和如必然的不確定產生。被用來從存在的資料中去導出新屬性的邏輯，或數學的原則（模式或內插法）可能是有瑕疵的，或可能包含導致循環性誤差的計算方法。當已被評估於不同實體或在不同支持中取樣的資料被結合，則在空間轉換的不同點可能是如此的強大，以致於樣本比較不能被操作。

最後，在結果的視覺呈現上，如果符號語言不清楚，如果顏色和陰影是不適當的，或是如果展示太擁擠，或是它只是太困難去得到一個清楚的結果，則使用者可以獲得誤差的印象。

一般視覺的誤差和不確定是因為它們是糟糕的，並不是不可避免的，無論如何，因為去瞭解誤差和不確定如何產生，它們可以如何被管理及可能地被減少，以及誤差和誤差傳播的知識如何可以被用來改善我們對空間格局（pattern）和過程的瞭解，這些是非常有用的。連結了空間不確定的良好瞭解於模式化和內插法的數字上的方法，可以提供有用的工作於最佳的取樣（藉以提供金錢上的價值），及於確認空間分析中較弱和較強的部分。誤差及誤差傳播的良好瞭解朝向於積極的品質控制。

很多 GIS 使用者處理資料分析的使用技巧說明在第 7、8 章，在絕對的假設下，所有資料是全部全部沒有誤差（error free）。藉“error free”不只是意謂有缺陷的調查或輸入所產生的缺乏實際的錯誤資料，也是意謂統計上的誤差，意指 free 來自變異。換句話說，加入兩張地圖的算術工作同時有著一個樣本疊圖的意思，指出來源地圖兩者可被討論為完美、完全地判定文件，使用一定的資料品質水準在所有研究地區。這個觀點是被加入一個大範圍，以大部分資料庫及修正語言（雖然它可能是用別的方法），這是由於資料品質資訊的缺乏、具體表現精確的概念、瞭解誤差是如何被傳播的連結、以及 GIS 工具對誤差評價的缺乏。

很多野外科家和地理學家從經驗瞭解清楚劃分於地圖中的邊界和等高線，是精確的錯誤表達逐漸、模糊的變化（Burrough and Frank 1996）。人們被限定去瞭解地球表面的變化，而那些變化是由地誌圖（清晰的邊界）的逐步功能或平滑地變化的數學表現（見第 2 章）來描寫，而且發現很難想像現實不是那樣。除了“結構”已由邊界或等高線來模式化，常有一個殘留的非製圖的變化，產生超過距離小於這些可分離的原始論述，此外，自然現象的空間變化不只是一個當地雜訊機能或收集更多資料、增加測量的精密度來移動的誤差，而常是一基礎觀點的本質，產生在所有的尺寸，如碎形提議者已指出的。（見 Mandelbrot 1982, Burrough 1983 a,b, 1984, 1985, 1993a, Goodchild 1980, Lam and De Cola 1993）

瞭解空間誤差的本質，及誤差可能對使用 GIS 分析品質產生之影響，這是非常重要的。

本章的第一部份探討在空間資料中的誤差來源，以及影響品質的因子，關於空間現象以實體為基礎和連續面為基礎的模式。第二部分解釋影響空間資料品質的因子，而第三部分包括了結合轉換實體和面基礎資料的誤差的發展和瞭解，藉由線的數位化和經由多邊形疊圖中從一個表現到另一個（向量-網格）。第 10 章則說明運用統計的方法去瞭解誤差傳播在數字的模式於第 6、7 章所敘述的空間分析的種類的關連，及顯示不確定的徹底瞭解如何能用於最適取樣和空間分析。



## 空間資料的誤差來源(Sources of errors in spatial data)

Box9.1 顯示控制因地理資訊處理所產生的誤差的主要因子，“誤差”被用在這裡，有廣泛的意思，不只包括“錯誤”也包括統計概念的誤差，即“變化”。“誤差”包括觀察和容易考證的錯誤，但有更多難以捉摸的誤差來源可能常常當操作資料時才被察覺。大部分困難的“誤差”來源為那些可提升為執行確定種類的處理結果，他們的發現不只是對資料，更對資料模式要求要有一個深入的知識，因此，他們有可能去阻撓大部分的使用者，很多這些“誤差”的外觀或更適當的“資料品質”是正經由國際協會提出。(cf. Aalders 1996)

## 內容的正確度(ACCURACY OF CONTENT)

內容的正確度不論是附上點、線、面的地理資料庫的屬性皆是問題，我們可以區分提到一般變數或符號是否正確的品質正確度（例如：土地使用圖中的一個地區可能被錯誤的編碼為“馬鈴薯”的“小麥”），以及提到在估計值分配的偏見水準（例如：一個差的 PH 計量器測量尺度可能從頭到尾估計的 PH 值都高 1 單位），確保正確度是有關於有可靠的、有來源證明的輸入和轉換程序的問題。

## 測量誤差(MEASUREMENT ERRORS)

差的資料可能是來自於不可靠、不精確、或有偏見的觀測者或工具，讀資料者應該清楚地瞭解正確度和精密度之間的區別，正確度是估計值接近正確值的程度，且通常由標準誤差來估計，在統計數語上，精密度是一個觀察的分佈（通常以標準偏差的觀點來測量）測量，大約是平均，精密度也提到電腦呈現數字於十進位數化的正確數目的能力。

### Box 9.1 影響空間資料品質的因子（原文 p.223）

1. **現階段**：資料是否為最新的？  
時間序列
2. **完善**：區域的適用範圍—是部分或全部？
3. **一致**：地圖比例尺  
標準的描述？  
適當
4. **易取得性**：格式  
著作權  
價格
5. **正確度及精密度**：觀測密度  
位置的正確度  
屬性的正確度—定性且定量的  
位向關係的正確度  
系統—何時聚集、藉由何者、如何進行？
6. **資料的誤差來源**：資料進入或輸出錯誤  
原始資料模式的選擇  
邊界位置和位向關係的自然變化和不確定  
觀察者偏見

處理：電腦中數字上的誤差、電腦呈現數字的限制

### 7. 導出資料和模式、分析結果的誤差來源：

與地圖疊圖有關的問題

分類和歸納的問題

分析模式的選擇

邏輯的不當使用

錯誤的傳輸

使用內插方法

## 野外資料(FIELD DATA)

測量員是在被輸入到很多 GIS 的資料的品質它的一個評論的因素。經良好設計的資料蒐集模式信賴直觀，例如：土壤或地理測量，是最重要的，直觀就是指解釋說明是在野外產生或從地景或次地表的佈局變化的航空照片或地震儀來得到。使用者應該瞭解某些觀察者他在本質上就較其它的觀察者來的更入微或努力—“土壤測量員的素質改變從不可靠的航空照片判讀者兩個細微的工作，到測量者的取樣計畫暗示他正栽植洋蔥。”（Smyth, quoted in Burrough 1969）。在地圖的外觀，非常大的不同可能起因於測量者的不同或製圖模式的不同，例如由 Bie 和 Beckett(1973)及最近的 Legros *et al.* (1996)對土壤測量的研究，及 Salome *et al.*(1982)對地形學已清楚地證明。

在大型機構裡，應該可以去對每個測量員的素質作判定及紀錄，額外的屬性可運用資料本身來儲存。如此的程序可能因評量尺度而被反抗，例如：在專業見解的含糊，但改善觀察者素質的最佳模式是去改善所有資料結合過程的外觀，例如：標準觀察技巧和資料記錄形式，以及藉由發展一個聯合約定，在測量管理和評量工作最可行的標準之間。

增加使用在很多野外學科的自動化取樣計畫連結電子資料伐木者，指的是如果全部都可能被操作，則資料的正確度和精密度就是好的。從電子取樣設計蒐集者的資料可被全自動地在螢幕呈現極端值指出不良的函數。在區域新的取樣設計，例如：地理工程（geoengineering）和污染科學，意指觀察可被製作在材料的 *situ*，否則必須在實驗室裡被分析（Rengers 1994）。

## 實驗室誤差(LABORATORY ERRORS)

直覺地，預期實驗室測定的品質超過那些在野外的測定品質，雖然在單一實驗室執行同樣程序的測定，結果可能是可重複的，相同的不可能被說為執行在不同實驗室的分析。主要遍及世界的實驗室，由在 Wagenungen 的 International Soil Reference and Information Centre 交換執行的計畫（van Reeuwijk 1982,1984），顯示相同土壤取樣的實驗室結果的變化，可能大於黏土量的 $\pm 11\%$ ，陽離子交換的 $\pm 20\%$ （黏土碎片的 $\pm 25\%$ ），基礎飽和的 $\pm 10\%$ 及 PH 的 $\pm 0.2$  單位的誤差，數字模式的結果所牽涉的是龐大的！實驗室分析應該被改善在可重複的由於實驗室設備更廣泛的使用上，但實驗室科技無總額將組成為簡陋的收集或簡陋的預備樣本。

## 位置的正確度(LOCATIONAL ACCURACY)

地理資料位置正確度的重要性強烈依賴考慮下的資料型態，地形資料通常會對位置正確

度做一個非常高程度的測量，這個正確度是要對物件良好適當的定義，例如：道路、房屋、土地包圍的邊界，和其它記錄的特徵，運用現代電子的測量技巧和 GPS 物件在土地表面的位置，現在可以被記錄到公釐的正確度（millimetre accuracy）。對照下，土壤或植被單元邊界的位置常反映了測量員對於在何處做一分割線的判斷，即使有，應該要被放置。常常，植被型分類為另一個時，有些超過瞭解的差異的考量，例如：由微氣候、土地高低和水的體制（water regimes）來源轉變測定的結果。坡度分類的改變或地下水體制的改變也是不可能總是發生在清楚地定義的邊界上。

位置的誤差可能是因為簡陋的野外研究，因原始紙製基本地圖的失真或縮小，或因網格掃瞄後低品質的向量化（Dunn et al. 1990, Bolstad et al. 1990）。局部性的誤差（local errors）常常可以由繪圖工作站的交互作用數位化來修正，當一般的位置誤差可以由各種轉換來修正，即一般所知的“橡皮伸張法（rubber-sheeting）”技巧，在第 4 章已經敘述過了。對於誤差發現結合現代硬體和軟體已大大改善了近幾年數位化的品質。

橡皮伸張法（rubber-sheeting）對修正幾何的失真成功依賴了被轉換的資料型式和轉換的複雜狀態。很多方法良好的運作於簡單線性轉換，但當複雜性減少必需被修正，則方法失敗。當原始地圖強大地組成連結、直線的線段，則方法不一定良好運作。例如：一些年前，努力嘗試在 the Netherlands Soil Survey Institute 使用橡皮伸張法（rubber-sheeting）去結合 19 世紀初地形圖數位化說明為土地使用，評定改變為目前 1：25000 的地形圖（topographical sheet）上。有問題地區的道路型態和準確的格狀結構相似，當提出橡皮伸張法（rubber-sheeting）程序，道路線不是被拉伸的，但結構被擠壓在得路的交叉點上，很多相同的方法，橋或 meccano 製作的起重機可能擠壓合而為一。

## 自然的空間變化(NATURAL SPATIAL VARIATION)

很多主題圖，特別是那些地景的自然所有物，例如土壤、植被、不考慮空間變化或起因於現象製圖中的短範圍改變“不去考慮（impurities）”的地區來源，這個問題已經是很多研究的主題，特別是在土壤測量、土壤物理學和地下水研究上（e.g. Beckett and Webster 1971, Bouma and Bbell 1983, Nielsen and Bouma 1985, Burrough 1993b）。問題和很多結合了空間地過份簡化的土壤分類和製圖的範例一樣，結合了沒有充分瞭解的土壤的自然變化。（Burrough et al. 1997）

製圖的協約迫使土壤科學家去製圖土壤為明確的界定、均質的地區。包括在邊界內逐漸改變的資訊和變更寬度的邊界不可能被呈現在傳統的彩色圖（conventional chorochromatic maps）。這些地圖已被努力的數位化，而且數位化的土壤多邊形已呈現給 GIS 的使用者像是一個空間良好定義般的地籍單元的實體。但是事實常常不是那樣—這些明確的多邊形實際上是粗糙的，但已合宜的接近，主要的問題牽涉有關這些模式和實質間的不同資訊連結。

在最初，傳統的土壤類型圖，S：1：25000-1：50000，顯示以不去考慮（impurities）的觀點單位描寫的特徵（Soil Survey Staff 1951），被認為是沒有超過 15-25%。不去考慮（Impurities）被定義為特定在土地圖例上沒有完全要求相同的觀察。很多研究（e.g. see Beckett and Webster 1971 or Burrough 1993b for a review）顯示不只 15% 荒地推測，也顯示 impurity 的概念的一些含意。藉由改變圖例，指示所有皆是一個結合觀察的定義且如此的純正，不能被隨意的操作。皆下來的工作證明了土壤的自然變化和顯示瞭解污染問題，或精密農業的最適土壤施肥的重要。（e.g. Burrough 1993b, Goode 1997）。增加資訊在土壤的變化

上，和其它自然現象上，例如：水質或可能只是對專家有用的物種組成。

瞭解看不建的空間變化的現象，如：土壤、lithology 或水質，是重要的。可以強烈促成因模式化和地圖疊圖所造成的相對和絕對誤差的結果。關於如何去估計這些誤差如何傳播通過數字模式的進一步細節，在本章下面將會介紹。

## 影響空間資料可否信賴的因素(Factors affecting the reliability of spatial data)

### 資料的年代(AGE OF DATA)

所有資料在同一時間蒐集用於計畫是很罕見的，除非計畫是一個特殊例子的研究。大部分規劃者和環境政府機構被迫去使用現有發表的地圖和報告的型式的資料，藉更接近現在的遙測影像（包括航空照片）和野外研究來填充為所需要的。Mead (1982) 發表談論“除了地質資料外，其它資料的可靠性會隨年代減少。”雖然這可能有幾分概略的事實，地質的改變是較土壤、水的體制、植被、或土地使用來的慢很多，但老舊資料是不適當的，這也是有可能的，因為它們是根據不同系統的標準去蒐集而來的，而那些標準沒有延續到現在被使用或接受。很多計畫去獲得老舊的資料是使用手寫的野外圖，和過時的術語已必需要被單純地捨棄，因為龐大的成本牽涉其中。

### 區域的適用範圍(AREAL COVERAGE)

一個研究地區的全部是合意的，即一個實驗的田或鄉鎮應該有一個不變的資訊覆蓋，如果不是如此的資源資料，加工業者必需要滿足資訊的部分水準，通過全球的數位化資料正有效的增加（例：是藉在網際網路的數位化圖），這仍然是一般的，甚至在已開發國家，除了在尺度上，對於要求的目的是太小，對於超過研究地區沒有完全包括的可靠種類的主題資訊，例如：很多國家仍然有破碎覆蓋區域比例尺 1：25000-1：50000 的土壤圖。此外，在 30-40 年間，土壤、植物和地質的主題分類的概念和定義已被改變，方法應該要被製圖或測量員自己去變動。歷史事實可能導致不足取的地圖單元，沿土地邊界，如果沒有進一步的測量，就難去解決。

如果覆蓋區域不是全部，則決定必需要是包括需要等質是如何被完成的。選擇是去蒐集更多資料，從遙測去獲得代用的資料，要注意“blow up 放大”歸納或小比例尺地圖資料去獲得需要的覆蓋區域，是非常不智的。

### 地圖比例尺及轉換(MAP SCALE AND RESOLUTION)

大部分地理資源地圖以生成且儲存於主題圖的型態中，且在最近幾年，隨著數位化資訊系統的發展，使得已有可能有原始的野外觀察變數做更長遠的計畫。大比例尺地圖不只顯現更多位向關係的細節（空間的資源），而且通常也有更多細部的圖例（例：比例尺 1：25000 的土壤圖和通常描述土壤系列更大的圖例單位，當土壤圖的比例尺為 1：250000，將只展現土壤的關聯而已—見 Vink 1963）。

來源地圖的比例尺符合研究的要求是重要的一小比例尺地圖可能有不足的細節，大比例尺的地圖可能包括太多資訊，而變成對資料容量的一個負擔。很多測量組織提供他們的地圖資訊在比例尺的範圍上，使用者應該選擇在控制下最適合的課題的比例尺。

## 觀測密度(DENSITY OF OBSERVATIONS)

很多有關於觀測密度已被寫，需要去支持地圖或內插（e.g. Vink 1963, Burrough 1993b, Webster and Burgess 1984），但仍然有組織生產地圖沒有附帶任何資訊為基礎在地面實際狀況。這種態度正在改變當中—the Netherlands Winand Staring Institute 提供它的契約測量委託人，附帶顯示所有土壤觀測位置和分類的地圖。The UK Land Resources Development Centre(UK 資源發展中心)以公開顯示測量的密度和位置的取樣點和穿越線（transects）的圖。（see for example the Reconnaissance Soil Survey of Sabah, Acres et al.1976）

雖然實際的觀測密度可能是對資料可靠性程度一個合理的一般指導原則，不是一個完全的測量，例土壤已顯現變化的統計研究。對需要去解決所給的型態的觀測密度其粗糙的指導原則被“取樣定理（sampling theorem）”，從電子的信號發現創造，指出至少（2 觀測/信號元素）需要被製作，以便單獨地確認它。有也已被瞭解的工作在航空測量去估計觀測密度，需要由航空照片去製作在立體繪圖機上，以便去支援可靠的數位化地形模式（Makarovic 1975, Ayeni 1982）。

簡而言之，取樣密度只是一個對資料品質粗淺的指導原則，知道是否取樣是一個最適密度可以去解決空間型態的興趣，這也是重要的。這個主題在第 6 章和接下來的章節可見。

## 適當(RELEVANCE)

不是所有應用於地理資訊程序的資料對它們被使用的目的是直接適當的，但已被選擇作為代用品，因為想要的資料不是不存在，就是對蒐集而言太貴了，主要的例子從被用來估計土地使用、生物量或濕度、或被用來預測土壤生產力、沖蝕感受性或濕度為基礎的土壤型態學的土壤系列觀測的遙測電子信號，提供代用品和想要的變數間的適當，已被完全的確定，然而，代用品可以是一個良好資訊的來源。

代用品的校準刻度是遙測科技的主要部分，簡短地來說，在影像上像元的數目是由使用一個像“訓練樣區（training set）”來選擇。每一個記錄頻率帶的反應變化被展現於直方圖的型態上，使形式去選擇一個回到勉強分類的像元訓練樣區（training set），這些訓練樣區（training set）像元被“真實狀況（ground-truth）”觀測的測定刻度，所以像元的組可能是和農作物形式、土壤單位或任何其它可限定的現象相等的。影像內剩下的像元接下來被指定為和訓練樣區（training set）相同的樣區。訓練樣區（training set）使用以分辨分析在原始頻率帶的多變量空間的最小距離、最大的可能性或平行六面體的分類者為基礎的分配演算法。（見例子，Estes et al. 1983, Lillesand and Kiefer 1987）

## 資料格式化、轉換及內插(DATA FORMAT, DATA EXCHANGE, AND INTEROPERABILITY)

三種重要的資料格式化。第一種是純粹的技術面，就是資料如何從一個電腦系統轉換到另一個寫在磁片上的技術。這包括了媒體方面(如數化磁片、軟碟、硬碟)、寫入資訊的密度(錄音帶質控、軌跡以及密度)、使用錄音帶特性(ASCII 或二進位)、以及記錄的長度。對資料鄰接，兩個電腦間轉換速度配合是基本的，但大部分的數據機會自動地確定。

第二種格式化涉及的方法是資料的安排，或在其他的字詞中資料本身的結構。是否當資料涉及空間的實體，記錄如點、線、面在相關的模式中，如在物件方位系統中的物件，或獨立連續區域如網格般被記錄呢？如果面區域是在網格格式中編碼的話，那每一個像元的尺度

是多少呢？而是否這些資料的組成連潔特定的電腦系統在沒有轉換下互相交換是困難的？舉例來說，許多商業的地理資訊系統有其自己本身的內部結構(見第 3 章)而其可能使直接轉換資料變得困難。系統內插的流動與網格上資料資源的可獲得性趨使人們發展更廣為可接受的、可互換的資料結構，以符合較廣泛範圍的工業及國際標準(Schell 1995a)。

第三種格式是涉及到位置與屬性資料、尺度、投影與分類。尺度與投影可由使用在座標資料中特有數學轉換而更容易達到轉換(如 Maling 1973)。配合來自不同資源的分類是困難的，以及 GIS 地方性的應用可能因不同部門有完全不同的記錄方式，以致於沒有相同的分類系統進行相似實體的分類如道路、公共設施。

資料轉換通常要求共同元素格式的再格式化而可由許多系統容易讀取。這些格式不是必然的，且大部分的硬體不是有效率也不合算。衛星資料的資料格式，有其商業性的格式標準、國家的標準(如英國、荷蘭與德國等等)以發展為國際地理資料的標準。空間資訊的編碼、轉換的一般標準由歐洲的標準委員會、US 聯盟資料標準委員會建立，以及最近形成的 GIS 國際協會(Schell 1995a)。互相操作的問題促使人們去思考資料轉換的首要問題，而不如集中技術的論據。

### 易取得性(Accessibility)

不是所有的資料都是容易取得的。關於土地資料在一個國家可能可自由取得，但在其他國家土地資料可能被視為機密。此外軍事的 GIS 資料(這裡直接思考為數化地帶模式)內部辦事處也可防止資料的流動。成本與格式問題也嚴重妨礙資料取得。最近幾年來一種新的經紀人、資訊經紀人的湧現，可能幫助數化檔案的資料搜尋。詳細的資料服務可由政府及國際機構(如 EUROGI)獲得。同樣地有更多的資訊可在網際網路上取得。

### 成本及著作權(COSTS AND COPYRIGHTING)

收集與新資料的輸入、轉換及舊資料再格式化的成本。對任何物件，物件的經理人應該能評估使用存在資料的成本與效益。因數化手繪地圖的細節或連結空間資料的屬性，數化的成本可能會特別高。掃瞄器可以提供儲存資料輪廓線與影像的輸入。對特別的公共設施機構來說，測量的媒介可能使省略數化工作變得比在室內使用職員較為便宜。同樣地，如果媒介只有在資料轉換時偶然實在地執行，或輸出結果需要昂貴的裝置，如高度品質的雷射印表機，使用服務辦事處會比較便宜。

當尋找數化圖或使用空間資料或商業應用時，著作權發表的地圖或空間資料會因國別而變化及視每個案例的法定情形。

### 電腦中的數字誤差(NUMERICAL ERRORS IN THE COMPUTER)

在資料中本來就有的問題和原先電腦中存在看不見的誤差一樣。這最容易被遺忘，甚至電腦過程中臨界的部分，是電腦能存取處理在要求精確程度的能力。記錄數字的電腦語言的正確度對算術操作與資料存取有重要的影響。

許多人不重視電腦變化與排列的使用不充足的精密度，導致計算上嚴重的誤差，特別是必須有加減乘除才能獲得 2 個大量數字的要求時。舉例來說，建立一組數字變化的速記理論包括增加所有的數字、符合數字間的結果與分歧。不變的是所有數字平方的總合可獲得誤差值的總合。BOX9.2 解釋了當許多大數字包括幾乎必然大型誤差的發生時，電腦中每一位數

字是不宜處理精密度的要求。

#### Box 9.2 由電腦指令覆蓋所產生的誤差 (原文 p.228)

有一個簡單的窗框(下列表)試驗計算的正確度，其可證明電腦指令長度如何影響計算的結果(Gruenberger 1984)。1.0000001 是被平方 27 次(等於提升 1.0000001 為 134 217 728'th 動力 ( power ) )。這表顯示在個人電腦計算使用 Microsoft 中 4-byte 或 8-byte 加工處理一個 80486 的結果。在平方 27 次後單精度的結果較雙精度結果超過 1300% 的計算誤差！顯而易見的，其程式必須避免依賴呈現精確計算結果中的位置。

#### 四捨五入的誤差(Rounding error)

當形式化統計計算在大型電腦中且當程式語言允許兩個精密度變化與排列定義時，四捨五入的誤差便不是問題。它們使用 16-bit 的大型電腦，特別是計算的速記理論 (shorthand methods) 使用是困難的。在之前的例子，第一次計算 1 組數字的平均數是比較聰明的，然後計算從每個平均數的誤差值最後總合誤差的平方。平方總和的推估理論並不是只有接近原來的限定變化的理論而已，並且避免了在減法過程中四捨五入的誤差。

在許多系統是用影像分析資料進行整數編碼。正確呈現在網格格式中多邊形的面積與周長見第 3 章。法蘭克林(1984)探索資料精密度在其它 GIS 運算中的問題，如尺度與座標系統的轉換當算術操作的結果是刪減為最接近的整數時。圖 9.1a 顯示由三個點組成的三角形其中 P 點在尺度不同的三角形所呈現不同的結果。圖 9.1b 則是在圓周上 P 點的移動。

明顯的方法避免問題是增加電腦呈現數字的精密度，如與真實的十進位數字一同配合出現。

如法蘭克林所證明的，這僅僅只是將問題推至另一個層次，而並不是消失。這個問題是一個緊密連結電腦呈現數字的方法。為電腦簡易算術的執行違反重要真實數字分配的、結合的、與交替的定理發現真實數字是可能的。

舉例來說，等式(associativity)：

$$(A+B)+C=A+(B+C)$$

這規則是違犯在電腦中存取少數少於 10 的重要數字如  $A=1.E10$ ， $B=-1.E10$ ， $C=1$ 。法蘭克林顯示當其本身帶來複雜的額外問題與需要發展或使用特別的非慣例 (subroutines) 算術操作時，這些問題可由不同的計算理論所修正。

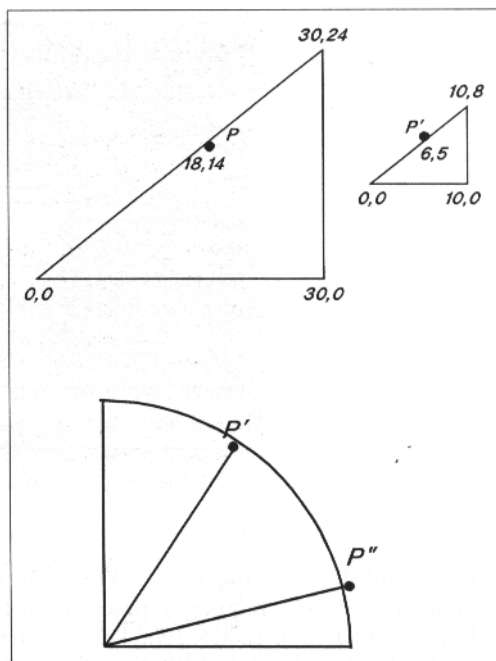


圖 9.1：整數算術、尺度或座標系統轉換可以使點接近邊界在內部或外部多邊形周圍（原文 p.229）

**地理座標系統與精密度(Geographical coordinates and precision)**

Chrisman (1984b)測驗了在其他 GIS 中硬體限制的作用，存取地理座標系統的命名以要求精密度的水準。當 16-bit 的機械 (machines) 在存取低解析的座標上呈現較少的問題，單一位置的大地衛星(LANDSAT)影像，由地籍系統要求高正確度，或垂直座標要求覆蓋大陸所得的數字，其數字因太大而無法由 16-bit 所記錄。所以這時 32-bit 甚至是 64-bit 就顯得有必要(表 9.1)。

幸而這不是一個太嚴重的問題。32-bit 在許多 GIS 中電腦流動使用頻繁，允許空間尺度由下列的精確性所記錄：

最大空間尺度 (公尺)Maximum dimension	最大可得精密度
10000.00	dddd.dx
100000.0	dddd.x
1000000	ddddd

在這裡 d 代表著好的資料，而 x 則是超過精密度需要避免模糊的位向關係(如圖 9.1)。當使用者要求一塊 1000\*1000km 的區域精密度大於 10m 時，則來自於如人造衛星 SPOT 計測器的資料其每 10 公尺的解析度，可能用來支持資源目錄的資料，這意謂著 32-bit 所創設的點呈現在 GIS 中是有所限制的。而且，在目錄中涉及背景控制點的位置以獲得較好的精密度是必要的。

Chrisman(1984b)、Tomlinson、及 Boyle(1981)指出當使用者希望分界那些習慣在不同的尺度與不同精密度的水準不同資料組的種類時，位置的精密度是具決定性的。當建立目錄時，這個問題就變的比較大且應是使用舊型 16-bit 系統的幾何學比當所有資料必須聚集在具體物件時來得更大，因為通常在後來的案例中，資料是從頭開始蒐集的。



表 9.1：在電腦指令長度，數化範圍與精確度之間的關係（原文 p.230）

透過硬體和軟體的國家及國際的協議和改變，資訊在數位化資料的品質上變成資料本身的重要部分。

## 有關空間實體真實性的前題假設所造成的誤差(Faults stemming from assumptions concerning the exactness of spatial entities)

如已經注意到的，一般被使用在地理資訊過程中大部分的程序假設 (a) ~ (f) 絕對的：(a)資料來源相同，(b)數位化過程確實，(c)地圖疊圖僅僅是相交邊界與再接合線段網路的分析，(d)邊界可以嚴格地界定及繪出，(e)所有的算術可被假設是在完全決定論的方法下進行操作的，以及(f)分類的區間對所有製圖屬性是最佳狀態。這個想法是源自於在資料中轉換方式的分類與製圖。他們出現了呈現大型技術的困難對 GIS 的設計者而言，但少有這些問題已經如在變化層面被發現影響觀察、記錄與製圖的問題。

許多地理資訊過程中的操作要求一個或更多空間網路的組合。空間網路的組合包含了線、規則網格、或不規則的多邊形。疊圖可能是為了資料轉換，例如轉換一個多邊形網由向量式至網格式藉由疊合網格而得到結果，或是為了資料組合或模式化的目的，例如當 2 個多邊形網是相交的，或當集水區的邊界是被用來分離經營單位疊合的區域，或當土壤多邊形的資料被輸入到農作物生產模式。

這個部分覆蓋的誤差可以由以下得到結果：(a)覆蓋空間實體如自向量式到網格式所呈現的多邊形；(b)自疊合與相交 2 個多邊形網路在空間實體是被確實定義的假設下。

## 來自於網格化向量式地圖的誤差結果(ERRORS RESULTING FROM RASTERIZING A VECTOR MAP)

**Grid cell are only approximations? 網格只是近似值?** 如圖 9.2 所顯示，轉換向量式三角形至單位像元結果導致資訊嚴重的損失。三角形的區域應該是 7 個單位，由計算邊緣得知，只能發現 6 個或 7 個單位。這個直角三角形斜邊應是 7 個網格長，如 4 個網格如斜邊的近似值，但只有 6 個如果我們選擇 3 個網格，2 個結果都高估。今天，網格近似值的誤差是較少有問題的，因為我們有更大更快的電腦資料庫。在製圖應用中如數化像片圖(plate1)及在許多雷射與噴墨印表機中網格化鉛筆繪製在向量平面圖更小更好—事實上，大部分的繪圖者今天也都使用網格的技術。只有當大的網格用來當基本的數據資料庫實體時，需要我們考慮到向量式與網格式空間描述不同的精密度。

## 混合像元(Mixed pixels)

當空間現象是由網格配置所呈現時，誤差會出現在 2 種方式。首先與大多可見的誤差來源是混合像元的問題，因為每一個網格只有含有一個單一的屬性值，這代表其值是由網格所導致的。在原來的大地衛星 LANDSAT 中，許多網格的尺度是 80\*80M，像元顯明的特徵是在所有網格區域的平均值；人造衛星 SPOT 其網格是 20\*20M，因此空間平均值較少。在網格尺度的不同代表著，如果大地衛星的部分覆蓋高度反射表面如水，這可能如此廣泛地代表反射，如給予一個水域的過度描述，與 SPOT 比較的話則可能記錄其他土地在 80\*80M 區域覆蓋的類型。這些分級的種類誤差可以引起不論網格尺度大於資訊描寫的特色。特別是當大區域網格用來記錄許多複雜景觀特色時就會出現問題(圖 9.3)。在向量—網格的轉換中，混

合像元的問題導致不論是否根據分類覆蓋幾何反射網格點的分類系統的兩端論或根據有系統的分級。在遙感與其他掃描影像中，問題更複雜因為網格值是區域計測器資訊尋找的權重平均值不但由像元所覆蓋而且也有鄰近地區的區域。

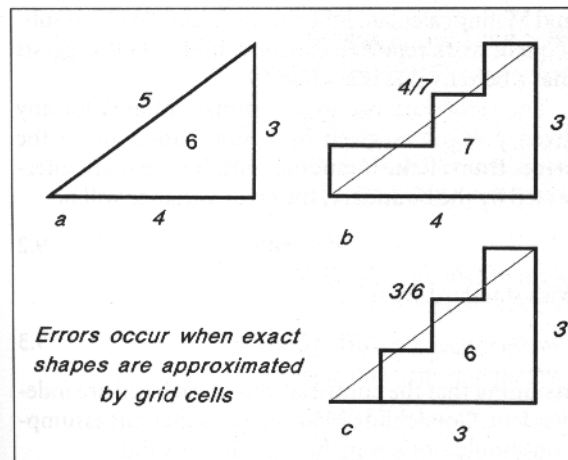


圖 9.2：當確實的投影是由網格近似值引起的誤差 (原文 p.231)

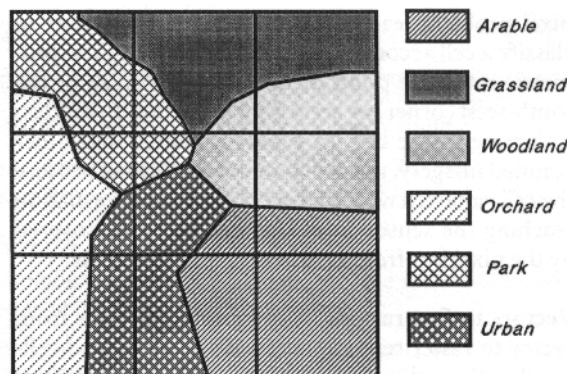


圖 9.3：當網格太大而無法解決空間的細節，則混合像元問題引起 (原文 p.231)

### 向量到精細的網格(Vectors to fine rasters)

當使用網格小於多邊形時轉換向量至網格的多邊形帶來位向關係配合的問題，當平滑的多邊形邊界是近似值時。儘管高品質的網格掃描與繪圖者經由網格轉換資訊而遺失有大量去掉，有許多的例子，如有主題資料、原來向量多邊形形式、需要配合網格的資料，諸如遙感所建立的或某些分析的例子見第 9 章。因此，去判斷不相配合所造成的嚴重問題就顯得必要了。Pinowar(1990)為向量至網格轉變進行好幾個算法的測驗，包括結果的品質、精密度、側面邊界的移置，及運算操作的速度。為此下了一個結論不是所有的算法都能運作相當好；某些快但會產生變形，而某些需要較長時間但可以得到較好的結果。

在接下來向量轉換為網格的討論中，多邊形會注意確實實體準確的位置邊界；轉變的誤差因此幾乎是由幾何學所呈現的地理區域。不好的識別誤差或無法確實的定義區域比較什麼而不是論述什麼。

### 統計應用在推估向量轉換為網格的誤差(Statistical approaches to estimating the errors of vector to raster conversion)

Frolov 與 Maling(1969)考慮到誤差出現的問題是當網格被確實的邊界線一分為二的時候。他們假設邊界線可以被視作直線繪製隨意橫過網格，這代表著規則每個一分為二邊界切

斷的部分，網格  $i$  (誤差變化) 可由式 9.1 所建立，

$$V_i = aS^4 \quad \text{式 9.1}$$

在這裡  $V$  是誤差變數， $S$  是線性網格的尺度， $a$  是常數。Frolov 與 Maling 計算的  $a$  值是 0.0452 但 Goodchild(1980) 後來計算出較好的值  $a=0.0619$ 。

誤差變數在區域的推估是由所有邊界網格的總合所得。如果  $m$  網格是由邊界所相交者，則誤差變數是

$$V = maS^4 \quad \text{式 9.2}$$

而標準誤差

$$SE = (ma)^{1/2} S^2 \quad \text{式 9.3}$$

假設每個網格的貢獻是獨立的。Goodchild(1980) 推測這個假設不總是正確的。

邊界網格  $m$  的數字，可以從多邊形的周長推估得知。Frolov 與 Maling(1969) 顯示  $m$  與是成比例  $\sqrt{N}$  的，當  $N$  是多邊形網格上數字的總和。 $m$  的標準誤差是由  $kN^{1/4}a^{1/2}S^2$  所推估的。因為多邊形  $A=N \cdot S^2$  的推估，標準誤差如推估的百分比與  $N^{-3/4}$  是成正比的(Goodchild 1980)

$$SE = ka^{1/2} N^{-3/4} \quad \text{式 9.5}$$

如果變數是網格邊長  $S$  代替  $N$ ，則百分比誤差依賴  $S^{3/2}$ ，Goodchild(1980) 研究指出這些關係已證實了。

不變的  $k$  依賴多邊形的投影，長薄形的投影比圓形投影邊界長。Frolov 與 Maling(1969) 為標準投影設定了一個  $k$  值，使用依賴直線假說。

## Switzer 理論(Switzer's method)

Switzer(1975) 描述一個一般由向量多邊形圖而成網格影像精密度推估問題的解決方法。他分析並不處理觀察或位置的誤差，但假設誤差是完全使用一系列點座落在網格中心以推估與原來地圖相似的網格說明的結果。Switzer 理論處理基本的標準色彩圖，如在相似製圖單元中被非常薄、尖的邊界所分立的主題圖。這個理論假設真實地圖的存在，對抗由取樣所建立的地圖進行比較。易於了解的真實的地圖通常不為熟識或了解，Switzer 顯示由應用某些假設及結論的統計值，不合的誤差可由推估或網格本身得知。

由假設開始的分析其地圖  $M$  被分割為相似製圖單元  $k$  或顏色。每一個製圖單元會呈現由一個或更多次區域組成。真實的地圖由  $n$  的配置推估得基本取樣網格。在這裡我們將只要考慮位置有取樣格子配置是規則與一致的結果，以及每個網格是由單一樣點的中心點所定義的。在真實地圖上製圖單元表示  $M_1, M_2, \dots, M_k$ ，以及在由  $M_1, M_2, \dots, M_k$  所推估的地圖。在推估地圖的每個網格是分配至製圖單元  $M_i$ ，如果取樣點落在真實地圖上製圖單元  $M_i$  的話。這通常向量多邊形網格轉換至網格格式的步驟。就我們所知這個分析的目的是要假設地圖所有區域的尺度是一致的。如  $A(M)=1$

推估地圖不當配合的程度是兩個獨立要素的函數，(a) 真實地圖的複雜度，(b) 取樣網的幾何學所有物。首先考慮地圖的複雜度，我們可以界定一個數量  $P_{ij}(d)$  的可能性，其可能性指的是隨機點在真實製圖單元  $I$  及網格中心點在製圖單元  $j$  是由距離  $d$  而分立的。Switzer 獲得以下百分比重疊部分  $O_{ij}$  每一對製圖單元  $ij$  的表達，

$$O_{ij} = 0.60 P_{ij} (n^{-1/2}) - 0.11 P_{ij} (2n^{-1/2}) \quad \text{式 9.5}$$

(係數值不同於 Switzer 所發表的公式；而 Goodchild 校正其值)

所有 k 製圖單元的總誤差則由下列式子得之

$$O = \sum_{i=j}^k O_{ij} \quad \text{式 9.6}$$

BOX 9.3 顯示 O 如何由練習中推估。

舉例說明，圖 9.4a 顯示簡單主題圖的邊界描述土壤或地質單元。假設他們形式化真實地圖，將關係著不合誤差從數化時使用不同尺度的網格中出現。如圖 9.4a、b 的 16\*16 與 32\*32 的網格一樣，分別地的網格？

表 9.2 得到其結果，一個網格測量 16\*16 的格子，在距離 d=1 時就有 960 個格子。格子數的總和跨越一個邊界導致推估每一製圖單元的頻率。當距離 d=2，堆數是 896。進入方程式 9.6 推估頻率導致每分 9.5 不合的推估。使用 32\*32 的網格導致每分 4.1 的推估，證明因素 4 在網格數的增加需要減少部分推估誤差。比較不合推估與良好的不合推估由測量在原來地圖上製圖單元區域所獲得。這代表 600dpi 的不適當配合幾乎是 300dpi 的一半。

Box 9.3 Switzer 理論 (原文 p.233)

$P_{ij}(n^{-1/2})$  與  $P_{ij}(2n^{-1/2})$  可能性是自下列頻率數的推估得知：

1. 在距離  $d = 1$  推估網格的組數，每一組數字距離是相同的。

$$NPAIRS = 4 * (P * Q) - 2 * d * (P + Q)$$

P = 行數，Q = 列數，第二次收集邊緣網格

2. 每一對製圖單元 i、j，上下網格數基於不同製圖單位 (TALLY<sub>ij</sub>)
3. 計算  $P_{ij}(n-1/2)$ 、TALLY<sub>ij</sub>/NPAIRS
4. 重複 1-3 的步驟，以推出  $P_{ij}(2n-1/2)$
5. 由方程式 9.5 計算  $O_{ij}$
6. 計算  $O_i$  不合的總和  $O_{ijs}$ ，且  $O_{ji} = O_{ij}$

計算  $O_{i,s}$  的總和

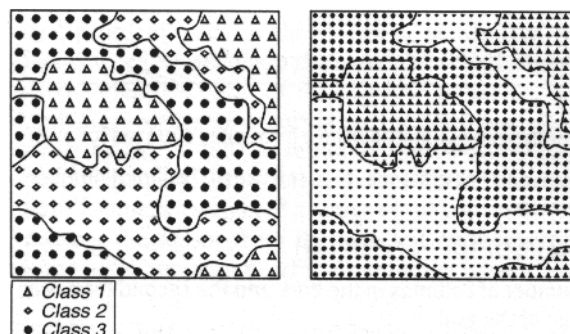


圖 9.4：網格化向量圖在 2 個網格尺度以推估網格誤差 (原文 p.234)

表 9.2：在圖 9.3 使用 Switzer 方法於地圖上的結果 (原文 p.234)

Brege et al.'s 方法

Bregt *et al.* (1991)對於估計因為結合向量與網格轉換的誤差，發展了一個較佳的方法，因為其必須對圖形進行網格化二次，故被稱作二倍轉換法 (the double conversion method)。首先，將向量轉成網格時必須先依據需求目標決定需求網格尺寸，這個結果便是大家所說的基礎網格 (base raster)。之後圖形被網格化為極小的格子，且兩者之間可以相互比較，這些小的網格與基礎網格的不同造成了基礎網格估計的錯誤。

Bregt *et al.*比較誤差得出一個參數，稱作邊界指數 (boundary index (BI))，其被定義為圖上每平方公分單位中的邊界長度(公分)。BI 的計算是藉由多邊形邊界的總長度區化導出。他們發現對一格約定區域尺度的網格化誤差(錯誤比率)是一個 BI 的線性函數。他們區別出以下兩種情況，(a)網格在基礎網格上的點位是根據多邊形的中央點進行分類，且(b)多邊形網格分類的方法支配及面積，表 9.3 表示出這樣的結果。

表 9.3：網格化方法、網格尺寸、及網格化誤差的關係 (原文 p.235)

網格化方法	網格尺寸(mm)	回歸方程式	變異說明(%)
Central 中央	1×1	L=2.4BI	99.8
Dominant 優勢	1×1	L=2.4BI	99.8
Central 中央	2×2	L=4.7BI	99.8
Dominant 優勢	2×2	L=4.6BI	99.8
Central 中央	4×4	L=9.0BI	99.4
Dominant 優勢	4×4	L=8.7BI	99.0

Bregt *et al.*比較了他們與 Switzer's 的方法，且明確的指示出這個方法對於網格化的誤差估計是較容易且好用的，指示需要計算 BI，其中 BI 值是獨立的使用單位。他的缺點在於回歸方程式需要在所有可能情況外操作，然而 Switzer's 方法是一般且不需要前面的工作。

## 數位化地圖或編碼整合時的誤差 (ERRORS ASSOCIATED WITH DIGITIZING A MAP, OR WITH GEOCODING)

十分顯著的，Switzer, Goodchild, and Bregt *et al.* 的方法估計錯誤是含蓄的假設一個“真實”圖的存在，即具有均值的圖像單位且邊界無限而明確。無論如何，在操作上即使是 best-drawn 圖也不是完美的，且在數化的過程中額外的誤差也會進入。Blakemore (1984), Bolstad *et al.* (1990), Dunn *et al.* (1990), and Polker (1982)已經指出，在向量格式的多邊形數化圖上，必須考慮邊界寬度與位置的問題(圖像單位的同質性將在稍後進行處理)，而數化也幾乎是以圖的形式展現。以下是兩個可能的誤差來源：(a)來源圖的整合誤差，及(b)數化圖的整合誤差。

(a)部分潛在原因為紙張延伸、印刷扭曲或來源檔所產生的誤差，這是可以糾正的，而邊界位置所產生的誤差卻相當容易發生，因為在繪製邊界時並不是無限的微小。一個 1 毫米寬的線在 1:1250 的圖上覆蓋的面積是 1.25 公尺寬，圖樣的線在 1:1000000 的圖上覆蓋面積為 100 公尺寬。一張 1:25000 詳細的土壤或地質圖測量了 400 x 600 毫米，其繪圖線也許最大已經涵蓋面積 24000mm 在 24 000 sq. mm 圖上，為其 1/10!一般認為真實的劃分線應該被繪至於繪圖線的中央點，很諷刺的，被邊界線覆蓋的面積會相當容易成為一個不確定、可能及混亂的面積。當這些邊界線被數化加以轉換時，因為操作人員手工數化將使得其無法永遠精確

的位於線的中間，且在掃描器方面，誤差會因為數據簡化的運算法則等因素增加。

(b)彎曲陰影的表現有賴於頂點數的使用(Aidred 1972: 5)，因此，數化直線的誤差將會較數化合成的曲線少許多，轉化一個圖上連續彎曲的線成為數化影像使其成為一個取樣的過程：沿著曲線的無數微小的可能點採樣(見圖 9.5; Smedley and Aldred 1980)。

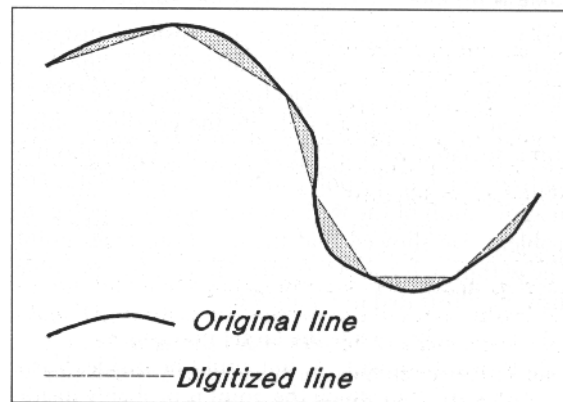


圖 9.5：數化一條線是一個取樣的過程（原文 p.235）

顯而易見的，主題圖的邊界被獨立的考慮，但在有一個整合錯誤帶或具信心的區域時例外。MacDougal (1975)建議估計整體錯誤的邊界以

$$H = \sum_{i=1}^N (h_i l_i) / T \quad \text{式 9.7}$$

其中  $h_i$  為  $i$  線段的水平誤差(in standard deviations 在標準差內)，長度為  $l_i$ ， $N$  為邊界線的數目，而  $T$  是圖形的總面積。假如所有的邊界線是相同型式(例如他們皆為土壤邊界或皆為土地使用邊界)，其方程式可簡化為

$$H = (hL) / T \quad \text{式 9.8}$$

總長度  $L$  是以原來在圖上的網格位置加以估計，並計算交叉點的數目  $K$ ，而其使用公式如下

$$L = (TK) / 0.6366 \quad \text{式 9.9}$$

其中 0.6366 是 Wentworth (1930)所述的常數，但今日總長度卻能夠藉由數據庫簡單的被計算出。

在一個經驗的研究中，Bolstad et al. (1990)標示誤差來自於人工數化 1:25000-1:50000 的土壤圖是如此的小，而對美國來說，這原因的重要性小於由於不確定的地理注記導致的位置誤差。

### 沿著數化線的誤差範圍：點在多邊形內的搜尋及何時整合網格及向量數據的問題(error bands around a digitized line: problems for point-in-polygon searches and when combining raster and vector databases)

Perkal (1966)建議一個“緩衝模糊帶 (epsilon)”距離應被定義為大約在製圖線周圍，作為一個客觀概括性的意義。Blakemore (1984)提出相反的概念以顯示整合誤差帶寬度所可能發生的混亂，是關於多邊形邊界數化型式與易懂的多邊形內點 (point-in-polygon) 問題間的相關性，他表示問題‘是否點  $P$  落於多邊形  $A$ ’，最後得到五個答案，圖形表示如圖 9.6。

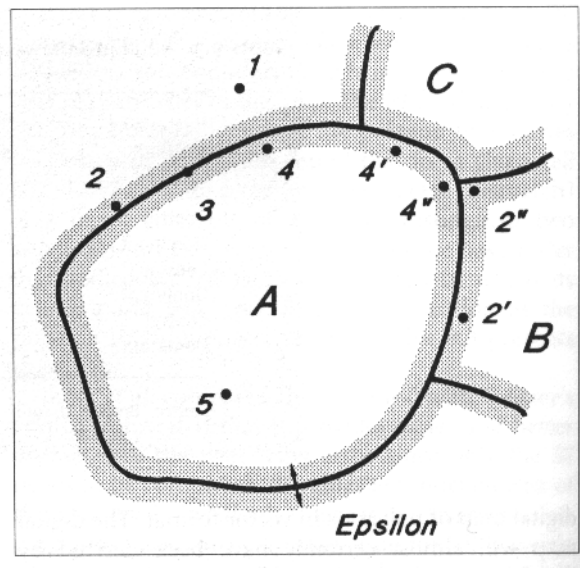


圖 9.6：沿著被數化線的一個緩衝模糊帶 (epsilon) 的 Perkal's 概念 (原文 p.236)

1. 目標多邊形 A 的外部明確
2. B or C 點可能位於 A 之外，但也可能在裡面。當變異為 2' 時，點可能在鄰近的 B 之內但也能夠在 A，而變異為 2'' 時，其點可能為於 A 之外，但在 B 或 C 之間。
3. 在邊界模糊不確定的區域。
4. 點可能位於 A 之內，但也可能在外面;其他的變異質為 4'，則點位於 A 之內，但可能鄰近 C，而變異值為 4'' 時，這'in'點可能也位於 B 或 C 兩個鄰近區域內。
5. Definitely in A. 明確的在 A 之內。

“明確的在內部 (Definitely in)” 表示在誤差帶內的核心面積；“也許在內部 (possibly in)” 表示一個點落於確定帶及多邊形的內半疊圖上。“也許在外部 (Possibly out)” 表使一個點落於確定帶外半部；技術上的說法，這個點將被回復為落於多邊形外部，但事實上，假若遇到數化或編碼錯誤時，他可能是落於“真實 (ture)” 多邊形內部。一個不確定點具有座標時，便能確實的符合數化邊界上的點—如此的點是稀有的必需盡力讓它發生 (do occur)。

Blakemore (1984)圖示說明當處理整合向量多邊形網路與正方形網格網路時，誤差種類的影響問題。他選擇的問題為在一張多邊形圖上 115 個英格蘭西北工作區域的工業機構，其疊合於 UK 工業部門 1 公里正方形基礎網格，對於一個 1 公里的正方形網格的圖形辨識率，其數據編碼共有 780 個項目，以測試在多邊形網路內的內容。一個 1 公里正方形網格會導出一個 0.7071 公里的緩衝模糊帶 (epsilon) 或確定區域。表 9.4 表示 Blakemore 的結果。

表 9.4：Epsilon 誤差結果 (原文 p.237)

種類	%
Possible out definite 可能在外部的定義	1.5
Possible in 可能在內部	4.4
Unassignable 不能移轉的	1.4
Possible in/out 2 polys.	29.8

可能在兩個多邊形內部/外部	
Possible in/out > 2 polys. 可能大於 2 個多邊形的內部或外部	6.7
Ambiguous 不明確的	1.2
Subtotal 小計	45.0
Definitely in 明確的	55.0
Total 總計	100.0

這個“可能在外部的定義 (possibly out definite)” 種類包括數據點落於全部工作區域的多邊形網路外。“可能在內部 (Possibly in)” 表示點落於多邊形網路內多邊形邊界上的誤差區域的內半部。“不能轉讓的 (Unassignable)” 表示點落於外部多邊形的外部邊界的外部誤差區域內。在一些情況下，多邊形內點 (point-in-polygon) 常被用於表示產業在海上的位置。“可能在 2 個多邊形內部/外部 (Possibly in/out 2 polys)” 表示點落於兩個緊鄰多邊形的裡面或外面；“可能在內部/外部>兩個多邊形 (possibly in/out > 2 polys)” 表示點可能落於超過兩個多邊形的內部或外部。“不明確的 (Ambiguous)” 表示這個點目前存在於數化多邊形的邊界上。這個研究表示在資料庫內的產業員工只有 60% 能夠被明確的整合在工作區域。這些整合的誤差及不確定性將會因為長度而增加，微小的多邊形和工作區域相較於大或廣闊圓形區域有狹窄的突起或鑲嵌。

這研究在於資料庫的確認和檢驗的許可數量，以確立誤差所帶來的網格點編碼的移動。Perkal 的模糊緩衝帶 (epsilon) 假設邊界為真實的，其問題僅在於一個已知的位置。有時他並不是這個位置，邊界的存在讓人懷疑，而必須使用其他的方法—見第 11 章。

## 隨著重疊兩個或更多的多邊形網路的整合誤差(ERRORS ASSOCIATED WITH OVERLAYING TWO OR MORE POLYGON NETWORKS)

一個區域內兩個或多個主題圖間的空間整合，通常被另一多邊形頂端上的外部線及同時發生的邊界展示或調查。在數化圖的時間之前，其過程成功的使用清楚的表格，而疊合的邊界被較粗的製圖線描述。因為所有的邊界被精確的進行編碼，使得數化圖的開始便允諾有更好的結果，但事實上新技術的結果在電腦製圖 (cartography) 造成更困難且更需要進行研究的問題。不只在技術上解決問題花費許多年的時間，而且調查疊合區的結果有更多尚未解決的數據品質及邊界不疊合的問題。

McAlpine and Cook (1971) 是第一個調查當工作在具土地資源的問題而他們的方法至今仍舊被保存。他們認為兩個同樣位置的圖包括了各自的  $m_1$  和  $m_2$  ( $m_1 \geq m_2$ ) 於開始的圖段 (polygons 多邊形)，其疊合後產生  $n$  圖段。為簡化這個問題，他們以具任意方向的單一六邊形進行試驗，並在鑲嵌的六邊形上進行置換，這樣的實驗被用在邊為 0.5、1、2 及 3 的鑲嵌六邊形邊上。他們發現導出的多邊形數  $n$  在導出圖上可被估計為

$$n = m_1 + m_2 + 2 \cdot \{m_1 m_2\}^{1/2} \quad \text{式 9.10}$$

對兩張圖而言其  $k$  圖可被一般化為

$$n = \left[ \sum_{i=1}^k m_i \right]^2 \quad \text{式 9.11}$$



McAlpine and Cook (1971)表示在衍生圖上圖形的疊合會使得相當大比率的小多邊形增加。他們應用他們的分析在一個案例研究中，疊合 3 張比例尺為 1:250000 的統計圖，呈現土地使用的強度及 Papua and New Guinea 土地系統，共包括 7、42 及 101 個最出各自的多邊形，共導出 304 個多邊形。(以等式 9.11 估計可導出 368 個多邊形，但 McAlpine and Cook 把這視為滿意的結果)這疊合的過程使得區域內 38% 被少於 3.8 sq.公里的多邊形覆蓋。

疊圖的結果是依據衍生的多邊形尺度與邊界的複雜性分類加以估計(例如多邊形單獨受限於最先的圖段，這些邊界只為土地使用及土地系統的界線，且受限於所有三種型式的邊界)。衍生多邊形 10% 的任意取樣是被三個決定是否同意的量度在原先與衍生多邊形種類間加以估計。在圖 9.7 表示，缺乏一致是衍生出最小多邊形的重要因素，而衍生圖內 30% 的面積因為多邊形與先前的描述較少或沒有一致性產生。

Goodchild (1978)擴大討論多邊形疊合問題，以顯示導出多邊形的數量比位在被疊合圖上的多邊形數量有更多複雜的邊界函數。他指出兩個疊合的多邊形有各自  $v_1$  及  $v_2$  的端點，當所有布林運算包括利用非 A 且非 B (NOT.A.AND.NOT.B)，則可由 three to  $v_1.v_2 + 2$  導出任意數量的多邊形。在 McAlpine and Cook 的案例中，當產生適量多邊形時，疊合的圖可表示統計上的獨立。無論如何，當來源圖上多邊形邊界具高度關聯，透過產生大量小而“偽 (spurious)”多邊形會產生嚴重的問題。在幾張圖裡，顯著而重要的特徵如區域邊界或河流也許會成為多邊形邊界的部分。這些相同邊界的幾種表現方式將會被個別的數化，但因為數化及其他的誤差，其將不會確切的重合。

這些偽多邊形包含兩個明顯的矛盾點。首先，為更精確每一個邊界是被數化在個別的圖上，而更多的座標的使用，便產生更多的偽多邊形。第二為繪圖時個別的方法，當採人工法進行疊合時，需設計以輔助肉眼，這樣的數化圖會產生大問題。

Goodchild (1978)分析偽造多邊形的情況，最可能透過兩個相同卻個別的數化弧、及  $n_1$  和  $n_2$  端點版本發生(圖 9.8)。Goodchild 正利用二元的符號進行統計，顯示偽造多邊形數量  $S$  是依據兩弧具有  $n_1$  和  $n_2$  端點範圍由

$$S \min = 0 \quad 9.12$$

到

$$S \max = 2 \min(n_1, n_2) - 4 \quad 9.13$$

具有一個任意的期望值

$$E(S) = [2n_1, n_2 / (n_1 + n_2)] - 3 \quad 9.14$$

假如沿著這個結合面積，標號連續而隨意的發生。 $S$  的最小值發生在疊合圖時，一個型式的標號存在一起，而  $S$  的最大值發生在最大的混合時。藉著模擬五種結合弧可能的情況，Goodchild 以等式 9.14 估計偽多邊形最高平均數量可能發生在 17%。實際的偽多邊形數量被發現從未超過  $S$  最大值的 71%。無論如何，越是小心的一張圖被數化，便有越大的  $n_1$  及  $n_2$  值，而更大量的偽多邊形將會產生。

事實上偽多邊形便同等於網格化多邊形錯誤面積的結果。當數化的精密度增加時，他們的總面積應該會減少，但更大的問題是，他們在最後的圖上為避免無意義而進行的移動。他們可為求最小面積，或兩端點可被直線連接及消除邊，篩選多邊形，並在隨機的基礎上消除一個邊。更多複雜的途徑是去考慮由複雜的弧內約定距離的所有點當作新線段位置的估計，而後藉由最少方形 (least squares) 或最大可能性的方法是合一個新的線段。除非數化邊界的一個型式可以被限定，高度相似的複雜線將由其位向上的“真實”位置被移動。疊合土壤圖

(沒有確切的邊界定位)和郡縣邊界圖(位向明確的邊界)的結果，也許位向上的邊界將會變的扭曲，除非使用者指定他們必須維持不變。

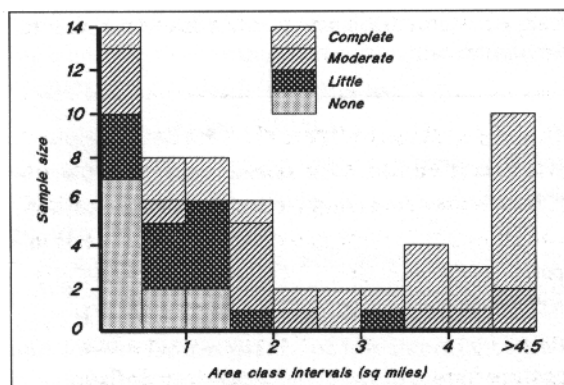


圖 9.7：在多邊形疊合後最初與衍生多邊形的測量 (原文 p.238)

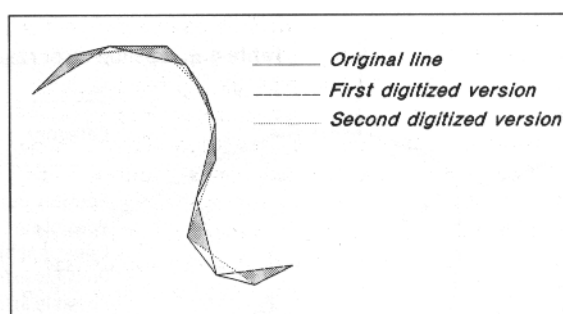


圖 9.8：當相同的線段被數化兩次，在圖上偽多邊形如何發生 (原文 p.238)

---

對於重疊的實體或交集，轉換一種表現法至另一種的問題，其採用精確的邊界或平滑的輪廓線，但這些方法估計誤差的存在會使得運作上更為困難。誤差的程度導因於空間現象轉為可能不恰當、精確、清楚定義的實體，其以被承認有較少的關注，但有主要未見的誤差來源及資訊的錯失。具有不確定邊界的地理現象背涵蓋於第 11 章(See also Burrough and Frank1996)。

---

### 結論:誤差與錯誤(Summary: errors and mistakes)

在每一個製造的過程中，不良的物質會導致不良的產物。可是空間資訊系統 (Spatial information systems) 中，假若在數據蒐集、塑造與分析上沒有適當的注意，他也有可能使良好的物質轉換為不良的產品。依照慣例，數據的品質關聯到地理座標上的精密度，但今天，精確的位置卻不是良好品質的一個方向。讀者也應該意識到有時人們期望一個高品質的產物勝過具相當的可能性或必需性，例如，對於自動導航系統 (autonavigation)，它相當需要幾何資料庫及真實事務的精密度，但對於一些行銷上的研究(e.g. Plates 2.5-2.8)特別是空間上的正確度卻不是唯一的需求，但卻威脅個人隱私。

### 問題(Questions)

- 1.回顧能夠被使用作決定空間誤差的不同方法。思考不同 GIS 應用範圍及在不同的應用中決定合適的誤差分析技術。
- 2.設計一個改變 meta 數據系統以記錄數據品質及誤差增加，作為空間數據設定的工作

方向。

3.回顧四個實際情況缺乏的有關臨界誤差資訊以承認 GIS 分析結果。

4.整理編譯第 6、7 章 GIS 分析的案例其誤差來源表，並分類本章中有關誤差的術語。  
成功的分析評論每個案例決定誤差來源。

### **未來閱讀的建議(Suggestions for further reading)**

1. DavrD, B., VAN DEN HERREWEGEN, M., and SALGB, F. (1996). Conceptual models for geometry and quality of geographic information. In P. A. Burrough and A. U. Frank (eds.), *Geographic Objects with Indeterminate Boundaries*. Taylor & Francis, London.
2. GOODCHILD, M., and GOPAL, S. (1989). *The Accuracy of Spatial Databases*. Taylor & Francis, London.
3. GUPTILL, S., and MORRISON, J. (eds.) (1995). *The Elements of Spatal Data Quality*. Elsevier, Amsterdam.
4. THAPA, K., and BOSSLER, J. (1992). Accuracy of spatial data used in Geographic Information Systems. *Photogrammetric Engineering and Remote Sensing*, 58: 841-58.

# 第十章 數值化模型之誤差擴增

## *Error Propagation in Numerical Modelling*

### 前言

本章節主要探討空間資料沒有保握的統計運算來源，特別是關於建立及使用連續性的區域及由數值化模型中之誤差累積。藉由闡述在不同尺度、取樣間隔及空間關聯性結構中造成的空間誤配之誤差如何累積，本章將誤差之擴增理論及工具呈現出來。誤差分析之蒙第卡羅方法(Monte Carlo methods)為接在誤差擴增統計的理論之後。內插法之比較研究及藉由連結於簡單的成本-效益(cost-benefit) 分析之迴歸模型分析誤差擴增，讓使用者知道如何選擇最好的技術以預測不同狀態重金屬污染。獲取誤差及誤差擴增之知識後，精明的 GIS 將可以核對 GIS 模擬結果之內在的品質，並可以告知使用者如何得到其要求的精確度結果。

### 數值化模型之誤差擴增之統計處理(Statistical approaches to error propagation in numerical modeling)

在大部分 GIS 之數值環境模擬中，屬於點、線、多邊形或網格的屬性被輸入於數值化模型中以計算導出成為新的多邊形或格子屬性。於計算中經常假設可以得到精確的結果，因為大部分 GIS 並未提供方法以驗證我們數入資料所得到結果之誤差。然而非常明顯的，數量化模型所得到的結果其品質取決於三個因子，

1. 資料之品質；
2. 模型之品質；
3. 資料與模型之作用方式。

為了取得較可信賴的結果，必須知道模型參數及資料擴增過程中不確定的因子。分析誤差擴增(對最終結果不明因子的累積及影響)，我們需要：

1. 誤差估計的來源；
2. 誤差擴增的理論；
3. 誤差擴增的工具。

### 統計上不確定的來源(Sources of statistical uncertainty)

資料統計不確定性之來源包括量測誤差及無法由物理模型解釋的關聯性空間變數。第六章顯示地理統計內插及狀態模擬(隨機繪圖 stochastic imaging)之方法，可以產生內插之誤差值。在 GIS 儲存資料的錯誤最簡單的方法是假設所有資料是常態分布並且其誤差可以由標準偏差正確呈現，這意味著實體或網格之所有屬性必須變成由兩個數字呈現，此紀錄或平均值(mean value)及標準偏差(standard deviation)；當然自動導致儲存成兩倍的空間屬性資料。連結 GIS 至統計事件可以說是將資料的參數統計計算。

### 關聯性空間資料錯配之影響(Effects of mismatches in spatial correlation structures)

一個非常重要但未被察覺在空間資料不確定性之來源，此或許是收集資料時之資料大小問題，特別是在資料來自不同來源而必須被結合在 GIS 中時。縱使資料共享同一網格，然而當以不同解析度取樣得到之空間形式時，不同組的空間資料亦會被錯配，簡單鋪蓋在多邊形或向量及網格或特徵平面上已相同幾何解析度並不能保證可以將不同圖層以相同方法疊合以得到空間資料。此種觀念指出觀察在大區域中空間差異未被察覺之部分是必須的，此種疏忽是非常普遍的，除了在少許出版物上有被注意到。(eg. Burrough 1993, Beckett and Webster 1971, Openshaw 1977, Openshaw and Taylor 1979).

關聯性空間及時間結構上的錯配或許是在結合不同資料來源或現象或屬性上之重要的問題，並且會導致使用 GIS 的理想破滅，這並不是 GIS 的問題，而是資料收集或結合時之問題。

關聯性結構是重要的 Correlation structures are important!

假使取樣調查不同屬性轉向為不同關聯性空間結構，得到合理的空間模型配對是非常困難的，此種問題並非 GIS 所有，而是在自然現象中空間變異平衡所造成的。分析空間共變數指數如變異圖可以使用於確定不同資料在空間上是相容的。

錯配可能經由下列情形發生：

1. 每個現象以不同量測方法(面積/體積)；
2. 每個現象有其不同內在的空間變異；
3. 許多現象在其他現象以外部資料不適當加入空間聚集區塊時直接被收集或集合時；
4. 不同現象的空間變異藉由不同操作尺度過程被控制著，相同的屬性在不同尺度下可能有不同關聯性空間結構。

下列例子(Ten Berge et al. 1983)舉例證明相同變數同時發生之測量法，在由不同提供之測量方法以轉換為不同關聯性空間結構時是困難的，此研究之目的為繪製橫跨 350 公尺寬之土壤表面溫度之空間變數地圖，為在 Ir. A. P. Minderhoudhoeve 試驗農田位於 Flevopolders, 荷蘭。此區域非常平坦，但是由於在農場內有不同之沉降作用形式，自 1980-2 每年早春以紅外線調查試驗田土壤表面溫度。

現場調查是此工作量的兩倍，第一，調查者希望所調查赤裸田地表面之特質與田地之溫度變數相關，最後去判定表面溫度變化連結到作物反應。第二個目標是反射輻射空中的測量反應到地面的測量，不僅大氣吸收作用不會影響並且此方法既便宜又容易於任何時間攜帶。

此試驗是在晴朗天氣狀態下於 1986 年五月 26 日中午進行，熱感應紅外線由飛行的 Daedalus DS 1260 多光譜掃描器使用 8-14  $\mu\text{m}$  光譜視窗(Daedalus Enterprises Inc., PO Box 1869, Ann Arbor, Michigan 48106, USA) 於輕型飛行器上(Plate 4.2)。此飛行器攜帶此掃描器之高度產生出一個之像素為 1.5  $\times$  1.5 公尺，亦即意味著在此區塊內之所有變化皆被平均為一個單一資料值(加權平均)輸出。

反射輻射之地面資料及土壤特性同時被飛機收集著。由此比較，我們僅使用由土壤表面散發出之熱紅外線感應並且此資料包含土壤 0~5mm 表面層結構。每隔 4 公尺收集一個定點資料，共兩個橫貫 200 公尺之資料(50 個觀測點資料)，此兩組資料成直角，以預測此區之溫度梯度。此橫貫之資料被明顯的標定，所以可以輕易座落於空中熱梯度中。土壤表面散發之紅外線熱能資料由架設在三腳架上手提式 Heimann KT-15 輻射計收集，此輻射計設定一特別解析度為 0.03 平方公尺(0.17  $\times$  0.17 公分)如同為光譜視窗上 8-14  $\mu$ 。土壤特性是由一個由九

個不同密度取樣單位鐵篩網測量於 1.00 x1.00 公尺區域而取得，因此有效的建立此 1 平方公尺內之土壤質地資料，除此之外，土壤表層 0~5 公分之溫度資料由一溫差電偶探針讀出數位資料。

簡單來說，此調查方法提供兩組互相垂直 50 點位置資料，而此資料為以前亦曾以不同形式調查過。空中紅外線掃描資料有著 1.5 x1.5 公尺之解析度，而 Heimann 紅外線資料有著 0.17 x0.17 公尺之解析度，並且土壤質地調查有效提供 1.00 x1.00 公尺之土質資料。所有第一次調查資料是常態的，並未發現嚴重非常態分布之資料。

當不同資料被繪製到同一空間變一時，不同支持大小之影響成為證據。圖 10.1 顯示原始資料點及其變異圖，可以被作為標準化的比較。非常明顯的，此兩種紅外線測量方法之間並無任何關係，空中資料反映出與田地有關之空間變異，但是 Heimann 資料顯示只有小範圍變異或干擾；變異圖模型無法對應這些資料。溫差電偶之觀察證明為何如此。此試驗田曾經被翻犁過並成 15-20 公分高之脊狀，部分為 25-30 公分。脊的一邊被五月斜射太陽照射使其溫度達到攝氏 14-15 度，因此小範圍之溫度差異導致 Heimann 讀取資料時有較大的波動。因為 Daedalus 掃描之像素較大足以平均掉小範圍之溫度差異，因此重新以此量度此田地較長範圍之資料。由空中收集之之土壤質地資料顯示其樣式類似於 Daedalus 儀器所收集的，此資料為橫跨田地土壤表面溫度漸進變化是與土壤表面質地之空間變異有關。

Lessons learned. 本段落提供一例子使清楚提供之尺度大小，以往曾經使用調查過之'點'資料或 bulked 取樣，這些資訊必須被提供為一個變換資料(metadata) 檔案，此檔案描述其來源及公家或商業資料庫品質(第 12 章)，但是大部分時候常常是無法取得這些資料。

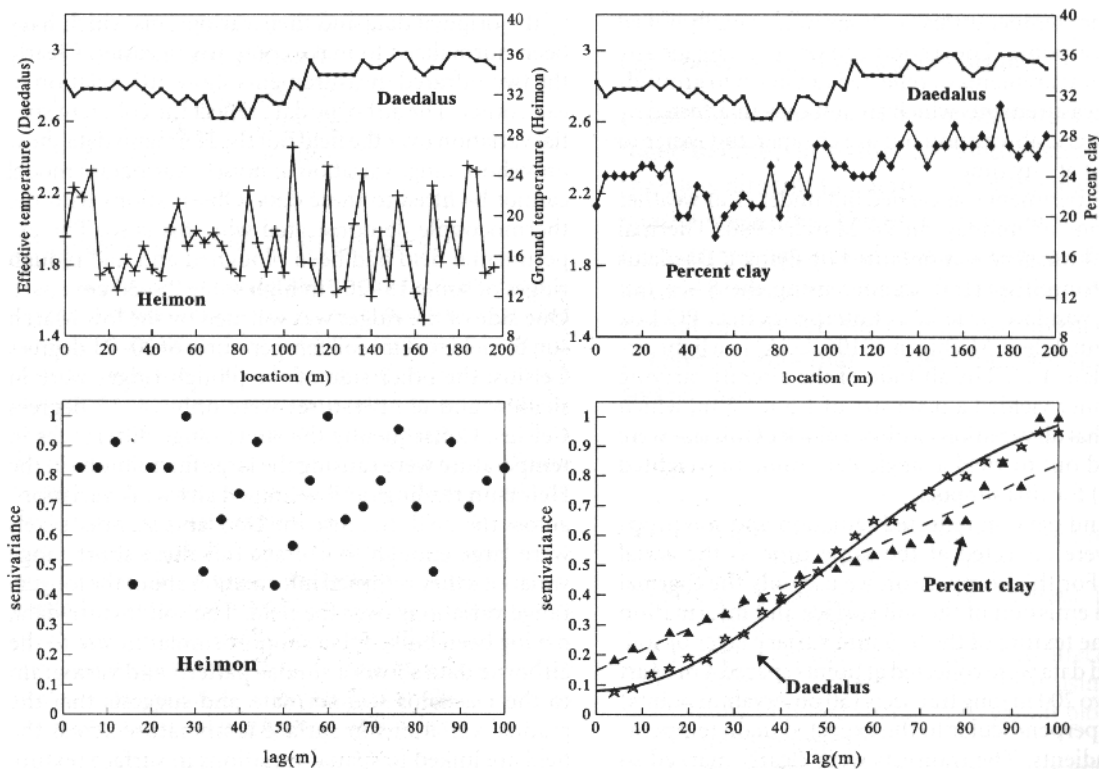


圖 10.1 (原文 p.244)

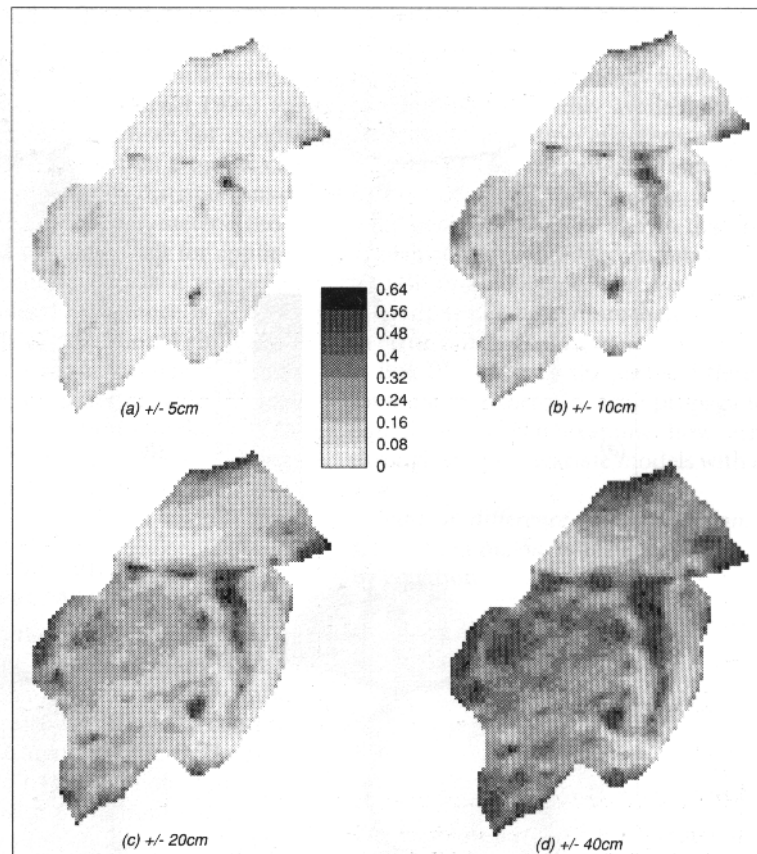


圖 10.2 (原文 p.245)

### 錯誤擴增理論(1)：蒙第卡羅模擬(Error propagation theory (1)：Monte Carlo Simulation)

為了得到我們必須(1)忽略資料統計可能出現的誤差及(2)我們使用較普通的統計法或幾何統計法，隨機模擬或回溯確認獨立的資料以平均掉誤差量，如何使用此資訊量化或減低不精確性，可以得到較精確之數值化模型。簡單化來說，假使一個新的屬性  $U$  定義為輸入  $A_1, A_2, \dots, A_n$  之功能作用，我們想知道與有關聯於  $U$  之誤差是什麼，以及從每個  $A_n$  所貢獻到誤差是什麼？如果我們可以解決此問題我們便可以將 pedigree 歸類於模型的結果並且比較此結果的信賴度。

最簡單的，而非電腦密集式處理的，誤差擴增為處理每個像素或網格屬性使其有 Gaussian (常態) 機率分布函數(probability distribution function, PDF) 為我們已知的平均值  $\mu$  及變方  $\sigma^2$ 。在這最簡單的例子中我們可以使用單一的 PDF 運用於所有的網格中，並且我們假設其為不變的。假使可以得到較豐富的連續性空間資訊，我們可以使用狀態模擬去估計特別網格的 PDFs，並且可以反映已知資料點的位置，及其屬性的關聯性空間結構(第六章)。

算數上的運算以導出新的資料是使用一個由 PDFs 得到的值運用到所有網格上，而並非是使用一個平均值，處理 PDFs 內的變方，重複許多次的計算得到的(至少 100 次)，以求獲得每個像素或網格之平均(mean)結果及其標準偏差(standard deviation)，此技術為眾所皆知的蒙地卡羅方法，此乃由於其運用了機率。

雖然蒙地卡羅方法為電腦增強之結果(一般所知的 "brute force" 技術)，其提供較有趣的資訊是有關於資料中所有可能出現的誤差如何影響到不同地理區域的部分運算的結果。此方

法於第八章有談到如何計算衍生出網格 DEMs，像是坡度。高度資料常常伴隨著特別的 RMS(root mean squared error)，為假設所有區域均為均勻的。不同等級的 RMS 誤差對計算坡度 DEM 之影響以及污水排水網位相關係可以由下列說明調查得來。平均值  $\mu = 0$  及變方  $\sigma^2$  之誤差為其標準偏差，我們稱為  $\pm 1.0$  公尺。這被加入原始 DEM 中並且其坡度是由所選擇的演算法所完成。此結果被儲存著，並且此計算重複 100 次，每個時間產生出不同等出的坡度，將此 100 次之結果平均產生出一條平均坡度圖，並且每個網格有其標準偏差圖，由平均坡度圖區分此標準偏差圖產生出地圖之相對誤差。其他變方值(values of  $\sigma^2$ )可以重複此程序。

圖 10.2 顯示出 Catsop 集水區(附錄 3)之 RMS 誤差為分別為 5, 10, 20 和 40 公分，非常明顯的顯示出其對較低起伏之相對誤差有非常強大的影響，意味著得到明顯區隔之結果是非常困難的。

不僅僅是坡度如此，其他衍生出如地區排水網路及上游集水區之貢獻對於海拔高度之誤差是非常敏感的。圖 10.3 顯示 RMS 之誤差如何影響到 DEM 內上游貢獻區域的大小之試驗結果，這只不過是以一個已得到的 RMS 誤差去估算誤差之程序，並將此誤差加入 DEM 中，並以一般之運算程序如 D8 運算法(第八章)去估算區域排水網路及上游貢獻區。重複模擬 100 次並將所得到之 RMS 誤差平均，此可以明顯的顯示上游貢獻區對其海拔高度所產生的誤差是如何敏感，在此區的部分地區其集水線是非常集中的，而其他部分之集水區域則是非常分散的。注意到此例子有小的 RMS 誤差為  $\pm 5$  或 10 公分(在耕犁中此種誤差與表面的崎嶇不平比較之下是可以忽略的)，可以影響較小起伏地區之大部分可能的流域面積。

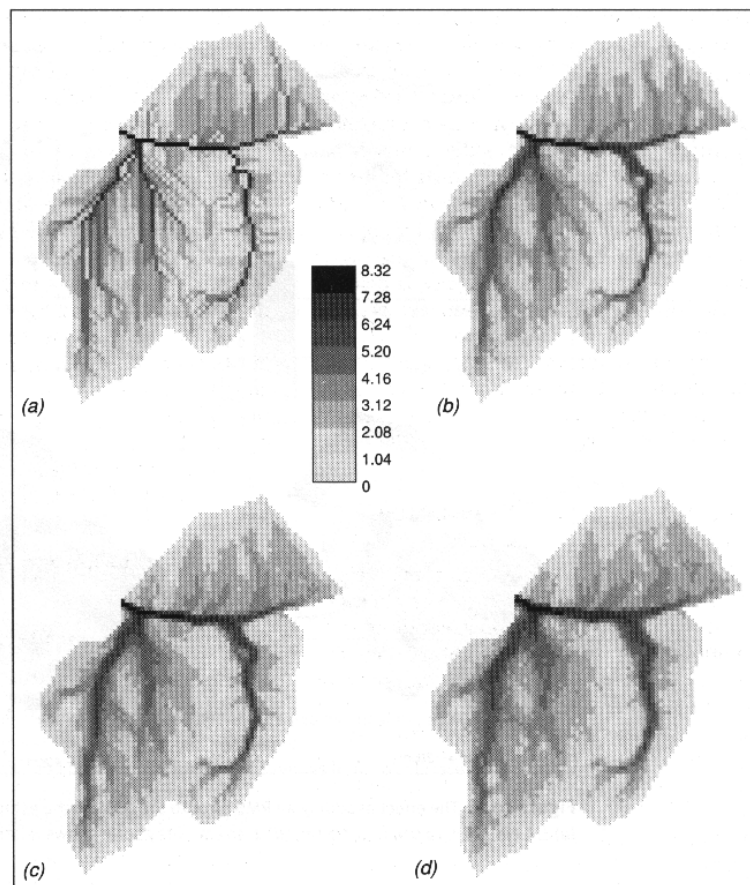


圖 10.3 (原文 p.245)



此類資訊主要用意是要指出數字表面流動模型之等級及確認(Desmet 1997)，如果小的相關誤差強烈影響到運算出之河流水道位置，並且使這些個別河流模型標準均一化，以從此較大可能範圍外取其最有效率最好之結果。標準化模型?? 假使標準化測量是由機率較高之位置得來，而此位置是與已知的上游貢獻區域相關的水流強度  $F$  得來，那麼此種測量方法是較有彈性的。如果此標準化測量是由機率較低之區域得來，那麼其預測與標準化測量趨於一致之機會將自然而然較小。這正解釋了為什麼在水文學上基於集水區之出水口之水流測量模擬是如此的 robust，此時隨意選擇集水區內任意地點作標準化測量時會造成一定的誤差。

明確的模型運用了簡單的數字取代機率分布可以產生未知強度的誤差，特別是相關的誤差較大時以及我們所未知的，假使此資料庫內所有單一數字呈現典型的或平均的狀態或者從平均值算起較大或較小之偏差之時候。這正是為什麼在水文學中隨機模式(stochastical modelling)中最近趨勢及環境科學之其他區域(De Roo et al. 1992; Fisher 1991; Journel 1996)朝向使用蒙地卡羅方法以探究全區域的輸出，由於輸入資料與模擬參數之機率分布之結果而形成模型(這些模型中的數字被應用於特定區域: 回歸模式之係數即為簡單之例子)。試驗為證明狀態模擬之方法，有助於有關於地質學、土壤、土地使用或其他控制變數之”軟性(soft)”資訊，提供最好的輸入方式於蒙地卡羅模擬中(Biekens 1994, Biekens and Burrough 1993a, 1993b, Gómez-Hernández and Srivastava 1990, Journel 1996)。

## 誤差擴增理論(2)：誤差擴增之分析方法(Error Propagation Theory (2): Analytical Approaches to Error Propagation)

雖然蒙地卡羅方法對於誤差分析較為簡單，並且可以適用於各種數值模型中，即使今日它們依舊需要大量電腦資源。例如使用蒙地卡羅方法(使用或不使用狀態模擬)將地上鑿孔資料繪圖於厚厚一層的 sub-surface 岩石層上相關誤差之計算，需要有 100 次模擬至此層頂端之厚度，以及 100 次模擬至此層底端之厚度，以及 10000 次計算所有可能配對之差異。此或許需要較簡單之方法以快速計算，很幸運的，使用於 GIS 上之許多計算現存實體或網格資料為新屬性之各種數值方法，可以使用標準統計理論計算誤差擴增(eg. See IParrat 1961, Taylor 1982)。

單一實體或網格之數值操作為我們所知的”點分析(point analysis)”，沒有包含空間的交互作用像是緩衝區(buffering)，位相關係(topological)或視窗操作(windowing operations)。在轉換演算僅僅包括算數關係時(+, -, ×, ÷, raising to powers, 取幕, 等等)，估計每個實體或網格之輸出誤差值  $U$  可被當作為輸入值  $A_i$  誤差之函數。

$$U = f(A_i) \quad 10.1$$

考慮在地圖上  $A_i$  之屬性值位置並不精確而是有一個伴隨著的誤差稱為  $\delta$ ，所以此屬性值沒有比以  $A \pm \delta a_i$  表示更好， $A_i$  可能為土壤繪圖單位上較易取得的含水數值，此土壤假設為均質狀態。我們希望將較易取得之土壤含水量結合於有效灌溉估計  $A_j$ ，若加入其誤差則成為  $A_j \pm \delta a_j$ 。假使屬性  $A_i$  與  $A_j$  獨立統計，並且  $\delta a_i$  與  $\delta a_j$  為分別為  $A_i$  與  $A_j$  之 20%，在計算  $U = (A_i + A_j)$  時所有含水誤差可以計算為  $\delta u$  為其 28%。製圖輪廓操作包含超過兩個步驟，誤差之增加可能 be explosive。

Box 10.1 顯示誤差擴增之簡單理論之偏微分方程式(partial differential equation)。以此方程式可以解釋誤差擴增如何通過簡單 bivariate 模式， $a_1 = 10 \pm 1$  以及  $a_2 = 8 \pm 1$ 。

### Sum or difference—no correlation

相加總和(sum)或者相減差距(difference)—沒有相關性。令  $u = a_i \pm a_j \pm \dots$ ，使得  $\delta u / \delta a_1 = 1$ ， $\delta u / \delta a_2 = \pm 1$ 。

藉由方程式

$$(10.4.2) S_u = \sqrt{(Sa_1^2 + Sa_2^2)} \quad 10.2$$

所以

$$u = 10 + 8 = 18$$

以及

$$S_u = \sqrt{(1+1)} = 1.414$$

$u$  的絕對誤差比  $a_1$  或  $a_2$  大，但於此之外的例子中，相對誤差(1.414 / 18 = 8%) 比原始變數(10 或 12%)小。就扣除(減法)方面來看，絕對誤差  $S_u$  是相同的，但其相對誤差則較大(1.414 / 2 = 70%)。然而加入兩個隨機數目，兩個地圖，可以認定為有利於有關誤差擴增之操作，扣除可以導致相對誤差的增加，特別是在  $a_1$  及  $a_2$  有著類似的數值時。

當  $a_2$  為常數時，例如  $u = a_1 + \text{常數}$ ，變數  $u$  與  $a_1$  之間則無差異，加或減一常數對於誤差皆無不良影響。

Box 10.1 (原文 p.248)

**Addition of correlated variables.**

增加相關變數：當變數  $a_1, a_2, \dots$  有相關性時，則方程式 (10.4.4) 必須包含於誤差  $u$  之計算，令  $u = a_1 + a_2$ ， $ra_1a_2$  在  $a_1$  與  $a_2$  間呈現出其相關性(-1 < r < 1)。

$$Su = \sqrt{\{Sa_1^2 + Sa_2^2 + 2Sa_1 \cdot Sa_2 \cdot ra_1a_2\}} \quad 10.3$$

以及

$$Su = \sqrt{\{1+1+2 \cdot 1 \cdot 1 \cdot ra_1a_2\}}$$

假使  $a_i$  與  $a_j$  呈現 100% 正相關，在  $u$  內之誤差可能也一樣多，但是不會多於  $a_i$  與  $a_j$  之誤差之和。假使  $a_i$  與  $a_j$  為負相關，在  $u$  及  $Su$  內之誤差則比個別的  $a_i$  與  $a_j$  之誤差小。

**Product or quotient---no correlation.**

乘積或商數---沒有相關。令

$$u = a_1^c \cdot a_2^d \quad 10.4$$

其中  $c$  和  $d$  假設為常數，則

$$\delta u / \delta a_1 = ca_1^{(c-1)} \cdot a_2^d \quad \text{以及}$$

$$\delta u / \delta a_2 = da_1^c \cdot a_2^{(d-1)}$$

所以方程式(B10.1.4)成為

$$Su = \sqrt{\{c^2 \cdot a_1^{2(c-1)} \cdot a_2^{2d} \cdot Sa_1^2 + d^2 \cdot a_1^{2c} \cdot a_2^{2(d-1)} \cdot Sa_2^2\}} \quad 10.5$$

因此假使

$$u = a_1 \cdot a_2$$

然後使得

$$u = 8 \cdot 10 = 80$$

以及

$$\begin{aligned}
 Su &= \sqrt{(a_2^2 \cdot Sa_1^2 + a_1^2 \cdot Sa_2^2)} \\
 &= \sqrt{(64 \cdot 1 + 100 \cdot 1)} \\
 &= \sqrt{164} \\
 &= 12.8
 \end{aligned}$$

相乘不僅會增加絕對誤差，亦會增加絕對誤差，在此例子中增加到  $12.8 / 80 = 16\%$ 。

當  $a_j$  為常數  $c$  時，例如  $u = a_1 \cdot c$ ，誤差擴增減少為：

$$Su = \sqrt{(c^2 \cdot Sa_1^2)} \quad 10.6$$

Raising to powers

$$\text{當 } u = Ca_1^c \quad 10.7$$

$C$  和  $c$  為常數，注意到  $a_1$  完全與其本身相關，所以其  $u$  的誤差  $Su$  為

$$Su = \sqrt{(C^2 \cdot c^2 \cdot a_1^{2(c-1)} \cdot Sa_1^2)} \quad 10.8$$

其中在  $u = a_1^2$  中  $a = 10 \pm 1$ ，

$$u = 10^2 = 100$$

以及

$$\begin{aligned}
 Su &= \sqrt{\{(2a_1)^2 \cdot Sa_1^2\}} \\
 &= \sqrt{(20^2 \cdot 1)} \\
 &= 20
 \end{aligned}$$

不僅絕對誤差增加，相對誤差亦加倍 ( $= 20 / 100 = 20\%$ )。

### Logarithmic and other relations.

對數及其他關係

$$\text{令 } u = C \ln a_i \quad 10.9$$

使得

$$\delta u / \delta a_i = C / a_i$$

所以

$$\begin{aligned}
 Su &= \sqrt{\{(C^2 / a_i^2) \cdot Sa_i^2\}} \\
 &= C \cdot Sa_i / a_i \quad 10.10
 \end{aligned}$$

方程式(10.24)顯示增加或減小誤差僅僅有賴於  $C : a_i$  之比率。

假使

$$u = C \sin a_i$$

使得

$$Su = C \cdot Sa_i \cdot \cos a_i \quad 10.11$$

其中  $Sa_i$  與  $a_i$  為弧度。

### 誤差擴增之簡單例子(Simple Examples of Error Propagation)

估計小麥淨值時之誤差擴增(Estimating error propagation in net returns for wheat)

種植小麥農民可能會希望估計其小麥田相關而無法確知之淨利，以知道其產量其花費管理，以及其種植後各時期之收穫，並且會接受到無法確知之價格。每塊田地平均每公頃有多少公噸產量，農民希望在其預測中對其誤差之評價

$$\text{Net value } (N) = \text{Yield } (Y) \times \text{price } (P) - \text{cost } (C) \quad 10.12$$

每一塊田地令其  $Y$  為  $6 \pm 2t \cdot ha^{-1}$ ，每一公噸為  $P:100 \pm 10$  貨幣單位，以及  $C:40 \pm 20$  貨幣單位，由於農民無相關之資料，因此每塊田地之花費假設為相同。

總量為

$$\begin{aligned} G &= Y * P \\ &= 6 * 10 \\ &= 600 \text{ 貨幣單位} \end{aligned} \quad 10.13$$

無法確知之總量為

$$\begin{aligned} S_G &= \sqrt{P^2 \cdot S_Y^2 + Y^2 \cdot S_P^2} \\ &= \sqrt{10000 * 4 + 36 * 100} \\ &= 116.62 \text{ 貨幣單位} \end{aligned} \quad 10.14$$

其淨值為

$$\begin{aligned} N &= G - C \\ &= 600 - 40 \\ &= 560 \text{ 貨幣單位} \end{aligned} \quad 10.15$$

而其未確定部分為

$$\begin{aligned} S_N &= \sqrt{S_G^2 + S_C^2} \\ &= \sqrt{13600 + 400} \\ &= 118.32 \text{ 貨幣單位} \end{aligned} \quad 10.16$$

### 在國際土壤流失方程式中之誤差(The error in the Universal Soil Loss Equation)

考慮到誤差擴增對第七章中其中一個模型試驗之影響，此模型為模擬肯亞 Kisii 地區土壤流失之研究，此研究之目的是為了估計土壤之沖蝕量，此地區為已經耕作 40 年玉米之田地。使用國際土壤流失方程式估算土壤侵蝕流失量(USLE---Wischmeier and Smith 1978)：

$$A = R * K * L * S * C * P \quad 10.17$$

$A$  為每年每公頃土壤流失量， $R$  為降雨侵蝕量， $K$  為土壤侵蝕力， $L$  為坡度公尺長， $S$  為坡度百分率， $C$  為耕作參數以及  $P$  為保護參數。

資料來源受限於降雨量研究及傳統地誌圖只有顯示土壤之描寫 'representative' 值與氣候變數。此外， $R$ 、 $L$  及  $S$  因子是他們自己使用許多回歸方程式導出。

**The R factor. R 因子。**  $R$  值是由 FAO 公式估計得出：

$$R = 0.11abc + 66 \quad 10.18$$

$a$  為每年平均降雨量(公分)， $b$  是在這兩年內之最大日降雨量(公分)， $c$  是兩年當中之最大陣雨總降雨量(公分)，在此研究地區所能取得之最好之  $a$ 、 $b$  及  $c$  資料(Wielemaker and Boxem 1982)為：

$$a = 172.5 \pm 20 \text{ 公分}, b = 5.41 \pm 1.1 \text{ 公分}, c = 2.25 \pm 0.5 \text{ 公分}$$

由於此地區有限的資料量之限制，使得其標準偏差較大。

公式 10.18 中， $R = 297$  公分，Box 10.1 中， $R$  之標準偏差為  $\pm 72$  公分/年。

**The K factor. K 因子。**在此研究地區外之磚紅壤沖蝕地區，有較多沖蝕數據建議其  $K$  值為 0.1 較為可能。土壤特性常常有較大的變化性(a . 50 % 的 CV 是常見的)，一個地圖單位標準偏差為  $\pm 0.05$  是不會不合理的。

**The L factor. L 因子。**坡長因子  $L$  是利用 Wischmeier 及 Smith (1978) 的公式計算得來：

$$L = (l/221)^{1/2} \quad 10.19$$

$l$  為斜坡長(公尺)，考慮到在此地區之典型的長坡，長度為 100 公尺  $\pm$  20 公尺是合理的，轉換到  $L$  值即為  $2.13 \pm 0.045$ 。

**The S factor. S 因子。**坡度因子  $S$  為使用拋物線迴歸方程式計算得到：

$$S = 0.0065s^2 + 0.0454s + 0.065 \quad 10.20$$

$s$  為坡度，以百分率(%)表示(Smith and Wilschmeier 1957)。當其坡度為  $10 \pm 2\%$  時方程式 6.57 及 6.47 產生一個  $S$  值為  $1.169 \pm 0.122$ 。

**The C factor. C 因子。**作物管理因子  $C$  之估計值為 0.63，給定這易變的估計  $C$  值至作物中並不能含括整年之土壤表面，並且 Rose (1975) 發現玉米作物之  $C$  值之變動範圍介於 0.4 與 0.9 之間，其誤差估計為  $\pm 0.15$ 。

**The P factor. P 因子。**沖蝕控制常規因子  $P$  估計為  $0.5 \pm 0.1$ 。

綜合上述所有的 USLE 因子值及其誤差值：

$$P = 297 \pm 72$$

$$K = 0.1 \pm 0.05$$

$$L = 2.13 \pm 0.045$$

$$S = 1.169 \pm 0.122$$

$$C = 0.63 \pm 0.15$$

$$P = 0.5 \pm 0.1$$

上述產生出每年土壤流失率為  $23 \pm 14.8$  公噸/公頃/年，相當於土壤表面降低了  $0.23 \pm 0.15$  公分/年，或者 40 年內降低了  $9 \pm 6$  公分。換句話說，給定完善的模型預測土壤之變化，95 % 的網格有氣候/坡度/特定的土壤區域，使土壤流失範圍介於 3 ~ 21 公分。這 95 % 的信賴區間限制  $A$  (公噸/公頃/年) 之範圍為  $-7.0 < 23.0 < 53$  (在  $r = 0$  時)， $-17.0 < 23.0 < 63$  (在  $r = 0.25$  時)，這顯示了小的正的相互關連關係對於誤差帶有不利的影響。

### 推薦之算數演算法 Recommended strategies for arithmetic algorithms

下列理論及例子提供降低誤差擴增之準則

1. 避免變數間互相關聯
2. 盡可能增加地點
3. 如果無法增加，則相乘或相除
4. 盡可能避免差距太遠，或增加變數。

對於自 GIS 中取得之環境程序的飼育模型資料，這些準則對其結果之品質有著相當大的涵義。他們證明了，為了減小點模型輸出之誤差，減少輸入之誤差是重要的。在空間模型中，輸入之誤差可以經由下列方法減少：

- (a) 使用適當的內插技術

- (b) 使用適合的取樣密度
  - (c) 移除或檢視離開本體之部分、子群部分、偏斜部分....等等
  - (d) 採取適當的分級
  - (e) 使用較敏感的模型
- 降低模擬參數的誤差以改進模型等級。

### 誤差擴增工具—(ADAM Error Propagation Tools—ADAM)

由以上所提供之由兩三個參數組合而成之單一程式可以很容易以口袋型電算機計算，但是如果必須從資料庫中計算大量網格或實體資料時則會顯得繁複，或者當數值模型包含了大量的屬性時亦如此，包括關聯性資料。最近，Gerard Heuvelink 及 Cees Wesseling (兩者位於 Utrecht 大學) 撰寫了稱為 ADAM 之電腦程式，此程式可以經由點模式(point mode) 中之複雜的數值模型追溯出多重網格圖中之實體屬性(Heuvelink 1993, Heuvelink et al. 1989)。

多重迴歸中  $a_i$  值為輸入網格地圖之屬性， $b_i$  值為係數。

$$u = b_0 + b_1 a_1 + b_2 a_2 + \dots + b_n a_n \quad 10.21$$

貢獻到輸出地圖  $u$  中之誤差來自此模式中之係數  $b_i$  之誤差以及空間變數之誤差與  $a_i$  輸入屬性之測量誤差(nugget)。

### 與模型係數有關之誤差(Error associated with the model coefficients)

判定與模型有關之誤差不適容易的，但在多重迴歸模型中誤差是由係數  $b_i$  之標準偏差而來，所有關係必須由好套裝統計提供。

### 由空間變數而來之誤差(Error accruing from spatial variation)

此方法可以由多種方法導出，包括直覺以及以傳統及空間統計為基礎之取樣，假使沒有 hard 資料，憑直覺估計屬性值之標準偏差可以提供誤差估計。對空間實體如土壤或植物多邊形取樣時可以提供在估計屬性及其相關屬性之標準偏差時一些有用的資訊。假使有充足之資料以使用像克利金(kriging)或狀態模擬(conditional simulation)方法對網格疊圖進行內插，而其網格之標準偏差正是我們所要找的不確定性之來源。整合克利金法與狀態模擬兩者方法(Deutsch and Journel 1992)可以利用於形成相關表面(correlation surfaces)，儘管假使資料是正確均勻分布於區域中，在不同屬性  $r_{ij}$  之間之簡單相關係數為我們所要的。

### 誤差之平衡(The balance of errors)

於輸出圖  $u$  的誤差  $\delta u$  是由所有來源得來，為它的係數或變方，這有助於確定哪個是最主要不確定來源。如果需要，這些是可以改良模型或資料收集。不同模型之誤差比較分析(eg. 單一迴歸對多重迴歸)可以允許使用者決定是否在第二屬性收集資料，是符合時間及經濟效益。比較模型與資料間之誤差貢獻，提供了模型分級及製圖有用且適當的調查方法。

**The theory behind ADAM.** 在 ADAM 背後的理論。上述公式(10.35)可以歸納為：

$$y = g(z_1, z_2, \dots, z_n) \quad 10.22$$

$g$  為從  $R^n$  到  $R$  之連續微分方程式， $g$  的自變量  $z_i$  由與實體或網格相關之  $A_i$  屬性及其模型係數  $b_i$  組成。此處之問題為如何由  $z_i$  之誤差算出  $y$  之誤差，我們假設  $y$  及  $z_i$  為隨機變數  $Y$  及

Z 之真值，我們定義向量  $\mu = (\mu_1, \mu_2, \dots, \mu_n)^T$  以及  $Z = (Z_1, Z_2, \dots, Z_n)^T$ 。

計算所關注的 Y 之平均及變方數量之最簡單的方法，為使用 g around  $\mu$  之泰勒展開式 (Taylor Series expansion)，忽略 (higher-order) 方面 (詳細請見 Heuvelink et al. 1989)。

當在參數 Z 為 0 時，Y 的變方為：

$$\delta^2 = \sum_{i=1}^n \left\{ \delta_i^2 \left( \delta_g / \delta_{Z_i} \cdot (\mu) \right)^2 \right\}$$

這意味著 Y 的變方為各部分的總和，每個貢獻一個 Zi，這分開部分的特性允許我們分析每個輸入部分對最後誤差之貢獻度，為迴歸係數或變量。

ADAM 使用此種技術應用於每個實體之誤差擴增計算，在網格地圖之例子中為格子之像素，但是在其他方面也可以為多邊形或向量形式之實體 (Wesseling and Heuvelink 1991)。迴歸模型係數及其誤差可以由取樣資料及相同之點取樣算出，此可以使用建立網格以輸入網格數及其誤差之地圖，也可以是從所有係數及所輸入之變方求得之誤差貢獻的空間變異的地圖 (圖 10.4)。

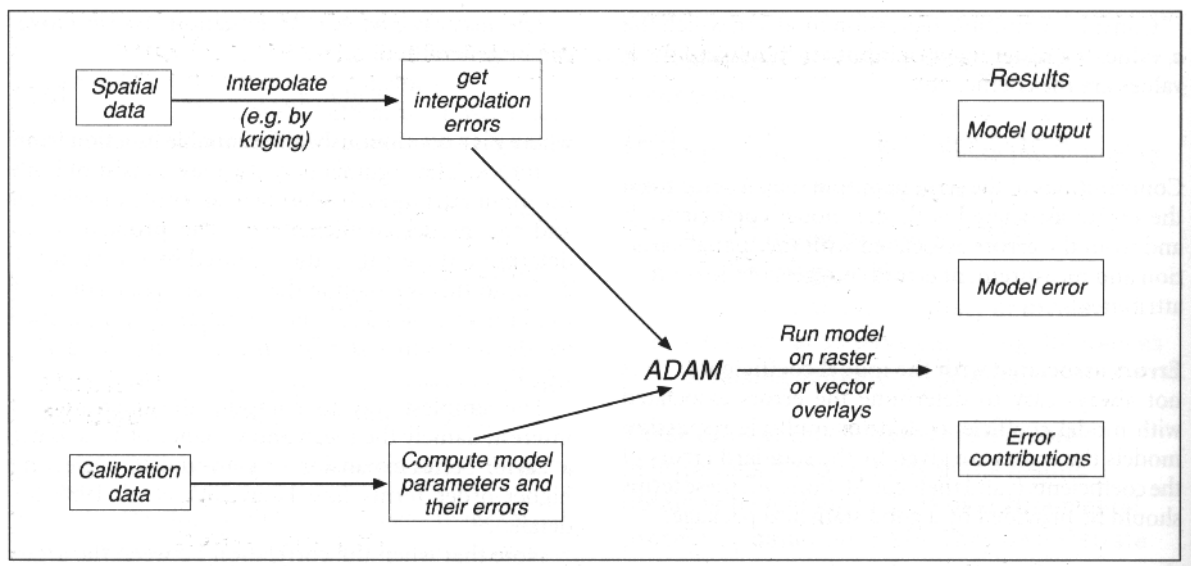


圖 10.4 (原文 p.252)

### 在迴歸模型中利用 ADAM 方法計算誤差擴增 (An Example of Computing Error Propagation in Regression Modelling With ADAM)

第 5 及第 6 章介紹部分的數值方法以預測荷蘭南方 Maas 河之氾濫區內重金屬污染之濃度。此方法包括基於重金屬濃度、至河之距離及相對海拔高度及以平常的克利金法或整合克利金法直接進行地理統計內插的多重迴歸分析模式。

雖然許多作者 (Lam 1983, Laslett et al. 1987, Leenaers et al. 1990, Englund 1990) 已經比較不同內插技術之功能，但是在成本—效益方面較少涉及。我們使用 ADAM 誤差擴增工具及獨立確認取樣以比較不同內插技術，為了方便了解哪個是較符合經濟效益的但是並不是適當的內插法。在這個例子中我們使用了與第 5、6 章相同的資料組 (但是是來自稍大之區域之資料)。此時這完整之 155 組土壤樣本被分為兩組，一組為 53 個樣本隨意取出作為對照，剩下 102 個樣本作內插分析。比較兩組之統計分析，結果顯示在這地區不同之取樣單位無顯著差異，對照資料提供較好之預測效果。

初步之分析(第 5 章)顯示鋅含量與'至河之距離'及'海拔高度'屬性有強烈的關聯性。因為鋅含量與'至河之距離'兩者之間有著顯著之對數--常態(log-normal)分布，它們是利用自然對數轉換且所有的統計分析是利用 log-normal 轉換資料，對於 ln(鋅)與 ln(距離)[-0.768]及 ln(鋅)與相對海拔高度進行相關性分析顯示兩者之間呈負相關，ln(距離)與相對海拔高度則呈正相關。此多重迴歸為

$$\ln(\text{zinc}) = B_0 + B_1 \cdot \ln(\text{distance}) + B_2 \cdot \text{elevation} + \varepsilon \quad 10.24$$

當樣本數為 102 個時，其迴歸  $R^2 = 0.74$ ，所以顯示鋅濃度依其他兩個屬性變化而變化之模式為合理的模式。

### Procedure 程序

我們比較下列選項以繪製汜濫平原土壤鋅含量：

(1) 使用土壤型態或汜濫頻率地誌圖及計算繪製單位之平均及標準偏差，假設在每個單位內所含之變異是來自取樣位置之平均及變異。

(2) 利用雙變數迴歸方程式(10.24)繪製鋅含量地圖，由次序性克利金法繪製至河之距離及相對海拔高度，距離及海拔高度之克利金誤差地圖提供屬性誤差之來源。

(3) 直接使用次序性克利金法內插鋅含量值。

圖 10.5 介紹第 2、3 種方法之程序。

每種方法之成本有賴於觀測之數量及所使用之取樣型態，距離取樣是比較便宜的，是由於可以直接從地圖上讀取或由緩衝區指令形成，海拔高度資料可以由田間資料累積得到，主要的成本為勞力成本。土壤重金屬之取樣是昂貴的，由於必須定點取樣，並且是需要使用雙手的，所以是昂貴的。因此土壤中的鋅含量資料之分析可能是由較不足之取樣數所得出，而此量測位置屬性是較節省成本的。

觀察結果之品質如何依賴於取樣之成員(members of samples)，102 個取樣單位資料組常常建立許多較小之資料組，那即是兩組 51 個及兩組 27 個取樣數及兩組 14 個取樣數。每一組資料之收集及製作成本為使用討論中的屬性商業率(commercial rate)。

每一種方法皆使用於所有資料組以預測 20 x20 公尺網格值，相同大小之資料組之結果是 pooled，使用 ADAM 法、克利金標準誤差及由 53 個測試取樣數獨立 validation 以計算誤差擴增，此種方式是 criteria 成功的。



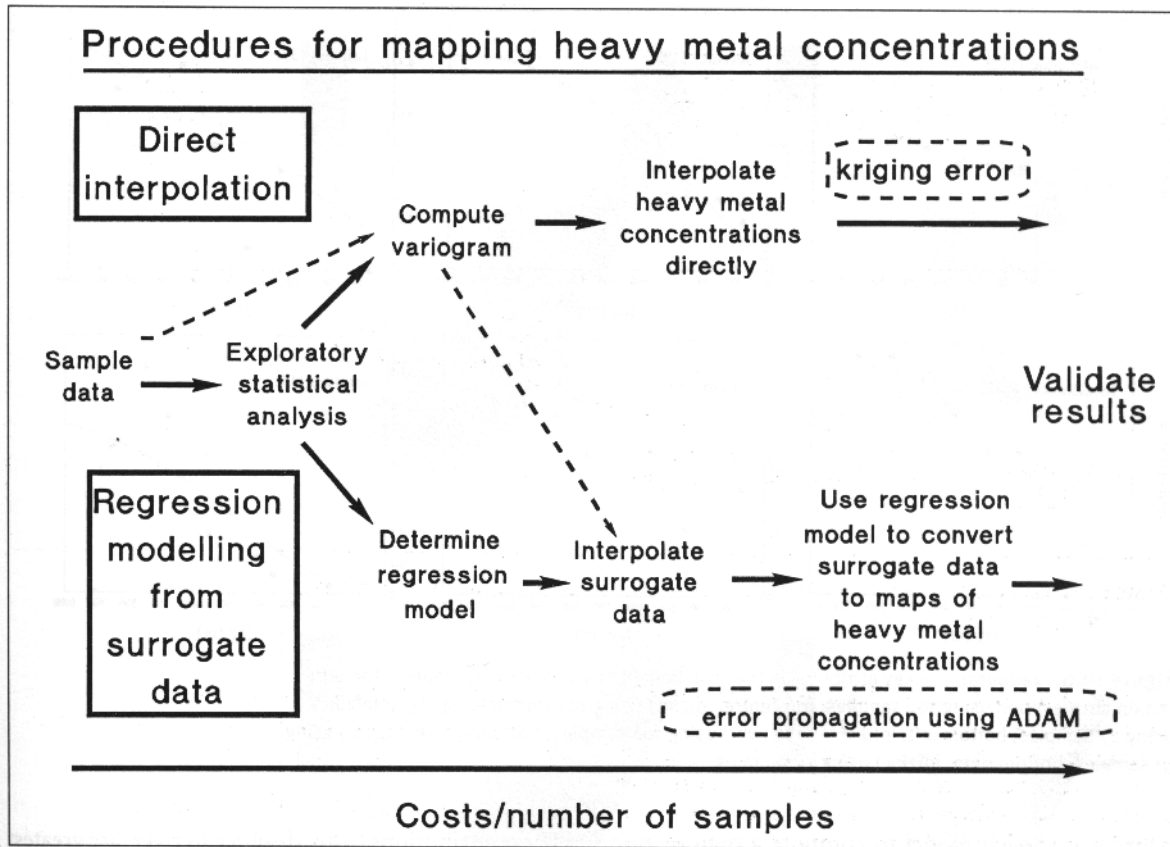


圖 10.5 (原文 p.253)

迴歸模式為假設所輸入之圖像是由 102 個位置內插得來，此兩種屬性之變異圖是由圖 10.6 得來。所有資料是 20x20 公尺方格以克利金法求得(圖 10.7)。係數及相關的迴歸模式是分別使用 14、27、51 或 102 個小資料組鋅資料求得，以提供不同程度適當範圍之模式。

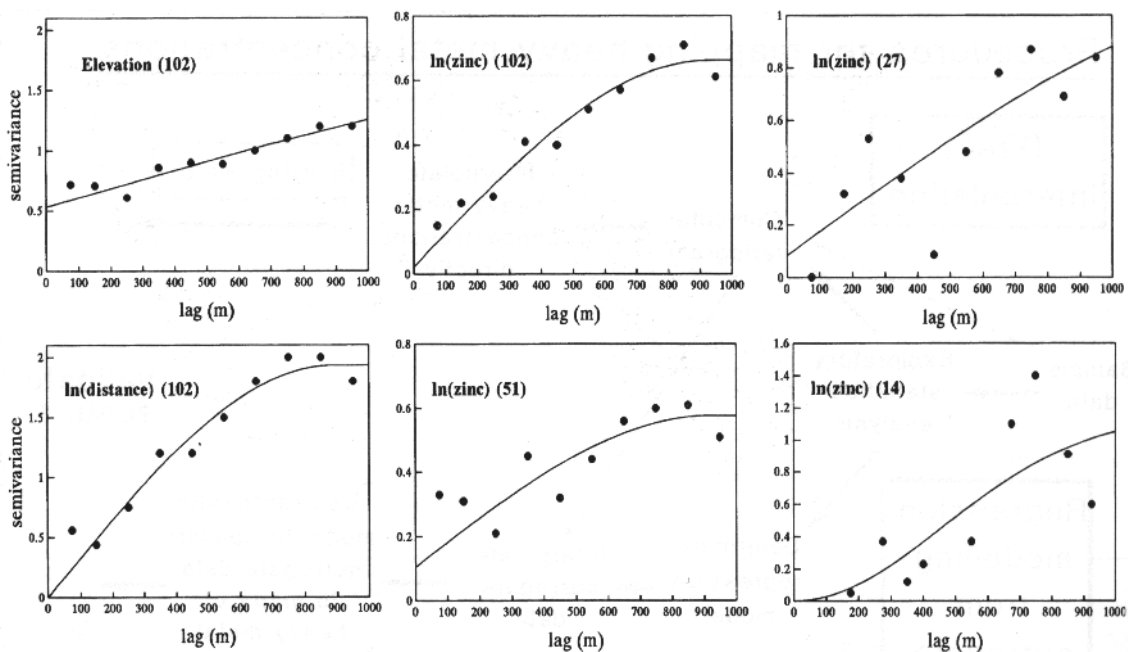


圖 10.6 (原文 p.254)

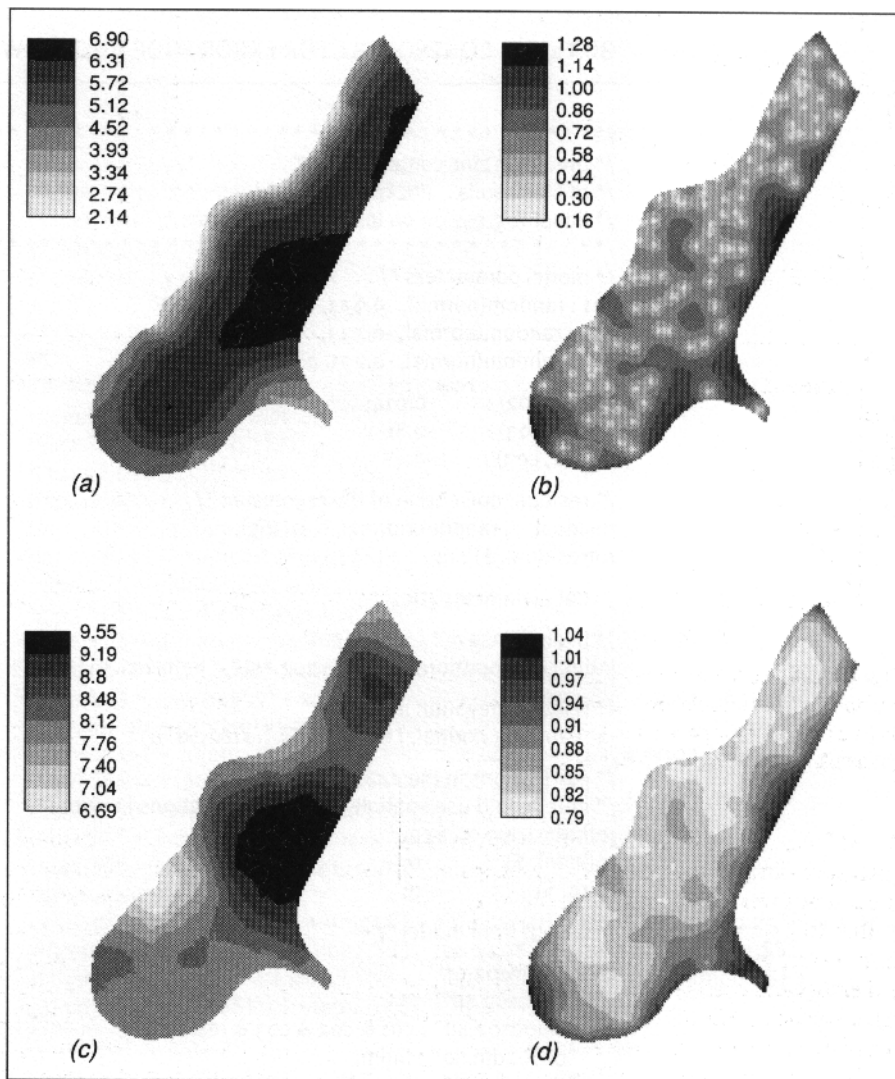


圖 10.7 (原文 p.255)

Box10.2 列出線性迴歸模型中模擬係數、標準誤差、相關性、模擬結果及其誤差表面之控制檔案。ADAM 建議使用之模型為二次泰勒展開式級數較適合。Box10.32 列出自動轉換控制檔案至 PCRaster 網格模型系統中一組命令之結果(Wesseling et al. 1996)；這顯示出嘗試以手算出誤差擴增之結果是非常困難的。

由兩組分別為 27 及 102 點之範例資料迴歸計算出其  $\ln(\text{鋅})$  結果分別顯示在圖 10.8b 及 10.8d。

Box 10.2 (原文 p.256)

### 誤差之平衡(The balance of errors)

ADAM 將誤差分配入'至河之距離'、海拔高及模擬係數屬性中，並且以連續性圖顯示出來(圖 10.9)。

這些地圖顯示網格結果與取樣位置具一致性，模型為誤差之最大來源，許多時候會達到 65 % 之誤差。誤差來自"至河之距離"之貢獻較大，這些屬性不曾被量測過，如位於此區之東邊邊緣地區。來自"海拔高"之貢獻度亦呈高峰，可能由於這些特性有較低之全面性變異，且在取樣位置上具較小之全面性突出，區域性偏差為明顯的，與對"至河之距離"取對數之誤

差比較顯現出較大之區塊。

迴歸模式估算誤差擴增是計算從圖 10.8f 及 10.8h 得來之 27 點極 102 點之樣本資料計算來的。

Box 10.3 (原文 p.257)

### 直接內插(Direct interpolation)

直接繪製 ln(鋅)圖，計算每一小組資料之變異圖以及將相同組之資料內插。在小組資料為 14 點資料時，沒有變異圖可以被模型化；而基於 27 個資料點之變異圖甚少定義---圖 10.6。

圖 10.8a 及 10.8e 與圖 10.8c 及 10.8 分別顯示由次序性克利金法分析 27 點及 102 點資料組之圖及誤差表面。圖 10.8 顯示由克利金法使用 102 個取樣點給予其空間解析度最好的組合，並且有著較低的預測誤差，但是如果是使用 27 個取樣點作分析，則明顯的每個迴規模是皆無法得到滿意的結果甚至是較差的結果。兩種迴歸模式所得出之結果差異不大，意味著將取樣數量由 27 增加至 102 點對於改良地圖幫助不大。

使用土壤型態及氾濫頻率地圖之類似的研究完成甚多，所有地圖是由計算一組 53 點資料之標準偏差  $VSD = \sqrt{\left\{ \frac{\sum (\hat{z}(x_i) - z(x_i))^2}{53} \right\}}$  得來。每個地圖是由內插或模型產生，平均預測標準偏差(prediction standard deviation PSD)是由圖 10.8e-h 之誤差表面計算得來。

### 成本與效益(Costs and benefits)

圖 10.10 顯示平均預測標準偏差(PSD)及 validation 誤差(VSD)如何隨氾濫頻率地圖、迴歸模式及鋅的次序性克利金法之成本而變化。圖 10.10 非常清楚顯示當取樣數較少時(<20)，可以提供簡單氾濫頻率地圖效果(預測標準誤差類似於 validation 誤差)，但是如果使用較多的資料以改善地圖單元之平均值及標準偏差之精確度時，並不能改善空間上之預測值。如同第 6 章顯示一樣，非常明顯的在氾濫頻率內有許多的空間 validation，這是無法由簡單的區域-平均(area-average)統計法計算得出。

使用超過 20 個取樣點資料作經驗式迴歸模式會優於使用氾濫頻率地圖(validation 結果比預測值好)，可是一旦曾經有正確的迴歸模式描述其空間趨勢，使用超過 35 點資料估計迴歸模式，對於預測或 validation 會得到較少的改良。

使用不足的資料(本例中，<50 取樣點)是無法模式化，從稀疏的資料作內插會帶來較大的內插誤差，由較低之 validation 誤差 confirmed。不足的資料通常意味著內插演算法無法對整個區域作預測，如同圖 10.8a 及 10.8e 顯示一般。當提供的資料超過 50 點，克利金內插法則顯得較適用，是由於有足夠的"hard"資料以解決在距離上之空間 validation，這是無法由總體迴歸模式解決的。我們順便注意到在 KT-universal 克利金模式中，將克利金與迴歸結合依然是較有效率之處理(Deutsch and Journel 1992)，並且適當的估計區域變異以將總體趨勢知識結合(第 6 章)。

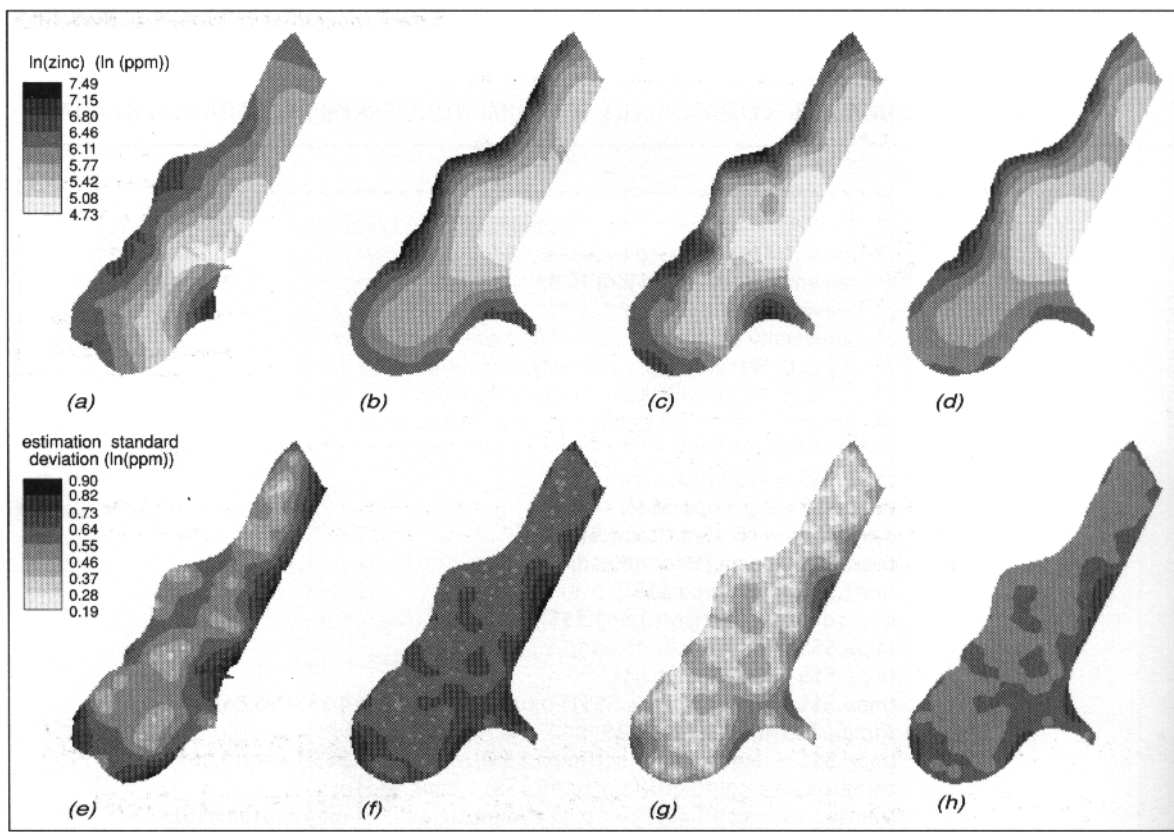


圖 10.8 (原文 p.258)

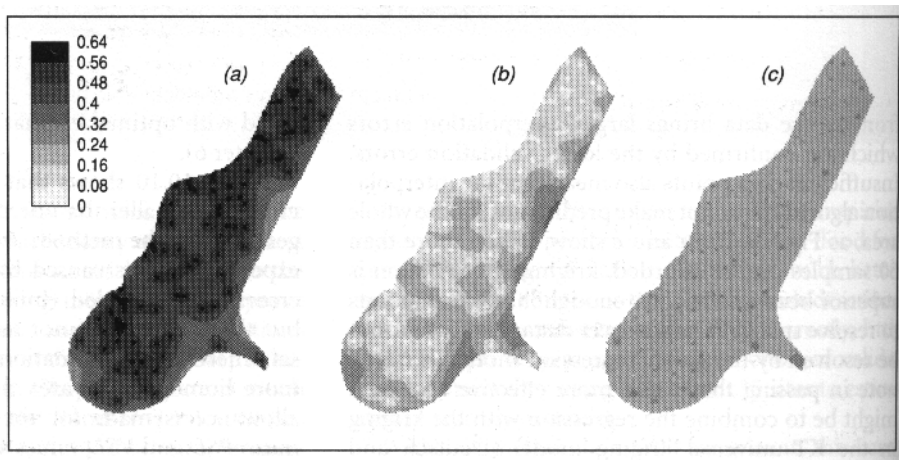


圖 10.9 (原文 p.258)

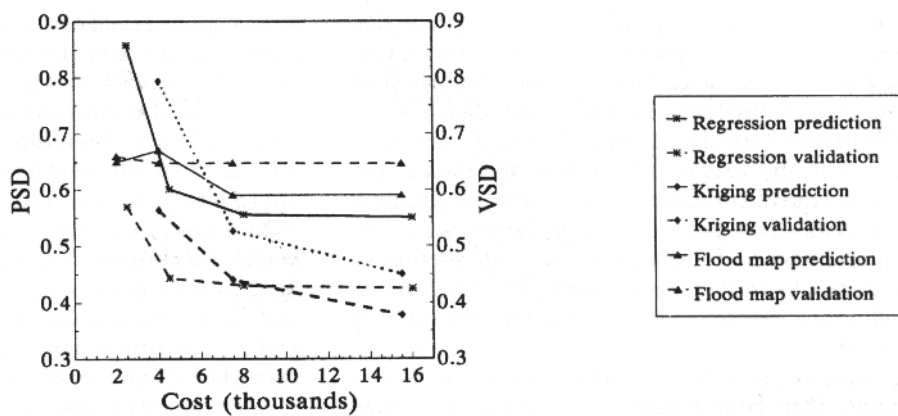


圖 10.10 (原文 p.259)

圖 10.10 顯示 validation 誤差通常比預測誤差小，所有方法顯示出比預期的好。這是由區域中未取樣部分所造成的較大的內插誤差，但此區域並不包括在 validation 組中。因此 validation 資料所取樣之空間資料比總體調查取樣自較多之均質的地區資料。假使這些未取樣區域是被允許的，那麼 PSD 及 VSD 之平均值則較為接近。

## 使用變異圖將取樣網絡適當化及降低空間解析度之成本下減少誤差 (Using the variogram to optimize sampling networks and reduce errors at the cost of reduced spatial resolution)

在本章節之最初部分，我們證明當它們轉向為自然尺度變異時，取樣網絡及取樣大小 (support) 會有較好之表現。這種由取樣原理 (sampling theorem) (由電子信號過程試驗導出) 為憑經驗的陳述，這種觀察必須由頻率及解析度產生，以符合實際現象之關聯性空間資料。假使取樣頻率及取樣大小為缺乏選擇性 (ill-chosen)，使得各種變異的形式無法被解決並會出現大量雜訊。

本章節將可學到的 The lessons to be learned

雖然所有河流及所有形式污染模型之結果不需要相同，本研究說明正確鑑別關聯性空間結構及其有關於物理程序之數值。它們證明將被繪製之屬性之關聯性空間結構之內插法及取樣密度轉換後之好處，以及顯示了使用合適的技術可以節省成本。

許多自然現象會隨不同的尺度而變 (如類似碎形之行為—Burrough 1983ab, 1993a, Mandelbrot 1982)，在第 5 章中我們指出在內插之前使用 bulking 於取樣以移除較小範圍之空間變異，第 6 章我們顯示使用區塊克利金法可以預測土地區塊比小的克利金變方要大許多。在第 6 章 (6.17)，我們也顯示區塊克利金是隨著 (1) 變異圖，(2) 區塊大小及 (3) 資料點的結構而改變，假使變異圖為已知的則我們可能可以導出區塊內的變異及在有限往格支取樣空間之量的關係。這些關係可以在有限的經費下使用於篩選最好的取樣空間以得到 PSDn。McBratney 及 Webster (1981) 為此撰寫了一個名為 OSSFIM 的電腦程式。

然而如果增加了區塊大小，減少了預測誤差，結果為，我們必須為解決土地大區塊而準備，如果我們將內插使用於數值模型中，內插表面網格大小以靈敏的符合於模型中之網格大小，這可以直接得出誤差擴增。區塊大小與誤差間之關係是重要的，當土地如何處理及其精確度如何解決。例如，當鋅污染被認為較高時，則此區之土壤必須被移除或處理。小的單位 (區塊) 是較為適合的，因為可以比大區塊較容易清除且較省錢。假使與小區塊有關之平均值之誤差較大時，有較合理的改變，降到可接受限度水準之下之區塊必須被處理並且在限度之上的區塊會遺失掉。

## 使用變異圖以最適化鋅知取樣網 (Using the variogram to optimize the zine sampling net)

對鋅資料之不同內插法之成本及效益比較是基於原始取樣形式之基礎，在所有例子中皆是如此 (除了“次—取樣”之外)。假使我們在作田間調查之前就已知道輸入資料 (海拔高、至河之距離、鋅值) 之變異圖，以及假使我們準備接受土地位置 B 之內插平均值以替代區域 (area) support 之大小，我們可以在任何程度之投資下計算出最好的取樣空間以繪製地圖。

圖 10.11(上) 顯示 ln 鋅及海拔高之常態變異圖，非常明顯的，鋅的關聯性空間結構強於

海拔高，顯示了較大之海拔高礦塊及從模式中估計半變異數得到較大的偏差。圖 10.11(中)顯示利用 OSSFIM 將兩種屬性之取樣空間對區塊大小比值圖，我們可以看到在有限的取樣數量，我們可以從鋅值及海拔高得到較好的結果，但是鋅之空間結構平均值是從測量之花費提供。所以假使我們要使用海拔高地圖以幫助我們預測鋅值，這有助於我們計算網格方格之取樣空間及區塊大小之改良。

圖 10.11(下)顯示此兩種屬性之常態克利金變異隨著取樣空間而變(成本也增加)。此證明了在此區域 50 平方公尺之海拔高取樣格及計算 100 公尺區塊平均產生了相對精確度，如同於在 150 公尺方格取樣鋅時一樣。這意味著提供足夠的鋅取樣資料可以選擇直接內插鋅值，並不計算含有海拔高模式之回歸，縱使海拔高紀錄於網格內，為如同鋅取樣網稠密。在有限的區域中增加鋅取樣密度比海拔高之效益較好。

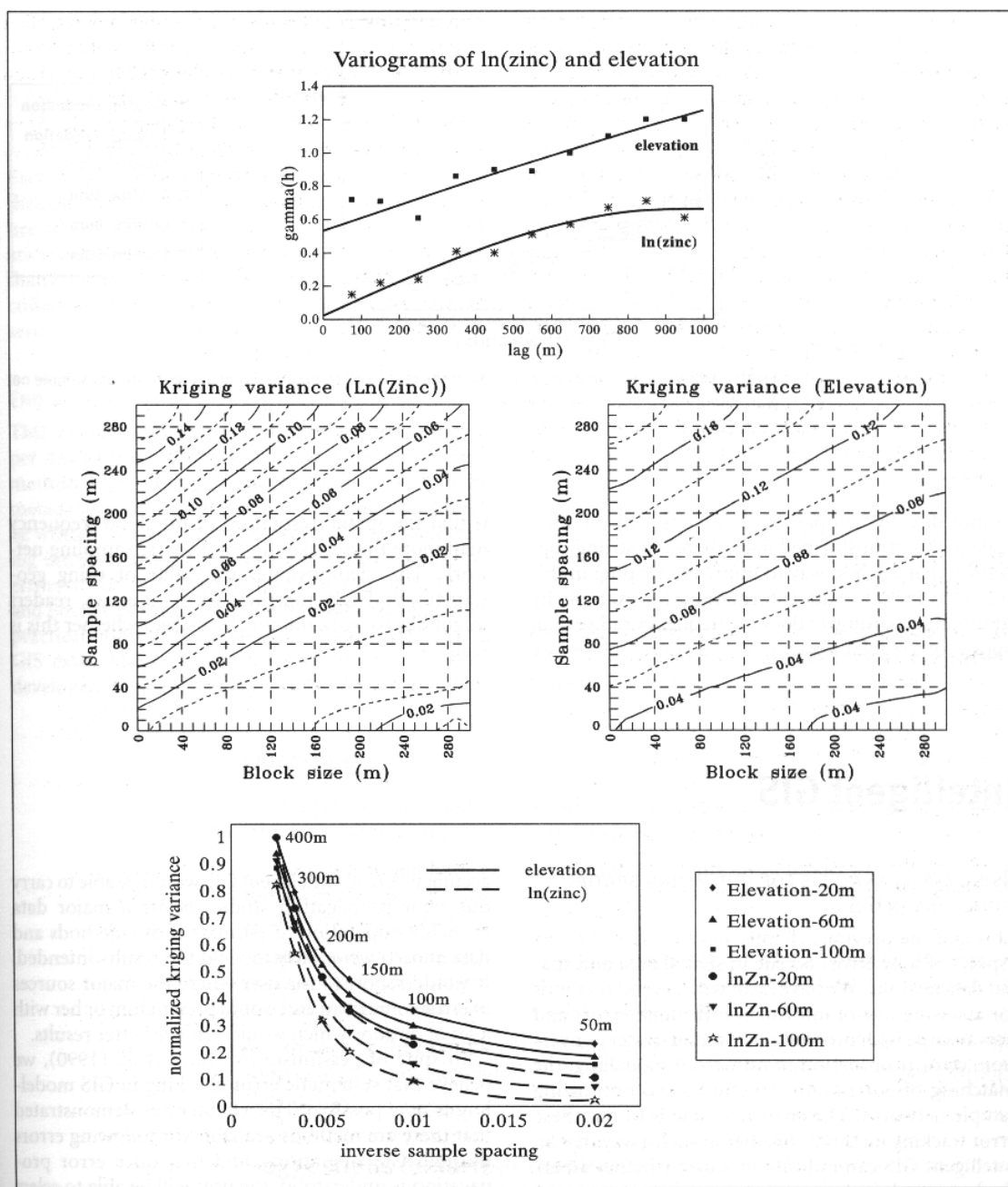


圖 10.11 (原文 p.261)

利用成本及取樣數量之知識，用於比較由原始取樣網(original sampling net)之克利金得到的 PSD 及 VSD 程度研究，由在一定範圍方格區域之不同區塊大小之最大的期望克利金標準偏差，類似於實際使用之取樣點(14、27、52、102)。這些結果顯示於圖 10.12。他們建議在此例子中對於方格最適化之取樣網與實際的點克利金內插法之 validation 比較並不會得到較好的效益。假使 validation 為未知，在稀少的方格取樣會阻礙變異圖估計，此並不會緊接於點以在短期間歇上計算半變異數。

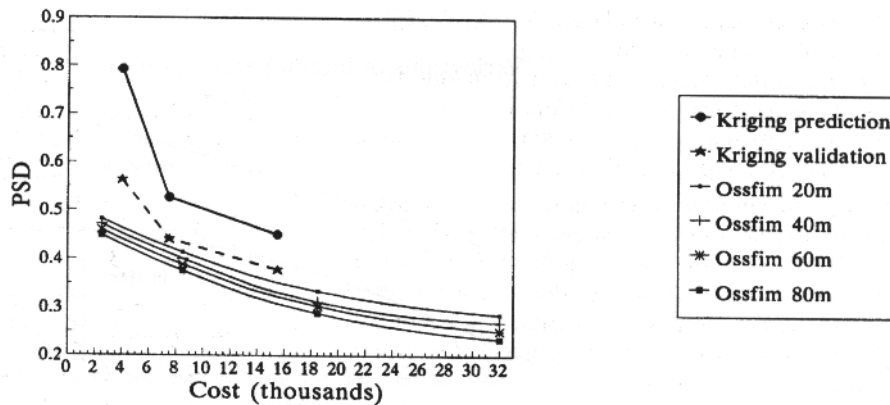


圖 10.12 (原文 p.262)

### 誤差擴增之地理統計(Geostatistics in error propagation)

地理統計之重要角色為減少 GIS 之數值模式中之誤差，地理統計可以提供未確定的內插之資料，以及可以提供適合屬性值估計給定空間區塊的大小及已知的變方。不管不同的資料是空間上的相容的，變異圖 皆可以調查。變異圖的知識可以將使用於內插地圖之標準的取樣形式幾何學適當化。藉由蒙第卡羅法與誤差擴增連結之地理統計，可以經由數值模型使用於誤差之擴增。

回到第 6 章之表 6.2，雖然我們無法直接比較未經對數轉換計算出之結果，改良此區地圖之最好的方法為將軟(soft)性資料包括進來，例如將氾濫頻率分級而不是選擇標準的方格。在這例子中，氾濫頻率地圖之空間解析度比組成稀少的取樣網絡要多，這顯示了同時使用地理統計及 GIS 之數值。不管是否真的如此所說，讀者可能希望由自己完成。

### 智慧的地理資訊系統(Intelligent GIS)

#### 使用誤差分析以使 GIS 中之空間模型最適化(Using Error Analysis for Optimizing Spatial Modelling in GIS)

本章即先前章節呈現出誤差如何在空間資料及空間資料分析中累積。我們已經提出許多方法以接受資訊經由誤差而遺失，以及如何應付向向量-網格轉換、數值化模型中之誤差擴增、關聯性結構之相配以及適合化取樣網絡之問題。現金之挑戰為將這些追蹤誤差之方法結合為一個有智慧的 GIS，此 GIS 可以導引使用者是否其特別的空間分析問題可以被解決至一定程度的品質，如果不行，亦可以提出步驟以得到正確結果。一套有智慧的 GIS 通常包括正規的規則以幫助使用者選擇最好的程序以及工具以求解決在資料比對、資料品質、成本及準確性的問題。假使這些方法及所選擇之資料有可能產生預期中之結果，一個真正有智慧的 GIS 在主要資料估計操作之前可以完成誤差擴增研究。這可以告知使用者主要誤差來源為何，並且可以對使用者呈現出較好的結果

除了 Dunn *et al.* (1990)較悲觀外，我們相信對 GIS 模型系統作有系統的檢測誤差適可能的。本章已經舉例證明有許多方法經由 GIS 之操作得到誤差，並且一旦了解誤差擴增，使用者可以選擇這些方法以得到較好的結果。改進的選擇包括：

1. 使用好的空間內插方法或使用數值方法取代簡單的邏輯；
2. 收集較多的資料及適合的取樣；
3. 收集不同的資料；
4. 使用較好的模型—使資料與模型相匹配；
5. 改良模型的等級；
6. 由相配之關聯性結構改良空間或時間的解析度。

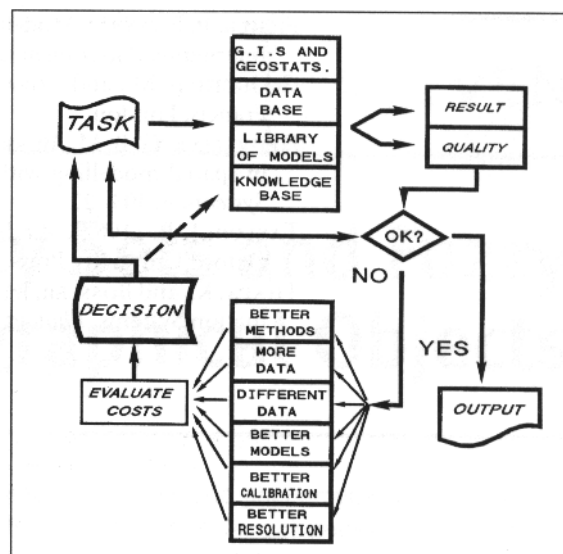


圖 10.13 (原文 p.263)

每一種選擇伴隨著成本的估計，所以可以做出合理的決定(圖 10.13)。此系統可以指出結果可以比預期中好之情形，在例子中儲存資料收集及過程可以確保無重大的資料遺失。

### 最後註記(End Note)

本章證明我們需要概念上的見解、資料收集之方法以及伴隨著所有資料來源之資料交替分析技術，並且需要選擇合適的標準。資料累積及 GIS 中誤差之角色愈行增加(例如 Guptill and Morrison 1995, Thapa and Bossler 1992)，雖然到目前為止不普及，誤差擴增統計之整合性的實際工具已經存在 GIS 之中。許多重要的組成已經發展，並且很快的被研究者接受，且許多研究者已在進行誤差擴增方面研究(Lanter and Veregin 1992, Lodwick *et al.* 1990, Heuvelink 1993)。

### 問題(Questions)

1. 討論使用 GIS 對環境進行模型化時，得到關聯性空間結構之價值及其困難？
2. 在任何模型輸出之誤差適依賴於(1)回歸方程式之未確定性？(2)資料中的空間差異？你如何調查此數量？
3. 你如何設計一個成本--效益研究以將預測誤差關聯於(1)模型等級點的數量？(2)使用



於內插之位置之數量？(3)使用於預測鋅含量的屬性種類？

4. 在第 7 及 8 章中之所有 GIS 分析的例子中，完成誤差來源及發展流動圖以顯示未確定性如何經由模型而擴增。

### **建議延伸閱讀(Suggestions for further reading)**

- Burrough, P. A. 1992. Development of intelligent geographical information systems. *International Journal of Geographical information systems* 6: 1-15.
- De Roo, A. J. P., Haxelhoff, L., and Heuvelink, G. B. M. 1992. Estimation the effects of spatial variability of infiltration on the output of a distributed runoff and soil erosion model using Monte Carlo methods. *Hydrological Preocesses* 6: 127-143.
- David, B. Van Den Herrewegen, M., and Salge, F. 1996. Conceptual models for geometry and quality of geographic information. In P. A. Burrough and A. U. Frank (eds.), *Geographic Objects with Indeterminate Boundaries*, Taylor & Francis. London.
- Goodchild, M., and Gopal, S. 1989. *The Accuracy of Spatial Databases*. Taylor & Francis, London.
- Guptill, S. and Morrison, J. (eds.) 1995. *The Elements of Spatial Data Quality*. Elsevier, Amsterdam.
- Thapa, K. and Bossler, J. 1992 . Accuracy of spatial data used in Geographic information Systems. *Photogrammetric Engineering and Remote Sensing* 58: 841-858.

# 第十一章 模糊集合及模糊地理物件

## *Fuzzy Sets and Fuzzy Geographical Objects*

### 前言

我們假設真實世界現象可以被模擬，不管是藉由精確定義或描述實體，如多邊形或平滑連續的區域。不確定性已被看做可能性，使用常見的統計方法及空間統計。本章我們描述不確定、複雜的、以及以模糊、部分重疊集合的模糊性的方法。替代機率理論，模糊集合理論使用公認的機率概念，以用來描述模糊隸屬函數。模糊隸屬函數允許個體具部分差異、重疊的集合。此集合可以由外部定義，使用語意輸入模式(Semantic Import model)或是由使用模糊 *k*-means 方法的多變數資料計算。假設原始值為點位置的量測，隸屬函數可以由內插製圖：貫穿不同模糊集合表面的區域是位於混亂的位置，且可以精確地提供明確的邊界。這種方法是解釋於土壤調查、土地分類、污染製圖及植被科學的應用。

第二章解釋為了模擬真實地理，首先必須將世界區分為不管是明確的實體(土地區塊、行政單位、土壤製圖單位、或是生態型)或是連續的區域(大氣壓、溫度梯度、地表水層級、族群密度)都是不連續的，通常 0 如固定網格。這些主要的空間實體一點、線、面、像元一由位置、屬性及位相來描述。原理組成(第二章，28-29)提供一般邏輯規則(如數學)操作空間實體的基本規則。其他形式資料的擷取與運作最主要工作為將分開的資料整理成群、成組：而這些實體為符合或不符合選擇準則。

在分類及修正的過程上，我們下意識地使用思考的基本法則，首先由 Aristotle 發展，命名為：

特性法則(任何事是什麼它就是什麼，房子就是房子)

非矛盾法則(某些事和它所否定的事是不能同時為真—房子不能同時為房子和非房子)

排中律；互斥定理(*the principle of the excluded middle*)：每個說明是對或錯—這房子是存在或不存在

排中律(互斥定理)確保所有常見對數的陳述可僅具兩個值—真(true)或偽(false)，並可以編碼成 0 或 1。這假設是大部分數學及電腦科學的核心區(Barrow 1992)，因此在計算及資料修正的範例自然是很深刻的連結至我們的工具。兩個原理：對數值產生分級重複，這是空間集合關係概念及空間真實可能性概念。

許多地理現象，並不是簡單明確的實體。在許多時間及空間尺度上的自然分析修正產生的格局及整體定義的事實是由許多互動屬性而來。因此，它通常是一個修改困難的實際問題用以分割真實世界為單一、非重疊集合。圖 2.2 顯示當我們察覺真實產生的某些事是既不是清楚的實體也不是連續區域時，我們使用簡單的實際主義，強迫觀測值為一個或其他形式，因為直至最近，我們在 GIS 內也沒有方法，藉由統計分離，處理明確(crisply)定義的實體，或是互相排斥的群體。

認識一個可以創造(虛構)所有差異的方法，符合本身形式的邏輯推論是具有一個釋放影響根據思考者對抗問題上像是公然爭論傳統形式

很多努力已經花費在如何更佳定義標準的地理資料，並討論讓軟硬體在多種品牌機器上能有意義溝通的組織中心狀態(見第十二章)。事實上，藉由邏輯理論的限制做二進位的判斷，我們限制資料分類的恢復的情形是可能完全的符合。在日常生活中，我們常做某些程度的妥協，當目標適合我們的計劃時—假如目標幾乎是我們所要的而我們將欣然接受。例如你要買一間新房子。你的計劃是它需具有三間臥室，具有一花園，最少需 300 平方公尺，位於具車站及購物中心 10 分鐘以內車程，不超過 20 分鐘至工作地點的路程，花費小於 150,000(英鎊)。這聽來完美精確，但是什麼是“10 分鐘車程”？它是“平均”的計算，假如是，又是怎麼計算來的？妳有考慮過，當在其他標準的都適合時，其他在這計劃外的選擇如房子具 290 平方公尺，價格 155,000(英鎊)、具四間臥房。

相似的問題不僅是困惑自然地理現象的分類的定義，如岩石類型、土壤或植被分類，但同樣也影響社會經濟群集的劃分，法律上判斷，國籍界定，甚至是國家的邊界(Burrough and Frank 1996)。

有罪、沒罪或沒被證明的，m'lud？有些法律系統(如蘇格蘭)提供不確定性，不同於其他僅具兩種判斷。

即使我們認為我們可以精確定義層級，但可能無法將實體精確的歸類到正確群組，因為原理是模糊不清或是量測值無法獲得足夠的正確性。不同的土壤調查者可能導致不同的地景資料判斷，增加本身的判斷可能在本質上是不合理的(見 Bie and Beckett 1973, Legros *et al* 1996)。解釋離散法的原理的差異以適於資料收集及資料持續分類，並具有一個直接顯示及解釋模擬結果的影響(第二章)。

## 不精確是一種思考方式(Imprecision as a way of thought)

因為在 Western 邏輯的形式思考分析具有傳統的使用範例，為真或為偽；是一個二進位或布林邏輯，我們在如何處理重疊的概念具非常少的訓練形式。此互斥法則及本身準則在數學驗證—歸謬法—在科學及哲學發展具極高重要性。使用電腦查詢語言的邏輯原理，是根據精確的思考如對或錯，暗示所有的資料庫物件都會適合某些範例。在這環境資料上並不是必須的。

普通的邏輯(conventional logic)有時引導至沒有解答的矛盾。情形如 Bongo 說他總是‘說謊’可以不管對或錯，因為假如事實是真的，它和總是在說謊是矛盾的。這種矛盾可以由允許部分真實或在自然發生在某些文化重疊的分類而解決(見 Barrow 1992, Burrough and Frank 1996)。同樣的，我們通常允許分類重疊及自然語言的部分真實，當我們沒有強迫精確概念的工作的科學及法定考量。

在自然語言中，我們可以自由增加修改，區分部分重疊分類及任何困難明確定義的中間區域情形。如，考慮區別高及矮的人。首先我們可能考慮到 190 公分的人是高的，150 公分的人是矮的(圖 11.1)。但是 170 公分是矮或是高？假如高/矮的界線在這兩種極端是不完全固定的，我們將無法區分，因此實際上我們使用中間區域並建立一個新的分類設定—矮—平均—高。產生 154 至 186 公分高的人的新的分類問題，但我們可以在重疊區域簡單的建立新的分級，處理在平均以下及以上的例子。這種過程可以連續延伸，直到達到我們一般可接受系統。相似定量的方法允許我們處理如枯燥忙碌的假日海灘或是氣候區域感覺的概念。

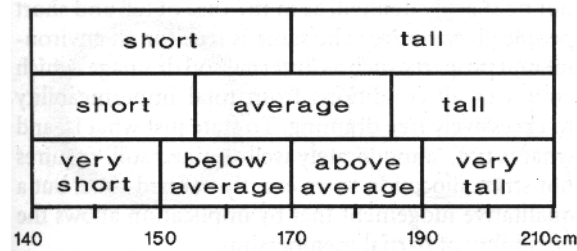


圖 11.1：自然語言藉由再定義及層級劃分解決模糊分類（原文 p.267）

在模糊集合，群體的分級是可由在 0 至 1 非常連續的尺度方面表現；許多地理資訊的使用者具有他們所需的清楚的概念(中心概念)。土地的評估及計劃者通常可定義特別土地種類利用的需求想法，但他們通常不確定關於穩定及非穩定土地的分類界線是否應該被描述。許多的情形下，他們同樣規劃他們一般的需求如哪些區域的土壤適於使用如具經驗的田野？、多少土地可由小農耕作玉蜀黍？，哪些區域是在氾濫區下？或是哪部分溼地遭受污染排放？。這些不嚴謹的形式必須轉變成有用的資訊的基本單位。未來，並不是在簡單模式下所提供的資訊就是精確的。許多由田野調查蒐集的資料通常是模糊不清的描述：土壤可被紀錄成排水不足、具適度的有效營養、看起來易受土侵蝕影響；植被可描述成充滿活力、部分活力等等。儘管人為標準定義這些具更精確性的關係，事實上，他們保留強烈不明確的定量。

在自然資源調查及模擬，允許使用者特別的分級定義，這是不合理，因為我們需要同意標準化分類的限制數量以促進訓練、研究，及一般資訊的轉換。顯然的，我們需要具體說明如何我們必須處理不明確資訊及邊界例子的方法。

## 地理現象及不精確(Geographical Phenomena and Imprecision)

地理現象比許多其他多變數定義的實體更複雜，需考慮包含屬性空間及地理空間的群體。屬性空間的群體所量測的全部實體是相同種類，基於屬性選擇的分類定義量測的群體，所有具清楚的分配準則可導致清楚、不重疊的分類。地理空間的群體量測是發生在相似種類實體所佔據的鄰近地區。

當經常設定天然或是人為實體，假設相似的實體將聚集在一起，我們僅集中由屬性定義的建築分類，包含絕對的、合理的。但在發展中及很多文獻用於窄範圍自然變數的研究在土壤、地表水質量、其他自然及人為地景方面變得特別劇烈。

以下的文章，我們首先測驗如何處理不精確的重疊屬性分類。然後解釋這種了解如何可以延伸至地理現象，以表示連續的田野或是如多邊形的實體。

## 模糊集合與模糊物件(Fuzzy sets and fuzzy objects)

雖然其他的群集分析具有群體重疊的概念(Barrow, 1992)，直到 19 世紀末 Western 思考沒有形式方法已獲取龐大的邏輯(monster)。廣泛建立多變值邏輯想法是在 1965 年開始，當 Zadeh(1965)引進模糊集合的想法來處理在可定義方法下的不確定的概念。模糊的關係曾被視為不適宜，因為它建議模糊的影像、非結構性思考，因此是非科學性及不好的。這是定義並非實例，但是名字具 stuck。自從 1960 年代，模糊集合理論才被發展成有用的論點、實用性工具，在其他學科的使用是具有價值的(Zadeh 1965, Kandel 1986, Kauffman 1975)。

模糊是一個不精確特性的分類形式，許多不同推論不可也不具有明確定義的邊界。這些

不精確的定義分類稱為模糊集合。模糊通常伴隨著複雜性。當我們處理數學或計算模式上在經驗現象的模糊不清、茫然、矛盾時，它適合使用模糊集合。空間分析和廣大空間尺度的互動的概念在方法上不能完全被預測(比較第六、第十章)，因此可以體會地理資訊上模糊概念的需要。模糊集合理論是一個歸納但不是一個取代，為一易於了解的抽象集合理論，通常指的是布林(Boolean)邏輯。

模糊不是一個機率屬性，是一個集合隸屬的程度，以統計定義機率函數連結。它是機率的使用，個體是一個集合的構成，或所給予的情形是真實的。機率的推測可以基於主觀、直覺(專家)知識或是偏愛，但同樣關係到具有機率理論基礎的不確定性的清楚定義。如不確定性的分配可能連結某些重要的量測誤差的機率。

## 明確集合(Crisp Sets)

傳統或明確集合允許僅為二進位隸屬函數(真或偽)一個體於任何設定是隸屬或非隸屬。然而模糊集合承認**部分隸屬**(PARTIAL MEMBERSHIP)的機率，因此明確設定分類邊界不行或不可以被明確定義的情形，如高矮的例子。對於環境特性也是相同的，如'內部的土壤排水'是包含所有全部不滲透至過度自由排水的情形。這情形只是什麼是、什麼不是，'一個適度的適當排水土壤'不需要嚴格劃分至精確的分類定義；但是量化判斷的牽涉，允許部分隸屬的可能性。

明確集合的設定是所有關係獲取分級概念，分級邊界是明確的。單獨觀測值  $z$  是一個集合關係是由隸屬函數(*Membership function*) $MF^B$  表現其程度，由明確(Boolean)集合可以獲得 0-1 的值。注意  $z$  是使用在一般的屬性值，如地區性變異理論(第六章)。形式上寫成

$$MF^B(z)=1 \text{ if } b_1 \leq z \leq b_2$$

$$MF^B(z)=0 \text{ if } z < b_1 \text{ or } z > b_2 \quad 11.1$$

$b_1, b_2$  是  $A$  集合的精確邊界定義。假如在'污染'及'適度污染'間的邊界，污染的土壤的集合， $b_1=50, b_2=100$  單位，然後 11.1 的隸屬函數定義所有為'適度污染土壤'。

## 模糊集合(Fuzzy Sets)

模糊集合是數學式的定義如下：

假如  $Z$  代表空間物件，則在  $Z$  的模糊集合  $A$  的等級配對的設定為：

$$A = (z, MF_A^F(z)) \quad \text{for all } z \in Z \quad 11.2$$

隸屬函數  $MF_A^F(z)$  是所知的'z 在 A 的關係等級'， $z \in Z$  表示  $z$  屬於  $Z$ 。通常  $MF_A^F(z)$  範圍為 0-1，1 表示充滿集合關係(如"典型輪廓"或"形式")及 0 為沒關係。 $z$  在  $A$  的關係等級反應等級種類並不是基於機率但是是公認可能性。 $z$  在  $A$  的  $MF_A^F(z)$  值可以表示為暗示集合  $A$  和物件  $z$  聯合的適合度；換言之， $z$  在  $A$  的  $MF_A^F(z)$  具體說明精確範圍  $z$  可以視為歸屬於  $A$ 。因此  $MF_A^F(z)$  值給我們一個方法給予問題一個等級答案"什麼程度是觀測值  $z$  為  $A$  分級的關係？"。圖 11.2 使用 Venn 方法由圖示以解釋明確及模糊集合的差異。

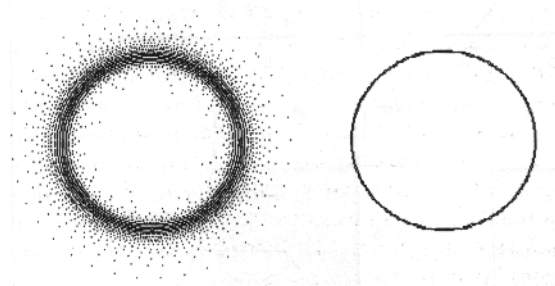


圖 11.2：左、模糊集合的影像；右、布林集合的影像（原文 p.269）

簡單來說，在模糊集合，等級關係是以尺度的形式表示，可以在 0-1 間連續修改(變動)。於隸屬函數中，個體接近至中心概念的值接近或等於 1；距離更遠的值越小。個體可以立即獲得圍繞在不包括中間的原則(但真實不是絕對的)。差異程度的問題，是比一個集合具更多的成員。問題是去量測明確的隸屬函數。

### 選擇基於屬性的分類(Choosing Classes Based on Attributes)

許多例子中，明確集合的邊界值的選擇不是(a)基於專家知識(如邊界值區別標準由習慣、法律或是外觀分類法)就是(b)使用數值形式的方法。基於專家知識的分類通常利用或引進分類，以建立不包含直接參考的當地資料集合。他們可能近似'自然的'區分，但是在任何統計判斷都不是適合的。僅有兩種參數，較高或較低的邊界值需要的。這些分類使用在實用科學及行政的大量(重要的)處理。

所謂自然的分類方法建立的分類是適合當地獲取的資料集合。實際上，選擇的分類法、參數值等，可以強烈反應分類的結果。數值形式的方法是被迫性資料的主要研究工具。

兩種選項同樣可能包含模糊集合。首先及最簡單的方法是使用個體可以設計成等級關係前的隸屬函數。所知如語意輸入法或模式。這是類似上面的(a)。

其次是類似群集分析及數值形式的隸屬函數的值，是一個使用分類的函數。這模式常使用說法是如模糊 k-means 法。

此兩種方法可以有效應用在環境資料，如本章其餘的解釋。首先描述顯示法模式和其他模糊邏輯的原理，其次是模糊 k-means 法。

### 選擇關係函數(Choosing the membership function)

#### 1. 顯示法(The semantic import approach, SI)

顯示法在使用者具有很好的歸類資料的質化想法情形下是有用的，但標準布林模式困難的相關，有許多不同的理由。選擇模糊集合並不意味著唯一的選擇。明確集合的邊界的選擇及分類區間可以是一個客觀或是主觀的過程(如 Burrough 1986, Evans 1977)。根據科學家同意的分類定義方法：變化通常多被認為是在分類間隔選擇明顯邊界。相同的是正如真實的分配模糊集合的隸屬函數。隸屬函數應該確保關係分級在中心集合是 1.00，在模糊邊界的近似方法它脫離至集合的值為 0 的區域外。隸屬分級=0.5 的點是稱作'交叉點'。隸屬函數需被定義在取得這些狀況下的方法，並不是所有函數都是可能的。

#### 使用顯示法的適合隸屬函數(Suitable membership function for use with the SI approach)

正如不同形式的機率分布(常態、對數、矩形、雙曲線、卜瓦松等等)，同樣的也有許

多不同種類的模糊隸屬函數。他們用來測定集合邊緣的隸屬值。多數普遍是線性的  $MF^F$  和”正弦(SIN)  $MF^F$ ”。

線性的  $MF^F$  是由一對處於高峰(MF=1)的傾斜線的集合的中心概念，c 和具有 MF 值在邊界為 0.5 (圖 11.3b)。線段的傾斜給予模糊轉換區域的寬度。注意這些在傾斜線內區域但是在布林矩形外是部分真實的區域。

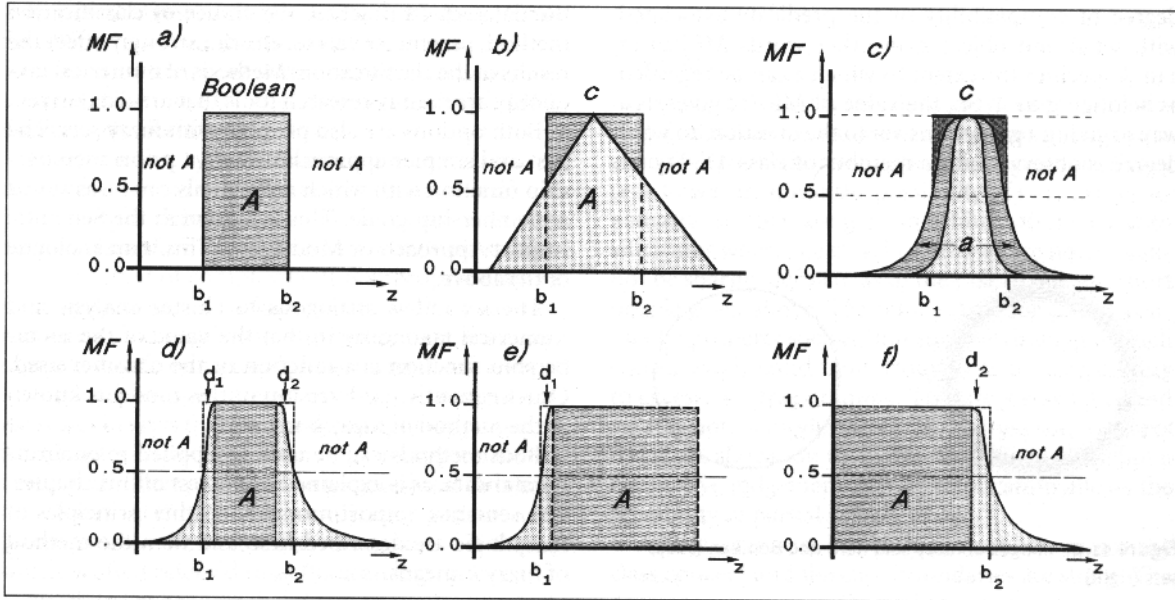


圖 11.3：顯示方法的布林及模糊隸屬函數 (原文 p.270)

正弦(SIN)  $MF^F$  如下獲得：

$$MF_A^F(z) = \frac{1}{(1+a(z-c)^2)} \quad \text{for } 0 \leq z \leq P11.3$$

A 是問題中的集合，a 是控制函數曲線的參數，c 是屬性 z 在中心概念的定義值。由 a 值的變化，隸屬函數的形式及交叉點位置可以趨近任何寬度的布林集合(圖 11.3c)。我們稱此隸屬函數為’模式 1’。

例如，考慮 0-200 公分倒轉的土壤深度，我們希望區別由淺至深的深的土壤。假如選擇 c=100 公分如完美的中心(標準的索引)是深的土壤的模糊集合，在 a=0.0004 時會得到表示在 50 公分及 150 公分交叉點較低或較高的值，而  $MF_A^F(z) = 0.5$ 。z 值側面的交叉上方點的中心概念(本例子為 50-100 公分及 100-50 公分)可以想成是圍繞在模糊集合的中心概念的轉換區域。

### 延伸隸屬函數的定義至適合中心概念範圍值的集合(Extending the definition of membership function to sets with a range of values that meet the central concept)

通常，它很明確的延伸一個模糊的中心概念以包含可能值的範圍而不是單一值，而這個想法在明確集合下的定義是絕對的(方程式 11.1—圖 11.3a)。方程式 11.3 可以簡單的模擬來獲取值範圍之上的中心概念；代替使用參數 a 來定義模糊隸屬函數的形式，在練習時更簡單使用 d<sub>1</sub> 到 d<sub>2</sub> 寬度的轉變區域。

圖 11.3d 顯示具中心區域及較高或較低的 d<sub>1</sub> 到 d<sub>2</sub> 寬度轉換區域的模糊集合。雖然較

高或較低的轉換區域至模糊集合可以是相同的寬度，這不是必須的，圖所顯示的較高或較低轉換區域也不是相等的寬度。我們稱這為模式 2。它像是直覺的去設置轉換區域，因此模糊集合的交叉點位在相同 Boolean 集合的邊界，因此保證位置剛好落於 Boolean 集合邊界的觀測值，同樣沒有完全隸屬中心集合。假如觀測值具有一個連結的誤差地帶(是指所想的紀錄的值可能位於 Boolean 集合內，真值可能在外面)，這是很明顯的。因此  $d_1$ 、 $d_2$  的值是實質的轉換區域的半寬度因為他們在 Boolean 邊界至  $z$  值內部的寬度和中心概念相對應。同樣由 Boolean 或模糊集合獲得的選擇在邊界  $b_1, b_2$ ， $MF^F = 0.5$  時可以嚴格比較。

模式 2 是由三個方程式定義

$$MF^F(z) = \frac{1}{1 + \left(\frac{z - b_1 - d_1}{d_1}\right)^2} \quad \text{if } z < b_1 + d_1 \quad 11.4a$$

$$MF^F(z) = 1 \quad \text{if } b_1 + d_1 \leq z \leq b_2 - d_2 \quad 11.4b$$

$$MF^F(z) = \frac{1}{1 + \left(\frac{z - b_2 + d_2}{d_2}\right)^2} \quad \text{if } z > b_2 - d_2 \quad 11.4c$$

$MF^F(z)$  是連續隸屬函數反應至屬性  $z$  的值。注意假如參數是 0，方程式(11.4)產生 Boolean 隸屬函數(方程式 11.1)。

### 非對稱隸屬函數(Asymmetric membership functions)

在許多情形下僅有較高或較低分類的邊界可能會是特別重要。這可能是真的，例如，我們僅希望了解土壤深度足夠所需目的的情形下—假如超出這深度 5-200 公分是無關緊要的。在此情形下，可以很簡單的使用方程式 11.4 的差異部分來描述較高或較低方面的隸屬函數(模式 3，圖 11.3e 或是模式 4，圖 11.3f)。

### 選擇轉換區域寬度的值(Choosing values for the width of the transition zones)

他們像是在選擇  $d_1$  及  $d_2$  不難及快速的法則，但它像是關係到轉換區域的寬度寬度，在所曉得關於預測量測的真實或問題物件的屬性。由點的資料量測來看，轉換區域的寬度可能會影響到所知量測技術的正確性：由 Kriging 內插的網格資料，轉換區域的寬度可以由 Kriging 的標準偏差得到。假如模糊隸屬函數是使用在表現如何擴散可能的地理邊緣(和 Perkal's epsilon—見第九章)，相關的隸屬函數的轉換區域的寬度至地理邊緣可以使用此領域的專家模式定義。

### 模糊地理物件(Fuzzy Geographical objects)

在成為特徵之前，必須提醒讀者，模糊集合的理論最初僅在處理擴散的邊界及屬性空間的分層套疊。應用地理實體的屬性因此是明確的。但在地理的資訊，我們必須處理擴散地理邊界，模糊的地理物件可以同樣應用在特徵值如多邊形的邊界，或是可以經由資料點內插的隸屬函數值的變異。在這例子，歸類連續性的變化表面如模糊的田野，是有用的。如所有的地理資料，緩和的模糊資訊可以根據空間相關性的結構製圖，如第六章的解釋。



### 一個實用的例子(A practical example)

為驗證 Boolean 及模糊集合的差異，圖 11.4a 顯示 Lacombe 試驗農場地表部分土壤(黏土)的碳含量圖，Alberta 由平常的 Kriging 區塊內插 (Burrough *et al.*1992)。量測 755 公尺\*745 公尺的圖籍區域包含 5 公尺的解析度。圖 11.4b 顯示區域獲得的 Boolean 分類限制為 $\geq 30\%$ 黏土，圖 11.4c 顯示使用模式 3 的模糊分類具有較低的轉換區域(具 5%黏土)寬度。Box11.1 給予假的編碼以計算這些圖籍，並顯示計算模糊隸屬函數的值是多麼簡單。

### 顯示法：重述要點(The Semantic Import Approach：recapitulation)

不像簡單的 Boolean 集合其關係值是分離的 0 或 1，顯示法轉換資料至連續的函數範圍是由 0 至 1。隸屬函數值產生實體屬於考慮中的集合的**程度**。模糊分類可以處理設有真正明顯邊界的物件，由限制範圍和指定定義界線的分類方式，即簡單及直覺的方式。和 Boolean 分類比較，僅有邊界值可選擇，顯示法帶來的額外問題是(a)在線性、正弦間的選擇或是評估分類隸屬的函數及(b)選擇分散指標  $b_1$ 、 $b_2$  的值。因此顯示法比普通的 Boolean 方法需要更多資訊，但這是由額外敏感的資料分析簡單的配對，如以下所將要做的驗證。

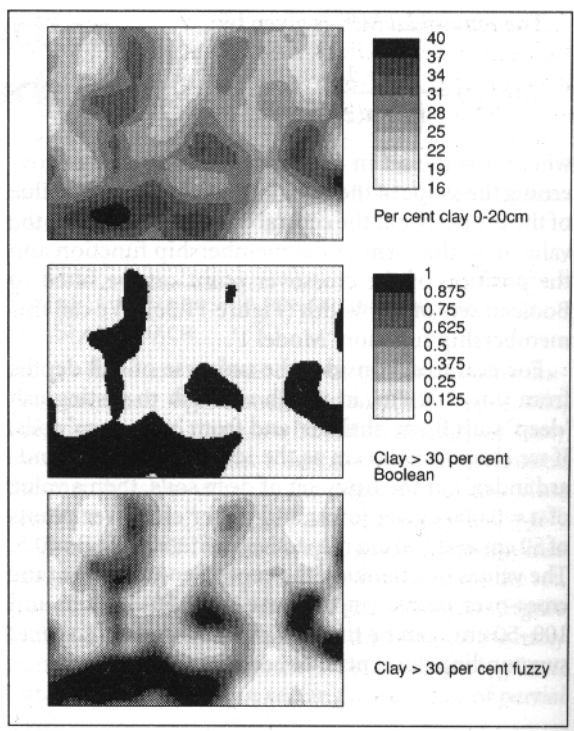


圖 11.4：土壤結構的布林及模糊校正 (原文 p.272)

Box 11.1：計算布林及模糊圖籍的程序 (原文 p.272)

### 多種模糊集合的作業(Operations on several fuzzy sets)

正如在 Boolean 集合，模糊集合資料的運用是使用邏輯疑問方法以選擇或是由幾個集合結合資料，傳統資料庫管理系統的標準疑問語言可以修改，以接受連續性的邏輯運算(如 Kollias and Voliotis 1991)。模糊次級集合的基本運算是和一般的 AND/OR/NOT/XOR 或是 Boolean 集合使用的運算相似。讀者可以由參考 Kandel(1986)、Kauffman(1975)或是 Kilr and

Folger(1988)得到幾近全部的運算。

Box11.2 是主要模糊集合的邏輯計算。注意整數的符號不表示'加總'，而是'物件 Z 的範圍'。我們的目的是多數使用的運算將結合'最大'、交叉'最小'及 context 結合(權重加總)。當連結的修改者使用如'本質'及'典型的'context combination 概念或是當不同的屬性可以補償其他屬性時，是有用的。這種想法特別適於在如我們所見的土地高程。結合、交叉或 convex combination 導致計算一個新的 MF 值(我們稱為 Joint membership function(加入隸屬函數)值或 JMF)。

Box 11.2：模糊集合的操作 Operation on fuzzy sets (原文 p.274)

現在它可簡單見到單獨的實體變成一個或更多集合的構成。圖 11.5 顯示兩個土壤的黏土質地鄰接分類的隸屬函數。觀測值 1 清楚在 A 分類，具有隸屬值在 A 分類為 1.0，在 B 分類為 0.01。觀測值 2 是接近分類邊緣：在 A 分類的隸屬值為 0.4 而 B 分類觀測值為 0.7。結合(OR)的 JMF 在觀測值 1 是 1.0，觀測值 2 為 0.7；交叉(AND)的 JMF 在觀測值 1 為 0.01 而在觀測值 2 為 0.4。注意假如這兩種分類具有不同的轉換寬度，所給予觀測的隸屬值在顯示法的總和不為 1。同樣的，在使用不同屬性的分類定義也是如此。

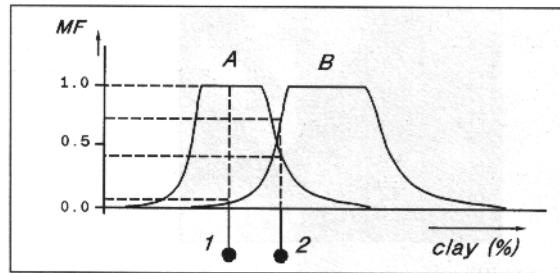


圖 11.5：兩個重疊的模糊隸屬函數 (原文 p.253)

### 聯合隸屬值的外插屬性(Joint membership value for interpolated attribute)

Box11.3 證明不同屬性如何進行結合(OR)及交叉(AND)工作。假設我們由土壤剖面獲得的資料，可將黏土包含在誤差 5%左右的三種分層—0-20 公分、30-40 公分、70-80 公分。讓我們定義一個'重要的土壤結構'如大於、等於 30%的黏土層為第一分層，大於、等於 35%的黏土層為第二分層、大於、等於 40%的黏土為第三分層。如之前我們使用具 5%分散值的正弦函數(模式 3)。在練習中，它可能必須去區別剖面在三範圍中僅有一個是'重要的'(黏土在剖面的任何地方)而非所有三層剖面都是'重要的'(黏土是遍布的)。最先的情形是結合(OR)，並需要選擇**最大** MF 值；第二是交叉(AND)，需要選擇**最小** MF 值。表格右邊的幾行顯示將被選擇的剖面，及被兩個類型拒絕的剖面。注意假如在深度量測誤差是±5%時，5%量測的黏土值小於 Boolean 邊界(反應至  $MF^F = 0.2$ )將是可能的"重要土壤"的集合的構成，因此他們將選擇完全拒絕 Boolean。

圖 11.6 顯示相同的分析，但是現在應用至 Lacombe 資料的土壤背景內插的表面以區別地區某些部分黏土層(圖 11.6a, b)至'黏土遍布剖面'(圖 11.6c, d)的狀況。很清楚的，在地理資訊系統的模糊分類是一個不重要運算，是為提供'製圖代數'。運算在黏土背景的三个分層中將是相同的本質，如由土壤多邊形表示，但結果的現象將受製圖單元的形式支配。

Box 11.3：黏粒最大及最小的 MF 值的例子 (原文 p.275)

保留相對較小的值；強調相對較大的值(Comparative minima are in hold；comparative maxima underlined)

Limits: layer 1  $\geq$  31% layer 2  $\geq$  35% layer 3  $\geq$  40%；width 5%

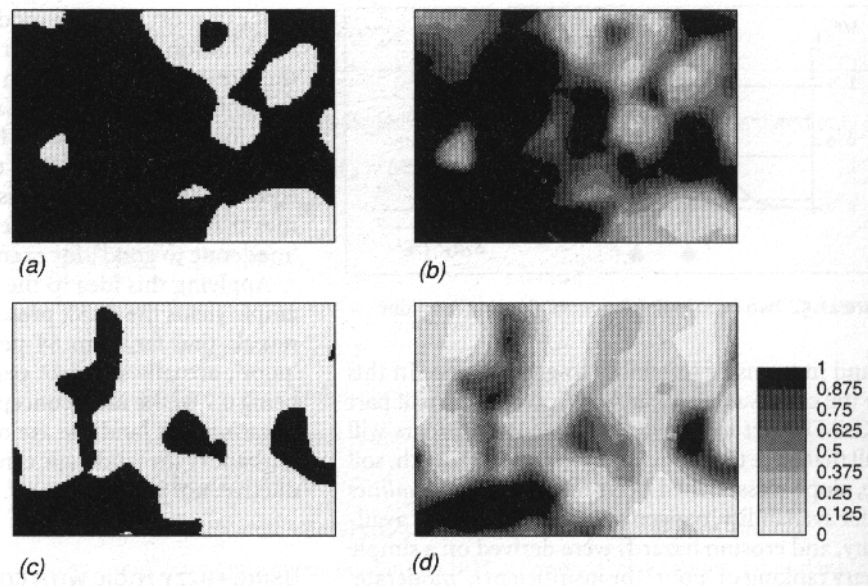


圖 11.6：(a, b)布林及模糊聯集的集合顯示在任一層黏粒的發生(c, d)布林及模糊的交集顯示黏粒發生於土壤剖面的情形（原文 p.276）

假如在點(support)的資料是原始的，必須決定是在獲得模糊分類前先將所有資料內插至一般網格，或是先分類剖面資料再內插至隸屬函數。第二種選擇於 Boolean 分類是困難的，除非由指標函數進行內插(見第六章)；但是模糊 MF 值則否，雖然他們可能是理論的問題(Gruijter *et al* 1997)，因為模糊 MF 並非常態分布或一個被迫具有 0 至 1 的範圍。在另一個模擬區域，計算的因素可能決定哪個路線是最好的。對於許多地理資訊系統的使用者，應用在已經內插好的資料的分類模式將會變得更簡單。

### 使用原始資料的模糊分類(Fuzzy classification with ordinal data)

在第七章顯示了常見的 Boolean 邏輯是如何在適當關係的總體的土地區域分類發展一個收穫(crop)。在這個例子，收穫是玉蜀黍，是位於肯亞的 Kisii District 的小部份位置(圖 7.10)。讀者需回憶已經製圖的土壤特性(土壤深度、土壤型、坡度分類)、土地特性(水的有效性，氧的有效性、營養的有效性、及侵蝕的機會)是在一個簡單的三分等級—'不好(或不足)'、'中等的(或只是大約足夠的)'及'好的(確實足夠)'。最後土地適合的分類多數限制因子的簡單規則來達成。換言之，對於所有地位或是網格單元，最差等級的任何四個土地特性決定最後的等級。最後結果是 84%的區域是明顯的'不好(或不足)'、5%是'中等的(或只是大約足夠的)'、11%是'好的(確實足夠)'。

檢視土地特性的四個圖(圖 7.10)顯示有效性水及營養的有效性是適合玉蜀黍最大限制的兩個土地特性，可是這兩個屬性可以更容易修改(由灌溉或增加肥料或是有機肥料)。如之前 Lacombe 土壤的黏土背景的例子，我們可以定義每個土地特性的模糊隸屬函數。在這例子，函數已經被定義，因此轉換區域至一半的等級是相等的。這種土地特性的區別是'由不好(或不足)至中等的(或只是大約足夠的)'至中等的(或只是大約足夠的)至好的(確實足夠)'的例子。

應用這種想法至 Kisii 的土地高程例子產生如圖 11.7 顯示的結果。這展現了遠至 84%的

區域是'不好(或不足)'、實際上某些 71%的分數比'好的(確實足夠)'理想概念的 0.7 更好。它牽涉的土地利用型是明顯的一包含農業的發展，地景可以被保護，或是建立，而不是只有書寫而已。

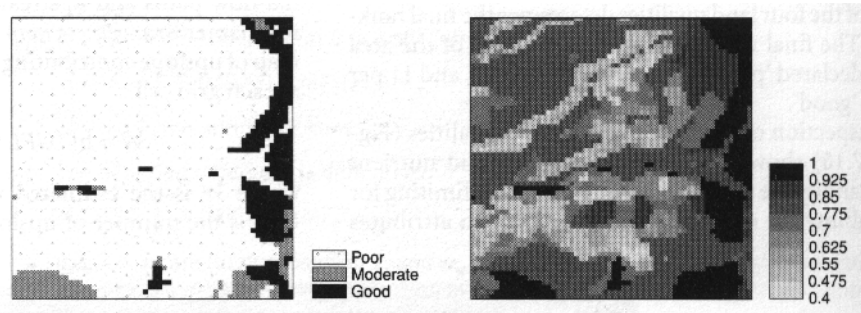


圖 11.7：對照使用常見(左)及模糊當量(右)的綜合土壤潛能分類的結果 (原文 p.276)

### 使用連續性資料的模糊邏輯如流域網系統(Using fuzzy logic with continuous data such as drainage nets)

如第八章所顯示的，由網格的連續面、坡度、潛在的水流路徑及上游區域，分離次要的資料是簡單的。如第八章，一張溼度圖可以由坡度上提供的區域圖及每個網格單元的坡度圖計算。

$$W = \ln(UPL * g / \tan \beta) \quad 11.5$$

W 是溼度於每個網格的推估值，UPL 是排除在網格單元中的坡度上要素的數量，g 是每個網格單元覆蓋區域， $\beta$  是每個網格上的土地坡度。

因為上游的分布區域圖及溼度指標是由他們的連續性變化獲得，他們沒有簡單的方法去擷取，可以稱為邊界的特徵值。當連續或群集的屬性的問題是沿著限制的路徑，甚至輪廓線也不可以判斷。模糊集合在導出這種資料(如以下的例子)是明確適當的。

圖 11.8 表示一個衍生的圖顯示較高位相值導出結合底土黏土背景的溼度指數。黏土背景的限制是取得如 40%包含 5%門檻值，溼度指數限制是 14，門檻值為 6。圖 11.8 顯示 Boolean 糾正包含位於底土黏土 $\geq 40$  的指數 $\geq 14$  的所有區域，圖 11.8c 顯示相等的(等值)模糊。注意模糊糾正保持充足黏土的流域網的全部區域，因此維持流域的連續位相，在 Boolean 解析上會遺失。

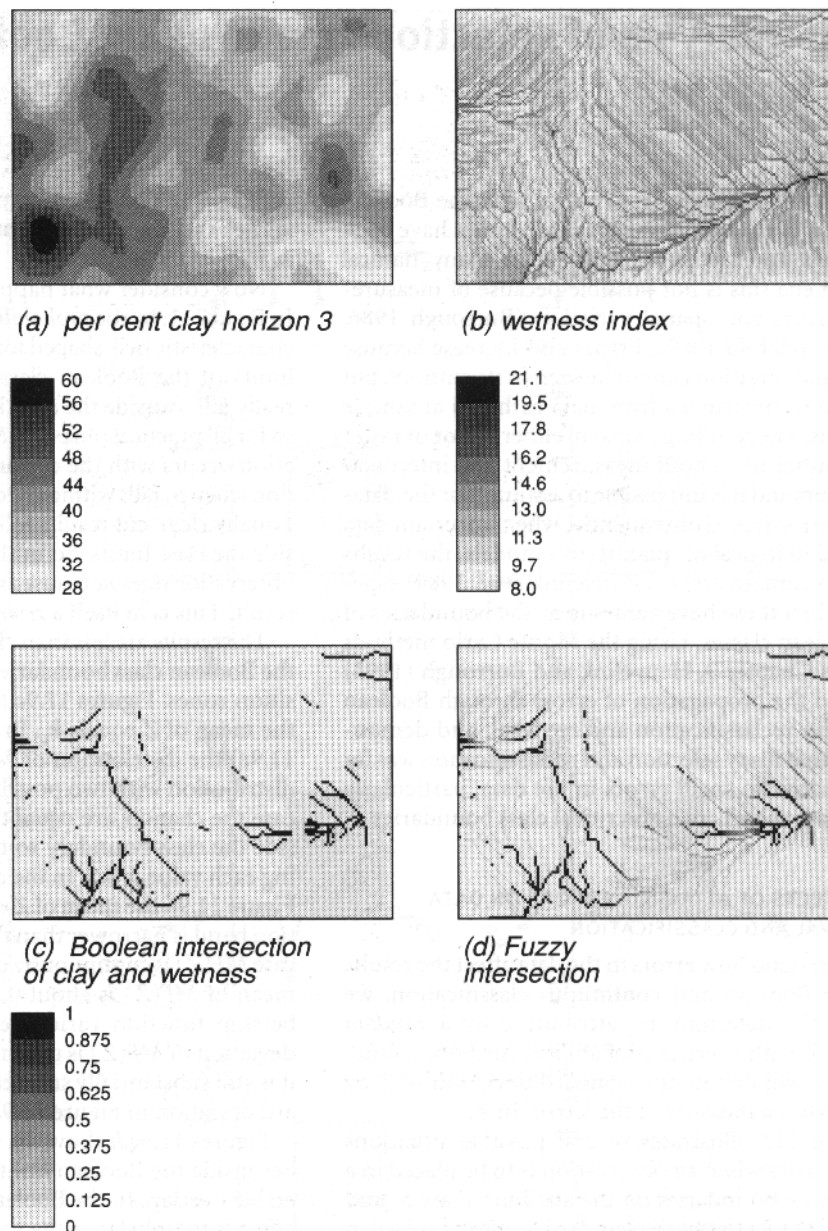


圖 11.8：對於有興趣的資料層級使用模糊方法以連續性資料詳細保存及布林選擇損失的連結（原文 p.277）

## 使用布林運算和模糊邏輯中選擇所造成的誤差分析(Error analysis of selections made using Boolean and fuzzy logic)

如我們所見，Boolean 模式主要問題是假設描述及量測的屬性是正確的。在許多自然現象中，這是不可能的，因為會有量測誤差及空間變異存在(Burrough 1986, 1989; Goodchild, 1989)。因為空間變異無法可以完全見到，所以誤差也同樣會增加，但是或許可由樣點位置獲取的資料重建。產生的圖(不是向量就是網格形式)因為量測及內插的誤差而更糟，所以不可能假設資料庫是隨意的誤差。通常，當不確定的資料是使用在邏輯或量化模式，結果將同樣會包含誤差(Heuvelink, 1989)，特別是當這些值接近 Boolean 分類的邊界。使用第九章的 Monte Carlo 方法，Heuvelink 及 Burrough(1993)經由 Boolean 及模糊再分類、糾正誤差增加的分析，說明模糊選擇及再分類是遠小於敏感的資料的小誤差，特別是當資料值是接近關鍵的分類邊緣。

## 資料糾正及分類上屬性誤差的影響(The effects of attribute errors on data retrieval and classification)

為了解資料的誤差如何影響 Boolean 及連續分類的結果，我們由隨機變數  $Z$  替換決定的屬性  $z$ ，包含一個確定的機率分布模式。我們將由  $\sigma_z$  表示  $Z$  的標準偏差(是一個在  $Z$  所量測的"誤差")。

圖 11.9 解釋許多可能的情形，可發生觀測值是在分類中替換，這些在屬性  $Z$  的邊界是  $b_1$  及  $b_2$ 。圖 11.9a 顯示沒有誤差時的標準 Boolean 例子，因此  $\sigma_z = 0$ 。屬性  $Z$  是決定的結果，因此單獨的觀測值不管是全部落在分類邊緣內(右手邊限制)或是落在外面(左手邊限制)。隸屬函數的反應值分別為 1 或 0。因為  $\sigma_z$  是 0，隸屬值  $MF(Z)$  同樣是隨意的誤差。圖 11.9b 顯示相同的情形—單獨觀測值目前是由連續隸屬函數來分類。在右手邊的觀測值是在 Kernel 分類中，因此  $MF(Z)=1$ 。左手邊觀測值是在 Kernel 及 Boolean 邊緣外，所以它會得到  $MF(Z)<0.5$  的值。

現在考慮當  $\sigma_z$  不是 0 時，什麼會發生。在圖 11.9c， $Z$  的機率密度( $P_z$ )取得鐘形形式的特性，並位於適當的 Boolean 分類的限制內。 $Z$  的機率真的落在分類限制外是消失不重要的，所以對於所有的實際目的  $MF(Z)=1$ 。當  $P_z$  落在 Kernel 分類中(圖 11.9d)，相同的情形會發生在連續的隸屬函數。相同的明確結果因  $P_z$  位於適當的分類限制外而產生。因此在這些例子中，觀測值的誤差不會造成分類結果的誤差。這於自身的結果是沒有任何價值的。

當  $P_z$  雜亂地伸展在 Boolean 分類邊緣或是連續轉換區域，結果是絕對不明確的。圖 11.9e， $f$  顯示當  $Z$  的平均值等於  $b_1$  的情形。在 Boolean 的例子(圖 11.9e)， $MF(Z)$  的分布變成具有兩個可能分離值(0 和 1)的分布。在這個例子， $Z$  落在分類內或外的可能性是相等的，因此包含另一個機率值為 0.5。圖 11.9f 顯示連續性例子， $Z$  的平均值等於  $b_1$  及分布地帶是比轉換區域更為狹窄。在這例子  $MF(Z)$  如同  $Z$  一般，是連續的分布。 $MF(Z)$  的平均約為 0.5，因為隸屬函數在  $b_1$  為急劇的變化， $MF(Z)$  的標準偏差比  $Z$  更大，雖然在圖 11.9e，它實質上是比 Boolean 標準偏差為小。

圖 11.9g，h 顯示當  $P_z$  大部分是位於在 Boolean 限制的  $b_1$  內，但仍舊具相當多的重疊。在 Boolean 例子(圖 11.9g)的結果是維持隸屬值在 0 或 1 的不平衡的機率；維持在 1 的值的機率是增加的。在連續的例子  $P_z$  處於 Kernel 分類，因此產生的分布是混雜的分布：它是包含連續及分離的分布。在某些例子中，隸屬值將精確至 1，這是一個明確的選擇，雖然小於 1 的值同樣會發生。雖然沒有顯示， $P_z$  比轉換區域更寬的情形將普遍地產生  $MF(Z)$  包含 0 或 1 的最高點的混雜的分布。

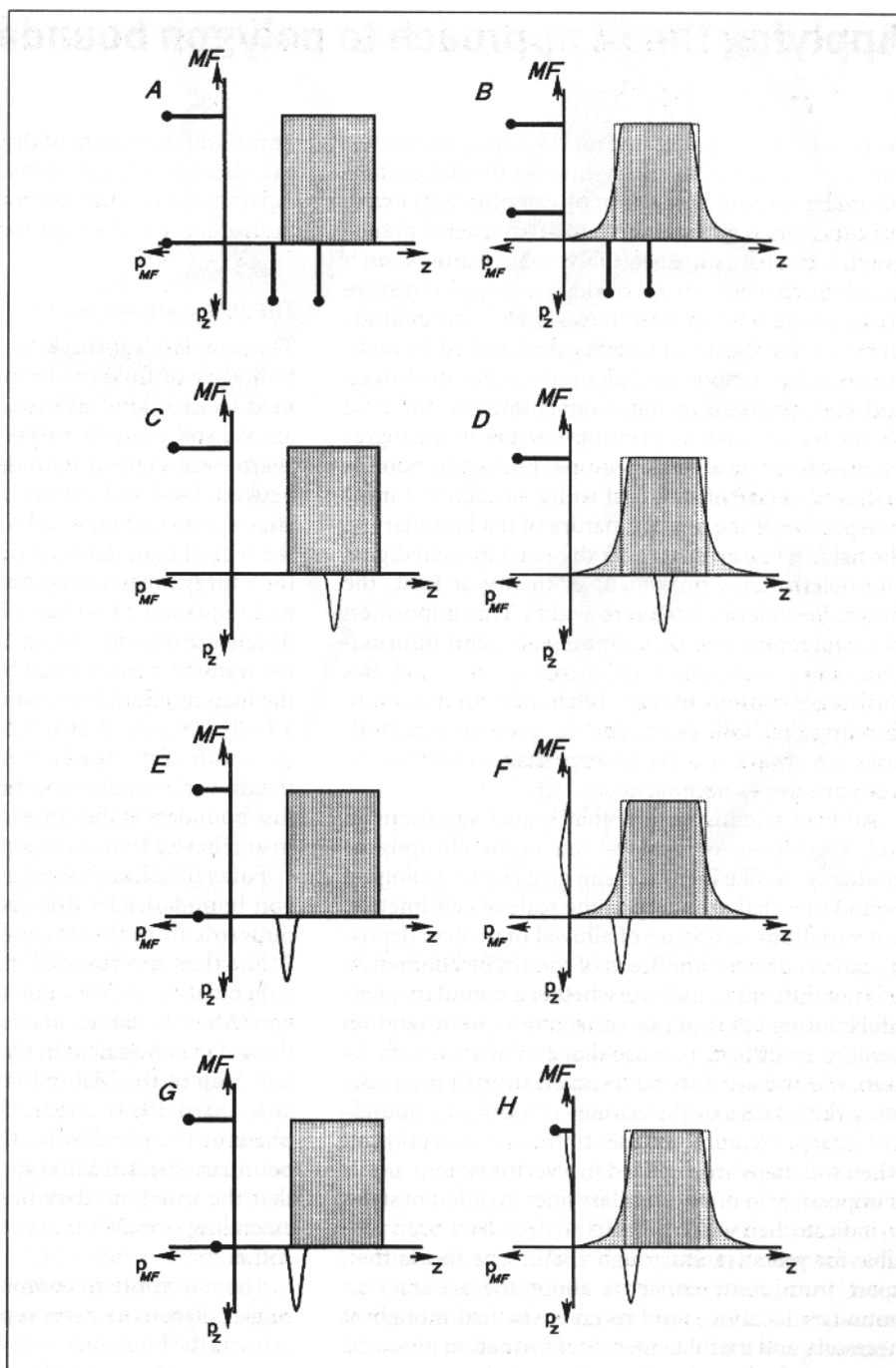


圖 11.9：量測誤差影響布林分類(左)及模糊分類(右)的結果（原文 p.279）

## 多邊形分界的標準方法 (Applying the SI approach to polygon boundaries)

目前為止，我們是集中在分類屬性，但許多種類的製圖牽涉到土地的相似區域的輪廓，通常藉由測驗它的表面的現象，或是在田野或由航空照片或衛星影像。顯著的特徵，如色彩的變化，坡度的斷點，或結構格局的邊緣，是使用推論的“邊緣”。輪廓首先是由鉛筆描繪，然後是由相似分類的區域中的觀測點來標記，最後再用墨水。通常，田野表單(sheets)是重複描繪如描繪所有土壤、生態型、具標準寬度線(通常為 0.2mm)的地理邊緣(不顧原始自然的田野邊緣)的製圖師熟練的複製。這些經後來人工的數化圖會產生數位地誌圖，最近的理論表示這些描繪的邊界幾乎沒有寬度。這些乾淨的製圖會抑制關於自然邊界改變的有用資訊，並

使一般使用者對於地誌圖產生錯誤想法，想成土壤、地覆、及地理邊界在任何地方都是明顯的，單位總是均值的。

所有田野科學家了解土壤、地覆、及地質的空間變化可以不連貫地或逐步的發生。壩的侵入在公分級的尺度可以增加急遽的地理邊界。但是沖積土或風的堆積的結構變異可以發生至數百公尺或公里上。這並不困難去指出不管是田野工作或是航空照片詮釋的邊界定義是急遽或是擴散的，因為這通常是可見的。同樣去附加相關屬性象徵或評估邊緣的寬度(急劇的、適中的、擴散的)在多邊形的弧度並不困難，當土壤圖是在向量系統數化時，也不可能描繪不同等級的邊界線來表明邊緣的寬度—事實上這些情形在過去幾年都是可能的(比較 Burrough 1980)。事實上，關於某些個別邊界位置的正確性，直至最近都很少有思考它所包含關於地誌圖的自然邊界資訊的必須性及有用性(Lagacherie *et al* 1996)。確實，這樣做的缺點是可能由具有急遽邊緣的多邊形區域來計算。

藉由多邊形模糊邏輯方法的使用，是非常簡單去合併關於自然邊界的資訊，並計算明顯的區域尺度。

### 圖籍單位方法(The map unit approach)

最簡單的方法是假設邊界形式(急遽或是擴散的)在每個不同種類或是多邊形可以是唯一的屬性值。如在範圍狹小的氾濫地區的土壤單位可能具有急劇的、容易定義的邊界，反之在黃土或粗糙的風成堆積邊界可能發生在數百公尺上。邊界形式的資訊可以轉換至模糊隸屬函數的 1、2、4 模式型式的參數(方程式 11.4a, b, c)，目前是應用在描繪邊界的距離。轉換區域的寬度需依自己想法選擇圖籍的尺度。如，在 1:25000 的圖籍所描繪的 0.2mm 的急遽的邊界代表 50 公尺厚，因此空間轉換區域的寬度在描繪邊界的中心上將是 25 公尺。在相同尺度可能延伸至 500 公尺的一個擴散邊界，具 250 公尺的轉換區域寬度。

模糊轉換區域首先可以由多邊形邊界，透過等向延伸及原始邊界向外擴展(參考第七章)，然後應用 SI 模式指出表面逐漸變化的多邊形內部至外部適當的 MF 值。如考慮圖 11.10 零碎的土壤圖例子(由 Mabura Hill Forest reserve in Guyana 取得的土壤圖，Jetten 1994)。圖 11.10a 顯示土壤基座穩定水準面。這單元具相當急遽的邊界，如一個最初的近似值讓我們假設真實的邊界具有延伸 500 公尺的不確定性位置，即轉換區域寬度為 250 公尺。



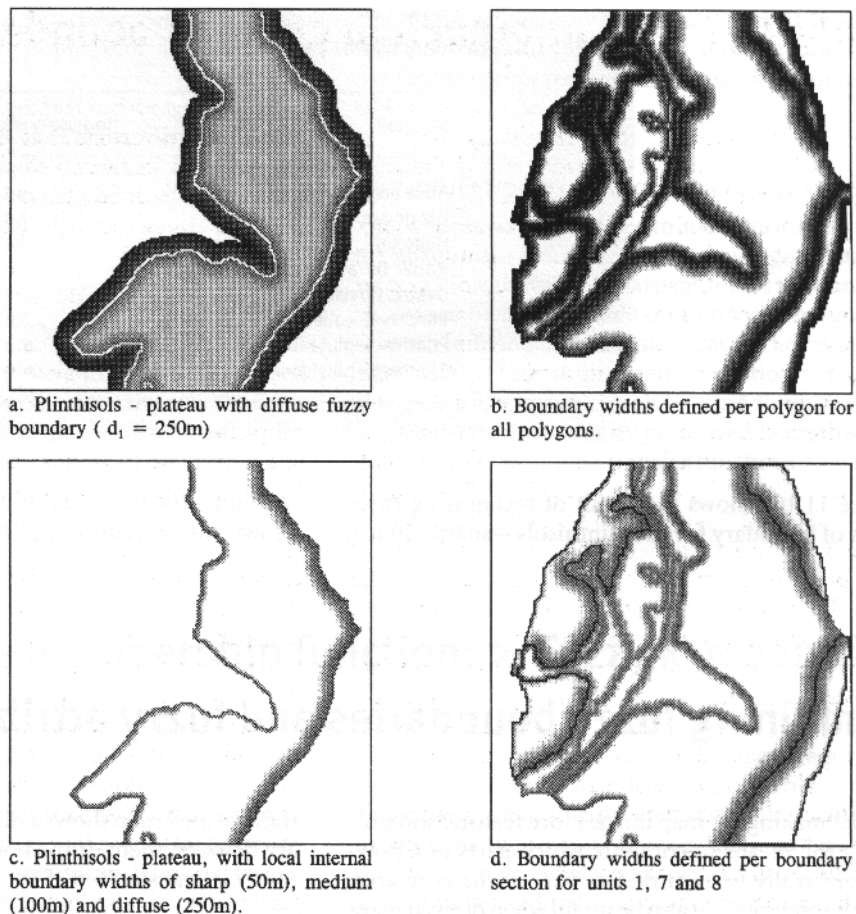


圖 11.10：使用模糊隸屬函數描述邊界寬度（原文 p.281）

單一的多邊形計算模糊邊緣的程序在網格的表示上是簡單的。一是取得具有過濾邊的邊緣，然後應用延伸的(spread)命令計算圍繞邊界的區域(第八章)。因為是選擇隸屬函數的參數，因此位置符合在交叉值(AND)原始描繪的邊界，即  $MF=0.5$ 。然後應用隸屬函數，伴隨著他們與邊界的距離位置，在位於適當的原始邊界內會得到 1 的隸屬值，在內部的但接近邊緣地位會得到 0.5-1 的隸屬值，在邊緣外的地位得到小於 0.5 的隸屬值。圖 11.10 顯示在白色部分的是原始邊緣。很清楚的，計算多邊形面積以符合任何水準的隸屬函數在現在是不重要的，因此以前對模糊評估多邊形面積懷疑的聲音已不在令人信服。

這程序可以重複於所有的圖籍單元，修改邊緣的寬度和單元相稱(表 11.1)。當所有多邊形由模糊邊緣展示，它在所有轉換單元的內部及外部展示是不明顯的，因為產生的圖很難解讀。因此較佳的是展示邊緣的 JMF 至內部的轉換區域(JMF 總能量 0.5)。圖 11.10b 顯示當邊緣寬度是根據表 11.1 設計時，不同圖籍單元的結果。

### 個別邊界法(The individual boundary approach)

通常，在周圍發生的所有方法，土壤或地覆面積並沒有相同種類邊緣。如高處的河流臺地可能在河流自身交叉的較高的一邊具有一個擴散的邊緣，而較低的一邊具有急遽的邊緣。其他的邊緣可能在長度部分是急遽的，在其他部分則為擴散。在邊緣區域增加屬性及單獨地延伸這些屬性並不困難。結果是後來結合的，而模糊隸屬函數則如之前的計算方法。圖 11.10c 顯示確認三種土壤基座邊緣種類—急遽(50m)、中等(100m)、連續(250m)的結果：圖 11.10 顯示這些程序在更多單元的應用。

表 11.1：圖籍單位的邊界寬度（原文 p.282）

## 結合模糊界線和模糊屬性(Combining fuzzy boundaries and fuzzy attributes)

除了使圖籍影像更易了解及提供方法計算可以真的被注意的多邊形中心區域部分，模糊邊緣在由原始土壤、地質學、地覆或土地利用型圖取得的圖同樣是有用的。一般而言，問題是一由轉換土地單元分類至一個數字或是編碼來代表所表示的屬性值所取得的圖籍。結果假如是鄰接的圖籍單元具有大的、不同的一個屬性值，則轉換的圖籍顯示它們之間的一個不連續邊緣。假如邊緣是急遽的，則結果是真實的，但假如結果是真的非常連續，則結果是一個由 pycnophylactic 內插移動的醜陋的邊緣(第五章)。

假如兩個多邊形的邊緣隸屬函數是使用於計算評估邊緣區域上屬性值的權重，它是可以提供關於邊緣上相異的兩個圖籍單元屬性逐漸變化的資訊。

$$Z_x = \frac{\sum (z_i * MF_i)}{\sum MF_i} \quad 11.7$$

圖 11.11 顯示使用計算穿過由 Plinthisol plateau 至高山旁土壤單元轉變的土壤溼度變數的權重程序，它具有 480mm/a 土壤濕氣的供應及分別為 400mm/a 及 250m 寬度的轉換邊緣，圖 11.12 顯示實地觀察和平常相等的 Boolean 分布(look up)表比較的結果。

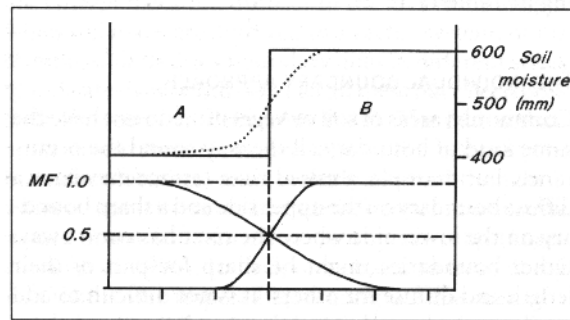


圖 11.11：使用模糊套疊計算屬性在模糊區域的平均權重（原文 p.282）

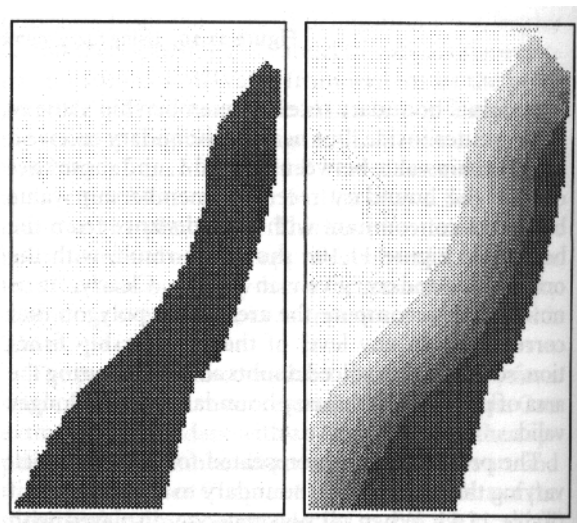


圖 11.12：傳統(左)及模糊(右)導出兩個具有逐漸、共享邊界的鄰接多邊形的土壤水分圖（原文 p.282）

## 重述要點：模糊多邊形邊界(Recapitulation: Fuzzy polygon boundaries)

SI 方法可以使用於增加不連貫的邊緣資訊至一個多邊形資料庫。關於如何急遽和擴散邊緣的資訊是可以在多邊形與多邊形間或是在單獨線的線段基礎來合併。結果得到關於橫越或是沿者邊緣的空間變化明顯的圖，可用來發展滿足簡單函數轉換的多邊形資料庫所獲得圖籍的資訊。

## 選擇隸屬函數：模糊 k-means(Choosing the membership function: 2.Fuzzy K-means)

雖然建立模糊分級的 SI 方法是非常有彈性的，但不是最適當的。使用者通常可能不知道它的分類是有用的、適當的或是可以建議適當、武斷的探究技術時，重疊分類可以是有用的。模糊 k-means 方法(如所知的模糊 c-means)是可以在地理水文學、土壤學(McBratney and de Gruijter 1993 and Odth *et al.* 1990)，及地覆製圖(Jetten 1994)使用的一個技術(Frapporti *et al.* 1993)。許多這些文獻同樣喜歡稱這方法為**連續分類**而不是**模糊分類**，因為它是在屬性空間的連續性分類；同樣的它也希望是在地理空間。

模糊 k-means 方法是起初是針對資料轉換及方便資訊轉換，發展野外的格局辨別(Bezdek 1981, Bezdek *et al.* 1984)。方法和常見的數值分類法具強大的關聯性(Sneath and Sokal 1973)。藉由轉換多屬性性質將 k(或 c)隸屬值轉換至 k 個分類或群集的資料轉換是可以理解的。群集在感覺上是適合的，因為在多變數內的群集變異是最小的。接近 0 的變異是指所有對象具有接近相等的屬性，表示他們之間在屬性空間上具高密度及小距離。相反的，大的變異在屬性空間是小密度及大距離。當邊緣在分類的屬性空間中，一個最適群集程序應該定義屬性空間的密集散佈如分類中心，應該是位於最低密度區域內。注意模糊 k-means 方法最初最適的、重疊的分類沒有談到關於地理上一連串的事物。

模糊 k-means 為一個反覆的程序作業，通常由目標的最初逢機位置開始至被分類的 k 個群集(增顯名字)。所給予的群集配置，每個群集中心(屬性值關係)是由目標屬性的平均計算。下一步驟，目標根據目標或是群集相似性的相關，重新配置在分級中。相似性指數通常是已知的量測距離：Euclidian、Diagonal(屬性排列具相同變異)或是 Mahalanobis(變數及共變數同時使用在距離排列上)，矩陣的使用很頻繁。重複配置是由重複開始直到在相似目標達到穩定的解析(答)，集合在同一群集中(圖 11.3)。

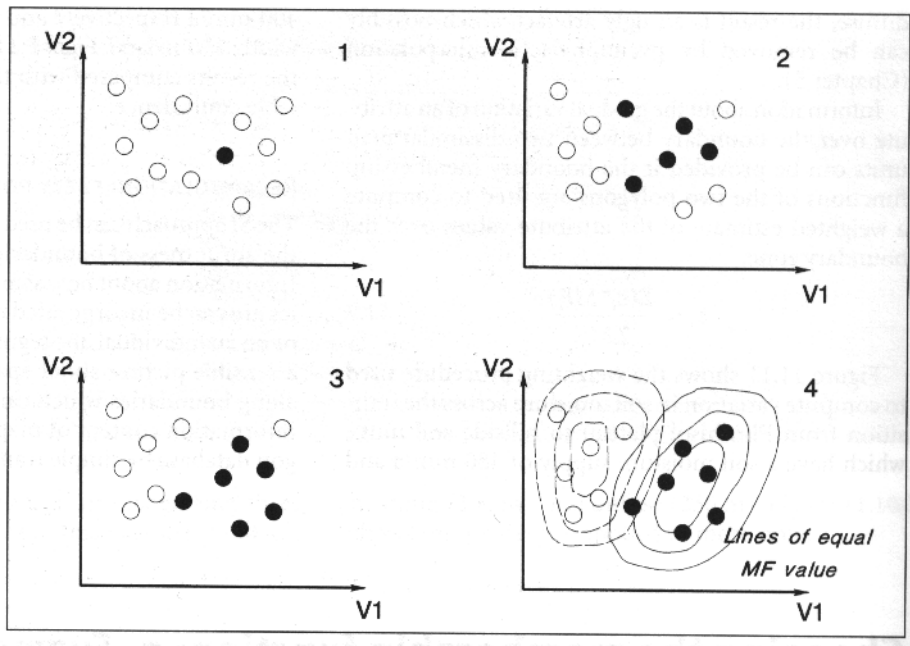


圖 11.13：群集組成使用模糊 k-means 重複 (原文 p.284)

在一般明確 k-means 的目標配置是到最接近的群集，這群集的 MF=1，其他則為 0。在模糊 k-means，隸屬值可能在 0-1 範圍間。Box11.4 列出使用計算 MF 及分類中心的演算法。在普通的 k-means 及每個目標的 k 關係總和為 1。這個例子缺少自由度，可能由建立一個特別等級(extracgrade class)而抵銷 (Gruijter and McBratney 1988)。參數 q 被稱為模糊指數 (Fuzzy exponent)，是測定群集模式的重疊數量。當 q 趨近於 1，配置是明確的，重疊是不被允許的。當 q 很大時是指完全重疊及所有群集是索引的。理想地 q 應該選擇來抓取通常是未知的正確的重疊數量，因為距離是相似的全面量測，背離屬性的分類中心可能由接近相符的另一個來補償。評估單獨屬性相對權重的影響是由適用量測距離的形式測定。

Box 11.4：使用模糊 k-means 群集的演算法 (原文 p.285)

模糊 k-means 的淨結果是單獨的多變數目標(點、線、多邊形)被設計為不同方面的 k 個重疊分類的 MF 值。資料方面，不同分類的質量中心是最適當的選擇。所使用技術的變異，MF 值加總會等於 1 而不同於單獨 SI 法的例子。表示不管所有集合都具有相同的值，如在 SI，其層級皆依據重要而定。這顯明的操作程序在 Box11.2 可見到，是指如交集及凸面的組合的程序需要小心地實行由模糊 k-means 所導出的集合。

## 模糊 k-means 及 Kriging 使用在重疊及多變數分類圖籍的例子 (Examples of using fuzzy k-means and Kriging to map overlapping, multivariate classes)

### (a) 在 mass 泛濫平原的重金屬污染(Heavy metal pollution on the mass Flood-plains)

第五及第六章以 Netherlands 南方的 Meuse 泛濫平原固定泛濫的河流區域取得的土壤樣品資料集合說明內差的方法，由不受重視的較大規模區域的資料是用在第十章解釋花費及效益。這裡同樣使用較大規模次集合資料，但包含所有的屬性(海拔高、至河流距

離、重金屬濃度(鎘、鋅、鉛、汞)及有機的原料)以解釋多變數的模糊 k-means 方法。

隨著最初的氾濫頻率圖(見第五章)，屬性資料被分類成三個分級，模糊重疊  $q=1.5$ 。模糊 k-means 分類在不同取樣地位產生不同於三個分級的一個新資料隸屬集合。分散樣區的三個分級的隸屬值(圖 11.14)顯示資料一般是適當分散在屬性空間中，僅有很少的觀測值取得中間的隸屬值。表 11.2 表示分級的摘要統計。

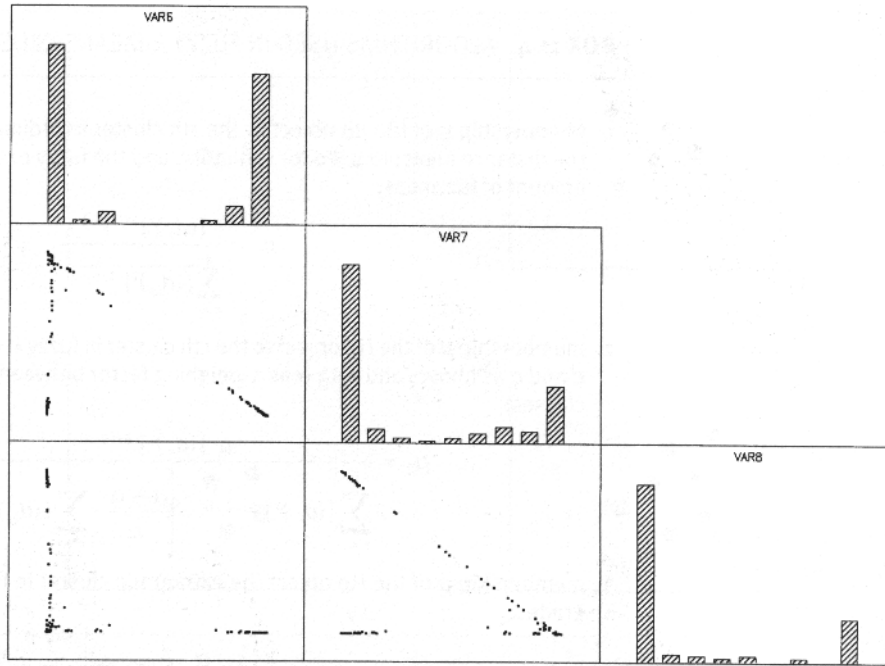


圖 11.14：土壤污染資料的隸屬值的分散(Scatter)直方圖 (原文 p.286)

表 11.2：模糊分級、氾濫污染的統計 (原文 p.286)

圖 11.5a, b, c 顯示三個分級的內插面。關於三種模糊分級間關係及氾濫頻率分級的一個強烈的建議，名為氾濫頻率分類 1，覆蓋類似模糊分級 3 區域，模糊分級 3 佔用區域是由南至北。注意這些在第五第六章使用的部分區域上的分級分布，是在圖 11.15e 的線外。在那區域僅有兩個模糊分級是真的重要，他們是達到小區域氾濫頻率製圖的多變數分析的推論之一。

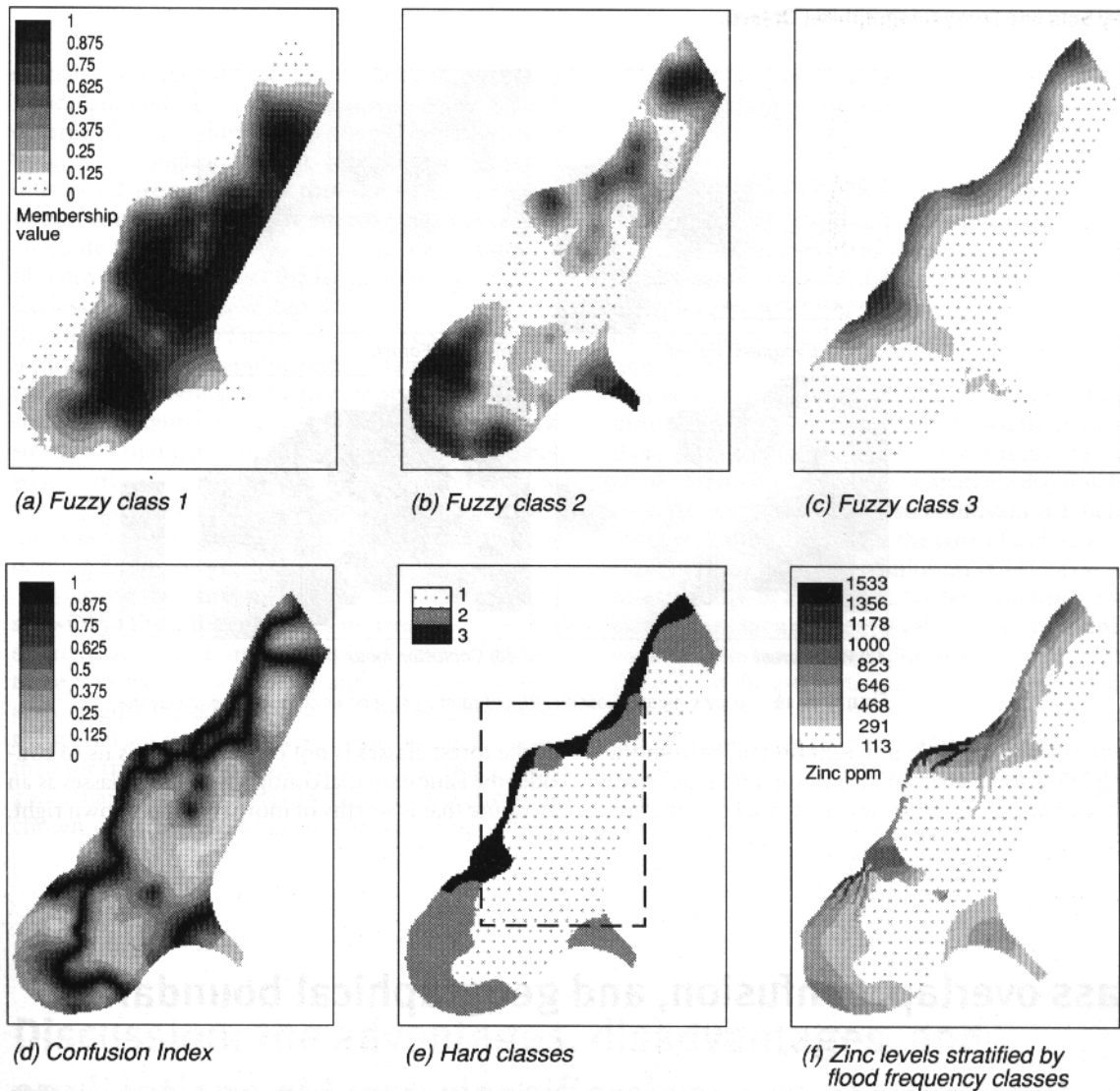


圖 11.15：圖(a, b, c)三種土壤污染等級的模糊 k-means 隸屬值的內插圖；(d)混亂指標圖；(e)每個分級是優勢的波狀圖，外框輪廓的虛線區分區域是用在第五及第六章以解釋內插；(f)鋅分級的圖籍包含和(e)相對的克利金分層

### (b) 熱帶雨林形式 Rainforest types

相似的練習是由 Guyana 熱帶雨林 100m\*100m 的網格每 0.05 公頃的 252 個地覆調查產生的結果(Jetten 1994)。最初在模糊 k-means 程序分析前，首先將存在/消失種的資料聚集至豐富度分級，產生三個分級表，名為'乾燥常綠森林'、'混淆林'、'潮濕森林'(Jetten 1994)。表 11.3 表示這些分級的統計摘要。

表 11.3：模糊分級及熱帶雨林形式的統計（原文 p.287）

分級關係是內插至 50m 網格，並顯示於圖 11.16。不像污染分級，這些分級並沒有適當的空間相關，可以由靜止的、合適的 variograms 礦物變異的比率看見。很清楚的，儘管屬性分類由氾濫資料最大、最小及平均隸屬值的關係的特性是小的差異，森林分級的空間分布是不清楚的。這導致我們考慮到模糊分級的空間系列的組織如屬性在更多本身的研究是有價值的。

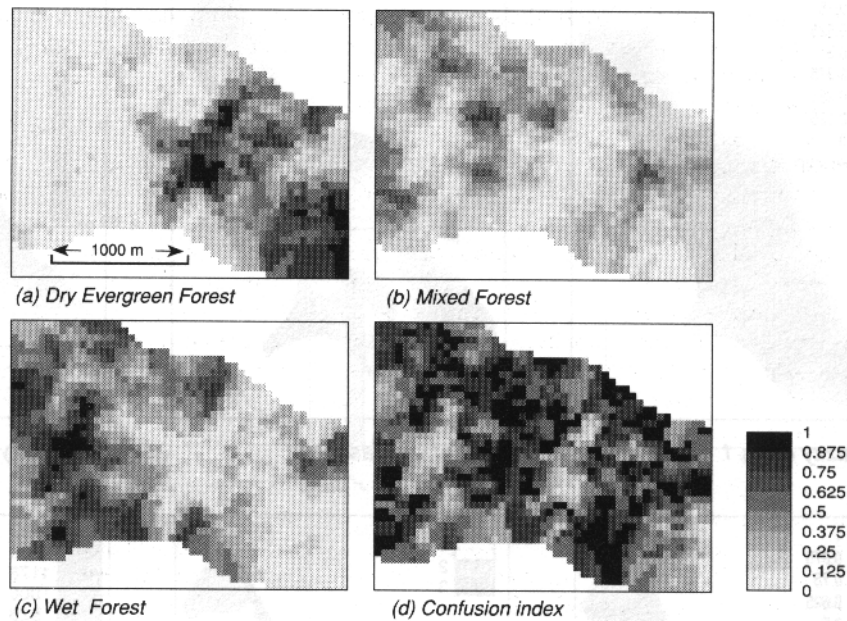


圖 11.16：模糊 k-means 於 Guyana 熱帶雨林樹種組成分類（原文 p.288）

## 疊合、融合和地理邊界(Class overlap, confusion, and geography boundaries)

熱帶雨林的資料證明本章開始所提到的現象，名為屬性空間的群集資料並不是必須在地理空間群集。在污染資料及雨林資料的模糊分級的 variograms 參數在氾濫過程的污染格局建立強大的空間相關是清楚的，但過程並不控制雨林林木的分布。這方法很簡單，沿著河流邊的污染區域的製圖是比森林林分製圖簡單很多。在後面的例子，簡單的獲得趨勢關於混亂的森林每個分級的中心區域是處於看起來是很多轉換區域的境地。

‘混亂的(confusion)’概念可能幫助我們結合在模糊隸屬的內插圖籍，使更簡單了解明確的區域。假如地位在模糊分級中隸屬值接近，它清楚的是屬於這個分級。但是假如 MF 值在兩個或是兩個以上的分級中是相似的，則其隸屬於哪一分類則不清楚。

這種情形是所謂的‘混亂’。讓我們定義屬性空間中每個觀測值、網格單元、或物件模糊指數的分級重疊程度；有兩個可能的形式

$$CI_1 = 1 - (MF_{\max} - MF_{\text{mx}2}) \quad 11.8$$

或是

$$CI_2 = MF_{\max} / MF_{\text{mx}2} \quad 11.9$$

使用 11.8 的方程式計算兩個資料的集合 CI 提供內插分級的一系列空間有趣的資料。氾濫平原的污染分級在大的 CI 產生適合的定義區域，大的 CI 是支配一個分級來取代另一分級。在這例子，大的 CI 區域是薄到可以簡單定義多邊形邊緣。影響的程序是計算這些單元具有的最大剖面及設法得到他們如 Boolean 資料形式。重新分類所有邊緣的單元會產生圖 11.15e 的地誌圖。包含最初氾濫頻率分級的三種製圖單元具顯著的相似性，是由具有氾濫頻率分層的 kriging 鋅的圖的圖 11.15e 的比較所提供(圖 11.15f)。很清楚的，當表面的連續隸屬函數是很強的一連串的事物，計算混亂指數描述給予的任何分級支配的區域是有用的方法。

在雨林資料上應用相同的方法顯示不同的分級在‘林木空間’於地理空間並不是很強烈的分離—確實這是值得考慮的重疊及少許的清楚‘邊緣’(圖 11.16d)。因為操縱程序，這些分級

的空間變異是比污染資料更弱。我們推斷出'林木分級的多邊形製圖是基於屬性的選擇，取樣的計劃在描述森林空間變異時將不是一個非常連續的方法。

### **混亂指標及 SI 方法(Confusion indices and the SI approach)**

儘管混亂指數的地理表示已經使用模糊 k-means 分析結果來表示，相同的方法可以使用在探索 SI 分級的空間表示。確實，由一個轉換區域寬度  $d_i$  的範圍重複 SI 分級，一是可以測驗地理位置及屬性空間分類中心。假如轉換區域寬度  $d_i$  是 0，它不是混亂的，因為它在屬性空間沒有重疊。假如重疊增加，則混亂將變得更廣泛。一個分級的中心面積將不會表示混亂的增加，如  $d$  增加：面積接近分級邊緣將是如此。因此，由計算混亂指數在  $d$  值的範圍(分類重疊)可以包含在屬性分級中心的地理位置圖，並可靠的預測分類重疊及邊緣的面積，及兩者在地理及屬性上的空間。結果假如每處都是低的 CI 的圖，我們可以推斷出描述是沒有用的。因此 CI 是一個 Perkal's epsilon 局部地轉換形式(第九章)。

### **討論：應用模糊分類的優、缺點(Discussion: the advantages, disadvantages, and applications of fuzzy classification)**

#### **優點(Advantages)**

本章表示的材料具有描述 SI 方法及模糊 k-means 分類在普通資料糾正及方類方法的優點。它顯示如何精確分類遺失資訊及當資料是由不精確的墮落時，增加分類錯誤的增顯，通常使用在環境資料的例子。它同樣顯示應用 SI 方法以精確地描述多邊形，可以提供他們的背景資訊，提供有價值的關於自然的急遽的或擴散的定義邊界的資訊。有一個研究中的文獻解釋 SI 方法於土地分類及土地評估的實用的優點 (Altman 1994, Davidson *et al* 1994, Wang *et al* 1990)，雖然在科學文獻應用的模糊 k-means 分類已經更被注意被承認：如土壤及環境科學見 Gaans *et al.*(1986), Vriend *et al.*(1988), Oden *et al.*(1990), Wang *et al.*(1990), Powell *et al.*(1991), McBratney and de Gruijter(1992), McBratney *et al.* (1992), Weiden *et al.*(1992), Gaans *et al.*(1992) and Frapporti *et al.*(1993), Burrough *et al.*(1997), Burrough and Frank(1996), Lagacherie *et al.*(1996), de Gruijter(1997)。

如和 SI 方法，重疊的分級及由模糊 k-means 產生的逐漸的邊緣比普通的方法看來好像更適合的真實。模糊隸屬值是由比明確分類獲得更多的最初資訊。(McBratney *et al.* 1992)。隸屬值可以在空間上簡單的內插，產生 variograms 分析及清楚的格局描述，假如所給予的屬性分級同樣具有一致的地理分布。

在對照重新定義的分類及分類邊緣的 SI 方法，模糊 k-means 方法產生局部最適分類是沒有必要基於線性的假設，不像普通的資料一轉換技術，如主成分及相似性分析。當線性相關可能在藉由適當的轉換函數的平均來增顯(對數、多項式等)，土壤、沈積物或是水的屬性通常是在很複雜的、固有地非線性方法共同修改。模糊 k-means 方法抓取這種非線性共變數 (de Gruijter and McBratney 1988, Oden *et al.*1990)。此外，連續及絕對的屬性資料可以簡單的合併(Oden *et al.*1990)。一般而言，分類技術於常態的屬性頻率分布的誤差是敏感的，包含有兩程式及極值(Vriend *et al.*1988)。因為模糊所允許，模糊 k-means 方法顯示由共變數至分離的群集幅度的終止的正確處理。

#### **缺點(Disadvantages)**



如多數參數方法，最大的困難是選擇控制參數的值以獲取最佳結果。以 SI 方法而言，使用者必須選擇隸屬函數、邊緣值、及轉換寬度的種類。模糊 k-means 方法則使用多屬性，在想像及評估適合的模糊群集的獲取是困難的。這有關 k 的選擇：分類數量；及 q：重疊或模糊允許的數量。進一步的困難是最適的模糊可能根據分類數或是 vice versa。使用特徵函數正式方法是可能的(Bezdek 1981；Roubens 1982)，但不總是成功的。科學上的洞察力或承諾對於存在的分類組合可能具有幫助(Vriend *et al.*1988；McBratney and de Gruijter1992；Frapporti *et al.*1993)。模糊 k-means 的其他問題是選擇屬性，是提供距離量測、量測模式程度及外形的模糊性和不同屬性的折衷辦法。它應該注意模糊性的程度在所有分級的假設總是相等的。

當模糊 k-means 程序可能看見在較少隨意外形的模糊邊界的結果，然後使用 SI 方法，這僅是一個問題因為這方法是很少正確。適當的製圖的技術來繪製 k 個分類(或 k+1 等級外)同時發生地持續發展的 SI 方法及模糊 k-means(但見 van den Wel and Hootamans 1993)。

### 結論—SI 或是模糊 k-means ? (Conclusions-SI or fuzzy k-means)

使用連續性分類的 SI 或模糊 k-means 法是根據問題所在及先前資訊的等級。在一適當定義及函數分類結構情形中，SI 方法提供較多的優點，但是模糊邊緣精確的定義需特別注意。這裡的研究報告顯示 SI 連續性分類比使用相同屬性的簡單的 Boolean 分類極端是更為健全及更少誤差。結果可以製圖顯示在資料空間及地理空間的屬性值逐漸的變化和隸屬函數的關係。

在模糊 k-means 方法，在基本分類中的準則區別是分析的結果而不是模式的輸入。這些準則可能於原始的屬性是複雜的非線性函數，因此介於隸屬值及屬性值間沒有清楚的相關。當具有定義屬性關係的物件，轉換一般都不是真實的，因為可能在屬性之中彼此補償。此後資料轉換是主要的理由之一，在使用分級關係代替最初屬性值，總是會面對某些資訊的遺失。

關於分級的數量及定義的資訊缺乏，模糊 k-means 方法是適當的。模糊 k-means 方法產生最適的集合、重疊分級可以同樣在資料空間及地理空間進行製圖。結果可以用在如(a) 在一個限制區域僅一次分析去定義(空間)變數的家系(lineage)(比較 Vriend *et al.*1988)，(b)如一個設定分級的方法，因此隸屬函數可以被內插，依次預測在未取樣位置的屬性值(McBratney *et al.* 1992)或是(c)如一個更好方法來定義簡單關係函數是為了用在直截了當地 SI 分析(圖 11.17)。在我們知識中，沒有一個是最近建立的例子，但是由 Burrough and Webster(1976)使用在數值分類及辨別分析類似的程序式設定及發展土壤調查的一個事先分類(Arces *et al.* 1975)

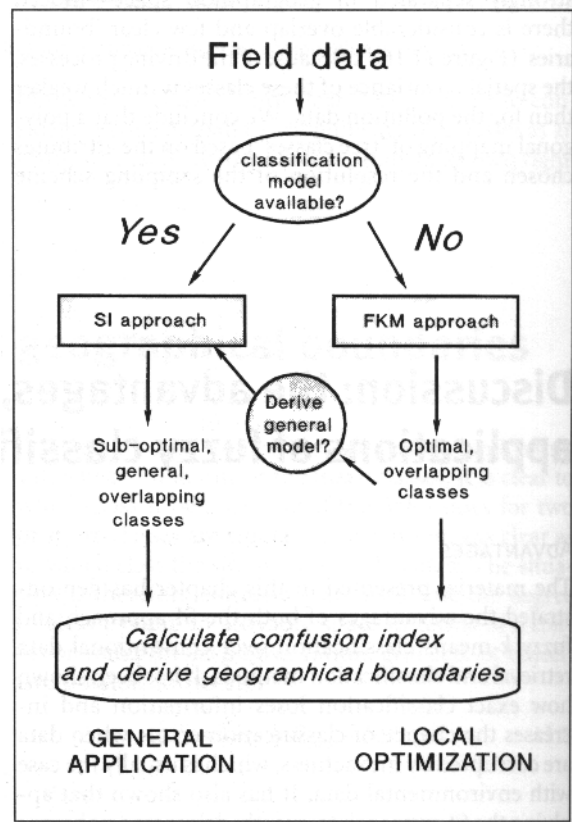


圖 11.17：使用 SI 或模糊分類指導模糊分類的流程圖（原文 p.290）

## 問題

- 1、列出 20 個不可由波動實體滿意的模擬的地理現象如第二章的解釋。建議這些現象的替代資料模式，並討論模糊集合方法在每個例子的值。
- 2、解釋你在練習中如何進行關於廣大地理邊界的量測。(a)在地景中(b)在城市中。
- 3、探討在連續的資料糾正操作及過濾製圖上使用模糊關係函數的值。
- 4、討論不確定性的或然性處理及模糊集合或然性方法的差異。何時是其中之一方法較為受到喜愛？

## 建議未來的閱讀

BARROW, J.D.(1992). Pi in the Sky. Oxford University Press, Oxford.

KANDEL, A.(1986). Fuzzy mathematical techniques with applications. Addison-Wesley, Reading, Mass

KAUFFMAN, A.(1975). Intorduction to the theory of fuzzy subsets. Academic Press, New York.

KLIR, G.J., and FOLGER, T.A.(1988). Fuzzy Sets, Uncertainty and information. Prentice Hall, Englewood Cliffs, NJ.

KOSKO, B.(1994). Fuzzy Thinking: The New Science of Fuzzy Logic. HarperCollins, London.

YAGER,R.R., and FILEV,D.P.(1994). Essential of Fuzzy Modeling and Control. John Wiley, New York.

## 目次 (第十一章)

不精確是一種思考方式.....	2
地理現象及不精確.....	3
模糊集合與模糊物件.....	3
明確集合.....	4
模糊集合.....	4
選擇基於屬性的分類.....	5
選擇關係函數.....	5
1. 顯示法.....	5
使用顯示法的適合隸屬函數.....	5
延伸隸屬函數的定義至適合中心概念範圍值的集合.....	6
非對稱隸屬函數.....	7
選擇轉換區域寬度的值.....	7
模糊地理物件.....	7
一個實用的例子.....	8
顯示法：重述要點.....	8
多種模糊集合的作業.....	8
聯合隸屬值的外插屬性.....	9
使用原始資料的模糊分類.....	10
使用連續性資料的模糊邏輯如流域網系統.....	11
使用布林運算和模糊邏輯中選擇所造成的誤差分析.....	12
資料糾正及分類上屬性誤差的影響.....	13
多邊形分界的標準方法.....	14
圖籍單位方法.....	15
個別邊界法.....	16
結合模糊界線和模糊屬性.....	17
重述要點：模糊多邊形邊界.....	18
選擇隸屬函數：模糊 $k$ -means.....	18
模糊 $k$ -means 及 Kriging 使用在重疊及多變數分類圖籍的例子.....	19
(a) 在 mass 泛濫平原的重金屬污染.....	19
(b) 熱帶雨林形式.....	21
疊合、融合和地理邊界.....	22
混亂指標及 SI 方法.....	23
討論：應用模糊分類的優、缺點.....	23
優點.....	23
缺點.....	23
結論—SI 或是模糊 $k$ -means?.....	24
問題.....	25
建議未來的閱讀.....	25

## 圖表目次 (第十一章)

Box 11.1: 計算布林及模糊圖籍的程序 (原文 p.272) .....	8
Box 11.2: 模糊集合的操作 (原文 p.274) .....	9
Box 11.3: 黏粒最大及最小的 MF 值的例子 (原文 p.275) .....	9
Box 11.4: 使用模糊 k-means 群集的演算法 (原文 p.285) .....	19
圖 11.1: 自然語言藉由再定義及層級劃分解決模糊分類 (原文 p.267) .....	3
圖 11.2: 左、模糊集合的影像; 右、布林集合的影像 (原文 p.269) .....	5
圖 11.3: 顯示方法的布林及模糊隸屬函數 (原文 p.270) .....	6
圖 11.4: 土壤結構的布林及模糊校正 (原文 p.272) .....	8
圖 11.5: 兩個重疊的模糊隸屬函數 (原文 p.253) .....	9
圖 11.6: (a, b)布林及模糊聯集的集合顯示在任一層黏粒的發生(c, d)布林及模糊的交集顯示黏粒發生於土壤剖面的情形 (原文 p.276) .....	10
圖 11.7: 對照使用常見(左)及模糊當量(右)的綜合土壤潛能分類的結果 (原文 p.276) .....	11
圖 11.8: 對於有興趣的資料層級使用模糊方法以連續性資料詳細保存及布林選擇損失的連結 (原文 p.277) .....	12
圖 11.9: 量測誤差影響布林分類(左)及模糊分類(右)的結果 (原文 p.279) .....	14
圖 11.10: 使用模糊隸屬函數描述邊界寬度 (原文 p.281) .....	16
圖 11.11: 使用模糊套疊計算屬性在模糊區域的平均權重 (原文 p.282) .....	17
圖 11.12: 傳統(左)及模糊(右)導出兩個具有逐漸、共享邊界的鄰接多邊形的土壤水分圖 (原文 p.282) .....	17
圖 11.13: 群集組成使用模糊 k-means 重複 (原文 p.284) .....	19
圖 11.14: 土壤污染資料的隸屬值的分散(Scatter)直方圖 (原文 p.286) .....	20
圖 11.15: 圖(a, b, c)三種土壤污染等級的模糊 k-means 隸屬值的內插圖; (d)混亂指標圖; (e)每個分級是優勢的波狀圖, 外框輪廓的虛線區分區域是用在第五及第六章以解釋內插; (f)鉍分級的圖籍包含和(e)相對的克利金分層 .....	21
圖 11.16: 模糊 k-means 於 Guyana 熱帶雨林樹種組成分類 (原文 p.288) .....	22
圖 11.17: 使用 SI 或模糊分類指導模糊分類的流程圖 (原文 p.290) .....	25
表 11.1: 圖籍單位的邊界寬度 (原文 p.282) .....	17
表 11.2: 模糊分級、氾濫污染的統計 (原文 p.286) .....	20
表 11.3: 模糊分級及熱帶雨林形式的統計 (原文 p.287) .....	21

## 第十二章 地理資訊系統目前狀況與未來趨勢

### *Current Issues and Trends in GIS*

#### 前言

GIS 是一個不斷在進步的技術，儘管在最近幾年他已經發展的相當成熟。他的成長受許多不同的力量所影響，包括一般資訊技術的進步，使 GIS 產業本身加入了極佳的轉變特性。目前諮商的範圍和研究乃在於開放 GIS (Open GIS)、相互間的可操作性 (Interoperability)、及空間數據的基礎結構 (infrastructure) 應該導出系統專屬性質的主要障礙及數據的交換問題是需要加以克服的。最近幾年，依據區域應用及國家使用，GIS 的使用者已經顯著的增加，但在數據使用、數據品質、及 GIS 使用上仍然存在著不同的問題。

在這本書中我們已經由 GIS 的基礎探究中，有了許多概念上、計算上、及數學上的空間數據分析原理。開始的時候，由地理學上人們察覺到的現象，建立空間模式的概念，進而展示在電腦中如何塑造不同實體和連續區域的確切數據模式，及如何能建立數化空間數據的資料庫。我們展示這些數據如何被恢復及分析，不同的空間分析方法如何由原始數據中導出新的資訊，及這些數據如何使用邏輯、數值的模式提供新的理解與見識，也測試新假說的成長。短時間內，我們已經見到電腦技術的集合，與地理數據所能提供使用者在自然、社會科學的分析上強而有利的工具，並大大的延伸了傳統製圖與影像的功能。

儘管本書仍然著重於討論 GIS 科目為 10 年前的主要項目，但第一次出版時 (Burrough 1986) 技術 (及研究區域的結合) 已經相當的成熟。例如對於大量地理數據的有效儲存技術發展，已經考慮至全部的空間資訊描述，但仍是一個持續研究的重要範圍。今天這個系統仍然在進步，且在最近幾年，數據的處理與產品的生產仍然隨著 GIS 的使用與被誰使用而改變。同時，大量的地理數據的供給與使用隨著增加，且其意識因重要的組織而成長。在許多研究中，對於人與環境間相互作用不斷了解的要求已經成為一個既成的動力，它需要處理大量的空間與位向數據 (Burrough 1996b, Schell 1995a, 1995b)。

這本書已經說了一些關於使用這些系統的人與所有在其週遭已經成長的產業。藉由 GIS 進行複雜的數據處理與分析正在增加，這表示他們被使用在公共發展、訂定決策、及在公共、經濟及藝術設置研究領域。這些系統較以前被更多的使用者所採用，且這個技術具有相當大的國際市場。GIS 在已開發國家的單獨領域並不長，許多開發中國家使用 GIS 在許多的應用上，包括地籍圖的製圖與資源分析。GIS 的使用與發展已被資訊的結構所支持，包括專門的書籍、記錄及雜誌、討論會與展覽，及被專業所認可的新理論、技術，且其應用與發展已經擴展至相當廣泛的使用。

在以後幾年 GIS 的改變將會被以下所支配 (a) 資訊技術的發展，(b) 不同的國家與產業標準化的協議，及 (c) 數據可用性的擴展與聯合組織的規定。在 GIS 的使用者團體與可見的技術上，這些都將會帶來轉變，以下將主要的趨勢作一總結。

#### 技術改變(Changes in technology)

十年前，商用、現貨供應 (off-the-shelf) 的 GIS 系統是一個完全統一的系統，其具有許多的使用指南，可展現不同的函數與性能，在這鍛鍊的時期中，他經常被人們使用。今天

許多 GIS 工具是不同的且專業的輔助系統所構成，例如製圖工具、資料庫的使用與回復、電子數據製表及空間數據分析工具、地理統計學插入推估方法、遙感探測圖像處理、或計算模式，這些被整合為一個較大或次要的程度。主要的數據儲存模式及處理工具，其廣泛的集成已經被網路通用的標準，成功的將其轉換為一個標準化模式，例如大型關聯數據處理系統（large relational database management systems）。非專業人士，如經營者及行政管理人員，現在也在他們的工作崗位上使用空間數據，不必藉由中間處理工具處理先前複雜的系統，換句話說，在具高度技術的建立 GIS 專業人士及具備複雜的空間分析技術與數量龐大的使用者皆在成長之中。

由於 IT 產業下快速且相對較經濟的電腦處理過程，支持了 GIS 數據處理的需要，使得 GIS 在發展上有相當的助益。個人電腦及網路已將 GIS 由一個背後支援（back-room）“不親和（techy）”的周邊活動轉為桌面管理系統的應用。潮流的趨勢朝向 64-bit 的電腦處理器，其將可以允許更快的數據處理及更大的記憶使用，甚至可以同時處理正在發生的事件。無庸置疑的，這樣的發展允許 GIS 的使用者作更複雜的、缺乏數據（data hungry）的業務，且更加迅速，而在未來，這系統中將有更多關鍵性的轉變。

直至最近，多數的 GIS 被設定為區域或單一的主要使用者所使用，或是在有限的範圍內工作。長久以來，地理資訊的蒐集、儲存、分析、呈現及散播受限於單一的學會、政府組織或事務單位、不同的軟硬體、及稍許重要的數據標準化（包括數據模式、數據安裝啟用、交流標準等）。對於個別的應用而言，'最佳的'系統取決於其審查的功能性、精度、速度及執行所需的價格與使用者需求間的關係。GIS 是被“基準測試（benchmark tests）”所估價的，如同他們在工業販賣部沒有比機械多，其賣主以不公開運算法則、數據結構及過程保護他們的利益，所以每個系統間仍然存在著很大的中斷（Burrough 1996b, Burrough and Masser 1997）。

沒有什麼好驚訝的，對於地理資訊團體的成員而言，當前的兩個主題為開放 GIS（Open GIS）及相互可操作性（Schell 1995a, 1995b），他們緊密的相關且需要建立系統以支援有效的描述、儲存、使用、及地理資訊數據的移轉成為組織、國家、甚至全球化。這些發展將允許更多的數據可被推展為具有更大量的使用者，促進地理資訊系統在解決問題上的使用，並在數據的攫取上導領出相當有效的使用。

許多國家透過政府運作（例如美國聯邦數據資料委員會 the US Federal Data Committee 及國家空間數據創始機構 the National Spatial Data Initiative--NSDI）支持開放 GIS（Open GIS）想法的發展，在國際方面另有透過組織的合作，例如 EUROGI（歐洲地理資訊聯合組織 European Umbrella Organization for Geographic Information）及商業事業（the Open GIS Consortium），其將允許 GI 的使用者獲得可進行高效轉換及共享於不同軟體的資料。OGC 關心的便是為了交換數據所發展中的技術，並連結更大的 IT 企業，所以正需要克服擁有系統數據資料的問題。透過組織的支持，數據及 GIS 的標準與技術將被真實的轉移。

相互可操作性目標的發展是為了移除當前正在使用的特殊硬體，與為特殊數據或作業的軟體的系統規定參數。例如藉由設定 GIS 資料庫的標準，不同賣主的軟體能在許多不同的電腦中用作查詢數據的基礎程序。許多目前的商用系統可以進行輸入，或至少可以展示，在地理數據廣泛的形式範圍。數據轉換的問題之後可透過技術的領域解決—我能夠讀取你的磁片，或你能夠輸入我發送在网络上的數據？—語意上的專業，這不只關係到每一個數據軟件，更包括他們如何被察覺、記錄及塑造，這確保其可以數據的形式被讀取是察覺不到閒置

的。

儘管網際網路與世界寬頻網路 (World Wide Web) 的動態、無法控制的性質使得趨勢有點困難去預知，其將持續影響 GIS 的使用及系統的發展。無庸置疑的，數據及軟體的可用性與他們的普及將作為繼續成長的重要資訊來源。在他們擁有的操作權限中，網路也似乎變成一個重要的 GIS'操作環境 (operating environment)' 並伴隨著更多動態交互影響工具的發展。編輯語言的開發，例如 JAVA (of Sun Microsystems)，在這樣的背景下，他們藉著主要網路瀏覽操作套裝軟體自動的活動並允許寫入的應用 (已知有 Java applets) 是相當有用的。這些網路語言 applets 是很容易在網路伺服器下載於區域的個人電腦中，所以確保藉著網際網路更快的交互作用。JAVA 語言的一個真正的利益是程序編碼 (program code) 直接可轉移為不同的電腦 platforms 以運轉於 UNIX、Macintosh、Windows 95、Windows NT 等系統之下。這將使得網路開發者免去為不同的 platform 需要寫不同的應用程序 application programs。

目前 GIS 在處理靜態及地理空間中容易確認的單元是相當成功的，而更複雜的數據組成與處理，特別是一些暫時 (temporal) 的元件，會繼續向 GIS 的開發者挑戰。在最近一些年，數據庫研究已經具優勢，藉由物件導向 object-oriented 的發展過程，對實體數據的結構化有利 (特別是簡單的定義原件，例如所有權的部份 property parcels，及工具的基礎組織 (utility infrastructure)、提出可能的時空暫時 (spatio-temporal) 的數據處理)，但應用他去連續也許不存在固定關係的區域數據仍有困難。

更多的機動性和效率也許能夠經由系統允許多尺度的 (multi-scale) 表示方法，使用網槽狀層級 (nested hierarchical) 過程在數據儲存與分析。這些支援的概括變化度與數據的抽象化似乎因為流行而受益，使我們更容易瞭解地理數據的空間尺度。邏輯基礎 (logic-based) 的數據庫發展也提供描述數據更直覺方式的潛力，但這仍是一個活耀的研究領域 (Worboys 1995)。

三度空間的 GIS 已經逐漸變成通用的 (cf. Plates 3.3, 3.4)，但表現地理資訊暫時性質的問題仍繼續存在為主要需克服障礙。目前許多 GIS 資料庫結構不只能夠表現靜態的 2D 資訊，並且可以一個區域內不同的資料群組系列或多暫時 (multi-temporal) 屬性表現暫時的變化。這導出一個數量的問題，包括冗長的資料庫及醜陋的集成在不同資料組的任何分析中 (Langran 1992)。從目前的發展與研究報告中明確的指出，在地理現象的塑造中，他仍然相當需要理論和技術的開發。

現在許多個人電腦 (human-computer) 的 GIS 界面包括在重複的作業中寫入高階的程序 (巨集指令 macros)。藉著努力有時他們被用作時空上的塑造，但通常在使用上是困難的。最近網格 GIS (Mitasova *et al.* 1996, van Deursen and Burrough 1998) 發展正提供有力且容易的動態塑造工具，他們本身具有高階的程序設計語言。

## 資料提供的改變(Changes in data supply)

GIS 資料在 GIS 採用和運用上的限制已是多年來的問題，雖然在近五年來主要的趨勢之一已在他們的預備下成長。當地理資訊產業發展，使用者瞭解到如果他們可以買賣空間資料，那麼節約成本是可以做到的：因為如果其它的機構已有你所需要的資料數位化說明，為什麼還要付出時間和金錢去重做這個工作呢？實際上，因為資料要轉換成有效的格式去提供使用，使得資料的轉換有時候反而較原始的數化花費來的更多，這在第四章已有討論到。直

到最近，很多政府，以及其它去執行資料轉換和資料品質標準的機構，在地理資料中破碎的市場裡，他們的方法已在大部分的案例中為特別的目的，以及確定政策決定的連結已朝向一個較少的控制。資料的使用者並不總是確定資訊的正確度（accuracy）和精密度（precision），以及資料在開始準備時透過地球網際網路所得到的資料來源，這些從不同來源所得到的資料，以不同品質、蒐集標準、地理相關研究等的觀點來看，或許已在技術上輕易地連結，但這些資料可能是不相容的，導致似是而非的分析結果。

不同新的創始（initiatives）是這些問題的開端，開放 GIS（Open GIS）的發展及資料互換的牽涉將被討論，同時各種空間資料的創始，在國家、大陸性的和全球性的尺度下，藉著為有條理的地理資料的互換和區分提出協議、指導方針、規則及政策，嘗試減低資料有效性及整合的問題。這些創始也包含著制定大型、一般實質、空間的資料庫，例如美國和澳洲等國家已準備好要進行基礎建設的國家空間資料。也有一些大陸性和國際的創始（例如 CORINE 1992, EUROSTAT 1996, UNEP/FAO 1994, Digital Chart of the World 1991）嘗試要從很多來源去蒐集資料，並使這些資料在幾何學和分類上是在共同層次（levels）上的（RIVM 1994）。在資料說明（metadata）（data about data 有關資料的資料）標準的發展，包括了有關來源、紀錄方法、解釋、資料蒐集、資料模式和資料型式的資訊，如果資料提供對於他們的目的是適合且可靠的，為了使用者的判定，他們正製作成更為簡單的資訊。

如果機構和環境的問題對於今日某些國家是所要面對、考量的，那麼這些發展是非常重要的。很多情況是源自於在一個國家中影響生活、和人們生活中自然資源的自然、產業、經濟和政治活動的結果，並不被行政的邊界（不論是區域或國家的）所限定的活動和意外。一個特別而深切，對於環境災害的國際性規模證明，發生在 1986 年的歐洲，因車諾比意外而漫遊在北半球的放射性殘留的雲，無視於國界和政府的命令，沈澱這些放射性殘留物於距來源地遙遠的地區。（Karaoglou *et al.* 1996）。這一點和在人類衝突的區域，不受控制的避難活動問題，需要透過定性、定量的分析，運用整合的空間資料組（datasets）去提出，這個空間組應包含所有影響區域，而不只在於一個特別的國家或省而已。

越過數個行政管理區域，例如國家或省，聚集一定的地理資料庫的困難及問題是 legion（c.f. Langas 1997）。地理研究的資料可能引用不同的座標系統、基本層（base levels）或地圖投影。屬性資料可能是使用不同取樣方法或樣本大小去蒐集而來的。實驗室方法如果是用來分析土壤、水、空氣或生物的物質，在不同的國家可能是不同的，或者由於地區的專門技術知識的不同而使得誤差的變化程度不同。在特別的社會資料或土地使用資料的某些資料可能會關連了被不同定義且有不同空間資源的行政管理單元。（Burrough 1996b, Burrough and Masser 1997）

除了有像資料所有權、法律責任和標準等的問題外，並沒有提及無數在政治和制度上所需要解決的爭議，對於標準資料蒐集和相關研究的創始，例如：判定地理位置、或高層資料、或自然物件（如：河、海岸、和湖）位置的主題資料，以及人為的特徵物（如：路、鐵路、城鎮和都市），這些常常會碰到在保衛家園領土的阻力。

對於產業的蒐集成本、儲存和互換資料遠超過資料開發而言，各種的發展自節約金錢和改善效率的觀點來看是重要的。（Burrough and Masser 1997）。當地理資訊的影響市場被開發，有時使得複雜和豐富大量的資料更寬廣的公開，且使決定獲得改善，更重要的，改善了我們對地球表面現象空間分類的知識和瞭解。

資料整合和轉換與語意有關的問題比剩下的技術層面更難去處理，問題首先是文化、語



言和知覺，擴大到空間和時間分析的範圍。接著，更具影響力的理論概念是需要更深入到目前商業的 GIS：這是 GI 科學的領域（Burrough 1996b, Burrough and Masser 1997）。

## 使用者的改變(Changes in the users)

GIS 使用者總是通過一個數量的訓練而被擴展，正反映出地理資料寬廣的應用。其系統已使用於動物、礦物及植物空間分佈的研究，就如同產業位置、選舉格局以及古老的與現代人類居所的研究一般良好。直到最近，大部分的 GIS 限制其範圍在空間或區域覆蓋的研究。其系統的趨勢在於精通軟體與重要技術、數據資料庫的專業者的使用。其產生的結果對 GIS 分析來說，通常轉換成文件的形式，以由大群人們廣泛地使用於其工作。

GIS 技術的趨勢致力於一系列專家子產品 subproducts，以增加數化資料的供應帶來關於使用者的改變。例如空間塑造（spatial modelling），在某些地區 GIS 繼續使用專家（有些也許說是傑出人才 elitist），然而，更專業的模組發展提供查詢與製圖能力，供應新的設定，使得使用者不必要知 GIS 技術為何。這容易使用（easy-to-use）的界面允許使用者（對資料有所了解的）與系統利用電子軟體直接相互作用。這代表他們不必要藉由第二者的幫助傳送他們的分析。區域網路（Intranet）的發展也提供組織一個供應地理資料與系統能力的有效方式，而此方式對於廣泛的使用者而言是最為親切的方法。對許多現今網際網路/區域網路的資料與技術的使用者來說，GIS 基本上是數化地圖與地名索引。這些使用者不一定要收集資料與繪製地圖，他們基本上只要一個新的使用者型式—“空間流覽器（spatial browser）”。

在使用團體中一個顯著的改變是 GIS 專業人才數量的增加。他們提供顧問與建議給使用者關於購買、維持、與使用系統以及了解 GIS 的觀念而非只是特定的系統類型。這些專業工作對於一個不同公司中各個部分的不同其組織顧問系統亦有所不同。某些 GIS 的專家是對特定區域有專業的知識如水文學、或市場潛力以及精通使用系統解決分析問題。各式 GIS 專家組織已經建立一個團體以提供為使用者與系統賣主相互作用的討論會。

GIS 使用團體現在有比網際網路發展更為寬廣世界性的趨勢。他們的應用在更廣泛的區域與項目，和透過更有趣的不同文化。這些系統的使用遇到了資料獲得及重要學會及社會制度變化程度，其變化程度比技術比身更具影響力。GIS 在組織上影響已經有一個數量的研究（如 Huxhold & Levinsohn 1995），且最近更多技術型式在組織思考影響已被研究（Bijker 等 1987）。

這個概念在 GIS 技術上確信的種類中，被具體的反應在世界上一些部分的人們所適應。這些想法的力量明顯地來自於這些想法被消化吸收的速度（如本書第一版本譯成日文、中文、泰文、及每個國家的 ARC-INFO）。取代提供人文的豐富度，帶來系統中不同想法的人文邊界的大量發展代表了可接受實際與應用無區分的有意義系統（Burrough 1996b）。技術（Technology），它包含的進步與發展的概念，通常見於一些必需一直被遵循的事物。

GIS 是用來解決問題的重要工具，其問題需要操作與地理資料的分析。許多今日提供實用的、社會的及環境改善共同的目標以及要求空間知識過程與形式的了解。在 Rio 的地球高峰會議中認為這應用應由政府提供許多活動包括反映空間資料與地理資訊在各個領域如土壤沖蝕、海洋污染、生物多樣性及畜牧、飲用水供應、大氣變遷、貧乏、都市化或人口統計變化。GIS 將持續提供答案但其成功是使用將依賴問題的提出、資料的獲得以及制度與政策的支持必然帶來關於任何的改變，在技能與洞察的過程。

## 未來閱讀的建議(Suggestions for further reading)

- BIJKER, W. E., HUGHES, T. P., and PINCH, T. I. (eds.) (1987) . The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology. MIT Press, Cambridge.
- BURROUGH, P. A. (1996) . A European view of Global Spatial Data Infrastructure. Proc. Emerging Global Spatial Data Infrastructure. A Conference held under the patronage of Dr Martin Bangemann, European Commission, EUROGI/Deutscher Dachverband fiir Geoinformation/Atlantic Institute/Open GIS Consortiuml Federal Data Committee/Federation Internationale des Geometres (Commission 3) , 4-6 Sept. 1996, Kijnigswinter, Germany.
- BURROUGH, P. A., and MASSER, F. I. (1997) . European Geographic Information Infrastructures: Opportunities and Pitfalls. Taylor and Francis, London.
- LANGRAN, G. (1992) . Time in Geographic Information Systems. Taylor & Francis, London.
- MOUNSEY, H. M. (1991) . Multisource, multinational environmental GIS: lessons learnt from CORINE. In D. J. MAGUIRE, M. F. GOODCHILD, and D. W. RHIND (eds.) Geographical Information System, ii: Applications, Longman Scientific and Technical, Harrow, pp. 185-200.
- SCHELL, D. (1995a) . What is the meaning of standards consortia! GIS World (August 1995) , 82
- (1995b) . Harnessing change. Editorial in Open GIS, Newsletter included in Geo Info Systems, May 1995.

## 目次 (第十二章)

技術改變.....	1
資料提供的改變.....	3
使用者的改變.....	5
未來閱讀的建議.....	6