

空間離散指標：

舊觀念、新公式

Spatial Dispersion Index: Old Conception, New Formula

翁培文*

蔡博文**

Pei-Wen Weng

Bor-Wen Tsai

摘 要

地物分布的集散情形是空間現象重要的研究議題，部份度量集散的指標是以空間自相關為基礎，如 Moran's I 和 Getis's G，部份指標則以其它方式加以權衡，如最近鄰法以最近鄰距離為主要依據，但這些指標都有一些侷限，因此，本研究嘗試建立一個新的指標：空間離散指標 (Spatial Dispersion Index, SDI)，該指標承襲統計學的離散觀念與公式，以統計單元大小、地物相對大小、鄰近程度高低為判定離散程度的三個原則，並以此來推導 SDI 的公式，該指標能應用在點資料與面資料的計算，但其最大值和數值分佈常因例而異，在判定 SDI 的集散意義時，部份方式計算較為複雜。

關鍵字：空間離散指標、地理統計、集中、分散

Abstract

One of the controversial issues in studying a spatial phenomenon is whether its distribution pattern is aggregated or dispersed. Some of the indexes are based on spatial autocorrelation, such as Moran's I and Getis's G. Some are not, such as Nearest Neighbor Index is calculated with nearest neighbor distance. However, all of them have some limitations. The present research tries to develop a new index:

* 國立臺灣大學地理環境資源學系碩士
Master, Department of Geography, National Taiwan University.

** 國立臺灣大學地理環境資源學系助理教授
Assistant Professor, Department of Geography, National Taiwan University.



Spatial Dispersion Index (SDI), following the concept and formula of dispersion from statistics. Three principles of SDI are adhered to: the area of statistical unit, comparative area of surface features and distance to neighbor. The SDI formula is deduced from those three principles. This index can calculate point and polygon features. But the SDI maximum value and value distribution will change by case. Estimating whether the SDI represents an aggregative or dispersive phenomenon, the calculations are more complex than other indexes.

Keyword: SDI, geostatistics, aggregative, dispersive.

前 言

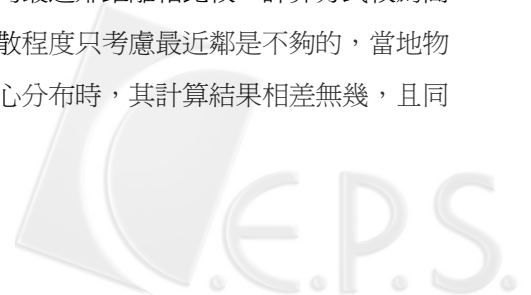
地理資訊系統 (Geographic Information System, GIS) 多用於空間資料的處理, 空間資料管理、空間資料整合、空間分析、視覺化 (Visualization)、地圖輸出等皆為 GIS 重要的功能, 而空間分析可說是其中最重要的功能之一, 其發展得力於地理統計 (Geostatistics) 頗多。地物分布的集散是地理統計重要的研究議題 (本文地物泛指具空間屬性的人、事、物等現象), 「在早期的土地利用研究中, 研究者以文字性的描述來展現土地利用的空間分佈, 隨著科技的進步, 計算工具及研究方法的不斷創新, 描述土地利用空間分佈的方式也從文字性的描述逐漸演變為以數學方式建立空間指標來描述空間分佈情形」(朱健銘, 2000:1), 集中、集中有散、散中有集、分散的文字描述無法解決許多問題, 集中是 10 個地物的集中? 還是 100 個地物的集中? 集中有散是 90%集中、10%分散? 還是 60%集中、40%分散? 即便是 90%集中、10%分散, 這 10%的分散, 是在 90%集中的周圍? 還是離 90%集中有一段距離? 這些問題皆不是文字描述所能處理, 使用數字描述除了使度量更加精確, 也讓空間分布的描述有了更多層次, 對集散現象的後續解讀則提供了更強的數學依據, 因此, 量化指標的建立確有其需求與必要。

文獻回顧

空間資料之分析與傳統的統計分析主要有兩大差異, 其中之一是空間資料在 d 維空間中具有某種空間相

關性, 也因為空間資料具有這項特點, 部份度量空間集散程度的指標是以空間自相關 (Spatial Autocorrelation) 為基礎, 空間自相關是指地物在空間上的分布具有空間相依性 (Spatial Dependence), 即距離近者性質相近, 距離遠者差異較大, 此觀念所衍生的指標可分為全域空間自相關 (Global Spatial Autocorrelation) 和區域空間自相關 (Local Spatial Autocorrelation), 前者以 Moran's I 為代表, 後者以 Getis's G 為代表, 相關的研究頗多, 有些較偏重學理的計算, 如 Jong (1984) 等人探討 Moran's I 的極值 (Extreme Value), 有些與實例研究結合, 如 Getis and Ord (1995) 將 Getis's G 應用在加州愛滋病的研究, 但二種指標或多或少都有一些侷限, Moran's I 和 Getis's G 對於地物的「存在」與「不存在」皆予以同等的重視, 但不少研究課題僅對地物的存在感到興趣, 對地物的不存在不感興趣, 以聚落的集散為例, 研究者通常關注的僅是建物的集散, 而 Moran's I 和 Getis's G 都會將非建物一併納入計算, 計算結果無法全然反應出建物的集散情形; Moran's I 以二空間單元相鄰與否做為公式中的一項參數, 但如此僅考慮到相鄰, 並沒有考慮到遠近, Getis's G 可用距離輔助判斷地物關聯, 但這樣的修正方式忽略了距離所形成的範圍內外, 皆仍有長短距離的差別。(黃信誠, 2000; Jong 等, 1984; Getis and Ord, 1992, 1995; 朱健銘, 2000)

最近鄰法 (Nearest Neighbor Index) 計算一個地物與最近地物的距離, 累計所有地物的最近鄰距離取平均值, 再與理論平均最近鄰距離相比較, 計算方式較為簡便, 但地物的集散程度只考慮最近鄰是不夠的, 當地物呈單核心與雙核心分布時, 其計算結果相差無幾, 且同



樣呈雙核心分布，二核心相距近或遠，計算結果亦相差無幾；方位角法 (Sector Method) 改良最近鄰法取單一最近鄰的缺點，對數個方位角取最近鄰，但仍難妥善修正其缺點。方格法 (Quadrat Index) 將統計單元切成網格，計算各網格的地物數 (或面積) 並與期望值相比較，再用 χ^2 檢定其分布情形，這種計算方式只考慮到網格內的情形，並沒有考慮到網格內與網格外的關聯，若單核心的地物平均分布在 n 個網格，雙核心的地物亦平均分布在 n 個網格，其計算結果並無法適切的反應出二者的差異。部份景觀生態學 (Landscape Ecology) 重要的指標可用來計算集散程度，景觀聚集度指數 (Contagion Index) 是用 i 綴塊 (Patch) 與 j 綴塊相鄰的機率來表示綴塊的聚集程度，但如此只考慮到相鄰的綴塊，並沒有考慮到不相鄰綴塊的集散情形；碎形 (Fractal, 或譯為分維、分形) 為不規則幾何形狀的非整數維度，主要是用來計測型態，亦可應用於度量集散程度，盒計數法 (Box Counting Method) 是以網格邊長與網格數的變化來描述地物的叢聚 (Cluster) 程度，但這種方法較針對個別的地物，對於地物間距離的偵測能力頗差。有部份指標導入統計學的觀念，平均距離 (Mean Distance) 將各地物的距離加總後平均，忽略了統計單元大小的影響，標準距離 (Standard Distance) 則是以全部地物的重心與各地物的距離來計算標準差，但忽略了同重心距離存在著有不同集散程度的可能。(梁蕪善, 1985; 鄔建國, 2003; Farina, 1998; 李介中, 2005; 邱逸民, 2001; Haining 2003)

以上針對幾種較重要來度量集散的指標進行初步的探討，這些指標或多或少都存在著一些侷限，其中大部份對地物間距離的偵測能力較弱，因此，若能改進此點，並發展出一個更適宜的指標，將有助於空間現象的描述及後續研究的進行。

SDI 推導

李介中 (2005) 認為，定義空間分布時應遵循的原

則有三：以統計單元的整體範圍做為尺度依據、判斷地物分布的範圍 (相對統計單元大小)、區分地物彼此間的鄰近程度，茲分述如下：

1. 以統計單元的整體範圍做為尺度依據是指，定義空間分布的集散須考慮到統計單元的大小，以圖 1 為例，二圖所要偵測的地物距離相同、大小相同，但統計單元不同，左圖較均勻的分布在整個統計單元，右圖較偏於統計單元一側，因此，右圖地物較為集中，而左圖較為分散。

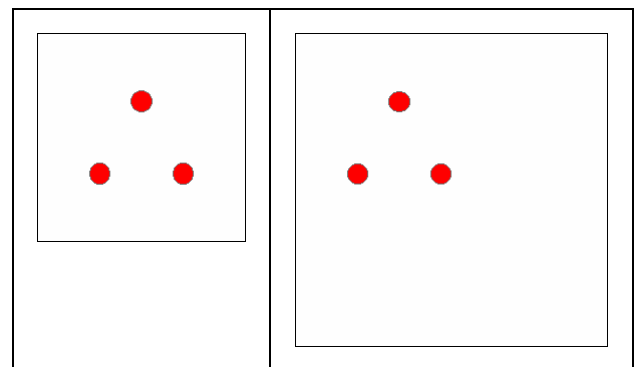


圖 1 統計單元大小影響集散示意圖

2. 判斷地物分布的範圍原意是指，定義空間分布的集散須考慮到地物分布的範圍，但相同的分布範圍，範圍內仍有集散程度之別，故此原則不予採用。以圖 2 為例，二圖所要偵測的地物距離相同 (就重心而言)、統計單元相同，右圖有一個地物較大，佔了地物面積的大部份，左圖三個地物等大，地物面積平均分散於三處，大面積地物可視為許多小面積地物的集合，地物面積越大表示越多地物的集合，面積越小表示越少地物的集合，面積較大比例的集中表示集中程度較高，較小比例的集中表示集中程度較低，因此，右圖地物較為集中，而左圖較為分散 (亦可解釋為集中於一處較分散於三處來的集中)。這種集散現象難以就地物自身的絕對大小加以研判，但可由不同地物的相對大小加以界定，故以「地物相對大小」為判定集散程度的第二原則。



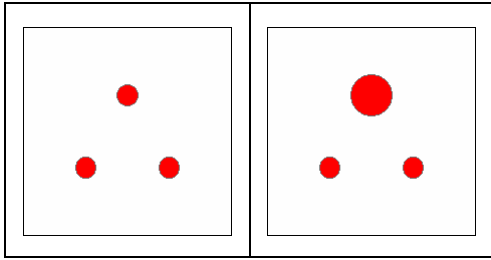


圖 2 地物相對大小影響集散示意圖

3. 區分地物彼此間的鄰近程度是指，定義空間分布的集散須考慮到地物距離，以圖 3 為例，二圖所要偵測的地物大小相同、統計單元相同，但地物距離不同，左圖地物距離較短，右圖地物距離較長，因此，右圖地物較為分散，而左圖較為集中。

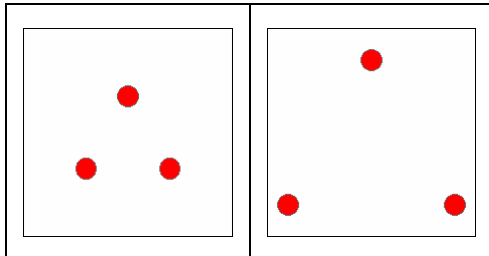


圖 3 鄰近程度高低影響集散示意圖

本研究承襲李介中 (2005) 第一和第三原則，對第二原則採不同的看法，因此，改以統計單元大小、地物相對大小、鄰近程度高低為度量集散程度的三個原則。三個原則若自成一獨立指標，則二案例在三個指標中互有集散時，比較上難免會有一些困難，因此，最好的方式是將三個原則整合成一個指標，各個原則在指標的公式中獲得相對應的權重，由計算結果即可判定不同案例的離散程度。

統計學已發展出多種計算數列離散程度的指標，如：離均差、離均差平方和、標準差等，上述三種方式皆是以「離均」程度來表示，這些指標亦可用來度量空間分布，唯二者空間維度不同，一維數列的計算方式應用二維圖形須做適當的轉換，才能將數列離均的觀念應用在空間分布的量測。圖形轉換為數列的方式是將地物面積轉換為分組次數，地物距離轉換為二組的差距，以圖 4 為例，二地物的面積分別為 2 和 5，距離為 10，若

以左方地物為原點，左方地物數值為 0，地物面積 2 轉換為分組次數 2，右方地物距離左方地物 10，可將右方地物轉換為數值 10，地物面積 5 轉換為分組次數 5，因此，圖 4 可轉換為 0 的分組次數 2、10 的分組次數 5，也就是 0、0、10、10、10、10、10，若改以右方地物為原點，可將圖形轉換為 0、0、0、0、0、10、10 (0 的分組次數 5、距離 10、10 的分組次數 2)，二數列的離均差、離均差平方和、標準差皆相同，以任一地物為原點其計算結果並無差異。一般而言，數列的離散程度較常以標準差來表示，標準差會使離均較大的數值獲得較大的影響力，地物的離散程度較不須如此權重，故計算式採離均差而非標準差。

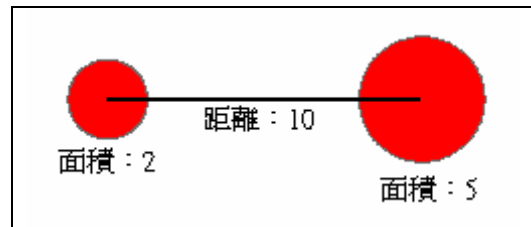


圖 4 二地物面積、距離示意圖

$$\text{公式爲：}\bar{d}_{(ij)} = \frac{a_i * d_i + a_j * d_j}{a_i + a_j} \dots\dots\dots\text{公式 (1)}$$

$\bar{d}_{(ij)}$ ： a_i 、 a_j 的離散程度(離均差)

a_i 、 a_j ： i 、 j 地物的面積

d_i 、 d_j ： a_i 、 a_j 與二地物重心的距離

a_i 與 a_j 二地物不相鄰，若統計單元內僅有一個地物， $\bar{d}_{(ij)}$ 判定為 0。

在二地物距離不變、面積改變的情況下，一地物面積越大，另一地物面積越小 (數值越集中)， $\bar{d}_{(ij)}$ 越小，離散程度越低，若兩地物面積越接近 (數值越不集中)， $\bar{d}_{(ij)}$ 越大，離散程度越高 (原則 2)；在二地物面積不變、距離改變的情況下，地物距離越短 (數值差異越小)， $\bar{d}_{(ij)}$ 越小，離散程度較低，若地物距離越長 (數值差異越大)， $\bar{d}_{(ij)}$ 越大，離散程度越高 (原則 3)；在地物大小、鄰近程度皆不同的情況下，離散程度隨 $\bar{d}_{(ij)}$ 的大小，也就是隨離均差的大小而改變。

點資料地物間的距離較為單純，面資料距離的種類較為多樣，有最短距離、最長距離、重心距離和平均距離等，最短距離和最長距離都是極端值，使用上較具爭議，若將二者平均，則忽略了最短距離與最長距離間的種種可能；重心距離是較常使用的一種，若地物分布呈同心圓，內圈與外圈既不相鄰，也不重疊，但重心距離為 0；平均距離的計算較為繁瑣，計算方式是將二個地物切成若干網格，不同地物的網格計算距離後取平均值，但平均距離會依網格的基點、網格的大小等而改變。本研究採重心距離，理由是絕大多數情形適用且計算方便。

空間離散程度的度量有時不能以離均來處理，以圖 5 為例，二圖地物面積相同、距離相同，若用離均的方式以所有地物的重心對各地物進行度量，二圖的離散程度相同，但二圖地物間的距離並不相同。

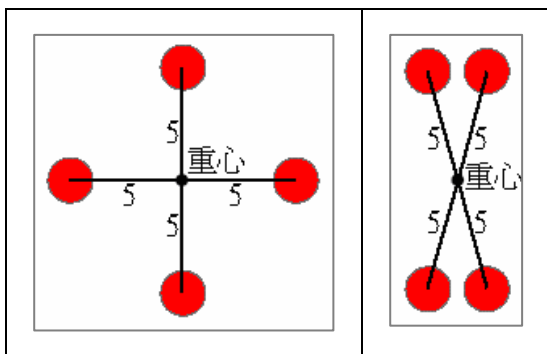


圖 5 相同離均、不同離散程度示意圖

本研究度量的方式是先以二個地物為一組，計算二地物的離散程度，再把這樣的計算過程擴及到所有地物不重複排列的可能，如此便能將較複雜的二維空間轉換為較單純的一維數列，避免二維空間離均的缺點，並具有應用一維數列離均的優點，以圖 6 為例， a_1 分別對 $a_2 \sim a_n$ 計算 $\bar{d}_{(1,j)}$ ， a_2 分別對 $a_1, a_3 \sim a_n$ 計算 $\bar{d}_{(2,j)}$ ，……，最後是 a_n 分別對 $a_1 \sim a_{(n-1)}$ 計算 $\bar{d}_{(n,j)}$ 。 a_1 對 a_2 計算過一次， a_2 也對 a_1 計算過一次，一組地物被等量的計算二次，只要除以 2，與 $a_1、a_2$ 僅計算一次的結果相同。 a_1 對其它 $n-1$ 個地物計算 $\bar{d}_{(1,j)}$ ，共計算 $n-1$ 次，其它地物亦對 a_1 計算 $n-1$ 次， a_1 總共計算 $2(n-1)$ 次。考量重複計算的影響，

公式修正為：

$$\frac{1}{2(n-1)} \sum_{i=1}^n \sum_{j=1}^n \bar{d}_{(i,j)}, i \neq j \dots\dots\dots \text{公式 (2)}$$

n ：地物個數

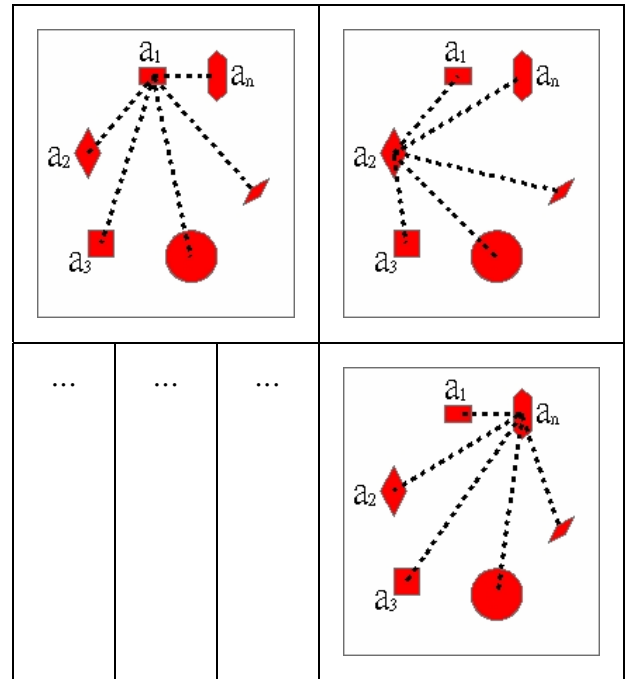


圖 6 地物對其它地物計算離散程度示意圖

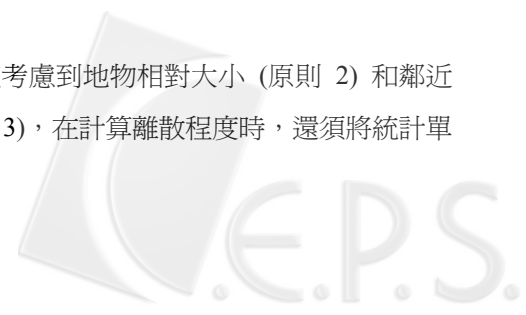
公式 (2) 以二個地物為一組計算 $\bar{d}_{(i,j)}$ ，因各組地物面積不同，各個 $\bar{d}_{(i,j)}$ 應獲得不同的權重，較大的地物獲得較多的權重，較小的獲得較少的權重。考量地物相對大小的影響，公式修正為：

$$\frac{1}{2(n-1)} \sum_{i=1}^n \sum_{j=1}^n (a_i\% + a_j\%) \bar{d}_{(i,j)}, i \neq j \dots\dots\dots \text{公式 (3)}$$

$a_i\%、a_j\%$ ： $i、j$ 地物面積佔總地物面積的%

公式 (3) 係針對面資料所撰寫， $\bar{d}_{(i,j)}$ 是用二地物面積佔總地物面積的比例來權重，點資料無面積大小之分，但有點數多寡之別，因此，將地物面積大小 ($a_i\%+a_j\%$) 改以點數多寡 ($2/\text{總點數}$) 代替，即可將公式 (3) 應用在點資料。

公式 (3) 僅考慮到地物相對大小 (原則 2) 和鄰近程度高低 (原則 3)，在計算離散程度時，還須將統計單



元大小納入考量 (原則 1)，表 1 左列為原圖的計算式，右列為原圖放大 k 倍的計算式，在沒有簡括化的影響下，圖紙放大、縮小並不改變地物的離散程度，二圖的離散程度應相同。地圖放大 k 倍對公式 (3) 的影響為 n、a_i%、a_j%不變，僅 $\bar{d}_{(ij)}$ 改變，在 $\bar{d}_{(ij)}$ 的部份，D 放大了 k 倍，a_i 與 a_j 放大了 k² 倍，其細部計算過程整理如表 1。

$$\bar{d}_{(ij)} = \frac{a_i * d_i + a_j * d_j}{a_i + a_j} = \frac{\Sigma d}{a_i + a_j}$$

$$\Sigma d = a_i |0 - M| + a_j |D - M|$$

$$M = (a_j * D) / (a_i + a_j)$$

Σd ：離均距離總和

M：二地物的重心

D：a_i 與 a_j 的距離

表 1 地圖縮放對 $\bar{d}_{(ij)}$ 影響表

原圖		原圖放大 k 倍
$\frac{a_j * D}{a_i + a_j}$ 設 $\frac{a_j * D}{a_i + a_j} = x$	M	$\frac{k^2 a_j * kD}{k^2 a_i + k^2 a_j} = \frac{k^2 a_j * kD}{k^2 (a_i + a_j)}$ $= k \frac{a_j * D}{a_i + a_j} = kx$
$a_i 0 - M + a_j D - M $ $= a_i 0 - x + a_j D - x $ 設 $a_i 0 - x + a_j D - x = y$	Σd	$k^2 a_i 0 - kx + k^2 a_j kD - kx $ $= k^3 a_i 0 - x + k^3 a_j D - x $ $= k^3 (a_i 0 - x + a_j D - x) = k^3 y$
$\frac{y}{a_i + a_j}$ 設 $\frac{y}{a_i + a_j} = z$	$\bar{d}_{(ij)}$	$\frac{k^3 y}{k^2 a_i + k^2 a_j} = \frac{k^3 y}{k^2 (a_i + a_j)} = kz$

原圖放大 k 倍後，統計單元面積由 A 放大為 k²A，表 1 計算結果放大 k 倍，因左右二圖離散程度相同，若將公式 (3) 的計算結果以 1/A^{1/2}修正，則二圖離散程度相同。考量統計單元大小的影響，公式修正為：

$$\frac{1}{\sqrt{A}} \frac{1}{2(n-1)} \sum_{i=1}^n \sum_{j=1}^n (a_i \% + a_j \%)\bar{d}_{(ij)}, i \neq j$$

.....公式 (4)

A：統計單元面積

公式 (4) 以統計學指標離均差為基礎，經改良後用來度量空間分布的離散程度，因此命名為「空間離散指標」(Spatial Dispersion Index)，簡稱 SDI， $\bar{d}_{(ij)}$ 是地物的離散程度，a_i%+a_j%是該 $\bar{d}_{(ij)}$ 的權重， $\Sigma \Sigma$ 及 i≠j 是地物對其它地物計算離散程度，1/2/(n-1) 是地物重覆

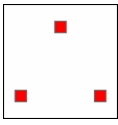
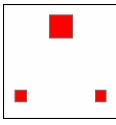
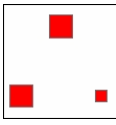
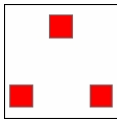
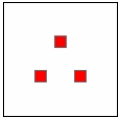
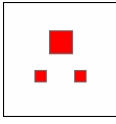
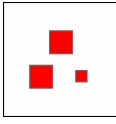
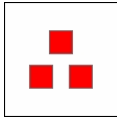
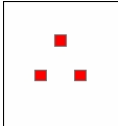
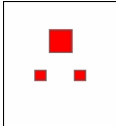
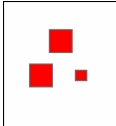
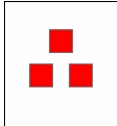
計算的次數，1/A^{1/2}則是對統計單元大小的修正。大抵而言，地物越集中，離散程度越低，SDI 也就越小，地物越分散，離散程度越高，SDI 也就越大。

指標比較與案例演算

表 2 係針對公式 (4) 與三個原則切合程度所做的測試，縱軸部份，a、b 二欄上下各圖的差別在於地物的鄰近程度高低 (測試原則 3)，b、c 二欄上下各圖的差別在於統計單元大小 (測試原則 1)；橫軸部份，各列的差別在於地物相對大小 (測試原則 2)。



表 2 SDI 與三個原則測試表

			
a-1 SDI : 0.35	a-2 SDI : 0.245	a-3 SDI : 0.28	a-4 SDI : 0.35
			
b-1 SDI : 0.175	b-2 SDI : 0.1225	b-3 SDI : 0.14	b-4 SDI : 0.175
			
c-1 SDI : 0.0875	c-2 SDI : 0.06125	c-3 SDI : 0.07	c-4 SDI : 0.0875

就縱軸而言，a、b 二欄上下各圖的統計單元相同、地物大小相同，但鄰近程度不同，a 欄地物間的距離皆為 7，b 欄地物間的距離皆為 3.5，在地物大小、統計單元相同的前題下，地物距離越遠越分散，因此，a 欄較為分散，b 欄較為集中，且 a 欄各圖地物間的距離皆為 b 欄同列各圖的二倍，因此，a 欄的 SDI 恰為 b 欄的 2 倍。b、c 二欄上下各圖的地物大小相同、鄰近程度相同，但統計單元大小不同，b 欄各圖的統計單元皆為 10*10，c 欄各圖的統計單元皆為 20*20，在地物大小、鄰近程度相同的前題下，統計單元越大分布越集中，因此，b 欄較為分散，c 欄較為集中，且 b 欄各圖的統計單元面積皆為 100，c 欄各圖的統計單元面積皆為 400，統計單元大小對 SDI 的修正為 $1/A^{1/2}$ ，b、c 二欄統計單元的修正值為 1/10:1/20，因此，b 欄各圖的 SDI 恰為 c 欄的 2 倍。橫軸部份，第一列各圖的地物平均分散在三處，第二列的地物集中於一處，第三列的地物集中於二處，第四列的地物平均分散在三處，與第一列同，唯地物大小不同，當地物集中於一處時 (第二列)，地物分布最為集中，SDI 最小，當地物集中於二處時 (第三列)，集中程度次之，SDI 次小，當地物平均分布於三處時 (第一列、第四列)，

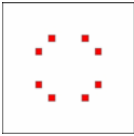
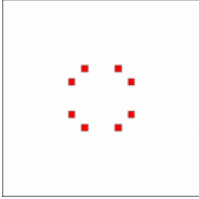
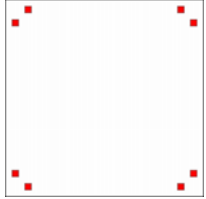
集中程度最低，SDI 最高，各欄的計算結果皆有相同的情形。

表 2 同欄第一列與第四列的 SDI 是相同的，以本研究界定離散的三個原則而言，二列統計單元大小相同 (原則 1)，地物相對大小相同 (原則 2)，地物鄰近程度相同 (原則 3)，因此，SDI 亦相同。不同的地物大小 (及距離) 組合有時會有相同的 SDI，因相同的 SDI 僅表示地物依統計單元大小修正後的離均距離相同，就數列來說，離均差為 2 的數列其可能組合有無限多種，如：A 數列 1,5、B 數列 3,7、C 數列 3,3,3,3,3,7,7,7,7、D 數列 2,3,3,3,4,6,7,7,7,8，這四個數列有不同的大小及距離組合，但若以離均差來論定離散程度，這四個數列的離散程度是相同的，而 SDI 也有著類似的情形，雖然表 2 第一列和第四列的地物大小不同，但二列 SDI 相同。值得注意的是，雖然不同的大小、距離組合會有相同的 SDI，但這並不意謂著相同的 SDI 具有相同的離散意義，這是因為受到統計單元型態、地物面積大小、地物個數多寡等因素的影響，各統計單元 SDI 最大值和各數值出現的機率大都不同，這部份會在文後 SDI 集散意義的判定加以說明。

表 3 是爲了比較 SDI 與其它指標所做的測試，前六種指標在地物鄰近程度的表現較差 (b、c 二列)，這是因爲 Moran's I、Getis's G、Contagion Index 將二地物相鄰與否做爲判定集散程度的依據，Nearest Neighbor Index 採計最近鄰距離，忽略了其它地物的距離，Quadrat Index 僅考慮到統計網格內的情形，沒有考慮到網格外的情

形，Fractal (Box Counting Method) 則是忽略了地物間距離的影響，第七、八種指標在統計單元大小的表現較差 (a、b 二列)，這是因爲二種指標並未對此做修正，而本研究所推導出的 SDI 將鄰近程度與統計單元納入考量，能反應出三個測圖離散程度的差異。

表 3 SDI 與其它指標比較表

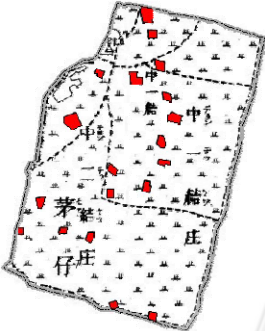
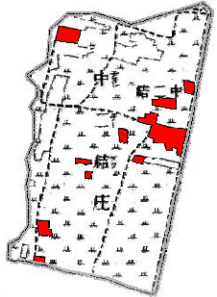
	a	b	c
			
Moran's I (z)	-2.83	-2.83	-2.83
Getis's G (z)	-2.83	-2.83	-2.83
Nearest Neighbor Index (χ^2)	0.8	0.53	0.53
Quadrat Index	24	64	64
Contagion Index	85.59	92.57	92.57
Fractal (Box Counting Method)	2	2	2
Mean Distance	7.37	7.37	24.83
Standard Distance	5.15	5.15	17.73
SDI	0.18	0.12	0.41

部份指標有參數須設定，以上數值僅爲某種設定下的結果

表 4 爲 SDI 的實例演算，分析所採用的地圖爲臺灣堡圖，中一結、中二結二土名皆位在宜蘭廳的茅仔藔堡，施添福 (1996b) 以居民點的戶數爲依據，累計十戶以上的居民點戶數佔總戶數的百分比爲集村度，但該指標忽略了居民點的分布情形，因此，中一結居民點雖多位在中西部，不致屬完全分散的型態，卻因所有居民點都過小，集村度爲 0，中二結則因十戶以上的居民點所佔比例較高，即使西北、西南二側建地距主聚落較遠，但集村度仍高達 0.85，而以堡圖上的建地爲資料計算 SDI，除了較符合以外顯地物判定集、散村的常例外，考量建地間的距離亦爲優勢之一，中二結大部份建地集中在主聚落附近，小部份則距主聚落較遠，相較於中一結建地分布較廣且零散，中二結的 SDI 雖仍低於中一結，較爲集中，但差異不若集村度來的大，分別爲 0.18 與 0.23，

較適合描述中二結集中有散 (集中於東部、有少部份分散) 與中一結散中有集 (大部份分散，但仍有其集中趨勢) 的分布差異。

表 4 中一結、中二結 SDI 與集村度比較表

中一結	中二結
	
集村度：0	集村度：0.85
SDI：0.23	SDI：0.18

SDI 集散意義的判定

SDI 可用於點資料與面資料集散程度的度量，但在判定 SDI 的集散意義時，部份方式計算較為複雜，僅就三種方式加以說明：

1.SDI

就二案例的 SDI 加以比較，這是最為簡便的一種方式。

2.SDI/SDI_{MAX}

統計單元內只有一個地物時，SDI 為 0 (SDI_{MIN})，這是最集中的情形，至於最分散的情形，由公式 (4) 加以推論，SDI 要獲得最大值，分子要盡可能大，分母要盡可能小，分子最大的情形為所有的權重 ($a_i\%+a_j\%$) 都落在一個最大的 $\bar{d}_{(ij)}$ ，也就是說，當統計單元內只有二個地物，並且這二個地物的 $\bar{d}_{(ij)}$ 最大，即是該統計單元最分散的情形，至於分母的部份，由於不同的統計單元常會有不同的 $A^{1/2}$ ，因此，各統計單元 SDI_{MAX} 常是不同的。統計單元內二個距離最長的地物可得該單元的 SDI_{MAX}，但地物大小和地物多寡都會影響到最長距離的長短，前者如較大的地物重心無法存在於統計單元的角落，從而降低了 SDI_{MAX}，後者如二個地物和十個地物的 SDI_{MAX} 是不同的，從而改變 SDI_{MAX} 的值，而統計單元型態亦對 SDI_{MAX} 有著不小的影響，在統計單元面積相同的前題下，當統計單元呈圓形，SDI_{MAX} 最小，當統計單元接近一條線，SDI_{MAX} 最大。相同的 SDI 在不同的統計單元常會有不同的意義，如統計單元 A 的 SDI 介於 0~0.8，地物的 SDI 為 0.5，統計單元 B 的 SDI 介於 0~1.2，地物的 SDI 為 0.6，雖然 B 單元的 SDI 大於 A 單元，但若將 SDI_{MAX}、SDI_{MIN} 納入考慮，則未必如此，轉換方式為「SDI/SDI_{MAX}」，A 的 SDI 約為 SDI_{MAX} 的 0.63，而 B 約為 SDI_{MAX} 的 0.5，以此相比較，A 單元的 0.5 反較 B 單元的 0.6 來得分散。

3.z (SDI) 或 PR (SDI)

z (SDI) 或 PR (SDI) 二種判定方式計算過程大致相

同，故一併說明。SDI 各數值出現的機率是不同的，以圖 7 為例，上圖地物的 SDI 與 SDI_{MAX} 有一段差距，但就機率而言，該統計單元內任取二點，離散程度要高於該圖的情形並不多，SDI 僅用 SDI_{MAX} 詮釋較為不足，用 z 分數 (z score, z) 或百分等級 (Percentile Rank, PR) 來詮釋較為完備，z (SDI) 是在某些條件下 (如：限定地物個數、地物面積)，推演地物可能的分布型態，每一種分布型態都有一個 SDI，也就是盡可能求出所有的 SDI 數值，由於地物可能的分布型態相當複雜，計算過程是在小部份簡化而不過度失真的要求下，將統計單元切成網格 (圖 7 下圖)，先判定該網格是否納入計算，在納入計算的 n 個網格中，以重覆排列、不重覆排列或不重覆組合選取出 m 個網格計算 SDI，其判斷依據為地物是否可在一地重覆出現，共計算出 n^m 、 P_m^n 、 C_m^n 筆 SDI，將地物的 SDI (樣本) 與 n^m 、 P_m^n 、 C_m^n 次 SDI (母群) 相比，即可將地物的 SDI 轉換為 z (SDI) 和 PR (SDI)。

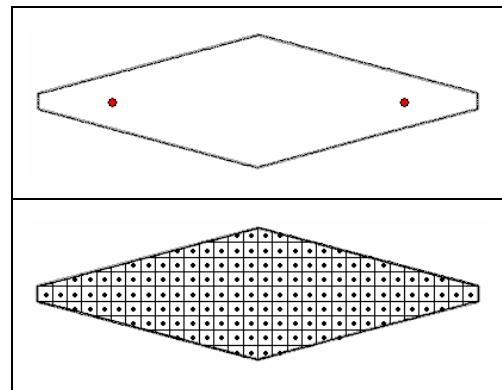


圖 7 統計單元切成網格示意圖

以圖 7 為例，下圖共切成 186 個網格 (網格切法將影響計算結果)，其中 150 個網格的中心點位在統計單元內，將這 150 個網格納入計算，若地物為同一個地點可重覆出現，則將 150 個網格採重覆排列計算，即 n_1 分別與 $n_1 \sim n_{150}$ 計算 SDI、 n_2 點分別與 $n_1 \sim n_{150}$ 計算、……、 n_{150} 點分別與 $n_1 \sim n_{150}$ 計算，產生 22,500 筆 (150^2) SDI，其數值分布情形如圖 8，y 軸為 SDI 值，x 軸為 SDI 由小而大排列累積次數所佔的百分比，該統計單元的 SDI 介於 0~1.18，平均值約 0.33、標準差約 0.21，圖 7 上圖的 SDI

約 0.82，約為 SDI_{MAX} 的 0.69 (0.82/1.18)，若將 0.82 轉換成 $z(SDI)$ ，其值約 2.33，若轉換成 PR (SDI)，其值為 96，因此，就 $z(SDI)$ 或 PR (SDI) 而言，圖 7 上圖的地物分布已相當離散。

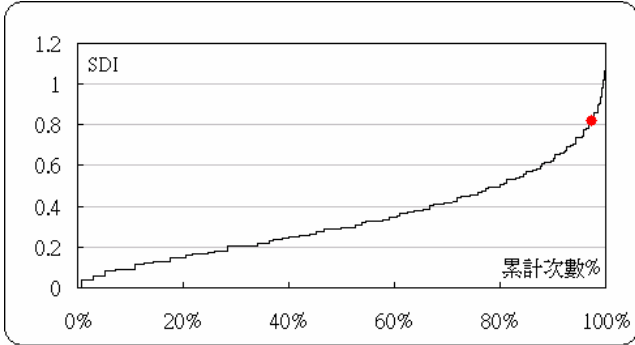


圖 8 圖 7 SDI 累計次數曲線圖

表 5 是四個面積相同、長寬不同的統計單元，a、b、c、d 的長寬分別為 6*6、9*4、12*3、18*2，這樣的設計是為了測試統計單元型態對 SDI 數值分布的影響，每個統計單元網格選取數皆為 2，地物為可重覆出現，以重覆排列計算，共產生 1296 筆 (36^2) SDI。統計單元內二個相距最遠網格的中心點和 SDI_{MAX} 有一些差異，網格越小，二者的差異越少，但為了說明、計算上的方便，僅將這幾個統計單元切成 1*1 的網格，不再做細切。計算

結果顯示，不同型態的統計單元有不同的 SDI_{MAX} 、SDI 平均值和 SDI 標準差，正方形的統計單元這三個數值較小，長寬比值越大的統計單元這三個數值越大。不同型態的統計單元其 SDI 數值分布並不相同，因此，相同的 SDI 在面積相同、型態不同的統計單元裡會有不同的意義，如 a、b、c、d 四個統計單元某地物的 SDI 皆為 0.5，在 a 單元轉換為 $z(SDI)$ 約 1.98、PR (SDI) 為 98，要出現高於 0.5 的機率並不多，在 b 單元轉換為 $z(SDI)$ 約 1.3、PR (SDI) 為 86，已不若前者離散，在 c 單元轉換為 $z(SDI)$ 約 0.66、PR (SDI) 為 72，離散程度屬中上，在 d 單元轉換為 $z(SDI)$ 約 -0.02、PR (SDI) 為 53，離散程度僅屬一般程度。

表 2 第一、四列雖有相同的 SDI，但其集散意義未必相同，就 SDI_{MAX} 加以說明，第一列的地物較小，第四列的地物較大，較小的地物重心能較接近統計單元的角落，地物間重心距離較長，有較大的 SDI_{MAX} ，反之則有較小的 SDI_{MAX} ，因此，第一列的 SDI_{MAX} 會略大於第四列，二列轉換為 SDI/SDI_{MAX} 的數值並不相同，轉換為 $z(SDI)$ 和 PR (SDI) 亦大抵如此，以上皆可說明表 2 第一、四列雖有相同的 SDI，但有不同的集散意義。

表 5 統計單元型態對 SDI 數值分布影響表

	統計單元型態	SDI_{MAX}	SDI 累計次數曲線圖
		Mean (SDI)	
		s (SDI)	
a		0.59	
		0.26	
		0.12	
b		0.71	
		0.29	
		0.16	
c		0.93	
		0.35	
		0.22	
d		1.42	
		0.51	
		0.35	



以上三種方式應用在點資料與面資料的度量時，計算過程稍有不同，第一種方式二者並無差異；第二種方式點資料最離散的情形是統計單元內二個距離最遠的點，面資料最離散的情形是統計單元內二個距離最遠的網格，網格切法可依地物面積的大小來考量，如：取平均值、負一個標準差、……、或是最小地物；第三種方式在 n 個網格取 m 個計算 SDI 時，如果是點資料，不須考慮相鄰網格是否為同一地物，如果是面資料，則須考慮相鄰網格是否為同一地物。

使用與後續研究建議

地理統計經過長時間的發展，已有不少度量集散的指標，而不論是指標的修正或應用，皆有賴更多研究的投入，以下是關於使用與後續研究的幾點建議：

1. GIS 的資料格式可分為向量式資料 (Vector) 和網格式資料 (Raster)，本文各公式係針對向量式資料所撰寫，網格式資料須先轉換成向量式資料後再行計算。

2. 可依研究需求，對公式 (4) 進行細步修改，如：SDI 以離均差來表示離散程度，必要時可改用離均差平方和或標準差來取代、SDI 地物間的距離是取重心距離，若有其它考量，可嘗試改用其它距離代替、有時可忽略統計單元大小的影響，改以公式 (3) 計算 SDI。

3. SDI 可用來計算空間聚集點 (Hot Spot)，如在涵蓋 60% 地物的前題下，SDI 較小的地區，唯計算次數頗多，較為不便。

4. SDI 對於地物面積的絕對大小較無偵測能力，這是公式較不足的地方，有二種修正方向，第一種是改進 SDI，第二種是以其它指標來輔助。

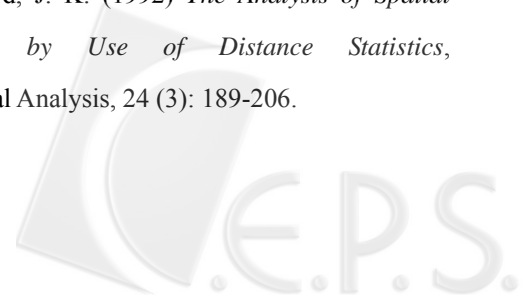
5. z (SDI) 和 PR (SDI) 的計算過程非常複雜，須撰寫程式輔助計算，電腦強大的運算能力除了計算單種地物的 z (SDI) 和 PR (SDI)，亦可於此分布型態下計算其它地物的 z (SDI) 和 PR (SDI)，對於判定不重覆出現地物的集散情形相當有幫助。

6. 更多研究的投入有助於指標的完備，冀望在使用

SDI 的同時，也請多思考 SDI 的優點與侷限，並嘗試改進並發展更好的指標。

引用文獻

- 朱健銘 (2000) 土地利用空間型態之研究，國立臺灣大學地理學研究所碩士論文。
- 李介中 (2005) 碎形理論應用於臺灣地區建地空間型態之研究，國立臺灣大學地理學研究所碩士論文。
- 邱逸民 (2001) 地理學初級統計，臺北：國立臺灣師範大學地理系。
- 施添福 (1996a) 臺灣堡圖：日本治臺的基本圖，臺北：遠流出版事業股份有限公司。
- 施添福 (1996b) 蘭陽平原的傳統聚落—理論架構與基本資料 (上)、(下)，宜蘭：宜蘭縣立文化中心。
- 梁蘄善 (2000) 地理學計量分析，臺北：中國文化大學出版部。
- 黃信誠 (2000) 空間統計簡介，自然科學簡訊，12 (3): 101-104。
- 鄔建國 (2003) 景觀生態學—格局、過程、尺度與等級，臺北：五南圖書出版有限公司。
- Farina, A. (1998) *Principles and Methods in Landscape Ecology*, London: Chapman & Hall.
- Chen, D. M., Getis, A. (1998) *Point Pattern Analysis (PPA)*, <http://www.geog.ucsb.edu/~dongmei/ppa/ppa.html> [2005.6.15].
- Unwin, D. (1981) *Introductory Spatial Analysis*, New York: Methuen.
- Getis, A. and Ord, J. K. (1995) *Local Spatial Autocorrelation Statistics: Distributional Issues and an Application*, *geographical analysis*, 27 (4) 286-306.
- Getis, A. and Ord, J. K. (1992) *The Analysis of Spatial Association by Use of Distance Statistics*, *Geographical Analysis*, 24 (3): 189-206.



Jong, P. de, Sprenger, C. and Veen, F. van (1984) *On Extreme Values of Moran's I and Geary's c*, *Geographical Analysis*, 16 (1): 17-24.

Haining, R. (2003) *Spatial Data Analysis: Theory and Practice*, University of Cambridge.

94年10月25日 收稿

95年03月27日 修正

95年04月28日 接受

