



國中學生視力不良率之校別特性分析

以空間統計解構國中學生視力不良¹率之校別特性，並結合教育部之學生視力保健策略，作為政策論述及資源分配之重要參考。

金允文、鄭冠宏、郭溫慈、程冠瑜（教育部統計處專業助理、科員、專員、專業助理）

壹、前言

國中生的視力問題近年來備受關注，主要原因包括升學壓力及其與 3C 產品的密切接觸，本文針對此議題進行深入探討，不僅運用多元線性迴歸方法建立模型，另外建立多尺度地理加權迴歸模型。透過空間統計分析，探討各校視力不良率與其特性之關聯性，並進一步分析空間上的差異性。本次研究成果於簽陳教育部（以下簡稱本部）部長後，回饋業務司（署）作為制定相關視力

保健政策之參考依據。

貳、研究方法

一、研究範圍界定

為進行空間分析並降低極端值出現的可能，剔除離島及檢測人數未達 10 人之學校，針對 111 學年符合條件之 931 間國中進行分析建模。

二、資料變數說明

在變數蒐集上，除考量學生家長之社經地位對學生 3C 產品之取得及用眼習慣存在相

關，並期能結合本部視力保健策略（近距離用眼每 30 分鐘休息 10 分鐘、每天戶外活動超過 120 分鐘），蒐集學生校內外運動與閱讀之相關資料。爰廣納校別及學校所在鄉鎮市區別之相關統計指標，並積極利用大數據碰檔，串接校內每生之特殊身分、會考成績、體適能表現及父母教育程度等，再依學校彙算占比或平均數值；另以空間統計技術，對學校進行環域分析，掌握其周圍之補習班家數及醫療資源。

為將各類資料轉換為建模

變數，由原始數值換算每生平均數值、調整環域半徑、挑選平均數或中位數、選擇體適能項目、計算父、母或雙方之教育程度等資料處理方式，綜合考量變數間之相關係數、共線性及合理性，並比較同類變數間之模型解釋力後，選定 15 項潛在影響變數，歸類為 5 個類別，各變數的操作型定義（表 1）及處理方式說明如下。

三、迴歸模型

由於相鄰學校的所在環境具有相似特性，因此學校與學校之間的視力不良率可能互有相關，且較近的學校與自身的關聯性更大。為了衡量此空間關聯性，本文會以莫蘭指數（Moran's I）來檢視視力不良率的空間自相關程度（Spatial Autocorrelation）（下頁圖 1）。

為了確認視力不良率是否須以空間統計方法建模，本文會先建立多元線性迴歸，檢定其殘差是否有空間自相關。

若有空間自相關，則建立地理加權迴歸以修正線性迴歸可能造成的模型偏誤。再對地理加權迴歸進行蒙地卡羅檢定，若各解釋變數與視力不良率的空間關係不同，則建立多尺度地理加權迴歸，並以赤池資訊量準則（Akaike information criterion, AIC）進行模型比較，以得到更接近真實的結果。

（一）莫蘭指數（Moran's I）

Moran's I 數值介於 -1 與 1 之間，絕對值愈趨近於

表 1 迴歸模型變數

類別	變數	單位	操作型定義	時間
應變數	視力不良率	%	學校視力不良學生占比	111 學年
1 學校特性	(1) 女學生占比	%	學校女性學生占比	111 學年
	(2) 班生數	人/班	學校班生數	111 學年
2. 社經地位	(1) 綜合所得	萬元	學校所在鄉鎮市區綜合所得中位數	110 年度
	(2) 母親碩士占比	%	學校學生生母教育程度碩士以上占比	111 學年
	(3) 原住民占比	%	學校原住民學生占比	111 學年
	(4) 新住民子女占比	%	學校新住民子女學生占比	111 學年
	(5) 中低收入占比	%	學校中低收入學生占比	111 學年
3. 閱讀考試	(1) 每生借閱冊次	冊	學校圖書館平均每生借閱冊次	110 學年
	(2) 會考成績	積分	學校學生平均會考成績	111 年
	(3) 補習班家數	間	學校周邊 500 公尺內，文理及外語類補習班家數	111 年 7 月底
4. 運動	(1) 每生運動面積	平方公尺	學校平均每生運動場所面積	111 學年
	(2) 跑走百分等級		學校學生平均跑走百分等級	111 學年
	(3) 體育場所面積占比	%	學校所在鄉鎮市區之體育場所面積占比	109-110 年
5. 區域特性	(1) 人口密度	萬人/平方公里	學校所在鄉鎮市區之人口密度	111 年底
	(2) 眼科家數	間	學校周邊 2 公里內，設有眼科之醫療機構家數	111 年底

資料來源：教育部統計處。

論述》統計 · 調查

圖 1 建模流程



資料來源：教育部統計處。

1 代表空間自相關程度愈高，若數值等於 0 則代表空間隨機。公式為：

$$\text{Moran's } I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n W_{ij}} \times \frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

其中 n 為校數， x 為學校視力不良率， \bar{x} 為學校視力不良率平均， W_{ij} 為學校 i 和 j 之間的空間權重。

(二) 多元線性迴歸 (Multiple Linear Regression, MLR)

$$Y = \beta_0 + \sum_{j=1}^k \beta_j X_j + \varepsilon \quad \varepsilon \sim N(0, \sigma^2 I)$$

其中， Y 為反應變數， X_j 及 β_j 分別為第 j 個解釋變數及其迴歸係數， ε 為誤差項。

(三) 地理加權迴歸 (Geographically Weighted Regression, GWR)

GWR 模型是一種延伸自傳統迴歸理論的分析方法，加入了空間座標作為加權變項，允許迴歸係數隨空間而變化，能夠具體呈現空間異質性 (張國禎等，2012)。迴歸方程式如下：

$$Y_i(u_i, v_i) = \beta_0(u_i, v_i) + \sum_{j=1}^k \beta_j(u_i, v_i) X_j(u_i, v_i) + \varepsilon_i(u_i, v_i)$$

其中， $Y_i(u_i, v_i)$ 為第 i 間學校之反應變數， (u_i, v_i) 為第 i 間學校的座標， X_j 及 β_j 分別為第 i 間學校的第 j 個解釋變數及其迴歸係數， ε_i 為誤差項。 β_j 會隨學校的不同而不同，因此空間異質性可由各校在

相同解釋變數上之迴歸係數差異觀察出。

帶寬 (Bandwidth) 是 GWR 中很重要的參數，它用來控制每間學校周圍有哪些學校須納入迴歸方程中。直觀上各校的空間關聯性會隨著距離增加而下降，而遠至關聯性為零的距離就為帶寬。帶寬的選取方式可分為固定核心及調適核心，前者是設定在固定距離內的樣本就納入迴歸方程中，如下頁圖 2 (a)；後者是設定在固定最近數個樣本就納入迴歸方程中，如下頁圖 2 (b)，其中最佳帶寬是根據各帶寬所建的模型之 AICc 大小來決定。

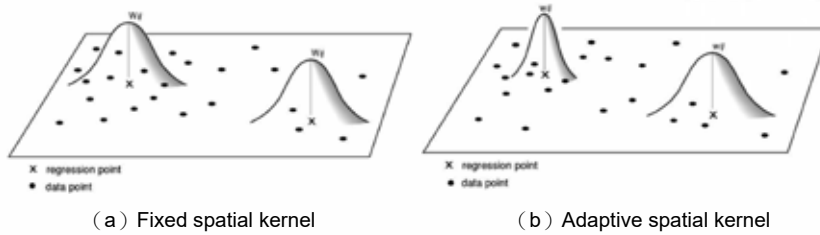
由於學校點位分布並不均勻，為了確保每間學校的迴歸方程有足夠多的樣本，本文採用調適核心選取帶寬，即每間學校周圍最近一定數目之學校會納入迴歸方程中。

(四) 多尺度地理加權迴歸

(Multiscale Geographically Weighted Regression, MGWR)

MGWR 為 GWR 的衍生模型，GWR 限制所有解釋變

圖 2 地理加權迴歸空間核心類型



圖片來源：Fotheringham et al. (2002)

數都要用相同的帶寬計算，因此忽略了不同解釋變數與反應變數可能存在不同的空間關係。為了去除 GWR 的限制，Fotheringham 等 (2017) 提出了 MGWR。相較於 GWR，MGWR 允許每個解釋變數使用不同的帶寬，從而反映不同的空間關係。迴歸方程式如下：

$$Y_i(u_i, v_i) = \beta_{bw_{0,0}}(u_i, v_i) + \sum_{j=1}^k \beta_{bw_{j,j}}(u_i, v_i) X_j(u_i, v_i) + \varepsilon_i(u_i, v_i)$$

其中， $\beta_{bw_{j,j}}$ 代表每個解釋變數的迴歸係數會根據各自的帶寬計算。

參、現況分析

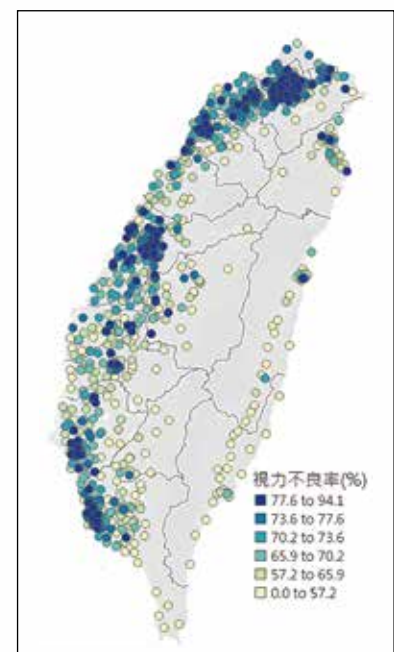
一、國中學生視力隨年齡增加而惡化

111 學年國中學生視力不良率為 73.2%，按年級觀察，國中 1 至 3 年級之學生視力不良率分別為 68.8%、73.4% 及 76.8%，顯示視力隨年齡增加而惡化。另依學生檔與全民健保檔碰檔結果，109 學年國中學生中，有 70.1% 曾於 109 年因眼與附器疾病就診，就診者中平均每人就診 2.9 次，進一步按病因觀察，國中學生中有 38.5% 曾因近視就診，於各教育階段中為最高，國小為 28.5%，高級中等學校為 19.1%。

二、學生視力不良率較高學校主要集中於六都及新竹市

觀察學生視力不良率較高學校之分布，以西部地區明顯多於東部地區，主要集中於六都及新竹市。另計算視力不良率的 Moran's I，得到 I 值為 0.247 (P 值 = 0.001)，表示視力不良率有空間聚集的現象。整體來說，視力不良率較高的學校，其周圍學校也具有高視力不良率的特性。(圖 3)

圖 3 國中學生視力不良率之學校分布



資料來源：教育部統計處。

論述》統計 · 調查

肆、模型配適

一、MLR

MLR 模型結果（表 2）顯示，係數檢定達顯著水準之變數計 10 項，其中再以標準化係數估計值觀察，會考成績影響學生視力不良率之正向程度最

大，而負向影響則以原住民占比及母親碩士占比之影響程度較大。

另以模型殘差計算 Moran's I， $I=0.02$ （P 值 <0.05 ）達顯著水準，顯示殘差分布具有空間自相關，適合建置 GWR 模型以修正線性迴歸之模型偏誤。

二、GWR

GWR 模型之最適帶寬為 501 間，並以此基礎進行模型推估，其分析結果（下頁表 3）顯示，母親碩士占比、原住民占比及會考成績於所有學校均顯著，而每生借閱冊次及眼科家數均不顯著，其餘變數則於部分學校顯著；另進行蒙地卡羅檢定，綜合所得、新住民子女占比及中低收入占比 3 項變數係數具有顯著的空間變異性。因 GWR 模型使用固定帶寬，會忽略不同解釋變數與應變數可能存在相異之空間關係，故繼續建置 MGWR 模型，以不同帶寬反映各變數之空間關係。

三、MGWR

計算各變數之最適帶寬，並以此基礎進行 MGWR 分析，其結果（第 80 頁表 4）顯示，女學生占比、班生數、母親碩士占比、原住民占比、會考成績、補習班家數及人口密度於所有學校均顯著，而每生借閱冊次及眼科家數均不顯著，其餘變數則於部分學校顯著。

排除在所有學校均不顯

表 2 多元線性迴歸模型係數估計值

變數	標準化 係數估計值	變異數 膨脹因子 (VIF)	P 值
常數	67.36		0.000 ***
女學生占比	1.33	1.03	0.000 ***
班生數	1.80	2.49	0.000 ***
綜合所得	0.94	1.73	0.006 **
母親碩士占比	-4.58	2.52	0.000 ***
原住民占比	-4.66	2.23	0.000 ***
新住民子女占比	-0.41	1.82	0.245
中低收入占比	-0.12	2.60	0.778
每生借閱冊次	0.08	1.06	0.761
會考成績	4.98	4.85	0.000 ***
補習班家數	1.18	1.76	0.001 ***
每生運動面積	-0.99	1.60	0.003 **
跑走百分等級	-0.63	1.07	0.020 *
體育場所面積占比	0.04	1.10	0.872
人口密度	1.41	2.83	0.001 **
眼科家數	-0.34	3.51	0.480

AIC=6516.71, $adj R^2=0.601$, 對殘差計算 Moran's I, $I=0.02$ (P 值 <0.05)

說明：* 表 P 值 <0.05 ，** 表 P 值 <0.01 ，*** 表 P 值 <0.001 。
資料來源：教育部統計處。

著之變數後，觀察各變數之帶寬分布情形，一般以總間數之 75% (698 間) 及 25% (233 間) 為門檻，班生數、母親碩士占比、中低收入占比、補習班家數及人口密度之帶寬較大，空間分布愈趨於穩定同質，而綜合所得之帶寬較小，呈異質空

間分布。

另外，本文將 MGWR 模型中各變數的係數值與 t 值整併於同張地圖，使能聚焦於顯著學校的係數分布情形。具體做法為利用實心圓形，呈現 t 值小於 -1.96 或大於 1.96 的學校，以突顯達顯著水準學校之

係數值，反之則以空心三角形表示。再併呈各變數之校別統計散布圖，便於比較及解讀。以下綜合考量變數之重要性及空間異質性，選取 4 項變數進行說明 (第 81 頁圖 4)：

(一) 會考成績

會考成績於所有學校均顯著。整體為會考成績愈高，其視力不良率愈高，以中南部學校影響程度²較大。

(二) 原住民占比

原住民占比於所有學校均顯著。整體為原住民占比愈高，其視力不良率愈低，以桃園新竹一帶影響程度最大。

(三) 每生運動面積

每生運動面積顯著之學校位於新竹以北及東部。大多為每生運動面積愈高，其視力不良率愈低，以校園運動空間相對較為缺乏之北北基學校影響程度最大，若作為資源分配參考，可優先於北北基學校增加校內運動空間。

(四) 中低收入占比

中低收入占比除南部部

表 3 地理加權迴歸模型係數估計值

變數	標準化係數估計值				蒙地卡羅檢定 P 值	顯著學校占比 (%)
	平均值	最小值	中位數	最大值		
常數	67.75	65.90	67.69	69.22	0.00	
女學生占比	1.30	0.77	1.28	2.26	0.49	96
班生數	2.02	0.65	2.11	4.62	0.50	85
綜合所得	1.15	0.22	0.82	3.45	0.02 *	33
母親碩士占比	-4.29	-5.97	-4.58	-3.08	0.20	100
原住民占比	-6.03	-7.37	-6.10	-2.54	0.13	100
新住民子女占比	-0.78	-3.11	-0.40	1.95	0.00 *	40
中低收入占比	1.58	-1.28	1.57	3.41	0.02 *	48
每生借閱冊次	-0.06	-0.85	-0.08	0.70	0.10	0
會考成績	4.91	3.14	4.75	7.18	0.26	100
補習班家數	1.04	0.24	1.04	1.81	0.18	45
每生運動面積	-0.83	-1.99	-0.61	0.42	0.36	42
跑走百分等級	-0.51	-1.98	-0.66	0.51	0.17	30
體育場所面積占比	-0.12	-0.47	-0.25	2.58	0.26	2
人口密度	1.23	-0.10	1.21	1.59	0.78	38
眼科家數	-0.42	-1.32	-0.42	1.44	0.21	0

帶寬 =501 間，AIC=6397.71，adj R²=0.64，對殘差計算 Moran's I，I= -0.01 (P 值 =0.886)

說明：*表 P 值 <0.05。
資料來源：教育部統計處。

論述》統計 · 調查

分學校外均顯著。整體為中低收入占比愈高，其視力不良率愈高，以北北基學校影響程度最大。

四、模型比較

針對 MLR、GWR 及 MGWR 模型配適結果進行比較，發現在 AIC 及 $adj R^2$ ，都以 MGWR

較佳，且其殘差已無空間聚集的情形，最終選取 MGWR 作為本文最適模型。觀察 3 個模型之變數顯著性及係數正負情形，將 MGWR 與 MLR 做比較，MLR 顯著的 10 項變數中，MGWR 有 7 項也是所有學校均顯著，且係數正負在 2 個模型中亦相同，而有 3 項 MLR

顯著變數在 MGWR 則是部分顯著且係數有正有負，另外有 2 項 MLR 不顯著之變數，在 MGWR 為部分學校顯著。綜上所示，MGWR 與 MLR 有著一致性但更細緻的模型結果。

五、結論

學生視力不良率較高學校有以下特性，原住民占比、母親碩士占比愈低、會考成績、班生數、女學生占比、補習班家數及人口密度愈高。其中原住民占比愈低這個特性以桃園新竹學校更為明顯，會考成績愈高這個特性以中南部學校更為明顯，其他 5 項則在全臺差異不大。

在特定區域下，北部、中部及東部視力不良率較高學校，有中低收入占比愈高之特性、新竹以北及東部學校則有每生運動面積愈低之特性。按變數類別觀察，雖顯著情形不一，但閱讀考試變數對於視力不良率為正向影響，運動類變數為負向影響，可作為支持本部視力保健策略之重要參據。

表 4 多尺度地理加權迴歸模型係數估計值

變數	帶寬 (間)	標準化係數估計值				顯著學校占比 (%)
		平均值	最小值	中位數	最大值	
常數	888	68.13	67.42	68.14	68.69	
女學生占比	563	1.32	1.03	1.37	1.98	100
班生數	930	1.74	1.63	1.69	2.06	100
綜合所得	137	1.32	-2.00	0.72	10.38	21
母親碩士占比	930	-3.91	-4.09	-3.91	-3.87	100
原住民占比	634	-5.58	-6.31	-5.49	-4.48	100
新住民子女占比	297	-0.67	-2.39	-0.47	2.13	34
中低收入占比	702	1.71	0.76	1.95	2.49	75
每生借閱冊次	919	0.02	-0.41	0.05	0.39	0
會考成績	611	5.27	4.53	5.66	5.84	100
補習班家數	930	1.23	1.10	1.23	1.35	100
每生運動面積	464	-0.86	-1.59	-0.85	0.36	47
跑走百分等級	381	-0.39	-2.17	-0.44	0.74	23
體育場所面積占比	304	-0.15	-0.73	-0.37	3.21	4
人口密度	930	1.20	1.12	1.19	1.36	100
眼科家數	913	-0.51	-0.68	-0.53	-0.18	0

AIC=6357.38, $adj R^2=0.65$, 對殘差計算 Moran's I, I = -0.02 (P 值 = 0.998)

資料來源：教育部統計處。

伍、結語

值此大數據時代浪潮，資料蒐集之面向及方式日趨多元及複雜，統計分析工作除著重於業務熟悉度及議題敏銳性之掌握外，亦須透過不斷深化及汲取統計理論及技術新知，得以釐析重要政策及其施政成果。本部統計處撰擬此篇分析，多

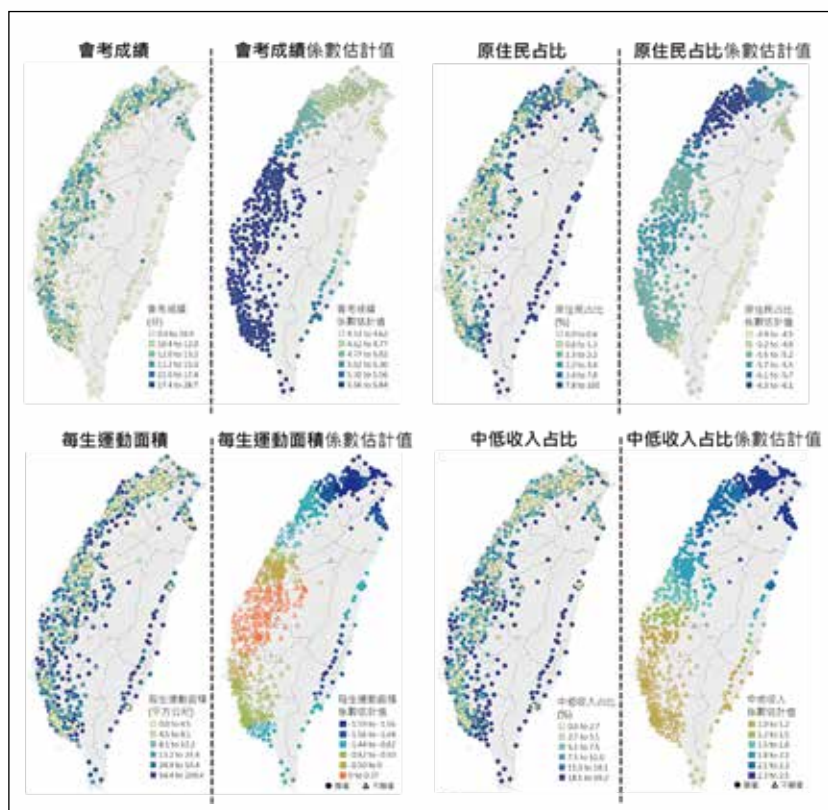
方蒐集學校特性、學生家長社經地位、閱讀及運動等相關解釋變數，建立多尺度地理加權迴歸模型，解構各校學生視力不良率與其特性之關聯性，並分析其於空間上之差異。非僅產出具學術基礎之統計結果，亦為支持當前視力保健政策之有力論述，並可提出教育資源分配之指引，供教育決策施政

參據，充分發揮「循證決策」功能及開展教育統計業務新篇章。

註釋

1. 視力不良係指學生其中一眼裸視視力未達 0.9。
2. 各變數於學校間之影響程度差異，係依各校迴歸方程中該變數迴歸係數估計值之絕對值大小進行評估。

圖 4 會考成績、原住民占比、每生運動面積、中低收入占比及其模型係數估計值分布



資料來源：教育部統計處。

參考文獻

1. 余清祥、梁穎誼、郭柔芸 (2022)，地理加權迴歸在視覺化分析之探討，中國統計學報，第 60 卷 4 期，208-232 頁。
2. 張文崧、陳嘉惠、張國楨 (2017)，應用空間統計於桃園地區土地利用變遷因素分析，地理研究，67 期，137-161 頁。
3. 張國楨、張文崧、曾露儀 (2012)，都市土地利用與人口老化對基層醫療資源分布之影響：以台北市為例，地理研究，56 期，25-40 頁。
4. Fotheringham, A. S., Brunson, C., and Charlton, M. (2002). Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. West Sussex: John Wiley and Sons Ltd. ❖