

# 應用地理人工智慧技術分析國小學區 NO<sub>2</sub> 濃度分布- 以嘉義市為例

王信榮<sup>1</sup> 吳治達<sup>2,3,4\*</sup>

## 摘要

二氧化氮 (NO<sub>2</sub>) 污染為都市重要公共健康議題，對兒童的負面健康影響更深遠。每日上午通勤時段為室外 NO<sub>2</sub> 排放量高峰期。然而有限監測站難以反映學童上學過程暴露的 NO<sub>2</sub> 污染濃度。為了準確掌握國小學童就學通勤時的 NO<sub>2</sub> 污染分布，本研究以嘉義市為例，運用地理人工智慧 (Geo-AI) 技術模擬 NO<sub>2</sub> 濃度分布。蒐集 2015-2020 年空氣污染監測數據，以及土地利用空間相關變數，並以機器學習演算法建立推估模型。結果顯示，主模型以及嘉義市皆有高等解釋能力 (分別為 R<sup>2</sup>=0.94 以及 0.93)，推估成果準確可靠。NO<sub>2</sub> 高濃度地區位於嘉義市中心偏南側，且西區濃度略高於東區。國小學區內道路及住宅區密度與 NO<sub>2</sub> 濃度呈正向關聯。

**關鍵詞：**二氧化氮、空氣污染、機器學習、地理人工智慧、國小學童

## 1. 前言

二氧化氮 (Nitrogen Dioxide, NO<sub>2</sub>) 是一種具刺激性與高活性的空氣污染物。NO<sub>2</sub> 在大氣中停留時間僅約數小時至數日，然而在短時間接觸含有 NO<sub>2</sub> 的廢氣，會造成呼吸不順、咳嗽等立即性不適症狀；長期的暴露更與糖尿病、心血管與呼吸道相關病變以及死亡風險相關 (Boogaard *et al.*, 2022)。除此之外，其容易與其他化學物質發生反應。在陽光中紫外線的照射下，NO<sub>2</sub> 會快速地被分解，最終形成 O<sub>3</sub> 等二次空氣污染物 (Leighton, 1961)，進一步影響人體呼吸系統以及免疫系統的健康。由於兒童的器官系統尚未發育完全，NO<sub>2</sub> 對於學齡兒童的健康損害更為嚴重長遠。NO<sub>2</sub> 不僅是誘發過敏的危險因子，高濃度的室外 NO<sub>2</sub> 濃度也與兒童過敏性鼻炎、呼吸道症狀以及氣喘發生率顯著相關；於高濃度交通污染周邊學校就讀之兒童，其認知發展的增長幅度低於低污染學校的兒童，顯示 NO<sub>2</sub> 的神經毒性可能影

響兒童的腦部發育及認知發展 (Shima & Adachi, 2000、Sunyer *et al.*, 2015)。

NO<sub>2</sub> 污染主要源自人為活動，尤以交通運輸排放為主因。在臺灣，上午 7 點至 9 點的上午通勤尖峰時段通常能觀察到 NO<sub>2</sub> 濃度達到當日的高峰值，使得民眾在主要外出的通勤過程暴露於大量室外空氣污染物 (Rivas *et al.*, 2016)。與此同時，這段短暫通勤過程卻導致兒童接觸到一天中很高比例的空氣污染暴露總量 (Dons *et al.*, 2012)。此外，臺灣通勤運輸模式以汽機車輛為主。多數學童上學過程暴露於汽機車廢氣中，可能對他們的健康產生長期負面影響。為了有效評估學童在上學通勤過程中的 NO<sub>2</sub> 暴露風險，需要掌握環境中的 NO<sub>2</sub> 濃度分布情況。

僅以有限的空氣污染監測站觀測數據，難以妥善描述日常周遭實際的空氣品質與污染物分布狀況。為此，過往研究已提出空間內插、土地利用迴歸 (Land Use Regression, LUR)、機器學習 (Machine

<sup>1</sup> 國立成功大學測量及空間資訊學系 碩士生

<sup>2</sup> 國立成功大學測量及空間資訊學系 教授

<sup>3</sup> 國立中興大學永續農業創新發展中心 教授

<sup>4</sup> 國家衛生研究院國家環境醫學研究所 兼任副研究員

\* 通訊作者, E-mail: chidawu@mail.ncku.edu.tw

收到日期：民國 113 年 03 月 15 日

修改日期：民國 113 年 05 月 13 日

接受日期：民國 113 年 05 月 22 日

Learning) 等方法模擬空氣污染物在時空尺度上的濃度分布 (Beelen *et al.*, 2009、Wu *et al.*, 2017、Requia *et al.*, 2020)。空間內插根據空氣污染濃度的空間自相關性，估算出連續面之污染濃度。地理空間資訊技術 (Geospatial Technologies) 的發展，使得環境科學與空氣污染研究能更全面地掌握大範圍內各種環境潛在因子的資訊。在此背景下，LUR 模型能綜合考量周遭土地利用、交通、污染源等多種影響因子，有助於辨識出重要影響變數與污染物濃度間的線性關聯，而被廣泛運用。隨著資訊科技與電腦運算能力的提升，機器學習演算法等人工智慧 (Artificial Intelligence, AI) 得以從龐大數據中捕獲污染物濃度與預測因子間的非線性特徵，甚至是潛在變數間交互影響關係，進而提高預測準確性。近年來，新興的地理人工智慧技術 (Geo-AI) 充分整合各方法學之優勢，首先透過地理空間資訊技術獲取環境影響因子並進行空間分析，接著運用機器學習演算法進行模型擬合與濃度推估。該方法已成功應用於苯、O<sub>3</sub> 等空氣污染的模擬研究 (Babaan *et al.*, 2023、Hsu *et al.*, 2022)。

因此本研究將以臺灣嘉義市為例，透過地理人工智慧 (Geo-AI) 推估技術，建置出 2015 至 2020 年間之二氧化氮時空推估模型。接著，透過所建模型模擬出上午通勤時段 NO<sub>2</sub> 濃度之分布情況。最終，分析 NO<sub>2</sub> 濃度在嘉義市境內時空分布以及各國小學區範圍內之空間差異，並探究其可能之原因。

## 2. 材料及方法

### 2.1 研究地區

嘉義市面積為 60.03 km<sup>2</sup>，位處臺灣西南部之嘉南平原，境內地勢平緩，四周由嘉義縣環繞。其人口密度達到 4380.20 人/km<sup>2</sup>，為臺灣第二高之縣市 (嘉義市政府民政處，2023)。同時，當地居民普遍搭乘汽、機車通勤代步，高度稠密的住宅配置與道路車流排放的廢氣，已為民眾的身體健康帶來隱憂。

外縣市的境外污染更是影響嘉義市空氣品質的主要因素。由於地理區位關係，雲嘉南空品區中鄰近的嘉義縣以及雲林縣等外縣市污染物會隨著

大氣環流、海風或季風挾帶進入嘉義市；與此同時，雲林縣六輕工業區因設立眾多高聳的煙囪，龐大的工業廢氣嚴重地惡化雲嘉南地區空氣品質。

### 2.2 研究資料庫

#### 2.2.1 環保署空氣污染監測資料

空氣污染資料取自 2015 至 2020 年行政院環保署設立之地面監測站數據。臺灣各區域內之空氣品質濃度，受當地產業、境內污染源、地形結構等因素影響，全臺依此區分成七個空品區 (Air Quality Zone)。由於嘉義市僅 1 個監測站，為了有效且準確地推估嘉義市境內空污濃度的分布狀況，本研究蒐集雲嘉南空品區內 11 個監測站之每小時監測資料 (圖 1)。考慮到早上 7 點至 9 點為上午的尖峰通勤時段，通勤車潮排放的高濃度二氧化氮廢氣為暴露到室外污染的主要時段，事先排除無觀測值之小時，計算出上午 7 至 9 點的平均二氧化氮監測濃度，共計約 2.3 萬筆數據。由於空氣污染物濃度存在空間上的自相依性，經由留一克利金 (Leave-One-Out Kriging Interpolation) 的內插方法計算出整個研究區的二氧化氮推估值，將其作為額外預測變數有助於妥善地捕捉二氧化氮濃度的時空變異 (Wu *et al.*, 2018)。除此之外，與 NO<sub>2</sub> 生成過程或化學反應相關的指標污染物，例如：細懸浮微粒 (PM<sub>2.5</sub>)、臭氧 (O<sub>3</sub>) 以及二氧化硫 (SO<sub>2</sub>) 等，亦經由上述流程計算出各個監測站上午通勤時間之平均數值。接下來，透過普通克利金法 (Ordinary Kriging) 內插法，獲得連續空間面上以及測站站址的濃度分布，並納為推估模型的預測因子。

#### 2.2.2 潛在影響因子資料庫

本研究蒐集與污染物濃度擴散、生成與反應相關的影響因子，根據種類可區分成：空氣污染物、氣象、地形、土地利用、道路、特殊排放源、遙測植生指標與其他因素。空氣汙染物資料已於 2.2.1 節說明之。氣象資料來自大氣水文資料庫 2015 至 2020 年之小時監測數據。本研究蒐集溫度、濕度、氣壓、降雨量、風速與風向等常見項目，參照 2.2.1 節之流

程計算出上午通勤時段之平均監測值，再透過普通克利金或反距離權重法等空間內插方式獲得連續面狀的推估數值，作為潛在的預測因子。地形結構資料來自內政部地政司於 2016 年建置之數值地形模型 (Digital Elevation Model, DEM)，其空間解析度為 20m × 20m。在資料前處理上，首先利用地理資訊軟體進行處理，將 DEM 網格的空間解析度調整成 50m × 50m，以利後續資料分析。接下來，透過空間分析功能，逐一地獲得高程、坡度與坡向等預測變數。土地利用資料來自內政部國土測繪中心辦理與測繪之國土利用調查資料。其記載著各土地利用類型的分布位置與面積，並且按照三階層之結構編定土地利用種類，能夠反映各地的產業結構、經濟發展以及國土使用現況。在本次研究中，採用了第二次 (2006-2015) 以及第三次調查資料 (2016-2019)，其中第二次調查資料的第一層級分為 9 大類、第二層級往下分為 41 類，而第三級再細分成 103 個子類別。考慮到近年的環境開發與社會經濟變遷影響，第三次調查資料針對第二以及第三級分類進行小幅度修改，分別調整為 57 與 101 類。道路資料使用交通部運輸研究所於 2013、2017 與 2020 年產製之數位道路路網資料，本研究根據道路等級，將其劃分成：主要道路、當地道路及所有道路三大類。特殊污染源又可分成大型排放源以及亞洲文化相關排放源：大型排放源包括工業區範圍、焚化爐位置與火力發電廠廠址。亞洲文化相關排放源則包含臺灣中式餐飲烹煮產生的油煙以及寺廟祭祀時的焚香燒金紙等行為。本研究從勤崙國際 POI (Point of Interest) 地標資料庫中，蒐集「中式美食」及「寺廟」類別之地標點資料。遙測植生指標資料取自 2015 至 2020 年間 NASA MODIS 提供之衛星影像。其透過衛星多光譜量測技術，經由公式(1)計算出紅光(Red)與近紅外光(NIR)波段之差異，獲得常態化差異植生指標 (Normalized difference vegetation index, NDVI)。其數值介於±1.0 之間，數值越高代表此處植被蓬勃生長，有助於辨識出大範圍之植被分布與生長狀況。除此之外，人口數與人口密度等資料，以及四季、平日周末等虛擬二元資料亦被納入考量。

$$NDVI = \frac{NIR-RED}{NIR+RED} \dots\dots\dots(1)$$

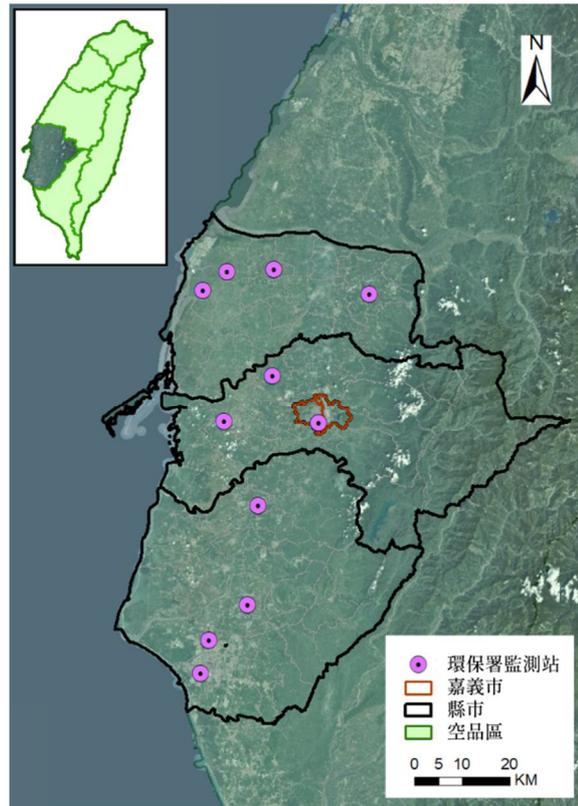


圖 1 雲嘉南空品區與環保署空污監測站分布圖

### 2.2.3 國小資料庫

為了瞭解嘉義市國小數目、所處位置與分布密度，本研究經由政府開放資料平臺獲取 109 年嘉義市國小的位置資訊。從圖 2 得知，嘉義市共有 20 間公立國小，其中東區與西區分別有 10 間。在空間分布上，國小較集中於嘉義市兩行政區交界偏南側。在現行學童的學校選擇權上，我國公立國小中主要採取「學區制」，均等地分配教育資源並提供學童就近入學的機會。各校學區依據「鄰里」劃分，根據學童戶籍地所在位置決定可就讀的學校。由於當地人口與校舍容納額度因素，部分鄰里可能同屬於數間學校的共同學區。一般而言，學區為當地多數學齡學童日常戶外活動與就學的範圍。本研究參酌 111 年嘉義市政府教育處公告之學區範圍，以「里」為單位區分出嘉義市各國小所屬學區，以利進行後續分析。

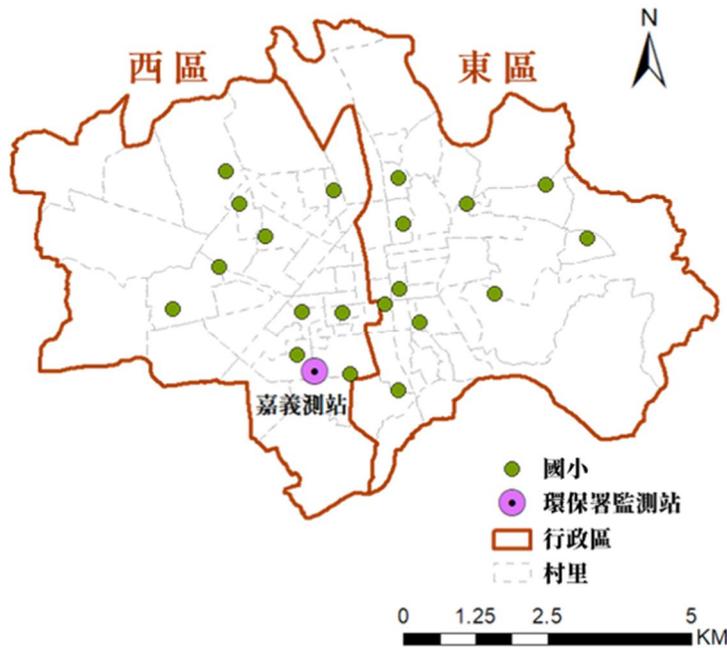


圖 2 嘉義市國小分布圖

## 2.3 地理人工智慧推估模型

本研究藉由地理空間資訊技術 (Geospatial Technologies) 獲取空間相關資料並且進行變數處理；接著透過人工智慧 (AI) 相關的演算法篩選重要變數，進行訓練擬合乃至建立地理人工智慧推估模型。經過模型驗證後，該推估模型將用於繪製 NO<sub>2</sub> 濃度推估圖與後續的統計分析。整體研究流程如圖 3 所示。

在地理空間資訊技術部分，本研究首先蒐集空氣污染物以及 2.2.2 節介紹之氣象、土地利用、遙測植生指標等八類潛在影響因子。接著依據資料格式，透過地理資訊軟體進行資料處理與空間分析，例如：計算各種污染源在監測站周遭 50 m 至 5000 m 不同方圓範圍內的分布密度，以及考慮鄰近道路與大型污染源的直線距離影響。統一資料格式後彙整為預測變數資料庫，總計納入 434 種變數。將 2015 至 2019 年間資料隨機取樣，其中 80% 資料 (15692 筆數據) 作為訓練集 (Training dataset)，用於訓練與建立推估模型；其餘 20% 資料 (3925 筆) 用於檢驗模型預測性能之測試集 (Testing dataset)。另外，2020 年資料 (3906 筆) 用於外部驗證 (External validation)，不納入模型建置流程。

本研究採用極限梯度提升 (eXtreme Gradient Boosting, XGBoost) 演算法建立 AI 推估模型。XGBoost 演算法多次使用訓練資料建立決策樹，從中加強學習錯誤率較高的樣本特徵。前次決策成果會影響後續決策樹的迭代訓練方向，逐漸降低模型的整體預測誤差。此外，該演算法具備對抗雜訊和極端數據影響的優勢，能兼顧良好模型預測性能與計算效率 (Chen *et al.*, 2015)。SHAP (SHapley Additive exPlanations) 值為新興的機器學習變數解釋指標，其概念為根據某變數納入或排除對於預測值的影響，量化該變數在模型中的重要性與影響效應 (Lundberg & Lee, 2017)。其計算公式如下：

$$g(z') = \phi_0 + \sum_{j=1}^M \phi_j \cdot z'_j \dots \dots \dots (2)$$

其中  $g(z')$  為某變數之 SHAP 值； $\phi_0$  表示平均值； $M$  為變數數量； $\phi_j$  為第  $j$  個變數之歸因值； $z'_j$  以 1 與 0 分別表示第  $j$  個變數存在與否。

在推估模型的建立流程中，本研究首先計算所有變數之 SHAP 值並由高而低排列，並依照此順序將變數逐一地納入 XGBoost 模型中。變數納入的標準為連續兩個解釋變數之 partial R<sup>2</sup> 皆  $\geq 0.005$ ；同時亦檢視變數影響的效應，並排除不符預期方向者。

接著，利用挑選之重要變數組合建立推估模型。透過 Python AutoML 套件自動化地調整模型超參數以及進行擬合訓練。最後，透過十折交叉驗證 (10-fold cross validation)、測試資料驗證、2020 年外部資料驗證、時間分層驗證以及空間分層驗證等方式對建立的模型進行驗證，藉此評估模型之預測穩定性和外推應用的適用性。

## 2.4 二氧化氮空間分布與統計分析

本研究透過建立之二氧化氮推估模型，繪製出嘉義市 2020 年每日上午通勤時段 (7 點-9 點) 的二氧化氮濃度分布圖，空間解析度為 50 m。接著，事先排除 1、7 與 8 月等寒暑假月份，計算出當年平日 (周一至周五) 的年平均濃度圖，藉此反映國小學童平日上學時段環境中 NO<sub>2</sub> 的空間分布趨勢。同時視覺化呈現 NO<sub>2</sub> 濃度的四季分布。為了解各學區內的 NO<sub>2</sub> 濃度差異，本研究首先使用非成對 t 檢定，評估嘉義市行政區間 NO<sub>2</sub> 濃度是否存在差異。接下來，本研究假設學童們皆在學區範圍內通勤就學，接著分別考慮學區全境或是只在道路上移動時暴露的平均 NO<sub>2</sub> 濃度，依據污染濃度高低將學校分成三種等級，最後以 ANOVA 檢定分析不同空污嚴重程度學區的統計顯著性。

## 3. 結果

### 3.1 二氧化氮監測值的敘述統計

上午通勤時段之二氧化氮監測數據取自雲嘉南空品區中 11 個環保署監測站，將每日上午 7、8、9 時監測數據平均處理而成。2015 至 2020 年共計蒐集 2,3522 筆數據，平均 NO<sub>2</sub> 濃度為 11.70±6.01 ppb，中位數為 10.53 ppb。空間分布上，從表 1 得知臺南市 (13.58±6.72 ppb) 與嘉義市 (13.68±6.68 ppb) 皆為雲嘉南空品區中濃度偏高的縣市；而雲林縣與嘉義縣平均濃度僅在 10 ppb 上下。時間趨勢上，平均 NO<sub>2</sub> 濃度值由 2015 年 12.78 ppb 平緩下降至 2020 年的 10.43 ppb。在雲嘉南空品區的四季逐年折線圖 (圖 4) 中，亦呈現出明顯地季節性濃度差異：整體而言，雲嘉南地區夏季時的平均濃度最低，約為 7 ppb；春季與秋季相近，冬季時濃度最高。用於建立主模型之訓練資料以隨機取樣形式，取出 2015-2019 年間 80% 之監測數據，藉從中學習到變數間的特徵關係。15692 筆訓練資料濃度之平均值與中位數分別為 11.96±6.11 ppb 與 10.77 ppb。在這之中，嘉義市監測站數據之平均值與中位數分別為 13.86±6.85 ppb 與 13.33 ppb。

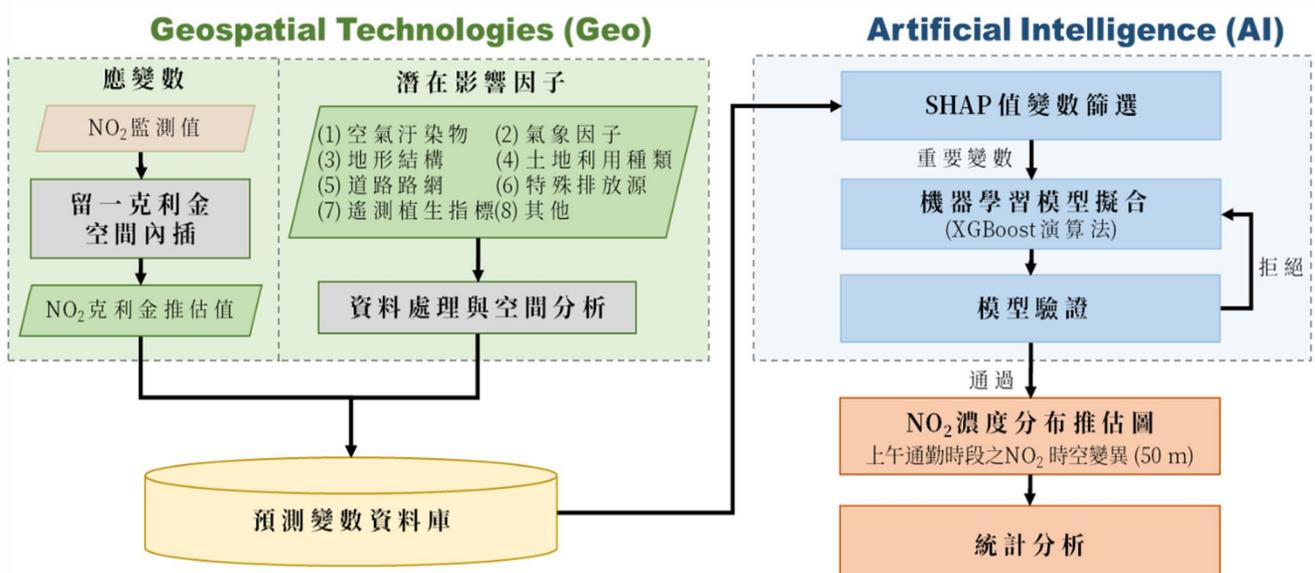


圖 3 地理人工智慧推估模型流程圖

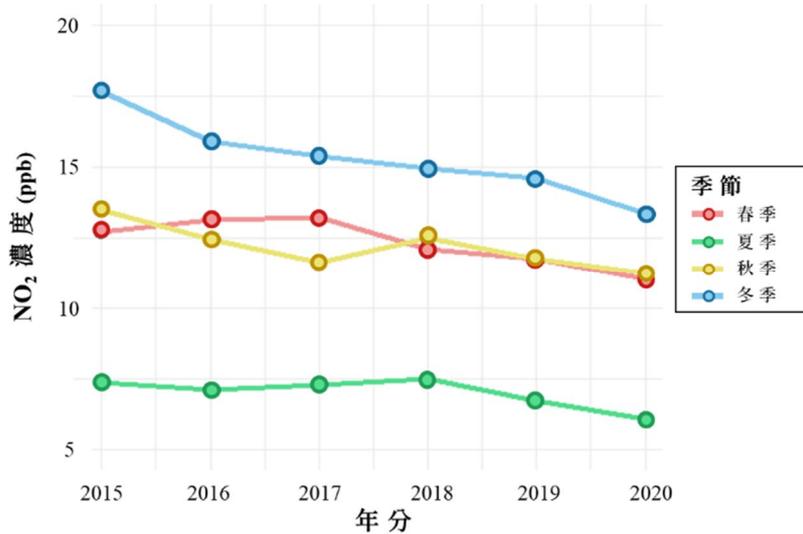


圖 4 雲嘉南空品區二氧化氮監測值之四季折線圖

表 1 二氧化氮監測值按縣市區分之敘述統計 (單位：ppb)

	平均值	標準差	第一四分位數	中位數	第三四分位數
2015-2020 年 雲嘉南空品區	11.70	6.01	7.10	10.53	15.33
臺南市	13.58	6.72	8.43	12.67	17.67
雲林縣	10.18	5.01	6.40	9.27	13.30
嘉義縣	9.98	4.41	6.73	9.10	12.33
嘉義市	13.68	6.68	7.90	13.10	18.67

### 3.2 重要預測變數

本研究計算各預測變數平均絕對 SHAP 值，依照數值降冪排序並篩選出解釋力高的變數，最終選出 10 個 NO<sub>2</sub> 濃度的重要變數。圖 5 (b) 展示各個預測變數與 NO<sub>2</sub> 濃度的關聯性效應，其中 X 軸 SHAP 值之正負號表示正向或負向關聯性；Y 軸則藉由紅與藍色呈現對應的 NO<sub>2</sub> 監測值大小。在這些重要變數中，土地利用以及污染源相關變數包括：750 m 方圓內稻田密度、1250 m 方圓內寺廟密度與當地道路密度。其中稻田變數與 NO<sub>2</sub> 濃度偏向負相關，表示在 750 m 環域範圍內稻田面積越多或密度越高，NO<sub>2</sub> 污染濃度越低；1250 m 環域範圍內寺廟與當地道路密度皆為正向關聯之變數。行政區人口數亦為正相關，當地人口越多，污染物濃度可能越高。此外，模型結果亦篩選出 NO<sub>2</sub>、SO<sub>2</sub>、O<sub>3</sub> 與 PM<sub>2.5</sub> 等指標污染物以及溫度與風速等氣象因子。PM<sub>2.5</sub>、SO<sub>2</sub> 與溫度皆與 NO<sub>2</sub> 濃度有正向關聯；O<sub>3</sub> 與 NO<sub>2</sub> 之間存在負相關性，表示 NO<sub>2</sub> 濃度越高時，周遭環境的 O<sub>3</sub> 濃度可能偏低。風速與 NO<sub>2</sub> 濃度略為呈現負相

關，代表風速越快，測站附近的污染物濃度會下降。

### 3.3 推估模型的預測成果

透過 XGBoost 演算法建立的推估模型，其主模型 R<sup>2</sup> 為 0.94、RMSE 為 1.57，表示該模型能解釋 94% 上午通勤時段雲嘉南地區 NO<sub>2</sub> 濃度的時空變異程度。表 2 展示各種模型驗證成果，十折交叉驗證、測試資料驗證與外部驗證之 R<sup>2</sup> 分別為 0.84、0.86 與 0.81，顯示該模型未有過度擬合疑慮且具良好預測能力。空間驗證中，各縣市的模型 R<sup>2</sup> 高於九成，尤其臺南市與嘉義市推估性能高達 95% 與 93%。接著，根據不同時間尺度檢視雲嘉南空品區內的模型推估能力，研究期間各年分 R<sup>2</sup> 中位數為 0.93；春、秋與冬季 R<sup>2</sup> 皆高於 0.90，夏季也有八成以上的良好表現。本研究亦針對平日 (週一至週五) 與周末 (週六、週日) 進行分層驗證，模型在平日 (R<sup>2</sup> 為 0.94) 與周末 (R<sup>2</sup> 為 0.93) 都有極佳的預測性能。綜合以上結論，建立的推估模型適用於推估嘉義市境內平日的 NO<sub>2</sub> 濃度。

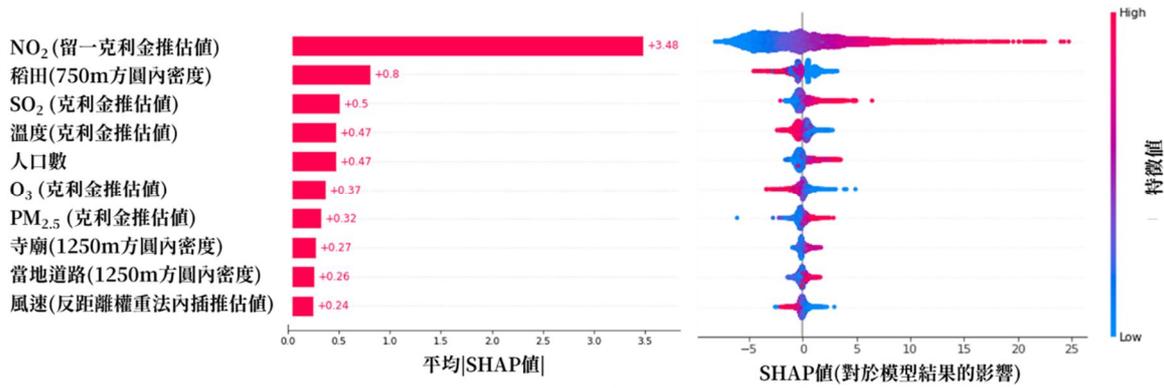


圖 5 重要預測變數之 (a) 重要性直條圖與 (b) 變數特徵散點圖

表 2 二氧化氮推估模型之驗證結果

驗證方式		R <sup>2</sup>	Adjusted R <sup>2</sup>	RMSE	MAE	
主模型 (80%數據)		0.94	0.96	1.57	1.19	
十折交叉驗證		0.84	0.84	2.45	1.81	
測試資料 (20%數據)		0.86	0.86	2.31	1.72	
外部驗證 (2020 年數據)		0.81	0.82	2.32	1.75	
空間(縣市)	臺南市	0.95	0.95	0.98	2.25	
	雲林縣	0.90	0.90	0.95	2.65	
	分層驗證	嘉義縣	0.89	0.89	0.94	2.21
	嘉義市	0.93	0.93	0.97	3.17	
雲嘉南空品區	春季	0.92	0.92	0.96	2.70	
	之時間(季節)	夏季	0.83	0.83	0.91	1.70
	秋季	0.90	0.90	0.95	2.84	
	分層驗證	冬季	0.94	0.94	0.97	2.64

### 3.4 二氧化氮空間推估分布結果

本研究透過上述地理人工智慧推估模型，繪製出 2020 年嘉義市境內平日的 NO<sub>2</sub> 濃度推估圖，其空間解析度為 50m × 50m (圖 6)，表示在每個 50m × 50m 網格範圍內均計算出對應的 NO<sub>2</sub> 濃度數值。在空間分布上，嘉義市中心地區 NO<sub>2</sub> 濃度偏高，嘉義監測站鄰近地區尤其較高；外圍道路較稀疏的郊區 NO<sub>2</sub> 濃度較低，整體 NO<sub>2</sub> 平均濃度為 11.68 ppb。四季分布上，春、秋季濃度與年平均相近，夏季平均濃度僅有 5.80 ppb，而冬季平均濃度為 14.12 ppb。此外，本研究比較兩行政區境內的 NO<sub>2</sub> 濃度差異，並且考慮到民眾在通勤過程，主要在道路上移動，於是進一步抽取道路位置之 NO<sub>2</sub> 濃度進行分析。東區與西區的平均 NO<sub>2</sub> 濃度分別為 11.81 與 11.53 ppb；而兩區域內的道路上 NO<sub>2</sub> 濃度略高於境內平均，分別為 11.99 與 11.55 ppb (表 3)。同時，在全境以及境內道路 NO<sub>2</sub> 濃度的非成對 t 檢定結果中，p 值都遠小於 0.01，達到統計上的顯著水準，結果顯示：西

區的 NO<sub>2</sub> 濃度顯著高於東區。同時，本研究也按照行政區劃分析境內各國小的學區範圍，其中西區與東區國小學區的平均涵蓋面積分別為 3.36 km<sup>2</sup> 與 5.83 km<sup>2</sup>。若探討不同行政區內上學途中的道路污染濃度情況，東區各國小學區之平均 NO<sub>2</sub> 濃度中位數為 11.70 ppb，西區則為 11.67 ppb。兩行政區所屬國小周遭的空氣品質相差無幾，皆遠低於環境部所訂定之年平均污染濃度標準 (30 ppb)。

接下來，本研究計算各國小學區範圍內道路之平均 NO<sub>2</sub> 濃度，並將學校依據污染程度區分成三個等級：高濃度 (12.01-14.00 ppb) 國小有 6 間、中等濃度 (11.51-12.00 ppb) 國小共 9 間，而低濃度 (10.63-11.50 ppb) 國小有 5 間。從表 4 之 ANOVA 檢定結果得知：三種污染等級之間存在顯著差異，顯示國小學區的地理位置以及境內空間相關污染源的分布，會影響學區的空氣品質與污染程度。圖 6 展示了嘉義市道路的 NO<sub>2</sub> 濃度分布以及國小位置，高濃度學校多位於嘉義市中心偏南側。接著，本研究分別以三間國小為例，探討國小學區內的綠地結

構 (例如綠地分布形狀、面積和密度) 以及周遭污染源的分布對學校周圍空氣污染程度的影響。以 A 國小為例, 該校位於嘉義市中央偏南側地區, 學區範圍為 3.74 km<sup>2</sup>。學區內的道路平均 NO<sub>2</sub> 濃度為 13.56 ppb, 屬於高濃度類型。圖 7 中觀察到學區範圍內住宅區密集且道路路網較為繁密。B 國小則位

於中等污染區域, 學區中道路平均 NO<sub>2</sub> 濃度為 11.69 ppb。學區面積為 2.67 km<sup>2</sup>, 且境內涵蓋較多住宅區以及製造業區域。C 國小的學區範圍為 5.89 km<sup>2</sup>, 道路平均 NO<sub>2</sub> 濃度僅為 10.63 ppb, 屬於低污染濃度的學校。該學區內以及鄰近地區由大面積稻田以及森林覆蓋。

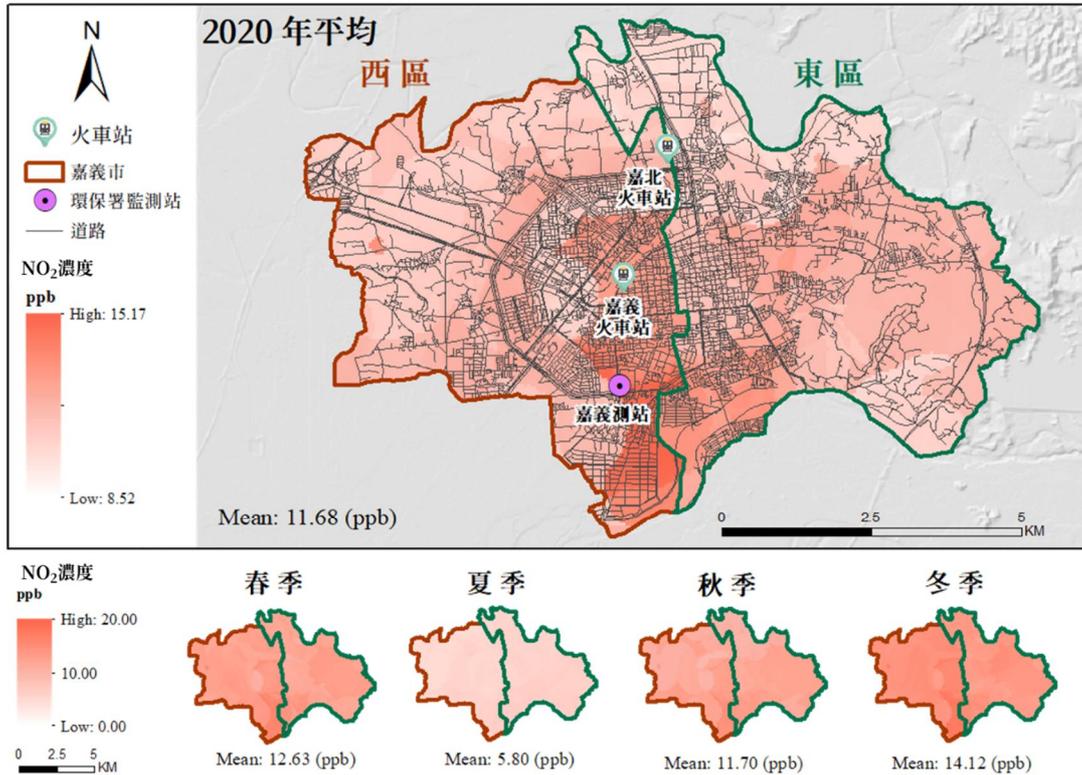


圖 6 嘉義市 2020 年平日上午通勤之 NO<sub>2</sub> 濃度年平均與四季分布圖

表 3 嘉義市依行政區分層下 NO<sub>2</sub> 濃度之非成對 t 檢定結果

	行政區	平均值	標準差	最小值	最大值	Median	p 值
全境 NO <sub>2</sub> 濃度	西區	11.81	1.13	8.75	15.17	11.52	0.00*
	東區	11.53	0.84	9.69	13.43	11.76	
境內道路 NO <sub>2</sub> 濃度	西區	11.99	1.19	8.70	15.17	11.82	0.00*
	東區	11.55	0.88	8.58	15.05	11.65	

\* p 值 <0.01

表 4 各國小學區依 NO<sub>2</sub> 濃度分層之 ANOVA 檢定結果

分層	學區內 NO <sub>2</sub> 濃度			學區內道路之 NO <sub>2</sub> 濃度		
	高濃度	中濃度	低濃度	高濃度	中濃度	低濃度
學校數量	6	9	5	6	9	5
平均值±標準差	12.81±1.16	11.58±0.76	11.18±0.67	12.79±1.16	11.64±0.81	11.21±0.70
中位數	12.97	11.78	11.11	12.90	11.79	10.99
p 值		0.00*			0.00*	
組別	高-中	高-低	中-低	高-中	高-低	中-低
差異統計值	1.63	1.23	0.40	1.58	1.15	0.42
p 值	0.00*	0.00*	0.00*	0.00*	0.00*	0.00*

\* p 值 <0.01

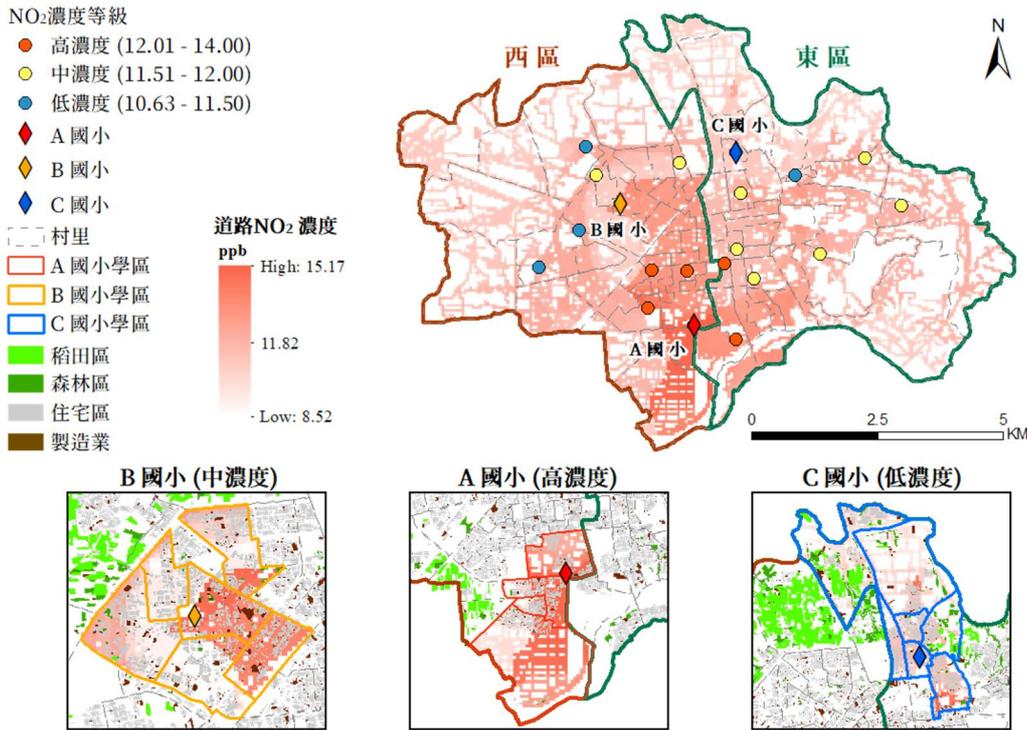


圖 7 嘉義市 2020 年平日上午通勤時段之 NO<sub>2</sub> 濃度與學區內國土利用分布

## 4. 討論

本研究採用地理人工智慧技術模擬出嘉義市境內在上午通勤時段的 NO<sub>2</sub> 濃度分布。藉由機器學習演算法辨識出關鍵排放源與影響因子，並且建置出精確的 NO<sub>2</sub> 推估模型。使用訓練資料建立之雲嘉南空品區主模型 R<sup>2</sup> 達到 0.94，且嘉義市的 R<sup>2</sup> 亦達到 0.93，顯示該模型具備高等的解釋能力。在應用空間技術推估二氧化氮濃度分布的相關研究中，Chen *et al.* (2020) 使用 2000 至 2013 年之臺灣本島月平均監測數據，分別建立出土地利用迴歸(R<sup>2</sup>=0.75)以及混合式克利金土地迴歸推估模型(R<sup>2</sup>=0.87)。Wong *et al.* (2021) 的研究開啟地理人工智慧技術應用的扉頁，該研究使用近 20 年之臺灣本島監測數據，首先以土地利用迴歸進行變數篩選，接著將重要變數組合分別以三種機器學習演算法(隨機森林、XGBoost 以及深度類神經網路) 進行擬合，最佳模型之 R<sup>2</sup> 為 0.91。整體而言，機器學習演算法能大幅提升模型預測性能，並能提供精確、高空間解析度的空氣污染推估成果，使得地理人工智慧成為近年主流的空氣污染模擬方法。

本研究透過評估各個變數在 XGBoost 演算法

模型中之 SHAP 值，辨識並篩選出影響嘉義市上午通勤時段 NO<sub>2</sub> 濃度的關鍵預測變數。人口數以及道路密度皆與 NO<sub>2</sub> 濃度呈正相關，當地經濟較蓬勃的地區人口數往往較多且住宅稠密，會產生較多民生相關污染；尤其在上午的尖峰通勤時段，龐大的車潮會排放大量的 NO<sub>2</sub> 廢氣。除此之外，臺灣的佛、道教廟宇在祭祀過程會焚燒金紙與線香，甚至燃放鞭炮，燃燒不全的濃煙中含有 PM<sub>2.5</sub>、NO<sub>2</sub>、CO<sub>2</sub> 以及多種刺激性物質等污染物，因此寺廟密度會與空氣污染濃度呈現正相關 (Wu *et al.*, 2017)。NO<sub>2</sub> 與其他指標空氣污染物濃度具有密切關聯。SO<sub>2</sub>、PM<sub>2.5</sub> 與 NO<sub>2</sub> 的濃度彼此皆為正相關，主要係因含氮或者含硫成分的燃料或廢棄物經過燃燒，進而生成上述指標污染物。此外，當 NO<sub>2</sub> 與大氣中物質產生反應，亦可能衍生出 PM<sub>2.5</sub> 與其他二次污染物。O<sub>3</sub> 濃度與 NO<sub>2</sub> 呈現負相關，由於 O<sub>3</sub> 為二次污染物，當 NO<sub>2</sub> 接觸陽光中的紫外線會發生光化反應，進而使得 NO<sub>2</sub> 被消耗而產生出更高濃度之 O<sub>3</sub> (Clapp & Jenkin, 2001)。風速與 NO<sub>2</sub> 濃度之間存在負向關聯，若該地區風速越快，空氣污染物越容易被吹向或擴散至其他區域，較不容易沉積於當地，因此 NO<sub>2</sub> 濃度會下降。

利用推估模型繪製出嘉義市 NO<sub>2</sub> 濃度的空間分布圖中，在嘉義車站附近的市中心，以及偏中南側鄰近嘉義測站等地出現高濃度群聚。此外，無論是行政區全境或境內道路上濃度，西區的平均 NO<sub>2</sub> 濃度皆高於東區。本研究分析影響變數數值之間的空間關係進而模擬出 NO<sub>2</sub> 的空間分布，由於嘉義市境內僅一個環境部中央測站提供室外 NO<sub>2</sub> 濃度的監測，因此鄰近嘉義測站的區域較可能推估出較高的污染濃度。為了有效管控境內污染源，嘉義市於 2023 年根據境內土地利用型態，劃設出博愛路一段與友愛路周邊、興業西路康樂街、市中心鄰近文化商圈等五個 PM<sub>2.5</sub> 污染的高濃度熱區（嘉義市政府環境保護局，2023）。從研究成果中觀察到 NO<sub>2</sub> 高濃度區域與上述 PM<sub>2.5</sub> 的熱區範圍大致重疊。由於 NO<sub>2</sub> 與 PM<sub>2.5</sub> 的生成來源相近，這也顯示本研究之 NO<sub>2</sub> 濃度推估成果能反映嘉義市內的空氣品質分布狀況，並可進一步供健康暴露評估分析所用。在季節分布上，夏季平均 NO<sub>2</sub> 濃度最低、冬季則為高濃度污染的季節。空污濃度的季節性差異主因受到大氣混合層高度變化、季風環流以及地形條件影響（Lee *et al.*, 2014、Wu *et al.*, 2019）。大氣混合層高度指的是大氣層中，污染物隨著氣流垂直向上混合的高度範圍。夏季時氣溫較高，使得混合層高度上升，頻繁的大氣垂直對流加上西南季風吹拂，使得 NO<sub>2</sub> 得以擴散至其他地區。冬季時，氣溫較低使得混合層高度下降，整體大氣對流狀態較穩定、擴散條件較差。再加上西南部地勢平坦且位於東北季風被風側，導致污染物更容易沉積於污染源附近。除此之外，較寒冷的天氣也會提升民眾開車通勤避寒的意願，使得汽車相關交通排放量增加。

嘉義市行政區間的分析結果顯示：西區的 NO<sub>2</sub> 濃度略高於東區，然而國小學區內的平均污染濃度則無明顯的行政區差異。另一方面，東區的國小普遍涵蓋較寬廣的學區範圍，平均就學通勤時間可能較西區學校高，導致學童在通勤過程中接觸的空氣污染總量增加。本研究並未觀察到學校人數跟污染濃度之間的明確關聯，這表示學生人數多的國小未必會暴露於較高的空氣污染中。除此之外，三種污染程度的學校案例證實了學區內的土地利用型態

以及綠地結構對室外 NO<sub>2</sub> 濃度的影響。當學區中的住宅區或製造業分布越密集，學童通勤時暴露到的空氣污染程度越高；而當周遭環境森林面積較大、較密集時，樹木能吸收環境中 NO<sub>2</sub> 濃度，從而改善空氣品質（Nowak *et al.*, 2006）。實際上，國小的選址通常會考量多重因素，包含周邊的交通條件、國土規劃設施、當地人口數、以及兒童身心健康等。校址會盡量遠離產生大量噪音以及空氣污染等工廠或嫌惡設施。另一方面，家長在選擇子女學校時，可能會考慮學校教育資源、地理位置或個人因素等條件，甚至可能會選擇「跨學區就讀」，使得學童上學的路程增加。相較於學校所在地的空污程度，學童上學過程中通勤路程距離以及通勤模式也是影響空污暴露量的重要因素（Wolfe *et al.*, 2021）。搭乘校車至較偏遠的學校，通勤暴露總量未必低於步行到學區內的學校。

在後續應用上，可以考慮步行、機車以及汽車等通勤模式差異，將所產製之 NO<sub>2</sub> 空間推估濃度分別轉換為對應的暴露劑量，藉此掌握實際的空污濃度影響程度，並應用於健康風險評估之相關研究。也可以將學童通勤路徑資訊與空間分析規劃功能結合，進而發展個人專屬的上學路徑規劃服務，提供空氣污染暴露總量較低或是兼顧健康與方便性的通勤路線以及通勤模式選擇。

本研究受到資料取得的限制，未能將交通量納為潛在影響變數。通勤尖峰時段經常是 NO<sub>2</sub> 濃度的排放高峰期，先前的研究也指出交通量對於 NO<sub>2</sub> 濃度的影響（Briggs *et al.*, 1997）。由於各縣市的交通量監測資料相對有限，本研究遂以道路路網密度作為替代指標。近年來，隨著感測設備的革新，多數道路段的即時交通流量資料已逐漸普及化，使得未來的空污相關研究能夠更有效地蒐集與運用即時交通資訊。除此之外，本研究納入的國土利用調查資料以及 POI 點位資料更新頻率以數個月或數年為單位，較難反映即時或短時間內的污染源影響。由於本研究在學區範圍劃分方面僅涵蓋至村里尺度，而實際國小的學區可能更加細緻，甚至細分至「鄰」層級。行政區劃的差異導致多個學校的學區在研究中涵蓋更大的範圍，可能導致各校學區內平均 NO<sub>2</sub>

濃度的估計值存在一定程度的誤差。

## 5. 結論

本研究藉由新興的地理人工智慧技術，建立出臺灣雲嘉南地區上午通勤時間的 NO<sub>2</sub> 濃度推估模型，模型能準確解釋雲嘉南境內 94%以及嘉義市內 93%的 NO<sub>2</sub> 濃度變異情況。研究成果顯示在空間分布上，嘉義市西區 NO<sub>2</sub> 濃度顯著高於東區，NO<sub>2</sub> 濃度由中央市中心向外側郊區遞減，同時也存在夏季低、冬季高的季節性差異。從國小學區的分析結果中，學區的空氣污染程度與行政區劃較無相關，境內的土地利用種類或污染源分布導致學區平均 NO<sub>2</sub> 濃度出現統計上顯著差異。研究成果可供有關單位針對高污染地區裝設空污監測設備以及推行管制措施。同時，也期望相關統計成果能提高民眾對於空污暴露議題的重視，並將其納入日常外出或通勤路徑的考量中。

## 參考文獻

嘉義市政府民政處，2023。嘉義市 111 年 12 月份人口統計資料，  
[https://civil.chiayi.gov.tw/News\\_Content.aspx?n=361&s=769891](https://civil.chiayi.gov.tw/News_Content.aspx?n=361&s=769891)，引用 2023/3/15。[Civil Affairs Bureau, Chiayi City Government, 2023. Jia yi shr 2022 nian 12 yue fen ren kou tung ji tz liau, Available at: [https://civil.chiayi.gov.tw/News\\_Content.aspx?n=361&s=769891](https://civil.chiayi.gov.tw/News_Content.aspx?n=361&s=769891), Accessed March 15, 2023. (in Chinese)]

嘉義市政府環境保護局，2023。嘉義市空氣品質感測物聯網維運及應用計畫，  
[https://www.chiayi.gov.tw/News\\_Content.aspx?n=8011&s=638270](https://www.chiayi.gov.tw/News_Content.aspx?n=8011&s=638270)，引用 2023/3/15。  
[Environmental Protection Bureau, Chiayi County, 2023. The maintenance project of internet of things for air quality monitoring in Chiayi City, Available at: [https://www.chiayi.gov.tw/News\\_Content.aspx?n=8011&s=638270](https://www.chiayi.gov.tw/News_Content.aspx?n=8011&s=638270), Accessed March 15, 2023. (in Chinese)]

[=8011&s=638270](https://www.chiayi.gov.tw/News_Content.aspx?n=8011&s=638270), Accessed March 15, 2023. (in Chinese)]

Babaan, J., Hsu, F.T., Wong, P.Y., Chen, P.C., Guo, Y.L., Lung, S.C.C., Chen, Y.C., and Wu, C.D., 2023. A Geo-AI-based ensemble mixed spatial prediction model with fine spatial-temporal resolution for estimating daytime/nighttime/daily average ozone concentrations variations in Taiwan, *Journal of Hazardous Materials*, 446: 130749, DOI: 10.1016/j.jhazmat.2023.130749.

Beelen, R., Hoek, G., Pebesma, E., Vienneau, D., De Hoogh, K., and Briggs, D.J., 2009. Mapping of background air pollution at a fine spatial scale across the European Union, *Science of the Total Environment*, 407(6): 1852-1867, DOI: 10.1016/j.scitotenv.2008.11.048.

Boogaard, H., Patton, A., Atkinson, R., Brook, J., Chang, H., Crouse, D., Fussell, J., Hoek, G., Hoffmann, B., Kappeler, R., Joss, M.K., Ondras, M., Sagiv, S.K., Samoli, E., Shaikh, R., Smargiassi, A., Szpiro, A.A., Van Vliet, E.D.S., Vienneau, D., Weuve, J., Lurmann, F.W., and Forastiere, F., 2022. Long-term exposure to traffic-related air pollution and selected health outcomes: A systematic review and meta-analysis, *Environment International*, 164: 107262, DOI: 10.1016/j.envint.2022.107262.

Briggs, D.J., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebre, E., Pyl, K., Van Reeuwijk, H., Smallbone, K., and Van Der Veen, A., 1997. Mapping urban air pollution using GIS: A regression-based approach, *International Journal of Geographical Information Science*, 11(7): 699-718, DOI: 10.1080/136588197242158.

Clapp, L.J., and Jenkin, M.E., 2001. Analysis of the relationship between ambient levels of O<sub>3</sub>, NO<sub>2</sub> and NO as a function of NO<sub>x</sub> in the UK, *Atmospheric Environment*, 35(36): 6391-6405, DOI: 10.1016/S1352-2310(01)00378-8.

Chen, T., He, T., and Benesty, M., 2015. Xgboost:

- Extreme gradient boosting, R Package Version 0.4-2, pp.1-4.
- Chen, T.H., Hsu, Y.C., Zeng, Y.T., Lung, S.C.C., Su, H.J., Chao, H.J., and Wu, C.D., 2020. A hybrid kriging/land-use regression model with Asian culture-specific sources to assess NO<sub>2</sub> spatial-temporal variations, *Environmental Pollution*, 259: 113875, DOI: 10.1016/j.envpol.2019.113875.
- Dons, E., Panis, L.I., Van Poppel, M., Theunis, J., and Wets, G., 2012. Personal exposure to black carbon in transport microenvironments, *Atmospheric Environment*, 55: 392-398, DOI: 10.1016/j.atmosenv.2012.03.020.
- Hsu, C.Y., Xie, H.X., Wong, P.Y., Chen, Y.C., Chen, P.C., and Wu, C.D., 2022. A mixed spatial prediction model in estimating spatiotemporal variations in benzene concentrations in Taiwan, *Chemosphere*, 301: 134758, DOI: 10.1016/j.chemosphere.2022.134758.
- Lee, J.H., Wu, C.F., Hoek, G., de Hoogh, K., Beelen, R., Brunekreef, B., and Chan, C.C., 2014. Land use regression models for estimating individual NO<sub>x</sub> and NO<sub>2</sub> exposures in a metropolis with a high density of traffic roads and population, *Science of the Total Environment*, 472: 1163-1171, DOI: 10.1016/j.scitotenv.2013.11.064.
- Leighton, P.A., 1961. *Photochemistry of Air Pollution*, Academic Press, New York.
- Lundberg, S.M., and Lee, S.I., 2017. A unified approach to interpreting model predictions, *Advances in Neural Information Processing Systems*, 30, in proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.
- Nowak, D.J., Crane, D.E., and Stevens, J.C., 2006. Air pollution removal by urban trees and shrubs in the United States, *Urban Forestry & Urban Greening*, 4(3-4): 115-123, DOI: 10.1016/j.ufug.2006.01.007.
- Requia, W.J., Di, Q., Silvern, R., Kelly, J.T., Koutrakis, P., Mickley, L.J., Sulprizio, M.P., Amini, H., Shi, L.H., and Schwartz, J., 2020. An ensemble learning approach for estimating high spatiotemporal resolution of ground-level ozone in the contiguous United States, *Environmental Science & Technology*, 54(18): 11037-11047, DOI: 10.1021/acs.est.0c01791.
- Rivas, I., Donaire-Gonzalez, D., Bouso, L., Esnaola, M., Pandolfi, M., De Castro, M., Viana, M., Álvarez-Pedrerol, M., Nieuwenhuijsen, M., Alastuey, A., Sunyer, J., and Querol, X., 2016. Spatiotemporally resolved black carbon concentration, schoolchildren's exposure and dose in Barcelona, *Indoor Air*, 26(3): 391-402, DOI: 10.1111/ina.12214.
- Shima, M., and Adachi, M., 2000. Effect of outdoor and indoor nitrogen dioxide on respiratory symptoms in schoolchildren, *International Journal of Epidemiology*, 29(5): 862-870, DOI: 10.1093/ije/29.5.862.
- Sunyer, J., Esnaola, M., Alvarez-Pedrerol, M., Forn, J., Rivas, I., López-Vicente, M., Suades-González, E., Foraster, M., Garcia-Esteban, R., Basagaña, X., Viana, M., Cirach, M., Moreno, T., Alastuey, A., Sebastian-Galles, N., Nieuwenhuijsen, M., and Querol, X., 2015. Association between traffic-related air pollution in schools and cognitive development in primary school children: A prospective cohort study, *PLoS Medicine*, 12(3): e1001792, DOI: 10.1371/journal.pmed.1001792.
- Wolfe, M.K., McDonald, N.C., Arunachalam, S., Baldauf, R., and Valencia, A., 2021. Impact of school location on children's air pollution exposure, *Journal of Urban Affairs*, 43(8): 1118-1134, DOI: 10.1080/07352166.2020.1734013.
- Wong, P.Y., Su, H.J., Lee, H.Y., Chen, Y.C., Hsiao, Y.P., Huang, J.W., Teo, T.A., Wu, C.D., and Spengler, J.D., 2021. Using land-use machine learning

- models to estimate daily NO<sub>2</sub> concentration variations in Taiwan, *Journal of Cleaner Production*, 317: 128411, DOI: 10.1016/j.jclepro.2021.128411.
- Wu, C.D., Chen, Y.C., Pan, W.C., Zeng, Y.T., Chen, M.J., Guo, Y.L., and Lung, S.C.C., 2017. Land-use regression with long-term satellite-based greenness index and culture-specific sources to model PM<sub>2.5</sub> spatial-temporal variability, *Environmental pollution*, 224: 148-157, DOI: 10.1016/j.envpol.2017.01.074.
- Wu, C.D., Zeng, Y.T., and Lung, S.C.C., 2018. A hybrid kriging/land-use regression model to assess PM<sub>2.5</sub> spatial-temporal variability, *Science of the Total Environment*, 645: 1456-1464, DOI: 10.1016/j.scitotenv.2018.07.073.
- Wu, C.H., Tsai, I.C., Tsai, P.C., and Tung, Y.S., 2019. Large-scale seasonal control of air quality in Taiwan, *Atmospheric Environment*, 214: 116868, DOI: 10.1016/j.atmosenv.2019.116868.

# Estimating Nitrogen Dioxide Concentration Distribution within Elementary School Districts Using Geo-AI Technology: A Case Study of Chiayi City

Hsin-Fen Wan<sup>1</sup> Chih-Da Wu<sup>2,3,4\*</sup>

## Abstract

Nitrogen dioxide (NO<sub>2</sub>) pollution is a concerned public health issue in urban areas. Children may experience more severe health effects when exposed to NO<sub>2</sub>. Furthermore, heavy traffic during the morning commuting time leads to peak outdoor NO<sub>2</sub> emissions. The limited number of monitoring stations poses a challenge in assessing NO<sub>2</sub> exposure during children's school commutes. To accurately depict the spatial distribution and variation of NO<sub>2</sub> concentration during elementary school children's commutes, this study estimated NO<sub>2</sub> distribution in Chiayi City using Geo-AI technology. Air pollution monitoring data during morning commuting time from 2015 to 2020, land use and potential geospatial-related variables were collected. Machine learning algorithm were then used for variable selection and model development. The results reveal that the main model and Chiayi City both had high explanatory power, with R<sup>2</sup> values of 0.94 and 0.93, respectively. The estimations are accurate and reliable. Higher NO<sub>2</sub> concentrations are clustered in the southern-central part of Chiayi City. The averaged NO<sub>2</sub> levels in Western District is slightly higher compared to the Eastern District. Furthermore, concerning the land use distribution patterns within elementary school districts, a positive correlation was observed between NO<sub>2</sub> concentrations around schools and road density and residential area density.

**Keywords:** Nitrogen Dioxide, Air Pollutant, Machine Learning, Geo-AI, Elementary School Children

---

<sup>1</sup> Master Student, Department of Geomatics, National Cheng Kung University

<sup>2</sup> Professor, Department of Geomatics, National Cheng Kung University

<sup>3</sup> Professor, Innovation and Development Center of Sustainable Agriculture,  
National Chung Hsing University

<sup>4</sup> Adjunct Associate Research Fellow, National Institute of Environmental Health Sciences, National Health Research Institutes

\* Corresponding Author, E-mail: chidawu@mail.ncku.edu.tw

Received Date: Mar. 15, 2024

Revised Date: May. 13, 2024

Accepted Date: May. 22, 2024