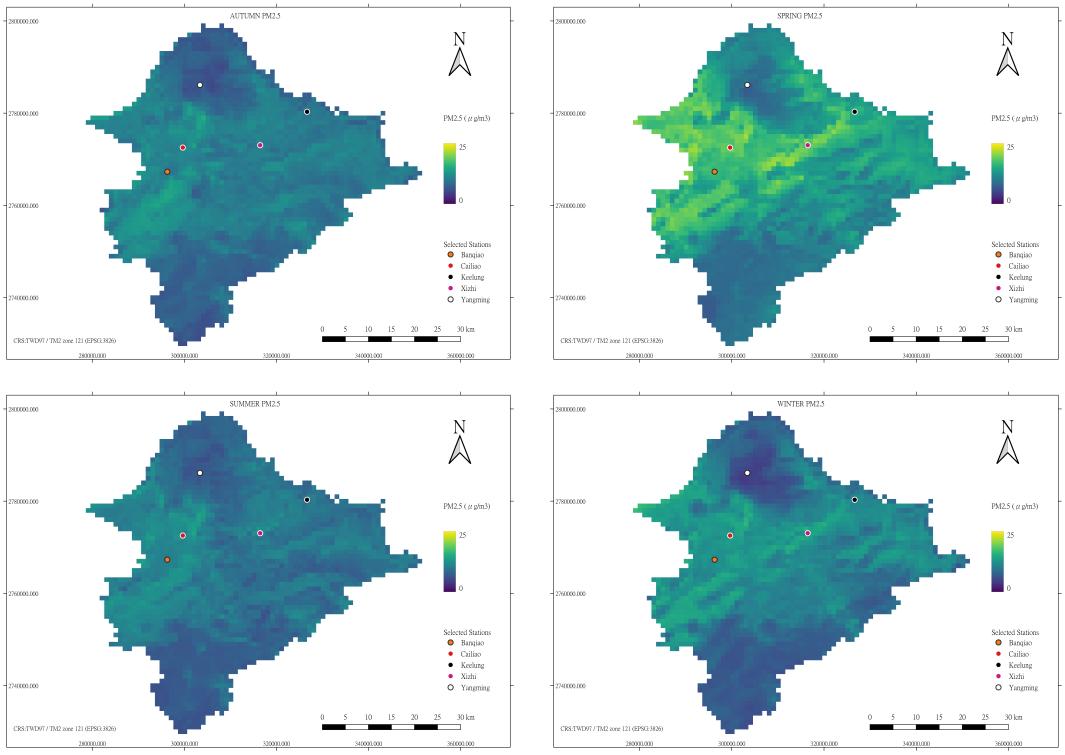
第 三 十一 卷 第 三 期
中華民國一一五年九月

ISSN 1021-8661
DOI: 10.6574/JPRS

航測及遙測學刊

中華民國航空測量及遙感探測學會出版



Journal of Photogrammetry and Remote Sensing

Volume 31 No.3 September 2026

Published by Chinese Society of Photogrammetry and Remote Sensing

航測及遙測學刊

Journal of Photogrammetry and Remote Sensing

發行人：蔡富安

出版者：中華民國航空測量及遙感探測學會

地址：台北市文山區羅斯福路五段113號三樓
電子信件：csprsmail@csprs.org.tw
網址：http://www.csprs.org.tw

PUBLISHER: Fuan Tsai

PUBLISHED BY: Chinese Society of Photogrammetry and Remote Sensing
Address: 3F, No.113, Sec.5, Roosevelt Road, Taipei, Taiwan
E-mail: csprsmail@csprs.org.tw
WebSite: http://www.csprs.org.tw

總編輯：楊明德

國立中興大學土木工程學系
電　話：886-4-2285-2168
電子信件：jprssubmit@gmail.com

EDITOR-IN-CHIEF: Ming-Der Yang
Department of Civil Engineering, National Chung Hsing University
Tel: 886-4-2285-2168
E-Mail: jprssubmit@gmail.com

編輯委員(依中文姓氏筆劃排列)

王素芬 (國立彰化師範大學)
王聖鐸 (國立臺灣師範大學)
史天元 (國立陽明交通大學)
何宗儒 (國立臺灣海洋大學)
林玉儂 (中央研究院)
林昭宏 (國立成功大學)
林唐煌 (國立中央大學)
邱式鴻 (國立政治大學)
洪榮宏 (國立成功大學)
徐百輝 (國立臺灣大學)
張智安 (國立陽明交通大學)
莊永忠 (逢甲大學)
郭俊麟 (國立東華大學)
陳偉堯 (國立臺北科技大學)
陳國華 (國立臺北大學)
陳朝圳 (國立屏東科技大學)
曾國欣 (國立中央大學)
黃倬英 (國立臺灣大學)
蔡富安 (國立中央大學)
蔡榮得 (國立中興大學)
蘇東青 (國立金門大學)
C. Stewart (The Ohio State University) (USA)
M. Matsuoka (Tokyo Institute of Technology) (Japan)
T. Kim (Inha University) (South Korea)

EDITORIAL BOARD

S. F. Wang (National Changhua University of Education)
S. D. Wang (National Taiwan Normal University)
P. T. Y. Shih (National Yang Ming Chiao Tung University)
C. R. Ho (National Taiwan Ocean University)
Y. N. Lin (Academia Sinica)
C. H. Lin (National Cheng Kung University)
T. H. Lin (National Central University)
S. H. Chio (National Chengchi University)
J. H. Hong (National Cheng Kung University)
P. H. Hsu (National Taiwan University)
T. A. Teo (National Yang Ming Chiao Tung University)
Y. C. Chuang (Feng Chia University)
C. L. Kuo (National Dong Hwa University)
W. Chen (National Taipei University of Technology)
K. H. Chen (National Taipei University)
C. T. Chen (National Pingtung University of Science and Technology)
K.H. Tseng (National Central University)
C. Y. Huang (National Taiwan University)
F. Tsai (National Central University)
J. D. Tsai (National Chung Hsing University)
T. C. Su (National Quemoy University)
C. Stewart (The Ohio State University) (USA)
M. Matsuoka (Tokyo Institute of Technology) (Japan)
T. Kim (Inha University) (South Korea)

封面照片說明 About the Cover

The (4) figures above represent $PM_{2.5}$ concentration maps across the four common seasons: winter (December – February), spring (March – May), summer (June – August), and autumn (September – November) for 2023. These seasonal $PM_{2.5}$ concentration maps were obtained by aggregating the monthly map products into seasonal regimes to observe consistency of $PM_{2.5}$ spatiotemporal patterns across these two (monthly and seasonal) temporal regimes. Five observation stations were sampled to show the concentration levels with their rationale discussed in the manuscript text. The interpretation of these map results was also discussed in the manuscript text.

航測及遙測學刊 第三十一卷 第 3 期 第 137 – 159 頁 民國 115 年 9 月
Journal of Photogrammetry and Remote Sensing
Volume 31, No.3, 2026, pp. 137 – 159
DOI：10.6574/JPRS.202609_31(3).0001

A Deep Learning Approach for Spatiotemporal Estimation of $PM_{2.5}$ Concentration in Urban Areas: A Case Study of Greater Taipei Area

Vincent Wiso [1*] Hsuan Ren [2] Tang-Huang Lin [3]

Abstract

Air pollution in urban areas remains a threat, with particulate matter being one of the key indicators of pollution. Numerous studies show that particulate matter with a diameter of 2.5 μm or smaller ($PM_{2.5}$) affects public health, hence the need for more accurate models. Traditional numerical methods fail to capture pollutants' spatiotemporal patterns; however, neural networks can learn from environmental data inputs to obtain correlations. In this study, a Temporal Convolutional Network–Bidirectional Long Short-Term Memory (TCN–Bi-LSTM) hybrid model is proposed. TCN- Bi-LSTM is for $PM_{2.5}$ monthly estimation in 2023 using Taiwan's Ministry of Environment air quality historical data from July 2015 to December 2022 – including weather parameters and land use data – to produce 1km $PM_{2.5}$ maps via k nearest neighbors – inverse distance weighting (KNN – IDW). The results compare the proposed model against classical models like Convolution Neural Network and Long Short-Term Memory. The proposed model performs better compared to the others with a metric evaluation of R^2 = 0.7960 and root mean square error (RMSE) = 2.1606 μg/m³ in the GroupKFold cross-validation. This study aims to contribute to air quality monitoring research by reducing reliance on aerosol optical depth (AOD) derived $PM_{2.5}$ estimates, which are usually affected by cloud cover gaps during satellite AOD retrieval. This study, therefore, addresses the missing data issue by producing continuous $PM_{2.5}$ concentration exposure maps. Future work aims to incorporate independent surface $PM_{2.5}$ observations to check whether new findings and results will be obtained.

Keywords: Particulate Matter (PM), Spatiotemporal, Temporal Convolutional Network, Long Short-Term Memory

1. Introduction

Outdoor air pollution, caused by various pollutants, remains a threat in urban areas. The World Health Organization (WHO) estimated that, in 2019, the majority of the world's population lived in areas that did not meet air quality guidelines, and this was projected to cause approximately 7 million premature deaths annually. This mortality is attributed to exposure to fine particulate matter (PM), which leads to multiple diseases like respiratory and lung cancer (WHO, 2024; WHO, 2025). Particulate matter refers to a blend of solid and liquid particles, of different sizes, that are inhalable and found in the air (US EPA, 2025; WHO, 2025). In Taiwan, studies have shown positive associations between long-term $PM_{2.5}$ exposure and cardiovascular and respiratory morbidity, including hospital visits and outpatient visits (Chien *et al.*, 2018; Wang *et al.*, 2021; Lo *et al.*, 2022). In line with this, it is crucial to estimate the $PM_{2.5}$ concentration accurately spatially and temporally. High-resolution $PM_{2.5}$ concentration estimates complement the monitoring stations' observations, aid in exposure assessment of cohort studies, and policy evaluation of air pollution.

Currently, properly calibrated air quality monitoring systems offer coarse resolution (spatial and temporal) surface-level measurements (Song *et al.*, 2019). The uneven distribution and sometimes few air quality monitoring stations make it difficult to accurately estimate $PM_{2.5}$ concentration at a high spatial resolution in huge geographical areas.

Modeling for surface-level $PM_{2.5}$ concentration includes a myriad of techniques like land use regression, numerical model simulations, geostatistical interpolation, and machine learning, with each having its pros and cons. Numerical models simulate the transformation, transport, and deposition of fine particulate matter (Li & Jiang, 2023). However, this method requires a high amount of computing resources, making it difficult to obtain fine spatial and temporal resolution $PM_{2.5}$ concentrations (Yu *et al.*, 2022). Geostatistical interpolation methods utilize spatial autocorrelation among observation data to produce

[1] Master Student, Centre for Space & Remote Sensing Research, National Central University
[2] Professor, Centre for Space & Remote Sensing Research, National Central University
[3] Distinguished Professor, Centre for Space & Remote Sensing Research, National Central University
[*] Corresponding Author, Tel: +886 981 038441, E-mail: vincletus2050@gmail.com
Received Date: Jun. 12, 2026
Revised Date: Jun. 29, 2026
Accepted Date: Aug. 19, 2026

continuous $PM_{2.5}$ concentration maps. However, this method assumes spatial stationarity and, in some cases, may have over-smoothing or underestimation, especially in sparsely covered areas (Lee *et al.*, 2012; Requia *et al.*, 2019). Land use regression (LUR) estimates $PM_{2.5}$ concentrations by modeling the relationship between land-use variables, such as cover types and distance to emission sources, and pollution levels. Low temporal resolution and low capacity to capture nonlinear relationships among predictors compromise their use (Zhang *et al.*, 2021b; Zhang *et al.*, 2018). Machine learning enables capturing of nonlinear relationships between surface-level $PM_{2.5}$ concentrations and variables like land use characteristics, weather parameters, and satellite aerosol optical depth (AOD), offering improved accuracy (Chen *et al.*, 2018). However, machine learning models may have limited capacity to learn deeper relations, for example, global spatial information and pollutant concentrations' temporal change (Qin *et al.*, 2019).

The majority of surface-level $PM_{2.5}$ concentration methods (apart from machine learning) adopt a linear approach and ignore spatiotemporal relationships. This makes it difficult to estimate $PM_{2.5}$ concentration, especially in areas with intricate atmospheric sources and diverse sources of pollution. On the other hand, deep learning utilizes neural networks to learn the (nonlinear) spatiotemporal relationships between $PM_{2.5}$ concentration and influencing variables like land use characteristics, meteorological parameters, and satellite AOD (Li *et al.*, 2025; Wu *et al.*, 2024).

In this study, we aim to use a hybrid TCN – Bi-LSTM model to estimate monthly $PM_{2.5}$ concentration in the Greater Taipei Area. The final product will be a 1km monthly $PM_{2.5}$ concentration map. The study aims to contribute to the bulk of air quality monitoring by mitigating the missing data issue, mainly caused by cloud cover on satellite observations during aerosol optical depth retrieval.

The report is organized into five sections, with part one being the introduction. Part two explores related work in this line of research, while part three focuses on the study area, dataset, and methodology. Part four discusses the results, while Part five concludes the report.

2. Related Work

The related work reviews approximately 30 articles from the last 30 years, as existing research. This section discusses the importance of $PM_{2.5}$ estimation, challenges associated with $PM_{2.5}$ estimation, data inputs employed, deep learning (and other) modeling techniques employed, results verification, and interpretation, from the perspective of existing research. This section also describes the gaps in knowledge from existing research, which will influence our research study's contribution to $PM_{2.5}$ estimation studies.

The applications of $PM_{2.5}$ estimation from the previous works were threefold. Some studies focused on modeling $PM_{2.5}$ estimation for health outcomes and on-hospital visits (Chien *et al.*, 2018; Wang *et al.*, 2021), epidemiological assessments for public health research (Jung *et al.*, 2018; Zhang *et al.*, 2018; Song *et al.*, 2019), and other uses, such as modeling biomass burning transportation (Lin *et al.*, 2014) and forecasting accuracy of $PM_{2.5}$ estimation (Ong *et al.*, 2016).

The majority of previous studies argue that mapping continuous $PM_{2.5}$ exposure surfaces enables the realization of local spatial heterogeneity, leading to the reconstruction of long-term spatiotemporal patterns for regional comparisons (Chen *et al.*, 2018; Yang *et al.*, 2020; Cheng & Hsu, 2019). Mapping continuous exposure surfaces has been a strong motivation factor, with some studies developing AOD – $PM_{2.5}$ models to estimate pollution in areas with limited air quality monitoring stations, various land use contexts, and complex terrain (Jung *et al.*, 2018; Wang *et al.*, 2025).

This is because output from air quality observation stations is point data, which deteriorates with an increase in distance (from the observation stations) when interpolated (Lee *et al.*, 2012). Furthermore, the observation stations in most areas are sparse and unevenly distributed due to various factors such as operational constraints and suitable site accessibility affecting their installation (Hacıoğlu *et al.*, 2016; Gupta *et al.*, 2018). Empirical evidence of continuous exposure mapping from some studies showed the realization of fine-resolution map products (Wei *et al.*, 2023; Li *et al.*, 2025; Wang *et al.*, 2025). This improvement leads to variations in spatial and temporal scale choices, with the specifications to be employed depending on the objectives to be achieved and tradeoffs to be made.

The previous works concur that multisource data is crucial, with the majority utilizing satellite AOD, chemical transport model (CTM) scaled AOD (reanalysis input data), weather parameters like precipitation, and land use context (Li *et al.*, 2017 & 2022; Zhang *et al.*, 2018; Berrocal *et al.*, 2020). This is because various factors influence $PM_{2.5}$ concentrations; for instance, wind speed affects the accumulation, dispersion, and transportation of $PM_{2.5}$ (Chang, 2019). The issue of satellite AOD missing data during retrieval was addressed mainly by imputing the AOD to obtain a gapless product (Wei *et al.*, 2023; Wang *et al.*, 2025). The former study (Wei *et al.*, 2023) further supplemented the AOD product with a CTM-scaled AOD to stabilize observations (trading off fine resolution for coarse resolution).

The literature review identified three major modelling approaches for $PM_{2.5}$ estimation. Some

adopted a geostatistical approach with a spatial regression aspect where $PM_{2.5}$ as a spatiotemporal input is encoded in a myriad of ways, such as spatial weights or distance, for instance, stratified land use regression and kriging interpolation (Zhang *et al.*, 2021b) and IDW (Wu *et al.*, 2024). Nonlinear function modelling was another approach, where (hybrid) neural network regressors were employed, for instance, geographic regression neural network (Li *et al.*, 2017), and a forest-based machine learning method (Requia *et al.*, 2019). Finally, deep sequence approaches, which model spatiotemporal dependence in their architecture, for instance, a hybrid CNN – LSTM model was used to estimate $PM_{2.5}$ concentration (Qin *et al.*, 2019), a Fast Fourier Transform (FFT) LSTM was employed to forecast $PM_{2.5}$ (Yu *et al.*, 2022). Furthermore, deep sequence approaches consider the observation station sparsity issue; for instance, one approach modeled $PM_{2.5}$ estimation using a correlation-based station selection and spatial attention to prevent some stations from dominating (Lin *et al.*, 2022).

Different authors handled model generalization and evaluation in different ways. Some opted for a cross-validation design, where samples are held out for testing (Chen *et al.*, 2018; Wang *et al.*, 2025), while others opted for the train, validation, and test splits design (Qin *et al.*, 2019; Lin *et al.*, 2022; Hu *et al.*, 2024). Model spatial generalization was explored in two ways, where one was based on the global-to-local framework, with evaluation focusing on localized method behavior under site-based splits using a geographically and temporally weighted nearest neighbor logic (Li *et al.*, 2022). The other exploration way was spatial generalization at the provincial level. However, such regionalization introduced boundary issues that made spatial generalization difficult (Chen *et al.*, 2018). On this note, with no empirical evidence, Qin *et al.* (2019) argued that single-city generalization is preferable to multi-city generalization; hence, the latter should be explored before warranting generalization claims. The temporal generalization perspective focused on the model performance with temporal regime changes. For instance, Zhang *et al.* (2021a) concentrated on the short-term and long-term behavior of $PM_{2.5}$ concentration, Yu *et al.* (2022), and Li & Jiang (2023) centered on forecasting, prioritizing predictive ability across time rather than the (unmonitored) station areas. Jung *et al.* (2018) addressed the seasonality and regime dependence of $PM_{2.5}$ concentrations across seasons and locations, and Lin *et al.* (2014) focused on the Southeast Asian springtime biomass-burning window. There was an observation that the temporal aggregation changes one's claim; for instance, daily estimation does not translate to monthly estimation reliability and vice versa. This is because coarse temporal aggregation can give the impression of improved results while masking local errors (Chen *et al.*, 2018; Wei *et al.*, 2023).

Three patterns emerged in the interpretation of $PM_{2.5}$ findings. First, interpretability by design, whereby model output is readable, for instance, in spatial statistical relationships, through which interpretability is realized in spatial guidelines, such as distance- conditioned blending that explains the reliability of a kriging surface (Lee *et al.*, 2012). Second, is attribution after fitting, whereby models rank variables influencing predictions. For instance, Chen *et al.* (2018) employed ensemble learners to check variables and interpret their influence, and Wei *et al.* (2023) used xAI screening to interpret large-scale aerosol-influencing variables. Third, a pattern to process narrative, whereby authors interpret the mapped output into plausible drivers by diagnostic structure rather than formal causal claims. For instance, Lin *et al.* (2014) framed springtime $PM_{2.5}$ elevation levels around biomass-burning influence and utilized emission comparisons to support attribution. Wang *et al.* (2021) and Wang *et al.* (2025) employed regime stratification as an interpretation method to observe behavior across pollution categories. Moreover, Basith *et al.* (2022) used mechanistic plausibility from biomedical synthesis to support particular exposure patterns in subsequent health contexts without claiming the exposure model itself to be causal.

The literature review of previous works showed two gaps in knowledge. First, the generalization under the station held-out validation design was not consistent because most authors relied on random splits, with few paying attention to the station's sparse areas. Second, model output uncertainty is not consistently showcased as a map product, with few authors producing an uncertainty surface grounded in held-out error behavior.

We endeavor to address these gaps in knowledge in different ways. Our study will adopt GroupKFold by station, which is station -holdout generalization to enable out-of-fold (OOF) samples to represent unseen locations. Moreover, the distance-stratified OOF error will be translated to an uncertainty surface grounded in evaluation, for instance, root mean square (RMSE) by distance.

3. Methodology

3.1 Study Area

The Greater Taipei Area is home to about 7 million people, according to the 2019 census. It is made up of Taipei, Keelung, and New Taipei cities, boasting an area of 2457km^2. This makes it one of the most densely populated areas in the country, at approximately 3000 people per square kilometer (New Taipei City Government, 2017).

New Taipei area has 120km of coastline

replenished by huge rivers like Xindian and Keelung, with terrain features like hills and mountains. The city experiences a warm and humid climate with high rainfall caused by typhoons and monsoons. New Taipei is the most populous municipality with 29 districts led by a mayor and an elected city council as the legislature. Major economic activities are tech and business firms and tourism, for instance, Yingge ceramics (New Taipei City Government, 2017).

Keelung, a natural deep-water harbor, is situated on the northeast coast of Taiwan. Keelung faces the Pacific Ocean, surrounded by offshore islets like Heping and Keelung Islets, and hills and mountains. The climate of the area is maritime with a wet winter influenced by the northeasterly monsoon. The city has 7 districts with multiple departments, like finance and public works. Economically, Keelung, being a port city, logistics and transport depends on the harbor (Keelung City Government, 2022).

Taipei City is located at the northeast tip of Taipei Basin, bordered by New Taipei on all sides. The terrain has the basin floor, overthrust ridges, and the Datun volcanic group, with Keelung, Tamsui, and Xindian rivers shaping the urban lowlands. The city experiences a subtropical monsoon climate with hot summers and mild winters characterized by heavy rainfall and seasonal typhoons. The city is the country's capital and functions as the economic, political, cultural, and financial center. As a municipality, it is divided into 12 districts led by a mayor and an elected council as the legislature (Taipei City Government, 2025).

The Greater Taipei Area has 20 observation stations under the custodianship of the Ministry of Environment (MoENV), with the majority located in the northern part and none in the South, as shown in Figure 1. Therefore, Taoyuan City and Yilan County were included within the study domain to incorporate observations from 8 MoENV stations for spatiotemporal $PM_{2.5}$ mapping, as shown in Figure 2. All model input data were masked to the study area shown in Figure 2. The resulting $PM_{2.5}$ maps were subsequently clipped to the Greater Taipei Area, and only results within this final area were reported, excluding Taoyuan City and Yilan County.

3.2 Data Collection

The choice of dataset was influenced by a review of previous works, as discussed above. The dataset contains static and dynamic variables, with the former remaining stable (changing over long periods of time) while the latter changing over short periods of time. The former includes population density, land use, slope, aspect, building height, building footprint, and road network, while the latter consists of satellite AOD, wind speed, temperature, precipitation, air pressure, relative humidity, and surface- level $PM_{2.5}$, with its data collection (monthly) from July 2015 to December 2023. However, only data from July 2015 to December 2022 were used for model development, with the 2023 data input channels being used exclusively to generate the station-assisted maps described in Section 3.8 and not used in the model training.

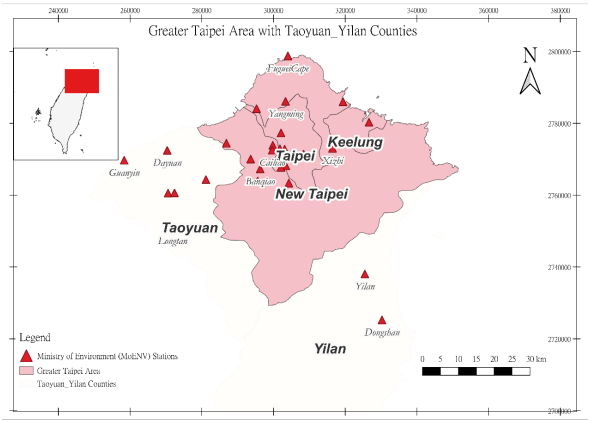


Figure 1 The study Area showing Ministry of Environment (formerly Environmental Protection Agency) stations

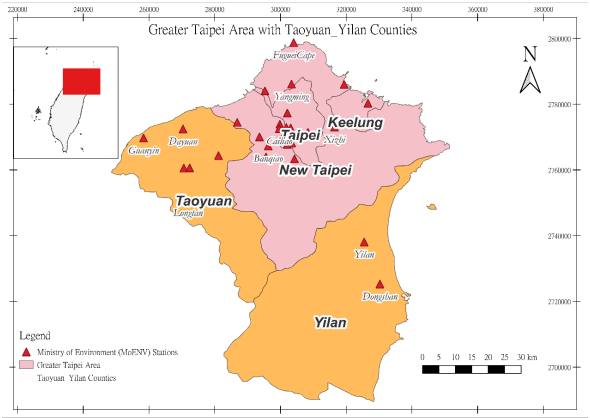


Figure 2 The study Area (including Taoyuan & Yilan) showing Ministry of Environment (formerly Environmental Protection Agency) stations

Satellite AOD is a valuable source of satellite-observed pollutants, which is a proxy for $PM_{2.5}$ to aid the model in learning spatiotemporal patterns. The Himawari product was chosen due to its rich temporal resolution (10 minutes), and a tradeoff in spatial resolution (5km) was made. The downloaded (daily) product was then aggregated into a monthly product and masked and clipped to the study area (including Taoyuan & Yilan areas).

The AOT daily valid count was obtained from aggregating satellite AOD to a monthly product to represent the days contributing to each pixel's monthly AOD, with a higher count indicating that the monthly AOD product is reliable. The AOD daily valid count, ranging from 0 to 31 days, was retained in its original scale rather than standardized to maintain the direct

interpretation (mentioned in beginning of this paragraph).

Meteorological parameters (wind speed, temperature, precipitation, air pressure, and relative humidity) generally influence the formation, loading, transportation, and dispersion of pollutant particles in the atmosphere.

Population density is a vital predictor of $PM_{2.5}$ concentration since it is an indicator of human activities generating emissions, for instance, industrial and commercial activity.

Different land use contexts exhibit distinct emission patterns, hence influencing $PM_{2.5}$ concentration. For instance, agricultural lands are associated with biomass burning and dust.

Slope and aspect influence $PM_{2.5}$ concentration, especially in areas with complex topography. For instance, slope affects how air and aerosol particles traverse the terrain, while aspect influences moisture conditions and solar radiation exposure, depending on the slope direction.

Building height and footprint are urban grain parameters that showcase the built environment conditions that affect emission concentration, airflow obstruction, and pollutant dispersion, hence influencing $PM_{2.5}$ concentration.

Road networks (roads and railways) are associated with suspended dust and exhaust fumes from vehicles, hence influencing $PM_{2.5}$ concentration.

Surface-level $PM_{2.5}$ observations from ground stations served as the dependent target variable for supervised model development, whereas the other 26 variables served as independent predictor inputs. The 2023 $PM_{2.5}$ observations were not used to train the deep-learning model but only for station-level tuning and anchoring of the final station-assisted maps, as described in Section 3.8.

Positional embeddings and temporal encodings were developed by the author for different purposes. The former was to provide the spatial location of each grid cell, enabling the proposed model to learn geographic patterns and location-dependent $PM_{2.5}$ variations. The latter was to represent cyclic temporal patterns, such as seasonality, while preserving continuity between periods, for instance, December and January being treated as temporally close.

3.3 Data Preprocessing

Resampling techniques were used to harmonize the multisource data to a 1km EPSG3826 grid (and monthly temporal resolution). Bilinear interpolation was used to resample continuous variables (satellite AOD, meteorological parameters, slope, aspect, building height, and footprint) to produce smooth transitions. AOT valid daily count was resampled using nearest neighbor to avoid creating artificial validity values.

The static and dynamic data channels were examined to check their distributions, including mean, minimum, maximum values, standard deviation, and percentiles. Dynamic variables were also checked to ensure that no monthly observations were missing. Based on this analysis, a scaling profile was developed to normalize the data inputs and prevent variables with large magnitudes from dominating model training. Satellite AOD, air pressure, air temperature, relative humidity, and slope were standardized using z-scores. Building height, population density, precipitation, road-rail density, and wind speed were transformed using log1p followed by z-score normalization. AOD daily valid count, aspect sine and cosine, building footprint, and land-use were retained in their original form since they were already standardized; for instance, aspect values ranged from -1 to 1, while that of building footprint fractional coverage ranged from 0 to 1. The scaling parameters were derived exclusively from 72 months (from July 2015 to June 2021) and subsequently applied, without refitting, to the other monthly time steps (till December 2022) and 2023 datasets, as will be discussed in the following subsections.

The data channels were stacked into a 147 x 113 x 90 x 26 THWC tensor representing grid height, width, 90 monthly time steps, and 26 channels. A channel-specific monthly mask identified valid and missing entries, retaining months with at least 24 channels of valid data while the masked values were left in the tensor with zero effective weight. The relaxed masking threshold was performed to prevent otherwise usable months from being discarded and limit the propagation of zero-weight and missing samples. The resulting tensor was then ready for model input.

3.4 Flowchart

The selection of the CNN–Bi-LSTM autoencoder, TCN, and Bi-LSTM was informed by a review of related work (Li *et al.*,2022; Wei *et al.*,2023; Hu *et al.*, 2024; Li *et al.*,2025; Wang *et al.*,2025). The specific model configurations, integration strategy, calibration workflow, and station-assisted mapping procedure were developed and adapted for the present study.

A CNN – Bi-LSTM autoencoder was trained on the first 72 months and validated on the subsequent 18 months using a masked Huber reconstruction loss. The autoencoder checkpoint with the lowest masked validation loss was selected, achieving a reconstruction performance of R^2 = 0.9687, RMSE = 0.1589, mean absolute error (MAE) = 0.0993, and mean bias error (MBE) = 0.0382. During training, the decoder reconstructed the initial 26 channels, and at the same time, the encoder learned a 64-dimensional

representation for each month pixel observation, with the selected encoder (with the lowest validation loss) applied to all the 90-time steps to produce a 90 x 64 latent sequence for each pixel for subsequent $PM_{2.5}$ modelling. The pretrained encoder was then fine-tuned (for 300 epochs) using (90 months) surface-level $PM_{2.5}$ observations (at station pixels) to enable learning features more directly related to $PM_{2.5}$ variability, and its (encoder) weights were frozen and used as a fixed feature extractor. Each (monthly) time step was weighted according to predictor availability (see Table 1), $PM_{2.5}$ availability, and monthly observation coverage raised to 0.58, with invalid and missing months receiving zero weight to prevent some months from dominating. The CNN – Bi-LSTM autoencoder specifications and hyperparameters are reported in the appendix.

Training samples for the TCN – Bi-LSTM (proposed) model were generated from the MoENV stations' core grid cells and nearby jittered cells. Temporal encodings (6) and 2 spatial offsets (dx and dy) were appended to each 64-dimensional latent vector. The 6 temporal encodings were: $\sin(2\pi m/12)$, $\cos(2\pi m/12)$, $\sin(4\pi m/12)$, $\cos(4\pi m/12)$, $\sin(\pi \cdot y_norm)$, and $\cos(\pi \cdot y_norm)$, where m is the month-of-year and y_norm is the normalized year position across the study period. The temporal cues represented the annual cycle, second seasonal harmonic, and normalized interannual position, respectively, using sine–cosine pairs, while dx and dy described each sampled pixel's displacement from its source station. The variables (6 temporal encodings and 2 spatial offsets) were not inadvertently entered twice. The original positional and temporal encodings (see Table 1) were included among the 26 channels processed by the autoencoder and compressed into the 64-dimensional latent representation. The six temporal cues were then explicitly appended to the latent features to preserve clear seasonal and long-term temporal information for the prediction heads. Likewise, dx and dy represented relative offsets of the jittered samples from their monitoring stations and were therefore distinct from the absolute x and y coordinate embeddings. This produced a final model input sequence of 90×72.

The TCN and Bi-LSTM regression heads used the derived latent-sequence samples and were trained using station-grouped GroupKFold cross-validation. Entire monitoring stations were withheld in each fold, and both models were optimized using a weighted masked Huber loss. The resulting station-held-out out-of-fold predictions from the two heads were linearly blended using weights estimated exclusively from the OOF results. Calibration was learned from the blended station-held-out OOF predictions after aggregating jittered samples to the station–month level. Candidate station and cluster-based bias or affine corrections were retained only when they improved RMSE without worsening absolute bias, followed by a small global bias adjustment to remove residual drift.

For 2023, rolling 90-month input windows were processed through the fine-tuned PM-aware encoder, after which the blended TCN–Bi-LSTM heads generated raw station and grid-level $PM_{2.5}$ estimates. Station predictions were tuned using the best-performing calibration function against 2023 observations to define station anchors. Monthly maps were then produced by interpolating the anchor-based correction deltas with KNN–IDW using up to three stations within 120 km, adding these corrections to the raw pixel estimates, and fixing station-core pixels to the tuned anchors. The resulting 1km GeoTIFF maps were finally clipped to the Greater Taipei Area. The following sections of the methodology will further explore the proposed model training, calibration, and station-assisted mapping.

Table 1 Dataset used for the proposed model input, with their original resolutions

Variables	No. of variables	Resolution (spatial/temporal)	Source
Satellite AOD: Himawari L3 Monthly(500nm)	1	5km/Daily	Himawari 8
AOT daily valid count	1	NA/NA	
Precipitation, Relative Humidity, Temperature, Atmospheric pressure,	4	2.5km/Monthly	CWA
Wind speed	1	2km/ Daily	TCCIP
Population Density	1	1km/NA	WorldPop Hub
Land Use	6	10m/NA	ESA WorldCover 2021
Slope/Aspect (cos &sin)	3	30m/NA	ALOS DEM
Building height	1	100m/NA	GHSL of EU-JRC
Building footprint	1	10m/NA	GHSL of EU-JRC
Road network	1	NA/NA (rasterized)	NLSC
Surface $PM_{2.5}$	1	NA/Hourly	MoENV
Positional embeddings (x & y)	2	NA/NA	
Temporal encodings: cos/sin (Month & Year)	4	NA/NA	

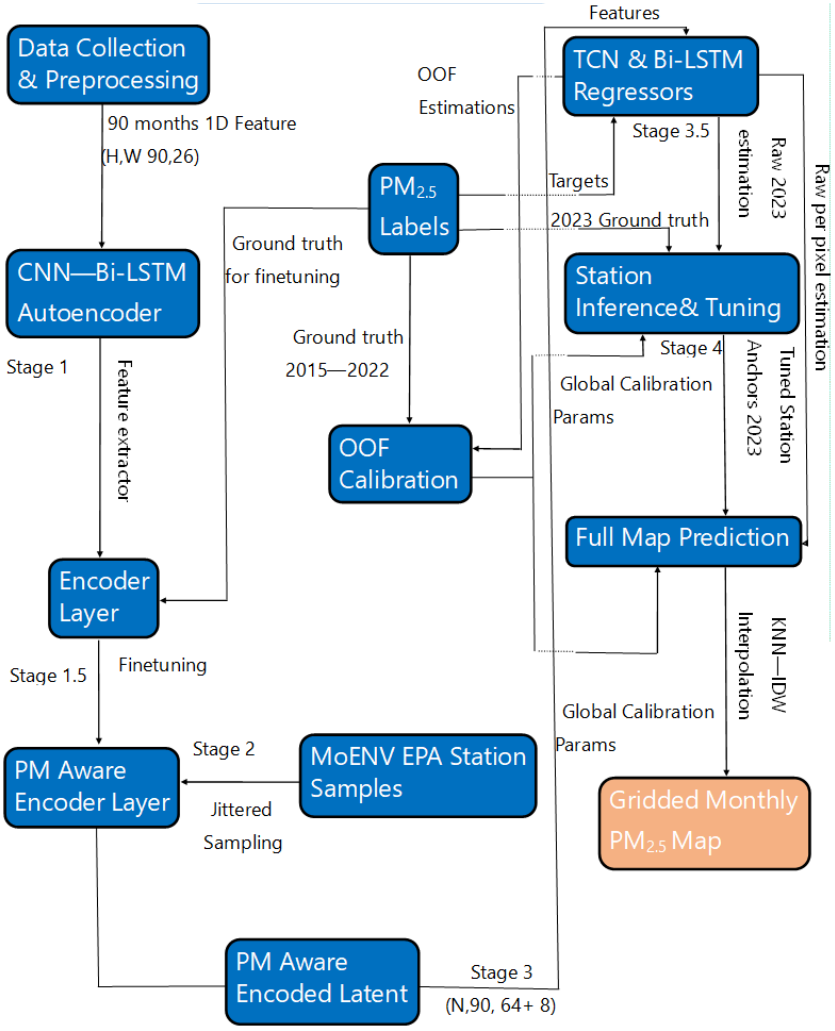


Figure 3 Overall workflow for $PM_{2.5}$ estimation and 2023 station-assisted spatial mapping

3.5 Model Architecture

The proposed model comprises parallel dual heads – TCN and Bi-LSTM, as shown in Figure 4. The architecture was selected based on the complementary strengths of the two models and their demonstrated effectiveness in related studies (as discussed earlier). The TCN head captures local and multi-scale temporal patterns using causal convolutions, whereas the Bi-LSTM captures longer-range dependencies in both temporal directions.

The TCN head is a causal, dilated 1D convolution block in residual stacks. Two cycles of four dilated residual blocks with dilations, d ∈ {1,2,4,8} were utilized. Each of these blocks contains two Conv1D layers (kernel size, k =3 and 64 filters) with ReLU, Dropout, Layer Normalization, and a 1 x 1 skip projection. The receptive field, R, is represented by equation (1).

$$R = 1 + (k-1)\sum_{l=1}^{L} d_l \quad \text{.....(1)}$$

where the sum includes all the Conv1D layers in the stack. L stands for the total number of all layers, while d is the dilation rate.

$$\sum_{l=1}^{L} d_l = 2(1+2+4+8) \times 2 = 60$$
$$\Rightarrow R = 1 + (3-1) \cdot 60 = 121$$

This enables the head to see 121-time steps in one forward pass with no future leakage and, therefore, can cover our historical period, 90 months, with ease.

The Bi-LSTM head is made up of two layers with 64 and 32 units of direction, respectively. For a monthly sequence, $X = \{x_t\}_{t=1}^{T}$, with $T = 90$, $x_t \in R^{72}$, per each time step, the $PM_{2.5}$ concentration is mapped by stacking the two bidirectional layers. For the l^{th} Bi-LSTM layer,

$$h_{t,\text{fwd}}^{(1)} = \text{LSTM}_{\text{fwd}}^{(1)}\left(z_t^{(l-1)}, h_{t-1,\text{fwd}}^{(1)}, c_{t-1,\text{fwd}}^{(1)}\right) \quad \text{.....(2)}$$

$$h_{t,\text{bwd}}^{(1)} = \text{LSTM}_{\text{bwd}}^{(1)}\left(z_t^{(l-1)}, h_{t+1,\text{bwd}}^{(1)}, c_{t+1,\text{bwd}}^{(1)}\right) \quad \text{.....(3)}$$

(2) and (3) concatenated to obtain (4);

$$h_t^{(1)} = \left[h_{t,\text{fwd}}^{(1)} ; h_{t,\text{bwd}}^{(1)} \right] \quad \text{.....(4)}$$

where $z_t^{(1)} = x_t$, $z_t^{(2)} = h_t^{(1)}$. After the second Bi-LSTM, layer normalization and dropout are applied, obtaining the normalized state $\widehat{h_t} = \text{Dropout}\left(\text{Layer Norm}\left(h_t^{(2)}\right)\right)$

$$\widehat{y_t} = W_o \, \widehat{h_t} + b_o, \qquad \widehat{y_t} \in R \quad \text{.....(5)}$$

where W_o is the readout weight of the Dense (1) layer while b_o is bias.

3.6 Model Training

Each 90×72 latent-sequence sample was screened before training. A sample-month was considered valid when at least 24 data channels and a $PM_{2.5}$ observation were available; invalid jitter–station associations and zero-weight samples were removed. Its effective weight was represented as

$$w_{i,t} = 1_{valid_{i,t}} \left(\frac{hour_c overage_{i,t}}{100} \right)^{0.58} s_i \quad \text{.....(6)}$$

where $1_{valid_{i,t}}$ is a time mask while $\left(\frac{hour\ _c overage_{i,t}}{100} \right)^{0.58}$ refers to coverage weight and s_i is spatial weight. The samples and time steps (month, in this case) were represented by $i \; and \; t$ respectively. Jitter sampling utilized a ring-based spatial weight that decreased with distance from the source station (r≤2 and a guarded r=3 for Yangming station) to ensure well-observed months and pixels nearer the station contributed more to the effective weight. The sampling radius r represents **±2 pixels** (approximately **±2 km** on the 1 R^2 km grid employed).

The dual heads were optimized by a weighted Masked Huber Loss (δ = 0.6) across valid samples and time steps

$$L = \frac{\sum_{i,t} w_{i,t} \, \rho_\delta\left(\widehat{y_{i,t}} - y_{i,t}\right)}{\sum_{i,t} w_{i,t} + \epsilon} \quad \text{.....(7)}$$

$$\rho_\delta(e) = \begin{cases} 1/2e^2, & |e| <= \delta \\ \delta(|e| - 1/2\delta), & |e| > \delta \end{cases} \quad \text{.....(8)}$$

where $\rho_\delta(e)$ is the Huber penalty applied to an individual prediction residual, with a cutoff δ, ϵ is a small numerical constant to prevent division by zero, $\widehat{y_{i,t}}$ is the model estimation while $y_{i,t}$ are the observations.

TCN & Bi-LSTM heads were trained using *Adam optimizer* with a learning rate of 1e-3. *L2* regularization was 2e-4, batch size 64, and a maximum of 200 epochs for OOF training (and up to 300 epochs for final heads refitting on all samples). *Early Stopping* was employed with (patience =8, restore best) and *ReduceLROnPlateau* (x0.5, patience =3, min lr=1e-5). Computations ran on float32 with fixed seeds for reproducibility, and during cross-validation, RMSE, MAE, MBE, and R^2 were logged on station- held-out valid months.

Model training and pre-calibration evaluation used 10-fold station-grouped cross-validation across 27 stations. 28 stations were initially considered (20 in the Greater Taipei Area plus 8 in Taoyuan and Yilan), however, Cape Fuguei was excluded, leaving 27 stations for the GroupKFold cross-validation. This exclusion was due to the sample-weighting issue identified at the beginning of model training, which made Cape Fuguei unsuitable for inclusion in the weighted training and evaluation framework. All samples associated with a station, including its jittered neighboring pixels, were assigned to the same fold, so that 2–3 complete stations were withheld in each iteration. For every fold, new TCN and Bi-LSTM heads were trained on the remaining stations using the weighted masked Huber loss, with Adam optimization, a batch size of 64, and a maximum of 200 epochs subject to early stopping and learning-rate reduction. The trained heads then generated predictions only for the withheld stations, and these fold-wise outputs were combined to form complete station-held-out OOF predictions for all samples, as shown

$$\widehat{y_i^{\text{ens}}} = w_{\text{TCNs,m}} \widehat{y_i^{\text{TCN}}} + \left(1 - w_{\text{TCNs,m}}\right) \widehat{y_i^{\text{BiLSTM}}}, \quad 0 \le w_{\text{TCNs,m}} \le 1 \quad \text{.....(9)}$$

where $\widehat{y_i^{\text{TCN}}}$ is the TCN OOF prediction, $\widehat{y_i^{\text{BiLSTM}}}$ is the Bi-LSTM OOF prediction, i denotes a valid OOF sample month prediction, $w_{\text{TCNs,m}}$ represents the TCN ensemble weight assigned for station, s and calendar month, m and $1 - w_{\text{TCNs,m}}$ is the Bi-LSTM ensemble weight.

Station–calendar-month-specific TCN weights, $w_{\text{TCN},s,m}$, were estimated from the corresponding historical OOF prediction–observation pairs using constrained least squares. A global TCN weight, $w_{\text{TCN,global}}$, was additionally estimated from all valid OOF pairs and used as a fallback when insufficient information was available for a station–month-specific weight.

$$w^*_{\text{TCN,global}} = \text{clip}_{[0,1]} \left[\frac{\sum_i \left(\widehat{y_i^{\text{TCN}}} - \widehat{y_i^{\text{BiLSTM}}} \right) \left(y_i - \widehat{y_i^{\text{BiLSTM}}} \right)}{\sum_i \left(\widehat{y_i^{\text{TCN}}} - \widehat{y_i^{\text{BiLSTM}}} \right)^2 + \varepsilon} \right] \text{..(10)}$$

Where $w^*_{\text{TCN,global}}$ represents the optimal global TCN ensemble weight derived from the historical OOF predictions, i indexes all valid historical OOF prediction–observation pairs, $\text{clip}_{[0,1]}$ constrains the weight to the interval [0,1] and ε prevents division by zero.

The resulting raw ensemble predictions were evaluated at both the sample-month level and, after weighted aggregation of jittered samples, the station-

month level. These results represent the predictive performance of the TCN–Bi-LSTM ensemble before calibration and specifically assess spatial generalization to unseen station locations rather than strict future time forecasting.

3.7 Calibration

The raw OOF ensemble predictions were first aggregated from the jittered sample-month level to the station-month level using the corresponding coverage and spatial weights. Stations were grouped into three clusters using six standardized descriptors: the observed $PM_{2.5}$ mean, standard deviation, seasonal amplitude, winter–summer difference, and two principal components derived from the station-associated latent features. Cluster-wise monthly affine parameters were subsequently estimated using ridge regression:

$$\widehat{y_{s,t}^{\mathrm{cal}}} = a_{k,m}\widehat{y_{s,t}^{\mathrm{raw}}} + b_{k,m} \quad (11)$$

Where s is the station, t $monthly\ step$, k denotes the station cluster and m is the calendar month corresponding to t. Residual winsorization, slope clipping, and shrinkage toward a unit slope were used to limit the influence of extreme residuals and prevent excessive correction.

For each station, five calibration alternatives were evaluated: identity, constant-bias, monthly-bias, shrink-affine, and cluster-wise monthly affine correction. Identity retained the raw prediction unchanged. A non-identity station-specific calibrator was considered only when at least 12 valid OOF months were available, except for the cluster-based correction, which could also support stations with fewer observations. The correction was retained only when it reduced RMSE by at least 8% and did not increase the absolute MBE. Among eligible alternatives, the calibrator with the lowest RMSE was selected, with absolute MBE used as the tie-breaker. The rationale for the choice of calibration functions is discussed in the appendix.

The selected station-level corrections were applied to the OOF predictions, after which a small global bias anchor was calculated from the remaining mean residual

$$B_{\mathrm{global}} = -\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{y_i^{\mathrm{cal},0}} - y_i\right) \quad (12)$$

where B_{global} represents the global bias, N is the number of valid OOF station-month prediction-observation pairs, $\widehat{y_i^{\mathrm{cal},0}}$ represents the calibrated $PM_{2.5}$ value before global bias and y_i as surface-level $PM_{2.5}$ value (observed). The final calibrated estimate was obtained as follows

$$\widehat{y_i^{\mathrm{cal}}} = \widehat{y_i^{\mathrm{cal},0}} + B_{\mathrm{global}} \quad (13)$$

where $\widehat{y_i^{\mathrm{cal},0}}$ represents the calibrated $PM_{2.5}$ value before global bias, and $\widehat{y_i^{\mathrm{cal}}}$ represents the final calibrated $PM_{2.5}$ value after global bias. The ensemble weight, station assignments, cluster-specific monthly parameters, selected station calibrators, and global bias were then saved and frozen.

After calibration had been derived from the OOF predictions, new final TCN and Bi-LSTM heads were trained from scratch using all available samples for up to 300 epochs, with early stopping and learning-rate reduction based on training loss. Their predictions were blended using the frozen OOF-derived station–month-specific TCN weights, with the global TCN weight used as a fallback, after which the frozen calibration rules were applied without further fitting to produce final calibrated estimates. Thus, calibration was **not incorporated into the neural-network training loss**; it was applied as a separate post-processing stage to the final blended predictions. Therefore, these results are calibration-adjusted OOF performance (and not a full independent validation outcome) since calibration functions were fitted and selected using the same OOF prediction-observation pairs used for evaluation.

3.8 Station-Assisted Mapping

The twelve-monthly data channel layers for 2023 were first stacked into a four-dimensional tensor of shape (12, H, W, C). Its channels were reordered to match the canonical channel sequence used during model development. Each channel was transformed using the scaling rule fitted from the 72-month (July 2015 – June 2021) training period used in the CNN–Bi-LSTM autoencoder, including *raw*, *z-score*, or *log1p–z-score* scaling. The corresponding means and standard deviations (in the 72 months) were reused unchanged, and invalid values were masked and replaced with zero only after scaling. Consequently, no 2023 observations were used to estimate preprocessing parameters, avoiding scaling leakage.

The scaled 2023 tensor was then combined progressively with the historical 90-month tensor to construct a rolling 90-month sequence for each 2023 month. For instance, January used the latest 89 historical months, including January 2023, whereas December used the latest 78 historical months, including all 12 months of 2023. For each station and jittered sampling location, the 90-month rolling window was passed through the frozen encoder to obtain a 90×64 latent sequence. Six temporal cues and the corresponding jitter offsets (discussed in *3.4 Flowchart*) were then appended, producing a 90x72 input required by the final TCN and Bi-LSTM heads.

The final heads, which had been refitted on all available historical samples, generated one prediction for the final timestep of each rolling window. Their

outputs were combined to obtain the 2023 pre-calibration model estimates:

$$\widehat{y_{s,m}^{\text{estimate}}} = w_{s,m}^{\text{TCN}}\widehat{y_{s,m}^{\text{TCN}}} + \left(1 - w_{s,m}^{\text{TCN}}\right)\widehat{y_{s,m}^{\text{BiLSTM}}} \quad \text{.....(14)}$$

The $w_{\text{TCN},s,m}$, derived from historical OOF predictions was applied to each station-associated sample. The frozen $w_{\text{TCN,global}}$, was used as the fallback when corresponding station-month weights were unavailable. No ensemble weights were estimated using the 2023 observations. The resulting jitter-level estimates were finally averaged to produce one pre-calibrated monthly estimate for each station.

This procedure should not be interpreted as future forecasting because the estimation data for the target month were included in the corresponding 90-month sequence. However, the 2023 surface-level $PM_{2.5}$ observations were not introduced to the 2023 encoder or the dual heads during this stage, but in the subsequent 2023 calibration.

The pre-calibrated monthly estimates were compared with the 2023 surface-level $PM_{2.5}$ observations, and several calibration functions were tested as tuning alternatives for each station. The function producing the highest R^2 score was chosen, and its adjusted values became the tuned 2023 station estimates. Accordingly, the tuned station estimates were model-derived values adjusted using the 2023 observations; they were not the observations themselves. These calibration functions included exponential moving (EMA) bias, affine EMA, EMA affine blend, change-point bias, change-point affine, and hybrid correction, with the rationale and their corresponding equations described in the appendix. Each station correction was clipped using the station-specific adjustment limit K_s associated with its selected 2023 calibration function; the tested limits did not exceed $5\mu g/m^3$.

For each month of 2023, the corresponding 90-month rolling predictor window was applied across every valid grid cell. The frozen encoder, Bi-LSTM, and TCN heads were used to generate an initial $PM_{2.5}$ estimate at each pixel, with its horizontal and vertical offset features set to zero because grid pixels have no jitter displacement. The two regression outputs were combined using the global OOF-derived TCN weight. At each grid cell, the TCN and Bi-LSTM outputs were combined using the frozen global OOF-derived TCN weight, $w_{\text{TCN,global}}$, because station–month-specific weights were not applicable to non-station grid cells. Station-specific corrections were then calculated from the difference between the tuned station estimates and their corresponding pre-calibration model estimates. Each station correction limit was constrained using the correction limit selected earlier.

The station corrections were spatially propagated across the study grid using (KNN – IDW), with each grid cell, up to the three nearest stations within a 120-km search radius, being selected. Their clipped corrections were weighted using inverse squared distance, and the interpolated correction field was calculated as

$$\Delta_{g,m}^{\text{IDW}} = \left(\sum_{s\in\mathcal{N}_{g,m}} w_{g,s}\,\Delta_{s,m}^{\text{clipped}}\right) \div \left(\sum_{s\in\mathcal{N}_{g,m}} w_{g,s}\right) \text{..(15)}$$

where g denotes the target grid cell, m denotes the 2023 month, s is the station anchor, $\mathcal{N}_{g,m}$ represents the three nearest stations with valid corrections in the search radius, and $w_{g,s}$ represents the inverse-squared distance assigned to stations s the relative to the grid cell g. The final station-assisted grid estimate was then obtained by adding the interpolated correction to the pre-calibrated model estimate

$$\widehat{y_{g,m}^{\text{final}}} = \widehat{y_{g,m}^{\text{initial}}} + \Delta_{g,m}^{\text{IDW}} \quad \text{.....(16)}$$

where $\widehat{y_{g,m}^{\text{final}}}$ denoted the final station-assisted grid estimate, $\widehat{y_{g,m}^{\text{initial}}}$ denotes the $PM_{2.5}$ estimate generated by the dual heads at that grid cell, and $\Delta_{g,m}^{\text{IDW}}$ represents the KNN – IDW interpolated station correction.

Where no valid station correction was available within the search radius, the pre-calibrated model estimate was retained unchanged. A representative anchor pixel was assigned to each station using the available core or audited jitter-sampling locations. After spatial correction, the value at each anchor pixel was replaced with the corresponding tuned station estimate. This hard anchoring ensured exact agreement between the mapped and tuned station values at the station locations.

Finally, pixels outside the valid study-domain mask were assigned the NoData value. The corrected monthly grids were exported as georeferenced GeoTIFF files using the original grid dimensions, coordinate reference system, spatial transform, and resolution. A final quality-control check compared the GeoTIFF values at the anchor pixels with the tuned station estimates to confirm that the hard-anchoring procedure was correctly applied (with the difference in values being 1e-8, which is negligible).

3.9 Model Evaluation

The model was evaluated using R^2 score, root mean square (RMSE), mean absolute error (MAE), and mean bias error (MBE). R^2 score significance is to showcase the model's spatiotemporal variance and is given by

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad \text{.....(17)}$$

where y_i are observations, $\hat{y}_i$ are predictions, $\bar{y}$ is the sample mean, and n is the number of evaluated pairs. RMSE monitors the error magnitude from the model estimations and is given by

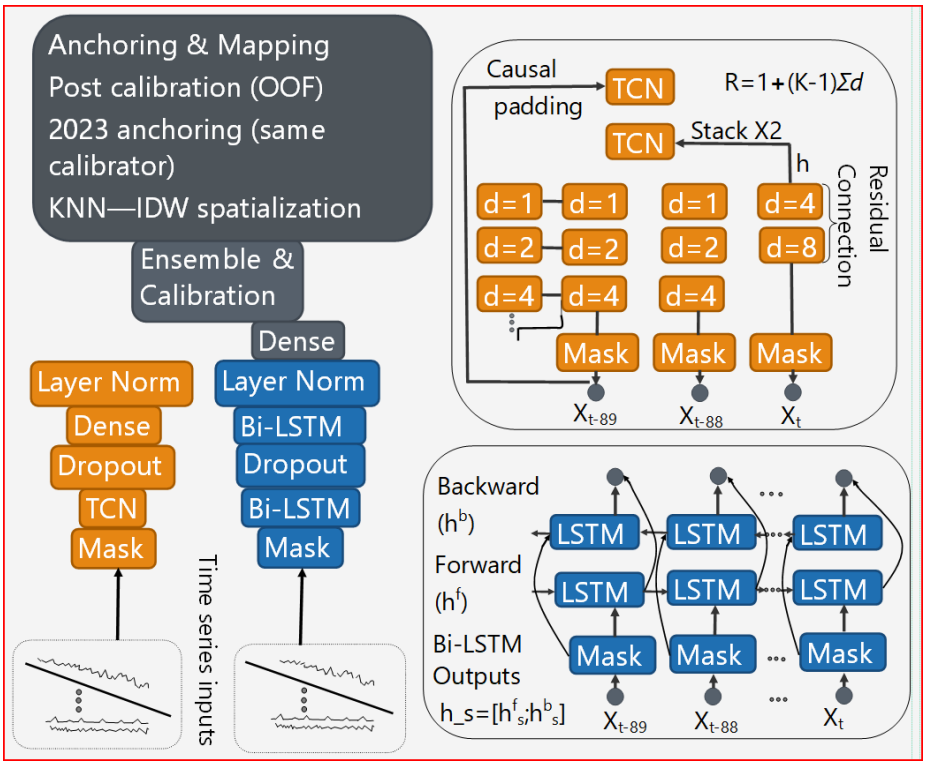


Figure 4 Architecture of the proposed model, calibration, and mapping framework, modified from Wang *et al.* (2025)

$$\text{MSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad \text{.....(18)}$$

MAE showcases the model average absolute error and is formulated as shown

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\, y_i - \hat{y}_i \,| \quad \text{.....(19)}$$

Model bias is reported by MBE using this formula with positive values signifying overestimation, and the reverse is true.

$$\text{MBE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \quad \text{.....(20)}$$

The model validation was from OOF estimations. Here, in each fold, all samples from the held-out station group form the validation set while the rest form the training set. This achieves spatial OOF estimation without station leakage.

The proposed model was benchmarked against classical models – a convolutional neural network (CNN) and long short-term memory (LSTM). The selection of the classical models was informed by previous works (Qin *et al.*,2019; Zhang *et al.*,2021a; Lin *et al.*,2022; Yu *et al.*,2022; Li & Jiang, 2023; Wu *et al.*,2024). The specific model configurations were developed and adapted for the present study. The classical models were developed under the same conditions as the proposed model, to be fair, and then all were evaluated using the four metrics discussed above. The hyperparameters and the specifications of the classical baseline models are reported in the appendix.

4. Results

4.1 Comparison with Classical Models

The proposed model outperformed the CNN and LSTM baseline models both before and after calibration. The proposed model, after calibration, led with R^2 = 0.7960 and RMSE = 2.1606 µg/m³. The calibrated CNN and LSTM models achieved R^2 values of 0.6581 and 0.6226, and RMSE values of 2.7971 µg/m³, and 2.9387 µg/m³, respectively, as shown in Tables 2 and 3.

Table 2 Models' GroupKFold performance before calibration

Model	R^2	RMSE ($\mu g/m^3$)	MAE ($\mu g/m^3$)	MBE ($\mu g/m^3$)
CNN	0.5978	3.0339	2.3205	-0.6738
LSTM	0.5835	3.0874	2.3434	-0.5586
TCN & Bi-LSTM (proposed)	0.7475	2.4035	1.8144	-0.0227

Table 3 Models' GroupKFold performance after calibration

Model	R^2	RMSE ($\mu g/m^3$)	MAE ($\mu g/m^3$)	MBE ($\mu g/m^3$)
CNN	0.6581	2.7971	2.1741	0.0134
LSTM	0.6226	2.9387	2.2696	-0.0102
TCN & Bi-LSTM (proposed)	0.7960	2.1606	1.6401	0.1428

Calibration improved the agreement between the OOF estimates and station observations by reducing systematic bias and overall error. The proposed model's R^2 score increased from 0.7475 to 0.7960, while RMSE decreased from 2.4035 μg/m³ to 2.1606 μg/m³, with comparable improvements observed for the CNN and LSTM models. The shift of MBE toward zero indicates that the systematic underestimation present in the pre-calibrated estimates was substantially reduced, while RMSE and MAE describe the remaining overall error magnitude.

Figures 5–7 visualize the calibrated results reported in Table 3 (rounded off to 3 decimal places), with the solid blue line denoting the 1:1 reference, the dashed orange line showing the fitted regression, and the color indicating sample density.

The proposed model showed the closest agreement with the 1:1 line, achieving the highest R^2 score and lowest RMSE after calibration, followed by the CNN and LSTM models. The fitted regression lines indicate that all models retained some prediction compression at higher $PM_{2.5}$ concentrations, although this effect was least pronounced for the proposed model.

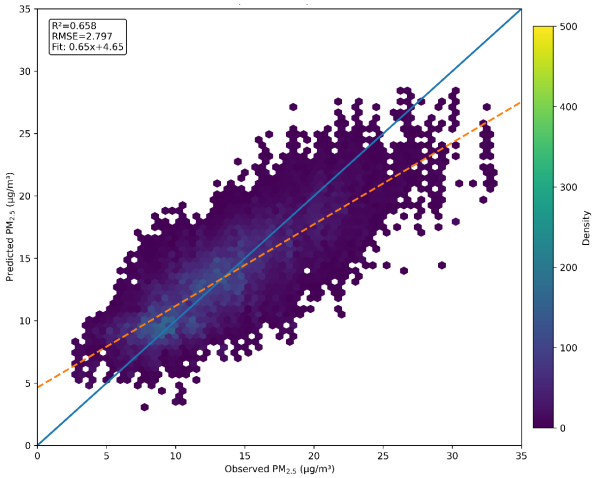


Figure 5 Observed versus calibrated OOF $PM_{2.5}$ estimates from the CNN model

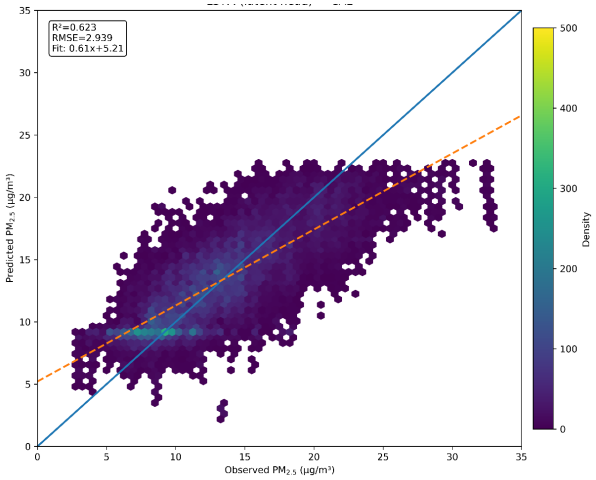


Figure 6 Observed versus calibrated OOF $PM_{2.5}$ estimates from the LSTM model

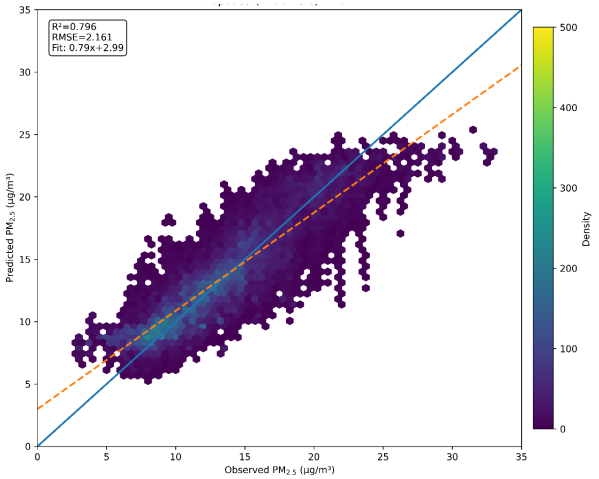


Figure 7 Observed versus calibrated OOF $PM_{2.5}$ estimates from the proposed model

4.2 $PM_{2.5}$ Concentration Maps

The calibrated monthly $PM_{2.5}$ maps, presented in Figure 8 using a common concentration scale, allow direct spatial and temporal comparison throughout 2023. $PM_{2.5}$ concentrations were generally higher from January to May, with peaks in March and April, declined from June to September, and increased slightly from October to December, as shown in Figure 9. This pattern is plausibly consistent with Taiwan's seasonal cycle, where winter and early spring commonly favor pollutant accumulation, whereas summer rainfall and stronger atmospheric dispersion reduce $PM_{2.5}$ concentrations.

The maps also reveal substantial spatial variability across the study area. Banqiao and Cailiao represent densely urbanized environments, Keelung provides a coastal reference, Xizhi represents the eastern urban corridor, and Yangming represents a higher-elevation environment. Although the five stations exhibited similar seasonal patterns, Yangming consistently showed lower concentrations plausibly due to its higher elevation.

The accompanying spatial uncertainty magnitude map in Figure 10 can be interpreted as a station-assisted support index derived from calibrated station-held-out OOF errors. Weighted RMSE and MAE values were first computed for distance bands and then assigned to each grid pixel according to its distance from the nearest station anchor. Lower values indicate stronger station-assisted support and lower expected error, whereas higher values indicate weaker support and greater expected error farther from monitoring stations. The layer is expressed in μg/m³ and should therefore be interpreted as an empirical error-support indicator, rather than a formal probabilistic uncertainty estimate.

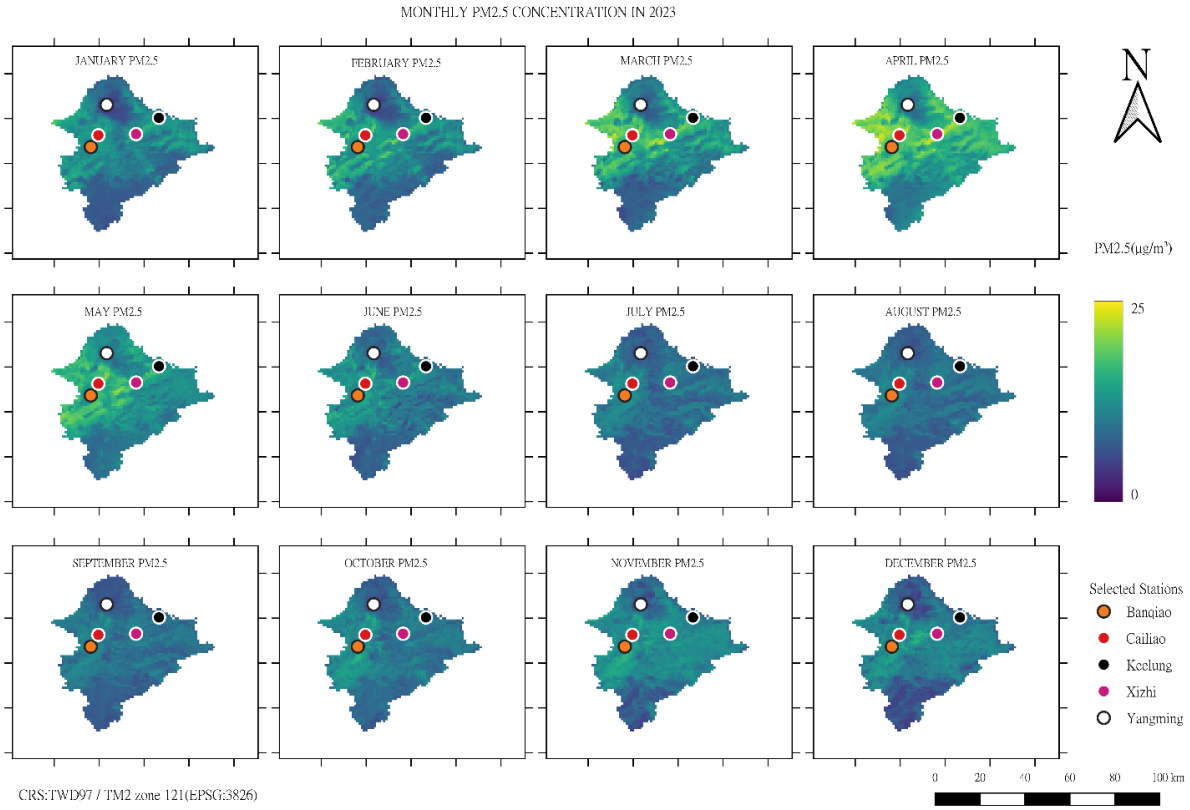


Figure 8 Calibrated monthly station-assisted $PM_{2.5}$ concentration maps for the Greater Taipei Area in 2023

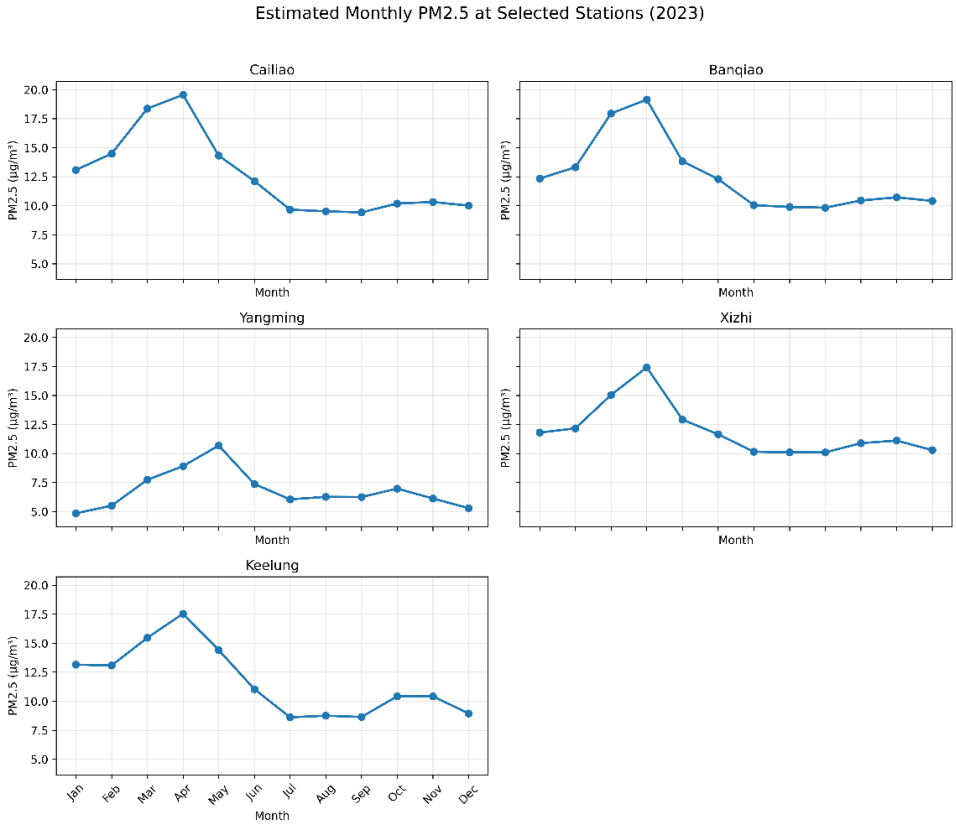


Figure 9 Monthly calibrated $PM_{2.5}$ estimates at the five selected monitoring stations in 2023

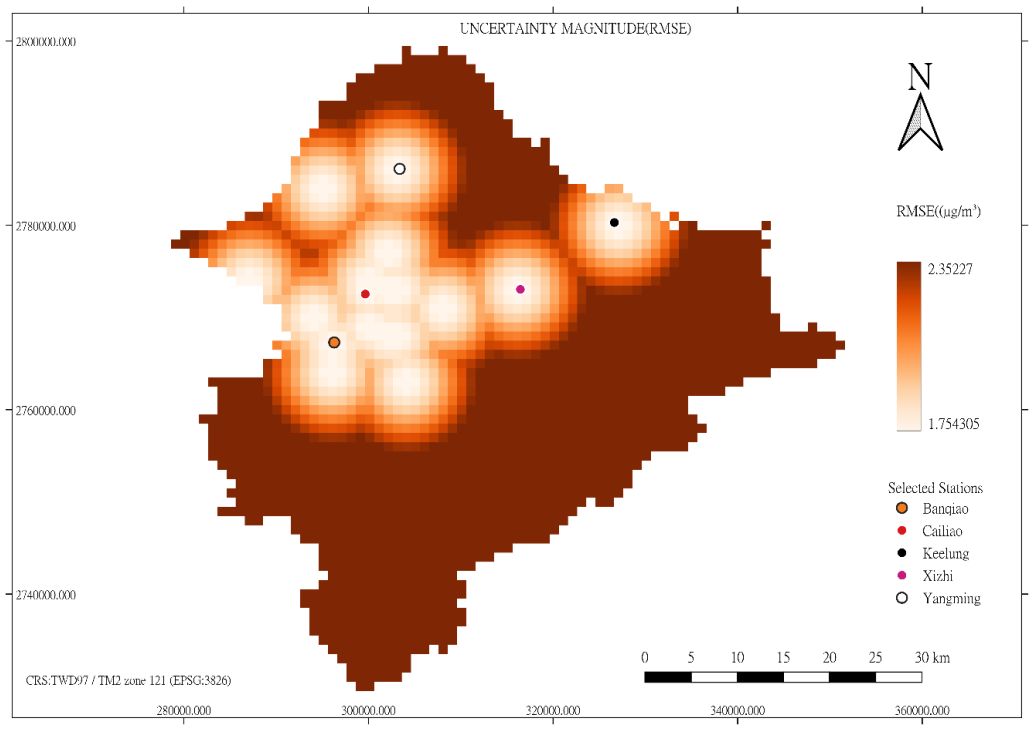


Figure 10 Empirical station-assisted support map for the 2023 $PM_{2.5}$ product, derived from calibrated OOF RMSE and distance to the nearest station anchor

4.3 Seasonal $PM_{2.5}$ Concentration Maps

The 2023 calibrated $PM_{2.5}$ estimates were aggregated by season to assess whether the monthly spatiotemporal patterns persisted. The seasonal definitions commonly adopted in previous studies were employed, with winter defined as December–February, spring as March–May, summer as June–August, and autumn as September–November. As shown in Table 4, concentrations were highest in spring, followed by winter, autumn, and summer. The similar mean and median values indicate that these seasonal patterns were not driven by a small number of extreme pixels, while the higher standard deviations in spring and winter reflect greater spatial variability. Lower summer concentrations are consistent with rainfall and wet scavenging, whereas autumn indicates the gradual return of pollutant accumulation before winter.

Table 4 Pixel-level summary for the calibrated 2023 seasonal $PM_{2.5}$ exposure maps

Season	Mean ($\mu g/m^3$)	Std deviation ($\mu g/m^3$)	Median ($\mu g/m^3$)	Valid pixels
Winter	10.489	2.651	10.840	2478
Spring	13.542	3.272	13.464	2478
Summer	9.520	1.711	9.532	2478
Autumn	10.000	1.797	10.201	2478

Although winter conditions generally favor $PM_{2.5}$ accumulation, the elevated spring concentrations shown in Figure 12 may also reflect peculiar transport processes in northern Taiwan. The region is influenced by long-range aerosol transport, especially biomass burning plumes from Southeast Asia, with previous studies showing Indochina-sourced aerosols reaching Taiwan from February to April. Moreover, wind speed and transboundary pollution during the winter-to-spring transitional period drive this seasonal aerosol accumulation (Lin *et al.*, 2014; Jung *et al.*, 2018; Cheng *et al.*, 2019; Chang, 2019).

Table 5 shows that calibration produced the largest positive adjustment in winter, indicating systematic underestimation in the pre-calibrated estimates. Spring required only a minor overall reduction but showed the greatest spatial variability in calibration-induced changes.

Table 5 Pixel-level summary of calibration-induced changes in the 2023 seasonal $PM_{2.5}$ exposure maps (post-calibration minus pre-calibration)

Season	Mean Δ ($\mu g/m^3$)	Std deviation Δ ($\mu g/m^3$)	Median Δ ($\mu g/m^3$)	Valid pixels
Winter	1.665	0.593	1.820	2478
Spring	-0.160	1.088	-0.038	2478
Summer	-0.663	0.651	-0.519	2478
Autumn	-0.283	0.799	-0.456	2478

Figures 11–14 show higher seasonal $PM_{2.5}$ concentrations across the western, central, and eastern study area, with lower levels in the north and south. Elevated western and central concentrations likely reflect urban emissions and pollutant accumulation, while eastern concentrations may be associated with transport along the urban corridor and weaker dispersion.

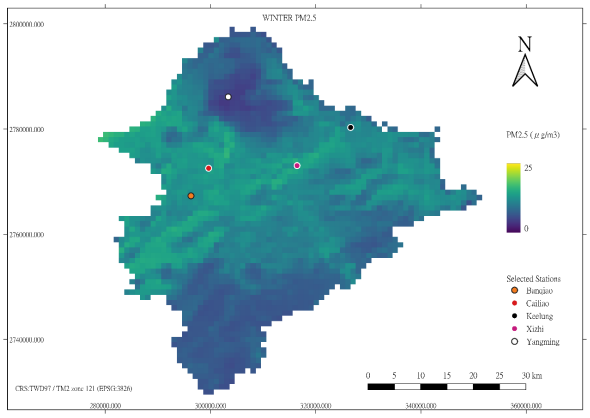


Figure 11 Calibrated winter $PM_{2.5}$ concentration map for 2023

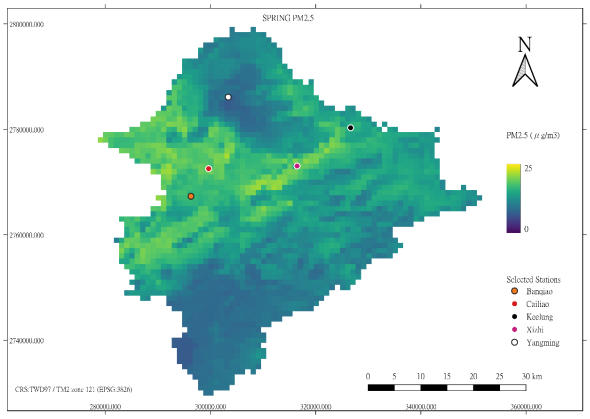


Figure 12 Calibrated spring $PM_{2.5}$ concentration map for 2023

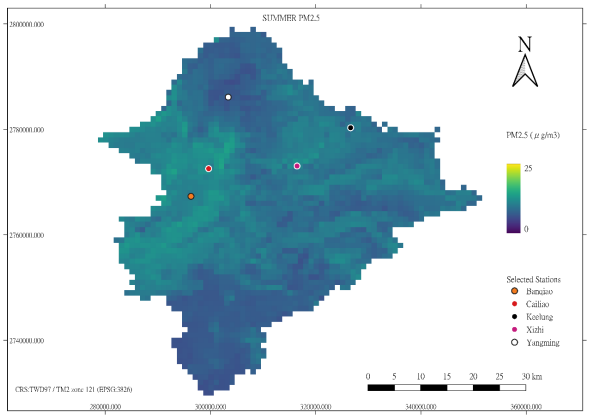


Figure 13 Calibrated summer $PM_{2.5}$ concentration map for 2023

Tables 6 and 7 summarize the proposed model's season-wise pooled OOF performance before and after calibration.

Table 6 Seasonal pooled OOF performance of the proposed model before calibration

Season	R^2	RMSE ($\mu g/m^3$)	MAE ($\mu g/m^3$)	MBE ($\mu g/m^3$)
Winter	0.7026	2.5557	1.9987	0.2150
Spring	0.6920	2.8779	2.1393	-0.6990
Summer	0.5250	1.8858	1.4335	0.1456
Autumn	0.6196	2.2331	1.7205	0.1924

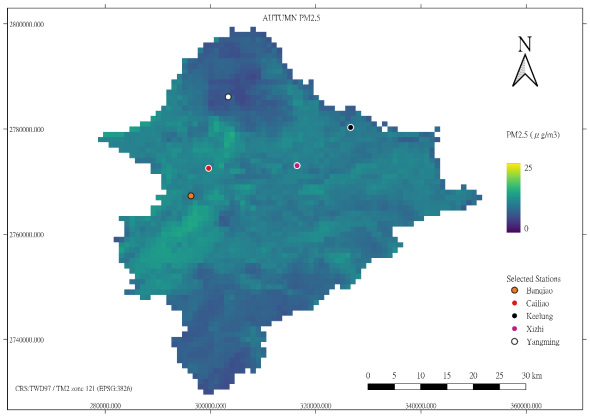


Figure 14 Calibrated autumn $PM_{2.5}$ concentration map for 2023

Seasonal performance varied by metric, with winter achieving the highest R^2, indicating the strongest variance capture, whereas summer produced the lowest RMSE and MAE, partly reflecting its lower $PM_{2.5}$ concentration range. Spring and autumn showed intermediate performance, as shown in Table 6. Despite spring having higher $PM_{2.5}$ concentrations, the proposed model captured winter's concentration variance more accurately, reflecting spring's spatially uneven concentration structure, which is also consistent with its elevated MBE and the patterns discussed in Tables 4 and 5. The concentration variance captured in summer and autumn was low, evidenced by the R^2 score. However, the overall reduced concentrations yielded better RMSE and MAE, indicating lower absolute errors despite weaker variance capture.

Table 7 Seasonal pooled OOF performance of the proposed model after calibration

Season	R^2	RMSE ($\mu g/m^3$)	MAE ($\mu g/m^3$)	MBE ($\mu g/m^3$)
Winter	0.7760	2.2181	1.7433	0.3002
Spring	0.7504	2.5903	1.9538	-0.1746
Summer	0.5736	1.7868	1.3446	0.1846
Autumn	0.6916	2.0107	1.5495	0.2339

In Table 7, the seasonal performance ranking remained unchanged after calibration. Seasonal performance varied by metric, with winter achieving the highest R^2, whereas summer produced the lowest RMSE and MAE, and the arguments in Table 6 still hold. A slight positive residual bias persisted in winter, summer, and autumn as the calibration prioritized correcting systematic bias through R^2, RMSE, and MAE rather than centering MBE around zero, consistent with the Table 3 discussion. Overall, Tables 6 and 7 confirm that the spatiotemporal patterns observed in the monthly $PM_{2.5}$ concentration maps persist across the aggregated seasonal map products.

5. Conclusion

This study developed a hybrid TCN–Bi-LSTM framework to estimate monthly $PM_{2.5}$ concentrations

and produce 1km exposure maps for the Greater Taipei Area. The main objective was achieved by integrating multi-source environmental data with station observations to generate continuous monthly $PM_{2.5}$ surfaces for 2023.

Based on pooled station-grouped OOF evaluation, the proposed model outperformed the CNN and LSTM baselines before and after calibration. After calibration, the proposed model achieved $R^2 = 0.7960$ and RMSE = 2.1606 μg/m³, compared with $R^2 = 0.6581$ and RMSE = 2.7971 μg/m³ for CNN and $R^2 = 0.6226$ and RMSE = 2.9387 μg/m³ for LSTM. Calibration reduced systematic bias and improved overall agreement between the estimates and observations. These results should, however, be interpreted as calibration-adjusted OOF performance rather than fully independent station-level validation because the PM-aware encoder refinement preceded fold separation.

The 2023 maps showed coherent monthly and seasonal patterns. Concentrations were highest in spring, followed by winter, autumn, and summer. Higher $PM_{2.5}$ levels were generally observed across the western and central urbanized areas, whereas lower levels occurred in elevated environments such as Yangming. Spring exhibited the greatest spatial variability and remained comparatively difficult to model, while winter showed the strongest seasonal variance capture. The consistency between the monthly and seasonal products indicates that the framework captured the broader temporal and spatial structure of $PM_{2.5}$ across the study area.

The final maps combined a deep-learning background field with spatialized station-level corrections. The deep-learning model represented the wider spatiotemporal concentration structure, while KNN–IDW propagated bounded corrections from available stations and preserved the tuned values at station pixels. Therefore, the 2023 products should be interpreted as station-assisted exposure maps rather than purely independent model extrapolations. The accompanying station-assisted support layer further identifies where the mapped surface is better supported by observations and where interpretation should be more cautious.

The principal contribution of this study is a practical framework that combines station-held-out evaluation, OOF calibration, deep-learning estimation, and observation-assisted mapping under limited station coverage. The resulting exposure maps can complement existing monitoring networks and provide continuous spatial information where satellite AOD retrievals are incomplete. Nevertheless, sparse station distribution, monthly temporal resolution, limited representation of fine-scale urban heterogeneity, and dependence on 2023 station tuning constrain the generalizability of the final products. Accordingly, the maps are most appropriate as decision-support and exposure-assessment products used alongside, rather than in place of, ground-based observations.

ABBREVIATIONS

The following abbreviations are used in this article:

AOD Aerosol Optical Depth
Bi-LSTM Bidirectional Long Short-Term Memory
CNN Convolutional Neural Network
CTM Chemical Transport Model
CWA Central Weather Administration
EPA Environmental Protection Agency
EC-JRC GHSL European Commission, Joint Research Centre, Global Human Settlement Layer
FFT Fast Fourier Transform
IDW Inverse Distance Weighting
KNN k Nearest Neighbor
LSTM Long Short-Term Memory
LUR Land Use Regression
MAE Mean Absolute Error
MBE Mean Bias Error
MoENV Ministry of Environment
NLSC National Land Surveying & Mapping Centre
OOF Out-of-fold
PM Particulate Matter
TCCIP Taiwan Climate Change Projection Information & Adaptation Knowledge Platform
TCN Temporal Convolutional Network
RMSE Root Mean Square Error
WHO World Health Organization

APPENDIX A

The specifications of the models used in this research study are listed below

CNN – Bi-LSTM Autoencoder

a) Conv1D (64,128 k=3)
b) SE channel = 128
c) Bi-LSTM (64 to 128, 32 to 64 'concat')
d) Dense (128,64, residual)
e) TimeDistributed Dense (26)
f) Temporal SoftMax
g) Epochs =120
h) Batch size = 64
i) Seed= 42
j) Early Stopping, patience=10
k) ReduceLROnPlateau, patience=5
l) Loss – Masked Huber (δ.=0.6)
m) Optimizer – Adam
n) Activation – ReLu, tanh, GeLu
o) Learning rate=0.001
p) Weight decay = 0.0001
q) Layer Norm = 0.000001
r) Dropout= 0.08
s) MHA head/dim= 4/16

CNN Model

a) Conv1D (64,32 k=5)
b) Conv1D (16 k=3)
c) Dense (1)
d) Epochs =300
e) Batch size = 32
f) Seed= 42
g) Early Stopping, patience=12
h) ReduceLROnPlateau, patience=6
i) Loss – Masked Huber (δ.=0.6)
j) Optimizer – Adam
k) Activation – ReLu, tanh
l) Learning rate=0.0003
m) Coverage power=0.58
n) Dropout = 0.15
o) L2 regularizer=0.00001

LSTM Model

a) Conv1D (64,32 k=5)
b) LSTM (64 'true')
c) TimeDistributedDense(1)
d) Epochs =300
e) Batch size = 32
f) Seed=42
g) Early Stopping, patience=12
h) ReduceLROnPlateau, patience=6
i) Loss – Masked Huber (δ.=0.6)
j) Optimizer – Adam
k) Activation – tanh
l) Learning rate=0.0003
m) Dropout=0.15
n) Coverage power =0.58
o) L2 regularizer=0.00001

TCN – Bi-LSTM Regressor

a) Conv1D (1,2,4,8 k=3) X 2
b) Bi-LSTM (64,32)
c) Dense (1)
d) Epochs =200
e) Batch size = 64
f) Seed=42
g) EarlyStopping, patience=8
h) ReduceLROnPlateau, patience=8
i) Loss – Masked Huber (δ.=0.6)
j) Optimizer – Adam
k) Activation – ReLu, tanh
l) Learning rate=0.001
m) Dropout=0.12
n) Coverage power =0.58
o) L2 regularizer=0.0002

APPENDIX B

The calibration functions were selected to address different forms of systematic prediction error while avoiding unnecessary post-processing. The *identity function* retained the raw estimate whenever no correction was warranted and therefore served as the baseline option. *Cluster-wise monthly affine* calibration applied month-specific slopes and intercepts estimated from groups of stations with similar temporal and latent characteristics, allowing stations with limited OOF support to benefit from a shared correction structure. *Constant bias* correction introduced a single station-specific offset to remove persistent overestimation or underestimation. *Shrunk affine* calibration adjusted both prediction scale and offset, while slope clipping and shrinkage toward unity were used to reduce instability and prevent excessive correction. *Monthly bias* calibration applied calendar-month-specific offsets to account for seasonally varying bias without directly modifying the prediction amplitude.

The 2023 station-assisted mapping used a separate set of observation-assisted station-tuning functions from those applied to the historical OOF predictions. The historical calibration was developed for the 2015–2022 station-held-out OOF setting and, therefore, used comparatively restrained functions to limit overfitting. In contrast, the 2023 procedure retrospectively adjusted monthly station estimates within the target year before spatial mapping, allowing station-specific short-term bias changes to be represented.

The candidate functions included EMA bias, which recursively updated an additive correction from recent residuals; affine EMA, which updated both slope and intercept to address changing level and scale bias; changepoint bias, which identified a shift in the within-year residual pattern and applied different additive corrections across regimes; and changepoint affine, which allowed both slope and intercept to differ between the detected regimes. A hybrid alternative selected between EMA and changepoint-bias corrections according to recent performance, while the blend combined EMA-bias and affine-EMA outputs using a rolling performance-based weight. For each station, the function producing the highest 2023 station-level (R^2) was retained, and its adjusted outputs were used as the tuned station estimates for the subsequent delta-based KNN–IDW mapping.

Appendix X. Historical OOF Calibration Functions

(2015–2022 station-held-out out-of-fold calibration)

Notation: $y_{s,t}$ is the observed $PM_{2.5}$ concentration for station s at time t, $\hat{y}^{\mathrm{pre}}_{s,t}$ is the corresponding pre-calibration OOF ensemble estimate, and $m(t)$ denotes the calendar month associated with time t.

A. Identity Calibration

The identity function retained the pre-calibration estimate without adjustment:

$$\hat{y}^{(I)}_{s,t} = \hat{y}^{\mathrm{pre}}_{s,t}$$

B. Constant-Bias Calibration

A station-specific mean residual was calculated as:

$$b_s = \frac{1}{N_s} \sum_{t \in \mathcal{T}_s} \left(y_{s,t} - \hat{y}^{\mathrm{pre}}_{s,t} \right)$$

$$\hat{y}^{(B)}_{s,t} = \hat{y}^{\mathrm{pre}}_{s,t} + b_s$$

where N_s is the number of valid OOF observations available for station s.

C. Monthly-Bias Calibration

A separate additive correction was estimated for each station and calendar month:

$$b_{s,m} = \frac{1}{N_{s,m}} \sum_{t \in \mathcal{T}_{s,m}} \left(y_{s,t} - \hat{y}^{\mathrm{pre}}_{s,t} \right)$$

$$\hat{y}^{(MB)}_{s,t} = \hat{y}^{\mathrm{pre}}_{s,t} + b_{s,\,m(t)}$$

where $N_{s,m}$ is the number of valid OOF observations for station s in month m.

D. Shrink-Affine Calibration

A station-specific slope and intercept were estimated using a regularized affine model:

$$(a^*_s, b^*_s) = \arg\min_{a_s, b_s} \left[\sum_{t \in \mathcal{T}_s} \left(y_{s,t} - a_s \hat{y}^{\mathrm{pre}}_{s,t} - b_s \right)^2 + \lambda_a (a_s - 1)^2 + \lambda_b b_s^2 \right]$$

$$\hat{y}^{(SA)}_{s,t} = a^*_s \hat{y}^{\mathrm{pre}}_{s,t} + b^*_s$$

The regularization terms shrink the slope toward 1 and the intercept toward 0; the fitted slope was constrained to the interval 0.70–1.40.

Appendix X. Historical OOF Calibration Functions (continued)

E. Cluster-Wise Monthly Affine Calibration

Stations were grouped into clusters, and a monthly affine correction was estimated using pooled OOF observations within each cluster:

$$(a^*_{k,m}, b^*_{k,m}) = \arg\min_{a,\,b} \left[\sum_{(s,t) \in \mathcal{D}_{k,m}} \left(y_{s,t} - a \hat{y}^{\mathrm{pre}}_{s,t} - b \right)^2 + \lambda_a (a - 1)^2 + \lambda_b b^2 \right]$$

$$\hat{y}^{(CA)}_{s,t} = a^*_{k(s),\,m(t)} \hat{y}^{\mathrm{pre}}_{s,t} + b^*_{k(s),\,m(t)}$$

where $k(s)$ denotes the cluster assigned to station s.

F. Calibrator Eligibility and Selection

A non-identity calibrator was retained only when it reduced station-level RMSE by at least 8% relative to the identity function and did not increase the absolute mean bias error:

$$\mathrm{RMSE}_s(c) \le 0.92\, \mathrm{RMSE}_s(I)$$

$$|\mathrm{MBE}_s(c)| \le |\mathrm{MBE}_s(I)|$$

$$c^*_s = \arg\min_{c \in \mathcal{C}^{\mathrm{eligible}}_s} \mathrm{RMSE}_s(c)$$

The candidate set was $\mathcal{C} = \{I, B, MB, SA, CA\}$.

G. Global Bias Anchoring

After applying the selected calibration functions, the remaining mean residual was calculated as:

$$B_{\mathrm{global}} = \frac{1}{N} \sum_{i=1}^{N} \left(y_i - \hat{y}^{\mathrm{cal},\,0}_i \right)$$

$$\hat{y}^{\mathrm{cal}}_i = \hat{y}^{\mathrm{cal},\,0}_i + B_{\mathrm{global}}$$

where $\hat{y}^{\mathrm{cal},\,0}_i$ is the calibrated estimate before global bias anchoring.

This appendix summarizes the historical OOF calibration used during model development. It is distinct from the separate 2023 observation-assisted station-tuning functions applied prior to station-assisted mapping.

Appendix X. Candidate Functions for 2023 Station Tuning

The equations below summarize the six candidate functions used to retrospectively adjust the 2023 pre-calibration station estimates before delta-based spatial mapping. For station s and month m, $y_{s,m}$ is the observed surface-level $PM_{2.5}$ concentration and $\hat{y}^{\mathrm{pre}}_{s,m}$ is the corresponding pre-calibration model estimate.

Common notation. Define $\mathrm{clip}(x; L, U) = \min\{\max(x, L), U\}$ and $C_{K_s}(x) = \mathrm{clip}(x; -K_s, K_s)$, where K_s is the station-specific maximum adjustment selected from the tested clipping grid. Residual updates were bounded to $\pm 6\,\mu g\,m^{-3}$.

A. Exponential moving-average (EMA) bias

The most recent residual updates a single additive bias term:

$$\tilde{e}_{s,m-1} = \mathrm{clip}\left(y_{s,m-1} - \hat{y}^{\mathrm{pre}}_{s,m-1}; -6, 6 \right), \tag{1}$$

$$b^{(B)}_{s,m} = (1 - \alpha_b) b^{(B)}_{s,m-1} + \alpha_b \tilde{e}_{s,m-1}, \tag{2}$$

$$\hat{y}^{(B)}_{s,m} = \hat{y}^{\mathrm{pre}}_{s,m} + C_{K_s}\left(b^{(B)}_{s,m} \right). \tag{3}$$

The bias state is initialized at $b^{(B)}_{s,0} = 0$.

B. Affine EMA

Affine EMA updates both the scale and additive bias from the previous month:

$$\rho_{s,m-1} = \mathrm{clip}\left(\frac{y_{s,m-1}}{\max\left(\hat{y}^{\mathrm{pre}}_{s,m-1}, 10^{-3} \right)}; 0.70, 1.40 \right), \tag{4}$$

$$\tilde{e}^{(A)}_{s,m-1} = \mathrm{clip}\left(y_{s,m-1} - a^{(A)}_{s,m-1} \hat{y}^{\mathrm{pre}}_{s,m-1}; -6, 6 \right), \tag{5}$$

$$a^{(A)}_{s,m} = (1 - \alpha_a) a^{(A)}_{s,m-1} + \alpha_a \rho_{s,m-1}, \tag{6}$$

$$b^{(A)}_{s,m} = (1 - \alpha_b) b^{(A)}_{s,m-1} + \alpha_b \tilde{e}^{(A)}_{s,m-1}, \tag{7}$$

$$\hat{y}^{(A)}_{s,m} = \hat{y}^{\mathrm{pre}}_{s,m} + C_{K_s}\left(a^{(A)}_{s,m} \hat{y}^{\mathrm{pre}}_{s,m} + b^{(A)}_{s,m} - \hat{y}^{\mathrm{pre}}_{s,m} \right). \tag{8}$$

The states are initialized at $a^{(A)}_{s,0} = 1$ and $b^{(A)}_{s,0} = 0$.

Interpretation: EMA bias represents gradually evolving additive error, whereas affine EMA also permits the prediction scale to change.

Appendix X. Candidate Functions for 2023 Station Tuning (continued)

C. Changepoint bias and changepoint affine

Let $e_{s,m} = y_{s,m} - \hat{y}^{\mathrm{pre}}_{s,m}$. A candidate changepoint separates the residual record into two regimes:

$$\tau^*_s = \arg\min_{\tau} \left[\sum_{m \le \tau} (e_{s,m} - \bar{e}_{s,1})^2 + \sum_{m > \tau} (e_{s,m} - \bar{e}_{s,2})^2 \right], \tag{9}$$

where $\bar{e}_{s,1}$ and $\bar{e}_{s,2}$ are the mean residuals before and after τ^*_s.
The changepoint-bias estimate is

$$\hat{y}^{(\mathrm{CPB})}_{s,m} = \hat{y}^{\mathrm{pre}}_{s,m} + C_{K_s}\left[\bar{e}_{s,1}\, \mathbb{I}(m \le \tau^*_s) + \bar{e}_{s,2}\, \mathbb{I}(m > \tau^*_s) \right]. \tag{10}$$

For changepoint affine, separate ridge-regularized slopes and intercepts are fitted to the two regimes:

$$\hat{y}^{(\mathrm{CPA})}_{s,m} = \begin{cases} \hat{y}^{\mathrm{pre}}_{s,m} + C_{K_s}\left(a_{s,1} \hat{y}^{\mathrm{pre}}_{s,m} + b_{s,1} - \hat{y}^{\mathrm{pre}}_{s,m} \right), & m \le \tau^*_s, \\ \hat{y}^{\mathrm{pre}}_{s,m} + C_{K_s}\left(a_{s,2} \hat{y}^{\mathrm{pre}}_{s,m} + b_{s,2} - \hat{y}^{\mathrm{pre}}_{s,m} \right), & m > \tau^*_s. \end{cases} \tag{11}$$

D. Hybrid correction

The hybrid uses changepoint bias only when its recent rolling mean-squared error is at least 8% lower than that of EMA bias:

$$I_{s,m} = \mathbb{I}\left[\mathrm{MSE}^{(\mathrm{CPB})}_{s,m} < 0.92\, \mathrm{MSE}^{(B)}_{s,m} \right], \tag{12}$$

$$\hat{y}^{(H)}_{s,m} = I_{s,m} \hat{y}^{(\mathrm{CPB})}_{s,m} + (1 - I_{s,m}) \hat{y}^{(B)}_{s,m}. \tag{13}$$

E. EMA–affine blend

The blend combines the EMA-bias and affine-EMA estimates using a rolling, performance-based convex weight:

$$\hat{y}^{(\mathrm{BL})}_{s,m} = (1 - \beta_{s,m}) \hat{y}^{(B)}_{s,m} + \beta_{s,m} \hat{y}^{(A)}_{s,m}, \tag{14}$$

$$\beta^*_{s,m} = \mathrm{clip}\left(\frac{\sum_{k \in \mathcal{H}_{s,m}} d_{s,k} \left(y_{s,k} - \hat{y}^{(B)}_{s,k} \right)}{\sum_{k \in \mathcal{H}_{s,m}} d^2_{s,k}}; 0, 1 \right), \tag{15}$$

$$d_{s,k} = \hat{y}^{(A)}_{s,k} - \hat{y}^{(B)}_{s,k}, \tag{16}$$

$$\beta_{s,m} = \mathrm{clip}\left(0.7 \beta^*_{s,m} + 0.3 \beta^{\mathrm{prior}}_{s,m}; 0, 1 \right). \tag{17}$$

Here, $\mathcal{H}_{s,m}$ contains the available recent months used to judge performance.

F. Station-specific selection

For each station, the candidate with the highest 2023 station-level coefficient of determination was retained:

$$k^*_s = \arg\max_{k \in C} R^2_s(k), \tag{18}$$

$$\hat{y}^{\mathrm{tuned}}_{s,m} = \hat{y}^{(k^*_s)}_{s,m}, \tag{19}$$

$$C = \{B, A, \mathrm{CPB}, \mathrm{CPA}, H, \mathrm{BL}\}. \tag{20}$$

RMSE, MAE, and MBE were calculated as supporting diagnostics but were not used to select the winning function.

Use note: the tuned values are observation-assisted model estimates, not the station observations themselves. They were subsequently used to derive the station correction deltas for KNN–IDW mapping.

APPENDIX C

Daily AOD was obtained from the JAXA Himawari Monitor Level 3 Daily Aerosol Property product (1DARP) at 500 nm. The *AOT_L3_Merged_Mean* field was used (derived from cloud-screened AOT retrievals), representing the daily mean of the Level 3 merged AOD product. Because this Level 3 product had already undergone JAXA retrieval, cloud-screening, and quality-control processing, no additional QA bit filtering was applied in this study. Missing or invalid values provided in the product were excluded, while no additional numerical AOD threshold was imposed. Monthly AOD for each pixel was calculated as the arithmetic mean of all available valid daily *AOT_L3_Merged_Mean* observations within the month. The number of valid contributing days was retained separately as the AOD daily-valid-count variable.

Land cover information was obtained from the ESA WorldCover 10 m 2021 v200 product. This is a global land-cover classification for the reference year 2021 derived from Sentinel-1 and Sentinel-2 observations. The WorldCover product contains 11 original land cover classes and was produced using a gradient-boosting decision-tree classifier (CatBoost), followed by expert-rule-based refinement. The product has a reported global overall accuracy of 76.7 ± 0.5%, while the reported overall accuracy for Asia is 82.1 ± 1.0%. In this study, the 10 m categorical map was converted to 1km fractional land cover layers. Five classes were retained explicitly: Tree cover (class 10), Grassland (30), Cropland (40), Built-up (50), and Permanent water bodies (80). The remaining WorldCover classes—Shrubland, Bare/sparse vegetation, Snow and ice, Herbaceous wetland, Mangroves, and Moss and lichen—were combined as 'Other.' Fractional coverage within each 1km pixel was calculated by averaging the corresponding binary 10 m class masks.

Population density was derived from the **WorldPop 2020 Taiwan population density raster** and adjusted using **2020 Ministry of the Interior (MOI) population totals** for Taipei, New Taipei, Keelung, Taoyuan, and Yilan to improve consistency with official local statistics. The resulting 2020 population density surface was treated as a static spatial covariate throughout 2015–2023. This assumption was adopted because population density patterns were expected to vary more gradually than the monthly environmental variables; however, it does not imply that population remained unchanged over the study period, and interannual demographic changes were therefore not explicitly represented, which is an acknowledged challenge.

Building footprint and building height were derived from the GHSL 2018 built surface and average building height products. These layers were resampled to the 1km study grid and treated as static urban form covariates throughout 2015–2023. This assumption reflects the relatively slow evolution of urban morphology compared with monthly environmental variables, while acknowledging that construction and redevelopment during the study period were not explicitly represented.

Road network density was calculated by combining road, conventional railway, and high-speed railway vector networks (as of 2021) and summing their line lengths within each 1km grid cell, expressed as km^2. The resulting transport network density was treated as a static spatial covariate during 2015–2023 because major transport infrastructure changes relatively slowly compared with the monthly environmental variables. Nevertheless, network additions or modifications during the study period were not explicitly represented.

APPENDIX D

A total of 28 monitoring stations were initially considered. Jitter sampling was performed on the 1km grid using offsets of ±2 km in both spatial directions, yielding a (5 x 5) neighborhood (25 candidate samples) for each standard station. Yangming used a ±3 km guarded neighborhood ((7 x 7), 49 candidates). Thus, a theoretical maximum of 724 candidates' locations was possible. Nearest-station ownership was enforced, duplicate station pixel locations were removed and samples with no valid samples were discarded, to remain with 590 samples. Each candidate sample retained its parent-station identity together with its horizontal and vertical offsets *dx*, *dy* and ring distance *r* from the station core. Candidate samples were subsequently screened using the temporal validity mask (≥24 valid channels), the nearest-station consistency guard, and the effective weight criterion, as discussed previously in chapter 3. Samples failing these checks were removed before model training to remain with 517 valid samples from 27 stations, with station-specific counts provided in the table below. There was no observation of nearest station violation from the samples.

Station	Initial Samples	Final Samples	Discarded Samples
Banqiao	16	16	0
Cailiao	12	12	0
Cape Fuguei	25	0	25
Datong	5	5	0
Dayuan	25	24	1
Dongshan	25	25	0
Guanyin	25	11	14

Guting	15	15	0
Keelung	25	14	11
Linkou	25	25	0
Longtan	25	25	0
Pingzhen	20	20	0
Sanchong	11	11	0
Shilin	19	19	0
Songshan	25	25	0
Tamsui	25	22	3
Taoyuan	25	25	0
Tucheng	20	20	0
Wanhua	12	12	0
Wanli	25	6	19
Xindian	25	25	0
Xinzhuang	22	22	0
Xizhi	25	25	0
Yangming	49	49	0
Yilan	25	25	0
Yonghe	9	9	0
Zhongli	15	15	0
Zhongshan	15	15	0

References

Basith, S., Manavalan, B., Shin, T.H., Park, C.B., Lee, W.S., Kim, J., and Lee, G., 2022. The impact of fine particulate matter 2.5 on the cardiovascular system: a review of the invisible killer, Nanomaterials, 12(15): 2656, DOI:10.3390/nano12152656.

Berrocal, V.J., Guan, Y., Muyskens, A., Wang, H., Reich, B.J., Mulholland, J.A., and Chang, H.H., 2020. A comparison of statistical and machine learning methods for creating national daily maps of ambient $PM_{2.5}$ concentration, Atmospheric Environment, 222: 117130, DOI:10.1016/j.atmosenv.2019.117130.

Chang, I.-C., 2019. Identifying leading nodes of $PM_{2.5}$ monitoring network in Taiwan with big data-oriented social network analysis, Aerosol and Air Quality Research, 9(12): 2844–2864, DOI:10.4209/aaqr.2019.11.0554.

Chen, G., Li, S., Knibbs, L.D., Hamm, N.A.S., Cao, W., Li, T., Guo, J., Ren, H., Abramson, M.J., and Guo, Y., 2018. A machine learning method to estimate $PM_{2.5}$ concentrations across China with remote sensing, meteorological, and land use information, Science of The Total Environment, 636: 52–60, DOI:10.1016/j.scitotenv.2018.04.251.

Cheng, F.Y., and Hsu, C.H., 2019. Long-term variations in $PM_{2.5}$ concentrations under changing meteorological conditions in Taiwan, Scientific Reports, 9(1): 6635, DOI:10.1038/s41598-019-43104-x.

Chien, L.C., Chen, Y.A., and Yu, H.L., 2018. Lagged influence of fine particulate matter and geographic disparities on clinic visits for children's asthma in Taiwan, International Journal of Environmental Research and Public Health, 15(4): 829, DOI:10.3390/ijerph15040829.

Gupta, S., Pebesma, E., Mateu, J., and Degbelo, A., 2018. Air quality monitoring network design optimisation for robust land use regression models, Sustainability, 10(5): 1442, DOI:10.3390/su10051442.

Hacıoğlu, H.İ., Arı, A., Özkan, A., Elbir, T., Tuncel, G., Yay, O.D., and Gaga, E.O., 2016. A new approach for site selection of air quality monitoring stations: Multi-criteria decision-making, Aerosol and Air Quality Research, 16(6): 1390–1402, DOI:10.4209/aaqr.2014.11.0273.

Hu, J., Jia, Y., Jia, Z.H., He, C.B., Shi, F., and Huang, X.H., 2024. Prediction of $PM_{2.5}$ concentration based on deep learning for high-dimensional time series, Applied Sciences, 14(19): 8745, DOI:10.3390/app14198745.

Jung, C.-R., Hwang, B.-F., and Chen, W.-T., 2018. Incorporating long-term satellite-based aerosol optical depth, localized land use data, and meteorological variables to estimate ground-level $PM_{2.5}$ concentrations in Taiwan from 2005 to 2015, Environmental Pollution, 237: 1000–1010, DOI:10.1016/j.envpol.2017.11.016.

Keelung City Government, 2022. Keelung City, Available at: https://www.klcg.gov.tw/en/klcg2, Assessed April 6, 2026.

Lee, S.J., Serre, M.L., van Donkelaar, A., Martin, R.V.,

Burnett, R.T., and Jerrett, M., 2012. Comparison of geostatistical interpolation and remote sensing techniques for estimating long-term exposure to ambient $PM_{2.5}$ concentrations across the continental United States, Environmental Health Perspectives, 120(12): 1727–1732, DOI:10.1289/ehp.1205006.

Li, J., Chen, J., You, R., and He, Q., 2025. $PM_{2.5}$ concentration prediction: Ultrahigh spatiotemporal resolution achieved by combining machine learning and low-cost sensors, Sensors, 25(17): 5527, DOI:10.3390/s25175527.

Li, T., Shen, H., Yuan, Q., and Zhang, L., 2022. A locally weighted neural network constrained by global training for remote sensing estimation of $PM_{2.5}$, IEEE Transactions on Geoscience and Remote Sensing, 60: 4102513, DOI:10.1109/TGRS.2021.3074569.

Li, T., Shen, H., Zeng, C., Yuan, Q., and Zhang, L., 2017. Point-surface fusion of station measurements and satellite observations for mapping $PM_{2.5}$ distribution in China: Methods and assessment, Atmospheric Environment, 152: 477–489, DOI:10.1016/j.atmosenv.2017.01.004.

Li, W., and Jiang, X., 2023. Prediction of air pollutant concentrations based on TCN-BiLSTM-DMAttention with STL decomposition, Scientific Reports, 13(1): 4665, DOI:10.1038/s41598-023-31569-w.

Lin, C.-Y., Zhao, C., Liu, X., Lin, N.-H., and Chen, W.-N., 2014. Modelling of long-range transport of Southeast Asia biomass-burning aerosols to Taiwan and their radiative forcings over East Asia, Tellus B: Chemical and Physical Meteorology, 66(1): 23733, DOI:10.3402/tellusb.v66.23733.

Lin, S., Zhao, J., Li, J., Liu, X., Zhang, Y., Wang, S., Mei, Q., Chen, Z., and Gao, Y., 2022. A spatial–temporal causal convolution network framework for accurate and fine-grained $PM_{2.5}$ concentration prediction, Entropy, 24(8): 1125, DOI:10.3390/e24081125.

Lo, W.-C., Ho, C.-C., Tseng, E., Hwang, J.-S., Chan, C.-C., and Lin, H.-H., 2022. Long-term exposure to ambient fine particulate matter ($PM_{2.5}$) and associations with cardiopulmonary diseases and lung cancer in Taiwan: A nationwide longitudinal cohort study, International Journal of Epidemiology, 51(4): 1230–1242, DOI:10.1093/ije/dyac082.

New Taipei City Government, 2017. New Taipei City, Available at: https://foreigner.ntpc.gov.tw/home.jsp?id=d127e0ce0f4f407b&utm, Assessed April 6, 2026.

Ong, B.T., Sugiura, K., and Zettsu, K., 2016. Dynamically pre-trained deep recurrent neural networks using environmental monitoring data for predicting $PM_{2.5}$, Neural Computing and Applications, 27(6): 1553–1566, DOI:10.1007/s00521-015-1955-3.

Qin, D., Yu, J., Zou, G., Yong, R., Zhao, Q., and Zhang, B., 2019. A novel combined prediction scheme based on CNN and LSTM for urban $PM_{2.5}$ Concentration, IEEE Access, 7: 20050–20059, DOI:10.1109/ACCESS.2019.2897028.

Requia, W.J., Coull, B.A., and Koutrakis, P., 2019. Evaluation of predictive capabilities of ordinary geostatistical interpolation, hybrid interpolation, and machine learning methods for estimating $PM_{2.5}$ constituents over space, Environmental Research, 175: 421–433, DOI:10.1016/j.envres.2019.05.025.

Song, Y., Huang, B., He, Q., Chen, B., Wei, J., and Mahmood, R., 2019. Dynamic assessment of $PM_{2.5}$ exposure and health risk using remote sensing and geo-spatial big data, Environmental Pollution, 253: 288–296, DOI:10.1016/j.envpol.2019.06.057.

Taipei City Government, 2025. Taipei City, Available at: https://english.gov.taipei/, Assessed April 6, 2026.

US EPA, 2025. Particulate Matter (PM) Pollution, Available at: https://www.epa.gov/pm-pollution/particulate-matter-pm-basics, Assessed April 6, 2026.

Wang, F., Chen, T., Chang, Q., Kao, Y.-W., Li, J., Chen, M., Li, Y., and Shia, B.-C., 2021. Respiratory diseases are positively associated with $PM_{2.5}$ concentrations in different areas of Taiwan, PLoS One, 16(4): e0249694, DOI:10.1371/journal.pone.0249694.

Wang, Z., Crooks, J.L., Regan, E.A., and Karimzadeh, M., 2025. High-resolution estimation of daily $PM_{2.5}$ levels in the contiguous US using Bi-LSTM with attention, Remote Sensing, 17(1): 126, DOI:10.3390/rs17010126.

Wei, J., Li, Z., Lyapustin, A., Wang, J., Dubovik, O., Schwartz, J., Sun, L., Li, C., Liu, S., and Zhu, T., 2023. First close insight into global daily gapless 1 km $PM_{2.5}$ pollution, variability, and health impact, Nature Communications, 14(1): 8349, DOI:10.1038/s41467-023-43862-3.

WHO, 2024. Ambient(outdoor) air pollution, Available at: https://www.who.int/news-room/fact-sheets/detail/ambient-%28outdoor%29-air-quality-and-health, Assessed April 6, 2026.

WHO, 2025. Air quality, energy and health, Available at: https://www.who.int/teams/environment-

climate-change-and-health/air-quality-and-health/health-impacts/types-of-pollutants, Assessed April 6, 2026.

Wu, Y., Xu, Z., Xu, L., and Wei, J., 2024. An improved deep learning approach considering spatiotemporal heterogeneity for $PM_{2.5}$ prediction: A case study of Xinjiang, China, Atmosphere, 15(4):460, DOI:10.3390/atmos15040460.

Yang, K., Shi, Y., Luo, Y., Liu, R., Sun, W., and Sun, M., 2020. Assessing spatiotemporal air environment degradation and improvement represented by $PM_{2.5}$ in China using a two-phase hybrid model, Sustainable Cities and Society, 59: 102180, DOI:10.1016/j.scs.2020.102180.

Yu, T., Wang, Y., Huang, J., Liu, X., Li, J., and Zhan, W., 2022. Study on the regional prediction model of $PM_{2.5}$ concentrations based on multi-source observations, Atmospheric Pollution Research, 13(4): 101363, DOI:10.1016/j.apr.2022.101363.

Zhang, J., Peng, Y., Ren, B., and Li, T., 2021a. $PM_{2.5}$ concentration prediction based on CNN-BiLSTM and attention mechanism, Algorithms, 14(7): 208, DOI:10.3390/a14070208.

Zhang, Y., Cheng, H., Huang, D., and Fu, C., 2021b. High temporal resolution land use regression models with POI characteristics of the $PM_{2.5}$ distribution in Beijing, China, International Journal of Environmental Research and Public Health, 18(11):6143, DOI:10.3390/ijerph18116143.

Zhang, Z., Wang, J., Hart, J.E., Laden, F., Zhao, C., Li, T., Zheng, P., Li, D., Ye, Z., and Chen, K., 2018. National scale spatiotemporal land-use regression model for $PM_{2.5}$, PM_{10}, and NO_2 concentration in China, Atmospheric Environment, 192: 48–54, DOI:10.1016/j.atmosenv.2018.08.046.

基於遙測之都會區 $PM_{2.5}$ 濃度估測之深度學習方法：以大臺北地區為例

魏文森 [1*] 任玄 [2] 林唐煌 [3]

摘要

都市空氣污染仍為重要環境威脅，細懸浮微粒($PM_{2.5}$)對人體健康具有不利影響，因此提升其估算精度十分重要。本研究提出時間卷積網路與雙向長短期記憶網路(TCN－Bi-LSTM)混合模型，使用臺灣環境部 2015 年 7 月至 2022 年 12 月歷史空氣品質、氣象及土地利用資料，估算 2023 年逐月 $PM_{2.5}$，並以 K 近鄰－反距離加權法(KNN－IDW)產製 1 km 濃度圖。GroupKFold 交叉驗證顯示，本模型優於 CNN 與 LSTM，R^2 = 0.7960，RMSE = 2.1606 μg/m³。本研究整合多源環境資料，以降低對衛星 AOD 直接推估 $PM_{2.5}$ 的依賴，並減少雲層造成的 AOD 缺值影響，建立連續 $PM_{2.5}$ 暴露圖。未來將納入獨立地表 $PM_{2.5}$ 觀測資料，以進一步驗證模型表現。

關鍵詞：懸浮微粒(PM)、時空變化、時間卷積網路(TCN)、長短期記憶(LSTM)、雙向長短期記憶(Bi-LSTM)

[1] 國立中央大學 太空及遙測研究中心 碩士生
[2] 國立中央大學 太空及遙測研究中心 教授
[3] 國立中央大學 太空及遙測研究中心 特聘教授
[*] 通訊作者，電話：0981-038-441, E-mail：vincletus2050@gmail.com

收到日期：民國 115 年 06 月 12 日
修改日期：民國 115 年 06 月 29 日
接受日期：民國 115 年 08 月 19 日

航測及遙測學刊　第三十一卷　第 3 期　第 161 – 171 頁　民國 115 年 9 月
Journal of Photogrammetry and Remote Sensing
Volume 31, No.3, 2026, pp. 161 – 171
DOI：10.6574/JPRS.202609_31(3).0002

Hydrocarbon Estimation in Northern Taiwan Using the LightGBM Regression Model

Thia Prahesti [1]　Aji Kusumaning Asri [2]　Yu-Ting Zeng [3]　Chih-Da Wu [4,5,6,7,8*]

Abstract

Total Hydrocarbon (THC) is a critical air pollutant composed of various hydrocarbon compounds, commonly emitted from vehicle exhaust, industrial activities, and fuel combustion. Accurate estimation of THC concentrations is essential for understanding pollution patterns and supporting environmental management. This study applies a Light Gradient Boosting Machine (LightGBM) model to estimate the spatiotemporal variation of THC levels in a part of the Northern Air Quality Zone (AQZ) from June 25, 2015, to December 31, 2020, with 2021 used for testing with unseen data. LightGBM, a machine learning algorithm known for its high accuracy and computational efficiency, achieved strong predictive performance with R^2 values of 0.84 for training, 0.82 for validation, and 0.77 for testing. The average THC concentration in the study area was 2.07 ppm. SHAP (SHapley Additive exPlanations) analysis revealed that nitrogen dioxide (NO_2) and nitrogen oxides (NOx) were the most influential predictors, highlighting their close association with combustion-related emissions. Other important variables included population density, land use types such as plantation forests and paddy fields, and wind speed. Spatially, THC concentrations were higher in the western part of the study area, which is more urbanized and industrialized, compared to the more rural eastern region. These results demonstrate the potential of LightGBM to model complex air pollution patterns and support data-driven air quality strategies.

Keywords: Light Gradient Boosting Machine Regression (LightGBM), Spatial Prediction Model, Total Hydrocarbon (THC)

1. Introduction

Total hydrocarbon (THC) concentration refers to the measurement of a group of organic compounds that play a significant role in air pollution and have been linked to adverse health effects (United States Environmental Protection Agency, 2005). These organic compounds primarily originate from sources such as vehicular emissions, industrial activities, and the use of solvents (Bishop & Stedman, 2008; Environment Canada, 2003). Monitoring and predicting THC levels is crucial for effective air quality management, particularly in urban areas where pollution levels are often higher. Studies have shown that the spatiotemporal variation of THC is influenced by various factors, including land use, meteorological conditions, and traffic patterns (Ambrose *et al.*, 2026). Therefore, understanding the distribution and sources of THC is essential for developing strategies to mitigate their impact on public health and the environment.

Light Gradient Boosting Machine (LightGBM) is a fast and accurate tree-based machine learning method that is well-suited for handling large datasets and complex variable interactions (Guo *et al.*, 2023a; Li *et al.*, 2024). For example, Lu *et al.* (2025) applied a space-time LightGBM model to estimate volatile organic compounds (VOCs) in Shanghai, producing high-resolution hourly prediction maps (1 km, next 24 h) with strong performance, achieving R^2 values ranging from 0.527 to 0.938 and MAE values between 41.5 and 126 ppb. Similarly, Thongthammachart *et al.* (2022) integrated LightGBM into a land use regression framework for NO_2 and $PM_{2.5}$ estimation, showing that LightGBM outperformed Random Forest and Extreme Gradient Boosting, particularly in capturing high concentration levels that are critical for health risk assessment.

While previous studies have applied machine learning models such as LightGBM to estimate air

[1] Ph.D Student, Department of Geomatics, National Cheng Kung University
[2] Postdoctoral Research Fellow, Department of Geomatics, National Cheng Kung University
[3] Assistant Researcher, Department of Geomatics, National Cheng Kung University
[4] Professor, Department of Geomatics, National Cheng Kung University
[5] Adjunct Research, National Institute of Environmental Health Sciences, National Health Research Institutes
[6] Adjunct Professor, Innovation and Development Center of Sustainable Agriculture, National Chung Hsing University
[7] Adjunct Research, Research Center for Precision Environmental Medicine, Kaohsiung Medical University
[8] Joint Appointment Research, Chronic Diseases and Health Promotion Research Center, Chang Gung University of Science and Technology
[*] Corresponding Author, E-mail: chidawu@mail.ncku.edu.tw
Received Date: Jul. 16, 2025
Revised Date: Sep. 17, 2025
Accepted Date: Aug. 21, 2026

pollutants (Lu *et al.*, 2025; Thongthammachart *et al.*, 2022), few have explored their application to total hydrocarbon (THC) concentrations, and the spatiotemporal variations of THC, especially in Northern Taiwan, remain underexplored. This study addresses these gaps by applying a LightGBM regression model to estimate THC concentrations using data from multiple monitoring stations, considering that THC primarily originates from sources such as vehicular emissions, industrial activities, and the use of solvents. The study also provides insights for air quality management, highlighting the novelty and significance of this research in advancing previous work on air pollution modeling (Lu *et al.*, 2025; Thongthammachart *et al.*, 2022).

2. Method

This research separates the workflow into three parts, the dataset preprocessing of feature points' density map, the comprehensive deep Q network model for path finding, and the post-processing for smoothing the path, shown as Figure 1 below.

2.1 Study area

Taiwan, historically known as Formosa, is located north of the Philippines, east of China, and south of Japan. For air quality management, the Ministry of Environment (MOENV) divides the island into seven Air Quality Zones (AQZs), one of which is the Northern Area that includes Taipei City, New Taipei City, Keelung, and Taoyuan (Risk Society and Policy Research Center, 2026).

The study area covers approximately 3,678 km² in northern Taiwan, a region with some of the highest population densities in the country, particularly in Taipei. Nevertheless, motorcycles and private automobiles remain the dominant travel modes and are major sources of THC, thereby intensifying regional air pollution (Bishop & Stedman, 2008).

2.2 Air pollution database

Hourly data on THC concentrations have been recorded since June 26, 2015, at 8 monitoring stations in Northern AQZ managed by MOENV (Figure 1). This study utilizes data spanning from 2015 to 2018, where 80% of the dataset is designated for model training and the remaining 20% is used for validation. Additionally, data from monitoring stations in 2019 were collected for testing to evaluate the model's reliability. The Ministry of Environment Taiwan adheres to standards and uses instruments comparable to those of the United States regulation (Department of Monitoring and Information, Ministry of Environment, Taiwan. (2024). To investigate potential correlations between THC and its co-pollutants, such as PM_{10}, $PM_{2.5}$, O_3, SO_2, NO_2, and NO_x, we also collected measurements for these pollutants from the Ministry of Environment's published database.

2.3 Land-use and geospatial databases

To support the development of the spatial prediction model, a wide array of environmental and geospatial explanatory variables were compiled from meteorological, land-use, topographic, and anthropogenic datasets (Hsu *et al.*, 2022). Meteorological variables including wind speed, wind direction, relative humidity, temperature, ultraviolet (UV) radiation, and precipitation were obtained from the Central Weather Bureau of Taiwan, as these factors significantly influence pollutant dispersion, chemical transformation, and atmospheric behavior of THC.

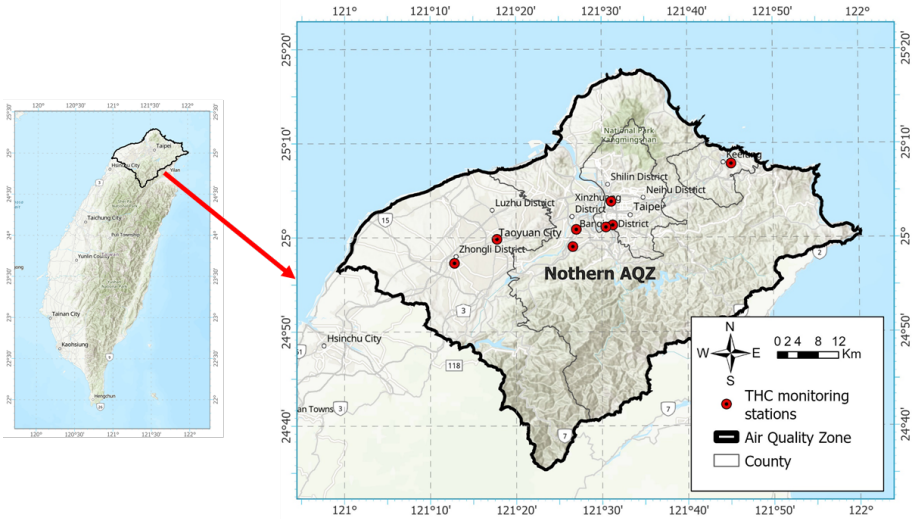


Figure 1 Distribution of Ministry of Environment monitoring stations measuring THC concentrations in Northern Area

Land-use and geospatial data were sourced from the National Land Surveying and Mapping Center (NLSMC), covering the years 2015–2018. These data included classifications of land types such as farmlands, forests, residential and industrial zones, airports, and water bodies. Vegetation characteristics were captured using the MODIS Normalized Difference Vegetation Index (NDVI), with a spatial resolution of 250 m × 250 m. Topographic variables, such as elevation and slope, were derived from a 20 m × 20 m Digital Terrain Model (DTM), which helped account for the influence of terrain on air movement and pollutant accumulation.

To characterize human activities and potential THC emission sources, we collected data from the Kingwaytek landmark database as a Point of Interest (POI), the Ministry of Finance's business tax registry, and the Ministry of Economic Affairs' industrial site records. These anthropogenic activities, including transportation, industrial operations, and fossil fuel combustion (Environment Canada, 2003; Stewart *et al.*, 2021), enable the identification of localized THC emission hotspots for modeling their impact on air quality. Social factors, including the locations of major Chinese temples and restaurants (Wu *et al.*, 2017) and annual township-level population data (2015–2021), were incorporated as predictive variables to capture anthropogenic and cultural influences on THC concentrations. Seasonal variations were also considered by deriving a seasonal variable from the date of each observation, allowing the model to capture changes in pollutant concentrations across different times of the year. Spatial predictor variables were then calculated within multiple circular buffer zones centered at each air quality monitoring station. Specifically, buffer radii included: 50 m, 150 m, 250 m, 500 m, 750 m, 1000 m, 1250 m, 1500 m, 1750 m, 2000 m, 2500 m, 3000 m, 4000 m, and 5000 m. These varying distances allowed for multi-scale analysis of land-use influence, enabling the model to capture both localized and regional factors affecting THC concentrations in Northern AQZ.

2.4 Study framework

To estimate THC concentrations, this study employed a spatiotemporal modeling approach using LightGBM regression as the primary machine learning algorithm. As illustrated in Figure 2, the workflow comprised four main steps: data preparation, model development, model validation, and spatiotemporal prediction of THC concentrations in the Northern Area of Taiwan.

2.5 The proposed LightGBM model approach

LightGBM is a powerful boosting technique tailored for handling large-scale datasets with high predictive performance. It incorporates strategies such as Gradient-based One-Side Sampling (GOSS), which accelerates the training process by focusing on instances with higher prediction errors, and Exclusive Feature Bundling (EFB), which simplifies high-dimensional data by combining features that do not overlap. Unlike conventional boosting methods, LightGBM adopts a leaf-wise tree growth approach, enabling more refined splits that improve prediction accuracy and reduce the risk of overfitting (Guo *et al.*, 2023b; Li *et al.*, 2024).

To enhance model performance while reducing the risk of overfitting, hyperparameter tuning was performed on key parameters, including learning rate, number of leaves, maximum tree depth, feature fraction, and bagging fraction (Tırınk, 2025; Vertes, 2020). The model's generalization ability was evaluated using an 80/20 train–test split and k-fold cross-validation (GeeksforGeeks, 2025). Additionally, regularization and controls on tree complexity, such as limiting tree depth and the number of leaves, were applied to prevent overly complex models (Bahmani, 2020).

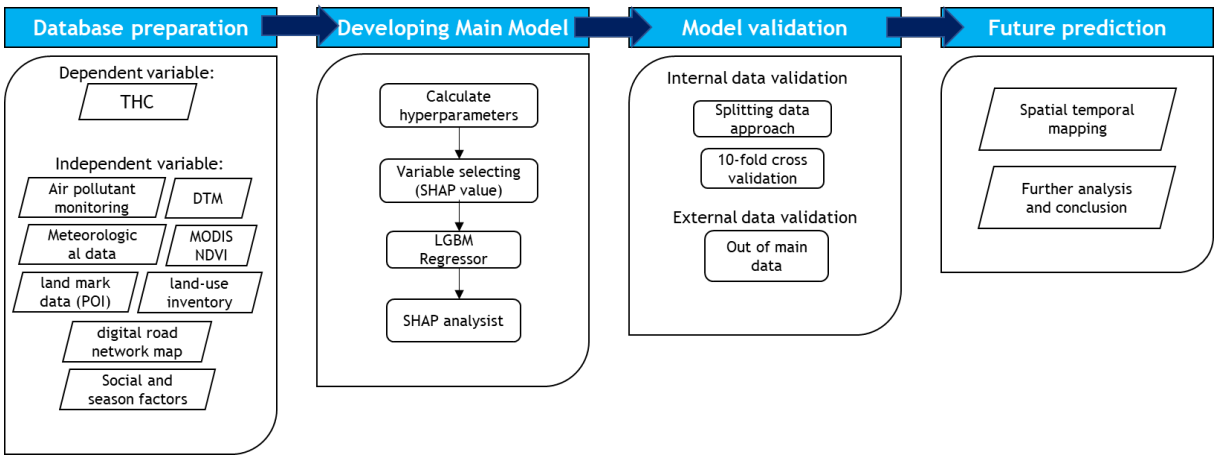


Figure 2 Study framework

This study developed prediction models to estimate monthly THC concentrations in Taiwan by machine learning algorithms. We applied machine learning algorithm (LightGBM) to enhance model performance, given their ability to capture non-linear relationships between THC levels and predictor variables. The process begins with hyperparameter tuning to identify the most effective combination of model parameters. Next, feature selection is performed to eliminate unimportant or redundant variables, ensuring the model focuses only on the most relevant inputs. With the optimal parameters and selected features, the model is then trained to achieve the best performance. Finally, SHAP (SHapley Additive exPlanations) analysis is used to interpret the model by showing how each feature contributes to the predictions. To simplify the hyperparameter tuning process, we focused on algorithms with automated tuning features. Model performance was evaluated by calculating the R^2 values from the 80/20 data split, cross-validation, and external validation. Each algorithm was tuned to optimize predictive accuracy, and the best-performing model was chosen to generate final predictions.

2.6 Model validation

We employed three categories of validation methods, encompassing three distinct approaches, to assess model reliability. These included an 80%/20% data split, 10-fold cross-validation and external validation (using 2019 data). In the 10-fold cross-validation procedure, the dataset was partitioned into ten equal subsets, with each subset used once as a validation set while the remaining nine subsets served as the training data. We recalculated R^2 and other metrics to confirm model stability. To evaluate model accuracy, we used correlation coefficient (r), R^2, and adjusted R^2 values, alongside mean squared error (MSE), mean absolute error (MAE), and root mean squared error (RMSE) to measure prediction accuracy, defined as follows (1-6). These metrics were used complementarily to provide a comprehensive evaluation of model performance, and no single metric was considered the sole indicator of model accuracy.

$$\boldsymbol{r} = \frac{\left(\sum(y_i - \bar{y})(\hat{y}_i - \hat{\bar{y}})\right)}{\sqrt{\sum(y_i - \bar{y})^2 \sum(\hat{y}_i - \hat{\bar{y}})^2}} \quad \text{.....(1)}$$

$$R^2 = 1 - \left(\frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}\right) \text{.....(2)}$$

$$Adjusted\ R^2 = 1 - \frac{(1-R^2)(n-1)}{n-p-1} \quad \text{.....(3)}$$

$$MSE = \left(\frac{1}{n}\right)\sum(y_i - \hat{y}_i)^2 \quad \text{.....(4)}$$

$$MAE = \left(\frac{1}{n}\right)\sum|y_i - \hat{y}_i| \quad \text{.....(5)}$$

$$RMSE = \sum\left(\left(\frac{1}{n}\right)\sum(y_i - \hat{y}_i)^2\right) \quad \text{.....(6)}$$

Where y_i is the observed THC concentration, $\hat{y}_i$ is the predicted THC concentration, $\bar{y}$ is the mean of the observed THC concentrations, $\hat{y}$ is the mean of the predicted THC concentrations, n denotes the number of observations, p denotes the number of predictor variables, and i represents the index of each observation.

2.7 Geographical mapping of the spatiotemporal variations of THC

Using the model with the best explanatory power, we generated spatiotemporal maps of THC concentrations at a 50 m × 50 m grid resolution. First, we converted the selected geographic predictors from raster format to ASCII format. Then, the ensemble model was applied to estimate predictions for each 50 m grid center using the Python platform. In the final step, the ASCII-format predictions were converted back into raster format to create the THC concentrations prediction map. For spatial data analysis, we utilized ArcGIS Pro, statistical analyses were performed in R 4.3.2, and machine learning models were developed using Python.

3. Results

3.1 Descriptive statistics of THC concentrations

Figure 3(a) presents the histogram of THC concentrations measured across all stations from 2015 to 2018, showing the overall distribution of observed values over the study period. Figure 3(b) illustrates the average THC concentration at each monitoring station, highlighting spatial differences in pollutant levels across the study area. As a result, the overall THC concentration in the dataset (n = 8,369) exhibited a mean of 2.10 ppm and a median of 2.07 ppm, indicating a relatively symmetric distribution. The concentrations ranged from a minimum of 1.6 ppm to a maximum of 4.23 ppm, with a standard deviation of 0.19 ppm, reflecting moderate variability across the study period.

At the station level, the average THC concentrations varied, with Keelung showing the lowest mean of 1.97 ppm and Zhongli the highest at 2.27 ppm. Other stations recorded intermediate values, including 2.05 ppm at Guting, 2.07 ppm at Tucheng, 2.09 ppm at Zhongshan, 2.10 ppm at Taoyuan, 2.12 ppm at Yonghe, and 2.15 ppm at Banqiao, highlighting spatial heterogeneity in THC levels across the study area.

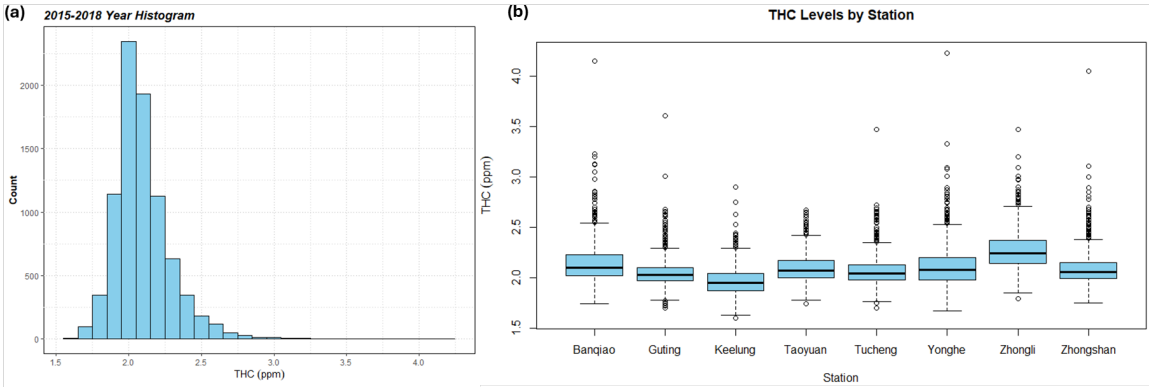


Figure 3 (a) histogram of THC concentrations from 2015 to 2018, (b) average concentration at each station

3.2 Explanatory variables of the LightGBM model and model performance

Figure 4 presents the statistical values, including R^2 for LightGBM. The model showed strong predictive performance across different evaluation stages. During training, the model achieved an R^2 of 0.84, indicating a good fit to the training data. In validation, the R^2 was slightly lower at 0.82, demonstrating the model's ability to generalize well. External testing with unseen data resulted in an R^2 of 0.77, confirming reasonable predictive accuracy on independent datasets. Overall, the combined total R^2 was 0.84, reflecting consistent performance. Additionally, cross-validation yielded an average R^2 of 0.80, supporting the model's robustness and stability.

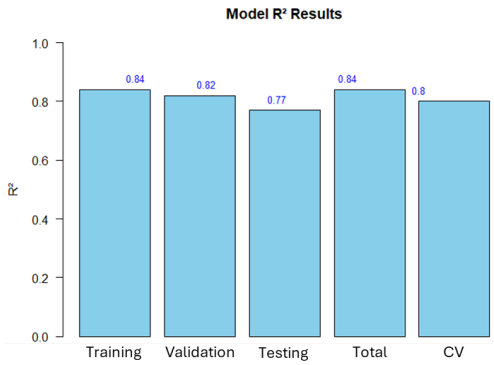


Figure 4 Coefficient of determination (R^2) of the LightGBM model for the training, validation, testing, total, and cross-validation datasets

Figure 5 presents the SHAP analysis from the LightGBM spatiotemporal model for THC mapping. Panel (A) shows the mean absolute SHAP values, representing the overall importance of each variable in the model, while Panel (B) shows the impact of each variable on the LightGBM model for THC mapping based on individual SHAP values. In Panel (B), positive SHAP values indicate that a variable contributes to higher predicted THC concentrations, whereas negative SHAP values indicate a contribution toward lower predicted THC concentrations. The color gradient represents the feature values, with blue indicating lower values and red indicating higher values.

The mean absolute SHAP values are used to rank the variables according to their relative contribution to the model predictions. NO_2 (ppm) shows the highest importance, with a mean absolute SHAP value of 0.055, indicating its strong effect on model predictions. This is followed by NOx (ppm) at 0.041, and Population (1750m) at 0.032, highlighting the relevance of air pollutants and population density in estimating THC levels. Plantation forest (4000m) and Wind Speed (m/s) also contribute significantly, reflecting the role of land use and meteorological conditions. Other important variables include Paddy fields (2000m), Plantation forest (2500m), and Park (1750m), underscoring the spatial influence of surrounding land use types. Overall, these findings demonstrate that both environmental pollutants and land characteristics are key drivers in the model's predictions.

3.3 Model validation

Table 1 show the validation of the LightGBM model includes eleven different validation processes based on several statistical values. These include the coefficient (r), R^2, and adjusted R^2 values, as well as MSE, MAE, and RMSE to measure prediction accuracy from 80%/20% data split (training and validation), 10-fold cross-validation and external validation. The model performs well across different stages. In training, it explains 84% of the variation in the data (R^2 = 0.84) with low errors (RMSE = 0.10, MAE = 0.06). Validation results are similar, showing the model generalizes well (R^2 = 0.82). Testing with unseen data is slightly lower (R^2 = 0.77), but still indicates good accuracy. Overall, the combined results show consistent performance (R^2 around 0.83-0.84). Cross-validation confirms the model's stability with an average R^2 of 0.80 and only a small increase in error (RMSE = 0.11). This means the model is reliable and effective for predicting the target variable.

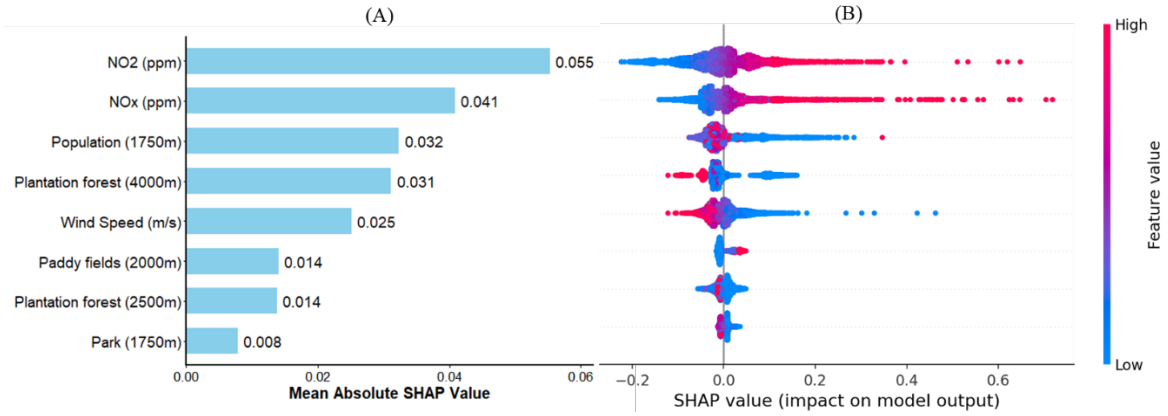


Figure 5 Impact of each variable on the LightGBM model for THC mapping : (a) mean absolute SHAP values, and (b) SHAP summary plot showing the direction and magnitude of each variable's contribution

Table 1 The validation of the LightGBM model

Result	R^2	adj R^2	RMSE	MSE	MAE
Training	0.84	0.84	0.10	0.01	0.06
Validation	0.82	0.82	0.10	0.01	0.06
Testing	0.77	0.77	0.10	0.01	0.07
Total	0.84	0.83	0.10	0.01	0.06
CV	0.80	0.80	0.11	0.01	0.06

3.4 Spatio temporal distribution of THC concentration

Figure 6 presents the prediction maps of THC concentrations estimated using the LightGBM model. Subfigure (a) shows the average THC concentrations for 2015–2018, while subfigure (b) presents seasonal maps for spring, summer, fall, and winter. The maps use dark green to indicate low THC concentrations and red to represent high concentrations. The predicted average THC concentration over the four years is 2.07 ppm. Among the seasons, spring exhibited the highest average concentration at 2.12 ppm, while summer had the lowest at 2.05 ppm. Notably, higher THC concentrations were observed in the western part of the Northern AQZ and lower concentrations in the eastern area.

4. Discussion

The LightGBM spatiotemporal model demonstrated robust predictive performance in estimating THC concentrations across different evaluation stages. During the training phase, the model achieved a high R^2 value of 0.84, indicating that it was able to explain 84% of the variability in the THC data. This strong fit was maintained during validation phase, with an R^2 of 0.82, which suggests that the model generalizes well to unseen data without substantial overfitting. The external testing dataset, which represents truly independent samples, yielded a slightly lower R^2 of 0.77, reflecting the expected reduction in predictive accuracy when applied to new data but still indicating reliable performance. The overall total R^2 of 0.84, along with a consistent RMSE of approximately 0.10, reinforces the model's robustness and stability in capturing the complex spatiotemporal dynamics of THC concentrations. Cross-validation further supported these findings, with an average R^2 of 0.80, confirming the model's consistent behavior across multiple subsets of data.

When comparing the performance of our LightGBM spatiotemporal model with previous studies, it is evident that our model achieves competitive accuracy in estimating THC concentrations. For instance, Su *et al.* (2010) utilized Land Use Regression (LUR) techniques to capture the spatial distribution of volatile organic compounds, including THC, in Toronto, reporting moderate R^2 values between 0.66 and 0.68. Similarly, Doris *et al.* (2024) applied LUR models to evaluate THC exposure near natural gas production areas in northeastern British Columbia, achieving a wider range of R^2 from 0.73 up to 0.91 depending on the specific site and time period. In contrast, our model consistently attained R^2 values of 0.84 during training and 0.82 in validation, with testing R^2 of 0.77, situating its predictive power well within or above the range of these previous LUR studies. This suggests that the integration of advanced machine learning techniques like LightGBM, combined with a comprehensive set of environmental, pollutant, and land use variables, can provide improved accuracy and robustness in THC estimation compared to traditional regression approaches. Such enhanced performance is particularly

valuable for capturing the complex spatiotemporal patterns of THC in diverse environments, ultimately supporting more precise exposure assessments and better-informed air quality management decisions (Lu *et al.*, 2025; Thongthammachart *et al.*, 2022).

Beyond traditional performance metrics, the SHAP (SHapley Additive exPlanations) analysis provided valuable insights into the contribution of individual features to the model's predictions. The analysis identified NO_2 (ppm) as the most influential variable, underscoring the critical role of nitrogen dioxide concentrations in driving THC levels. NOx (ppm) was also found to have a significant impact, highlighting the broader importance of nitrogen oxides in air pollution chemistry relevant to hydrocarbons (Bishop & Stedman, 2008; Geddes *et al.*, 2021; Guaman *et al.*, 2022; Environment Canada, 2003). Population density within a 1750-meter radius emerged as another important predictor, suggesting that human activities and urbanization strongly affect local THC concentrations (Coffin *et al.*, 2021). Land use variables such as plantation forests and paddy fields at varying distances (e.g., 2500 m to 4000 m) were also influential, indicating that the type and extent of vegetation and land cover affect the dispersion and accumulation of hydrocarbons (Kansal, 2009; Zhou *et al.*, 2023). Plantation forest at the 4000 m buffer showed a higher ranking than at 2500 m because a larger buffer of point density captures broader spatial patterns, which better represent pollutant dispersion and regional variation, while smaller buffers mainly reflect localized effects with lower explanatory power. Additionally, meteorological factors, particularly wind speed, played a notable role, consistent with its known effect on pollutant transport and dilution Terrain variables, including higher elevations and steeper slopes, were generally associated with lower THC concentrations due to enhanced airflow and pollutant dispersion, whereas valleys and low-lying areas tended to accumulate more THC (Bose & Chowdhury, 2023; Cheng, 2002).

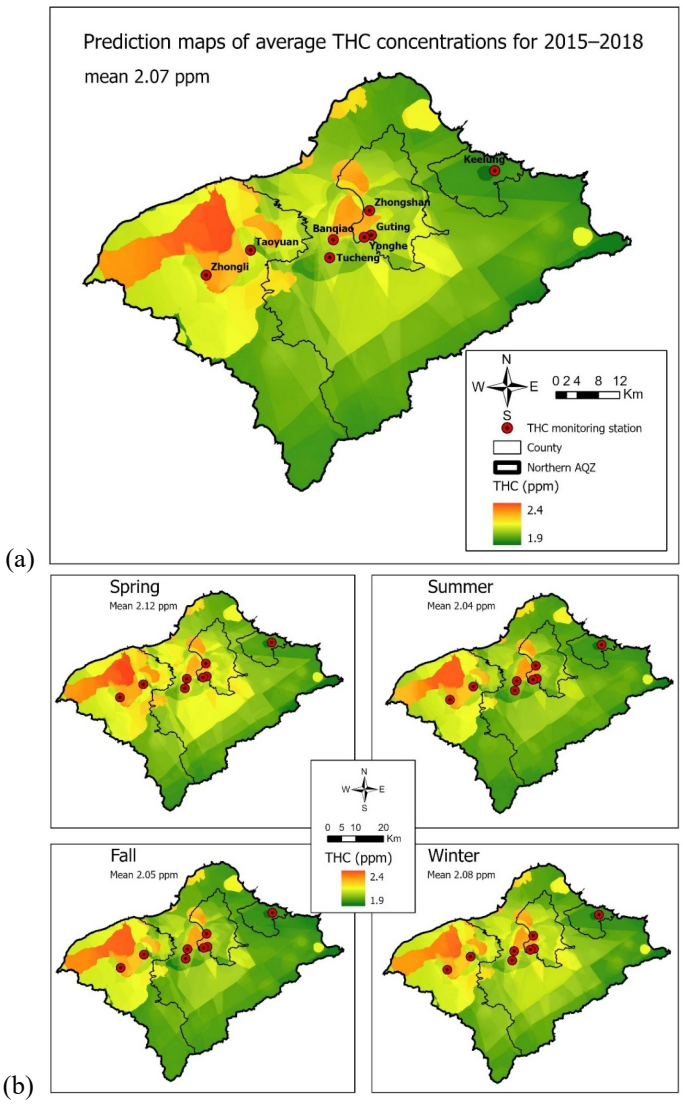


Figure 6 (a) Prediction maps of average THC concentrations for 2015–2018 using LightGBM, (b) the maps for each season (spring, summer, fall, and winter)

These combined results emphasize that THC concentrations are governed by a complex interplay of both anthropogenic emissions and environmental conditions. The ability of the model to integrate air pollutant measurements, population data, land use characteristics, and meteorological variables allows for a more comprehensive and accurate estimation of THC distribution in space and time. This holistic approach is critical for developing effective air quality management strategies and for understanding the sources and dynamics of hydrocarbon pollution in the studied region. Future research could further improve model performance by incorporating additional variables such as traffic intensity or industrial emissions, as well as by exploring temporal trends with finer resolution data.

Seasonal variations in predicted THC concentrations are influenced by meteorological conditions that affect pollutant dispersion (Bose & Chowdhury, 2023; Li *et al.*, 2021). In Northern Taiwan, spring showed the highest average THC concentration (2.12 ppm) due to relatively stable atmospheric conditions and lower rainfall, which limit pollutant dispersion. In contrast, summer exhibited the lowest concentration (2.05 ppm) as higher temperatures, stronger winds, and more frequent precipitation enhance atmospheric mixing and pollutant washout. Fall and winter presented intermediate values, reflecting the combined effects of local emissions and seasonal weather patterns.

From a regulatory and health perspective, THC concentrations are not regulated as a criteria air pollutant in Taiwan, and no specific ambient air quality or health-based standard has been established for THC. Hydrocarbon-related compounds are instead addressed within the broader VOC framework through source-specific emission controls (Ministry of Environment, 2023). Because THC represents a mixture of different hydrocarbon compounds with varying toxicological properties, a single health-based threshold cannot be directly applied. Therefore, the THC concentrations estimated in this study are interpreted as an indicator of atmospheric hydrocarbon levels rather than against a specific regulatory or health-based threshold.

5. Conclusions

In conclusion, the LightGBM spatiotemporal model demonstrated strong predictive capability in estimating THC concentrations, with R^2 values of 0.84 for training, 0.82 for validation, and 0.77 for testing with unseen data, indicating reliable and consistent performance across datasets. The average THC concentration in the study area was 2.07 ppm. SHAP analysis identified NO_2 (ppm), NOx (ppm), and population density as the most influential variables driving THC levels, alongside land use factors such as plantation forests and paddy fields, and meteorological variables like wind speed. These findings highlight the combined effects of air pollution sources, human activity, and environmental conditions in shaping THC distribution. Overall, the model provides a robust tool for accurately capturing the spatial and temporal variation of THC, which is essential for effective air quality monitoring and management strategies.

References

Ambrose, I.S., Offiong, N.-A.O., Ayua, T.J., Ebong, G.A., Ambrose, E.I., Silas, I.I., and Saturday, U.U., 2026. Seasonal dynamics of atmospheric heavy metals and total hydrocarbon content at the Edonwhii Beach, Nigeria, Discover Environment, 4: 200, DOI:10.1007/s44274-026-00712-5.

Bahmani, M.J., 2020. Understanding LightGBM parameters (and how to tune them), Available at: http://dev.to/kamil_k7k/understanding-lightgbm-parameters-and-how-to-tune-them-14n0, Assessed March 11, 2026.

Bishop, G.A., and Stedman, D.H., 2008. A decade of on-road emissions measurements, Environmental Science & Technology, 42(5): 1651–1656, DOI:10.1021/es702413b.

Bose, A., and Chowdhury, I.R., 2023. Investigating the association between air pollutants' concentration and meteorological parameters in a rapidly growing urban center of West Bengal, India: A statistical modeling-based approach, Modeling Earth Systems and Environment, 9(2): 2877–2892, DOI:10.1007/s40808-022-01670-6.

Cheng, W.-L., 2002. Ozone distribution in coastal central Taiwan under sea-breeze conditions, Atmospheric Environment, 36(21): 3445–3459, DOI:10.1016/S1352-2310(02)00307-2.

Coffin, A.W., Ouren, D.S., Bettez, N.D., Borda-de-Água, L., Daniels, A.E., Grilo, C., Jaeger, J.A.G., Navarro, L.M., Preisler, H.K., and Rauschert, E.S.J., 2021. The ecology of rural roads: Effects, management, and research (Issues in Ecology Report No. 23), Ecological Society of America (ESA).

Department of Monitoring and Information, Ministry of Environment, Taiwan, 2024. Air quality monitoring daily average values (air pollutants), Ministry of Environment Environmental Information Open Platform, Available at: https://data.moenv.gov.tw/en/dataset/detail/AQX_P_19, Assessed March 17, 2026.

Doris, M., Daley, C., Zalzal, J., Chesnaux, R., Minet, L., Kang, M., Caron-Beaudoin, É., MacLean, H. L., and Hatzopoulou, M., 2024. Modelling spatial & temporal variability of air pollution in an area

of unconventional natural gas operations, Environmental Pollution, 348: 123773, DOI:10.1016/j.envpol.2024.123773.

Environment Canada, 2003. Future Canadian measures for reducing emissions of volatile organic compounds (Report No. En14-153/2003E), Available at: https://publications.gc.ca/collections/collection_2014/ec/En14-153-2003-eng.pdf, Assessed March 17, 2026.

Geddes, J.A., Wang, B., and Li, D., 2021. Ozone and nitrogen dioxide pollution in a coastal urban environment: The role of sea breezes, and implications of their representation for remote sensing of local air quality, Journal of Geophysical Research: Atmospheres, 126(18): e2021JD035314, DOI:10.1029/2021JD035314.

GeeksforGeeks, 2025. Cross-validation and hyperparameter tuning of LightGBM Model, Available at: https://www.geeksforgeeks.org/machine-learning/cross-validation-and-hyperparameter-tuning-of-lightgbm-model/, Assessed March 11, 2026.

Guaman, M., Roberts-Semple, D., Aime, C., Shin, J., and Akinrem, A., 2022. Traffic density and air pollution: Spatial and seasonal variations of nitrogen dioxide and ozone in Jamaica, New York, Atmosphere, 13(12): 2042, DOI:10.3390/atmos13122042.

Guo, J., Yun, S., Meng, Y., He, N., Ye, D., Zhao, Z., Jia, L., and Yang, L., 2023a. Prediction of heating and cooling loads based on light gradient boosting machine algorithms, Building and Environment, 236: 110252, DOI:10.1016/j.buildenv.2023.110252.

Guo, X., Gui, X., Xiong, H., Hu, X., Li, Y., Cui, H., Qiu, Y., and Ma, C., 2023b. Critical role of climate factors for groundwater potential mapping in arid regions: Insights from random forest, XGBoost, and LightGBM algorithms, Journal of Hydrology, 621: 129599, DOI: 10.1016/j.jhydrol.2023.12959.

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.

Kansal, A., 2009. Sources and reactivity of NMHCs and VOCs in the atmosphere: A review, Journal of Hazardous Materials, 166(1): 17–26, DOI:10.1016/j.jhazmat.2008.11.048.

Li, S., Dong, X., Ma, D., Dang, B., Zang, H., and Gong, Y., 2024. Utilizing the LightGBM algorithm for operator user credit assessment research, arXiv preprint arXiv:2403.14483, DOI:10.48550/arXiv.2403.14483.

Li, Z., Ho, K.-F., Chuang, H.-C., and Yim, S.H.L., 2021. Development and intercity transferability of land-use regression models for predicting ambient PM_{10}, $PM_{2.5}$, NO_2 and O_3 concentrations in northern Taiwan, Atmospheric Chemistry and Physics, 21(6): 5063–5078, DOI:10.5194/acp-21-5063-2021.

Lu, B., Zhang, S., Liu, C., Ganbat, G., Herrmann, H., Li, X., and Zhao, Y., 2025. Machine learning-enabled estimation and high-resolution forecasting of atmospheric VOCs, Atmospheric Environment, 358: 121364, DOI:10.1016/j.atmosenv.2025.121364.

Ministry of Environment, 2023. Volatile Organic Compounds Air Pollution Control and Emission Standards, Available at: https://oaout.moenv.gov.tw/laW/LawContent.aspx?id=FL015377, Assessed August 18, 2026.

Risk Society and Policy Research Center, 2026. Air pollution governance and its distance with people, National Taiwan University, Available at: https://rsprc.ntu.edu.tw/web/research/research_in.jsp?lang=en&rp_id=RP1729156478027, Assessed March 17, 2026.

Stewart, G.J., Nelson, B.S., Drysdale, W.S., Acton, W.J.F., Vaughan, A.R., Hopkins, J.R., Dunmore, R.E., Hewitt, C.N., Nemitz, E., Mullinger, N., Langford, B., Shivani, Reyes-Villegas, E., Gadi, R., Rickard, A.R., Lee, J.D., and Hamilton, J.F., 2021. Sources of non-methane hydrocarbons in surface air in Delhi, India, Faraday Discussions, 226: 409–431, DOI:10.1039/d0fd00087f.

Su, J.G., Jerrett, M., Beckerman, B., Verma, D., Arain, M.A., Kanaroglou, P., Stieb, D., Finkelstein, M., and Brook, J., 2010. A land use regression model for predicting ambient volatile organic compound concentrations in Toronto, Canada. Atmospheric Environment, 44(29): 3529–3537, DOI:10.1016/j.atmosenv.2010.06.015.

Thongthammachart, T., Araki, S., Shimadera, H., Matsuo, T., and Kondo, A., 2022. Incorporating Light Gradient Boosting Machine to land use regression model for estimating NO_2 and $PM_{2.5}$ levels in Kansai region, Japan, Environmental Modelling & Software, 155: 105447, DOI:10.1016/j.envsoft.2022.105447.

Tırınk, S., 2025. Machine learning-based forecasting of air quality index under long-term environmental patterns: A comparative approach with XGBoost, LightGBM, and SVM, PLOS One, 20(10): e0334252, DOI:10.1371/journal.pone.0334252.

United States Environmental Protection Agency, 2005. Conversion Factors for Hydrocarbon Emission Components (EPA420-R-05-015 / NR-002c), US EPA Report.

Vertes, D., 2020. Beyond grid search: Using Hyperopt, Optuna, and Ray Tune to hypercharge hyperparameter tuning for XGBoost and LightGBM, Available at: https://druce.ai/2020/10/hyperparameter-tuning-with-xgboost-ray-tune-hyperopt-and-optuna, Assessed March 11, 2026.

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 $PM_{2.5}$ spatial-temporal variability, Environmental Pollution, 224: 148–157, DOI:10.1016/j.envpol.2017.01.074.

Zhou, Z.-Y., Zhang, X.-X., Wang, F., Zhu, C.-D., and Pan, J.-J., 2023. Vertical pollution characteristics and source analysis of soil PAHs in different land use types, Environmental Science, 44(3): 1583–1592, DOI:10.13227/j.hjkx.202204071.

台灣北部地區使用 LightGBM 回歸模型進行碳氫化合物估計

帕斯蒂[1] 古曼寧[2] 曾于庭[3] 吳治達[4,5,6,7,8*]

摘要

總碳氫化合物(THC)是重要的大氣污染物，主要來自交通、工業活動及燃料燃燒。本研究採用 LightGBM 模型估計 2015 年 6 月至 2020 年 12 月台灣北部空氣品質區的 THC 濃度，並以 2021 年資料進行外部驗證。模型訓練、測試及外部驗證的 R^2 分別為 0.84、0.82 及 0.77，研究區域平均 THC 濃度為 2.07 ppm。SHAP 分析顯示，NO_2 與 NOx 是主要影響因子，其次為人口密度、林地、水田及風速。空間分析發現，西部 THC 濃度高於東部，可能與都市化及工業活動的空間差異有關。本研究顯示，LightGBM 結合 SHAP 可有效估計 THC 濃度並辨識重要影響因子，為空氣品質評估及環境管理提供參考。

關鍵詞：Light Gradient Boosting Machine Regression(LightGBM)、空間預測模型、總碳氫化合物(THC)

[1] 國立成功大學測量及空間資訊學系 博士生
[2] 國立成功大學測量及空間資訊學系 博士後研究員
[3] 國立成功大學測量及空間資訊學系 研究助理
[4] 國立成功大學測量及空間資訊學系 教授
[5] 國家衛生研究院國家環境醫學研究所 兼任研究員
[6] 國立中興大學永續農業創新發展中心 兼任教授
[7] 高雄醫學大學精準環境醫學研究中心 兼任研究員
[8] 長庚科技大學慢性疾病暨健康促進研究中心 合聘研究員
[*] 通訊作者, E-mail: chidawu@mail.ncku.edu.tw

收到日期：民國 114 年 07 月 16 日
修改日期：民國 114 年 09 月 17 日
接受日期：民國 115 年 08 月 21 日

航測及遙測學刊　第三十一卷　第 3 期　第 173- 191 頁　民國 115 年 9 月
Journal of Photogrammetry and Remote Sensing
Volume 31, No.3, 2026, pp. 173 – 191
DOI：10.6574/JPRS.202609_31(3).0003

整合深度學習與資料整合技術於地表變遷偵測

陳聿新 [1]　朱宏杰 [2*]

摘要

本研究探討深度學習模型於地表變遷增益損失偵測之應用，採用雙時期影像(bi-temporal image)為輸入，透過端對端架構直接分析變遷前後之差異。研究主要以 Densely Attentive Refinement Network (DARNet)模型為基礎，發展多類別辨識模型架構，並引入統計匹配與超解析技術以提升不同來源影像(Landsat 與 Sentinel-2)間的一致性。以海岸線變遷為案例，評估比較多種深度學習變遷偵測模型於多類別地景變化資料集的偵測表現，分析不同模型架構於變遷類別識別能力、邊界判釋精度與整體分類效果的差異，作為後續模型選用與改良之依據。研究結果顯示多分類 DARNet 結合資料整合技術於細尺度地物變遷上具優勢，考量超解析處理與統計匹配能提升模型精度與泛化性，有助於變遷偵測之實務效能。

關鍵詞：衛星影像、深度學習模型、增益損失、變遷偵測

1. 緒論

1.1 研究動機

地表變遷監測長期以來是遙測領域的重要研究課題(Zhu *et al.*, 2022)，尤其面對氣候變遷、都市擴張、自然災害與人為干擾等因素所導致的地景快速變化，如何透過有效方法掌握環境變遷情形，進而提供決策支持，是學術與實務界共同關注的議題。衛星遙測技術因具備大範圍、定期更新、多光譜與高時間解析度等特性，使其成為觀測地表環境的關鍵工具。特別是雙時態影像(bi-temporal imagery)於變遷偵測中的應用(Zhang *et al.*, 2020)，可有效捕捉某區域於不同時間點的變異性，對於短期災後分析與長期環境監測皆具高度應用價值。然而，在實際應用中，雙時態影像常會遇到第一時刻影像缺失或品質不佳的問題，原因可能來自雲霧遮蔽、資料儲存不全、衛星重訪頻率低等限制。此外，隨著多源衛星影像的取得變得更為便利，不同解析度影像的整合與應用也成為研究焦點。例如，Landsat 與 Sentinel-2 影像在解析度、光譜涵蓋與重訪頻率上各具優勢，若能有效結合並克服解析度不一致問題，將可進一步提升變遷偵測的時效性與空間精度。因此，本研究針對雙時態變遷偵測中面臨的兩大實務挑戰：第一為由二分類延伸至多分類的類別推估，第二為多解析度影像間的融合應用，分別提出相應方法並驗證其可行性與效益，以強化深度學習模型於地景變遷監測上的應用潛力。

1.2 研究目的

本研究旨在探討深度學習技術於多類別地景變遷偵測任務中的應用潛力，特別針對實際應用中經常遭遇之資料取得不完整、影像解析度不一致等問題，提出具有可行性之模型設計與資料處理方法。透過多組實驗設計與比較分析，期望能建立一套兼具準確性與實務可行性的深度學習變遷偵測流程，提升地表變遷監測的效率與精度。具體而言，本研究目標可分為以下四項：第一，評估並比較多種深度學習變遷偵測模型於多類別地景變化資料集上的預測表現，選定最適合於本研究應用場景之模型架構。透過引進不同模型架構，分析其於變遷類別識別能力、邊界判釋精度與整體分類效果的差異，

[1] 國立成功大學測量及空間資訊學系 碩士生
[2] 國立成功大學測量及空間資訊學系 教授
[*] 通訊作者, E-mail: honejaychu@geomatics.ncku.edu.tw

收到日期：民國 115 年 01 月 07 日
修改日期：民國 115 年 02 月 09 日
接受日期：民國 115 年 07 月 01 日

作為後續模型選用與改良的依據；第二，將以雙時態影像為主要輸入來源之變遷偵測模型應用於海岸線中的多類別變遷辨識與分類任務，建立能同時偵測無變化與多類別變遷(如沙洲堆積與流失等)之模型訓練架構，以驗證深度學習技術於多樣性變遷任務中的可行性與適用性；第三，針對多解析度影像的整合與應用進行深入探討，分別使用 Landsat 8 與 Sentinel-2 衛星影像作為雙時態輸入來源，比較不同解析度影像對模型預測結果之影響。此外，應用深度學習超解析模型提升 Landsat 8 影像之空間解析度，並透過統計匹配方法調整光譜分佈(Heo & FitzHugh, 2000)，減少不同資料源之光譜差異對模型判釋造成的影響，提升資料一致性與分類精度；第四，提供具備實務操作參考價值的變遷偵測資料處理流程與模型輸入設計方案，並驗證其於實際案例場域(如海岸線變遷)中的應用成效，以利未來地景變遷監測作業之推廣與應用。

本研究透過比較多種變遷偵測深度學習模型，選出最適合的模型架構。並在多類別與多解析度影像等實務情況下，對模型進行相對應的調整並分析模型的表現。

(1) 多類別推估：由原先常見二分類(目標變化與不變化)延伸至多分類，將目標與非目標的無變化與增加或損失等變遷過程更細緻的呈現出來。

(2) 結合多解析度影像應用：考量 Landsat 8 與 Sentinel-2 衛星影像在光譜涵蓋、空間解析度與資料取得頻率等層面的差異，設計跨解析度資料整合實驗，探討不同來源影像作為模型輸入時的預測差異與潛在誤差來源。並進一步應用深度學習超解析技術提升 Landsat 8 影像解析度，並輔以統計匹配方法修正其光譜分布，使其更趨近於 Sentinel-2 影像的特性，以強化資料一致性與預測精度，促進多源資料的互補整合應用。

2. 文獻回顧

2.1 變遷偵測技術發展概述

遙測變遷偵測技術自 20 世紀中葉以來持續演進，從早期的影像目視解譯逐步轉向自動化的數位分類與演算法判釋，逐漸發展出各種適應不同應用需求與資料特性的方法(Tewkesbury *et al.*, 2015)。變遷偵測的核心目標在於透過分析多時期影像資料，辨識地表土地覆蓋類別於不同時間點的差異，藉此推估土地使用變遷、自然災害影響、生態環境變化等動態現象。為改善傳統像素為基礎方法(pixel-based approach)易受影像雜訊干擾的問題，Blaschke (2010)推廣物件導向影像分析(Object-Based Image Analysis, OBIA)，透過分割操作將影像切分為具有光譜與空間一致性的區塊，並結合紋理、形狀等空間特徵進行分類與變遷判釋。此類方法在高解析度影像中具備更佳的解析能力，但亦受限於分割參數設定敏感度與分類特徵挑選的主觀性。在處理時間序列資料方面，Verbesselt *et al.* (2010)提出 Breaks For Additive Season and Trend (BFAST)方法，能夠在影像時間序列中偵測趨勢變化與突變事件，進而廣泛應用於森林監測與長期變遷偵測上。隨著遙測影像資料的數量與光譜空間維度持續增加，傳統方法逐漸難以因應龐大且異質的資料結構，在影像分類與變遷判釋上表現出處理效率與精度的侷限。因此，深度學習逐漸成為變遷偵測技術發展的重要方向，不僅具備端對端訓練能力，亦能自動學習具有高度語意與空間脈絡的影像特徵，提升模型對地景變遷的辨識靈敏度與泛化能力。

2.2 長期變遷偵測

衛星影像因具備大範圍、週期性與多光譜觀測的特性，長期以來被廣泛應用於地表環境監測與地貌變遷分析，成為遙測研究與實務應用中最重要的資料來源之一。藉由多時期影像的整合與比較，可有效掌握土地利用變化、水體分布動態、都市擴張進程以及生態環境變異，進而作為環境治理、風險評估與資源管理的重要依據。Li & Gong(2016)利用 Landsat 影像進行西佛羅里達州泥濘海岸的變遷動態監測，採用穩健指標來表示動態海岸線的年平均位置，透過像素的每年淹水面積來代替傳統的線性海岸線表示方法，並利用所提出的模型估算每年海岸像素的淹水面積，研究進一步使用統計模型對時

間序列淹水面積的長期趨勢進行分析，其結果證明了使用時間序列 Landsat 資料進行連續海岸線動態監測的可行性。另一種應用方式為 Aman & Chu (2023)利用 Landsat 影像與多尺度時序分析，提出一個結合像素級水頻率與時間趨勢的框架來監測河川地貌的長期變化，並應用於湄公河與恆河流域之案例研究中，結果顯示該框架能有效捕捉河川範圍的細微變化，並具備高度的時空精度。Bishop-Taylor *et al.* (2021)利用長期 Landsat 影像，並結合次像素水線提取技術與基於像素的潮汐建模方法，繪製澳洲海岸線的長期變遷狀況，結果顯示該方法能將長期的海岸變化率在大陸尺度上得到準確的量化。由上述研究得知 Landsat 影像適合觀察長期的變遷動態觀測，並提供本研究有關於 Landsat 影像在變遷偵測上的應用。

2.3 深度學習模型於變遷偵測

近年來，隨著遙測影像解析度提升與計算資源進步，深度學習技術逐漸成為地表變遷偵測研究中的主流方法。相較於傳統以指標差異或影像分類為基礎的變遷分析方法，深度學習模型具有更強大的特徵擷取與空間語意理解能力(LeCun *et al.*, 2015)，能自動從大規模資料中學習影像之時序與語意變化，特別適合處理高解析度與多來源的雙時態影像資料。

目前多數變遷偵測模型採用雙分支結構作為基礎設計，分別處理變遷前(T1)與變遷後(T2)的影像，透過共享或非共享的權重提取特徵，再進行差異分析與變遷區域判別。以本研究改良的 Densely Attentive Refinement Network (DARNet)為例(Li *et al.*, 2022)，其架構奠基於 UNet 框架，並結合多尺度殘差模組與注意力機制，有效提升模型對於邊界與小面積變遷區域的解析度，適用於如沙洲演替、崩塌生成等細部變化的辨識任務。在其他典型模型中，不同設計重點也展現出各自優勢，例如 Deeply Supervised Image Fusion Network (DSIFN) (Zhang *et al.*, 2020)結合特徵差異建模與注意力強化策略，特別注重變遷區域內外特徵差異的對比性學習；而其引入多層監督機制，能使誤差訊號有效回傳至網路深層，進一步強化模型對多尺度特徵的掌握。類似的概念亦可見於 Spatial Relationship Concerned Network (SRCNet) (Chen *et al.*, 2024)，此模型強調影像中空間關係的連續性與一致性，並設計互動模組加強時序間對應位置的感知能力，提升模型於空間結構變化之判別準確性，特別適用於建物、道路等具明顯空間形態的變遷分析。另有學者針對災害前後環境變遷，提出將時間序列建模與空間語意融合納入變遷偵測流程的架構，例如 MDPrePost-Net (Jamaluddin *et al.*, 2021)融合了時間特徵(透過 ConvLSTM)與空間特徵(透過多尺度卷積模組)的聯合學習，能有效捕捉短期劇烈地景變動，並產出多分類變遷圖。該方法特別強調模型對於局部災損區域的敏感性，並應用於紅樹林退化或颱風影響範圍之描繪，顯示深度網路在處理突發性環境變遷資料上的潛力。

總體而言，深度學習於變遷偵測中的應用已從早期的分類輔助工具，發展兼具時間、空間與語意資訊解析能力的綜合性架構。根據研究目標與資料特性之不同，模型可彈性整合多種模組與監督策略，發展出具針對性與泛化性的解決方案，未來亦可持續拓展至多時期、多來源與多尺度的地景變遷監測任務中。

3. 研究方法與資料

圖 1 為本研究流程圖。本研究整體流程可分為五階段，依序為資料建置、資料前處理、模型訓練、模型評估與結果分析。資料建置目的是彙整與建立適用於海岸線的多類別變遷偵測之衛星影像資料集，包括 Landsat 與 Sentinel-2 兩種不同影像來源的資料集。資料前處理為依據模型需求進行裁切(影像大小：256×256 pixels)，針對解析度不同的影像進行重採樣處理。此外，針對 Landsat 8 影像應用超解析模型提升其解析度，並透過統計匹配修正光譜分布，以縮小與 Sentinel-2 影像之間的光譜差異，以確保多源影像的融合預測比較。模型訓練為在先讓多組不同的深度學習架構進行訓練與比較後，以 DARNet 為本研究改良重點，進行多類別推估以及

多解析度影像等實驗。為客觀評估不同模型與不同輸入情境下之分類效能，模型評估採用整體準確率(Overall Accuracy, OA)、F1 Score、交集比(Intersection over Union, IoU)與 Kappa 係數等指標評估模型效能，其中 F1 Score 與 IoU 特別針對變遷類別(如沙洲堆積與流失)進行計算，以表示模型於微小變化區域之辨識能力。結果分析為透過視覺化輸出與量化指標綜合評估不同實驗組別，檢驗多類別推估與多解析度融合方法在實務應用中的優劣與潛力。

3.1 資料集與前處理

本研究為探討地表海岸線於多時間尺度下之變遷狀況，建立一套涵蓋多時期與多解析度的海岸線變遷資料集，作為深度學習變遷偵測模型之訓練與驗證依據。此資料集內含兩種不同來源之衛星影像，分別為美國之 Landsat 衛星系列與歐洲哥白尼計畫之 Sentinel-2 衛星，兩組不同資料橫跨時間範圍分別為 1990 年至 2024 年與 2019 年至 2024 年，區域為臺灣沿海地區之最大沙洲外傘頂洲、曾文溪口與珠江三角洲之羅鬥沙島，具備高度代表性與分析潛力。

3.1.1Landsat 影像資料集與前處理

本研究所採用之影像資料來源為 Landsat 5 及 Landsat 8 第二代產品(Collection 2)中品質較高的 Tier 1 類別，其屬於 Level 2 處理層級，已完成大氣與幾何校正，能提供地表反射率(Surface Reflectance, SR)及地表溫度(Surface Temperature, ST)等重要參數。這些影像具有高度空間定位精度，並適合用於多時期影像的比較與分析。本研究透過 Google Earth Engine(GEE)平台將從 1990 年至 2024 年的 Landsat 5 及 Landsat 8 年合成影像(陳聿新等，2025a)與常態化差異水體指標(Normalized Difference Water Index, NDWI)(McFeeters, 1996)影像下載下來。NDWI 指標的計算方式如下：

$$NDWI = \frac{Green-NIR}{Green+NIR} \quad (1)$$

其中 Green 為綠光、NIR 為近紅外光。Landsat 5 及 Landsat 8 年合成影像皆包含藍光、綠光、紅光、近紅外光、短波紅外光 1、短波紅外光 2 等 6 個波段，而 NDWI 影像則使用 Otsu 法(Liao *et al.*, 2001)轉換成二值化影像，Otsu 法是常見的影像分割技術，它能使組間變異最大(組內變異最小)，來自動選擇一閾值(陳聿新等，2025b)，接著再將不同年份的二值化影像相互進行判斷產製出水體不變、非水體不變、水體增加與水體減少共四分類的標籤。

3.1.2 Sentinel-2 影像資料集與前處理

本研究所採用之 Sentinel-2 衛星影像資料，經資料處理單位標準化後，屬於已完成大氣校正與地形校正之 Level 2A 層級產品，具備地表反射率資訊、高時間解析度與空間解析度。影像經過去雲與遮蔽處理，適合用於不同時期影像間的疊合分析與變遷判釋。本研究透過 GEE 平台將從 2019 年至 2024 年的 Sentinel-2 年合成影像與 NDWI 影像下載下來。Sentinel-2 年合成影像同樣取對應 Landsat 資料集的 6 個波段，包含藍光、綠光、紅光、近紅外光、短波紅外光 1、短波紅外光 2 等波段且各波段解析度統一輸出為 10 m，標籤製作方式與前述 Landsat 影像資料集相同，最後會產製出水體不變、非水體不變、水體增加與水體減少共四分類的標籤。表 1 為 Landsat 影像資料集與 Sentinel-2 影像資料集的數據比較。

表 1 Landsat 影像資料集與 Sentinel-2 影像資料集比較

	Landsat 影像資料集	Sentinel-2 影像資料集
波段	6 波段	12 波段，本研究取對應 Landsat 之 6 波段
影像類型	年合成影像	年合成影像
空間解析度	30 m	10 m
影像年份	1990-2024	2019-2024
影像張數	1740 張	360 張
影像大小	256×256 pixels	256×256 pixels

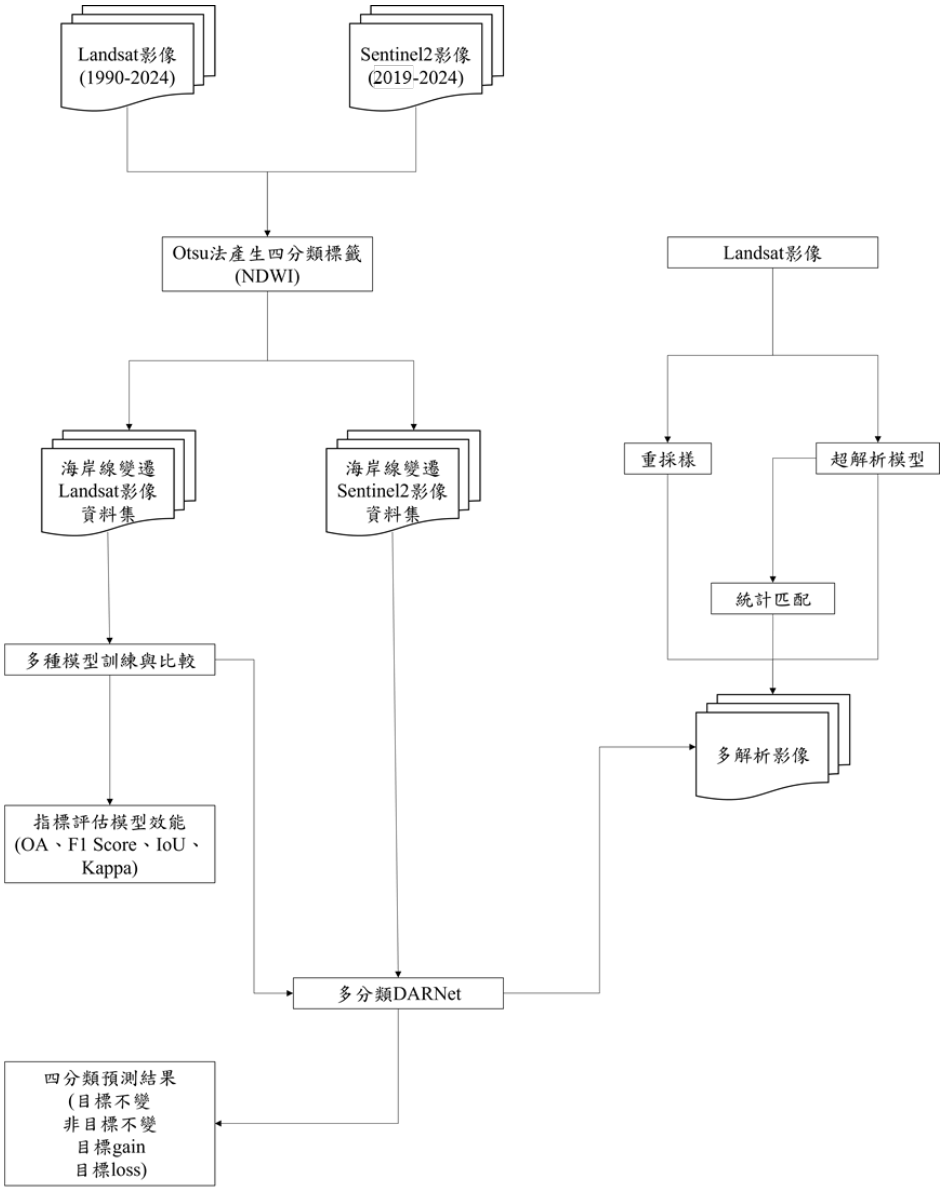


圖 1 研究流程圖

3.2 變遷偵測模型

深度學習模型近年廣泛應用於遙測影像的變遷偵測領域，其具備自動學習高階特徵的能力，能有效克服傳統方法需人工設計特徵所帶來的侷限。在本研究中，深度學習模型可透過大量已標記的時序影像資料進行訓練，自動提取空間與光譜資訊中潛藏的變化模式。與傳統基於物件或像素的分類方法相比，深度模型更能掌握複雜的非線性關係與多尺度特徵結構。常見的應用架構如卷積神經網路或編碼器－解碼器(Encoder-Decoder)結構，可用於像素層級的語意分割與變遷區域的提取。透過模型訓練與驗證，能進一步評估其在不同地區或資料條件下的泛化能力與變遷偵測準確度，並與其他模型進行比較分析，找出最適合的遙測變遷偵測策略。

3.2.1DARNet 與 Multiclass DARNet

在本研究中，改良源自於 DARNet (Li *et al.*, 2022)作為主要的變遷偵測模型，該模型屬於典型的編碼器－解碼器架構，並於原始設計中整合多尺度特徵抽取、時空注意力機制與逐層細化預測模組，展現了在遙測影像變遷偵測任務中的高度辨識能力。DARNet 主要由三大模組構成：特徵編碼器、特徵解碼器與遞迴式細化模組 (Recurrent Refinement Module, RRM)，並輔以密集跳接模組 (Dense Skip Connection Module, DSCM)與混合注意

力模組(Hybrid Attention Module, HAM)，以提升模型對時空變遷的表徵能力與分類精度。

在編碼器部分，DARNet 採用對稱式的孿生卷積神經網路(Siamese CNN)，用以從雙時態影像中分別抽取多層次特徵。每一個子網路由五個卷積區塊組成，分別對應不同解析尺度的特徵輸出，其通道數由淺至深分別為 16、32、64、128 與 256。每一區塊皆由兩組 3×3 卷積、Batch Normalization 與 ReLU 激活層組成，並於區塊之間插入最大池化層以進行特徵降維與提取。透過孿生架構，模型可確保兩個時間點的特徵在相同空間尺度下進行有效對齊與比對。

為了能充分整合來自不同尺度與時間的影像特徵，DARNet 導入 DSCM 模組，該模組不僅保留 UNet 架構中原有的高低層特徵串接，並進一步將來自不同層級的特徵統一重採樣至相同解析度後整合，使模型得以兼具語意抽象與空間細節，進而提升變遷區域的定位與邊界判別能力。緊接著，混合注意力模組(HAM)將多尺度特徵輸入進行空間、時間、通道三維度的非區域性注意力建模。該模組由空間時間注意力模組(Efficient Spatial–Temporal Attention Module, ESTAM)與通道注意力模組(Channel Attention Module, CAM)組成，前者可捕捉像素間的時空依存關係，而後者則強化與變遷相關的通道特徵，並抑制背景與非關聯特徵的干擾。在解碼器階段，模型會逐層融合來自編碼器與 HAM 輸出的多層次特徵，並搭配遞迴式細化模組 RRM 將初步的粗略預測不斷修正為空間解析度更高的變遷圖。每一層細化單元皆接收上一層預測圖與當前特徵圖，進行融合與重新判別，最終輸出一系列從低解析至高解析的變遷結果圖。為確保每一層輸出皆具監督訊號，DARNet 採用深度監督機制(Deep Supervision)，於每一階段輸出皆計算損失函數，並依不同階段加權後共同參與模型訓練。

本研究針對原始 DARNet 進行若干調整，以使其更適用於本研究所需之「四分類變遷偵測任務」。不同於原始設計中僅區分「變遷」與「未變遷」的二元分類模式，本文將變遷情形細分為四類，藉由標註類別間的轉換關係，進一步提升變遷偵測的解析能力。四類標籤分別為類別 A 不變，表示地物類別維持不變，例如非目標地物持續存在；類別 B 不變，代表目標類別(如沙洲)在兩時態間皆存在，無發生變化；類別 A 變為類別 B，對應於目標地物的新出現，例如沙洲新堆積；以及類別 B 變為類別 A，則表示目標地物遭移除或消退，例如沙洲流失。此種四分類設定較傳統的二分類(變與不變)更具描述力與應用彈性，可依不同研究對象靈活調整標註邏輯，涵蓋多樣地景變遷的樣態。在模型訓練階段，四分類標籤不僅提供更豐富的語意資訊，有助於模型學習類別轉換的細節特徵，亦可針對特定類型變遷進行針對性評估與解析，有效提升變遷偵測任務的實務應用價值。

為因應多類別分類需求，本研究將原本設計為二分類損失的目標函數，改採多分類下的混合損失函數。具體而言，本文整合了交叉熵損失(Cross Entropy Loss)與多類別 Dice 損失(Multiclass Dice Loss)，構成混合型目標函數，以平衡類別不均與提升邊界精度為目標。交叉熵損失可有效處理多類別分類任務中各類別之預測偏差，Dice 損失則透過測量預測圖與標籤圖之重疊程度，有效緩解資料集中各類別樣本數不均所造成的偏差，尤其能強化邊界處的預測能力。綜合兩者優勢可大幅提升整體分類準確率與空間判別性能。改良後 DARNet 多分類模型架構如圖 2 所示。

3.2.2其他模型

為驗證本研究所提出之多分類變遷偵測模型的效能，本文選用三種具有代表性的現有架構作為對照模型，分別為 DSIFN、SRCNet 與 MDPrePost-Net。

DSIFN (Zhang *et al.*, 2020)為一種基於雙時態影像融合的變遷偵測架構，透過共享參數的深層特徵擷取網路對變遷前後影像進行多尺度特徵提取，並利用差異判別模組融合兩時態的語意特徵。為強化模型對變遷區域的辨識能力，DSIFN 引入通道與空間注意力機制，進一步聚焦於重要變遷資訊；同時，搭配多層監督策略，有助於改善梯度消失問題並提升訓練穩定性。

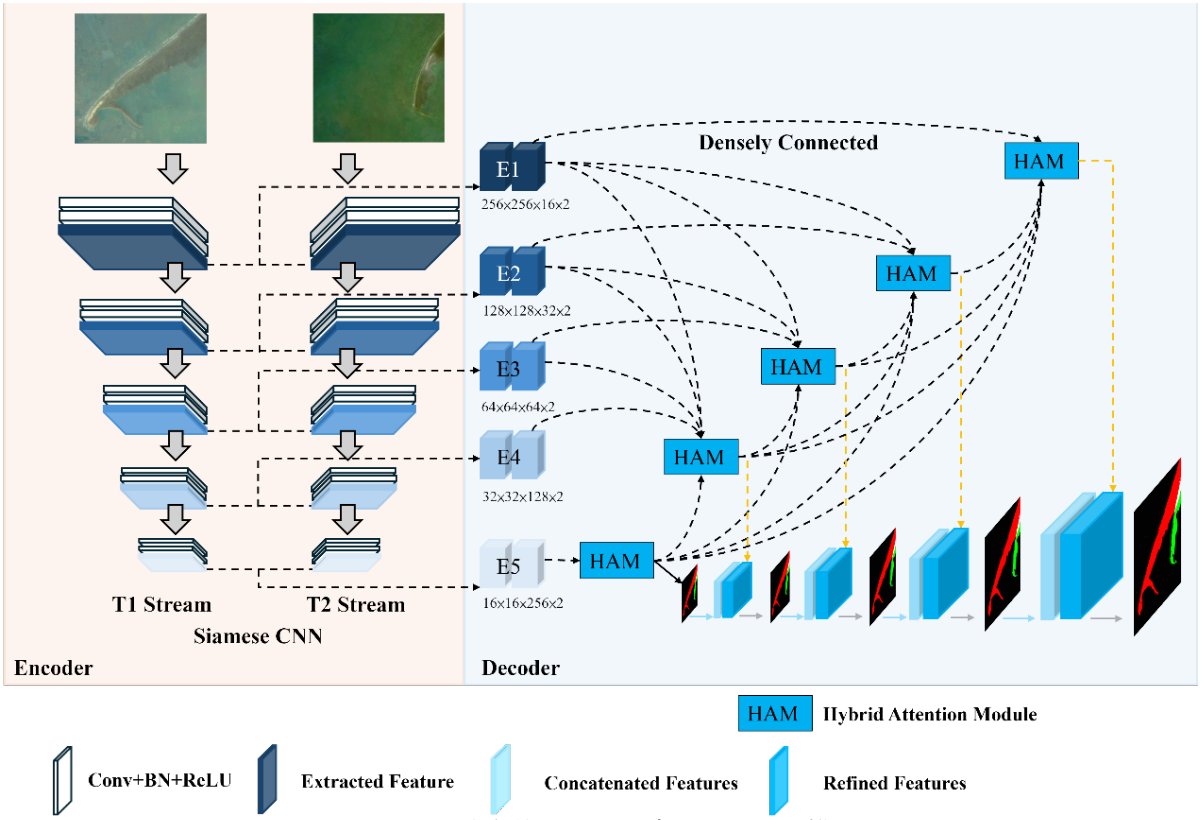


圖 2 改良後 DARNet 多分類模型架構圖

SRCNet (Chen *et al.*, 2024)強調雙時態影像間的空間結構關係，設計了感知與互動模組(Perception and Interaction Module, PIM)以及塊模式聯合特徵融合模組(Patch–mode Joint Feature Fusion Module, PJFFM)，前者可強化影像間的對應關係感知能力，後者則透過局部區塊方式進行細緻特徵融合，避免傳統融合方式可能造成的資訊損失。整體架構強調在特徵抽取與融合階段持續保留與強化影像間的空間一致性，有助於提升模型對變遷細節的識別能力。

MDPrePost-Net(Jamaluddin *et al.*, 2021)專為災前災後紅樹林環境監測任務設計，結合時間序列特徵與空間光譜資訊進行分類。其核心由兩部分組成：前後時序特徵擷取模組利用 ConvLSTM 結構捕捉時間差異，輔以 Inception 模組加強災後影像中細節變化的辨識能力；全卷積分類模組則以 UNet 為基礎架構，保留空間解析度並結合語意資訊，最終產出災前災後環境變遷的分類結果。此模型特別適合用於分析颶風等極端事件後的環境變遷情境。綜上所述，三種模型皆具備處理雙時態遙測影像的能力，且各自針對特徵提取、注意力機制、時間序列處理等方面有所著墨，為本研究所提出之模型提供關鍵的比較參考依據，所有模型訓練超參數如表 2 所示。

表 2 模型訓練超參數

Model	Learning Rate	Batch Size	Epoch
DARNet	1e-5	2	50
DSIFN	1e-5	2	50
SRCNet	1e-5	2	100
MDPrePost-Net	1e-4	2	100

3.3 多解析度影像

本研究針對實際應用中常見之衛星影像解析度不一致問題，進行多解析度影像之應用測試與方法設計。由於時序變遷模型多仰賴雙時態影像作為輸入資料，若在某時間點缺乏與訓練資料解析度相符之影像，則需考量是否可透過解析度轉換來替代使用。為探討此一問題對模型預測能力之影響，本研究以已使用 Sentinel-2 衛星影像所訓練之變遷模型為基礎，設計一系列輸入資料為不同解析度 Landsat 8 影像之替代實驗。此實驗架構分別採用重採樣處理與超解析度模型(Super Resolution)兩種方

法來處理解析度落差，並將結果與原始 Sentinel-2 影像進行比較分析，以評估不同解析度影像對預測準確性之影響及可行性。

3.3.1重採樣

為使 Landsat 8 影像能作為 Sentinel-2 的替代輸入，並與既有模型維持相同的輸入維度與空間尺度，本研究在 GEE 平台中對原始 Landsat 8 Level2 產品進行重採樣處理。Landsat 8 原始影像波段解析度為 30 m，僅少數全色或熱紅外波段達 15 或 100 m，為提升空間一致性，本研究選取與 Sentinel-2 相對應之六個多光譜波段(藍光、綠光、紅光、近紅外光、短波紅外光 1、短波紅外光 2)，並透過最近鄰插值法將其統一重採樣至 10 m 解析度。重採樣後之 Landsat 8 影像具有與 Sentinel-2 相同的空間分辨率與波段數量，並可直接作為模型輸入用於推論階段。該處理方式雖無提升實際空間細節，但可確保模型接收影像格式與解析度一致之資料，避免模型因輸入維度差異而無法運作。

3.3.2超解析深度學習模型

除了重採樣處理外，許多研究以深度學習提升衛星影像之解析度(Wang *et al.*, 2022)，本研究亦嘗試以深度學習方式提升 Landsat 8 衛星影像之解析度，藉此接近 Sentinel-2 所提供之高解析度品質。本研究採用一套預訓練之超解析度模型，採用以 UNet 為基礎之全卷積神經網路，並結合 Landsat 8 之多光譜與高解析度全色影像作為輸入，進行影像重建與空間細節增強。

首先，超解析模型將問題形式化為影像退化模型的逆向學習問題，即視 Landsat 8 影像為經由某種退化函數對 Sentinel-2 影像進行空間與光譜轉換後所得之結果。為了反推此退化過程，模型透過學習一組深度神經網路參數，實現從低解析 Landsat 8 影像重建出高解析且光譜一致之 Sentinel-2 對應影像。超解析模型架構如圖 3 所示，本研究採用的預訓練之超解析模型為 Sambandham *et al.* (2024)釋出之深度為五層的模型，在網路架構方面，核心採用經典 UNet 結構，具備編碼器與解碼器對稱配置，並透過跳接連結保留多尺度特徵。輸入資料由經雙三次插值(bicubic interpolation)放大後的 Landsat 8 多光譜影像與其全色影像組成，並經過通道串接後送入編碼器。解碼器端則依照不同設計選擇三種上採樣方式：「Up + Conv」即先上採樣後卷積、「Shuffle」為通道擴張後進行子像素卷積、「Conv Transpose」則為轉置卷積操作。實驗結果顯示「Up + Conv」與「Shuffle」兩種方式能有效提升重建影像品質，而「Conv Transpose」則因產生方格狀偽影而不被採納。

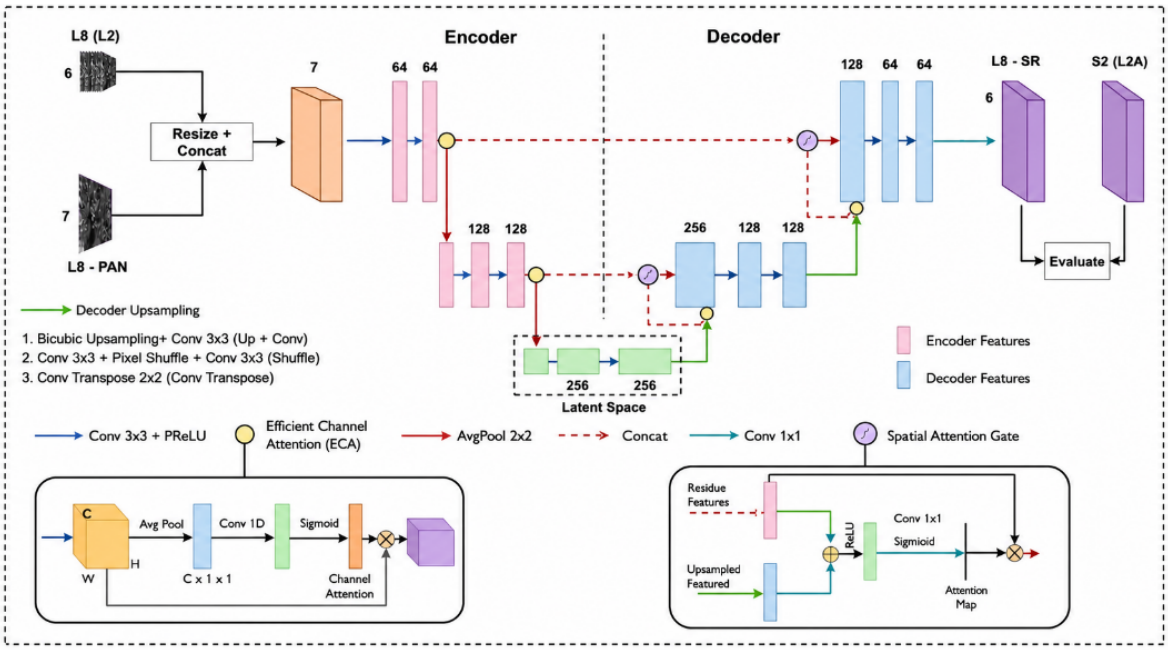


圖 3 超解析模型架構圖 (改繪於 Sambandham *et al.*, 2024)

此模型可自動重建較為細緻之紋理與邊界資訊，並將原始 30 m 解析度之 Landsat 8 影像升級至 10 m 解析度。透過此方式生成之高解析度 Landsat 8 影像，不僅在空間維度上達到與 Sentinel-2 相當，亦具備部分光譜與紋理細節的增強效果，有助於模型辨識地物變遷邊界與微小區塊。然而，須特別注意的是，超解析模型所生成之細節屬於推估結果，並非原始觀測資訊，因此仍需搭配後續分析驗證其影像特性與模型預測效能。

3.3.3統計匹配

在多源影像整合應用中，資料的一致性是深度學習模型能否有效進行特徵學習與分類判釋的關鍵因素之一。由於不同衛星平台具備不同的感測器響應特性、光譜配置與輻射校正方法，即使觀測同一地區、同一時間，所產製之影像在光譜分佈與數值尺度上仍可能出現顯著差異。此種異質性將直接影響模型對跨來源影像的辨識與推論能力，亦可能造成特定地物類別在一個來源影像中表現穩定、但在另一來源中呈現偏移甚至錯誤分類的現象。

為解決上述問題，本研究對超解析處理後之 Landsat 8 影像進行統計匹配處理，透過統計匹配將超解析處理後之 Landsat 8 影像的光譜分布進行調整，使其更趨近於 Sentinel-2 影像的光譜特性，進而提升變遷偵測模型於異質影像資料上的泛化能力與預測穩定性。統計匹配的執行流程如圖 4 所示。首先，對來源與目標影像進行波段標準化處理，將每個波段的像素值轉換至 0 到 1 的範圍，以消除不同影像間原始像素數值尺度的差異。標準化方法為最小最大正規化，即每個波段的像素值以該波段內最小值與最大值進行線性縮放。對 Sentinel-2 影像進行標準化後，計算每個波段的平均值與標準差，作為統計匹配的目標特徵。這些統計值用以引導來源影像的分佈調整，使其能在標準化空間中模擬出與 Sentinel-2 影像一致的分佈特性。本研究採用迭代法，在標準化空間中對每個來源影像波段進行平均值與標準差的調整。具體作法為將每個來源波段視為一個常態分佈，使用 z score 標準化後，再依據目標平均與標準差進行重新縮放。本研究與學者(Heo & FitzHugh, 2000、Du *et al.*, 2002)所提出之相對輻射正規化具有相同目的，旨在降低影像間由感測器與觀測條件造成之輻射差異；然而，本研究以全域統計分布進行匹配(標準化處理是針對所有影像，且來源影像與目標影像涵蓋相同的時/空間範圍)，而非考量擬不變特徵(Pseudo-Invariant Features, PIF)樣本。當所有波段於標準化空間完成統計匹配後，以 GeoTIFF 格式輸出標準化的匹配後的影像。

3.4 驗證方法

為全面評估本研究之深度學習變遷偵測模型於多類別地物變遷任務中的預測效能，特設計一套包含混淆矩陣與多項常用分類評估指標之驗證方法。該方法除考量整體分類表現(如整體準確度、Kappa 係數)，亦針對變遷類別進行強化，特別計算變遷類別(如沙洲流失與堆積等)之 F1 Score 與 IoU，以強化模型對細微變遷區域之解析能力的評估。

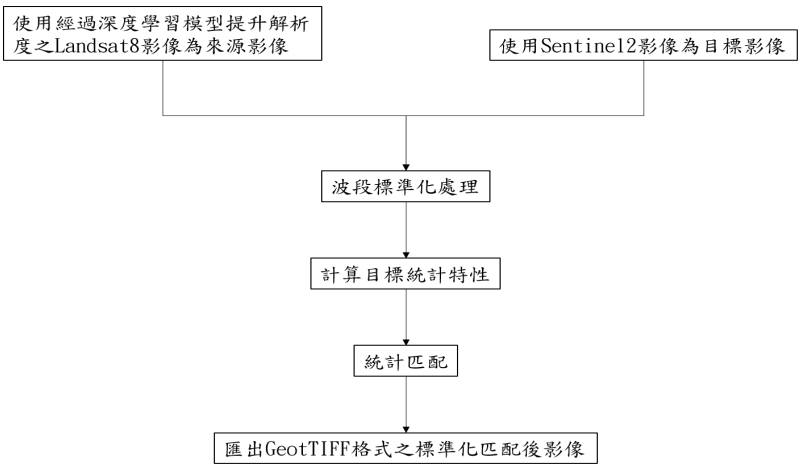


圖 4 統計匹配流程圖

3.4.1混淆矩陣(Confusion Matrix)

混淆矩陣是一種基本且直觀的分類驗證工具，主要用於記錄實際類別與模型預測結果之對應關係，並可延伸出多項性能指標。針對本研究所處理之四分類問題，混淆矩陣結構如表 3 所示。對應類別說明如下，以沙洲流失與堆積舉例，類別 0 為無變化類(第一時刻與第二時刻皆為沙洲的區域)；類別 1 為無變化類(第一時刻與第二時刻皆為水體的區域)；類別 2 為沙洲變為水體的區域，即沙洲流失；類別 3 為水體變為沙洲的區域，即沙洲堆積。

表 3 四分類混淆矩陣示意表

真實\預測	類別 0	類別 1	類別 2	類別 3
類別 0	n_{00}	n_{01}	n_{02}	n_{03}
類別 1	n_{10}	n_{11}	n_{12}	n_{13}
類別 2	n_{20}	n_{21}	n_{22}	n_{23}
類別 3	n_{30}	n_{31}	n_{32}	n_{33}

其中 nij 代表實際為類別 i 而被預測為類別 j 的像素數量。該矩陣可延伸計算出多項整體與類別層級之分類指標。

3.4.2評估指標

為精準反映模型於多類別變遷偵測上的表現，本研究針對四分類情境設計多項評估指標，涵蓋整體準確度(OA)、Kappa 係數、F1 Score 及 IoU 等。其中，F1 Score 與 IoU 為針對變遷類別(類別 2 與類別 3)之區別能力進行加權計算，以更貼近實際應用中變遷識別的需求。整體準確度為所有正確分類像素占總像素數量的比例，計算公式如下：

$$OA = \frac{n_{00}+n_{11}+n_{22}+n_{33}}{N_{total}} \quad (2)$$

其中N_{total}為總像素數量。

Kappa 係數可用來衡量分類結果與隨機分類之差異程度，其計算公式如下：

$$\kappa = \frac{p_o-p_e}{1-p_e} \quad (3)$$

其中p_o表示觀察到的準確率，即 OA；p_e表示期望隨機一致率，計算如下：

$$p_e = \frac{\sum_{i=0}^{3}(\sum_j n_{ij})(\sum_j n_{ij})}{N_{total}^2} \quad (4)$$

F1 Score 為一種綜合考量精確率(Precision)與召回率(Recall)的指標，能有效衡量模型在特定類別上的分類性能，特別是在資料不均衡的情境下依然具有良好代表性。若模型對某類別的精確率或召回率 偏低，F1 Score 會顯著下降，呈現模型在該類別的不足之處。若需計算每一類別的 F1 Score，需先定義精確率和召回率，其公式如下：

$$Precision_i = \frac{n_{ii}}{\sum_j n_{ji}} \quad (5)$$

$$Recall_i = \frac{n_{ii}}{\sum_j n_{ij}} \quad (6)$$

每一類別的 F1 Score 為：

$$F1_i = \frac{2 \cdot Precision_i \cdot Recall_i}{Precision_i + Recall_i} \quad (7)$$

總 F1 Score 可採用平均方式：

$$F1_{mean} = \frac{1}{4}\sum_{i=0}^{3} F1_i \quad (8)$$

IoU 則是衡量預測區與實際區之重疊程度的空間指標，代表模型預測區塊與真實區塊的交集與聯集比例。IoU 值越高，表示模型在空間邊界上的判斷越準確，對於遙測應用中特別重視邊界精準度的任務(如沙洲流失與堆積等)具高度參考價值。

每一類別 IoU 的計算方式為：

$$IoU_i = \frac{n_{ii}}{\sum_j n_{ij}+\sum_j n_{ji}-n_{ii}} \quad (9)$$

總 IoU 可採用平均方式：

$$IoU_{mean} = \frac{1}{4}\sum_{i=0}^{3} IoU_i \quad (10)$$

由於本研究重點在於探討變遷區(即類別 2 與 3)之預測準確性，因此針對這兩類設計，將類別 2 與類別 3 統一視為變遷區，其他類別(0 與 1)視為非變遷區，精確率和召回率定義如下：

$$Precision_i = \frac{\sum_{i=2}^{3} n_{ii}}{\sum_{i=2}^{3}\sum_{j=0}^{3} n_{ij}} \quad (11)$$

$$Recall_i = \frac{\sum_{i=2}^{3} n_{ii}}{\sum_{i=0}^{3}\sum_{j=2}^{3} n_{ij}} \quad (12)$$

針對變遷類別的 F1 Score 為：

$$F1_{change} = \frac{2 \cdot Precision_i \cdot Recall_i}{Precision_i + Recall_i} \quad (13)$$

針對變遷 IoU 的計算方式為：

$$IoU_{change} = \frac{\sum_{i=2}^{3} n_{ii}}{\sum_{i=2}^{3}\sum_{j=0}^{3} n_{ij} + \sum_{i=0}^{3}\sum_{j=2}^{3} n_{ij} - \sum_{i=2}^{3}\sum_{j=2}^{3} n_{ij}} \quad (14)$$

本研究所採評估方式兼顧整體準確度與變遷類別之精細辨識能力，能有效掌握模型在多類別變遷偵測任務下之實用性與敏感度。相較傳統以整體分類為核心之評估架構，本方法針對實際地景變遷監測之應用需求，提出精細化之評估設計，特別強調模型於變遷區域的辨識能力，並可反映不同類別間混淆所帶來的影響，對於日後模型改良與應用策略制定，提供具體且量化之參考依據。

4. 研究成果分析

4.1 模型比較-海岸線變遷分析

為評估所提深度學習方法於實際遙測應用中的表現，本研究以海岸線變遷偵測作為主要應用場景，選取自建之海岸線資料集進行實驗，並分別於四種具代表性的時序影像變遷偵測模型上進行訓練與比較，包括：DARNet、DSIFN、SRCNet 與 MDPrePost-Net。此比較實驗旨在探討各模型於處理四分類變遷任務下的準確性、空間判釋能力與泛化能力，並藉此驗證所提出方法的實用性與穩定性。

為確保實驗結果的公平性與代表性，所有模型皆以相同的資料與分割方式進行訓練與測試。海岸線資料集 (大小：256×256 pixels) 依 8：1：1 的比例劃分為訓練集、驗證集與測試集，並確保各時態、各類別分布盡可能均衡。所有模型皆以雙時態 Landsat 合成影像作為輸入，輸入波段包含藍光、綠光、紅光、近紅外光、短波紅外光 1、短波紅外光 2 等 6 個波段，輸出則為四分類的預測結果圖。使用的標籤採用四分類策略進行地物變遷定義。黑色為在第一時刻與第二時刻皆為水體的區域、白色為第一時刻與第二時刻皆為陸地的區域、紅色為第一時刻為陸地但第二時刻為水體的區域(代表沙洲流失)、綠色為第一時刻為水體但第二時刻為陸地的區域(代表沙洲新堆積)。此四分類設計可更清楚揭示海岸線動態變遷的空間位置與方向性，並優於傳統的二分類方法(僅判定變與不變)，使模型能針對不同型態的變遷進行細部判別與分析。

圖 5 顯示各模型於外傘頂洲測試資料上之預測結果，由 a 至 g 分別為第一時刻 RGB 影像、第二時刻 RGB 影像、海岸線變遷的四分類標籤圖(c)，以及 DSIFN、SRCNet、DARNet、MDPrePost-Net 四個模型之預測結果。由圖中可觀察到，DARNet 在沙洲變遷區域(紅色與綠色區域)的邊界判定最為準確，能清晰地對應至實際沙洲消失與新堆積位置。相較之下，DSIFN 與 SRCNet 在小區域沙洲堆積的偵測上略顯保守，部分新生成的沙洲區塊未能正確預測；而 MDPrePost-Net 表現雖亦穩定，但在邊界處仍有模糊現象。

為量化各模型在海岸線變遷任務中的表現，本研究採用四項常見分割評估指標包含 OA、F1 Score、IoU、Kappa，其中 F1 與 IoU 指標特別針對變遷類別(沙洲流失與沙洲堆積)進行計算，以評估模型對小區塊變化區之敏感度與區分能力。各個海岸線變遷模型的表現如表 4 所示。

表 4 海岸線變遷各模型評估指標比較

Method	OA (%)	F1	IoU	Kappa
DSIFN	99.04	0.9392	0.8863	0.9928
SRCNet	98.67	0.9123	0.8454	0.9900
DARNet	99.55	0.9716	0.9447	0.9966
MDPrePost-Net	99.51	0.9690	0.9402	0.9963

由表 4 得知 DARNet 在所有評估指標上皆為表現最優，尤其於 F1 Score 與 IoU 指標上領先幅度顯著，顯示其於沙洲變遷區的分類能力優於其他模型。而 MDPrePost-Net 亦具高度準確性，其結合預先與後處理結構，對整體分類穩定性有所貢獻，但在邊界處仍略遜於 DARNet。在視覺結果與數值評估雙重佐證下，DARNet 為最適用於本研究資料與任務需求之模型。

本研究亦將訓練完成的 DARNet 模型應用至羅鬥沙地區之影像進行推論測試(圖 6)。羅鬥沙島位於中國廣東省珠江口西岸海域，地處珠海市香洲區與

中山市南部之間，屬於珠江三角洲外緣的沿海沙洲島嶼。由結果可知，模型能成功辨識沙洲於時序影像間的明顯變化，明確標示出堆積與流失的空間區位，顯示其具有良好的泛化能力與實務應用潛力。

在多時期變遷預測方面，亦可將已訓練完成之 DARNet 模型，針對不同時間點的影像進行兩兩配對並推估變遷結果。圖 7 為 DARNet 在多時期的羅門沙地區預測三組前後不同時期(1990~2007、2007~2023、1990~2023)結果，由左至右分別為第一時刻、第二時刻、模型的預測結果。圖 7a 至 c 三組第一時刻分別為 1990 年、2007 年、1990 年，第二時刻 a 至 c 三組分別為 2007 年、2023 年、2023 年。三組前後不同時期(1990~2007、2007~2023、1990~2023)的預測結果顯示，DARNet 在三組不同時期的組合皆能將沙洲的流失與堆積預測出來。

此外，從 a 至 c 三組預測結果亦可觀察到羅門沙地區在過去三十年間沙洲的整體變遷趨勢：原本分布於東北與西南兩端的沙洲逐漸向中央集中，反映出沙洲動態變遷過程的空間特徵。上述結果顯示，DARNet 不僅可應用於單一時間區間之變遷監測，亦具備處理長期變遷分析的潛力，對於地貌變遷追蹤與環境監測應用具有參考價值。

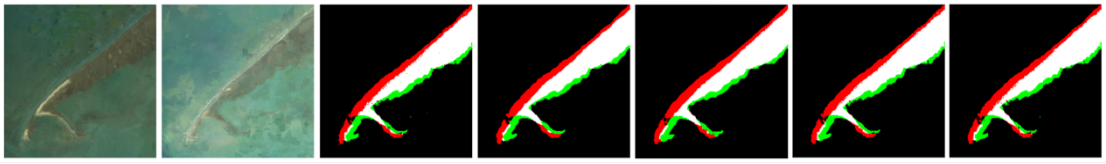


圖 5 外傘頂洲海岸線變遷各模型預測結果比較圖 (a)第一時刻的 RGB 影像、(b)第二時刻的 RGB 影像、(c)海岸線變遷的四分類標籤圖，以及 (d) DSIFN、(e) SRCNet、(f) DARNet、(g) MDPrePost-Net 四個模型之預測結果



圖 6 羅門沙變遷 DARNet 預測圖：黑色為在第一與第二時刻皆為水體的區域、白色為第一與第二時刻皆為陸地的區域、紅色為第一時刻為陸地但第二時刻為水體的區域(代表沙洲流失)、綠色為第一時刻為水體但第二時刻為陸地的區域(代表沙洲新堆積)

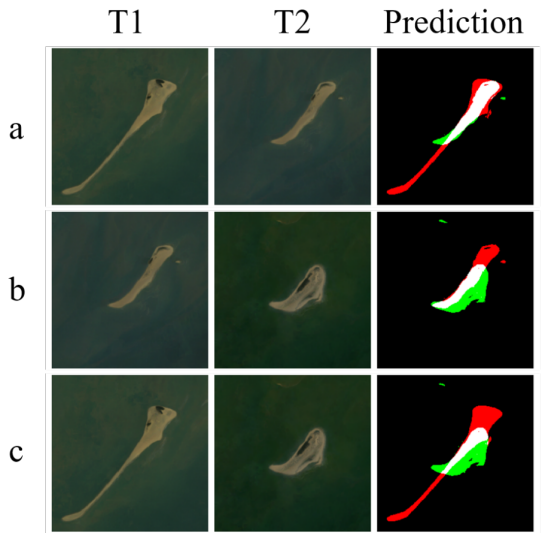


圖 7 羅門沙多時期變遷 DARNet 預測圖 (a)第一時刻為 1990 年、第二時刻為 2007 年的組別，(b)第一時刻為 2007 年、第二時刻為 2023 年的組別，(c)第一時刻為 1990 年、第二時刻為 2023 年的組別

4.2 海岸線變遷分析-Sentinel-2 資料集

在先前的實驗中，本研究以 Landsat 5 及 Landsat 8 所構建之海岸線四分類資料集作為主要研究對象，針對沙洲變遷類別進行多模型比較與效能驗證。雖然 Landsat 系列影像具有良好之時序延續性與光譜資訊，但其空間解析度受限於 30 m，在辨識小尺度沙洲變遷區塊或細緻邊界時，可能存在模糊或不完整的情況。

為進一步驗證所使用之深度學習模型在高解析度資料條件下的表現差異與適應能力，並探討解析度提升對變遷辨識準確性的影響，本研究另外建置一組以 Sentinel-2 衛星影像為基礎的海岸線變遷資料集。Sentinel-2 擁有最高 10 m 的空間解析度，並涵蓋可見光、近紅外光、短波紅外光等多個光譜波段，能提供比 Landsat 更細緻的地物紋理與空間細節，有助於提升模型對地表細微變化的辨識能力。

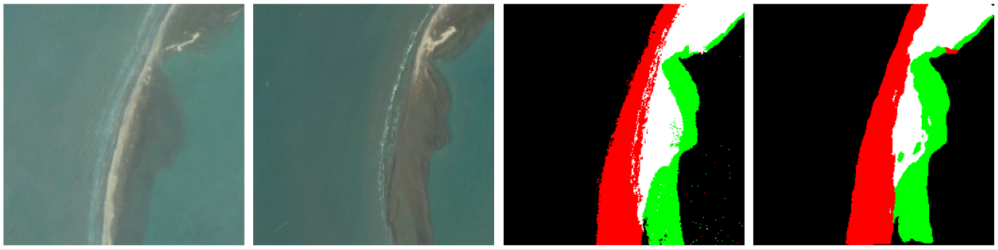
本實驗採用多分類的 DARNet 模型進行訓練與測試。資料集依 8：1：1 的比例劃分為訓練集、驗證集與測試集，模型的輸入一樣包含藍光、綠光、紅光、近紅外光、短波紅外光 1、短波紅外光 2 等 6 個波段，且所有波段在 GEE 平台上進行重採樣處理後解析度統一為 10 m 作為輸入，輸出則為與標籤對應的四分類預測圖。圖 8 顯示 DARNet 經過訓練後在測試資料上的預測結果，由 a 至 d 依序為第一時刻的 RGB 影像、第二時刻的 RGB 影像、海岸線變遷的四分類標籤圖、DARNet 的預測結果。從圖 8 中可觀察到，模型能成功辨識主要沙洲變遷區域，並對邊界區塊進行較為精細的判斷。然而，在特定變遷類別上仍存在預測偏差，顯示即便輸入解析度提升，仍可能受限於類別光譜混淆與樣本分佈不均等問題。

DARNet 模型於測試資料的 OA(%)、F1 Score、IoU、Kappa 等各項分類評估指標分別為 95.56%、0.8929、0.8078、0.9702，從整體指標來看，DARNet 在高解析度影像下依然維持良好的分類表現。特別是 F1 與 IoU 指標，雖較使用 Landsat 資料時略為下降，但仍達可接受水準，顯示模型具備相當的穩定性與泛化能力。Kappa 值亦達 0.97，代表預測結果與標籤間具有高度一致性。

進一步觀察模型於不同類別間之分類誤差情形，可發現在沙洲流失的部分，模型偶有將實際不變的沙洲區錯誤分類為流失區；在沙洲堆積類別中，模型預測偏差更為明顯，部分區域將原本不變的沙洲誤判為新堆積區，或漏判實際發生堆積之區域。此現象顯示，即使使用空間解析度較高的 Sentinel-2 資料，若模型本身對變遷區特徵抽取未具足夠區辨性，仍可能發生誤判或漏判。因此，在處理高解析影像時，模型不僅需具備較深的網路結構與更強的空間學習能力，也需搭配適當的資料增強、類別平衡或語意引導機制，以提升對複雜地貌變遷的判斷精度。儘管如此，從整體結果而言，DARNet 模型於四個變遷類別上皆能產生合理預測，並具備穩定的區域分類能力。其結合殘差連接與遞進式預測策略，有助於強化空間與語意特徵融合，即使僅有雙時態輸入，亦可維持良好之變遷表示效果。

4.3 海岸線變遷分析-多解析度影像

以雙時態影像為輸入的模型，若遇到前一時刻無同一解析度的影像，可否使用其他解析度的影像來替代並預測？本研究以使用 Sentinel-2 海岸線資料集訓練好的變遷模型來測試此一情況，原本海岸線資料集的 Sentinel-2 影像包含藍光、綠光、紅光、近紅外光、短波紅外光 1、短波紅外光 2 等共 6 個波段，本研究在 GEE 平台上將原本 Sentinel-2 影像的 6 個波段，以 Landsat 8 相對應之 6 個波段替換掉，最後在 GEE 平台上統一重採樣輸出成解析度為 10 m 的 Landsat 影像。本研究另外以能提升 Landsat 8 衛星影像解析度之深度學習模型，進一步將 Landsat 8 影像之解析度提升後同樣讓模型進行預測，該深度學習模型能提升 Landsat 8 影像之解析度，使其更接近 Sentinel-2 影像的品質。而在找尋合適的海岸線變遷地點時，觀測到曾文溪口地區，在 2019 年至 2024 年間有明顯的沙洲新堆積與沙洲流失變化，因此選擇此地點為測試區域。



(a) (b) (c) (d)

圖 8 外傘頂洲海岸線變遷 DARNet 預測圖(Sentinel-2 資料集) (a)第一時刻的 RGB 影像、(b)第二時刻的 RGB 影像、(c)海岸線變遷的四分類標籤圖、(d) DARNet 的預測結果

圖 9 為海岸線變遷模型的曾文溪口預測結果，由左至右分別為第一時刻(2019 年)的 RGB 影像、第二時刻(2024 年)的 RGB 影像、模型的預測結果。a 組為第一時刻與第二時刻皆使用 Sentinel-2 影像的組別，b 組為第一時刻使用 Landsat 8 影像、第二時刻使用 Sentinel-2 影像的組別，c 組為第一時刻使用經過深度學習模型提升解析度之 Landsat 8 影像，第二時刻使用 Sentinel-2 影像的組別。由圖 9 可看出 a 組的結果表現最佳，模型能夠準確捕捉沙洲的新堆積與流失區域，顯示在訓練與推論階段使用相同來源(Sentinel-2)資料時，模型的預測能力最為穩定。b 組的結果略遜於 a 組，儘管第一時刻的影像來源被替換為解析度匹配後的 Landsat 8 影像，模型仍能辨識出大部分沙洲的變遷，但可見在流失區域的邊緣出現部分錯誤預測。c 組的表現則不如 b 組，尤其在沙洲新生區域的附近，模型出現許多沙洲流失的錯誤預測，顯示經過深度學習模型提升解度之 Landsat 8 影像，即便解析度有所提升，但受到的影響反而比直接使用 Landsat 8 影像還大，使模型對變遷的敏感性降低。經由圖 9 中 a 至 c 之結果，可觀察到雖然 Landsat 8 影像在經過深度學習模型提升解析度後，雖然相較於原本的 Landsat 8 影像更為清晰，但在將其使用於模型預測變遷方面，則會出現許多錯誤預測。理論上來說 Landsat 8 影像在提升解析度後，使用在模型上應該會預測的不錯，但經過本研究將 Sentinel-2 影像、Landsat 8 影像及深度學習模型提升解析度之 Landsat 8 影像的各個波段，在標準化至 0 到 1 的範圍後，計算這些波段的平均值與標準差並觀察，各波段統計值如表 5 所示。

從表 5 可發現，經由深度學習模型提升解析度之 Landsat 8 影像在短波紅外光 1 與短波紅外光 2 的波段上，能明顯看出提升解析度後的 Landsat 8 影像相較於原本的 Landsat 8 影像與 Sentinel-2 影像的波段統計值相差更多。為使提升解析度後的 Landsat 8 影像與 Sentinel-2 影像具備一致的光譜特性，本研究針對這兩張影像進行統計匹配處理，將提升解析度後的 Landsat 8 影像光譜分佈調整為與 Sentinel-2 影像相近，以降低不同來源資料造成的光譜不一致問題，提高變遷模型於多解析度影像上的泛化能力與分類一致性。

本次實驗以圖 9 區域之第一時刻為例，將經過深度學習模型提升解析度之 Landsat 8 影像作為來源影像，並以 Sentinel-2 影像作為匹配目標。圖 9 中的 d 組為第一時刻進行統計匹配後的組別，由此預測結果可看出，在經過統計匹配後，模型的預測結果有顯著的改善。原先 c 組在沙洲新堆積區域的附近有許多地方出現沙洲流失的錯誤預測，而 d 組的結果則能排除 c 組原本關於沙洲流失的錯誤預測，整體預測也與 a 組較為貼近。綜合本研究於多解析度影像實驗之結果可知，經過統計匹配處理的提升解析度之 Landsat 8 影像，確實能在一定程度上改善模型預測表現，特別是在消除誤判沙洲流失區域方面，相較於未經統計匹配者，預測結果顯得更為接近實際變遷情形。然而，與原始 Sentinel-2 影像(a 組)及重採樣的 Landsat 8 影像(b 組)相比，統計匹配後的高解析度 Landsat 8 影像(d 組)預測結果仍存在一定誤差，顯示若其光譜特性與訓練資料仍存差異，可能影響模型對變遷類別的正確判斷。因此，在應用不同解析度或經過處理的衛星影像進行時序變遷偵測時，仍應審慎評估其與訓練資料在空間與光譜層面的差異性，並視情況進行額外調整與優化，以確保模型推論的準確性與穩定性。

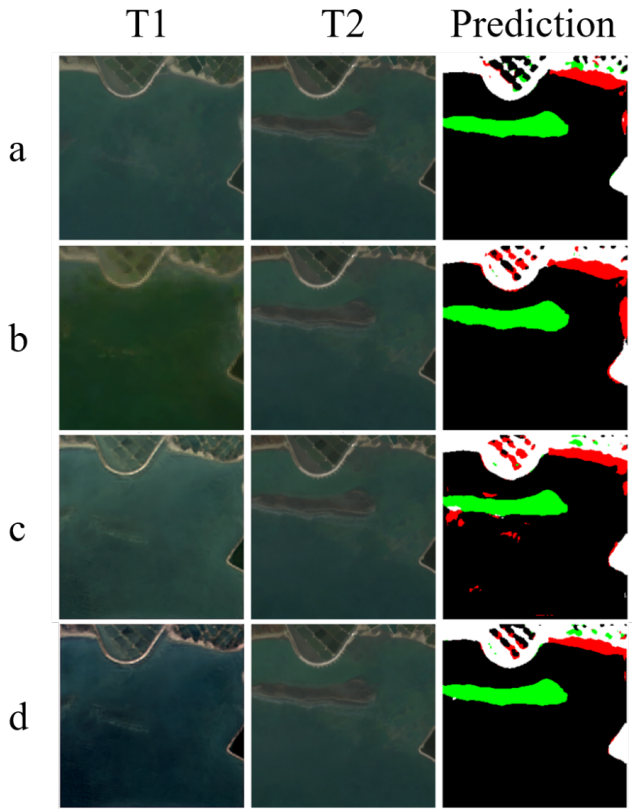


圖 9 曾文溪口區域預測結果比較圖 (a)第一與第二時刻皆使用 Sentinel-2 影像，(b)第一時刻使用 Landsat 8 影像、第二時刻使用 Sentinel-2 影像，(c)第一時刻使用經過深度學習模型提升解析度之 Landsat 8 影像，第二時刻使用 Sentinel-2 影像，(d)第一時刻使用經過深度學習模型提升解析度之 Landsat 8 影像並考量統計匹配，第二時刻使用 Sentinel-2 影像

表 5 第一時刻各影像波段統計值表

Band	Sentinel-2	Landsat 8	Landsat 8_enhance
Blue-Mean	0.211915	0.213621	0.299420
Blue-Std	0.066849	0.161918	0.077704
Green-Mean	0.192156	0.242270	0.265192
Green-Std	0.080115	0.142934	0.081880
Red-Mean	0.105049	0.199971	0.203327
Red-Std	0.104132	0.175523	0.083281
NIR-Mean	0.073730	0.118756	0.120166
NIR-Std	0.139256	0.159504	0.133996
SWIR1-Mean	0.063003	0.085972	0.111305
SWIR1-Std	0.150744	0.146387	0.130027
SWIR2-Mean	0.056335	0.075592	0.091340
SWIR2- Std	0.134421	0.134430	0.115087

5. 結論與研究限制

本研究以深度學習變遷偵測模型及多源衛星影像融合為核心，建立一套應用於多類別地景變遷任務的分類架構，針對不同時期與來源的遙測資料進行整合處理與系統性分析，涵蓋海岸線變遷等主題，比較多種深度學習變遷偵測模型，並結合多來源、多解析度之衛星影像資料進行系統性評估。本研究根據實驗結果與模型評估，提出以下結論與研究限制，最後提出未來研究建議。

(1) 多類別 DARNet 變遷偵測任務中具穩定之表現：本研究評估比較多種深度學習變遷偵測模型於多類別變遷資料集上的預測表現，證明將原始 DARNet 的損失函數修改為能因應多類別的 CE + Dice 損失函數組合，而多類別變遷偵測相較於原先的二分類，能將目標變化與不變化更清

楚的分類，以目標不變、非目標不變、目標增加與目標損失等類別呈現，實驗結果顯示，相較於其他模型，多分類 DARNet 表現更佳，且在變遷任務中(如沙洲流失與堆積)展現良好的空間判別能力。

(2) 統計匹配提升異質影像間之光譜一致性，增進融合應用穩定性：面對多源(Landsat 與 Sentinel-2)與多解析度影像所造成的光譜不一致問題，本研究透過統計匹配處理流程，將超解析後之 Landsat 8 影像光譜調整為與 Sentinel-2 影像相近。處理後模型預測結果顯示，統計匹配有效減少因光譜分布差異造成的分類偏誤與邊界雜訊，提升整體模型泛化能力。

(3) 模型具備跨時期預測能力，可應用於地景演替分析：本研究進一步將訓練完成之多分類 DARNet 模型應用於外傘頂洲及羅鬥沙地區進行多時期影像推論實驗，結果顯示模型能有效辨識長時間尺度下沙洲堆積與流失之空間變遷趨勢，具備一定的時間泛化能力。此能力可應用於地貌演替研究、歷史變遷重建與環境監測中，顯示深度學習架構在長期序列影像應用中之潛力。

儘管本研究在模型應用及調整與資料整合方面取得具體成果，仍有以下幾項限制：

(1) 資料品質受雲層與季節性變異影響：雖然本研究透過 GEE 平台進行年合成與雲遮蔽處理，惟 Landsat 與 Sentinel-2 影像仍可能因雲影、水氣等干擾，導致部分區域資料不連續或異常，進而影響模型學習效果與預測準確性。

(2) 變遷標籤產製倚賴自動閾值法：本研究標籤主要仰賴 NDWI 指標配合 Otsu 法進行分類，缺乏人工標記輔助，使得部分邊界模糊區域或複雜變遷類型可能存在誤分類情形，影響模型評估的準確性與公平性。

(3) 異質資料整合仍受限於感測器差異與幾何校正誤差：儘管本研究已導入統計匹配與超解析技術，仍無法完全克服來自不同感測器(如光譜響應、幾何偏移等)所造成的資料差異，部分測試場景仍可觀察到模型預測出現偏移或錯誤定位。

為進一步強化模型於實務應用中之適應能力與推廣性，未來可朝以下方向發展：

(1) 引入主動式遙測資料以補足光學影像限制：本研究現階段模型訓練與推論皆使用光學衛星影像，易受雲層遮蔽與日照條件限制。未來可考慮整合 SAR(合成孔徑雷達)等主動式遙測資料，其不受氣候與光照影響，能與光學影像形成互補，進一步提升模型在多天候條件下的應用穩健性。

(2) 改善標籤製作方式以提高訓練資料品質：本研究標籤使用 Otsu 法自動分類與 NDWI 指標產製，雖具效率但仍可能對複雜地景邊界判釋不足。未來可採用半自動標註流程，結合自動分類初標與人工修正方式，以提升標籤邊界之準確度與變遷類別之代表性，進而改善模型訓練成效。

(3) Batch size 調整：本研究在現階段模型訓練時，Batch size 設定為 2，雖能獲得不錯的訓練結果，但未來可嘗試採用更高效能的計算設備，或透過適度縮小影像輸入尺寸，將 Batch size 提升至 16、32 等較大數值。較大的 Batch size 有助於模型在訓練過程中獲得更具代表性的批次統計資訊，降低由小批次樣本所造成的梯度波動。此外，較大的 Batch size 不僅有助於提升梯度估計的穩定性，亦可減少每個 epoch 所需的參數更新次數，因此在硬體資源允許的情況下，能有效提升訓練效率。

參考文獻

陳聿新、朱宏杰、盧韋勳，2025a。以衛星大數據觀察長期國土變遷-以外傘頂洲為例，國土測繪與空間資訊，13(2)：215–226。[Chen, Y.X., Chu, H.J., and Lu, W.S., 2005a. Observing long-term territorial changes through spatial big data: A case study of Waisanding sand bar, Taiwan Journal of Geoinfomatics, 13(2): 215-226. (in Chinese)]

陳聿新、簡浩丞、朱宏杰，2025b。海岸線自動偵測與 AI 模型，航測及遙測學刊，30(2)：73–83，DOI：10.6574/JPRS.202506_30(2).0001。[Chen,

Y.X., Chien, H.C., and Chu, H.J., 2025b. Coastline automatic detection and AI model, Journal of Photogrammetry and Remote Sensing, 30(2): 73-83, DOI:10.6574/JPRS.202506_30(2).0001 (in Chinese)]

Aman, M.A., and Chu, H.J., 2023. Long-term river extent dynamics and transition detection using remote sensing: Case studies of Mekong and Ganga River, Science of the Total Environment, 876: 162774, DOI:10.1016/j.scitotenv.2023.162774.

Bishop-Taylor, R., Nanson, R., Sagar, S., and Lymburner, L., 2021. Mapping Australia's dynamic coastline at Mean Sea level using three decades of Landsat imagery, Remote Sensing of Environment, 267: 112734, DOI:10.1016/j.rse.2021.112734.

Blaschke, T., 2010. Object based image analysis for remote sensing, ISPRS Journal of Photogrammetry and Remote Sensing, 65(1), 2–16, DOI:10.1016/j.isprsjprs.2009.06.004.

Chen, H., Xu, X., and Pu, F., 2024. SRC-Net: Bi-temporal spatial relationship concerned network for change detection, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17: 11339-11351, DOI:10.1109/JSTARS.2024.3411622.

Du, Y., Teillet, P.M., and Cihlar, J., 2002. Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection, Remote Sensing of Environment, 82(1): 123-134, DOI:10.1016/S0034-4257(02)00029-9.

Heo, J., and FitzHugh, T.W., 2000. A standardized radiometric normalization method for change detection using remotely sensed imagery, Photogrammetric Engineering and Remote Sensing, 66(2): 173-181.

Jamaluddin, I., Thaipisutikul, T., Chen, Y.N., Chuang, C.H., and Hu, C.L., 2021. MDPrePost-Net: A spatial-spectral-temporal fully convolutional network for mapping of mangrove degradation affected by Hurricane Irma 2017 using Sentinel-2 data, Remote Sensing, 13(24): 5042, DOI:10.3390/rs13245042.

LeCun, Y., Bengio, Y., and Hinton, G., 2015. Deep learning, Nature, 521(7553): 436-444, DOI:10.1038/nature14539.

Li, W., and Gong, P., 2016. Continuous monitoring of coastline dynamics in western Florida with a 30-year time series of Landsat imagery, Remote Sensing of Environment, 179: 196-209, DOI:10.1016/j.rse.2016.03.031.

Li, Z., Yan, C., Sun, Y., and Xin, Q., 2022. A densely attentive refinement network for change detection based on very-high-resolution bitemporal remote sensing images, IEEE Transactions on Geoscience and Remote Sensing, 60: 4409818, DOI:10.1109/TGRS.2022.3159544.

Liao, P.S., Chen, T.S., and Chung, P.C., 2001. A fast algorithm for multilevel thresholding, Journal of Information Science and Engineering, 17(5): 713-727, DOI:10.6688/JISE.2001.17.5.1.

McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features, International Journal of Remote Sensing, 17(7): 1425-1432, DOI:10.1080/01431169608948714.

Sambandham, V.T., Kirchheim, K., Ortmeier, F., and Mukhopadhaya, S., 2024. Deep learning-based harmonization and super-resolution of Landsat-8 and Sentinel-2 images, ISPRS Journal of Photogrammetry and Remote Sensing, 212: 274-288, DOI:10.1016/j.isprsjprs.2024.04.026.

Tewkesbury, A.P., Comber, A.J., Tate, N.J., Lamb, A., and Fisher, P.F., 2015. A critical synthesis of remotely sensed optical image change detection

techniques, Remote Sensing of Environment, 160: 1-14, DOI:10.1016/j.rse.2015.01.006.

Verbesselt, J., Hyndman, R., Newnham, G., and Culvenor, D., 2010. Detecting trend and seasonal changes in satellite image time series, Remote Sensing of Environment, 114(1): 106-115, DOI:10.1016/j.rse.2009.08.014.

Wang, P., Bayram, B., and Sertel, E., 2022. A comprehensive review on deep learning based remote sensing image super-resolution methods, Earth-Science Reviews, 232: 104110, DOI:10.1016/j.earscirev.2022.104110.

Zhang, C., Yue, P., Tapete, D., Jiang, L., Shangguan, B., Huang, L., and Liu, G., 2020. A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images, ISPRS Journal of Photogrammetry and Remote Sensing, 166: 183-200, DOI:10.1016/j.isprsjprs.2020.06.003.

Zhu, Z., Qiu, S., and Ye, S., 2022. Remote sensing of land change: A multifaceted perspective, Remote Sensing of Environment, 282: 113266, DOI:10.1016/j.rse.2022.113266.

Leveraging Deep Learning and Data Integration Techniques for Land Surface Change Detection

Yu-Xin Chen [1] Hone-Jay Chu [2*]

Abstract

This study investigates the application of deep learning models for gain–loss detection in land surface change using bi-temporal satellite imagery. An end-to-end framework is adopted to directly capture semantic and spectral differences between pre- and post-change images. The proposed approach is based on the Densely Attentive Refinement Network (DARNet) and extends it to a multi-class change detection architecture. To improve consistency between multi-source imagery, including Landsat and Sentinel-2, statistical matching and super-resolution techniques are incorporated into the preprocessing workflow. Coastal change detection is selected as the case study to evaluate and compare the performance of several state-of-the-art deep learning models on a multi-class land-cover change dataset. The analysis focuses on class discrimination capability, boundary delineation accuracy, and overall classification performance to provide guidance for future model selection and development. Experimental results demonstrate that the multi-class DARNet integrated with statistical matching and super-resolution achieves superior performance in detecting fine-scale land changes. Furthermore, the proposed data integration strategy enhances both model accuracy and generalization, indicating its strong potential for practical applications in land change detection.

Keywords: Satellite Imagery, Deep Learning, Gain and Loss, Change Detection

[1] Master, Department of Geomatics, National Cheng Kung University
[2] Professor, Department of Geomatics, National Cheng Kung University
[*] Corresponding Author, E-mail: honejaychu@geomatics.ncku.edu.tw

Received Date: Jan. 07, 2026
Revised Date: Feb. 09, 2026
Accepted Date: Jul. 01, 2026

航測及遙測學刊　第三十一卷　第 3 期　第 193- 211 頁　民國 115 年 9 月
Journal of Photogrammetry and Remote Sensing
Volume 31, No.3, 2026, pp. 193 – 211
DOI：10.6574/JPRS.202609_31(3).0004

應用低成本光達於大葉桃花心木人工林測計之可行性評估研究

郭哲源 [1]　魏浚紘 [2]　陳建璋 [3*]

摘要

隨科技日新月異，光學雷達感測器已內建於部分智慧型行動裝置，本研究測試 iPhone 15 Pro 與 viDoc RTK Rover 兩設備於大葉桃花心木人工林進行林下定位及掃瞄可行性試驗。結果顯示，該設備能使手機定位誤差縮小至公分級，可替代傳統 RTK GNSS 接收儀於林下定位採點；有效點雲資料需連接 viDoc 掃瞄，標記地面控制點可提高準確度，胸徑量測以重疊度 90% 為最佳，立木位置則為重疊度 75%，經邊際效應分析運算時長與各測計結果，建議以連接 viDoc 於重疊度 70% 並標記地面控制點為最佳組合。此設備於減少成本同時取得有效點雲資料，然仍受限於掃瞄範圍且對立木胸徑小於 5 cm 偵測能力較弱。

關鍵詞：低成本光達、智慧型行動裝置、點雲分析、RTK GNSS

1. 前言

現今諸多領域已廣泛藉由航空攝影測量(Aerial Photogrammetry, AP)、近景攝影測量(Close Range Photogrammetry, CRP)及光學雷達(Light Detection and Ranging, LiDAR)等方式收集三維資料，相較傳統地面調查可於減少外業成本下，透過非破壞性方式取得可數位化資料(Balenović *et al.*, 2021、Tavani *et al.*, 2022、Martino *et al.*, 2024、Yan *et al.*, 2024)；光學雷達又稱光達，主要由雷射器、掃瞄儀及全球衛星導航系統(Global Navigation Satellite System, GNSS)接收器組成；作業原理為雷射器主動發射出雷射脈衝，並依飛行時間法(Time of Flying, ToF)、相位差(Phase Difference)或三角原理(Triangulation)計算脈衝接觸物體後反射回儀器之距離，推算出三維空間座標(X, Y, Z)，並可搭載於衛載、空載及地面(固定式或移動式)等平台(Platform) (林建忠，1999、魏浚紘，2014) ，相較 AP 及 CRP 具穿透性強、幾何精度穩定、環境適應性高等優勢。

光達於林業上可推估大面積蓄積量或生物量，亦可單獨計算單株立木之林木性態值，提供碳盤查方面極大的優勢(Balenović *et al.*, 2021)，然 LiDAR 系統之購買、資料處理、維護保養等綜合成本較高且儀器體積普遍較大，現今仍難以全面普及(Bi *et al.*, 2021)。而隨近年感測元件發展迅速及諸多關鍵性元件及技術上之突破，如固態光達(Solid-State LiDAR)、互補式金屬氧化半導體(Complementary Metal Oxide Semiconductor, CMOS)、垂直共振腔面射型雷射(Vertical-Cavity Surface Emitting Laser, VCSEL)、單光子雪崩二極體(Single-Photon Avalanche Diodes, SPADs)等，LiDAR 感測器已內建於部分智慧型行動裝置，目前蘋果公司自 2020 年推出 iPhone Pro/Pro Max 及 iPad Pro 系列產品市占率最高(Martino *et al.*, 2024)，再藉由多樣開源軟體進行掃瞄、處理、分析和計算等作業，整體使用成本相較於傳統光達系統便宜，故該類型光達又有低成本光達(Low-cost LiDAR)或消費級光達(Consumer-level LiDAR)之稱，

[1] 國立屏東科技大學森林系 碩士
[2] 國立屏東科技大學森林系 副教授
[3] 國立屏東科技大學森林系 教授
[*] 通訊作者, E-mail: janchang17@gmail.com

收到日期：民國 115 年 01 月 02 日
修改日期：民國 115 年 03 月 27 日
接受日期：民國 115 年 08 月 13 日

其原理與移動式光達相似亦有即時定位與地圖建構技術(Simultaneous Localization and Mapping, SLAM)，近期已於工程、林業、文史、地球科學等領域進行試驗，探討其應用成果及替代傳統光達系統之可行性(Gollob *et al.*, 2021、Mokroš *et al.*, 2021、Tavani *et al.*, 2022、Büyükmıhçı *et al.*, 2023、Chioni et al., 2023、Tamimi & Toth, 2024)。

儘管該類型低成本光達已證實部分林木性態值之掃瞄量測結果不亞於地面式光達，然受限於設備 5 m 之有效掃瞄範圍，於部分環境如 5 m 內出現複數目標、立木徑級變化較大之林地、樹皮複雜度高、環境光強度變化大及立木歪斜等，仍易出現掃瞄異常或錯誤，且對於樹高及樹冠之林木性態值資料收集仍較為困難(Bobrowski *et al.*, 2023、Brach *et al.*, 2023)，若強化行動裝置之 SLAM 演算法或配合 RTK GNSS 設備，可於一定程度提升 LiDAR 感測器掃瞄範圍及準確度並使點雲資料具更精確定位資訊，亦可能與空中獲取之資料結合，以取得更全面之森林或樣區點雲資料(Mokroš *et al.*, 2021、Brach *et al.*, 2023、Tatsumi *et al.*, 2023、Tamimi & Toth, 2024)；PIX4D 與 vigram GmbH 共同研發並推出 viDoc RTK Rover，透過藍牙連接行動裝置並搭配 PIX4Dcatch 軟體，可獲取公分級定位之影像與高品質點雲，該設備目前僅支援內建 LiDAR 感測器之 iOS 系統行動裝置，現今已有研究應用於古蹟建築、戶外道路、地球科學等領域中進行測量及試驗(Büyükmıhçı *et al.*, 2023、Suleymanoglu *et al.*, 2023、Martino *et al.*, 2024、Tamimi & Toth, 2024)。

大葉桃花心木(*Swietenia macrophylla* King)為臺灣中南部平地主要栽植樹種，於日治時期引進至恆春林試所進行栽植試驗，經試驗後發現其高度中南部適應氣候 (馮豐隆等，2010)，因過度採伐與天然更新困難，已於 2002 年列入《瀕危野生動植物種國際貿易公約》(CITES)附錄 II 並被限制出口及貿易，其永續經營成全球關注議題。

目前在臺灣鮮少有應用智慧型行動裝置於林業領域之研究及相關文獻，故本研究以 iPhone 15 Pro 與 viDoc RTK Rover 作為研究設備，於大葉桃花心木平地人工造林地進行：(1)評估 viDoc RTK Rover 於大葉桃花心木人工林之定位準確度；(2)評估智慧型行動裝置掃瞄之點雲資料測計準確度；(3)建立智慧型行動裝置之最佳掃瞄組合等 3 項掃瞄可行性試驗，以作為未來於林業應用之基礎。

2. 材料與方法

2.1 研究地點

本研究地點位於屏東縣萬巒鄉 185 縣道之萬金營區旁大葉桃花心木平地人工林，該造林地隸屬於臺灣糖業股份有限公司屏東區處屏東農場管理，於 1997 年配合政府林業政策種植，位置如圖 1 所示，共設立 3 個樣區並由北自南分別為 A1、A2 及 A3。

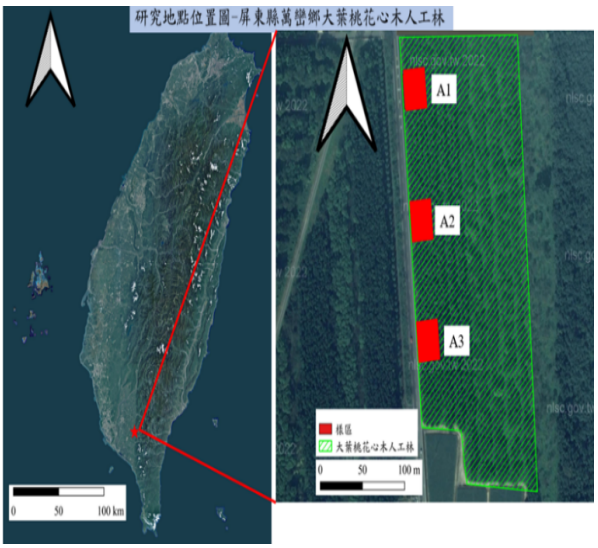


圖 1 研究地點位置圖

本研究地點海拔高度為 55 m，坡度為 5–10°，根據中央氣象署氣候觀測資料查詢服務(CWB Observation Data Inquire Severic, CODiS)統計，距離本研究地點最近之氣象站為赤山氣象站，設址於萬巒鄉建興路 40 號(22.592333, 120.614306)，海拔高度 32 m，本研究統計 2015 年 4 月至 2025 年 3 月之氣象資料，該地區年均溫約 25.7 ℃，平均年雨量約 3,019 mm，屬亞熱帶季風氣候類型，4–10 月為特濕期(Perhumid)，11–3 月為相對乾旱期(Relative Drought)，如圖 2 所示。

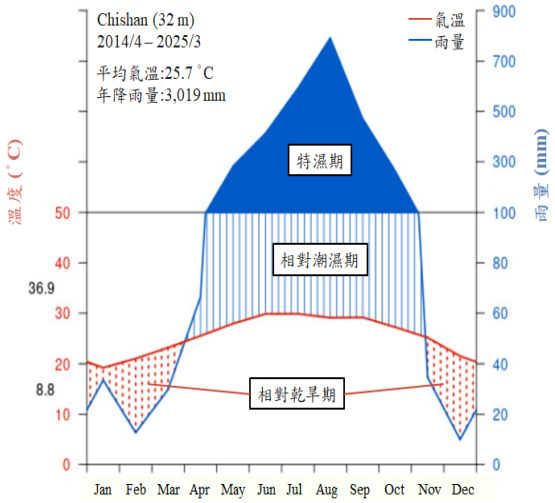


圖 2 研究地點生態氣候圖

2.2 研究對象

大葉桃花心木為楝科(Meliaceae)桃花心木屬(Swietenia)之大型落葉木本植物，原產自中南美洲，好生於低海拔之熱帶或亞熱帶地區，一般樹高可至 30 – 35 m，屬中淺根性植物且常有板根，葉為偶數羽狀複葉且於三、四月落葉，花為黃綠色不完全花，果實為卵形木質蒴果，種子為褐色扁長型且帶有長翅；枝葉生長茂密適合作為庭園或行道樹種，其木材性質優良，可用於製造家具、樂器及船艦等(馮豐隆等，2010)。

本研究於 2025 年 2 – 3 月設置共 3 個 25x 20 m 樣區，經地面調查後各樣區之資訊包括，A1、A2 及 A3 樣區之平均胸徑分別為 24.9、22.7 及 23.6 cm，最小胸徑為 4.0、3.5 及 6.4 cm，最大胸徑為 40.5、42.5 及 41.9 cm，立木株數為 39、46 及 55 株，立木密度為 780、960 及 1,100 trees/ha；將地面調查各樣區各立木胸徑值繪製成以 0 – 45 cm 為範圍並以 5 cm 為分級之胸徑級圖，直觀了解各樣區之立木胸徑概況，如圖 3 所示，其中立木胸徑小於 5 cm 共 3 株，分別為 A1 樣區 1 株及 A2 樣區 2 株。

2.3 研究設備

2.3.1 iPhone 15 Pro

蘋果公司之「iPhone Pro/Pro Max」及「iPad Pro」系列智慧型行動裝置均內建 LiDAR 感測器，採用直接式 ToF(directly Time of Flying, dToF)進行掃瞄，以 VCSEL 發射雷射脈衝，由 SPADs 接收並產生電流以記錄光子飛行時間，從而計算目標物空間位置；相比採樣頻率每秒數百萬點傳統光達系統的飛行時間法，dToF 採樣頻率僅每秒數十萬點，主要針對小範圍、室內或緩慢移動物體且視角固定，雖掃瞄範圍較短，然功耗低、環境光抗性增強且提高目標物細部成像；本研究以 iPhone 15 Pro 作為研究設備，其 LiDAR 感測器具深度感測能力，有效掃瞄範圍約 5 m，主相機鏡頭使點雲上色，如圖 4 所示，結合慣性測量單元(Inertial Measurement Unit, IMU)數據，經視覺里程計(Visual-Inertial Odometry, VIO)演算法獲取手機之姿態及位移資訊，A17 Pro 晶片可穩定進行掃瞄，1TB 容量可儲存資料，主機總重量為 187 g (Apple, 2024)。

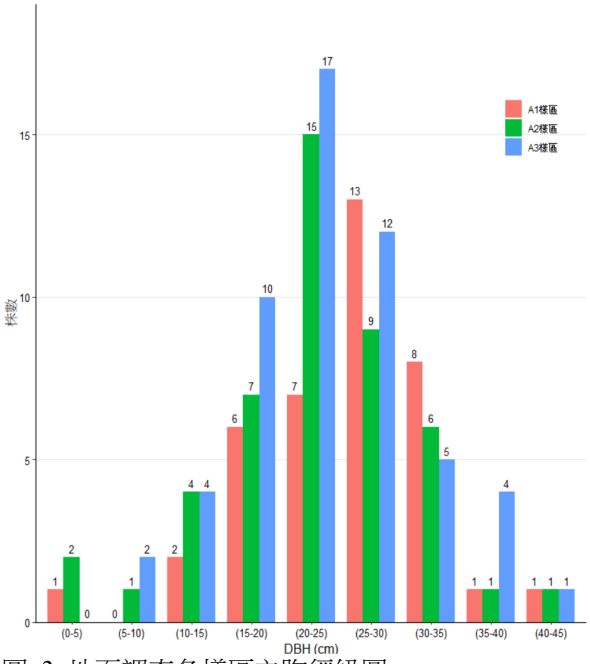


圖 3 地面調查各樣區之胸徑級圖



圖 4 iPhone 15 Pro 鏡頭配置圖

2.3.2viDoc RTK Rover

PIX4D 與 vigram GmbH 於 2021 年共同研發推出 viDoc RTK Rover，其設計目的在於使智慧型行動裝置具備公分級定位能力及透過自身 LiDAR 感測器進行掃瞄並立即生成點雲模型，用以取代傳統 RTK GNSS 基站或流動站及昂貴的光達系統，使建築工程人員能更輕鬆、快速且安全地完成決策流程及工程作業(viDoc, 2024)，其總重量約 261 g，目前主要搭配之智慧型行動裝置以 iOS 系統且內建 LiDAR 感測器之系列為主，設備外觀及各部位名稱如圖 5 所示。



圖 5 viDoc RTK Rover 外觀及各部位名稱

該設備進行作業需搭配 PIX4Dcatch 應用程式，當開機運行後需以藍牙方式與智慧型行動裝置連接，並透過固定端口連接行動裝置，方可於軟體介面中選擇 viDoc RTK Rover 以進行連線，連線後須設定網際網路傳輸 RTCM 協定(Network Transport of RTCM via Internet Protocol, NTRIP)以進行兩項作業：(1) RTK 定位，藉由 GNSS 接收器接收 GPS(含 L1C/A 及 L2C)、BeiDou(含 B1I 及 B2I)、Galileo(含 E1-B/C 及 E5b)、GLONASS(含 L1OF 及 L2OF)等四種衛星星系之訊號進行初步定位，再經由網路以 NTRIP 協定連結基地台傳輸 RTCM 即時校正資料，最後設備將自動校正大氣、多路徑及衛星軌道等多項誤差，最終於最大測量傾斜角 ± 20° 內取得約 1 cm ± 1 ppm 之定位精度 (viDoc, 2024)；(2)點雲模型生成，掃瞄原理為透過軟體內建之 GeoFusion 演算法，藉由行動裝置自身 GNSS 或 viDoc 取得 RTK 定位資訊即時結合 IMU、VIO 及 LiDAR 感測器之數據，並藉由連續影像估算設備位置，類似於視覺 SLAM 技術，當 短暫失去 RTK 訊號時，可藉由 IMU 與 VIO 之相對軌跡進行即時補償，以確保移動軌跡與姿態之連續性與精度 (PIX4Dcatch, 2026)；掃瞄時主鏡頭依影像重疊度或設備姿態進行影像觸發，LiDAR 感測器用於捕捉環境深度數據，影像及數據均具備經 GeoFusion 校正後之精確空間地理資訊，再經由密集影像匹配技術(Dense Image Matching Technology)即時生成點雲資料及模型，掃瞄完成後可匯出影像(.JPG)、點雲資料(.PLY)、網格模型(.OBJ)、紀錄檔(.LOG)及精度報告(.PDF)。

2.3.3 Stonex S9II RTK GNSS

本研究為進行 viDoc 於林下定位之準確性，以義大利 Stonex 公司之第二代 RTK GNSS 接收儀作為基準設備，其由全鎂合金外殼堅固且輕量化，可接收 7 種星系訊號，平面精度達 ±(8 mm + 1 ppm)，垂直精度為 ±(15 mm +1 ppm)，可藉由網路、外置電台、內置電台及 WiFi 四種方式連接。

2.3.4電腦設備

本研究生成和處理點雲模型及計算林木性態值所使用之電腦設備系統及硬體資訊如下，作業系統為 Windows 10 專業版 22H2，處理器為 Intel® CoreTM i7-7700 CPU @ 3.60GHz、24.0 GB 記憶體、NVIDIA® Corporation Quadro P2000/PCIe/SSE2 圖形處理器及 64 位元作業系統；因點雲資料具高密度數據特性，其資料生成、處理和計算之效率與電腦硬體效能呈正相關，增強硬體規格可有效減少運算時長。

2.3.5軟體設備

本研究所使用之軟體可分為外業掃瞄、模型生成及林木性態值提取等三方面，使用 PIX4D 公司開發之 PIX4Dcatch 應用程式進行掃瞄，並藉由 PIX4Dmatic 將影像生成為點雲模型，再透過 LiDAR360 提取所需之林木性態值，最後以 R 4.3.2 進行統計分析及繪製統計圖表，各軟體之版本及相關資訊，如下列所述：

(1) PIX4Dcatch 2.10.3，適用於內建 LiDAR 感測器之智慧型行動裝置的攝影測量軟體，藉由 GeoFusion 演算法與多感器器數據融合，以生成包含影像、點雲資料、網格模型及精度報表等資料。

(2) PIX4Dmatic 1.71.0，由 PIX4D 開發專用於處理無人機與地面測繪之影像拼接軟體，生成高精度點雲模型及網格模型，藉由 GPU 加速技術有效縮短運算時長，可於生成模型時進行地面控制點或手動標記連結點，進而提高資料之準確度；本研究應用其拼接 PIX4Dcatch 所拍攝之影像，並生成點雲模型，使用軟體時需先行率定步驟，方可選擇包含密集點雲、網格模型、數值地表模型等後續計算，本研究使用率定和密集點雲後即匯出精度報表及點雲模型。

(3) LiDAR360 v2.1，為專用於處理及分析來自地面及空載光達點雲資料之軟體，可進行裁減、躁點去除、自動化分類等運算，其設有專用於林業處理之模組，可進行單木分割、提取林木性態值及計算等功能；本研究應用其將生成之點雲模型進行處理並標記欲提取林木性態值之立木，以產出被標記立木之胸徑值及立木點位之.CSV。

(4) R 4.3.2，為專用於統計計算與圖形繪製之開源軟體環境，可於 UNIX、Windows 和 macOS 等多種平台運行，除提供多樣的統計模型及快速繪製相關圖表外，亦可藉由指令修訂及編輯圖表。

2.4 研究方法

本研究之研究目的為定位準確度及點雲資料準確度，並確立於人工林下之最佳掃瞄組合；本研究設置 3 個 25 × 20 m 樣區，每樣區內有 5 列立木，於每列立木中設置 2 個 GCP 及額外設置 1 個於樣區外作為掃瞄路徑之起終點，共計 11 個，如圖 6 如所示，GCP 除作為定位比較外，亦用於後續 PIX4Dmatic 拼接影像之標記點，設置完成後即進行地面調查及掃瞄試驗。

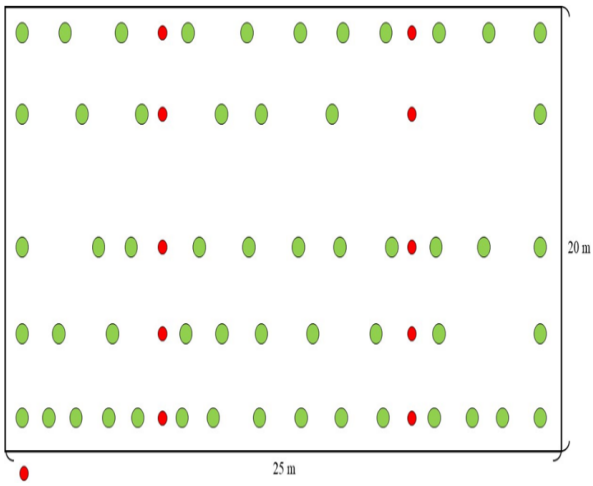


圖 6 樣區立木及 GCP 位置示意圖，綠色點為立木，紅色點為 GCP

2.4.1定位資料之準確度

本研究以三角錐作為 GCP，並於上方放置圓形球體作為球型共軛目標，如圖 7(a)所示。將三角錐放置完畢後，使用腳架搭載 S9II RTK GNSS，將腳架垂直置入三角錐之孔洞中後藉由腳架本身之水平氣泡儀定平，水平座標系統依 TWD97 EPSG:3826 設置，垂直座標系統依 Ellipsoidal height over the GRS 1980 ellipsoid 設置，當設備接收至固定解時，依每 2 秒採點 1 次共 20 次之設置進行 RTK 定位，如圖 7(b)所示；使用 viDoc RTK Rover 定位時，需於軟體上新增點位資料夾，座標系統與 S9II 使用相同設定，定位時行動裝置螢幕中綠色圓圈會顯示設備當前姿態，取得固定解後依每 1 秒採點 1 次共 15 次進行 RTK 定位，如圖 7(c)所示。使用 S9II 採點後會扣除標竿高，得實地高程資訊；viDoc RTK Rover 採點需開啟「底部雷射」，其底部會發射雷射光，高程是以雷射光接觸之目標物，非設備當下高度。

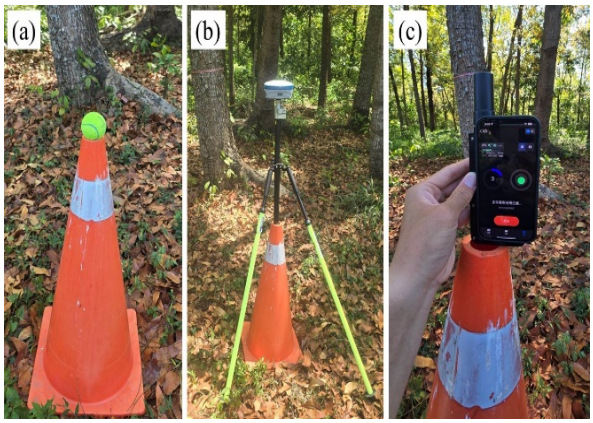


圖 7 GCP 設置與設備定位圖 (a)GCP 及共軛目標；(b) S9II 定位；(c) viDoc 定位

定位完成後，以 S9II 測量值為基準值，與 viDoc RTK Rover 測量值依北緯(以 N 表示)、東經(以 E 表示)及高程(以 h 表示)進行比較，計算逐點之差值及 RMSE、二維和三維之平均值、標準差，並透過成對樣本 t 檢定及 Pearson 相關係數(Pearson Correlation，以 r 表示)進行統計分析，各計算公式如下：

$$\begin{bmatrix}\Delta E\\ \Delta N\\ \Delta h\end{bmatrix}=\begin{bmatrix}E_{s9II}-E_{viDoc}\\ N_{s9II}-N_{viDoc}\\ h_{s9II}-h_{viDoc}\end{bmatrix} \quad (1)$$

$$\begin{bmatrix}\Delta_{2D}\\ \Delta_{3D}\end{bmatrix}=\begin{bmatrix}\sqrt{\Delta E^2+\Delta N^2}\\ \sqrt{\Delta E^2+\Delta N^2+\Delta h^2}\end{bmatrix} \quad (2)$$

$$RMSE=\sqrt{\sum_{i=1}^{n}\frac{d_i^2}{n}} \quad (3)$$

$$\overline{x}=\frac{\sum x_i}{n} \quad (4)$$

$$\sigma=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n}} \quad (5)$$

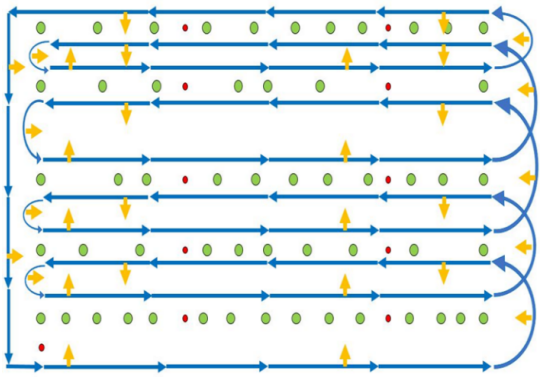
式中，$\overline{x}$為平均值(m)；i 為第 i 項差值；n 為總 GCP 數；d_i 為二維差值(Δ_{2D})或三維差值(Δ_{3D})，以 m 為單位。

2.4.2點雲資料測計準確度

本研究透過文獻回顧已知，藉由該類型設備於 5 m 外之林木性態值難以收集，故本研究以胸徑(Diameter at Breast Height, DBH)、立木位置(Tree Position, TP)、立木偵測率(Tree Detect Rate, TDR)、漏授率(Omission)及誤授率(Commission)進行比較，以地面調查值為基準值。

(1) 地面調查獲取林木性態值，本研究使用 Tanda 2.0M 胸徑尺於離地面高 1.3 m 處進行立木胸徑量測，以 cm 為單位；使用 Stonex S9II RTK GNSS 於樣區內每株立木旁並盡可能貼近樹幹進行立木位置採點，水平座標系統依 TWD97 EPSG:3826 設置，定平後依每 1 秒採點 1 次共 10 次進行 RTK 定位，將所獲取之立木位置資訊，分別以北緯及東經計算，以 TP_N 與 TP_E 表示，以 m 為單位；立木偵測率以目視方式，給與立木編號及統計每樣區立木數量。

(2) 掃瞄路徑設立，欲建置立木完整樹幹點雲資料，需至少以 2 個方向針對同一株立木進行影像獲取，本研究以手持設備離地面高約 1.3 m 處，鏡頭方向與行進方向夾 90° 角，始終將主鏡頭面朝左側並垂直正對立木，於樣區內每列立木繞行，如圖 8 所示，綠色圓圈為立木位置，紅色圓圈為 GCP 位置，藍色箭頭為行進方向，橘色箭頭為鏡頭方向。


圖 8 掃瞄路徑示意圖

(3) 掃瞄條件設立，本研究於每樣區中各使用 PIX4Dcatch(依手機自身 GNSS 定位，後以 PIX4Dcatch 表示)和 PIX4Dcatch + viDoc RTK Rover(依設備取得 RTK 定位，後以 viDoc 表示)進行掃瞄，影像觸發方式依 70%、75%、80%、85%及 90%共 5 種重疊度設定，而兩種定位模式及 5 種重疊度之每種組合於各樣區重複 3 次進行試驗；使用 viDoc RTK Rover 時，水平座標系統依 TWD97 EPSG:3826 設置，垂直座標系統依 Ellipsoidal height over the GRS 1980 ellipsoid 設置。

(4) 掃瞄過程，以不遮蔽相機鏡頭且朝向欲掃瞄方向，步行速度每秒約 0.5 – 1.0 m 平穩且勻速慢步前行；當連接上 viDoc RTK Rover 後，可藉由 PIX4Dcatch 介面知悉 GNSS 解算狀態，於固定解時螢幕上方會呈現綠色且顯示「固定解」，接收至固定解後再進行掃瞄，移動過程可能遇 RTK 信號中斷，此時螢幕上方會呈現紅色且跳出「遺失 RTK 定位訊號」，同時顯示「浮動解」或「GNSS 單點解」，當步行至上空開闊地或遠離障礙物後即恢復 RTK 訊號；螢幕左下方會顯示當前影像數量，如圖 9 所示。

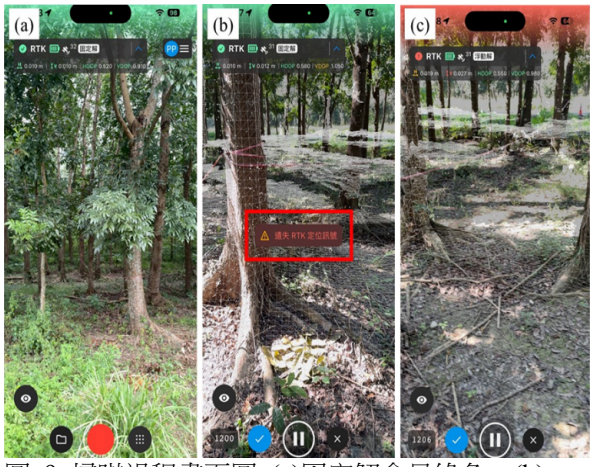


圖 9 掃瞄過程畫面圖 (a)固定解會呈綠色；(b) RTK 訊號中斷會出現警訊；(c) RTK 訊號中斷會呈紅色且顯示當前連線狀態

(5) 影像生成點雲模型，使用 PIX4Dmatic 作為影像拼接之軟體，其專為 PIX4Dcatch 建立處理範本，水平座標之座標系統為 TWD97 EPSG:3826，垂直座標之座標系統為 Ellipsoidal height over the GRS 1980 ellipsoid；點雲資料之生成需進行「率定」、「密集點雲」及「匯出檔」，率定之設定為使用 PIX4Dcatch 範本、依信任的位置及姿態計算、影像比為 1/1、自動計算特徵點、內方位參數可信度為低等，密集點雲之設定為使用硬體加速計算、影像比為 1/2、點雲密度最佳化、最小匹配數為 3、使用雜訊過濾等；匯出檔部分則勾選處理之精度報表及點雲模型。

(6) PIX4Dmatic 影像處理時進行標記，可行地面控制點(Ground Control Points, GCP)及手動連結點(Manual Tie Points, MTP)，兩者相同處均為透過運動恢復結構(Structure from Motion, SfM)建立點雲模型，計算原理為於每張影像中提取易辨認之特徵點，依有相似特徵點之影像進行初始分類，再藉由幾何約束進一步篩選相同或相似特徵之影像以提高建模準確性，而使用率較高之幾何約束包括光束法平差(Bundle Adjustment)、基礎矩陣(Fundamental Matrix)、RANSAC(RANdom SAmple Consensus) 和本質矩陣(Essential Matrix)等多種單個或多個組合方式，將篩選後影像之當時拍攝位置及姿態(本研究之研究設備掃瞄時即透過自身 IMU 感測器自行記錄相關數據)解算及分析其掃瞄路徑，再藉由三角化推估並建立點雲模型；兩者之不同處在於，GCP 為匯入已知精確之真實三維座標，藉由絕對幾何約束強制平差各特徵點，以使點雲模型符合現實座標，僅少量點即可提高準確度，MTP 則為於影像間建立虛擬三維座標之相對連結點，藉由相對約束以輔助提高模型相對精度，若於環境複雜地則需標記大量虛擬點；此外亦有驗證模型定位精度之檢核點(Check Points)，該點不參與平差及校正模型，僅根據真實三維座標計算與標記後特徵點之三維位置誤差、重投影誤差、精度及平均均方根位置誤差(PIX4Dmatic, 2025)，研究已證實結合 GCP 標記相比於單用 RTK 定位，可降低誤差(Tamimi & Toth, 2024)。本研究設置樣區時所放置之三角錐，除用於定位比較外，亦用於影響點雲資料準確度變因之一，探討有無標記 GCP 是否能糾正點雲資料以提高準確度。匯入 GCP 點位後，即可於單張影像中標定單個特徵點(本研究以圓形球體作為共軛目標)，每 GCP 各標定 20 張影像，其餘處理設定條件均相同；於標定時如圖 10，黃色十字為標記之特徵點，藍色十字為該真實點於影像中之位置，綠色十字則為標定後重投影推估位置。

(7)點雲資料提取林木性態值，藉由 LiDAR360 分別進行點雲資料裁切、躁點去除、分類地面點、根據地面點常態化及 TLS 種子點編輯。點雲資料裁切主要移除非目標點雲，降低檔案大小及運

算負擔；躁點去除為針對每個點計算到鄰近設定點之平均距離，並計算中位數及標準差，移除其餘大於設定標準差倍數之點雲，本研究設定鄰近點數為 10，倍數為 5；分類地面點為計算重要基礎步驟之一，旨在區分地面及其他物體之屬性，可有效減少計算量與誤差並高分析效率與精度，該軟體由改良式漸進加密三角網格濾波(Improved Progressive TIN Densification, IPTD)進行地面點分類，本研究設定最大建物尺寸 20 m、最大地形坡度 88°、迭代角度 10°及迭代距離 1.3 m 進行分類；常態化可降低因地形起伏造成之誤差，該軟體可根據 DEM 及地面點進行計算，本研究選用根據地 面點最後以 TLS 種子點編輯，設定斷面高度為 1.25–1.35 m，其會以 2D 呈現該高度之點雲，再根據地面調查之立木編號依序將點雲圈選，並使用最佳擬合圓(Fit Circle)計算胸徑，每筆資料完成後即可匯出每株立木之編號、二維座標及胸徑之.CSV 檔；二維座標部分，LiDAR360 提取立木性態值時，會以最佳擬合圓之圓心座標表示。



圖 10 標定特徵點過程圖，黃色標記點，藍色為真實點，綠色為重投影點

(8)評估準確度，本研究以地面調查值為基準值，變數為胸徑、立木位置(TP)及立木偵測率(TDR)，前 2 者分別計算平均值、標準差、變異係數、MAE 及 RMSE，後者以百分比表示；透過成對樣本 t 檢定及 Pearson 相關係數，計算各組合與地面調查值之精準度與系統性差異；此外為消除點雲量測可能存在之系統性偏差並提升資料精確度，本研究將採用普通最小平方法(Ordinary Least Squares, OLS)建立簡單線性迴歸校正模型(Calibration Models)，各計算公式如下：

$$\overline{x} = \frac{\sum x_i}{n} \quad (6)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n}} \quad (7)$$

$$CV(\%) = \frac{\sigma}{\overline{x}} * 100 \quad (8)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (9)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (10)$$

$$y_i = \beta_0 + \beta_1 X_1 + \varepsilon_i \quad (11)$$

式中，$\overline{x}$為平均值；x_i為第 i 株立木數值；n 為總株數；y_i為第 i 株立木真實值；$\hat{y}_i$為第 i 株立木推估值；β_0為迴歸模型截距；β_1為迴歸模型斜率；ε_i為誤差項。

2.5 評估最佳掃瞄組合

為從 2 種定位模式、5 種重疊度及建立點雲模型時是否標記 GCP 共 3 種變因評估出最佳掃瞄組合，本研究應用邊際效應(Marginal effect, ME)作為評估方式，將其定義為軟體運算單位時間增加所帶來的精度改善率(Incremental Accuracy Gain per Unit Time)，將依立木胸徑及立木位置之 RMSE 與各組合之運算時長進行分析以找出最佳點，並應用 R 繪製成圖，其公式如下：

$$ME = \frac{RMSE_{i-1} - RMSE_i}{\text{運算時長}_i - \text{運算時長}_{i-1}} \quad (12)$$

式中，$RMSE_i$為第 i 項組合之立木胸徑或位置 RMSE 值；運算時長$_i$為第 i 項組合之運算時長(min)。

2.6 研究流程圖

本研究之目的分為評估定位準確度及探討最高效且準確之掃瞄流程，前者與傳統 RTK GNSS 進行定位比較，後者之變因包含 GNSS 和 RTK、5 種重疊度及影像拼接時是否標記 GCP，架構如圖 11 所示。

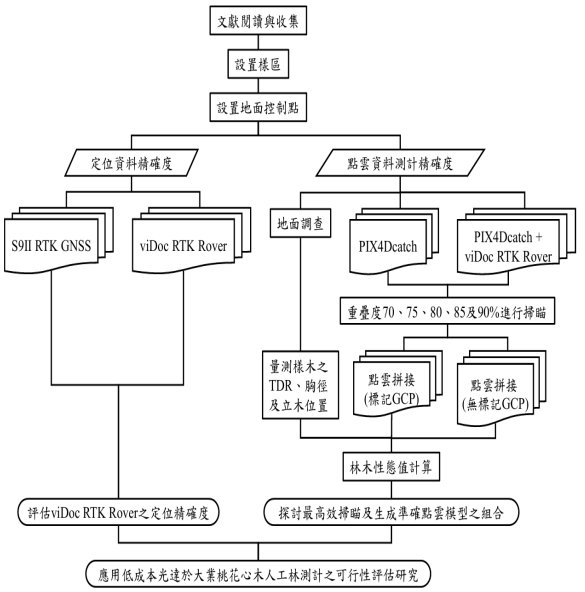


圖 11 本研究之流程圖

3. 結果與討論

3.1 定位準確度

本研究以 S9II 之定位採點值做為基準值，驗證 viDoc RTK Rover 作為 RTK GNSS 定位之準確度，於各樣區之 GCP 進行定位採點，將各方向性及誤差彙整成表 1，表中 N 為北緯、E 為東經、h 為高程(本研究僅以橢球高進行兩設備之相對精度驗證)、2D 為二維平面方向而 3D 為三維空間方向。由表可知在 N 及 E 兩方向之總平均差值均小於 0.003 m，且各標準差均大於平均值，說明誤差分布較隨機，而 h 方向各樣區均為負偏差且總平均差值為 -0.019 m，表示高程之測量存在微小的系統性偏差；3 樣區中 N 及 E 兩方向測量差值有減少之趨勢，而 h 方向則逐漸增加，二維平均差值中 A1 最高(0.026 m)，A3 最低(0.016 m)，三維平均差值中 A1 最高(0.043 m)，A2 最低(0.036 m)；RMSE2D 於各樣區分別為 0.027、0.022 及 0.018 m，與 N 及 E 兩方向相同呈減小趨勢，而 RMSE3D 為 0.045、0.040 及 0.048 m，說明高程方向於 A3 樣區影響最為明顯。

將 3 樣區中各方向性之測量值依成對樣本 t 檢定進行統計分析，N 方向之 S9II 與 viDoc 平均值分別為 2499639.378 及 2499639.077 m，E 方向分別為 211018.077 及 211018.076 m，h 方向分別為 55.971 及 55.989 m，N、E 及 h 之差值分別為 0.301、0.001 及 0.018 m，同時 r 均大於 0.998，表示兩儀器之定位數據呈高度相關；經 t 檢定後發現，N 及 E 方向 t 值分別為 0.671 及 0.539，P-value 均大於 0.05，顯示定位結果於二維平面方向無顯著差異，而 h 方向之 t 值為 -3.241，P-value 為 0.003，具有顯著差異，雖平均值差僅 0.018 m，然並非隨機誤差所導致，而為系統性誤差。

依據各 GCP 之各方向性位移數值繪製成盒鬚圖，如圖 12 所示，(a)、(b)及(c)分別為樣區 A1、A2 及 A3，由圖中可知 ΔN 在各樣區之中位數均趨近於 0，ΔE 之變異數稍大於 ΔN，且於 A1 樣區中出現離群值，然測量結果仍優於 Δh，因此 Δ2D 數值相較穩定；Δh 在各樣區中變異性明顯，為主要造成 Δ3D 位移量明顯之主要因子。

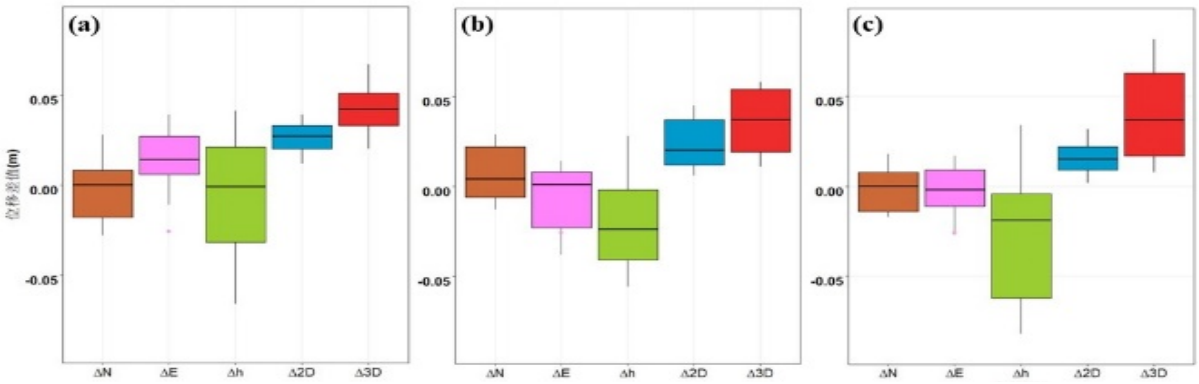


圖 12 各樣區 GCP 方向性差值之盒鬚圖與折線圖，ΔN 為棕色，ΔE 為粉色，Δh 為綠色，Δ2D 為藍色及 Δ3D 為紅色 (a)為 A1 樣區；(b)為 A2 樣區；(c)為 A3 樣區

表 1 各樣區ΔN、ΔE、Δh、Δ2D 及Δ3D 之差值及誤差表

樣區	A1	A2	A3	Total
ΔN (m)	-0.010 ± 0.018	0.007 ± 0.014	0.0 ± 0.012	0.002 ± 0.015
ΔE (m)	0.013 ± 0.018	-0.005 ± 0.017	-0.003 ± 0.015	0.002 ± 0.018
Δh (m)	-0.007 ± 0.036	-0.021 ± 0.027	-0.029 ± 0.036	-0.019 ± 0.033
Δ2D (m)	0.026 ± 0.009	0.019 ± 0.013	0.016 ± 0.009	0.020 ± 0.011
Δ3D (m)	0.043 ± 0.014	0.036 ± 0.016	0.042 ± 0.025	0.040 ± 0.019
RMSE2D (m)	0.027	0.022	0.018	0.023
RMSE3D (m)	0.045	0.040	0.048	0.044

本研究以傳統 RTK GNSS 接收儀 Stonex S9II 為基準值，應用 iPhone 15 Pro 連接 viDoc RTK Rover 於人工林下進行定位準確度評估，在二維及三維之 RMSE 測量結果與 Tamimi(2022)每約 203 m 誤差為 0.3 m 及 Demirel & Türk(2024)之 13 個 GCP 於二維和三維 RMSE 分別為 0.049 m 和 0.056 m 兩研究結果相符，然經統計分析後具有顯著差異，觀察各樣區數值可發現，當立木密度增加時，高程誤差亦隨之增加，造成此誤差可能與樹冠對 GNSS 訊號的遮蔽程度以及地形起伏有關，魏浚紘(2014)研究表明樹冠遮蔽可能導致部分衛星訊號接收不穩定，尤其影響高程方向的精度，或受周圍環境物體(如建物、地形、立木等)反射而產生多條傳播路徑現象之多路徑效應有關，該效應會造成設備同時接收直接訊號與延遲之反射訊號(Dong & Chen, 2017、Bian & Xu, 2024)，而 Stonex S9II 作為傳統測繪級接收儀具備較大之天線開口面與優異之多路徑效應抑止能力，其於高程測量展現較高之穩定度；viDoc 可於森林環境中進行定位採點，其中二維資料可信度高，而高程資料較不穩定，本研究建議實地應用時需待設備之垂直精度衰減因子(Vertical Dilution of Precision, VDOP)數值最低或穩定時再行採點，該數值可於連接之設備螢幕所見；或於林外訊號穩定處架設 RTK GNSS 設為固定站，以 viDoc 作為流動站連接固定站訊號，可於遮蔽環境下取得精確定位資訊。

3.2點雲資料測計準確度

3.2.1點雲資料測計準確度

本研究分別比較兩種定位模式，依行動裝置自身 GNSS(PIX4Dcatch)或外加設備以 RTK(viDoc)訊號進行掃瞄，兩種定位模式之點雲資料立木外觀具明顯差異；於 A1 樣區及重疊度 90%進行掃瞄並標記 GCP 之點雲資料，分別以樣區視角及離地面 1.3 m 高之水平截面進行比較，應用 PIX4Dcatch 且無論重疊度及是否標記 GCP，其點雲資料均如圖 13(a)所見之立木外觀分岔、重疊或重複顯像嚴重，1.3 m 截面均如圖 13(b)，難以簡單且快速辨別單株立木，對後續應用 TLS 種子點編輯進行林木性態值提取造成困難，故本研究後續將不予計算及討論 PIX4Dcatch 之點雲資料，而藉由 viDoc 且無論重疊度及是否標記 GCP，其點雲資料均如圖 13(c)所見之立木外觀完整且排列整齊，而 1.3 m 截面均如圖 13(d)，胸徑點雲分布完整且容易辨別，減少後續人工或自動化提取及計算之時間及錯誤率，故後續以 viDoc 掃瞄之點雲資料作為主要探討測計準確度。

3.2.2胸高直徑(DBH)

本研究針對 140 株立木，以地面調查值(後以 Ground 表示)作為基準值，與 viDoc 各組合之點雲資料依平均值、標準差、CV、MAE 及 RMSE 計算並進行成對樣本 t 檢定統計分析，將結果彙整於表 2，由表可知 Ground 之平均值、標準差及 CV 分別為 23.7 cm、7.6 cm 及 32.0%，viDoc 各掃瞄組合之平均值最高及最低為 viDoc-90-GCP 及 viDoc-70-GCP，標準差之最高為 viDoc-70，最低為 viDoc-80 和 viDoc-90，CV 最高為 viDoc-70，最低為 viDoc-80 和 viDoc-90，MAE 之最高及最低為 viDoc-80 及 viDoc-90-GCP，RMSE 之最高及最低為 viDoc-80 及 viDoc-90-GCP；成對樣本 t 檢定統計分析後可知，各組合之 r 值介於 0.899 - 0.987 之間，表量測結果與地面調查具高度線性關係，t 值介於 4.672 – 10.126， P-value 均小於 0.001，表示無論重疊度及標記 GCP 與否，均有極顯著差異且有系統性偏差。為評估並校正點雲資料之系統性偏差，進一步將各組合建立 OLS 迴歸校正模型及誤差分析，由表 3

迴歸校正後可知，各組合絕對誤差呈縮減趨勢，有效降低系統性偏差所造成的量測誤差，其中以 viDoc-90-GCP 為最佳，RMSE 為 1.2 cm。

剖析每筆立木胸徑資料發現，未標記 GCP 時，小於 10 cm 者有較大誤差且多為高估，大於 35 cm 者亦有較大誤差且多為低估，介於 15–35 cm 之間者則誤差值較小；標記 GCP 後各重疊度之平均誤差值均明顯下降，亦可於表 2 及表 3 之 MAE 及 RMSE 所見差異，由此可證其可降低量測誤差，具有糾正點雲資料且提高準確度之功能。

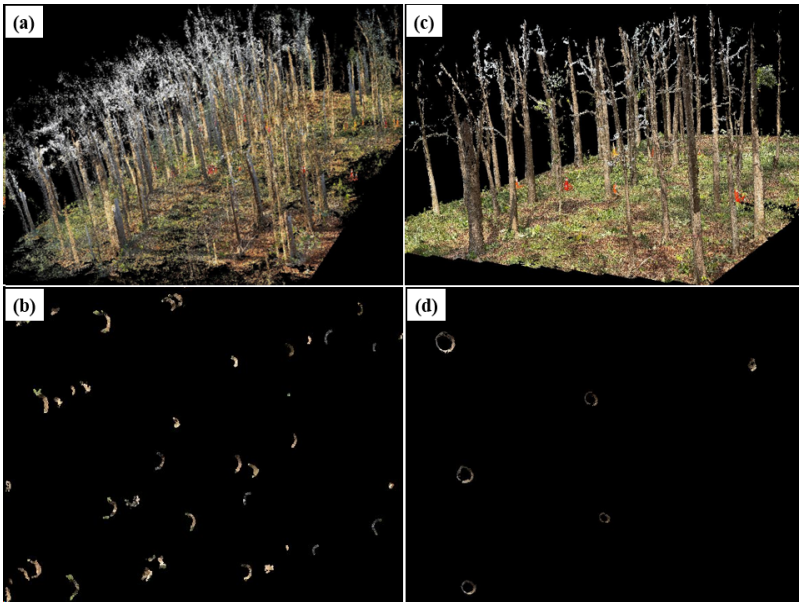


圖 13 相同樣區及重疊度之不同定位模式點雲資料圖 (a) PIX4Dcatch 樣區視角；(b) PIX4Dcatch 1.3m 高截面；(c) viDoc 樣區視角；(d) viDoc 1.3m 高截面

表 2 地面調查與各掃瞄組合掃瞄之胸徑結果表

Group	平均值 (cm)	標準差 (cm)	CV (%)	MAE (cm)	RMSE (cm)	r	t (df=139)	P-value
Ground	23.7	7.6	32.0	-	-	-	-	-
viDoc-70	22.0	7.5	34.1	2.4	3.0	0.948	8.222	<0.001***
viDoc-75	22.2	7.4	33.2	2.3	3.0	0.937	6.745	<0.001***
viDoc-80	22.4	6.8	30.3	2.8	3.5	0.899	4.672	<0.001***
viDoc-85	22.3	6.9	31.1	1.8	2.5	0.965	8.513	<0.001***
viDoc-90	22.4	6.8	30.3	1.7	2.2	0.977	9.141	<0.001***
viDoc-70-GCP	21.9	7.3	33.6	2.2	3.0	0.949	9.193	<0.001***
viDoc-75-GCP	22.4	7.4	33.0	2.1	2.8	0.944	6.094	<0.001***
viDoc-80-GCP	22.1	7.1	31.9	2.0	2.5	0.968	9.895	<0.001***
viDoc-85-GCP	22.5	7.0	31.3	1.5	1.9	0.982	10.126	<0.001***
viDoc-90-GCP	22.7	7.2	31.8	1.4	1.6	0.987	9.959	<0.001***

註：*為 p<0.05、**為 p<0.01、***為 p<0.001

表 3 各掃瞄組合之簡單線性迴歸模型表

Group	線性迴歸模型	MAE (cm)	RMSE (cm)	r
viDoc-70	$y_i = 2.635 + 0.9572X_i$	1.8	2.4	0.899
viDoc-75	$y_i = 2.2822 + 0.9654X_i$	1.9	2.6	0.879
viDoc-80	$y_i = 1.1636 + 1.0064X_i$	2.3	3.3	0.810
viDoc-85	$y_i = 0.1846 + 1.0569X_i$	1.3	2.0	0.932
viDoc-90	$y_i = -0.7192 + 1.0918X_i$	1.2	1.6	0.956
viDoc-70-GCP	$y_i = 2.3045 + 0.9799X_i$	1.6	2.4	0.901
viDoc-75-GCP	$y_i = 2.0157 + 0.968X_i$	1.9	2.5	0.891
viDoc-80-GCP	$y_i = 0.6995 + 1.0415X_i$	1.2	1.9	0.937
viDoc-85-GCP	$y_i = -0.0442 + 1.0584X_i$	1.0	1.4	0.965
viDoc-90-GCP	$y_i = 0.1484 + 1.0397X_i$	0.9	1.2	0.975

3.2.3立木位置(TP)

為探討 viDoc RTK Rover 獲取立木位置之準確性，本研究應用其各組合掃瞄之點雲資料提取立木位置，並藉由 S9II 以靠近樹幹方式定位為基準值進行成對樣本 t 檢定，將結果彙整於表 4，表中因各組合存在不同程度點漏授問題而缺少 1–3 株立木資訊，使 df 值不相同，在北緯(TP_N)及東經(TP_E)方向各掃瞄組合均呈高度線性相關，TP_N 僅 viDoc-75-GCP 和 viDoc-85-GCP 與 S9II 無顯著差異，而 TP_E 為 viDoc-80、viDoc-75-GCP 和 viDoc-85-GCP 三種組合無顯著差異。將各掃瞄組合與 S9II 量測值相減得差值，再依平均值、標準差、MAE 及 RMSE 計算並整理成表 5，由表可知 TP_N 平均值最高及最低為 viDoc-80 及 viDoc-85，標準差之最高及最低為 viDoc-90 及 viDoc-70 -GCP，MAE 之最高及最低為 viDoc-85 及 viDoc-70，RMSE 之最高及最低為 viDoc-85 及 viDoc-75-GCP，TP_E 平均值最高及最低為 viDoc-80 及 viDoc-85，標準差最高及最低為 viDoc-90 及 viDoc-70 -GCP，MAE 最高及最低為 viDoc-85 及 viDoc-80-GCP，RMSE 最高及最低為 viDoc-85 及 viDoc-75-GCP。標記 GCP 之誤差均有降低，不同於胸徑結果之影像重疊度增加則誤差降低，而此兩方向均於較低重疊度時有較低之量測誤差，尤為重疊度 75% 時均有最低，重疊度 85% 且未標記 GCP 均有最大測量誤差。

將立木位置之誤差分為有或無標記 GCP 兩組，橫軸為 TP_N 誤差和縱軸為 TP_E 誤差繪製成圖 14，圖 14(a)為未標記 GCP，各重疊度之 TP_N 誤差介於 -4.0–2.0 m 之間，TP_E 誤差介於 -2.0–2.0 m 之間，其中重疊度 70 % 出現 1 最大誤差點，而經標記 GCP 後如圖 14(b)所示，整體誤差有明顯下降，各重疊度之 TP_N 誤差介於 -2.0–2.0 m 之間，TP_E 誤差介於 -1.0–1.0 m 之間，對於重疊度 70% 之校準效果最為顯著。

表 4 各掃瞄組合與地面調查值之立木位置統計分析表

Group	df	TP_N			TP_E		
		r	t-value	P-value	r	t-value	P-value
viDoc-70	137	0.999	-2.822	0.006**	0.999	2.827	0.005**
viDoc-75	137	0.999	5.462	<0.001***	0.999	-7.322	<0.001***
viDoc-80	137	0.999	7.646	<0.001***	0.999	1.187	0.237
viDoc-85	137	0.999	-7.853	<0.001***	0.999	-5.767	<0.001***
viDoc-90	138	0.999	2.373	0.019*	0.999	-5.181	<0.001***
viDoc-70-GCP	137	0.999	5.121	<0.001***	0.999	-2.337	0.021*
viDoc-75-GCP	137	0.999	-1.292	0.197	0.999	0.214	0.831
viDoc-80-GCP	138	0.999	7.297	<0.001***	0.999	-3.695	<0.001***
viDoc-85-GCP	138	0.999	0.639	0.524	0.999	1.370	0.173
viDoc-90-GCP	139	0.999	2.967	0.004**	0.999	-4.559	<0.001***

註：*為 p<0.05、**為 p<0.01、***為 p<0.001

表 5 地面調查與各掃瞄組合掃瞄之 TP_N 及 TP_E 結果表

Group	TP_N				TP_E			
	平均值 (m)	標準差 (m)	MAE (m)	RMSE (m)	平均值 (m)	標準差 (m)	MAE (m)	RMSE (m)
viDoc-70	-0.091	0.380	0.165	0.390	0.036	0.162	0.113	0.164
viDoc-75	0.211	0.460	0.339	0.505	-0.124	0.182	0.148	0.219
viDoc-80	0.256	0.398	0.326	0.472	0.110	0.275	0.169	0.295
viDoc-85	-0.297	0.450	0.360	0.538	-0.187	0.368	0.289	0.422
viDoc-90	0.100	0.499	0.330	0.507	0.096	0.300	0.209	0.314
viDoc-70-GCP	0.147	0.341	0.257	0.370	0.054	0.141	0.103	0.147
viDoc-75-GCP	-0.034	0.314	0.251	0.315	0.013	0.146	0.112	0.143
viDoc-80-GCP	0.241	0.392	0.317	0.459	-0.038	0.167	0.095	0.171
viDoc-85-GCP	0.021	0.388	0.277	0.387	0.063	0.258	0.165	0.264
viDoc-90-GCP	0.099	0.395	0.260	0.406	-0.066	0.195	0.123	0.218

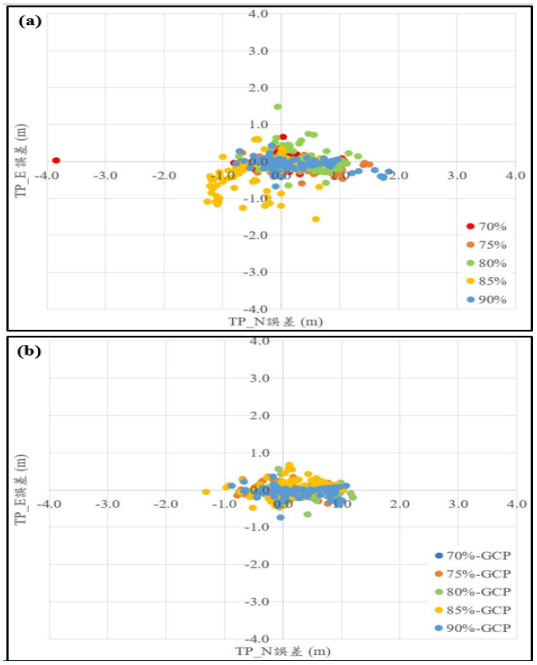


圖 14 標記 GCP 與 TP 誤差散佈圖(a)無標記 GCP；(b)有標記 GCP

3.2.4立木偵測率(TDR)、漏授率及誤授率

立木偵測率、漏授率及誤授率均是依照點雲資料中所出現之立木株數做比較，TDR 為統計所出現之總株數比例，Omission 為統計實際存在於現實而未被掃瞄到或偵測到導致資料中無顯現之比例，Commission 為統計不存在於現實卻於資料中顯現之比例。將各樣區 140 株立木掃瞄結果依掃瞄組合進行分類，並整理於表 6，由表中可知各重疊度均無完全被偵測，TDR 介於 97.9 – 99.3 % 之間，漏授率介於 0.7 – 2.1 % 之間，而誤授率均為 0 %，本研究發現於各樣區中共有 3 株胸徑小於 5 cm 之立木，分別為 3.5、4.0 及 4.4 cm，3.5 cm 之立木於各重疊度下均 100 % 未被偵測，而其餘二者均有機率被漏偵測；普遍而言，重疊度增加時可降低立木被漏授之機率，尤為 90 % 且標記 GCP 時有最低比例，而標記 GCP 可提高重疊度大於 80 % 之立木偵測能力；本研究發現若立木胸徑小於 4.0 cm，無論何種組合均無法被偵測，若介於 4.0 – 5.0 cm 之間則有機率被漏授，而立木大於 5.0 cm 者均可被 100 % 偵測。

表 6 各掃瞄組合之立木偵測率、漏授率及誤授率比例表

Group	TDR (%)	Omission (%)	Commission (%)
viDoc-70	97.9	2.1	0.0
viDoc-75	97.9	2.1	0.0
viDoc-80	97.9	2.1	0.0
viDoc-85	97.9	2.1	0.0
viDoc-90	98.6	1.4	0.0
viDoc-70-GCP	97.9	2.1	0.0
viDoc-75-GCP	97.9	2.1	0.0
viDoc-80-GCP	98.6	1.4	0.0
viDoc-85-GCP	98.6	1.4	0.0
viDoc-90-GCP	99.3	0.7	0.0

3.2.5綜合討論

點雲資料準確度為本研究第二探討面向，以 2 種定位模式、5 種影像重疊度及標記 GCP 與否進行試驗。結果顯示僅依靠設備自身 GNSS 定位無法取得有效之點雲資料，造成此結果可能包含：(1)智慧型手機不同於傳統測繪級 GNSS，其天線高度較短，故 GNSS 定位誤差介於 1–5 m，因手機定位之不穩定性而影響影像之座標，造成點雲資料成像困難(Mokroš *et al.*, 2021、Bian & Xu, 2024)；(2)行動裝置本身 SLAM 功能不足應用於該面積之掃瞄(Chioni *et al.*, 2023)，若降低掃瞄面積或立木株數可減少後續誤差；(3)樹冠遮蔽和樹葉反射會降低接收衛星訊號，導致信號重疊和不穩定，從而造成累積誤差(魏浚紘，2014；Dong & Chen, 2017)，此項因子影響程度次於 GNSS 天線高度，尤其樹冠結構不固定時，影響更明顯(Bian & Xu, 2024)；(4)林下環境複雜，包含立木間距過近可能造成重複掃瞄(Mokroš *et al.*, 2021、Gollob *et al.*, 2021、Brach *et al.*, 2023)，或立木樹皮複雜、光照量及角度不均等其他環境因素造成(Bobrowski *et al.*, 2023、Chioni *et al.*, 2023)。而藉由 RTK 定位輔助後點雲資料品質有顯著性地提升，除立木點雲均完整且易辨識外，亦發現最大高度可至 8 m，超文獻所提之 5 m 掃瞄範圍(Gollob *et al.*, 2021、Guenther *et al.*, 2024)。於不同重疊度及標記 GCP 與否之組合中，提高影像重疊度可提取越多特徵點，因此有助於提升胸徑測量精度，藉由標記 GCP 可有效校正點雲之整體位置和方向偏差，進一步提高胸徑資料之準確度，而本結果

與 Tamimi & Toth(2024)研究結果相符，藉由 RTK 定位並加入 GCP 進行運算校正，能顯著降低點雲誤差，本研究最低誤差組合為使用 viDoc RTK Rover 與 90 %影像重疊度及標記 GCP，然而經過成對樣本 t 檢定後發現，各組合均與地面調查值存在極顯著差異，表明存在系統性偏差，造成此因可能與掃瞄路徑有相關，因相機與立木僅呈垂直掃瞄，故將缺乏立木兩側之影像，致使點雲資料建模時缺乏直接性依據，導致立木量測均普遍出現低估之情形，如圖 15 中可直觀了解。

於立木位置將其拆分為北緯及東經兩方向性進行分析，RMSE 方面不同於立木胸徑量測結果，重疊度與其較無規律性之關係，共同之處在於標記 GCP 後各組合之誤差均有降低，再證明其具糾正點雲模型之功能，且各組合之量測值均與地面調查有高相關係數；然經成對樣本 t 檢定統計後發現，僅部分組合與地面調查值呈無顯著差異，該結果存在系統性偏差，與 Brach *et al.* (2023)研究相符均受地面調查之限制，因應用 S9II 難以取得立木真實中心點位，而利用點雲資料提取之立木位置為透過最佳擬和圓方式所推估正中心或靠近中心之點位，應藉由其他傳統光達系統並於相同軟體及條件下處理，較易進行比較及驗證。

立木偵測率部分，於 A1 及 A2 樣區中各有 1－2 株胸徑小於 5.0 cm 之立木，隨影像重疊度之提高，在其餘條件相同下每筆資料可偵測到立木之比例將有所提高，且標記 GCP 及提高影像重疊度可提高立木偵測率，各組合於 A2 樣區中之最小立木 100%漏授，胸徑為 3.5 cm，此與 Brach *et al.* (2023)表明內建於智慧型行動裝置之 LiDAR 感測器對於直徑小於 5.0 cm 物體解析度不足相符，本研究建議可嘗試藉由多角度掃瞄以提高每株立木之被掃瞄面積。

於 Mokroš *et al.* (2021)、Gollob *et al.* (2021)、Brach *et al.* (2023)及 Chioni *et al.* (2023)等研究中，均於人工林或都市林內進行掃瞄試驗，其試驗地為上空開闊且之立木間距較大之都市林或於經營完善且林木整齊之人工林地進行掃瞄，不依靠 viDoc RTK Rover，立木胸徑之準確度即與地面光達量測值相近(RMSE 介於 4.2－24.2 cm)，然 Guenther *et al.* (2024)提及當立木密度提高及林下植被越茂密時，測量誤差均會隨之提高，本研究各組合立木胸徑 MAE 介於 1.4－2.8 cm，而 RMSE 介於 1.6－3.5 cm，均低於所回顧之文獻，證實該設備於人工林下進行掃瞄時可有效提高點雲資料之準確性，同 Tamimi (2022)研究提及，單獨使用智慧型行動裝置之 LiDAR 感測器有效掃瞄範圍有限，難以收集遠距離物體之數據，未來研究可探討如何結合不同視角掃瞄或使用其他技術來提取高度相關之林木性態值及提高對小徑木的偵測能力。

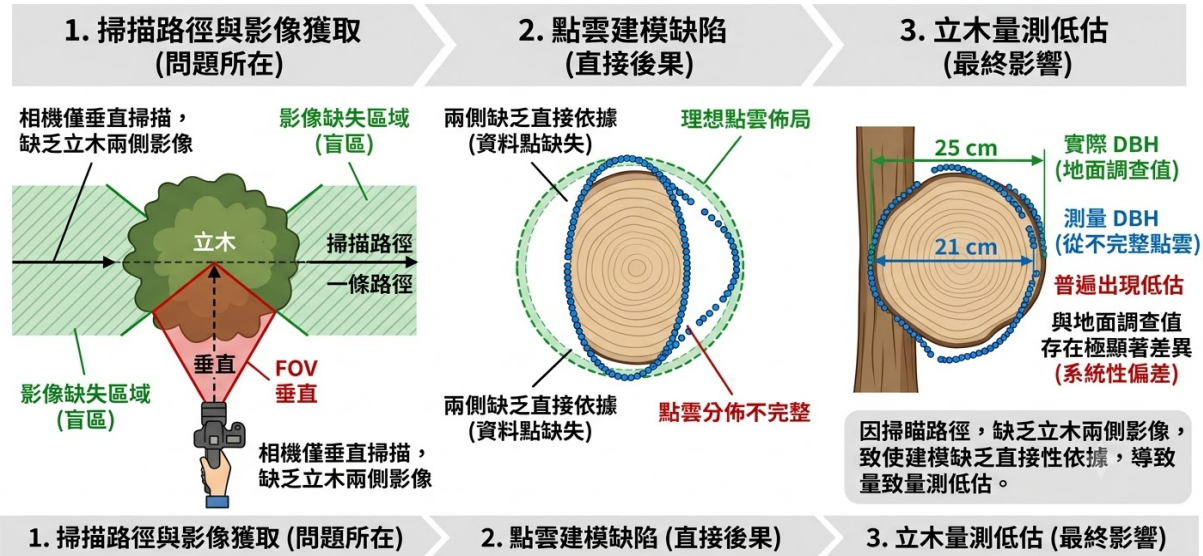


圖 15 立木點雲資料產生系統性誤差之核心原因與演變過程圖

3.3探討最佳掃瞄組合

3.3.1viDoc 各組照片數及運算時長

應用 PIX4Dmatic 建模點雲資料時產出之精度報表中，包含所處理之照片張數與進行率定及點雲模型兩項處理所花費之時長(以分鐘計算)，將各樣區各組合之影像張數、率定運算時長、點雲模型運算時長及兩者處理時長之總和依是否標記 GCP 進行分組，並分別計算平均值及標準差，並繪製如圖 16，由圖可知重疊度、影像張數及運算時長互為正相關，而有標記 GCP 之運算時長普遍高於未標記 GCP，且無論標記 GCP 與否，重疊度 90 %之運算時長需花費其他約 2－10 倍；此外本研究於每樣區針對 11 個 GCP 共標記 220 張影像，無論重疊度為何均平均花費約 20 分鐘時間進行標記作業，故標記 GCP 之總和運算時長加上標記所花費時間，方為整項作業之時間成本。

3.3.2最佳掃瞄組合

最佳掃瞄組合係指花費最少時間取得最準確資料，亦可視為運算時間與測量誤差之最佳平衡，而應用 viDoc 為首要條件；重疊度為主要影響運算時長之因子，普遍而言重疊度提高時其資料誤差越小；最後為 GCP 標記，經文獻回顧及本研究結果均證實標記後可有效提高點雲資料準確度，亦可增加小徑木之偵測能力。

為探討最佳掃瞄組合，本研究將各重疊度及標記 GCP 與否之組合，藉由胸徑 RMSE 和兩方向性立木位置 RMSE 分別與運算時長進行邊際效應評估，於各誤差及時間成本間決策出最佳之掃瞄組合，藉由 R 進行計算並繪製成圖 17，如圖 17 所示橫軸為運算時長，縱軸為 RMSE，圖 17(a)胸徑、圖 17(b)TP_N 及 圖 17(c)TP_E，經分析後發現立木胸徑及兩方向立木位置之最佳掃瞄組合均為藉由 viDoc RTK Rover 和重疊度 70 % 進行影像獲取並標記 GCP。

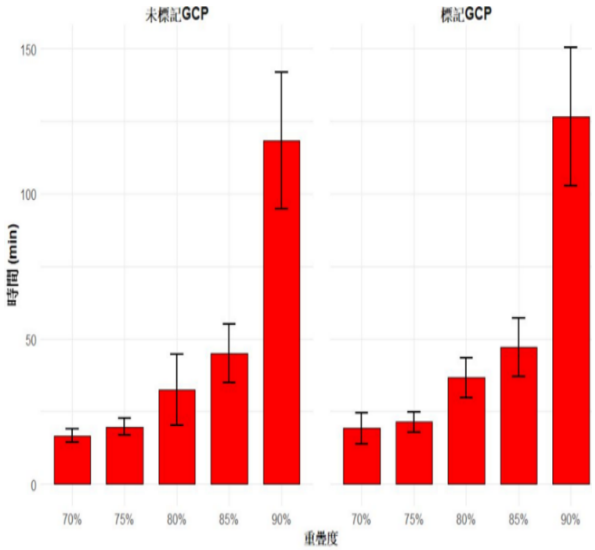


圖 16 各影像重疊度與標記 GCP 與否之運算時長圖

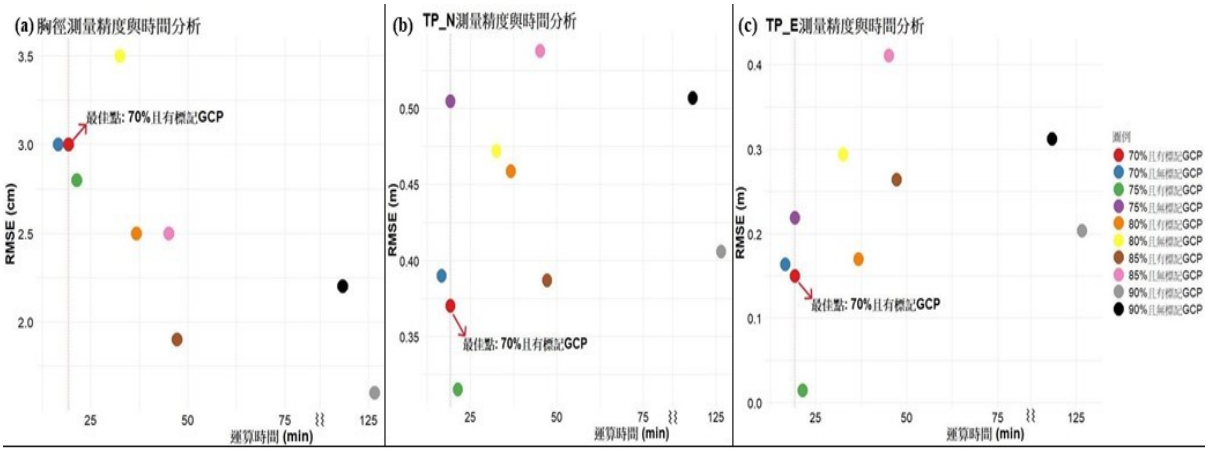


圖 17 各數值之精度與時間分析圖 (a)胸徑；(b)TP_N；(c)TP_E

藉邊際效應評估出各誤差與時間成本之最佳掃瞄組合，然不可忽略各組合仍各有優缺點，故本研究依自身研究結果進行幾項建議之應用分類，若需大量樣區掃瞄以取得全林分之胸徑及立木位置之點雲資料，精度之要求較低且以縮短作業時間為主，建議以重疊度 70 %進行掃瞄，並額外標記 GCP 以提高點雲資料之準確度；若於少量樣區掃瞄，以點雲資料精度為首要要求，不考量時間成本下，以重疊度 90 %且標記 GCP 為首要選擇，雖運算時長為其於組合之 2–10 倍，然其誤差為最低，而若考量時間成本，則建議重疊度降低；應用於上空開闊之都市林，其 RTK 之訊號接收應較為穩定，重疊度 90 %以下測量之精度應高於本研究(本研究掃瞄時，半數時間均於林下，故訊號中斷時長較長)，若考量時間成本，建議以重疊度 70 %，若上空遭遮蔽又欲取得高精度之測量結果，重疊度 90 %且標記 GCP 應為最佳組合。

4. 結論與建議

本研究應用 viDoc RTK Rover 於大葉桃花心木人工林進行定位準確度之結果與傳統RTK GNSS 資料無顯著差異，當其連接網路並取得固定解進行採點後，可取得平面(0.023 m)及空間(0.044 m)之高精度結果，於應用時需額外注意 VDOP 之數值，其為主要影響定位誤差之因子，然該設備之輕巧及便攜等優勢，未來可作為替代傳統 RTK GNSS 接收儀之潛力，此外本研究僅以橢球高進行兩設備之純粹相對精度驗證，若未來應用於工程實務，可透過 TWVD2000 大地水準面模型轉換為正高使用；點雲資料測計準確度方面，依靠 iPhone 15 Pro 自身 GNSS 定位進行掃瞄產出之點雲資料無法使用，而藉由 viDoc RTK Rover 取得公分級定位則顯著輔助提升點雲資料品質及準確性，儘管掃瞄過程遺失 RTK 訊號，起始時確保可連接上，仍可取得高精度資料；當重疊度增加時，立木胸徑之精度及偵測率均可有效提高，而立木位置則較無明顯關聯，然當重疊度為 90 % 時將花費最長運算時長，為其他重疊度約 2–10 倍，標記 GCP 後將增加些許運算時長，然可確保其誤差與不標記相比更小；無論何種組合於立木胸徑和位置均與地面調查相比多具有顯著差異且存在系統性誤差，然均具有高相關性且各 RMSE 與其他文獻相當或更勝於。於最佳掃瞄組合部分，建議均使用 RTK 定位以獲得可用之點雲資料，標記 GCP 能有效糾正點雲資料，明顯降低誤差，雖邊際效應評估後得重疊度 70 %且標記 GCP 為最佳組合，然本研究建議仍應考量掃瞄目的之精度及時間再選擇合適組合，若不考量時間成本則建議重疊度 90 %，因其可取得最高精度之點雲資料且可高小徑木之偵測率。本研究證實 iPhone 15 Pro 連接 viDoc RTK Rover 是可應用於平地人工林收集 8 m 內之部分林木性態值，然胸徑小於 5 cm 之立木及超 8 m 範圍之目標物有高機率無法順利掃瞄獲取，此情形將可能降低測量精度和提高誤授、漏授及誤差，而對於收集冠層與樹高相關資訊須仰賴與其他技術之結合，如空載光達或航空攝影等；建議未來研究可探討多角度掃瞄或結合其他軟硬體設備，此外亦需於不同環境及林相進行試驗，以進一步提高該類型設備於森林應用之可行性。

參考文獻

林建忠，1999。雷射測距技術與研究現況，光電產業與技術情報，19：30-33。[Lin, J.Z., 1999. Laser ranging technology and research status, Optolink, 19:30-33. (in Chinese)]

馮豐隆、張愷玲、張鈞媛，2010。大葉桃花心的生物、生態與利用，生物科學，52(2)：15-24，DOI： 10.29981/CB.201012.0002。[Feng, F.L., Chang, K.L., and Chang, J.Y., 2010. Biology, ecology and usage of big-leaf mahogany, Chinese Bioscience, 52(2):15-24, DOI:10.29981/CB.201012.0002. (in Chinese)]

魏浚紘，2014。應用光達技術於人工林之經營與監測，國立屏東科技大學生物資源研究所，博士學位論文。[Wei, C.H., 2014. Application of LiDAR system for monitoring and management of artificial forest, PhD Thesis, National Pingtung

University of Science and Technology, Taiwan, ROC. (in Chinese)]

Apple, 2024. Compare iPhone Models, Available at: https://www.apple.com/iphone/compare/?modelList=iphone-15,iphone-15-pro, Assessed October 22, 2024.

Balenović, I., Liang, X., Jurjević, L., Hyyppä, J., Seletković A., and Kukko A., 2021. Hand-held personal laser scanning–current status and perspectives for forest inventory application, Croatian Journal of Forest Engineering, 42(1): 165-183, DOI:10.5552/crojfe.2021.858.

Bian, X., and Xu, Z., 2024. Research on multipath error detection model in forest based on Signal-to-noise ratio, Journal of Physics: Conference Series, 2849(1): 012055, DOI:10.1016/j.ufug.2022.127815.

Bobrowski, R., Winczek, M., Zięba-Kulawik, K., and Wężyk, P., 2023. Best practices to use the iPad Pro LiDAR for some procedures of data acquisition in the urban forest, Urban Forestry and Urban Greening, 79: 127815, DOI:10.1016/j.ufug.2022.127815.

Brach, M., Tracz, W., Krok, G., and Gąsior, J.,. 2023. Feasibility of low-cost LiDAR scanner implementation in forest sampling techniques, Forests, 14(4): 706, DOI:10.3390/f14040706.

Büyükmıhçı, G., Akşehirlioğlu, A., and Koca, M., 2023. Advantages and disadvantages of laser scanning with ViDoc Rtk antenna: Ali Cafer Kümbeti, Advanced LiDAR, 3(2): 41-46.

Chioni, C., Maragno, A., Pianegonda, A., Ciolli, M., Favargiotti, S., and Massari, G.A., 2023. Low-Cost 3D virtual and dynamic reconstruction approach for urban forests: The Mesiano University Park, Sustainability, 15(19): 14072, DOI:10.3390/su151914072.

Demirel, Y., and Türk, T., 2024. Assessment of the location accuracy of points obtained with a Low-Cost Lidar scanning system and GNSS method, Mersin Photogrammetry Journal, 6(2):60-65, DOI:10.53093/mephoj.1540159.

Dong, P., and Chen, Q., 2017. LiDAR Remote Sensing and Applications, CRC Press, USA, DOI:10.4324/9781351233354.

Gollob, C., Ritter, T., Kraßnitzer, R., Tockner, A., and Nothdurft, A., 2021. Measurement of forest inventory parameters with Apple iPad pro and integrated LiDAR technology, Remote Sensing, 13(16): 3129, DOI:10.3390/rs13163129.

Guenther, M., Heenkenda, M.K., Morris, D., and Leblon, B., 2024. Tree diameter at breast height (DBH) estimation using an iPad Pro LiDAR scanner: A case study in boreal forests, Ontario, Canada, Forests, 15(1): 214, DOI:10.3390/f15010214.

Martino, A., Lingua, A.M., and Maschio, P., 2024. Affordable sensors for speditive and accurate documentation of built heritage: First tests and preliminary results, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 48: 487-493, DOI:10.5194/isprs-archives-XLVIII-1-2024-487-2024.

Mokroš, M., Mikita, T., Singh, A., Tomaštík, J., Chudá, J., Wężyk, P., Kuželka, K., Surový, P., Klimánek, M., Zięba-Kulawik, K., Bobrowski, R., and Liang, X., 2021. Novel low-cost mobile mapping systems for forest inventories as terrestrial laser scanning alternatives, International Journal of Applied Earth Observation and Geoinformation, 104: 102512, DOI:10.1016/j.jag.2021.102512.

PIX4Dcatch, 2026. How Does PIX4Dcatch Achieve Professional Accuracy?, Available at: https://www.pix4d.com/blog/advantages-of-pix4dcatch/, Assessed Augest 6, 2026.

PIX4Dmatic, 2025. Tie points - PIX4Dmatic, Available at: https://support.pix4d.com/hc/en-us/articles/360045074152#Check_errors, Assessed May 18, 2025.

Suleymanoglu, B., Tamimi, R., Yilmaz, Y., Soycan, M., and Toth, C., 2023. Road infrastructure mapping by using iPhone 14 Pro: An accuracy assessment, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 48: 347-353, DOI:10.5194/isprs-archives-XLVIII-M-1-2023-347-2023.

Tamimi, R. 2022. Relative accuracy found within iPhone data collection, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 43: 303-308, DOI:10.5194/isprs-archives-XLIII-B2-2022-303-2022.

Tamimi, R., and Toth, C., 2024. Experiments with combining LiDAR and camera data acquired by UAS and smartphone to support mapping, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 48: 619-627, DOI:10.5194/isprs-archives-XLVIII-1-2024-619-2024.

Tatsumi, S., Yamaguchi, K., and Furuya, N., 2023. ForestScanner: A mobile application for measuring and mapping trees with LiDAR-equipped iPhone and iPad, Methods in Ecology and Evolution, 14(7): 1603-1609, DOI:10.1111/2041-210X.13900.

Tavani, S., Billi, A., Corradetti, A., Mercuri, M., Bosman, A., Cuffaro, M., Seers, T., and Carminati, E., 2022. Smartphone assisted fieldwork: Towards the digital transition of geoscience fieldwork using LiDAR-equipped iPhones, Earth-Science Reviews, 227: 103969, DOI:10.1016/j.earscirev.2022.103969.

viDoc, 2024. viDoc Official Website, Available at: https://vidoc.com/en/use-cases, Assessed October 23, 2024.

Yan, X., Chai, G., Han, X., Lei, L., Wang, G., Jia, X., and Zhang, X., 2024. SA-Pmnet: Utilizing close-range photogrammetry combined with image enhancement and self-attention mechanisms for 3D reconstruction of Forests, Remote Sensing, 16(2): 416, DOI:10.3390/rs16020416.

Feasibility Study of Using Low-cost LiDAR to Measure the *Swietenia macrophylla* King Plantation's Characteristics

Che-Yuan Kuo [1] Chun-Hung Wei [2] Jan-Chang Chen [3*]

Abstract

With technological advancements, LiDAR sensors have been integrated into some Smart handheld devices. This study assesses the feasibility of using the iPhone 15 Pro paired with the viDoc RTK Rover for under-canopy positioning and scanning within a *Swietenia macrophylla* plantation. The results indicate that this setup can reduce smartphone positioning error to centimeter-level accuracy, demonstrating its potential to replace traditional RTK GNSS receivers for under-canopy point collection. To obtain valid point cloud data, connection to the viDoc for scanning is necessary, and marking ground control points (GCPs) can improve accuracy. Regarding specific measurements, a 90% overlap rate yielded optimal results for Diameter at Breast Height (DBH) estimation, while a 75% overlap was optimal for determining tree positions. Based on a marginal effect analysis of processing time and various measurement outcomes, the recommended configuration is to connect the viDoc with a 70% overlap rate and mark ground control points. This equipment combination effectively reduces costs while acquiring usable point cloud data, though it remains constrained by scanning range and shows limited detection capability for trees with DBH less than 5 cm.

Keywords: Low-cost LiDAR, Smart Handheld Devices, Point Clouds Analysis, RTK GNSS

[1] Master, Department of Forestry, National Pingtung University of Science and Technology
[2] Associate Professor, Department of Forestry, National Pingtung University of Science and Technology
[3] Professor, Department of Forestry, National Pingtung University of Science and Technology
[*] Corresponding Author, E-mail: janchang17@gmail.com

Received Date: Jan. 02, 2026
Revised Date: Mar. 27, 2026
Accepted Date: Aug. 13, 2026

中文稿件投稿注意事項（英文稿件請參閱下頁 Author Guidelines）

本學刊定期出版攝影測量及遙感探測領域之學術論著、技術短文(按中文及英文摘要、關鍵字、前言、材料及方法、結果、討論、參考文獻等撰寫)及應用性報導。收錄之稿件 皆經本編輯委員會安排之專家學者審查程序，並由通訊作者簽署著作權讓渡書。作者可自行聲 明文稿和所屬類別，本編輯委員會將依此及審查委員意見決定收錄 文稿之類別。本學刊不接受已經刊載過之文稿，於研討會發表過之論文，經適當增補其內容者為例 外。

投稿地址為 jprssubmit@gmail.com，請以附件寄送稿件之 WORD 檔及 PDF 檔各一份。為便利編審作業，請於 PDF 稿件中移除作者資料（如姓名、服務機關、電話、及電子郵件信箱等）。稿件撰寫要點如下：

1. 頁面：採用 A4 規格頁面。
2. 字體：(1) 中文均以**新細明體**橫式打字；(2) 英文均以 **Times New Roman 字體**橫式打字。
3. 單位：內文、圖表及符號中之度量單位皆採用 SI 公制單位。
4. 首頁：首頁應含內容依次為「題目、作者姓名、摘要、關鍵詞、及作者服務機關及職稱（置於頁尾）」，接續文稿內文。
5. 末頁：末頁應含內容與首頁相同，但皆以英文書寫。
6. 題目：論文題目宜簡明，字體大小為 20 粗體，英文題目每字第一位字母以大寫字體打印。
7. 作者：作者姓名列於題目之下方，字體大小為 14 粗體，其服務機關及職稱以 $^{1,2,\cdots}$ 上標編號註記於頁 尾，字體大小為 9。
8. 摘要：摘要字數以 250 字為限，字體大小為 10 粗體。
9. 關鍵詞：列出三至五個關鍵詞，字體大小為 12 粗體。
10. 章節及標題：(1)章節標題置中，編號一律為阿拉伯數字，如 1，2，...，標題字體大小為 18 粗體；(2)小節標題靠左，編號一律為 1.1，1.2，1.1.1，1.1.2……，字體大小為 14 粗體。
11. 內文：內文之長度(含圖表)以不超過 22 頁為原則，各段落第一行應縮排，字體大小為 10。
12. 數學式：所有數學式均須標式號於圓括弧內，數學式內所有符號須於符號第一次出現時，應於數學式下方逐一加以定義。
13. 圖及影像：(1)圖及影像均須數位化置於文稿中適當位置；(2)所有圖及影像下方須附圖流水編號及說明；(3)本刊接受彩色圖及影像，但紙版印刷一律採灰階印刷；(4)結尾毋加「。」及「.」。
14. 表格：(1)表格均須數位化置於文稿中適當位置；(2)所有表格上方須附表流水編號及說明；(3)本刊接受彩色表格，但紙版印刷一律採灰階印刷；(4)結尾毋加「。」及「.」。
15. 參考文獻(中、日文須英譯)：(1)參考文獻應列於文末，文稿中未直接引用者不得列入；(2)中文文獻需加英文翻譯，於結尾加上(in Chinese)，日文書寫同中文文獻，於結尾加上(in Japan)。先排列中、日文參考文獻， 再排列英文參考文獻，其他文字參考文獻列於英文參考文獻之後；(3)中文參考文獻以第一作者 姓名筆劃順序排列，英文參考文獻以第一作者之姓依英文字母順序排列。當作者相同時，依發表年代排列，作者及發表年代完全相同時，依日期在年代後加以 a, b, c 順序。第一作者姓名相同時，先排列單一作者之參考文獻，其次為多作者文獻；(4)參考文獻之格式如下列：
    - 期刊論文：作者姓名，年代。文章標題，期刊名稱，卷數(期數)：起始頁數－結束頁數。
    - 書籍：作者姓名，年代。書名，出版者，出版地。
    - 研討論文集：作者姓名，年代。篇名，論文集名稱，發表地點，卷(期)：起始頁數－結束頁數。
    - 網路引用：作者姓名，年代。篇名，取自網址，檢索日期。
16. 參考文獻引用：(1)作者為一人時，以作者全名(中文文獻)或姓(英文文獻)及發表年代，如楊明德(2007)或(楊明德，2007)、Yang (2007)或(Yang, 2007)；(2)作者為二人時，以二作者全名(中文文獻)或姓(英文文獻)及發表年代，如楊明德及蘇東青(2014) 或(楊明德及蘇東青，2014)、Yang &Su (2014)或(Yang & Su, 2014)；(3)作者為多人時，以第一作者全名(中文文獻)或姓(英文文獻)，之後加「等」(中文文獻)或「*et al.*」(英文文獻)及發表年代，如(楊明德等，2016)或楊明德等(2016)、Yang *et al.* (2016)或(Yang *et al.*, 2016)，「*et al.*」請以斜體；(4)作者及發表年代完全相同時，依日期在年 代後加以 a, b, c 引用，如 Yang & Su (2008a)、Yang & Su (2008b)。

Author Guidelines

The Journal of Photogrammetry and Remote Sensing periodically publishes academic articles, technical short papers, and application reports addressing topics in photogrammetry and remote sensing. Acceptance of submitted manuscripts for publication is based on a peer-reviewing process arranged by the editorial board. A manuscript will not be formally accepted until a copyright agreement is received. Although authors can indicate their submitted manuscripts belonging to one of the categories of academic articles, technical short papers or application reports, the editorial board will make the final decision based on the suggestions of reviewers. This journal accept no paper ever been published, but a revised appropriately paper published in conference proceedings is an exception.

The manuscript submission address is jprssubmit@gmail.com. Please send your manuscript in both WORD and PDF formats via e-mail attachments. However, for blind reviewing, please eliminate author information in the PDF file. Instructions for authors to prepare a manuscript are as follows:

1. Typing format: Manuscripts should be typed single-column and double-space on A4 page size using font-type of 10pt **Times New Roman**. The pages should be numbered.
2. Units: Manuscripts must use SI (Metric) units in text, figures, or tables.
3. Cover page: The cover page should contain manuscript title, author names, abstract, key words, and author affiliations.
4. End page: The end page contains the same information as the cover page, but in Chinese.
5. Title: The title of the manuscripts should be concise and informative. Please use font size of 20pt,Bold and capitalize the initial letter of each word in the title.
6. Authors: Authors names (using font size of 14pt,Bold) are listed below the title, affiliations (using font size of 9pt,Bold) are listed at the end of the first page marked with [1, 2,...].
7. Abstract: The abstract should be shorter than 250 words and typed using font size of 10pt..
8. Keywords: Keywords (no more than 5 words) should be given and placed below the abstract (using font size of 12pt,Bold).
9. Sections and Sub-sections: Section heading (using font size of 18pt,Bold) should be centered and numbered in 1, 2… etc. Sub-sections (using font size of 14pt,Bold) are justified to the left in 1.1, 1.2, 1.1.1, 1.1.2…etc.
10. Paper length: Published papers are limited to 22 pages.
11. Equations: Type formulae and equations clearly in black and numbered in curved brackets. All symbols in equations should be clearly defined when they firstly appear in the manuscripts.
12. Figures and Tables: Figures and tables are numbered and described in short. When papers are accepted, original figures, tables, and photographs are expected for reproduction. Color figures and tables are acceptable in digital version, but they will be printed in gray levels in hard copies.
13. References: Only works cited in the text will be included. References should be cited in the text by giving the last name of the author(s) and the year of publication of the reference, such as Yang and Su (2014) or (Yang and Su, 2014). With three or more names use the form "Yang *et al*. (2016) or (Yang *et al*., 2016)" in the text. When two or more references have the same text identification, please distinguish them by appending "a", "b", etc., to the year of publication. References should include the complete title(s), the number of the volume and the page(s) of the information cited. References should be listed in alphabetical order, according to the last name of the first author, at the end of the paper.

Volume 31 No. 3
September 2026

Journal of Photogrammetry and Remote Sensing

CONTENTS

Regular Papers

A Deep Learning Approach for Spatiotemporal Estimation of $PM_{2.5}$ Concentration in Urban Areas: A Case Study of Greater Taipei Area

(Vincent Wiso, Hsuan Ren, Tang-Huang Lin)..... **137**

Hydrocarbon Estimation in Northern Taiwan Using the LightGBM Regression Model

(Thia Prahesti, Aji Kusumaning Asri, Yu-Ting Zeng, Chih-Da Wu)..... **161**

Leveraging Deep Learning and Data Integration Techniques for Land Surface Change Detection

(Yu-Xin Chen, Hone-Jay Chu) **173**

Feasibility Study of Using Low-cost LiDAR to Measure the *Swietenia macrophylla* King Plantation's Characteristics

(Che-Yuan Kuo, Chun-Hung Wei, Jan-Chang Chen)..... **193**

第 三十一 卷 第 三 期
中 華 民 國 一 一 五 年 九 月

航測及遙測學刊

目錄

學術論著

基於遙測之都會區 $PM_{2.5}$ 濃度估測之深度學習方法：以大臺北地區為例
(魏文森、任玄、林唐煌) **137**

台灣北部地區使用 LightGBM 回歸模型進行碳氫化合物估計
(帕斯蒂、古曼寧、曾于庭、吳治達) **161**

整合深度學習與資料整合技術於地表變遷偵測
(陳聿新、朱宏杰) **173**

應用低成本光達於大葉桃花心木人工林測計之可行性評估研究
(郭哲源、魏浚紘、陳建璋) **193**