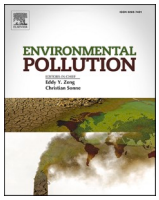
Environmental Pollution 383 (2025) 126879

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol



Assessing 3-D variability of ultrafine particle using a Geo-AI modelling approach: A case study in Zhunan-Miaoli, Taiwan☆

Chia-Wei Hsu [a], Yinq-Rong Chern [b], Jun-Jun Su [a], Pei-Yi Wong [c], Aji Kusumaning Asri [a], Candera Wijaya [d], Yu-Cheng Chen [e], Shih-Chun Candice Lung [c,f], Ta-Chih Hsiao [c,g], Tee-Ann Teo [h], Yi-Liang Shih [i], Chih-Da Wu [a,e,j,k,l,*]

[a] Department of Geomatics, National Cheng Kung University, 70101, Tainan, Taiwan
[b] Department of Electrical Engineering, National Cheng Kung University, 70101, Tainan, Taiwan
[c] Research Center for Environmental Changes, Academia Sinica, 11529, Taipei, Taiwan
[d] Agricultural Engineering Research Center, 32061, Taoyuan, Taiwan
[e] National Institute of Environmental Health Sciences, National Health Research Institutes, Miaoli, 35053, Taiwan
[f] Department of Atmospheric Sciences, National Taiwan University, 10617, Taipei, Taiwan
[g] Graduate Institute of Environmental Engineering, National Taiwan University, 10617, Taipei, Taiwan
[h] Department of Civil Engineering, National Yang Ming Chiao Tung University, 30010, Hsinchu, Taiwan
[i] National Center for High-Performance Computing, 30076, Hsinchu, Taiwan
[j] Innovation and Development Center of Sustainable Agriculture, National Chung Hsing University, 40227, Taichung, Taiwan
[k] Research Center for Precision Environmental Medicine, Kaohsiung Medical University, 80708, Kaohsiung, Taiwan
[l] Chronic Diseases and Health Promotion Research Center, Chang Gung University of Science and Technology, 61363, Chiayi, Taiwan

ARTICLE INFO

Keywords:
3-D distribution
Air pollution
Geo-AI
Ultrafine particulate matter ($PM_{0.1}$)
Unmanned aerial vehicle (UAV)

ABSTRACT

Previous air pollution modeling studies have predominantly emphasized horizontal distributions, overlooking the critical vertical variability of pollutant concentrations in urban environments. Therefore, the three-dimensional (3-D) behavior of air pollutants, and of ultrafine particulate matter ($PM_{0.1}$) in particular, is insufficiently characterized. This study examined the 3-D distribution of $PM_{0.1}$ in the Zhunan and Toufen regions in Miaoli, Taiwan. Using a hexacopter drone, $PM_{0.1}$ concentrations were measured at 12 locations, at altitudes of 40, 60, and 100 m. A geospatial-artificial intelligence (Geo-AI) model was developed to estimate 3-D $PM_{0.1}$ concentrations, incorporating databases such as land use, meteorology, and 3-D building data as predictor variables. SHapley Additive exPlanations (SHAP) analysis for variable selection showed that key predictors were building height, temperature, carbon dioxide, nitric oxide, forest coverage and 3-D spatial distance from buildings. Five machine learning algorithms were used for modeling. Extreme Gradient Boosting Regressor (XGBR) achieved the best performance with a training R^2 of 0.98. The model's robustness was further examined through 10-fold cross-validation and stratified validation, which yielded R^2 values exceeding 0.85, indicating a strong ability to capture the spatial variation of $PM_{0.1}$ across different environmental conditions. These findings underscored the crucial role of vertical pollutant dispersion in urban environments and the need to incorporate detailed 3-D measurements into urban planning and public health policies.

1. Introduction

Ultrafine particulate matter ($PM_{0.1}$) refers to particles with an aerodynamic diameter of less than 0.1 μm. This nanoscale size enables $PM_{0.1}$ to penetrate deeply into the respiratory tract and reach the bloodstream, which has a significant effect on health (Bhargava et al., 2018; Leikauf et al., 2020). Exposure to $PM_{0.1}$ is associated with chronic respiratory and systemic ailments, including chronic obstructive pulmonary disease, atherosclerosis, and neurodegenerative conditions in neuronal cells (Calderón-Garcidueñas et al., 2025; Jiang et al., 2024; Zhang et al., 2024). $PM_{0.1}$-induced pathogenesis involves multiple mechanisms, including inflammation, oxidative stress, autophagic cell

☆ This paper has been recommended for acceptance by Admir Créso Targino.

* Corresponding author. Department of Geomatics, National Cheng Kung University, 70101, Tainan, Taiwan.
E-mail address: chidawu@mail.ncku.edu.tw (C.-D. Wu).

https://doi.org/10.1016/j.envpol.2025.126879
Received 9 April 2025; Received in revised form 21 July 2025; Accepted 22 July 2025
Available online 23 July 2025

death and dysregulation of proteins that are essential for cellular homeostasis and viability (Hammond et al., 2022; Kelley et al., 2024). $PM_{0.1}$ is not a currently targeted pollutant that is monitored by environmental authorities in most countries, so there is a lack of information as to its distribution and effect on health. Monitoring $PM_{0.1}$ is crucial for maintaining public health and preventing chronic diseases and systemic damage.

Numerous studies have demonstrated the spatial heterogeneity of air pollution, with many assessments focusing on the horizontal distribution of air pollutant concentrations (Asri et al., 2024; Lee et al., 2017; Wu et al., 2017), although vertical variations have also been documented (Li et al., 2022; Pohorsky et al., 2025). In densely populated urban areas where people live and work in high-rise buildings, individuals at different heights within the same area may experience significantly different pollution levels. The vertical variation in air pollution is primarily affected by wind patterns that are generated by street canyons and by changes in wind speed and direction, which significantly affect the three-dimensional (3-D) dispersion of pollutants (Huang et al., 2019). An increase in wind speed enhances the dispersion of pollutants, accelerates their movement in residential areas, and reduces their accumulation. Low wind speeds result in pollutant deposition and accumulation. When a building is aligned parallel to the prevailing wind direction, pollutant penetration into residential areas is reduced. In contrast, buildings with other orientations tend to promote pollutant accumulation (Li et al., 2019). Therefore, vertical variability, in addition to two-dimensional (2-D) patterns, is essential for accurately assessing urban air quality.

To study the variation in air pollution in the vertical direction, previous studies used different measurement strategies (Cichowicz & Dobrzanski, 2021; Eeftens et al., 2019; Xu et al., 2022). These methods typically use interpolation techniques or establish the ratio of concentration between different altitudes and ground-level measurements to predict pollutant concentrations at different elevations. However, these studies assumed uniform vertical pollution gradients across spatial domains and did not consider that different land use patterns and building morphologies generate differential wind fields, which produce site-specific vertical concentration profiles.

Air pollution monitoring typically relies on both stationary monitoring networks and portable sensing technologies (Idrees & Zheng, 2020; Tessum et al., 2021). While stationary monitoring stations offer standardized and quality-assured measurements, their limited spatial coverage and fixed locations constrain their ability to observe variations in pollution levels across diverse spatial environments. Portable devices are cheaper and offer enhanced spatial resolution of air pollution measurements, but they are limited in terms of locations and transportation modes, which restrict monitoring to a 2-D plane. Some studies used an unmanned aerial vehicle (UAV) platform to measure air pollution (Lugassi et al., 2022; Pochwała et al., 2020). UAV is a highly mobile and flexible platform that can reach areas inaccessible to humans, such as high altitudes, industrial sites, or disaster areas. UAV provides comprehensive 3-D measurements.

Despite advances in monitoring techniques, the inherently point-based observations cannot adequately characterize the spatial gradients of air pollution for a large area, motivating the development of spatial modeling approaches. Land-use regression (LUR) models and machine learning algorithms have been widely used in air pollution studies (Li et al., 2023; Lung et al., 2021; Xu et al., 2018). LUR models utilize various land-use variables that influence air pollution concentration to develop a linear regression model for measuring air pollution levels. However, LUR models assume a linear relationship between land-use variables and air pollution levels (Azmi et al., 2023; Wu et al., 2017), so the complex, nonlinear interactions in real-world environments are not considered. Machine learning algorithms are utilized in conjunction with LUR models to develop Geospatial Artificial Intelligence (Geo-AI), which significantly enhances the predictive capabilities of the models (Asri et al., 2024; Babaan et al., 2024; Wong et al., 2021). Unlike conventional linear regression, machine learning techniques better capture nonlinear relationships between variables and air pollution concentrations, providing more accurate predictions.

This study used a UAV system for $PM_{0.1}$ measurements in 3-D space. A Geo-AI spatial model was combined with land-use/land-cover information to estimate the distribution of $PM_{0.1}$ concentration in the vertical and horizontal planes. The measured vertical variation in air quality enabled a more accurate estimation of $PM_{0.1}$ concentration in 3-D space, informing control strategies specific to different heights.

2. Materials and methods

2.1. Study area

This research was conducted in Zhunan Town and Toufen City in Miaoli County, Taiwan. The Zhunan-Toufen area covers approximately 90.86 km^2 and has a population of 195,123 people (Department of Household Registration, 2023; National Land Surveying and Mapping Center, 2023). Zhunan and Toufen have a variety of industrial plants in large industrial areas, such as the Zhunan Industrial Zone, Toufen Industrial Zone, and the Hsinchu Science Park Zhunan Base, which serves the electronics, semiconductor, and chemical industries.

The diverse topography and building morphology in Zhunan and Toufen create complex airflow patterns that influence the diffusion of air pollutants. In these areas, the combination of high population density and high-rise buildings necessitates detailed 3-D exposure measurements to assess particulate pollutant distributions at various elevations and locations. The study area and the distribution of the sampling locations are shown in Fig. 1.

2.2. Database

Five datasets were used to determine the predictor variables. These were selected to ensure a comprehensive analysis of the factors that affect air pollution in 3-D space. The database and its contents are shown in Table S1.

To process variables, the spatial density of each land-use feature within several buffer zones was calculated, for values from 25 to 2000 m. Pollutant and meteorological data were spatially interpolated using the Kriging method with an hourly resolution, the finest available temporal resolution. Values for each variable for each sampling location represented the surrounding environmental conditions during sampling. For 3-D building data, the volume of buildings within buffer zones ranging from 25 to 2000 m and the distances to different types of structures in both 2-D and 3-D space were calculated to capture terrain and environmental characteristics around each sampling location.

2.2.1. Ultrafine particulate matter database

P-Trak Ultrafine Particle Counter (model 8525, TSI Incorporated, USA) is a condensation particle counter that is used to measure the concentration of particles within a diameter of 0.02–1 μm. The maximum detection limit is 500,000 particles/cm^3, and P-Trak has demonstrated bias-corrected precision within 10 %, making it suitable for both indoor and outdoor exposure studies (Chen et al., 2017; Lin et al., 2022). The sampling interval for the P-Trak ranged from 1 s to 1 min. In this study, a 1 s interval was used to capture each airborne sample. P-Trak measures vehicle emissions, industrial pollution, vehicle exhaust migration, and smoking area assessments (Baquero et al., 2015).

3-D measurements were taken using a six-rotor hexacopter (model A610 SE, Asia UAV AI Innovation Application R&D Center, Taiwan). This drone was modified to carry air pollution monitoring equipment and was equipped with a 24.68 kg battery with a capacity of 32,000 mAh. This configuration allowed a flight time of more than 20 min. For this study, the position of air pollution sampling instruments on the UAV was selected to minimize the effect of airflow over the UAV. Previous studies demonstrated that placing sensors directly beneath rotors

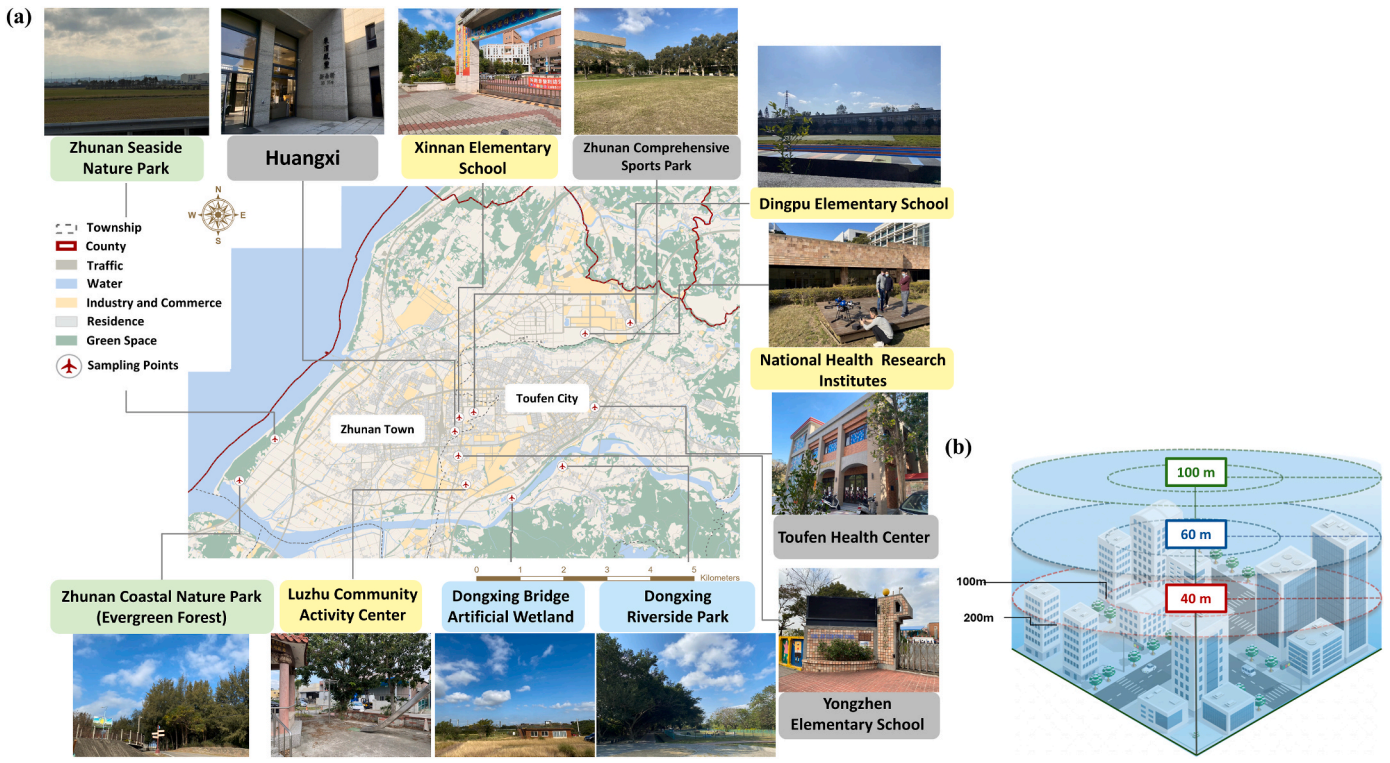


Fig. 1. (a) Study area sampling locations, with colors representing different land uses: green for forests, yellow for industrial and commercial areas, gray for transportation infrastructure, and blue for water bodies; (b) UAV sampling path. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

compromised measurement accuracy because rotor downwash induces mixing and dispersion of particulate matter (Han et al., 2024; Roldán et al., 2015; Villa et al., 2016). To mitigate this aerodynamic interference, the P-Trak was mounted on top of the six-rotor A610 SE hexacopter to minimize exposure to these flow disturbances (Fig. S1). This configuration increased measurement precision and minimized systematic bias (Han et al., 2024; Roldán et al., 2015).

Prior to each flight, zero-point calibration was performed, and the instrument's results were compared with those of a ground-based device to ensure data consistency and minimize instrumental bias. To minimize flight-induced variability, UAV trajectory, velocity and altitude were determined using ground-based mission planning software and were controlled autonomously by the onboard flight management system.

The study area was divided into four major land-use types. Each type was allocated 2–3 sampling locations, resulting in a total of 12 sampling points in the entire area, to ensure a balanced representation (Fig. 1). $PM_{0.1}$ measurements were taken at altitudes from the ground to 40, 60, and 100 m above sea level. The highest key altitude is governed by Taiwanese regulations that limit the maximum altitude of drones to 121.92 m. The other two key altitudes are evenly distributed between 0 and 100 m. The key altitudes for some sampling locations were adjusted to account for the presence of obstacles. At each of these three key altitudes, the drone also performed circular flight trajectories with radii of 100 and 200 m to collect planar samples. The flight paths are shown in Fig. 1.

The sample collection was conducted over two periods to determine the seasonal variation in $PM_{0.1}$. The first round of winter samples was collected over two periods, from January 12 to 17 and on February 8, 2023. These concentrations were found to be higher than those observed during summer in Taiwan (Hsu & Cheng, 2019). The second stage of the sampling was conducted in the summer, from July 15 to 18, 2023. Concentrations were measured to be lower during this period.

Sampling was conducted at 2–3 points per day for 1–1.5 h, which included setting up and measurement. The time resolution for the sampling data was one record per second. After filtering for erroneous sampling data and missing data (approximately 4.56 % of the entire dataset), outliers were excluded using the 3 × Interquartile Range (IQR) rule (values below Q1 − 3 × IQR or above Q3 + 3 × IQR). The P-Trak measurements demonstrated a high degree of variation, so this criterion helped avoid misclassifying normal values as outliers. A total of 35,523 records were used for analysis. This sampling data was the dependent variable for the machine learning model.

2.2.2. *Land-use inventory and road network distribution*

The land-use inventory data for this study were obtained from the Land Surveying and Mapping Center (LSMC) of the Ministry of the Interior. The dataset categorized land use into nine major types, 48 minor types, and 93 detailed types. This study primarily focused on residential areas, green spaces, agricultural land, industrial zones, artificial structures, and water bodies to identify the predominant land-use type at each sampling location and to assess the effect of land use on $PM_{0.1}$ distribution during the Geo-AI modeling process.

The digital road network map for this study was released by the Transportation Science Research Institute of the Ministry of Transportation in Taiwan. This database is updated annually and details the distribution of roads across the country in GIS format. After density calculations, the data were incorporated into the Geo-AI model to determine the relationship between road networks and $PM_{0.1}$ distribution.

2.2.3. *3-D building information*

The distribution of buildings affects the vertical variation in air pollution. Previous studies have demonstrated that building volume and height can impact the dispersion and distribution of nearby pollutants (Jiang et al., 2021; Zhu et al., 2021). The 3-D building models used in this study were obtained from the Ministry of the Interior's LSMC, which has maintained and updated them in Taiwan since 2019. These models range from coarse to detailed levels, which are classified as Level of

Detail 1 (LOD1), Level of Detail 2 (LOD2), Level of Detail 3 (LOD3), and Level of Detail 4 (LOD4). This study used LOD1-scale 3-D building models. These 3-D building models were integrated with land-use inventory data to quantify the volume and height of residential, commercial, agricultural, and industrial buildings. The distance from each sampling location to the nearest building was also measured to support model development.

2.2.4. Other air pollutants and meteorological database

Studies have shown a strong correlation between surface-level pollutants and their vertical distributions (Chen et al., 2022; Chen et al., 2024). Most 3-D air pollution samples for this study were collected at altitudes below 100 m, making surface-level concentrations reasonable proxies for the vertical variation in air pollution. With direct measurements at different heights often unavailable, ground-level gaseous pollutant data were used to approximate conditions affecting vertical pollutant dispersion. The ground-based air quality monitoring data for this study were from the Air Quality Monitoring Network of Taiwan's Ministry of Environment, which includes data from 76 monitoring stations. The measured pollutants included carbon monoxide (CO), nitric oxide (NO), nitrogen dioxide (NO_2), nitrogen oxides (NO_x), ozone (O_3), particulate matter with a diameter of 10 μm or less (PM_{10}), $PM_{2.5,}$ and sulfur dioxide (SO_2).

Meteorological data, including temperature, humidity, wind speed, wind direction, and solar radiation, were obtained from the Central Weather Administration, based on observations from approximately 453 weather stations across Taiwan. These data served as a reference for real-time environmental conditions.

2.2.5. Landmark database

This study utilized data from Taiwan's Nationwide Business Tax Registry to identify the geographic locations of various merchants and enterprises that are potential sources of $PM_{0.1}$ emissions. This database covers a wide range of Point of Interest (POI) categories, including industrial facilities (general industry, petrochemical plants, paper and textile mills, and wood and stone processing sites), agricultural and food-related businesses (cereal production facilities, food stalls, and garden centers), infrastructure and service sites (transportation hubs and funeral services), and emission-related sources (volatile substance manufacturers, waste treatment facilities, power plants, incinerators, and smokestacks).

2.3. Methodologies

This study developed a 3-D air pollution estimation model using Geo-AI modeling techniques. The process included important variable selection, Geo-AI modeling, validating model performance using Automated Machine Learning (AutoML), and visualizing $PM_{0.1}$ distribution using advanced 3-D estimation techniques. Spatial data were processed using ArcGIS Pro 3.2, and data analysis and machine learning were performed in Python 3.7 on the Jupyter Notebook platform. The study framework is shown in Fig. 2.

2.3.1. Descriptive statistics for $PM_{0.1}$ pollutants that are measured using a UAV

The spatial distribution and vertical variability of UAV-collected $PM_{0.1}$ concentrations were evaluated using descriptive statistics. Instrument-flagged invalid readings were filtered out to ensure accuracy and reliability. $PM_{0.1}$ concentration values were then summarized using the mean, median, standard deviation, minimum value, and maximum value. The data were divided based on two dimensions: the sampling seasons, S1 (cold season) and S2 (warm season), and the land-use types: forests, industrial and commercial areas, transportation infrastructure, and water bodies. Based on these groups, line charts were used to visualize the relationship between $PM_{0.1}$ concentration and altitude at various sampling locations. This method provided an initial assessment of vertical concentration changes across different locations. The results identified areas and elevations where potential contaminants tended to accumulate or disperse, providing key patterns for spatial modeling.

2.3.2. Determination of the importance of variables using SHAP

SHapley Additive exPlanations (SHAP) is an advanced interpretive machine learning technique that explains the variable importance of a model (Asri et al., 2024). SHAP analysis was used to quantify the contribution of each predictor variable to the prediction results and to determine the effect of each predictor variable on the prediction results for the model.

This study utilized SHAP values to rank variables by their impact on $PM_{0.1}$ concentrations (Chang et al., 2025), from the highest to the lowest. Variables were then added sequentially to the machine learning models based on this ranking. For each added variable, the change in the model's R^2 value was calculated. If the inclusion of a variable increased R^2 by 0.001 or more, it would be retained in the model (Babaan et al., 2024). This stepwise approach ensured that each added variable contributed meaningfully to the accuracy of $PM_{0.1}$ predictions.

2.3.3. Geo-AI model development

Five commonly used machine learning algorithms for air pollution estimation were employed to develop the model. These included a Categorical Boosting Regressor (CBR), a Gradient Boosting Regressor (GBR), a Light Gradient Boosting Machine Regressor (LGBMR), a Random Forest Regressor (RFR), and an Extreme Gradient Boosting

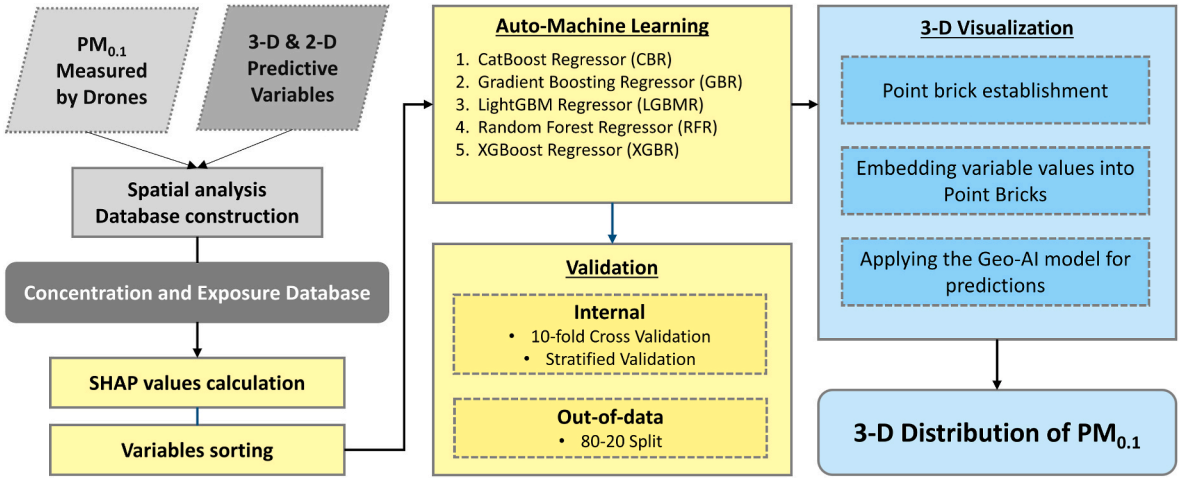


Fig. 2. Study framework, including database development, Geo-AI model development and validation, and 3-D visualization of air pollution.

Regressor (XGBR). CBR processes categorical features, which is critical when data types vary, making it suitable for mixed datasets (Dorogush et al., 2018). GBR is an ensemble learning method that uses a series of progressive weak learners, typically in the form of decision trees, to progressively improve the model's learning performance at each iteration (Friedman, 2001). LGBMR employs a leaf-wise tree-growing strategy that is both flexible and scalable, utilizing a histogram-based approach (Ke et al., 2017). RFR is a regression technique that is reliable and prevents overfitting (Breiman, 2001). It builds multiple decision trees during the training process to process scattered data. XGBR processes various data using regularization methods to prevent overfitting and provides advanced functions, such as parallel tree construction and approximate tree pruning (Chen & Guestrin, 2016).

This study determined the importance and order of variables, then incrementally added them to the model, using the screening method described in Section 2.3. The AutoML suite was used as a modeling tool, and the machine learning algorithms previously mentioned were used to model the nonlinear relationship between variables and air pollution concentration (He et al., 2021). Model performance was evaluated using several indicators, including R^2, adjusted R^2, RMSE, MSE, and MAE, to select the best-performing algorithm as the final model.

2.3.4. Model evaluation

The model evaluation process for this study included both internal and out-of-sample validations to ensure the reliability and robustness of the predictive model. Internal validation adopted the 10-fold cross-validation (10-fold CV) method, which divided the entire dataset into ten equally sized subsets. The model was trained on nine subsets and tested on the remaining one. This iterative process was repeated ten times, so that each subset served as test data once. This iterative validation assessed the model's stability across different portions of the dataset.

A stratified validation was also conducted to evaluate the model's performance across different environmental characteristics groups. These categories included elevation, day of the week, median $PM_{0.1}$ concentration, land-use type, administrative town, and proximity to major roads. Each subgroup was used in turn as the test dataset to verify the model's predictive ability under varying conditions and to assess its robustness across different environmental contexts.

Out-of-sample validation was performed using an 80-20 split, where the entire dataset was randomly divided into a training set (28,419 samples) and a testing set (7104 samples). The training set, which accounted for 80 % of the total data, was used for model training, while the testing set (20 %) was used to evaluate the model's performance on unseen data. Since the testing data were not exposed during model training, it served as an indicator of the model's generalization capability.

2.3.5. 3-D visualization of $PM_{0.1}$ prediction

To visualize the 3-D distribution of $PM_{0.1}$, the advanced 3-D mapping capabilities of ArcGIS Pro were utilized. 3-D point bricks were generated to cover the entire mapping domain. For each point within these bricks, the variables selected through SHAP analysis were re-applied in the trained model to generate a localized prediction. Temporal variables, such as temperature and nitric oxide (NO), were incorporated using their corresponding values at the target prediction time to reflect the actual environmental conditions at that moment. This ensured that the predicted concentrations reflected the actual environmental conditions and minimized discrepancies caused by temporal variability, thereby improving prediction reliability. The "Empirical Bayesian Kriging 3D" tool in ArcGIS Pro was then employed to perform a 3-D interpolation of the predicted $PM_{0.1}$ concentration across the study area. The resulting map presented a semi-transparent cubic concentration surface, with red denoting higher concentrations and blue denoting lower concentrations. The overall 3-D visualization workflow is illustrated in Fig. S2.

3. Results

3.1. UAV $PM_{0.1}$ measurements

The mean concentration (mean ± standard deviation) of $PM_{0.1}$ was 6025 ± 4970 particles/cm^3, ranging from 768 to 15,100 particles/cm^3. Fig. 3 shows the study area and the vertical variation of measured $PM_{0.1}$ concentration for Season S1/S2 and four land-use categories. The concentration was greater during the cooler season (8534 ± 5654 particles/cm^3, Season S1) than the warmer season (3322 ± 1667 particles/cm^3, Season S2). In terms of land-use groups, the overall $PM_{0.1}$ concentration in forests was slightly lower than in other categories. As altitude increased, the observed pollutant level for each category varied, and there was a non-linear relationship between $PM_{0.1}$ and height.

3.2. Geo-AI performance

This study used SHAP values to produce a comprehensive variable selection and ranking order to determine the contribution of each predictor to the 3-D $PM_{0.1}$ concentration model. The SHAP values for each selected variable are shown in Fig. 4. The colored bars represent the eigenvalues for the predictor variables: red for higher eigenvalues and blue for lower values. Variables such as Carbon Monoxide, Public Recreational Space, and temperature were the most influential predictors. Variables that pertain to height and buildings, such as the distance between the sampling location and nearby structures, were critical for the model and demonstrated the multifaceted nature of urban air pollution dynamics.

To estimate the 3-D concentration of $PM_{0.1}$, five machine learning algorithms were applied. As shown in Table 1, the training R^2 values for these models ranged from 0.92 to 0.98, and the RMSE values varied between 583 and 1398 particles/cm^3. The model performance and overfitting rate were confirmed using training R^2, testing R^2, and 10-fold CV R^2. "Training R^2-Testing R^2" and "Training R^2-10-Fold CV R^2" comparisons were used to determine the model's consistency. Of the five algorithms, XGBR achieved the highest R^2 values for both training (0.98) and testing (0.88), and the lowest MSE (3.40×10^5) and RMSE (583 particles/cm^3) values during training. It produced minimal prediction errors and allowed robust generalization. This stability across metrics showed that XGBR was the most reliable and robust algorithm.

We validated the performance of the XGBR model using out-of-data validation, 10-fold CV, and stratified validation. The results for validation are shown in Fig. S3. The testing R^2 value for the XGBR was 0.88, and the RMSE was 1744 particles/cm^3. The 10-fold CV had an R^2 value of 0.83. The validation of the model across different data subgroups is shown in Fig. S3, for which the error metrics (MSE, RMSE, MAE) are normalized to a 0–1 scale to allow consistent comparison. The R^2 values for XGBR for these subgroups were all more than 0.85, and the actual RMSE values ranged from 262 to 2757 particles/cm^3, which were small errors. These results showed that the Geo-AI model didn't exhibit significant overfitting and accurately predicted 3-D air pollution variations for diverse spatial contexts.

3.3. 3-D visualizations of $PM_{0.1}$

The results for January 12, 2023, at 8:00 a.m. and 6:00 p.m., and July 15, 2023, at 8:00 a.m. and 6:00 p.m. were used for 3-D $PM_{0.1}$ estimation and visualization. On January 12, the mean temperature in Zhunan was 17.5 °C, the mean wind speed was 0.8 m/s, and the mean relative humidity was 95 % (Central Weather Administration, 2023). On July 15, the mean temperature was significantly higher at 30.2 °C, the mean wind speed was 1.2 m/s, and the relative humidity was 80 % (Central Weather Administration, 2023). The 2023 annual means for Zhunan (22.4 °C for temperature, 1.8 m/s for wind speed, and 80 % for relative humidity) showed that the weather conditions for January 12 reflected cooler-season conditions, and those for July 15 were

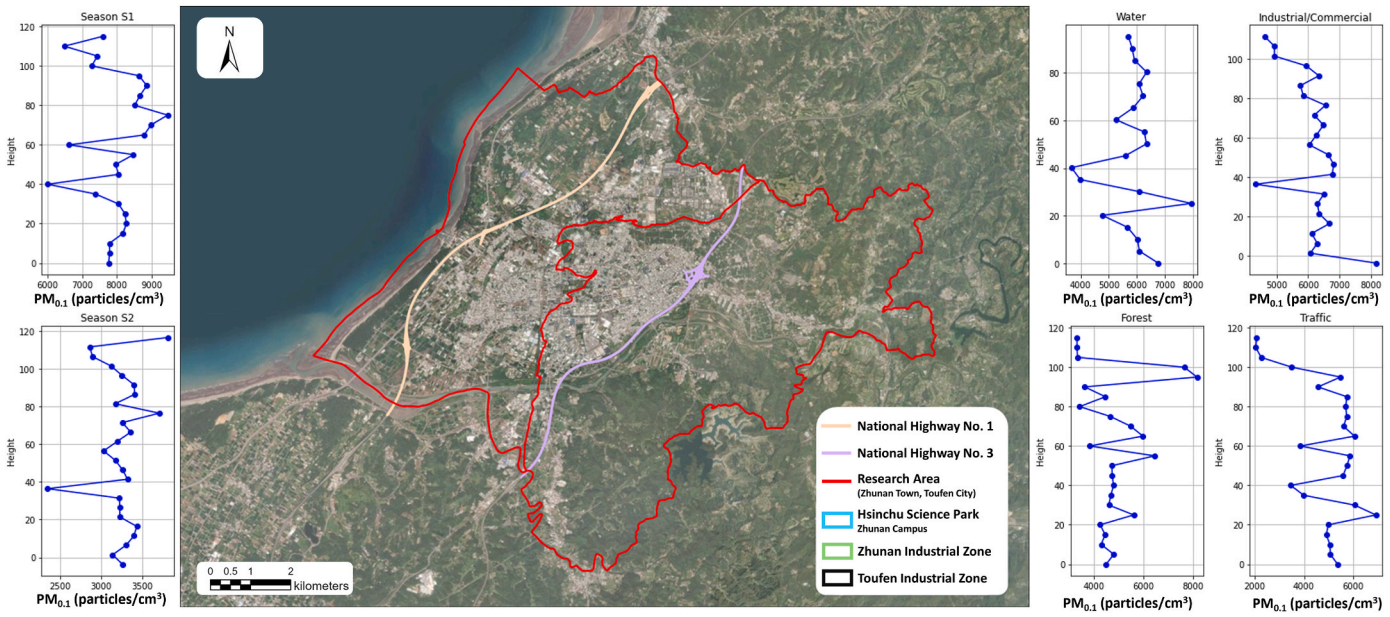


Fig. 3. Study area and vertical distribution of concentration measured using a UAV for each land-use category.

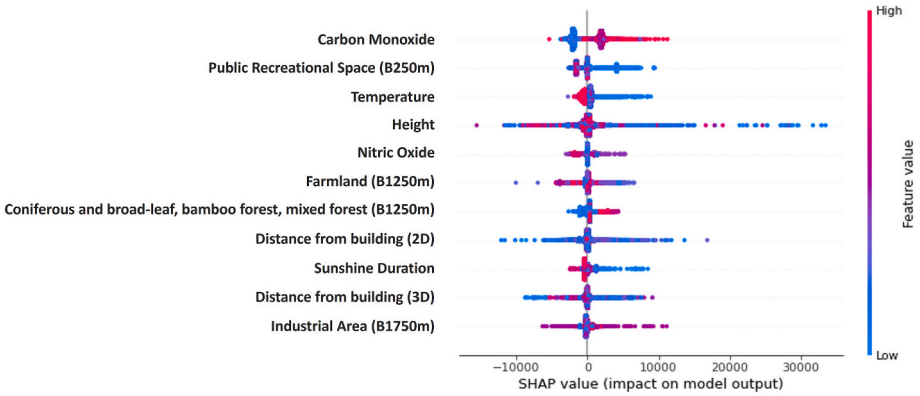


Fig. 4. Important variables selected using SHAP values.

characteristic of the warmer season (Central Weather Administration, 2023). These results allowed a comparison between seasonal conditions and diurnal variations. The selected locations for visualization were B3 (Luzhu Community Sports Center) and W0 (Dong Xing Bridge Constructed Wetland), representing residential communities and rural natural areas, respectively. The results are shown in Fig. 5.

In B3, which is primarily composed of residential areas, there was a high concentration of $PM_{0.1}$ on the northern side. The concentration gradually decreased with altitude and toward the south. The maximum concentration occurred in areas with dense buildings. Wetland areas (W0) exhibited a lower $PM_{0.1}$ concentration and better air quality. This result demonstrated clear differences in $PM_{0.1}$ concentrations across different land-use types. The pollution distribution for residential communities exhibited larger spatial and vertical changes than the distribution for rural natural areas.

The geographical relationship between pollutant levels and the surrounding land uses at both locations showed that there was a lower pollution concentration closer to rivers and forests (Fig. 5). Close to buildings and industrial zones, the concentration of pollutants increased, especially around waste disposal sites and factories. In terms of seasonal trends, the predicted $PM_{0.1}$ values were higher during the cooler season (January 12) than during the warmer season (July 15). High concentrations of $PM_{0.1}$ were more likely to accumulate at lower altitudes (below 120 m) during winter than during summer, and this phenomenon was particularly evident in W0. Pollution concentrations also exhibited diurnal variations. In both areas, $PM_{0.1}$ concentrations at 6 PM were higher than those at 8 AM This comparison highlighted the influence of land-use types on the diurnal variation of pollution.

4. Discussion

This study presented an innovative air pollution monitoring system that addressed the limitations of existing monitoring stations, which primarily provided point-specific, horizontal observations and no vertical profile information. Traditional LUR methods produced 2-D estimates (Cowie et al., 2019; Wu et al., 2017), which were insufficient for residents living in high-rise buildings in metropolitan areas. By integrating UAV-based 3-D air pollution monitoring technology with 3D-GIS, this study enhanced the conventional 2-D LUR methodology to produce a 3-D approach. Using machine learning algorithms, this study developed a Geo-AI-based 3-D air pollution estimation model to provide crucial data with high accuracy on air pollutant concentration estimates at different vertical heights and a more detailed perspective of urban air quality to improve public health and benefit urban planning.

Few studies have provided a 3-D estimation of air pollution distribution. Methods such as ground-based vehicles, drones, and instruments

Table 1
Model performance and ranking for five algorithms: CBR, GBR, LGBMR, RFR and XGBR.

Algorithm	Validation	R^2	Adjusted R^2	MSE	RMSE	MAE	Training-Testing / Training-10-Fold	Ranking					
								Training< R^2	Testing< R^2	10-fold CV< R^2	Training< -Testing	Training< −10-Fold	Total
CBR	Training	0.95	0.95	1.27×10^6	1126	248		4	2	3	2	2	13
	Testing	0.87	0.87	3.14×10^6	1772	276	0.08						
	10-Fold CV	0.82	0.82	4.40×10^6	2098	281	0.13						
GBR	Training	0.97	0.97	6.51×10^6	807	22		2	3	2	3	4	14
	Testing	0.87	0.87	3.17×10^6	1781	132	0.1						
	10-Fold CV	0.83	0.83	4.25×10^6	2062	155	0.15						
LGBMR	Training	0.95	0.95	1.17×10^6	1080	175		3	5	4	5	3	20
	Testing	0.84	0.84	3.89×10^6	1973	219	0.11						
	10-Fold CV	0.81	0.81	4.60×10^6	2144	222	0.14						
RFR	Training	0.92	0.92	1.95×10^6	1398	111		5	4	5	1	1	16
	Testing	0.85	0.85	3.71×10^6	1926	171	0.07						
	10-Fold CV	0.81	0.81	4.64×10^6	2153	195	0.11						
XGBR	Training	0.98	0.98	3.40×10^6	583	69		1	1	1	4	5	12
	Testing	0.88	0.88	3.04×10^6	1744	145	0.11						
	10-Fold CV	0.83	0.83	4.21×10^6	2051	165	0.16						

that are mounted on buildings were used to estimate 3-D air pollution (Cichowicz & Dobrzanski, 2021; Eeftens et al., 2019; Xu et al., 2022). These studies had limitations. Cichowicz et al. (2021) used drones for PM_{10} vertical measurements, but the analysis relied solely on spatial interpolation and did not consider the effects of surrounding land uses. Xu et al. (2022) performed a 3-D exposure assessment using $PM_{2.5}$ data collected from sensors deployed at multiple floor levels. A Generalized Additive Model (GAM) was employed to quantify the relationship between air pollution and ground elevation, while a 2-D Land Use Regression (LUR) model was integrated to capture spatial variability, thereby enabling a comprehensive 3-D exposure evaluation. Eeftens et al. (2019) used a decay constant 'k' for the LUR model to represent the exponential decrease in NO_2 concentration with height. However, both studies characterized the relationship between air pollution and height using a statistical modeling approach and did not examine the vertical variation in air pollution in terms of land use.

Compared to previous studies, our research introduced several key advances. Firstly, we focused on simulating the distribution of $PM_{0.1}$, which was crucial due to its smaller size and ability to penetrate deep into the respiratory system, posing significant health risks. Then, our study incorporated 3-D building information such as 3-D distance from buildings and height. These 3-D indicators were also selected as important variables through SHAP value calculation. This approach allowed for a more precise consideration of how ground facilities affect air pollution in both horizontal and vertical dimensions. This study used Geo-AI models in horizontal and vertical space to fit the relationship between air pollution and emission source factors, in order to more accurately measure non-linear variations in air pollution with elevation than previous studies, which primarily used an exponential statistical model (Eeftens et al., 2019). This study gave a more nuanced understanding of 3-D air pollution dynamics and significantly more accurate results than previous research methodologies. The Geo-AI model for this study achieved a significantly higher R^2 value of 0.98. The 3-D LUR model with GAM that was designed by Xu et al. (2022) achieved an adjusted R^2 value of 0.91, and the LUR model with an exponential statistical model that was developed by Eeftens et al. (2019) achieved an R^2 value of 0.84. This study achieved improved results.

Using the variables selected using SHAP values, this study identified factors that significantly impact $PM_{0.1}$ distribution in 3-D space. Building height, temperature, carbon dioxide, nitric oxide, forest coverage, and 3-D spatial distance from buildings were included. Building-related factors (building height and 3-D spatial distance from buildings) were important for the 3-D spatial model because these had a significant impact on the transmission and diffusion mechanism for air pollution. Yang et al. (2020) showed that the 3-D shape of urban space significantly affects the flow and accumulation of pollutants. Building-related factors, such as height and the degree of enclosure, are particularly important for pollutant retention. Buildings can act as barriers to pollutant dispersion, leading to the accumulation of pollutants around their structures. Moreover, by altering the local wind field, buildings can influence both the propagation direction and the spatial distribution of pollutant concentrations. In a high-density environment, the arrangement and height of buildings create wind speed deceleration areas and vortex areas, which trap pollutants in specific areas and further deteriorate air quality (Kim et al., 2022; Yang et al., 2020). This study used the height and distance of buildings in 3-D space as important factors so that the Geo-AI model could better capture the complex relationship between buildings and air pollution.

In terms of temperature, an important variable, this study showed that the predicted $PM_{0.1}$ values had different distributions at different temperatures, particularly in winter and summer. High-concentration $PM_{0.1}$ pollutants also exhibited a broader vertical distribution at less

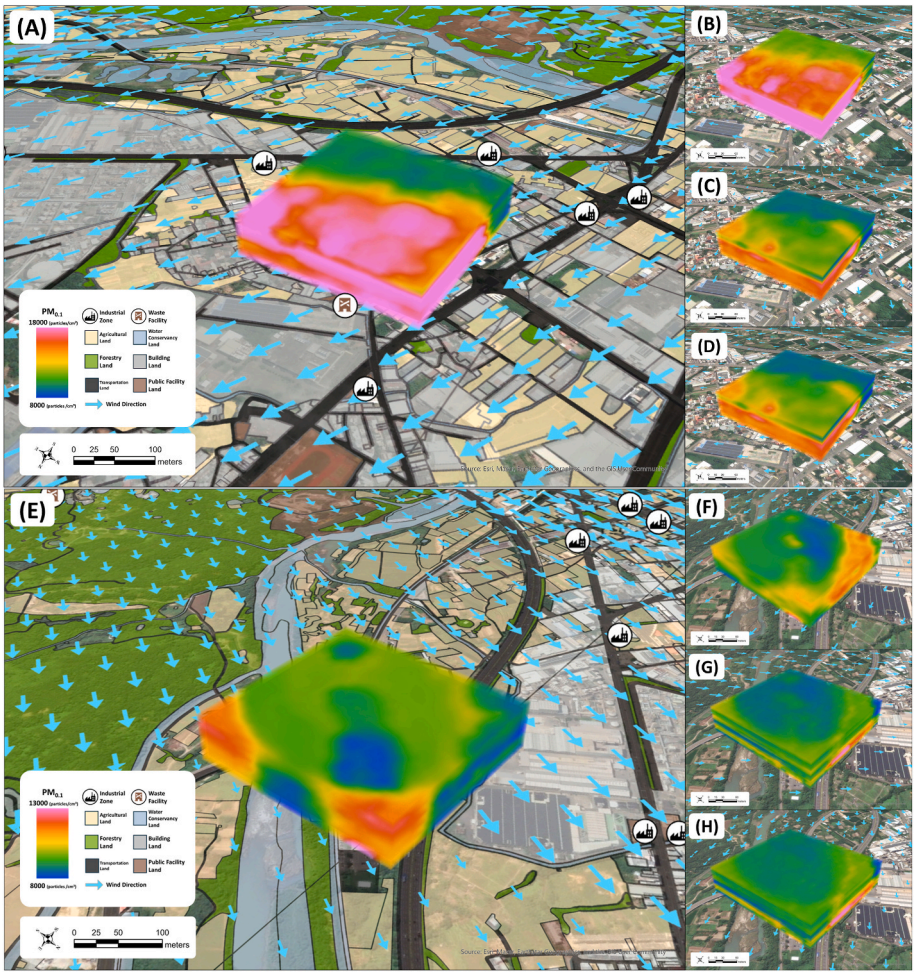


Fig. 5. 3-D visualizations of $PM_{0.1}$ concentrations at B3 (Luzhu Community Sports Center) and W0 (Dong Xing Bridge Constructed Wetland) during different seasons and times of day: (A) B3 on a winter morning (2023/1/12 08:00), (B) B3 on a winter evening (2023/1/12 18:00), (C) B3 on a summer morning (2023/7/15 08:00), (D) B3 on a summer evening (2023/7/15 18:00), (E) W0 on a winter morning (2023/1/12 08:00), (F) W0 on a winter evening (2023/1/12 18:00), (G) W0 on a summer morning (2023/7/15 08:00) and (H) W0 on a summer evening (2023/7/15 18:00).

than 120 m in winter, particularly at the measuring point W0, primarily a wetland environment. This phenomenon may be related to dispersion and Mixing Layer Height (MLH) for air pollution. In winter, the MLH is lower than in summer, so pollutants accumulate more at lower altitudes (Duc et al., 2022; Salvador et al., 2020). The restricted vertical MLH during winter increased the concentration of pollutants near the surface, which created air quality issues in these regions.

The combustion of fossil fuels in vehicles and industrial processes releases CO_2 and NO_2, which directly contribute to the formation of fine particulate matter (Awaworyi Churchill et al., 2021; Ramacher et al., 2020; Sun et al., 2019). CO_2 and NO_2 were important variables, so a high $PM_{0.1}$ concentration was probably related to vehicular and industrial emissions. Forest coverage also plays a crucial role in absorbing and filtering out pollutants and reducing air pollution (Zhai et al., 2022). For this study, areas with forest coverage exhibited low predicted concentrations of $PM_{0.1}$, but industrial and commercial areas with limited green space had no natural pollution control, so the $PM_{0.1}$ concentration was greater.

Wind speed and relative humidity are widely recognized as key meteorological factors that affect the spatial distribution of air pollutants (Xie et al., 2022; Zender-Świercz et al., 2024), but they were not selected as predictors for the final model for this study following SHAP-based variable selection. A Pearson correlation analysis comparing these unselected variables and those retained in the final model was performed. The results showed high collinearity between the u-component of wind speed, which is the horizontal component of wind velocity in the east-west direction, and both air temperature ($r = 0.74$) and sunshine duration ($r = 0.69$). Relative humidity was also strongly correlated with air temperature ($r = -0.80$) and sunshine duration ($r = -0.64$). These results were in close agreement with those of previous studies (Fan et al., 2018; He et al., 2020; Zateroglu, 2021), which demonstrated strong interrelationships among meteorological variables. Therefore, air temperature and sunshine duration, which were both retained in the final model, may implicitly represent the effects of wind and humidity, so they were excluded from the SHAP-based variable selection process.

This study had some limitations. Due to Taiwan's flight height regulations, drone sampling was limited to a maximum altitude of 121.92 m. In the future, the sampling altitude could be increased to allow the measurement of changes in air pollution at higher altitudes to improve high-level air quality monitoring and modeling. In terms of sampling regions, current drone regulations designate some areas as restricted flight zones. This restriction forced some sampling points to be located away from buildings and reduced the sample size for industrial and traffic land-use areas. This study selected alternative sampling locations that are representative of the surrounding environment to increase the robustness of the results, despite the limitations in terms of sample density.

Another limitation was the inconsistency in sampling times at each point due to the use of drones. Because of the high operational costs of

drone flights, this study used only one drone for sampling, so there was no synchronous data for each sampling point. Data collection was performed sequentially for each location, so meteorological and pollution data for each sampling period were incorporated into the database for calibration. This approach minimized the impact of atmospheric conditions at different times of sampling. Data quality was subject to limitations, so this study developed a methodological framework for 3-D air pollution assessment, rather than achieving highly precise measurements. Future research could simultaneously deploy multiple drones or install sensors on different floors of a building to capture more accurate and synchronized data at various sampling points and increase the accuracy of a 3-D air pollution assessment.

For this study, "Empirical Bayesian Kriging 3D" was used to visualize the predicted spatial distribution of $PM_{0.1}$, but several limitations existed. First, to mitigate computational constraints, a voxel size of 5 m was selected instead of a finer 1 m resolution, which substantially reduced processing time. However, this coarser resolution decreased spatial detail and may have obscured critical micro-scale variations, particularly in areas with steep gradients. Second, as the interpolation relied on predicted concentrations, any model errors could propagate and influence the 3-D spatial visualization. Lastly, Empirical Bayesian Kriging is known to be sensitive to outliers, which may result in unrealistically high or low estimation values (Zaresefat et al., 2024). While these limitations may affect the fine-scale precision of the visualizations, the results remained informative for identifying overall patterns and comparing locations and seasons.

In the future, this method can be applied to estimate gaseous pollutants and other air quality parameters. Its outputs may inform regulatory frameworks, such as policies requiring vertical installation of air pollution sensors on buildings. The resulting 3-D pollution profiles can guide urban planning to optimize building placement and green space allocation, reducing public exposure. This approach also helps identify high-risk zones, enabling targeted interventions and advisories, and supports the development of healthier, more sustainable urban environments.

5. Conclusions

Variable selection using the SHAP value indicated important factors, such as building height, temperature, carbon dioxide, nitric oxide, forest coverage, and 3-D spatial distance from buildings. The Geo-AI model achieved strong fitting, with an R^2 value of 0.98 and an RMSE value of 583 particles/cm^3. Techniques such as 10-fold CV and stratified validations also showed the stability of the model. The combination of UAV data and the Geo-AI model comprehensively captured the spatial variability of air pollution and revealed the complex distribution characteristics of $PM_{0.1}$ concentration in 3-D space, especially in urban environments with dense high-rise buildings. 3-D visualization identified pollution hotspots and revealed the variation in air quality with height, providing stakeholders with information to mitigate the effects of pollution and improve urban air quality management strategies in 3-D space.

CRediT authorship contribution statement

Chia-Wei Hsu: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yinq-Rong Chern:** Writing – review & editing, Software, Methodology, Investigation, Data curation. **Jun-Jun Su:** Writing – review & editing, Methodology, Investigation, Data curation. **Pei-Yi Wong:** Writing – review & editing. **Aji Kusumaning Asri:** Writing – review & editing. **Candera Wijaya:** Writing – review & editing, Resources. **Yu-Cheng Chen:** Writing – review & editing, Resources, Conceptualization. **Shih-Chun Candice Lung:** Writing – review & editing, Resources, Conceptualization. **Ta-Chih Hsiao:** Writing – review & editing, Resources, Methodology, Conceptualization. **Tee-Ann Teo:** Writing – review & editing, Resources, Conceptualization. **Yi-Liang Shih:** Writing – review & editing, Resources. **Chih-Da Wu:** Writing – review & editing, Supervision, Resources, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements and Funding

This study is supported by the National Science and Technology Council, Taiwan (NSTC 114-2119-M-006-010 -; NSTC 113-2121-M-006-010-; NSTC 112-2121-M-006-015-; NSTC 113-2123-M-001-013-; NSTC 113-2121-M-006-001-; 112-2314-B-038-118-MY2) and the Higher Education Sprout Project of the Ministry of Education (MOE) in Taiwan. This work is supported partially by the Research Center for Precision Environmental Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan under the auspices of The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project of the Ministry of Education (MOE) in Taiwan and by a Kaohsiung Medical University Research Center Grant (KMU-TC114A01).

This research is also funded by the Division of Research Planning and Development, Department of Research and Development, Taoyuan General Hospital, Ministry of Health and Welfare (PTH114058 and PTH114118), College of Oral Medicine, Taipei Medical University (TMUCOM202501), Chang Gung University of Science and Technology Foundation (ZRRPF6N0011) and Chang Gung Medical Research Program Foundation (CMRPF6P0041, CMRPF6P0042, and CMRPF6P0043).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envpol.2025.126879.

Data availability

The data that has been used is confidential.

References

Asri, A.K., Lee, H.Y., Chen, Y.L., Wong, P.Y., Hsu, C.Y., Chen, P.C., Lung, S.C.C., Chen, Y. C., Wu, C.D., 2024. A machine learning-based ensemble model for estimating diurnal variations of nitrogen oxide concentrations in Taiwan. Sci. Total Environ. 916, 170209. https://doi.org/10.1016/j.scitotenv.2024.170209.

Awaworyi Churchill, S., Inekwe, J., Ivanovski, K., Smyth, R., 2021. Transport infrastructure and CO_2 emissions in the OECD over the long run. Transport. Res. Transport Environ. 95, 102857. https://doi.org/10.1016/j.trd.2021.102857.

Azmi, W.N.F.W., Pillai, T.R., Latif, M.T., Koshy, S., Shaharudin, R., 2023. Application of land use regression model to assess outdoor air pollution exposure: a review. Environ. Adv. 11, 100353. https://doi.org/10.1016/j.envadv.2023.100353.

Babaan, J., Wong, P.Y., Chen, P.C., Chen, H.L., Lung, S.C., Chen, Y.C., Wu, C.D., 2024. Geospatial artificial intelligence for estimating daytime and nighttime nitrogen dioxide concentration variations in Taiwan: a spatial prediction model. J. Environ. Manag. 360, 121198. https://doi.org/10.1016/j.jenvman.2024.121198.

Baquero, T., Shukrallah, S., Karolia, R., Osammor, O., Inkson, B.J., 2015. Quantification of airborne road-side pollution carbon nanoparticles. J. Phys. Conf. 644 (1), 012023. https://doi.org/10.1088/1742-6596/644/1/012023.

Bhargava, A., Tamrakar, S., Aglawe, A., Lad, H., Srivastava, R.K., Mishra, D.K., Tiwari, R., Chaudhury, K., Goryacheva, I.Y., Mishra, P.K., 2018. Ultrafine particulate matter impairs mitochondrial redox homeostasis and activates phosphatidylinositol 3-kinase mediated DNA damage responses in lymphocytes. Environ. Pollut. 234, 406–419. https://doi.org/10.1016/j.envpol.2017.11.093.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Calderón-Garcidueñas, L., González-Maciel, A., Reynoso-Robles, R., Cejudo-Ruiz, F.R., Silva-Pereyra, H.G., Gorzalski, A., Torres-Jardón, R., 2025. Alzheimer's, parkinson's, frontotemporal lobar degeneration, and amyotrophic lateral sclerosis start in pediatric ages: ultrafine particulate matter and industrial nanoparticles are key in

the early-onset neurodegeneration: time to invest in preventive medicine. Toxics 13 (3).
Central Weather Administration, 2023. *Central weather administration OpenData*. Ministry of transportation and communications, MOTC. https://www.cwa.gov.tw.
Chang, H.T., Chern, Y.R., Asri, A.K., Liu, W.Y., Hsu, C.Y., Hsiao, T.C., Chi, K.H., Lung, S. C.C., Wu, C.D., 2025. Innovating Taiwan's greenhouse gas estimation: a case study of atmospheric methane using GeoAI-Based ensemble mixed spatial prediction model. J. Environ. Manag. 380, 125110. https://doi.org/10.1016/j.jenvman.2025.125110.
Chen, B., Song, Z., Pan, F., Huang, Y., 2022. Obtaining vertical distribution of $PM_{2.5}$ from CALIOP data and machine learning algorithms. Sci. Total Environ. 805, 150338. https://doi.org/10.1016/j.scitotenv.2021.150338.
Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.
Chen, Z.H., Li, B.W., Li, B., Peng, Z.R., Huang, H.C., Wu, J.Q., He, H.D., 2024. Identification of particle distribution pattern in vertical profile via unmanned aerial vehicles observation. Environ. Pollut. 348, 123893. https://doi.org/10.1016/j.envpol.2024.123893.
Chen, T.T., Wu, S.M., Ho, S.C., Chuang, H.C., Liu, C.Y., Chan, Y.F., Kuo, L.W., Feng, P.H., Liu, W.T., Chen, K.Y., Hsiao, T.C., Juang, J.N., Lee, K.Y., 2017. SUV39H1 reduction is implicated in abnormal inflammation in COPD. Sci. Rep. 7 (1), 46667. https://doi.org/10.1038/srep46667.
Cichowicz, R., Dobrzanski, M., 2021. 3D spatial analysis of particulate matter (PM_{10}, $PM_{2.5}$ and $PM_{1.0}$) and gaseous pollutants (H_2S, SO_2 and VOC) in urban areas surrounding a large heat and power plant. Energies 14 (14), 4070. https://doi.org/10.3390/en14144070.
Cowie, C.T., Garden, F., Jegasothy, E., Knibbs, L.D., Hanigan, I., Morley, D., Hansell, A., Hoek, G., Marks, G.B., 2019. Comparison of model estimates from an intra-city land use regression model with a national satellite-LUR and a regional bayesian maximum entropy model, in estimating NO_2 for a birth cohort in Sydney, Australia. Environ. Res. 174, 24–34. https://doi.org/10.1016/j.envres.2019.03.068.
Department of Household Registration, M. o. t. I., 2023. Population statistics by administrative division. https://www.ris.gov.tw/.
Dorogush, A.V., Ershov, V., Gulin, A., 2018. CatBoost: Gradient Boosting with Categorical Features Support. *ArXiv, abs/1810.11363*.
Duc, H.N., Rahman, M.M., Trieu, T., Azzi, M., Riley, M., Koh, T., Liu, S., Bandara, K., Krishnan, V., Yang, Y., Silver, J., Kirley, M., White, S., Capnerhurst, J., Kirkwood, J., 2022. Study of planetary boundary layer, air pollution, air quality models and aerosol transport using ceilometers in new south wales (NSW), Australia. Atmosphere 13 (2).
Eeftens, M., Odabasi, D., Flückiger, B., Davey, M., Ineichen, A., Feigenwinter, C., Tsai, M. Y., 2019. Modelling the vertical gradient of nitrogen dioxide in an urban area. Sci. Total Environ. 650, 452–458. https://doi.org/10.1016/j.scitotenv.2018.09.039.
Fan, Q., Liu, S., Wang, K., 2018. Detecting the auto-correlation between daily temperature and relative humidity time series. Fractals 27 (2), 1950003. https://doi.org/10.1142/S0218348X19500038.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29, 1189–1232.
Hammond, J., Maher, B.A., Gonet, T., Bautista, F., Allsop, D., 2022. Oxidative stress, cytotoxic and inflammatory effects of urban ultrafine road-deposited dust from the UK and Mexico in human epithelial lung (Calu-3) cells. Antioxidants 11 (9).
Han, T., Xie, C., Liu, Y., Yang, Y., Zhang, Y., Huang, Y., Gao, X., Zhang, X., Bao, F., Li, S. M., 2024. Development of a continuous UAV-Mounted air sampler and application to the quantification of CO_2 and CH_4 emissions from a major coking plant. Atmos. Meas. Tech. 17 (2), 677–691. https://doi.org/10.5194/amt-17-677-2024.
He, B.-J., Ding, L., Prasad, D., 2020. Outdoor thermal environment of an open space under sea breeze: a Mobile experience in a coastal city of Sydney, Australia. Urban Clim. 31, 100567. https://doi.org/10.1016/j.uclim.2019.100567.
He, X., Zhao, K.Y., Chu, X.W., 2021. AutoML: a survey of the state-of-the-art. Knowl. Base Syst. 212. https://doi.org/10.1016/j.knosys.2020.106622. Article 106622.
Hsu, C.H., Cheng, F.Y., 2019. Synoptic weather patterns and associated air pollution in Taiwan. Aerosol Air Qual. Res.
Huang, Y.d., Hou, R.-W., Liu, Z., Song, Y., Cui, P.-Y., Kim, C.N., 2019. Effects of wind direction on the airflow and pollutant dispersion inside a long street canyon. Aerosol Air Qual. Res.
Idrees, Z., Zheng, L., 2020. Low cost air pollution monitoring systems: a review of protocols and enabling technologies. Journal of Industrial Information Integration 17, 100123. https://doi.org/10.1016/j.jii.2019.100123.
Jiang, Y., Zhu, X., Shen, Y., He, Y., Fan, H., Xu, X., Zhou, L., Zhu, Y., Xue, X., Zhang, Q., Du, X., Zhang, L., Zhang, Y., Liu, C., Niu, Y., Cai, J., Kan, H., Chen, R., 2024. Mechanistic insights into cardiovascular effects of ultrafine particle exposure: a longitudinal panel study. Environ. Int. 187, 108714. https://doi.org/10.1016/j.envint.2024.108714.
Jiang, Z.W., Cheng, H.M., Zhang, P.H., Kang, T.F., 2021. Influence of urban morphological parameters on the distribution and diffusion of air pollutants: a case study in China. J. Environ. Sci. 105, 163–172. https://doi.org/10.1016/j.jes.2020.12.035.
Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. Lightgbm: a Highly Efficient Gradient Boosting Decision Tree. Neural Information Processing Systems.
Kelley, E.E., McCallen, H., Nurkiewicz, T., Pilkerton, J., Hussain, S., Devallance, E., 2024. Abstract 1119: ultrafine particle inhalation initiates vascular inflammation via hepatic-derived xanthine oxidoreductase-mediated epigenetic alterations. Arterioscler. Thromb. Vasc. Biol. 44 (Suppl. 1_1), A1119. https://doi.org/10.1161/atvb.44.suppl_1.1119. A1119.
Kim, J.W., Baik, J.J., Han, B.S., Lee, J., Jin, H.G., Park, K., Yang, H., Park, S.B., 2022. Tall-building effects on pedestrian-level flow and pollutant dispersion: large-Eddy simulations. Atmos. Pollut. Res. 13 (8), 101500. https://doi.org/10.1016/j.apr.2022.101500.
Lee, M., Brauer, M., Wong, P.L.N., Tang, R., Tsui, T.H., Choi, C., Cheng, W., Lai, P.C., Tian, L.W., Thach, T.Q., Allen, R., Barratt, B., 2017. Land use regression modelling of air pollution in high density high rise cities: a case study in Hong Kong. Sci. Total Environ. 592, 306–315. https://doi.org/10.1016/j.scitotenv.2017.03.094.
Leikauf, G.D., Kim, S.H., Jang, A.S., 2020. Mechanisms of ultrafine particle-induced respiratory health effects. Exp. Mol. Med. 52 (3), 329–337. https://doi.org/10.1038/s12276-020-0394-0.
Li, B., Li, X.B., Li, C., Zhu, Y., Peng, Z.R., Wang, Z.Y., Lu, S.J., 2019. Impacts of wind fields on the distribution patterns of traffic emitted particles in urban residential areas. Transport. Res. Transport Environ. 68, 122–136. https://doi.org/10.1016/j.trd.2018.01.030.
Li, C., Liu, M., Hu, Y., Wang, H., Xiong, Z., Wu, W., Liu, C., Zhang, C., Du, Y., 2022. Investigating the vertical distribution patterns of urban air pollution based on unmanned aerial vehicle gradient monitoring. Sustain. Cities Soc. 86, 104144. https://doi.org/10.1016/j.scs.2022.104144.
Li, Y., Hong, T., Gu, Y., Li, Z., Huang, T., Lee, H.F., Heo, Y., Yim, S.H.L., 2023. Assessing the spatiotemporal characteristics, factor importance, and health impacts of air pollution in Seoul by integrating machine learning into land-use regression modeling at high spatiotemporal resolutions. Environ. Sci. Technol. 57 (3), 1225–1236. https://doi.org/10.1021/acs.est.2c03027.
Lin, T.C., Chiueh, P.T., Griffith, S.M., Liao, C.C., Hsiao, T.C., 2022. Deployment of a Mobile platform to characterize spatial and temporal variation of on-road fine particles in an urban area. Environ. Res. 204, 112349. https://doi.org/10.1016/j.envres.2021.112349.
Lugassi, R., Blank, A., Rogozovsky, I., Ohneiser, K., Ansmann, A., Linzon, Y., Chudnovsky, A., 2022. From laboratory to in-situ 3D measurements of complex pollution states in the city: introducing a general concept using compact multisensory assemblies on UAVs. Atmos. Environ. 281, 119146. https://doi.org/10.1016/j.atmosenv.2022.119146.
Lung, S.C.C., Lee, H.Y., Chern, Y., Wu, C.D., Chen, Y.C., Zeng, Y.T., Chen, N.T., Wong, P. Y., Su, H., 2021. Using a land use regression model with machine learning to estimate ground level $PM_{2.5}$. Environ. Pollut. 277, 116846. https://doi.org/10.1016/j.envpol.2021.116846.
National Land Surveying and Mapping Center, 2023. Administrative area statistics. https://www.nlsc.gov.tw/.
Pochwała, S., Gardecki, A., Lewandowski, P., Somogyi, V., Anweiler, S., 2020. Developing of low-cost air pollution sensor—measurements with the unmanned aerial vehicles in Poland. Sensors 20 (12), 3582. https://www.mdpi.com/1424-8220/20/12/3582.
Pohorsky, R., Baccarini, A., Brett, N., Barret, B., Bekki, S., Pappaccogli, G., Dieudonné, E., Temime-Roussel, B., D'Anna, B., Cesler-Maloney, M., Donateo, A., Decesari, S., Law, K.S., Simpson, W.R., Fochesatto, J., Arnold, S.R., Schmale, J., 2025. In situ vertical observations of the layered structure of air pollution in a continental high-latitude urban boundary layer during winter. Atmos. Chem. Phys. 25 (6), 3687–3715. https://doi.org/10.5194/acp-25-3687-2025.
Ramacher, M.O.P., Matthias, V., Aulinger, A., Quante, M., Bieser, J., Karl, M., 2020. Contributions of traffic and shipping emissions to city-scale NO_x and $PM_{2.5}$ exposure in Hamburg. Atmos. Environ. 237, 117674. https://doi.org/10.1016/j.atmosenv.2020.117674.
Roldán, J.J., Joossen, G., Sanz, D., del Cerro, J., Barrientos, A., 2015. Mini-UAV based sensory System for measuring environmental variables in greenhouses. Sensors 15 (2), 3334–3350. https://doi.org/10.3390/s150203334.
Salvador, P., Pandolfi, M., Tobías, A., Gómez-Moreno, F.J., Molero, F., Barreiro, M., Pérez, N., Revuelta, M.A., Marco, I.M., Querol, X., Artíñano, B., 2020. Impact of mixing layer height variations on air pollutant concentrations and health in a European urban area: madrid (Spain), a case study. Environ. Sci. Pollut. Control Ser. 27 (33), 41702–41716. https://doi.org/10.1007/s11356-020-10146-y.
Sun, W., Zhou, Y., Lv, J., Wu, J., 2019. Assessment of multi-air emissions: case of particulate matter (dust), SO_2, NO_x and CO_2 from iron and steel industry of China. J. Clean. Prod. 232, 350–358. https://doi.org/10.1016/j.jclepro.2019.05.400.
Tessum, M.W., Sheppard, L., Larson, T.V., Gould, T.R., Kaufman, J.D., Vedal, S., 2021. Improving air pollution predictions of long-term exposure using short-term Mobile and stationary monitoring in two US metropolitan regions. Environ. Sci. Technol. 55 (6), 3530–3538. https://doi.org/10.1021/acs.est.0c04328.
Villa, T.F., Salimi, F., Morton, K., Morawska, L., Gonzalez, F., 2016. Development and validation of a UAV based System for air pollution measurements. Sensors 16 (12). https://doi.org/10.3390/s16122202. Article 2202.
Wong, P.Y., Su, H.J., Lee, H.Y., Chen, Y.C., Hsiao, Y.P., Huang, J.W., Teo, T.A., Wu, C.D., Spengler, J.D., 2021. Using land-use machine learning models to estimate daily NO_2 concentration variations in Taiwan. J. Clean. Prod. 317, 128411. https://doi.org/10.1016/j.jclepro.2021.128411.
Wu, C.D., Chen, Y.C., Pan, W.C., Zeng, Y.T., Chen, M.J., Guo, Y.L., 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. Environ. Pollut. 224, 148–157. https://doi.org/10.1016/j.envpol.2017.01.074.
Xie, J., Sun, T., Liu, C., Li, L., Xu, X., Miao, S., Lin, L., Chen, Y., Fan, S., 2022. Quantitative evaluation of impacts of the steadiness and duration of urban surface wind patterns on air quality. Sci. Total Environ. 850, 157957. https://doi.org/10.1016/j.scitotenv.2022.157957.
Xu, S., Zou, B., Shafi, S., Sternberg, T., 2018. A hybrid Grey-Markov/LUR model for PM_{10} concentration prediction under future urban scenarios. Atmos. Environ. 187, 401–409. https://doi.org/10.1016/j.atmosenv.2018.06.014.

Xu, X.Y., Qin, N., Zhao, W.J., Tian, Q., Si, Q., Wu, W.Q., Iskander, N., Yang, Z.C., Zhang, Y.W., Duan, X.L., 2022. A three-dimensional LUR framework for $PM_{2.5}$ exposure assessment based on mobile unmanned aerial vehicle monitoring. Environ. Pollut. 301. https://doi.org/10.1016/j.envpol.2022.118997. Article 118997.

Yang, J., Shi, B., Shi, Y., Marvin, S., Zheng, Y., Xia, G., 2020. Air pollution dispersal in high density urban areas: research on the triadic relation of wind, air pollution, and urban form. Sustain. Cities Soc. 54, 101941. https://doi.org/10.1016/j.scs.2019.101941.

Zaresefat, M., Derakhshani, R., Griffioen, J., 2024. Empirical bayesian kriging, a robust method for spatial data interpolation of a large groundwater quality dataset from the Western Netherlands. Water 16 (18).

Zateroglu, M.T., 2021. Statistical Models For Sunshine Duration Related To Precipitation And Relative Humidity [Güneşlenme Süresinin Yağış Miktarı Ve Nisbi Nem İle İlişkili İstatistiksel Modelleri]. Avrupa Bilim ve Teknoloji Dergisi(29) 208–213. https://doi.org/10.31590/ejosat.1022962.

Zender-Świercz, E., Galiszewska, B., Telejko, M., Starzomska, M., 2024. The effect of temperature and humidity of air on the concentration of particulate matter - $PM_{2.5}$ and PM_{10}. Atmos. Res. 312, 107733. https://doi.org/10.1016/j.atmosres.2024.107733.

Zhai, C., Bao, G., Zhang, D., Sha, Y., 2022. Urban Forest locations and patch characteristics regulate $PM_{2.5}$ mitigation capacity. Forests 13 (9).

Zhang, W., Chen, B., Yoda, Y., Shima, M., Zhao, C., Ji, X., Wang, J., Liao, S., Jiang, S., Li, L., Chen, Y., Guo, X., Deng, F., 2024. Ambient ultrafine particles exacerbate oxygen desaturation during sleep in patients with chronic obstructive pulmonary disease: new insights into the effect spectrum of ultrafine particles on susceptible populations. Sci. Total Environ. 947, 174519. https://doi.org/10.1016/j.scitotenv.2024.174519.

Zhu, L.Y., Ranasinghe, D., Chamecki, M., Brown, M.J., Paulson, S.E., 2021. Clean air in cities: impact of the layout of buildings in urban areas on pedestrian exposure to ultrafine particles from traffic. Atmos. Environ. 252, 118267. https://doi.org/10.1016/j.atmosenv.2021.118267.