Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Machine-Learning Approach for Improving MODIS Aerosol Optical Depth Retrieval Using Ground-based Observations over Taiwan

Jelita Citrawati Jihan
National Central University
Tang-Huang Lin
thlin@csrsr.ncu.edu.tw
National Central University
Nguyen Van Hieu
Thai Nguyen University of Agriculture and Forestry

Research Article

Keywords: aerosol optical depth (AOD), MODIS, AERONET, ANN

Posted Date: May 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9421875/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A Machine-Learning Approach for Improving MODIS Aerosol Optical Depth Retrieval Using Ground-based Observations over Taiwan

Jelita Citrawati Jihan[1,2*], Tang-Huang Lin[2,3†] and Nguyen Van Hieu[4]

[1*]Graduate Program of Environmental Science and Technology, National Central University, Taoyuan, 320, Taiwan. [2]Center for Space and Remote Sensing Research, National Central University Taoyuan City 320, Taiwan. [3]Center for Astronautical Physics and Engineering, National Central University, Taoyuan City 320, Taiwan, [4]Geoinformatics Research Center, Thai Nguyen University of Agriculture and Forestry, Thai Nguyen, Vietnam.

*Corresponding author(s). E-mail(s): thlin@csrsr.ncu.edu.tw;
Contributing authors: its.mejiehan@gmail.com, hieu_tn2002@yahoo.com;
†These authors contributed equally to this work.

Abstract

Taiwan's diverse land cover provides an excellent testbed for validating satellite and ground-based aerosol optical depth (AOD) measurements. This study validates MODIS (AQUA and TERRA) Level 2 AOD products against AEronet RObotic NETwork (AERONET) measurements across Taiwan during during 2018–2022 for model training, with independent temporal validation using 2023–2024 observations. We developed an Artificial Neural Network (ANN) model incorporating meteorological indicators, land cover characteristics, and temporal patterns to improve AOD retrievals in areas with and without ground-based observations. The ANN model demonstrated strong performance with determination coefficients (R^2) of 0.95 and 0.94 for TERRA and AQUA respectively, and root-mean-square errors (RMSE) of 0.05 and 0.07. Additionally, we implemented a novel Land Cover Learning (LCL) framework that enables cross-validation between different AERONET sites based on similar environmental characteristics, effectively addressing spatial gaps in AOD data. The LCL framework achieved 91% of cross-validated retrievals within the expected error bounds, demonstrating effective gap-filling over the 31,500 km^2 reconstructed domain. This approach demonstrates potential for application in other regions with similar geographical diversity and limited ground-based monitoring networks.

Keywords: aerosol optical depth (AOD), MODIS, AERONET, ANN

1 Introduction

Atmospheric aerosols exhibit diverse emission patterns over time from both natural and anthropogenic processes. According to the Intergovernmental Panel on Climate Change (IPCC) report on Climate Change and Land) IPCC (2022) agriculture, forestry and other land use (AFOLU) is a significant net source of greenhouse emissions (high confidence) at different species. For example, anthropogenic aerosol emissions have increased substantially due to urban development across industry, transportation, and energy sectors, raising major global issues related to public health and climate change Boucher et al. (2021), Pöschl (2005).

Aerosol properties exhibit significant spatial and temporal variations. Ground-based monitoring networks provide accurate columnar Aerosol Optical Depth (AOD) measurements at high spectral resolution but are inherently limited in spatial coverage. To compensate, satellite remote sensing provides spatially continuous AOD estimates representing aerosol scattering and absorption of solar radiation. However, space-based AOD measurements face limitations including low temporal resolution of polar-orbiting satellites and delayed launch times of geostationary satellites Cohen & Justice (1999). Aerosol retrieval under cloud cover becomes challenging due to insufficient satellite coverage.

Recent investigations Levy, Remer, Shi, & Kleidman (2023) have if focused on long-term correction of Level 2 data (AODlv.2) using ground-based measurements with 10 km × 10 km gridded areal observations. The Moderate Resolution Imaging Spectroradiometer (MODIS) instrument aboard AQUA and TERRA satellites from NASA provides measurements at three visible wavelengths (0.47, 0.55, and 0.66 μm) for AOD retrieval Sayer, Hsu, Bettenhausen, Jeong, & Meister (2015). The spectral variation across wavelengths enables interpretation of atmospheric species through scattering and absorption indices. The spectral variation across wavelengths enables interpretation of aerosol types through scattering and absorption indices. Missing values (NaN) in AODlv.2 occur due to fully and partially cloudy conditions, as cloud cover prevents satellite AOD retrieval. Unlike satellite sensors, AERONET sun photometers also require direct solar beam access and cannot provide AOD under overcast conditions Holben et al. (2006).

Several studies have established integration techniques for estimating aerosol concentrations in areas with missing satellite data. Yang & Hu (2018) developed spatiotemporal gap-filling methods for missing AODlv.2 data; however, this framework focused primarily on data imputation rather than physical aerosol characterization. While spatiotemporal of ground observations provides stability near domain boundaries Liang, Daniels, Biancardi, & Zhou (2023a), these studies used only AOD imagery for missing pixel comparison. Ground-based correction schemes influenced by meteorology and

surface conditions Lanzaco, Olcese, Palancar, & Toselli (2017) have focused primarily on land areas.

State-of-the-art integration of observational and model data requires new methods for handling diverse spatial and temporal sampling, collectively called data integration. Previous approaches used fusion models combining image and non-image data to assess aerosol variations on a global scale Remer et al. (2005). Our work focuses specifically on AOD over land surface areas of Taiwan, where supervisor's laboratory has demonstrated the feasibility of satellite-derived AOD estimation Liu, Chen, Lin, & Kuo (2002), including gap-filling methods that align with IPCC framework guidelines.

The main novelty of this study lies in the direct integration of modified aerosol imagery from complete AODlv.2 volumes from clear sky and cloud sky, combining space-based and surface-based observations through machine learning algorithms. We developed a machine learning model that accurately predicts Level 3 AOD products (AODlv.3) lab scale as ground-based measurements and full scale over Taiwan, utilizing spatiotemporal remote sensing data. Other studies have comprehensively discussed practical options for machine learning methods, particularly artificial neural networks (ANN), to correct satellite AOD retrievals and improve uncertainty estimation Kang, Kim, Kang, Cho, & Im (2022), Lary, Remer, MacNeill, Roscoe, & Paradise (2009).

Our first analysis focuses on real-time hourly aerosol signature data over the 2018–2022 training period, comparing TERRA and AQUA AODlv.2 products against AERONET measurements to derive corrected AODlv.3 using ANN at 550 nm. The trained ANN model is then independently validated against 2023–2024 observations to assess temporal generalizability. We evaluate the Land Cover Learning (LCL) framework, which transfers aerosol model knowledge from AERONET-monitored stations to unmonitored areas with similar land cover and EVI characteristics. Our second analysis applies species baseline combined with ANN correction to reconstruct AOD across all of Taiwan under various aerosol-cloud conditions. Both approaches together produce Level 3 merged AOD products (AODlv.3). Model performance is evaluated using the expected error bounds (EE = ±[0.05 + 0.15 × AOD_AERONET]) following Sayer, Hsu, Bettenhausen, & Jeong (2013).

2 Materials and Methods

2.1 Study Area and Data

The focus area encompasses AOD measurements across Taiwan using AERONET ground-based instruments operated by NASA (National Aeronautics and Space Administration) and PHOTONS (Photométrie pour le Traitement Opérationnel de Normalisation Satellitaire). Ten monitoring stations were selected to evaluate satellite aerosol products. The distribution of monitoring stations is shown in Figure 1, where stars represent AERONET locations described in Table 1, and red dots show locations of Central Weather Administration (CWA) stations. Urban areas are defined as densely developed city centers; rural areas as sparsely populated regions; and coastal areas as stations positioned along Taiwan's coastline. Western Taiwan represents the most populated area.

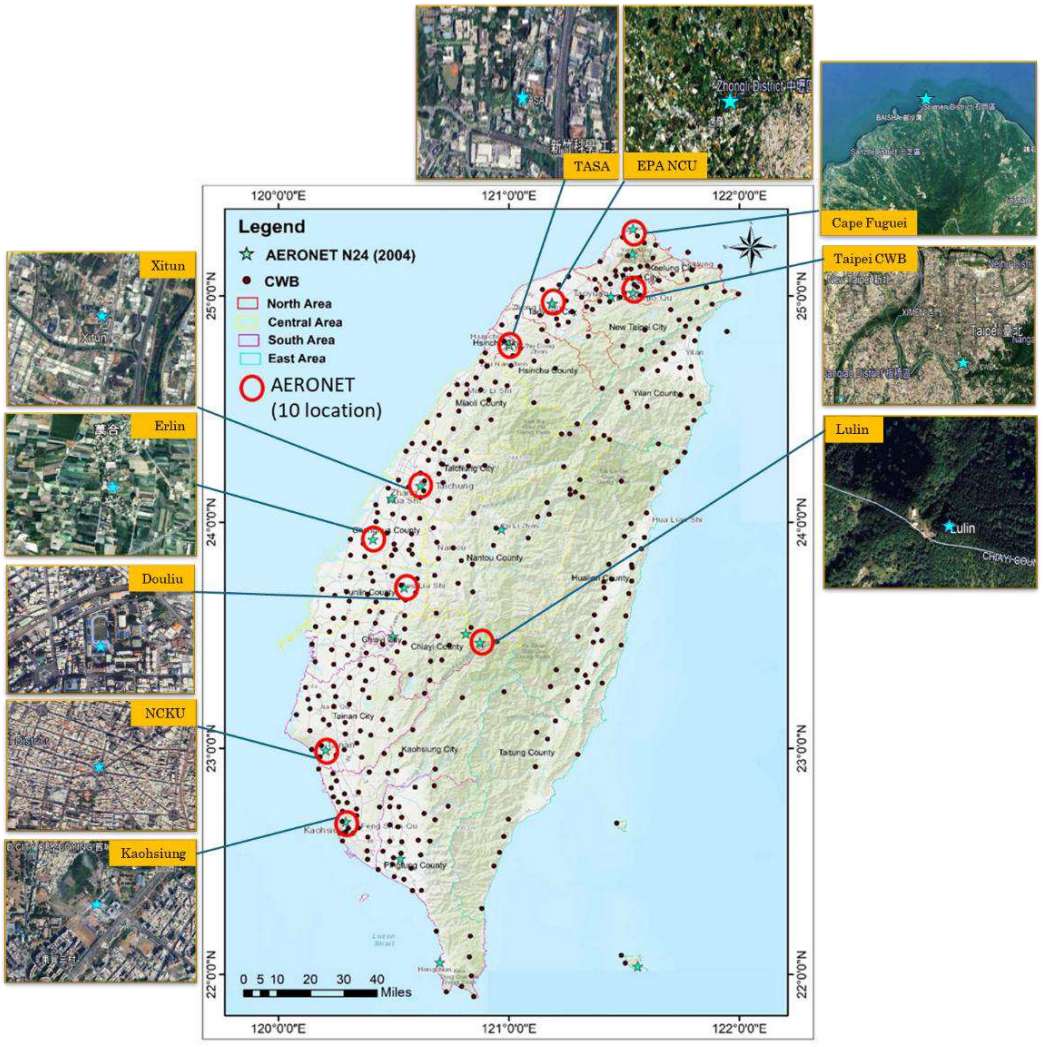


Fig. 1 Distribution of meteorological monitoring and aerosol observation sensors across Taiwan, showing 10 AERONET stations (circle) and CWA stations (red dots). Stations are classified as urban (U), suburban (SU), coastal (C), and rural (R).

2.1.1 Ground-based AOD Measurements

Table 1 AERONET station locations and characteristics across Taiwan

Location		Lat (N)	Long (E)	Ground-based Name	Application Areas / Land Cover Code	Data Availability TERRA| AQUA
North						
1.	New Taipei City	25°17'	121°32'	Cape_Fuguei_ Station	Coastal Park (C)/ 11	104 | 83
2.	Taipei City	25° 0'	121°32'	Taipei CWB	Elementary School (U)/13	173 | 163
3.	Taoyuan	24°58'	121°11	EPA-NCU	University (R)/13	388 | 348
4.	Hsinchu City	24°47'	121° 0'	TASA	Primary School (U)/13	321 | 276
Central						
5.	Taichung City	24° 9'	120°37'	Xitun	School (U)/13	446 | 422
6.	Changhua County	23°55'	120°24'	Erlin	School (R)/12	165 | 175
7.	Yunlin County	23°42'	120°32'	Douliu	Senior High School (SU)/13	64 | 40
8.	Nantou	23°28'	120°52'	Lunlin	Mountain /5	284 | 115
South						
9.	Tainan City	22°59'	120°12'	Chen Kung Univ	University (C)/13	203 | 185
10	Kaohsiung City	22°40'	120°17'	Kaohsiung	School (C)/13	314 | 317

To analyze AODlv.2 products, Table 1 provides information on ground-based stations categorized as urban (U), suburban (SU), coastal (C), and rural (R) areas, selected from educational and public facilities. The ten AERONET sites span northern, central, and southern Taiwan regions. Complete AERONET data are available at https://aeronet.gsfc.nasa.gov/. AERONET provides three data quality levels: Level 1.0 (unscreened), Level 1.5 (cloud-screened and quality controlled), and Level 2.0 (cloud-screened and quality-assured). This study uses Level 2.0 Version 3.0 AOD data Giles et al. (2019) at 15-minute intervals. The 2018–2022 dataset was used for ANN model training and validation, while the 2023–2024 dataset was reserved exclusively for independent temporal testing. Since AERONET does not provide AOD values at the 550 nm band, we calculated AOD at 550 nm using the Ångström power law between 500 nm and 675 nm bands Eck et al. (1999). The Ångström Exponent (α) represents the spectral dependence index between 500-675 nm, calculated using equations (1) and (2):

$$\alpha = -\frac{\ln\left(\tau_{\lambda_1} / \tau_{\lambda_2}\right)}{\ln\left(\lambda_1 / \lambda_2\right)} \tag{1}$$

$$\tau_{0.55\mu m} = \tau_{\lambda_2}\left(\frac{\lambda_1}{\lambda_2}\right)^{-\alpha} \tag{2}$$

Where, α = index 500 and 675nm, λ1 = AOD at wavelengths 500, λ2= AOD at 675.

2.1.2. Meteorological Data from Central Weather Administration (CWA)

Taiwan's Climate Observation Data Inquiry Service (CODIS) (https://e-service.cwa.gov.tw/HistoryDataQuery/index.jsp) provides time series meteorological data from manual observation stations across Taiwan. The distribution of CODIS stations includes flat ground and mountain locations (120°E-122°E, 21.5°N-26°N). Since meteorological factors significantly influence AOD concentrations, we used hourly data for temperature, wind direction, wind speed, and relative humidity from January 2018 to December 2024 for each ground-based location and corresponding AODlv.2 pixels. Meteorological data from 2018–2022 supported ANN model training; data from 2023–2024 supported independent temporal validation.

2.1.3. MODIS Space-based AOD Data (AODlv.2)

MODIS data were obtained from the TERRA and AQUA satellites using the Scientific Data Set (SDS) "Deep_Blue_Spectral_Aerosol_Optical_Depth_Land" from NASA (https://ladsweb.modaps.eosdis.nasa.gov/). NASA launched TERRA into sun-synchronous orbit (10:30 local standard time) in 1999 and AQUA (13:30 local standard time) in 2002. MODIS features 36 spectral bands with wavelength ranges from 0.4 to 14.4 micrometers (μm). During Earth overpasses, satellites capture surface strips approximately 2,330 kilometers wide at an altitude of approximately 705 kilometers King, Kaufman, Menzel, & Tanre (1992). We used the Deep Blue (DB) algorithm with Quality Assurance (QA=3) from January 2018 to December 2024 over Taiwan, with 10 km spatial resolution.

2.1.4. Land Cover Type Classification

Land cover characteristics significantly contribute to aerosol species formation through physical and chemical processes. We used Version 6 of the MODIS land cover product (MCD12Q1) with 500 m spatial resolution from https://modis.gsfc.nasa.gov/data/dataprod/mod12.php. Land cover properties were derived from five years of TERRA and AQUA MODIS observations to identify aerosol models across all sites, providing five global land cover classification systems: the IGBP, University of Maryland, LAI, BIOME-BGC, and Plant Functional Type schemes Friedl et al. (2010), Sulla-Menashe, Gray, Abercrombie, & Friedl (2019).

2.1.5. Global Enhanced Vegetation Index (EVI)

High vegetation cover areas offer resistance to wind-driven aerosol transport and enhance aerosol dry deposition Y J Kaufman & Tanre (1992). EVI was selected over NDVI for this study because of its superior sensitivity in high-biomass regions and its built-in atmospheric resistance correction for residual aerosol effects Huete et al. (2002). We obtained 16-day EVI data at 250 m resolution from MOD13Q1 (https://modis.gsfc.nasa.gov/data/dataprod/mod13.php). Long-term Vegetation Index (VI) data availability enhances vegetation monitoring with improved sensitivity in high biomass regions S. W. Running et al. (1999) for aerosol modeling

across 10 ground-based cross-validation (CV) sites within MODIS AODlv.2 10 km pixels. EVI analysis distinguishes between flat-ground and mountain land surface groups across the 10 AERONET sites. When simulations collect all species likely introduced to account for differences in land surface reflectance, land-use characteristics, and atmospheric aerosol influence. To sustained reconstruct and extracion AOD based on the VI interval of 0.04 from -1.0 to +1.0.

2.2 Methods

The techniques can be grouped into two classes: Statistical methods and spatial distribution related modify AOD is reconstruct aerosol based on real-time hourly. To do this, we simulated cross-validation lab-scale studies (stage 1) is 10 model-based aerosol follow intrinsic values of aerosol particle size distribution from variability land cover. Furthermore, full scale studies we compare and classified together EVI between raws and column image twin MODIS with each AERONET to be consistent EVI behaves for next analysis (stage 2). Precisely, algorithm AODlv.2 spectral direction uses wavelets at 0.55μm find coordinate point estimates, we use spectral to generate input feature in each grid cell projected in cylindrical map projection to reconstruct area AODlv.2 and no pixel data pairs with model-based aerosol and structure-environment (stage 3). Think of mapping has been conduct with arithmetic over Taiwan Boucher et al. (2021, Levy et al. (2023), Owili, Lien, Muga, & Lin (2017). This body of work is significant because it provides a systematic approach to missing aerosol, improving our understanding of aerosol distribution and their impact on the environment. The complete workflow is illustrated in Figure 2.

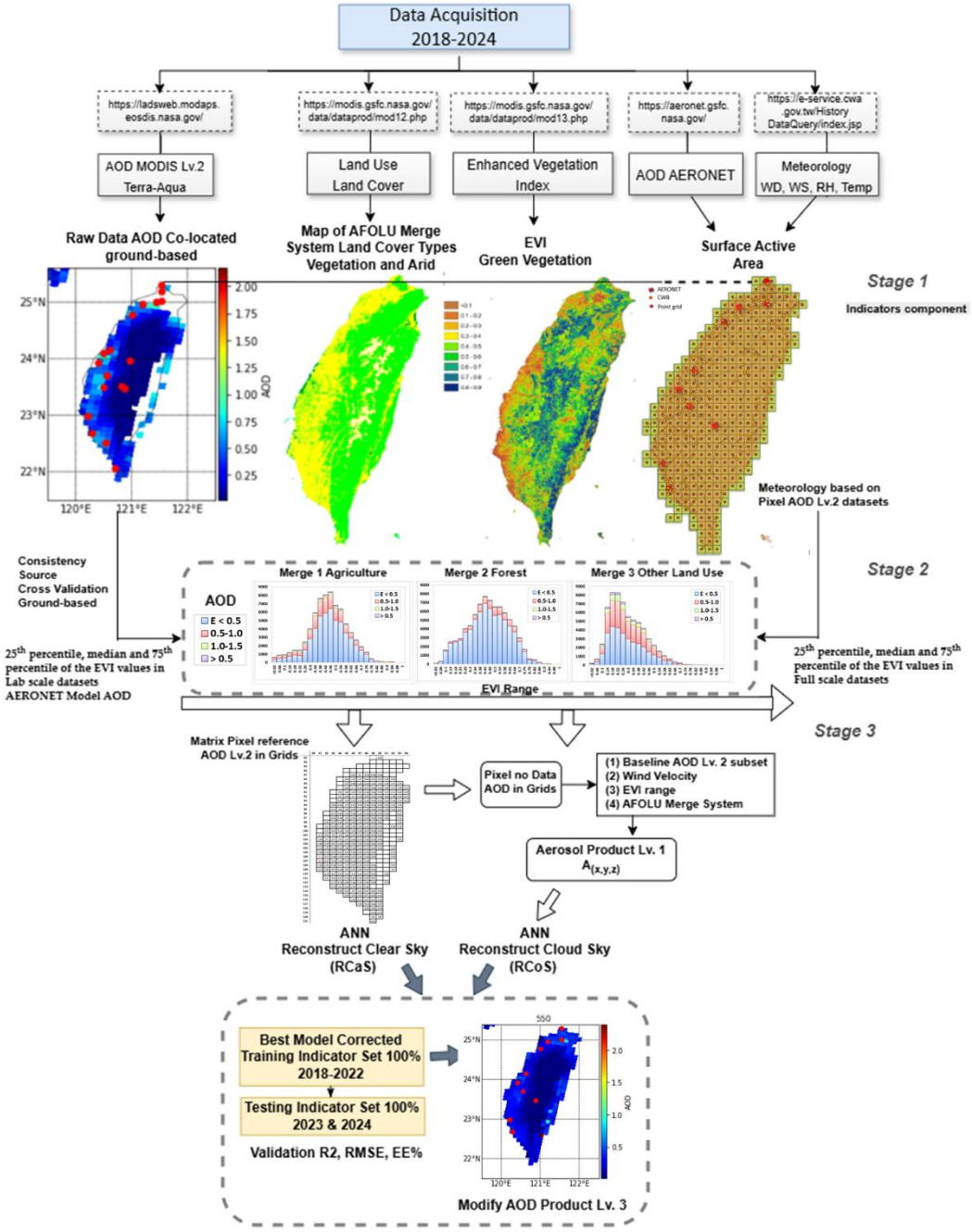


Fig. 2 Comprehensive flow diagram showing the procedure for obtaining corrected AOD values, processing 2,685 TERRA and 2,789 AQUA pure pixels extracted from January 2018 to December 2024 over Taiwan.

2.2.1 Artificial Neural Network (ANN) Correction Model

The Artificial Neural Network functions as a brain-inspired information processing system, chosen as a non-linear tool following Basheer & Hajmeer (2000). Neural activation functions are mathematical functions essential for neural network learning. Machine learning models for time series aerosol data attempt to improve predictions with ANN, where the entire training set is predicted simultaneously Hornik, Stinchcombe, & White (1989). The regression function performance can be improved by correcting uncalibrated MODIS AODlv.2 against AERONET Model Aerosol. We developed a correction system for MODIS AODlv.2 using available structure-environment indicators (temperature, relative humidity, wind speed, wind direction, and modified day of year) to learn non-linear system models using multiple layers between input and output layers.

We used ANN models to predict AOD from MODIS against each ground-based station and for areas without ground-based coverage. For suitable output, we selected a feedforward multilayer perceptron (MLP) architecture. Training data were used to train MLP neural networks using backpropagation with multiple hidden layers, adapting procedures for large datasets as described by Rumelhart, Hinton, & Williams (1986). Our framework introduces a novel three-layer feed-forward architecture (input, hidden, and output), with each layer using forward propagation and weight update phases to calculate error as the difference between target and output values. We applied systematic trial and error for function combination building blocks with linear independence for machine learning applications.

Model development simulates input dimensions using Rectified Linear Unit (ReLU) activation functions. Sigmoid activation functions in hidden layers provide values ranging between 0-1, while output layers use linear activation functions. Our framework contributions include: (1) MLP adaptive development for conditional and unconditional prior data considering compositions and size distributions; (2) Algorithm validation as valuable prediction functions with minimal error reaching actual output values; and (3) Modified AOD from indicator factors associated with targets to analyze and improve uncertainty performance.

The mathematical formulation (3) for the MLPN model is:

$$y_j = f\left(\sum_{i=1}^{N} w_{ji} \cdot x_{ji} + b_j\right) \tag{3}$$

where the definition of yj is the output of node j, f is the activation function, xji represents the input from node i to node j, wji denotes the weight associated with the connection between node i and j, bj stands for the bias of node j. Detailed procedures for AODlv.2 and Aerosollv. 1 construction are described below

1) Algorithm AODlv.2 colocations with reference 10 ground-based AERONET that has established that temporal on hourly. Efficient way, the combined algorithm obtain aerosol value from pixel-by-pixel using MODIS hourly image from 2018 to 2022.
2) Input a physical model environment AOD, day of year (DOY) and meteorological data estimate reconstruct AODlv.2 as output data. This output model based on linear regression model expressed as equation (4)

$$C\ AOD_{TERRA/AQUA} = l^* \text{ Model Aerosol} + a \tag{4}$$

Where Corrected (C) $AOD_{TERRA/AQUA}$ (AODlv.3) are correction existed value of satellite-retrieved AODlv.2, validating with AERONET Model Aerosol. l and a are regression slope and intercept.

3) We developed a simplified Land Cover Learning (LCL) framework that assumes building blocks teach machines to learn location data from model-based aerosol entering another model environment aerosol. Cross-validation enables simultaneous evaluation of collocated AERONET as model aerosol techniques, filling AOD gaps and improving corrections through $Y_{correctionlearning}$.

2.2.2 Enhanced Input Modified Day of Year (MDOY)

We evaluated time series AOD differences between observed and predicted values, combining with one-hour averaged meteorological data. Modified DOY significantly enhances model development accuracy Lanzaco, Olcese, Palancar, & Toselli (2016). The modification scales annual average meteorological trends from level 1 on January 1st to level 183 on July 1st, then decreases back to level 1 on December 31st, capturing seasonal cyclical patterns.

2.3 Enhanced AOD Reconstruct Clear Sky (RCaS) and Cloud Sky (RCoS) Coverage

2.3.1 Advanced Processing of Species Baseline for Cloud Sky

Under cloud-free conditions, mean AOD values for each pixel center located within grid cells are provided by MODIS aerosol products. Our simplified procedure uses spatial databases where missing AOD species values filling gaps using spatially neighboring data prediction considering spatial pixel pattern AOD clear sky measurements Liang, Daniels, Biancardi, & Zhou (2023b). In contrast to previous studies that rely on multi-day temporal inputs, this study utilizes hourly AOD data within the same day, which provides stronger temporal continuity and more consistent atmospheric conditions. The pixel baseline formula is:

$$Aerosollv.1_{(x,y,z)} = \sum_{k=1}^{K} \alpha_k AOD_{Lv.2} + \left(EVIxWS\right) + LCL_{M1/2/3} \tag{5}$$

where the coefficient $\alpha_k$is determined by the vegetation index range classification:

$$\alpha k = \begin{cases} \alpha_1, VI < 0.02 \\ \alpha_2, 0.02 \le VI \le 0.06 \\ \alpha_3, 0.06 \le VI \le 0.10 \\ : \\ \alpha_K, 0.98 \le VI \le 1.00 \end{cases} \tag{6}$$

with interval spacing

$$\Delta VI = 0.04 \tag{7}$$

From equation (5), $AODlv.2_{(x,y,z)}$ is the species baseline values at point (x,y). *EVI* and *WS* (Wind Speed) is value at pixel no data AOD, *LCL* is Land Cover Learning from Merge System code categories. Our approach focuses on each species was constructed with structure-environment using ANN, the model was validated. Therefore, to prove our experience the potential modify maps can be used for

monitoring of species.

2.3.2 Enhanced Modify AOD Framework

Before detailing procedures for modifying spatial patterns from predictive multisensory integration, RCaS and RCoS capture more accurate image features as described below. The challenge of spatial changes, coupled with limited ground-level observations, severely constrains our ability to systematically monitor complex values. We developed a correction model using cross-validation design with randomized trials from ground-based studies to examine validity and potentially construct AOD of clear sky coverage. We then applied the best quality results from LCL for well-generalized matched pairs, achieved by controlling both particle size distribution and aerosol loading. To apply this model for data analysis control strategies, we first constructed a spatial database containing prediction aerosol information (Figure 3). We developed a coordinate system notation RCoS to denote gap-filling methods to facilitate spatial position Aerosollv.1, operating in two processes. First, we estimated AOD spatial databases from species baseline neighborhood. Second, we included summation of all land cover across Taiwan's surface considering the contribution of aerosol observations within specific vegetation conditions, the empirical formulation can be improved by introducing Vegetation Index (VI) interval set using MLP. Reconstruct Cloud Sky (RCoS) can be expressed as:

$$AOD_{RCOS} = \sum_{k=1}^{K} \alpha_k AOD_{EVI_P}(Aerosollv.1_{(x,y,z)})(X_1) \tag{8}$$

The reconstructed aerosol optical depth under clear sky (RCaS) conditions can be formulated as:

$$AOD_{RCAS} = \sum_{k=1}^{K} \alpha_k AOD_{EVI_P}(X_2) \tag{9}$$

where, the X_1 is AODlv.2, DOY, T, RH, WD. X_2 is AODlv.2, LU, DOY, T, RH, WS, WD.

Indicator description $EVIP_{25,k}$,$P_{50,k}$,$P_{75,k}$ = 25th percentile, median, and 75th percentile of ground-based AOD for the k-th vegetation index range. The coefficient α_k is determined by the vegetation index range classification.

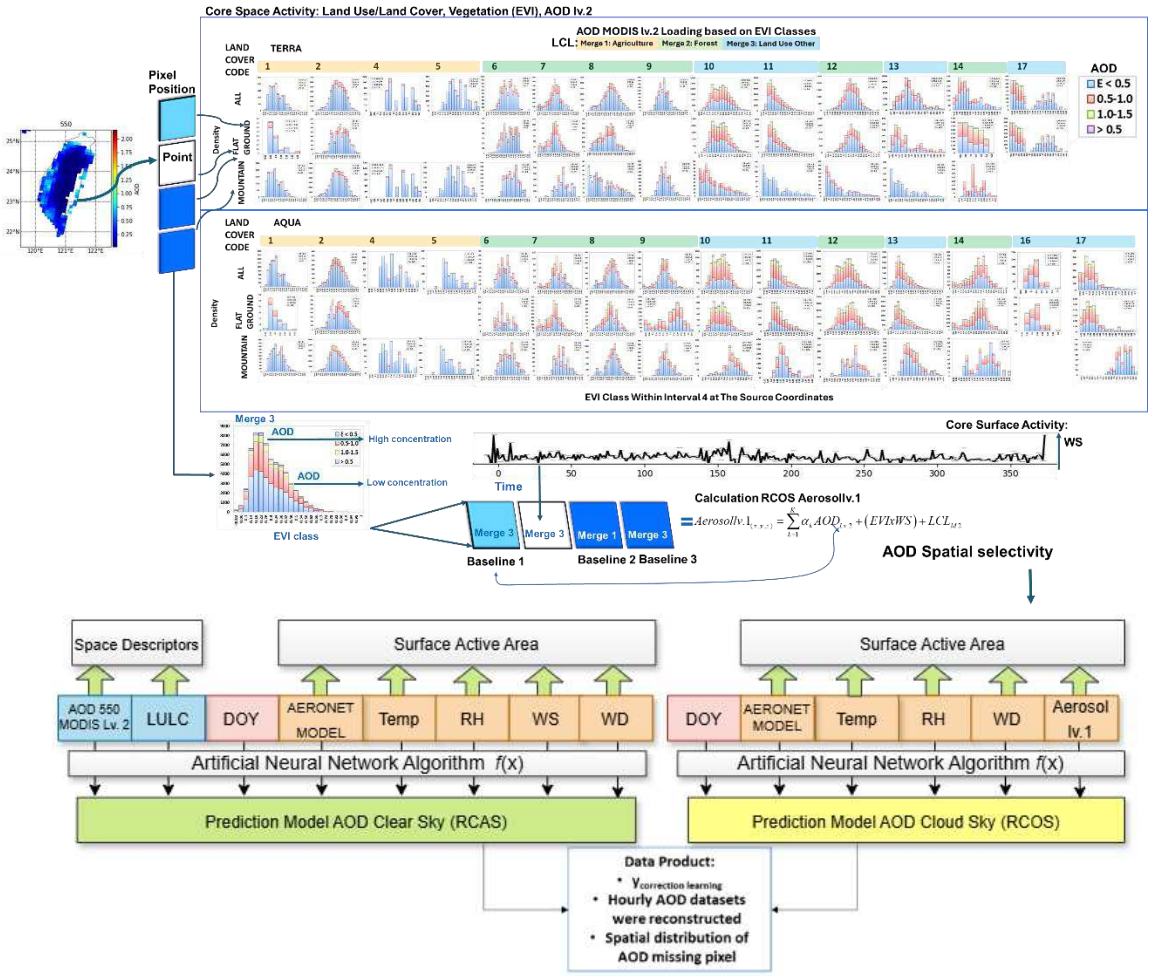


Fig. 3 Top, the flow information from inputs (AOD lv.2 loading from raw data with each EVI, Land Use) to outputs (Aerosollv. 1). The input indicator of the core space and surface at single time point. At the bottom, modified AOD framework guided by provided coordinate structure without ground-based measurements, facilitated by machine learning ANN f(x).

Direct aerosol management from area changes transforms pollutants with exactly the same AOD characteristics. These models offer effective approaches for monitoring and predicting aerosols using readily accessible data sources while maintaining acceptable accuracy levels, thereby overcoming limitations posed by data gaps and enabling more robust environmental assessments. This approach remains consistent with real-world applications, storing combined ground reference data with land use information through data transformation functions. Machine learning models present promising solutions to mitigate substantial constraints posed by different species estimates.

The process begins with clear sky conditions, using structure-based space descriptors to initialize AOD 550 MODIS and Land Use Land Cover data. Next, input surface active area data from real meteorological distributions. Finally, combining corrected AOD under clear sky and cloudy conditions to better construct AOD

information prediction.

2.4 Data Splitting and Performance Metrics

Model training used data from January 2018 to December 2022 (5 years). Independent temporal validation used the entire 2023–2024 dataset, which was withheld from training. This temporal split evaluates the model's ability to generalize to unseen years and provides evidence of temporal robustness beyond the training period. Model prediction accuracy was evaluated using the coefficient of determination (R^2), root mean square error (RMSE), and the Expected Error (EE) envelope, is defined as:

$$R^2 = \frac{\left(\sum_{t=1}^{n}(x_t - \bar{x})(y_t - \bar{y})\right)^2}{\sum_{t=1}^{n}(x_t - \bar{x})^2 \sum_{t=1}^{n}(y_t - \bar{y})^2} \tag{10}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(x_t - y_t)^2}{n}} \tag{11}$$

$$EE = \pm[0.05 + 0.15 \times AERONET\ AOD] \tag{12}$$

Where ȳ EE is quantitative evaluation of the AOD retrieval uncertainty. The Percentage Expected Error define confidence envelopes validated Deep Blue data against AERONET Sayer et al. (2013)**.** In equation (12), indicates a need to understand uncertainty the significant property for the retrieval MODIS AOD product. For first discussion of expected error retrieve aerosols MODIS, should be able to within 20-30% relates to vegetated and semivegetated land surfaces Y. J. Kaufman et al. (1997).

3 Results

3.1 ANN Correction Performance at 10 Ground-based Stations

The comprehensive corrected aerosol value framework in Figure 3 formulates AOD distribution across Earth's surface using mathematical equations 4. ANN describes evaluation processes for validating species retrieved from satellites related to ground-based measurements on an hourly basis. By identifying particular location ground-based features, we can correct aerosol optical properties even in areas without ground-based coverage for long-term monitoring using structural environment information. This method, known as reconstructed area AOD prediction, provides important advantages for aerosol model validation.

Our primary goal involved structure-environment predictor correction for all temporal MODIS AODlv.2 data collocated with AERONET, shown in Figures 4, 5 and 6 for 10 ground-based AOD stations. Figure 4 illustrates prediction results from ANN methods after correction at Lulin AQUA, where the proportion of forest area is highest, with 98% of retrievals falling within EE, the lowest R^2 of 0.85, and RMSE of 0.08. These AODlv.3 ANN prediction results are reliable due to similar annual averages from 2018 to 2022: AERONET 0.062 and AODlv.2 0.063.

Moreover, the lowest correlations between TERRA AODlv.3 and AERONET were found in Xitun (urban area), with R^2 0.84, RMSE 0.08, and 75% of retrievals within EE. However, in significant rural ground-based areas, we observe similar

frequency signals for AODlv.3 above 1.0 at Cape Fuguei, Douliu, and Lulin. For Douliu, the small number of observations may result from land surface conditions at the school field site, making it difficult for TERRA to collect high-quality data due to coarse spatial resolution surface ground-based in school field becomes difficult for TERRA collecting high quality data considered to coarse spatial resolution.

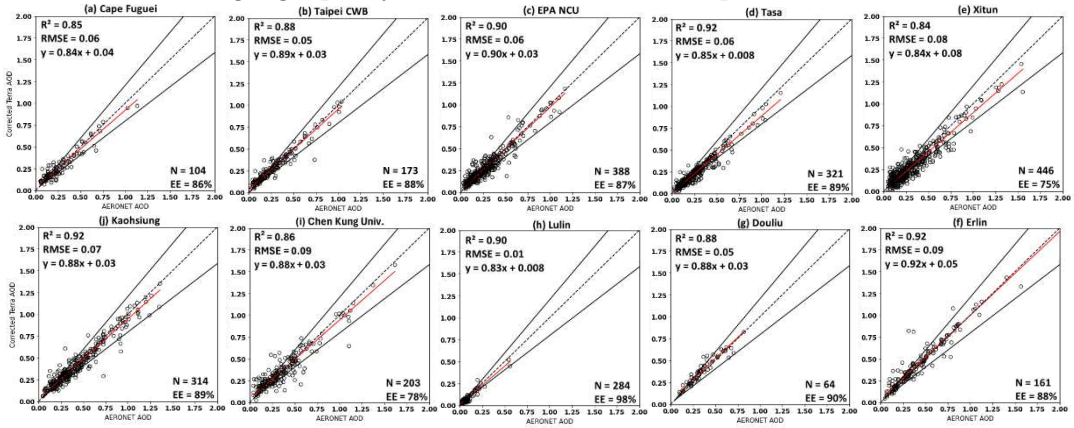


Fig. 4 Scatterplots of predicted ANN AODlv.3 TERRA against AERONET for datasets at each ground-based station. Dashed black lines represent 1:1 relationships, red lines show EE envelopes ± (0.05 + 15%) from January 2018 to December 2022. Statistical metrics include R^2, RMSE, and percentage within expected error bounds.

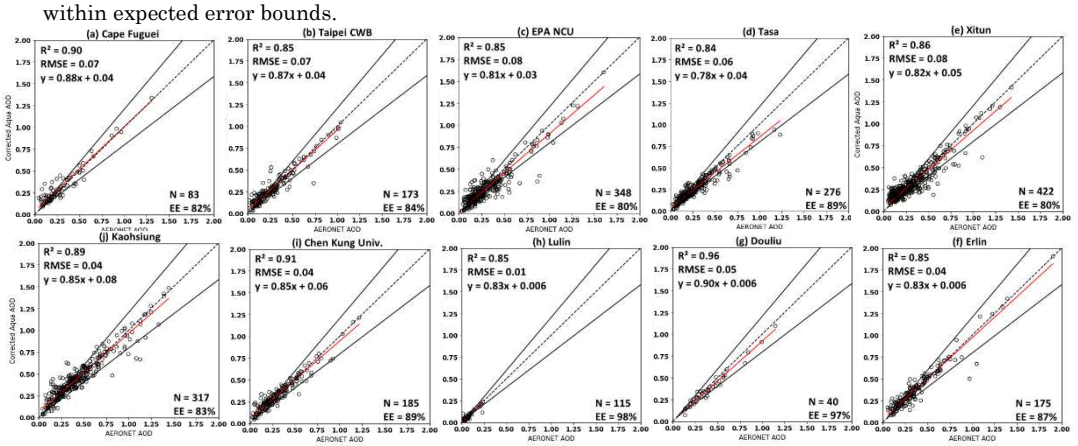
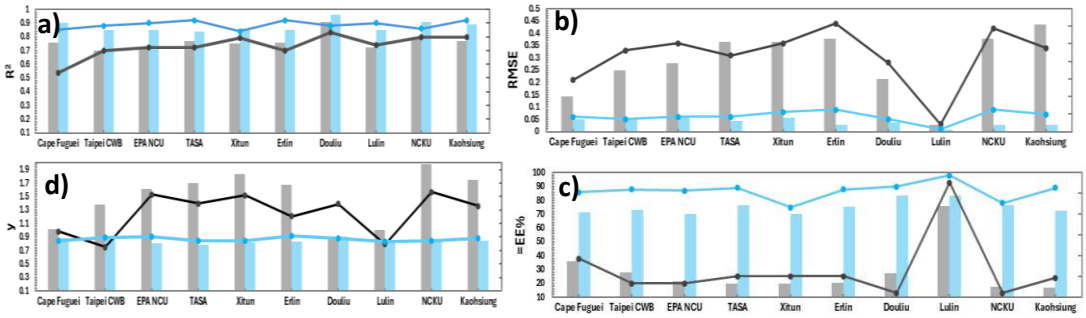


Fig 5. Scatterplots of predicted ANN AODlv.3 AQUA against AERONET for datasets at each ground-based station. Dashed black lines represent 1:1 relationships, red lines show EE envelopes ± (0.05 + 15%) from January 2018 to December 2022. Statistical metrics include R^2, RMSE, and percentage within expected error bounds.

Corrected AQUA shows better performance across 10 ground-based stations (Figure 5). In Kaohsiung, correlations (R^2 0.89, RMSE 0.04, 83% retrievals within EE) were lower than TERRA (R^2 0.92, RMSE 0.07, 89% retrievals within EE). High AODlv.3 variations occur at land surface areas with urban and built-up characteristics in Taipei CWB, EPA NCU, Xitun, TASA, NCKU, and Kaohsiung, plus Erlin in rural school areas. The first ground-based station in Taiwan's capital city shows low AODlv.3 variation, influenced by surrounding rural and mountain areas, similar to EPA NCU. TASA, Xitun, and NCKU still exhibit high AODlv.3 variation

caused by buildings, roads, and human activities. Additionally, peak AOD AQUA overpass AERONET in Erlin was influenced by industrial areas and agriculture flat ground with a maximum value of 1.9.



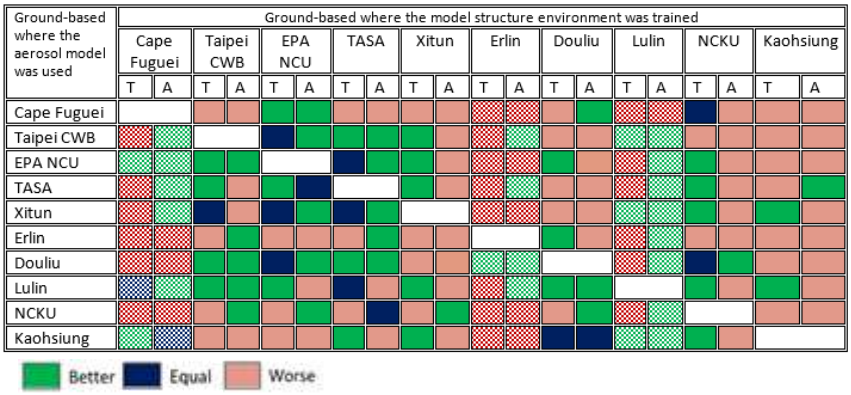
AERONET Taiwan

Fig 6. Lab scale studies: Performance at 10 Ground-based were collected from 2018 to 2022. The Quality Assurance (QA3) MODIS Terra (Black line) and Aqua (Grey bar). The Blue after correction is based on ANN and the validation statistics for (a and b) R^2 and the root mean square (RMS) error, (c and d) the agreement with AERONET within ±EE and slope.

In general, the relationship correlation data show the greatest stability through the time series and the highest quality of calibration statistic. As before, smaller errors are found for vegetated sites (11, 5) with a divergence first becoming apparent around Cape Fuguei, than arid sites (12, 13), and arid sites show a larger difference in comparison quality between calibration versions. All model slopes are less than 1, indicating minimal overestimation for high AOD values and minimal underestimation for low AOD values.

To construct accurate AODlv.3 across the region, Table 2 shows Land Cover Learning observations across different environments for evaluating AODlv.2 against other AERONET stations in Taiwan. Results were compared with original R^2, slope, and RMSE values for overall model performance validity. When TERRA overpass NCKU, the model environment was trained with model-based aerosol Cape Fuguei, resulting in nearly identical R^2 values (0.85 and 0.82 respectively). These ground-based stations have similar average AODlv.3 values: NCKU 0.3 and Cape Fuguei 0.31 respectively. Standard deviations at these two stations were 0.22 and 0.2 respectively. Additionally, structure-environment conditions near surfaces and aerosol absorption by vegetation contribute to these similarities.

Table 2. Enhanced performance of LCL between corrected MODIS AODlv.2 and AERONET in rural representative grid cells. Solid colors represent urban and built-up areas.



3.2 ANN Correction under Clear Sky Conditions over Taiwan

We combined dynamic meteorological observations across Taiwan with target uncertainty reduction to improve correction performance. Spatial data relationships with temporal features are important considerations. Key input data in Figure 7 were meteorological factors from hourly historical observations, with goals to predict AODlv.2 MODIS (MOD), AERONET (AER), Land Use Changes (LUCs), temperature (Temp), relative humidity (RH), wind speed (WS), wind direction (WD), and day of year (DOY). Figure 7 shows the correlation matrix among ANN input parameters for TERRA (N=65,799) and AQUA (N=32,391). Among all predictors, temperature showed the strongest correlation with relative humidity (r = 0.84/0.85 for TERRA/AQUA). This positive temperature-humidity correlation reflects Taiwan's dominant maritime influence: during summer monsoon months, warm and moist oceanic air masses are simultaneously advected over the island, producing co-occurring high temperature and high humidity conditions. EVI also showed notable correlation with AOD, indicating that vegetation structure influences aerosol loading patterns.

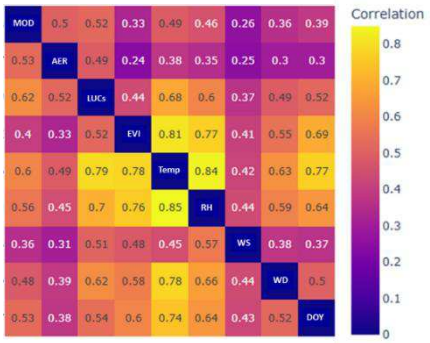


Fig. 7 Correlation coefficient matrix between ANN input parameters for TERRA (upper) and AQUA (lower) datasets, showing pairwise correlation coefficients and statistical significance levels.

To improve surface aerosol monitoring from twin MODIS sensors, we used EVI to monitor surface vegetation by capturing sensitivity in high biomass areas needed from canopy background signals Steven W Running et al. (2004). Table 3 presents aerosol sample data across MODIS land cover merge systems by EVI range. Table 4 summarizes EVI statistics and maximum AOD by station. Mountain ground-based Lulin covers the largest reconstructed area (14,000 km^2) with the highest median EVI of 0.385. The lowest median EVI (0.074) was found at NCKU (industrial/coastal), which also recorded the highest maximum AOD (3.5 for both TERRA and AQUA). This inverse EVI-AOD relationship is physically consistent: dense vegetation reduces direct aerosol emissions, enhances dry deposition, and suppresses boundary layer mixing that concentrates aerosols Y J Kaufman & Tanre, (1992).

Aerosols in the atmosphere can effect satellite-derived vegetation indices by altering the surface reflectance collected in observation ground-based (Figure 8 top) vs species baseline over Taiwan (Figure 8 bottom). The reconstruct AOD clear sky and cloud sky used dynamic variability of aerosol abundance is strongly associated with vegetation characteristics across different land cover types, make it more reliable than NDVI in areas with high atmospheric aerosol concentrations Sayer, A. M., Hsu, N. C., Bettenhausen, C., Jeong, J., & Meister (2015). In arid situations based on Table 3 elevated AOD levels across land cover types are found in arid environments.

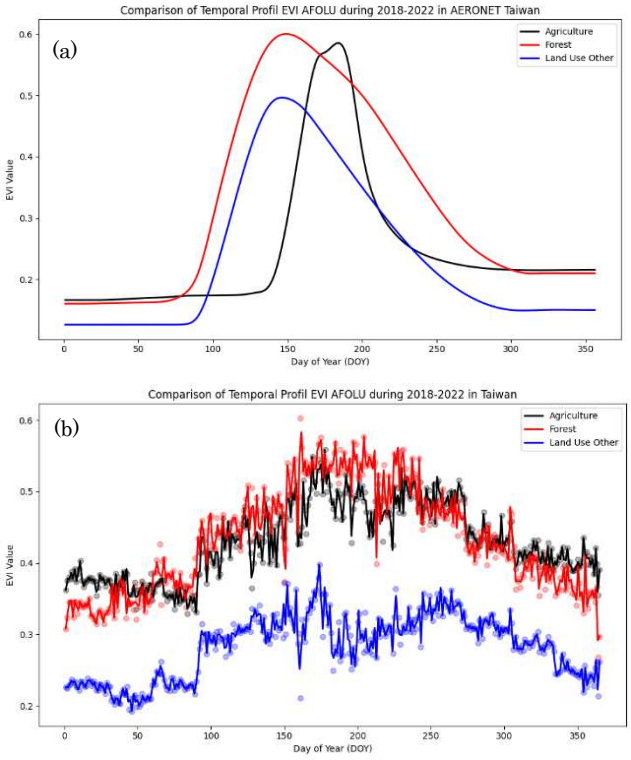


Fig. 8 Temporal Profile EVI by surface phenology AFOLU Classes at an (a) AERONET and (b) pixel-averaged across Taiwan in 2018–2024.

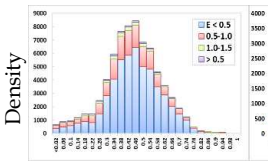
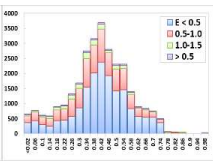
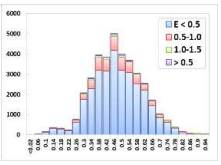
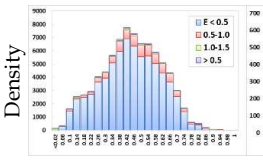
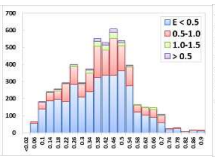
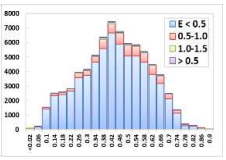
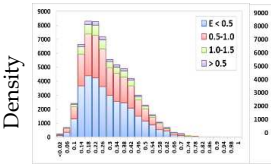
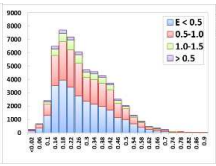
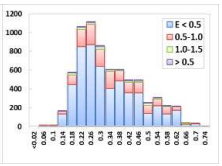
The first Figure 8 represents ground-based EVI measurements at a specific

location (AERONET ground-based), showing a smooth and well-defined seasonal pattern. EVI values increase sharply starting around DOY 100, peak between DOY 140–180, and then gradually decline toward the end of the year. Among land cover types, forest exhibits the highest peak EVI (~0.9), followed by agriculture (~0.85), while other land uses remain lower (~0.8). The curves are smooth, indicating reduced noise and strong seasonal vegetation dynamics at the local scale.

In contrast, the second Figure 8 shows spatially averaged EVI across Taiwan (pixel-based), which appears noisier and has a lower overall magnitude. The seasonal trend is still visible—EVI increases from early year, peaks around mid-year (DOY ~150–220), and declines afterward—but the variation is less pronounced. Peak values are significantly lower (~0.55 for forest, ~0.5 for agriculture, and ~0.35 for other land use), reflecting spatial averaging effects. The fluctuations in the curves indicate heterogeneity across different regions, land cover types, and environmental conditions.

Overall, the ground-based EVI shows stronger, clearer seasonal signals and higher peak values, while the pixel-averaged EVI over Taiwan is smoother in trend but noisier in detail and lower in magnitude, due to spatial heterogeneity and mixed land cover effects.

Table 3. Sample data pixel of Aerosol species for TERRA observed over merge system Land Cover Types via Enhanced Vegetation Index range.

Merge System and Land Cover Code	All Sample	Flat Ground	Mountain
Merge 1 Agriculture 6, 7 : Shrublands 8, 9 : Savannas 12 : Croplands 14 : Natural Vegetation			
Merge 2 Forests 1 : Evergreen needleleaf 2 : Evergreen broadleaf 4 : Deciduous needleleaf 5 : Deciduous broadleaf			
Merge 3 Land Use Other 10 : Grasslands 11 : Permanent Wetlands 13 : Urban and built-up 17 : Water			

Density

EVI Range

TERRA AQUA

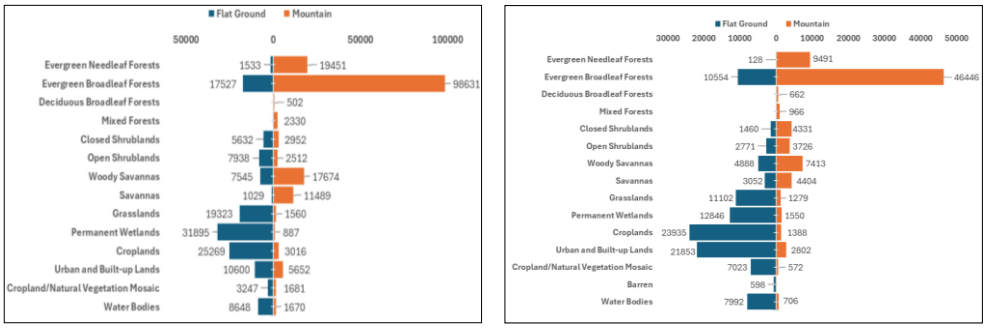


Fig. 9 Comparison datasets of AOD Clear Sky TERRA and AQUA across Taiwan in 2018-2024. The Land cover variability is displayed at the right side.

Mountain ground-based Lulin covers the largest area (14,000 km²) with vegetated spaces, reaching high median EVI of 0.385 based on data Figure 9. The second largest median EVI values are Cape Fuguei and EPA NCU at 0.34 and 0.311, respectively. The lowest median EVI (0.074) was associated with potentially highest AOD at NCKU in industrial and coastal areas, indicating larger direct aerosol emissions than other sectors Lin et al. (2007).

Under aerosol clear-sky, index values below 0 generally indicate non-cloud pixels. In mountainous regions, these negative values are predominantly related to the land cover types Merge A and Merge LO. Conversely, in Merge F negative index values are mainly observed over flat ground. When aerosol effects are present, the negative index values in Merge LO appear to be influenced by water, while those in Merge A associated with bare soil or agricultural water bodies.

Table 4 Enhanced 25th percentile, median and 75th percentile of the EVI values (2018–2024) and maximum MODIS AOD values for the models corresponding to each ground-based also distance sea boundaries, together with reconstructed areas for Terra satellites in Taiwan.

Model AEROSOL	EVI 25th percentile	EVI Median	EVI 75th percentile	AOD TERRA max	AOD AQUA max	Reconstructed Areas (10x10km)	Merge System LCTypes
NCKU	0.0686	0.0747	0.0895	3.5	3.5	1,500	1,2
EPA NCU	0.2557	0.3115	0.3713	3.462	3.4	1,500	1,2,3
Xitun	0.1662	0.2096	0.2741	3	2.85	1,500	1,2
Kaohsiung	0.1789	0.2175	0.2572	2.978	3.372	2,000	1,2,3
Erlin	0.1844	0.2087	0.3298	2.284	2.514	3,600	1,2,3
Taipei CWB	0.2312	0.2593	0.2886	2.622	2.297	1,000	1,2
Lulin	0.343	0.3855	0.4383	2.411	2.218	14,000	2,3
TASA	0.1796	0.1975	0.2329	1.62	1.753	1,600	1,2,3
Cape Fuguei	0.2771	0.3403	0.3965	1.586	1.289	4,000	1,2,3
Douliu	0.1393	0.1518	0.2071	1.452	1.453	800	1,2
Total						31,500	

Figure 10 shows full-dataset ANN model performance. Variation between TERRA and AQUA (panels a and b) is attributed to diurnal changes in atmospheric conditions at their different overpass times 10:30 vs 13:30 local time Yoram J

Kaufman et al. (2000), Levy, Remer, & Dubovik (2007). Scatter plots of Corrected AOD TERRA lv. 2 at 550nm and AQUA corresponding AERONET Model Aerosol. The black dashed are 1:1 line. The difference data of AOD (f) TERRA and (g) AQUA follow a Gaussian pattern before correction (red) show a wider spread. After correction (blue), the distribution narrows, suggesting improve consistency in the aerosol measurements. Panels (g, h, i, j) TERRA and (k, l, m, n) AQUA show independent 2023–2024 temporal validation: the model maintains comparable performance to the training period, confirming temporal generalizability. These four plots show the performance of satellite AOD correction compared with AERONET Model aerosol data. Validation Strategy: Training–Testing based on same location and time but different dates. In this study, the validation of the trained model is conducted using a spatiotemporal consistency approach, where the training and testing datasets share the same observation locations and local overpass time, but are separated by different acquisition dates.

The evaluation of satellite-derived Aerosol Optical Depth (AOD) from (d) TERRA and (c) AQUA against ground-based observations demonstrates substantial improvement after the correction process (blue line). The original Level 2 data (red line) exhibit considerable dispersion and systematic deviation compared to ground model aerosol (black line), particularly during high AOD events.

For TERRA, the original dataset shows moderate correlation ($R^2 = 0.60$), with a regression slope (y) of 1.31, indicating overestimation, and a relatively high RMSE of 0.33. After correction, the performance improves significantly, with R^2 increasing to 0.95, slope reduced to 0.94 (closer to unity), and RMSE decreasing sharply to 0.05. This indicates a substantial reduction in both systematic and random errors.

Similarly, for AQUA, the original data exhibit lower agreement with ground observations ($R^2 = 0.56$), a slope of 0.87, and a high RMSE of 0.46, reflecting large variability and underestimation tendencies. Following correction, the agreement improves markedly, with R^2 increasing to 0.94, slope approaching unity (0.92), and RMSE reduced to 0.07. Overall, the correction method effectively enhances the consistency between satellite-derived and ground-based AOD measurements for both TERRA and AQUA. The significant increase in correlation, reduction in RMSE, and adjustment of regression slopes toward unity confirm that the corrected datasets provide a more accurate representation of aerosol variability and are more suitable for atmospheric and climate-related analyses.

In addition, the fraction of data falling within the expected error (EE) envelope shows a substantial improvement after correction. For both TERRA and AQUA datasets, only about 30% of the original Level 2 retrievals fall within the expected error range, indicating low reliability and large deviation from ground-based observations.

After applying the correction, this proportion increases significantly to approximately 90%, demonstrating that the majority of the corrected AOD values lie within the acceptable uncertainty bounds. This indicates a strong enhancement in retrieval accuracy and consistency, as the corrected data not only reduce systematic bias and random error but also meet the expected uncertainty criteria commonly used in satellite validation studies. The increase from 30% to 90% within the expected error range confirms that the correction method substantially improves the quality and reliability of both TERRA and AQUA AOD products, making them more robust

for scientific and operational applications.

The table 6 presents the Artificial Neural Network (ANN) hyperparameter configuration applied to ground-based, TERRA, and AQUA datasets. A consistent network structure was used, comprising one hidden layer with 128 neurons and a single output node. Activation functions include ReLU-linear combinations in the first layer, sigmoid in the hidden layer, and linear activation in the output layer. Dropout rates were tuned for each dataset (0.01–0.88) to mitigate overfitting, while the Adam optimizer was employed for all experiments. Training epochs varied depending on data complexity, ranging from 150–1300 epochs for ground-based data and up to ~3750 epochs for satellite datasets. Model training was performed on a GPU environment (Google Colab, ~600 compute units), resulting in efficient computation times of approximately 2–16 ms for ground-based data and 9–28 ms for satellite data.

Overall, the ANN framework maintains a consistent architecture while adapting hyperparameters to optimize performance across different data sources.

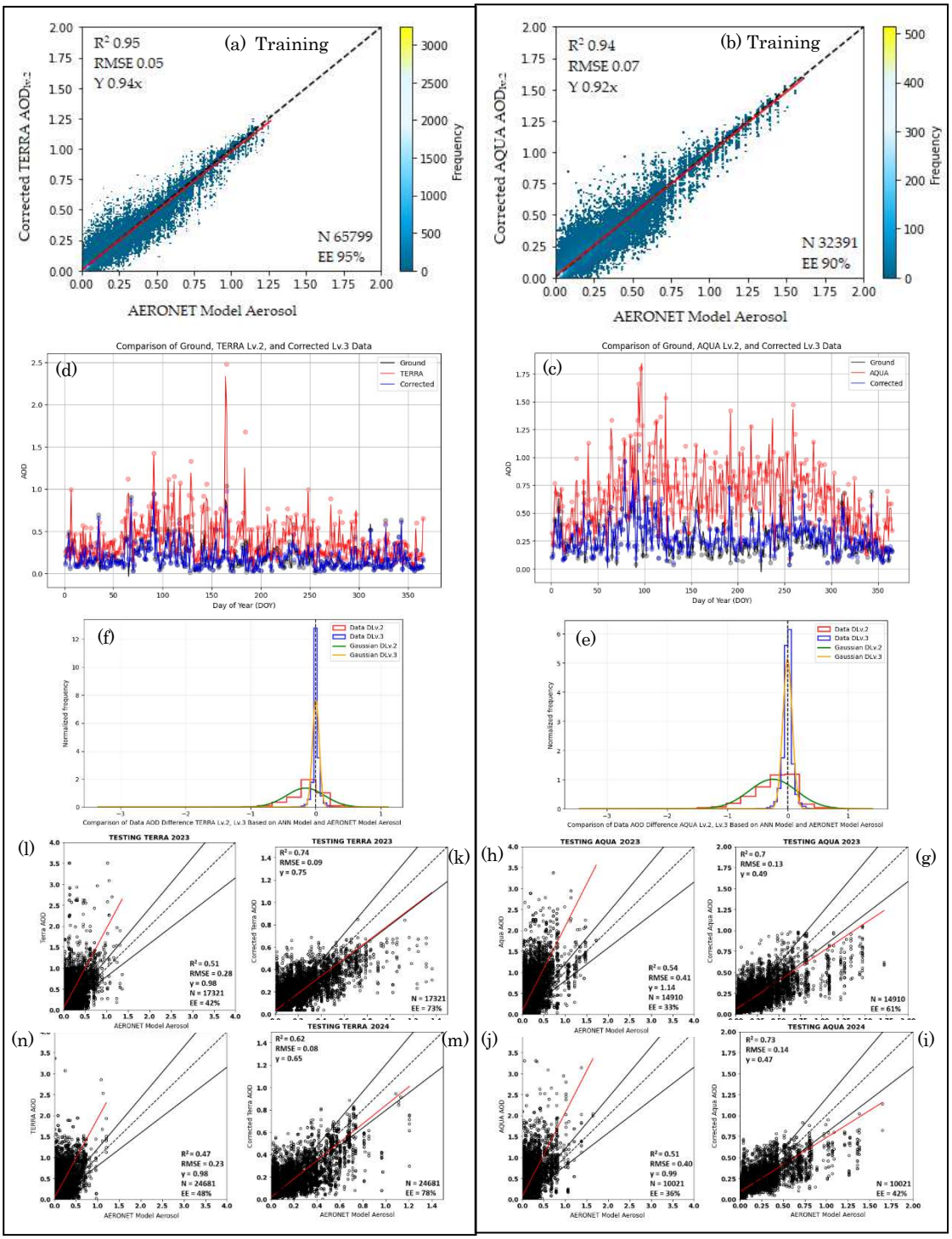

(a) Training
R² 0.95
RMSE 0.05
Y 0.94x
N 65799
EE 95%
Corrected TERRA AOD
AERONET Model Aerosol
Frequency
(b) Training
R² 0.94
RMSE 0.07
Y 0.92x
N 32391
EE 90%
Corrected AQUA AOD
AERONET Model Aerosol
Frequency
Comparison of Ground, TERRA Lv.2, and Corrected Lv.3 Data
(d)
Ground
TERRA
Corrected
AOD
Day of Year (DOY)
Comparison of Ground, AQUA Lv.2, and Corrected Lv.3 Data
(c)
Ground
AQUA
Corrected
AOD
Day of Year (DOY)
(f)
Data DLv.2
Data DLv.3
Gaussian DLv.2
Gaussian DLv.3
Normalized frequency
Comparison of Data AOD Difference TERRA Lv.2, Lv.3 Based on ANN Model and AERONET Model Aerosol
(e)
Data DLv.2
Data DLv.3
Gaussian DLv.2
Gaussian DLv.3
Normalized frequency
Comparison of Data AOD Difference AQUA Lv.2, Lv.3 Based on ANN Model and AERONET Model Aerosol
(l)
TESTING TERRA 2023
R² = 0.51
RMSE = 0.28
y = 0.98
N = 17321
EE = 42%
Terra AOD
AERONET Model Aerosol
(k)
TESTING TERRA 2023
R² = 0.74
RMSE = 0.09
y = 0.75
N = 17321
EE = 73%
Corrected Terra AOD
(h)
TESTING AQUA 2023
R² = 0.54
RMSE = 0.41
y = 1.14
N = 14910
EE = 33%
Aqua AOD
AERONET Model Aerosol
(g)
TESTING AQUA 2023
R² = 0.7
RMSE = 0.13
y = 0.49
N = 14910
EE = 61%
Corrected Aqua AOD
(n)
TESTING TERRA 2024
R² = 0.47
RMSE = 0.23
y = 0.98
N = 24681
EE = 48%
TERRA AOD
AERONET Model Aerosol
(m)
TESTING TERRA 2024
R² = 0.62
RMSE = 0.08
y = 0.65
N = 24681
EE = 78%
Corrected Terra AOD
AERONET Model Aerosol
(j)
R² = 0.51
RMSE = 0.40
y = 0.99
N = 10021
EE = 36%
AQUA AOD
AERONET Model Aerosol
(i)
TESTING AQUA 2024
R² = 0.73
RMSE = 0.14
y = 0.47
N = 10021
EE = 42%
Corrected Aqua AOD
AERONET Model Aerosol

Fig. 10 Scatterplots of ANN AOD against AERONET Model Aerosol for (a) Terra and (b) Aqua, including statistical metrics. The comparison mean and difference data MODIS AODlv. 2 is shown in red. Data AODlv.3 shown in Blue for (c, e) TERRA and (d, f) AQUA. Performing testing validation based on training model twin MODIS in 2028-2022, TERRA (g) and AQUA (h) in 2023 and 2024 respectively.

Table 5 Statistics of TERRA and AQUA Comparison for Training Data based on ANN Model (2018-2022)

No.	Input	R^2	Y	RMSE	=% EE	+% EE	-% EE	Mean	Std. Dev
1	TERRA Lv.2	0.6	1.31	0.33	39	52	7	0.31	0.33
	Lv. 3 Corrected	0.95	0.94	0.05	95	3	2	0.15	0.19
2	AQUA Lv2	0.56	0.87	0.46	30	60	10	0.48	0.43
	Lv. 3 Corrected	0.94	0.92	0.07	90	5	5	0.24	0.23

Table 6 Artificial Neural Network Hyperparameters GPU 600 compute units (Colab)

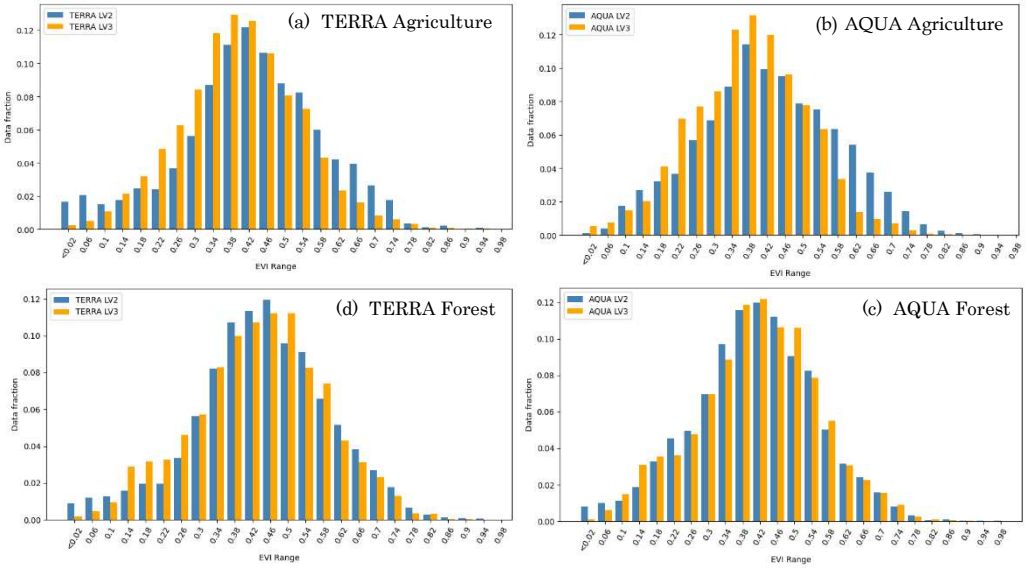
	Input Indicators					
Activation/Kernel	Ground-based		TERRA		AQUA	
Activation Function	Training	Cross Validation	Training	Testing	Training	Testing
Layer 1/100/uniform	ReLu	ReLu/ Linear	ReLu		ReLu	
Layer 2/128/uniform	Sigmoid	Sigmoid/ Linear	Sigmoid		Sigmoid	
Output/1/uniform	Linear	Linear	Linear		Linear	
Dropout range	0.01,0.5	0.01,0.88	0.01,0.05		0.01	
Optimizer	Adam					
Epoch	150-700	200-1300	3750		3500	
Time	02-16ms		09-28ms			

Table 7 Statistics of TERRA and AQUA Comparison for Testing Data based on ANN Model (2023-2024)

No	Input	N^2		R^2		Y		RMSE		=% EE		+% EE		-% EE		Mean		Std. Dev	
		2023	2024	2023	2024	2023	2024	2023	2024	2023	2024	2023	2024	2023	2024	2023	2024	2023	2024
1	TERRA Lv.2	17321	24681	0.51	0.47	0.98	0.98	0.28	0.23	42	48	52	43	6	8	0.25	0.19	0.29	0.24
	Lv. 3 Corrected			0.74	0.62	0.75	0.65	0.09	0.08	73	78	14	14	9	7	0.11	0.08	0.11	0.10
2	AQUA Lv2 N	14910	10021	0.54	0.51	1.14	0.99	0.41	0.40	33	36	62	60	5	4	0.41	0.39	0.40	0.39
	Lv. 3 Corrected			0.7	0.73	0.49	0.47	0.13	0.14	61	42	24	48	14	10	0.17	0.19	0.13	0.13

Figure 11 presents histograms of AOD twin distribution by EVI range. For (a. b) agricultural areas, both sensors exhibit a bell-shaped distribution of AOD fraction with a peak at moderate EVI values (0.3–0.5). At low EVI values (<0.14), AFOLU-type AODlv.3 values show a decreasing trend in TERRA while AQUA exhibits relatively small sample representation. Approximately 84% of total TERRA AOD values in agricultural areas fall below 0.5. Mean EVI decreases from approximately 0.45 in winter to 0.35 in spring, reflecting seasonal vegetation dynamics. In (c, d) forested areas, AODlv.2 is concentrated at higher EVI values. For heterogeneous land cover types such as (e, f) urban surfaces, permanent wetlands, and grasslands, the aerosol fraction is distributed toward lower to moderate EVI values.

Figure 11 shows acceptable twin MODIS ANN model performance when TERRA (a) AODlv.3 equals 1.25 with AERONET model aerosol, and when twin AODlv.3 values less than 0.25 equal AERONET model aerosol. AOD overestimation at low values can result from land use activities such as developed areas, road density, and vegetation restoration Lin et al. (2007). To address underestimation effects on AOD distribution, we applied expected error (EE) concepts for Deep Blue over land, predicted as ±(0.05+15%) for regions with low confidence, implying uncertainty Hsu et al. (2013), Sayer et al. (2014). By performing testing validation consistent location and ground-based reference data, the variation is attributed to diurnal changes in atmospheric conditons due different satellite overpass times.



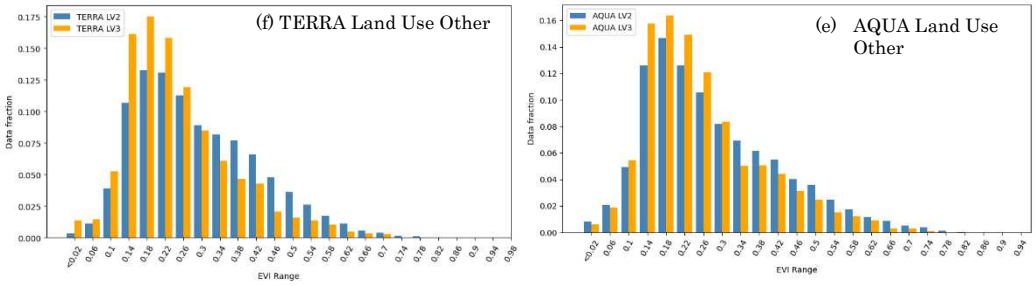


Fig 11. Histogram of the MODISlv.2 and MODISlv. 3 corrected values for AFOLU merge system of the dates and locations at 550nm by AOD inventories for <0.05-1.0 class in EVI range over Taiwan.

The distribution of aerosol corrected fraction (y-axis) as a function of MODIS twin was succecfull analyzed using the Enhanced Vegetation Index range across different land cover types according to Lanzaco et al. (2017). For predictived agricultural areas, twin MODIS exhibit a bell-shaped distribution of aerosol fraction with a peak occuring at moderate EVI values (approximately 0.3-0.5). By diagnostics the range EVI data from <0.02 to 0.14 for AFOLU twin AODlv.3 values reduce except AQUA A and TERRA LU relatively small sample representation. Around 84% of total AOD TERRA A values below 0.5 class, where 36% shows a decreasing trends compare to AQUA increase. In addition, all observations for 16% of the total shows a consistent trend decreased, approximately data variation from flat ground area in spring. Also, the mean EVI decreases from 0.45 in winter to 0.35 in spring. However, through interpretation of the data TERRA EVI range 0.14 to 0.26, AQUA 0.16 to 0.46 demonstrated AODlv. 3 increased compared to AOD lv. 2. Both, TERRA AQUA AODlv. 3 lower than AODlv. 2 dropping from 0.5.

In forest AOD lv. 2 MODIS twin is more concentrated at higher EVI values, reflecting the dense vegetation structure typically associated with forest cover. The TERRA data present a relatively symmetric distribution, whereas AQUA observations show slightly sharper peak. By examining the data land use and other, the aerosol fraction tends to be distributed toward lower to moderate EVI values, which correspond to sparsely vegetated surface or heterogeneous land cover types such as urban surfaces, permanen wetlands, grasslands and water.

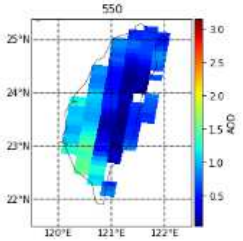
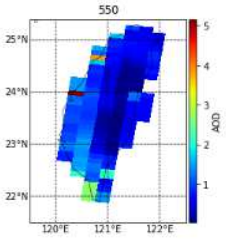
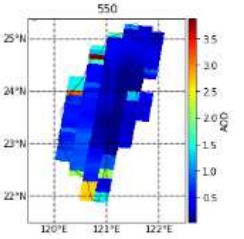
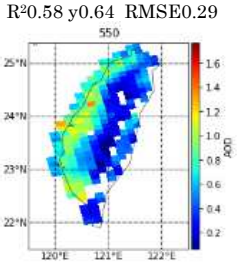
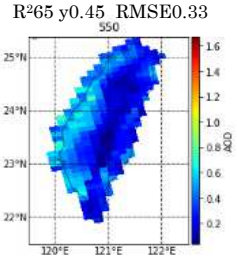
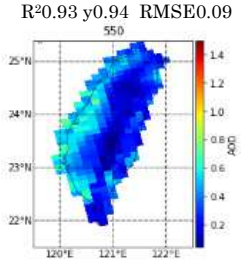
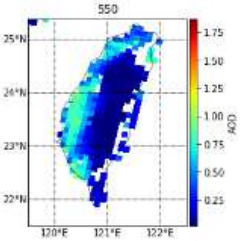
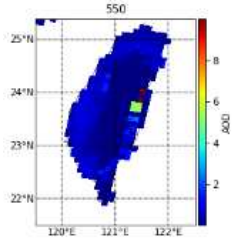
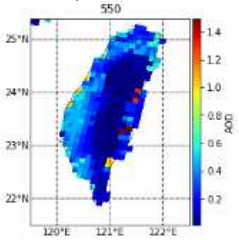
3.3 Spatial Analysis of Real-time Hourly based on ANN Reconstruction

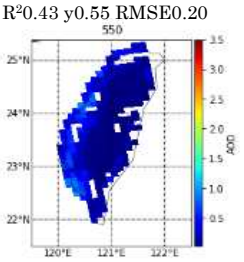
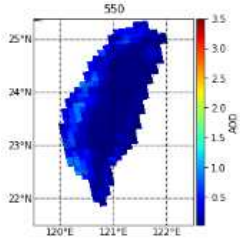
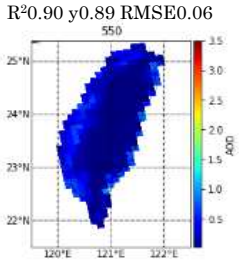
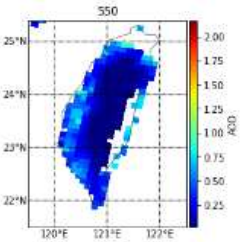
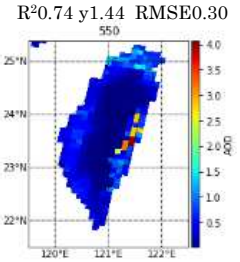
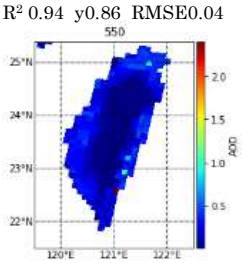
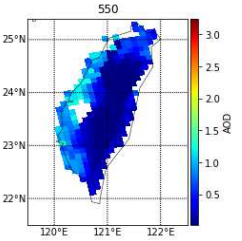
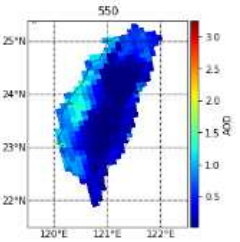
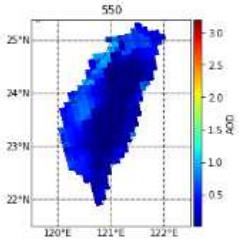
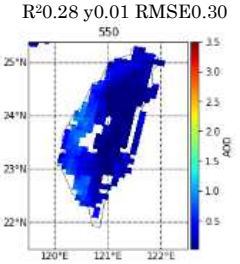
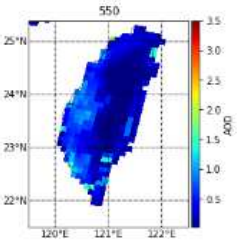
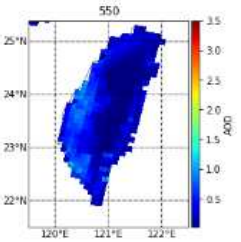
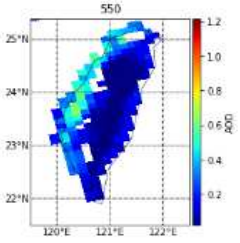
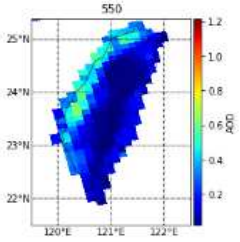
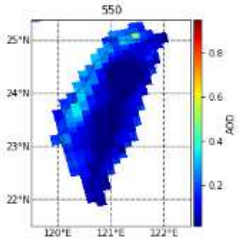
We conducted comprehensive performance analysis across different operational spacecraft platforms. AQUA observations are less dense than TERRA across all years and observation sites, attributed to generally higher cloud coverage in afternoons (Table 8 MOD) compared to mornings (Table 8 MYD)) Yoram J Kaufman et al. (2000), Levy et al. (2007). For surface cover requirements, we resampled MCD12Q1 (Table 8) to 10 km spatial resolution across 5-year periods. AOD MODIS Lv.2 & Aerosol Lv.1 column shows detail improves visibility, still EE not correctness. RCaS and RCoS show improvement from real-time hourly AERONET model aerosol using ANN. We computed missing AOD (Aerosollv.1) effects on cloud cover for each 10×10 km pixel, combined with AODlv.2 in Column 4.

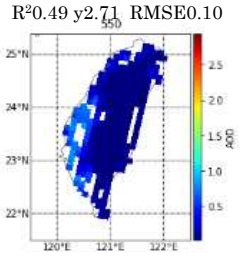
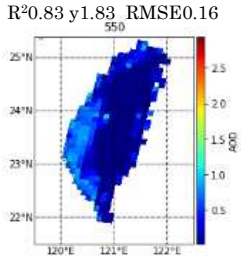
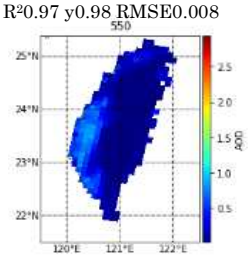
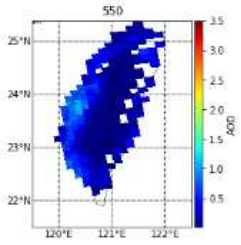
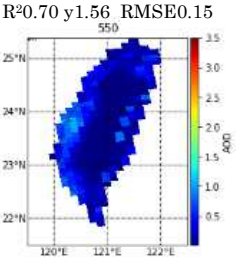
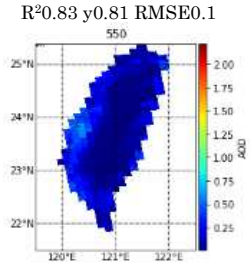
For cloud-free visualization and comparisons, ANN applied to the full Taiwan

domain demonstrates that vegetated and mountain areas show relatively low AODlv.3, except for northern urban corridors. Western Taiwan shows high AODlv.3, consistent with industrial activity and coastal aerosol sources. These spatial patterns are physically consistent with global AERONET validation results Ichoku et al. (2002): vegetated mountain environments such as Lulin in Taiwan are analogous to sites like Maricopa (Arizona, USA) and Solar Village (Saudi Arabia), where vegetated or open-terrain conditions produce systematically lower AOD compared to industrial coastal sites.

Table 8 Comprehensive spatial distribution analysis: Column 2: Spatial Distribution of AOD from MODIS lv. 2. Column 3: Spatial Distribution of AODlv.2 with Aerosol lv.1. Column 4: Spatial Distribution of modify AOD from RCaS and RCoS from March 2018 (DOY 077), November 2019 (314), January 2020 (031), February 2021 (032), October 2022 (284).

Year	Day of Year	AOD MODIS Lv.2	AOD MODIS Lv.2 & Aerosol Lv.1	RCaS & RCoS
2018	MOD 077 N 162 MODlv.2 135 Alv.1 27	$R^2$0.57 y0.97 RMSE0.35  -%EE11 =%EE53 +%EE36	$R^2$71 y0.94 RMSE0.54  -%EE11 =%EE36 +%EE53	$R^2$0.94 y0.97 RMSE0.07  -%EE5 =%EE94 +%EE1
	MYD 077 N 283 MYDlv.2 233 Alv.1 50	$R^2$0.58 y0.64 RMSE0.29  -%EE51 =%EE34 +%EE15	$R^2$65 y0.45 RMSE0.33  -%EE9 =%EE30 +%EE61	$R^2$0.93 y0.94 RMSE0.09  -%EE16 =%EE82 +%EE2
2019	MOD 314 N 369 MODlv.2 304 Alv.1 65	$R^2$0.23 y0.40 RMSE0.37  -%EE4 =%EE23 +%EE73	$R^2$0.50 y0.83 RMSE0.31  -%EE12 =%EE31 +%EE57	$R^2$0.90 y0.77 RMSE0.10  -%EE5 =%EE85 +%EE10

	MYD 314 N 335 MYDlv.2 275 Alv.1 50	R²0.43 y0.55 RMSE0.20  -%EE25 =%EE29 +%EE46	R²0.67 y0.69 RMSE0.25  -%EE25 =%EE29 +%EE46	R²0.90 y0.89 RMSE0.06  -%EE6 =%EE83 +%EE11
2020	MOD 031 N 361 MODlv.2 310 Alv.1 53	R²0.28 y0.01 RMSE0.30  -%EE30 =%EE20 +%EE50	R²0.74 y1.44 RMSE0.30  -%EE2 =%EE38 +%EE60	R² 0.94 y0.86 RMSE0.04  -%EE1 =%EE96 +%EE3
	MYD 031 N 358 MYDlv.2 275 Alv.1 50	R²0.45 y1.23 RMSE0.30  -%EE17 =%EE21 +%EE61	R²0.65 y1.15 RMSE0.15  -%EE5 =%EE41 +%EE54	R²0.89 y0.93 RMSE0.07  -%EE2 =%EE96 +%EE2
2021	MOD 032 N 374 MODlv.2 300 Alv.1 74	R²0.28 y0.01 RMSE0.30  -%EE54 =%EE35 +%EE12	R²0.70 y1.23 RMSE0.23  -%EE5 =%EE41 +%EE54	R²0.85 y0.95 RMSE0.04  -%EE2 =%EE96 +%EE2
	MYD 032 N 280 MYDlv.2 230 Alv.1 50	R²0.45 y1.23 RMSE0.37  -%EE40 =%EE20 +%EE20	R² 0.65 y1.15 RMSE0.15  -%EE1 =%EE46 +%EE53	R²0.83 y0.95 RMSE0.07  -%EE2 =%EE80 +%EE18

2022	MOD 284 N 384 MODlv.2 230 Alv.1 50	R²0.49 y2.71 RMSE0.10  -%EE0 =%EE73 +%EE27	R²0.83 y1.83 RMSE0.16  -%EE1 =%EE63 +%EE36	R²0.97 y0.98 RMSE0.008  -%EE2 =%EE95 +%EE3
	MYD 284 N 290 MYDlv.2 232 Alv.1 58	R²0.47 y1.56 RMSE0.52  -%EE51 =%EE49 +%EE0	R²0.70 y1.56 RMSE0.15  -%EE18 =%EE45 +%EE37	R²0.83 y0.81 RMSE0.1  -%EE8 =%EE82 +%EE10

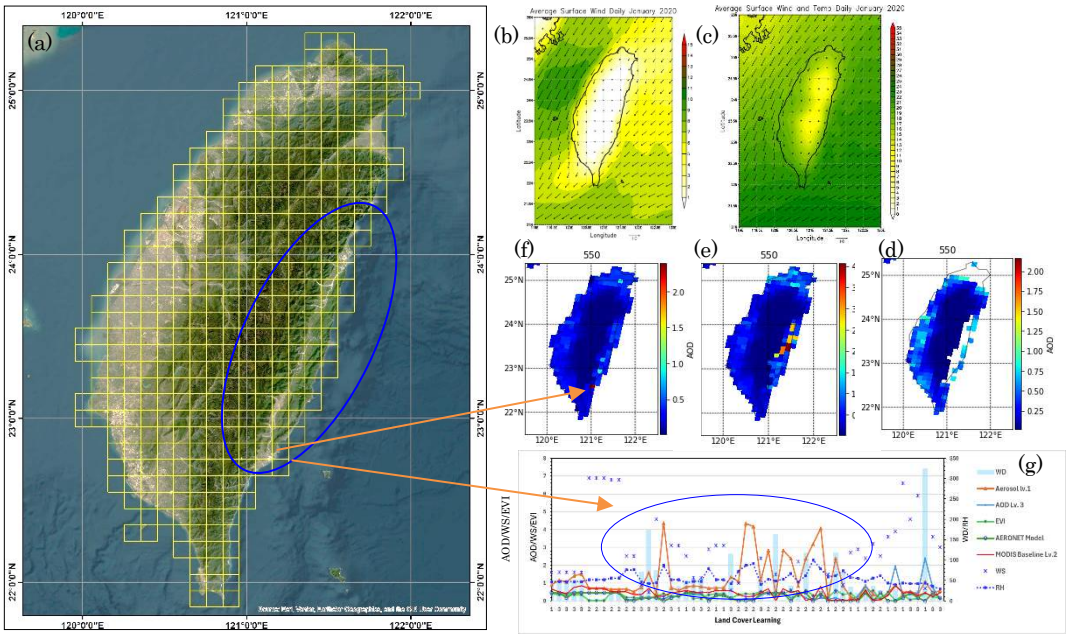


Fig. 13 Full scale studies: spatiotemporal analysis of atmospheric conditions and AOD retrieval over Taiwan during February 2020 under winter monsoon influence. Integrated analysis of (a) Wind, (b) Temperature, (d) AODlv. 2 and cloud contamination on satellite retrievals, (e) AOD lv. 2 and Aerosollv. 1, (f) RCaS and RCoS, (g) Graphic coupling with near-surface atmosphere.

RCOS approach (g): There is no Cloud without Aerosol. Extensively used because they describe realistically, based on comparison with field data and species baseline we found a decrease in EVI (green line) is generally associated with an increase in

AOD, indicating that reduced vegetation cover contributes to higher aerosol loading.

The effect of wind velocity on AOD is dual, as it can enhance aerosol concentrations through transport and resuspension, or reduce them via dispersion and dilution processes. Furthermore, land cover plays a fundamental role as the primary control of aerosol sources and emission intensity. The result indicate that cloud condition produce relatively high values, reconstruction outputs also showing elevated magnitudes. This pattern is especially evident when compared to non-vegetated land cover types. In addition, Aerosollv. 1 distribution is primarily concentrated in the eastern region, as illustrated in the graph.

The relative humidity (RH) is a measure of the amount of water vapor in the atmosphere and defined as the ratio between the partial water vapor pressure and the saturated water vapor pressure Seinfeld and Pandis, (2006). These RH outperforming the amount of wet particles in the atmosphere since the hygroscopic growth increases with increased RH (especially for RH > 90%) Gupta et al. (2006).

Natural sources release aerosols into the atmosphere independent of human activity and are responsible for ~90% of the total aerosol mass in the atmosphere NASA, (2014). In reality the diameter of a cloud condensation nuclei (CCN) is around 0.2 μm, a cloud droplet around 20 μm and water vapor at molecule size level (~ 0.2 nm). Then, the RH values were input into the graph (Figure 13), RH is not explicitly included in the formula, it was incorporated into the analysis. High RH typically above 70-80% indicates elevated atmospheric water vapor, leads to increased AOD.

Relative humidity (RH) plays a critical role in modulating aerosol optical depth (AOD) through its influence on aerosol hygroscopic growth. Under high RH conditions (typically above 70–80%), aerosol particles absorb water vapor, leading to an increase in particle size and scattering efficiency, which in turn enhances AOD values. This effect is particularly pronounced in regions influenced by moist air masses, such as during monsoonal conditions. However, elevated RH is also closely associated with cloud formation, introducing significant challenges for satellite-based AOD retrievals. The presence of clouds and sub-pixel cloud contamination can obscure or distort aerosol signals, often resulting in missing or biased AOD values in satellite products. Consequently, regions with high RH tend to exhibit both artificially elevated AOD near cloud edges due to hygroscopic growth and increased data gaps due to cloud screening. These combined effects highlight the inherent limitations of satellite retrievals under humid conditions and emphasize the need for reconstruction approaches that incorporate meteorological information to improve AOD estimation.

AOD overestimation at low values (AODlv.3 ≤ 0.25) can result from land use activities such as road density and vegetation restoration Lin et al. (2007). AOD underestimation at high values (AODlv.3 ≥ 1.25) reflects inherent regression toward the mean in ANN-based correction. Other physical processes, including subsurface dynamics and cloud interactions, remain challenging to model due to limited observations. Combined evaluation of AODlv.3 quantity and quality datasets should account for spatial and temporal variabilities. Land use features follow their own variability patterns, which may be separate from natural and/or anthropogenic emission fluctuations. In practical terms, however, they tend to be closely interlinked. ANN provides continuous and detailed aerosol distribution to fill

observation gaps in areas no data AOD.

4 Discussion

4.1 Model Performance in Context

The ANN-based AOD correction achieved R^2 = 0.95 and 0.94 for TERRA and AQUA respectively, which is comparable to or exceeds results from similar machine learning studies. Lanzaco et al. (2017) reported R^2 improvements of 2–79% using ANN and SVR for MODIS AOD correction in South America, with up to 91% of retrievals within expected error bounds in their best-case scenario. Our LCL framework achieves 91% within EE across cross-validated station pairs, confirming that transferring aerosol knowledge between stations with similar EVI and land cover characteristics is a viable approach to spatial gap-filling. Lary et al. (2009) showed that ML-based bias correction of MODIS AOD reveals systematic links between retrieval bias and surface type — consistent with our finding that urban stations (Xitun, NCKU) exhibit lower ANN performance than rural or coastal stations. Kang et al. (2022) reported R^2 = 0.89–0.93 for ML-based AOD retrieval over East Asia, comparable to our results in a similar regional aerosol environment.

The independent temporal validation on 2023–2024 data (Figure 10 g–m) provides critical evidence of model generalizability beyond the training period. This multi-year temporal validation step is rarely demonstrated in the satellite AOD correction literature, where most studies validate only within the same data period. The consistent performance across both periods indicates that the ANN has captured robust seasonal and land-cover-driven aerosol signals rather than overfitting to specific inter-annual variability.

4.2 Role of EVI in the Land Cover Learning Framework

EVI proved to be an effective auxiliary indicator for the LCL framework. The inverse relationship between EVI median and maximum AOD (Table 4) — NCKU showing the lowest EVI (0.074) and highest AOD max (3.5) while Lulin shows the highest EVI (0.385) and lower AOD max (2.4) — is physically consistent with the role of dense vegetation in aerosol dry deposition, boundary layer suppression of transport, and direct emission reduction Y J Kaufman & Tanre (1992). Huete et al. (2002) demonstrated that EVI maintains sensitivity in high-biomass regions where NDVI saturates, which is particularly important for the Lulin mountain station with dense forest canopy. The combination of EVI and MCD12Q1 land cover classification Friedl et al. (2010), Sulla-Menashe et al. (2019) allowed the LCL framework to reliably identify station pairs with similar aerosol regimes, enabling effective spatial knowledge transfer for gap-filling over Taiwan's 31,500 km² reconstructed domain.

4.3 Limitations and Future Work

Several limitations should be acknowledged. First, the full-Taiwan spatial real-time hourly validation is demonstrated for a single representative date (Table 8). Future work should present seasonal average maps to comprehensively characterize reconstruction performance. Second, the LCL framework currently applies pairwise

station transfer; a generalized approach using EVI and land cover type as direct continuous predictors could extend gap-filling simultaneously to all unmonitored pixels. Third, this study uses MODIS Collection 6 Deep Blue (QA=3) Hsu et al. (2013), which was superseded by Collection 6.1 in 2023; future updates should assess whether the newer collection Sayer et al. (2014) further improves ANN correction. Finally, while this study focuses on Taiwan's island geography, the methodology is transferable to other regions with limited ground-based monitoring, such as the Indonesian archipelago or Southeast Asian island systems.

5 Conclusion

This study presents a machine-learning approach for improving MODIS AODlv.2 retrieval over Taiwan by developing an ANN correction framework trained on 2018–2022 collocated MODIS–AERONET data across 10 ground-based stations and independently validated on 2023–2024 observations. At each training step, real meteorological data as inputs and model-based aerosol validation for missing pixel interactions. Advanced machine learning with parallel ANN simulations proved robust for AOD variations even with large datasets. We provided EVI values across different surface covers to study AOD satellite-ground relationships Reconstruct Clear Sky (RCaS) and Reconstruct Cloud Sky (RCoS) measurements.

The novel Land Cover Learning (LCL) framework enables AOD correction at locations cloud contamination monitoring by transferring model knowledge from species baseline, stations sharing similar EVI and land cover characteristics. The LCL cross-validation achieved 91% of retrievals within the expected error bounds, demonstrating effective spatial transferability. The combined clear-sky and cloudy-sky reconstruction produced the Level 3 merged AOD product (AODlv.3), providing spatially real-time hourly complete aerosol maps over the full 31,500 km² domain of Taiwan.

We developed deep generative models trained for matched ten ground-based stations and environment aerosol conditions, followed by cross-model applicability area corrected AOD in continuous space. Reliability of structured-environment reconstruction is fundamental to this work. Both ANN models consisting of structured-environment aerosol, day of year, Land Use Land Cover changes, and representing surface AERONET model aerosol were paired as neural network output targets. TERRA (65,799 pairs) achieved higher correlation (R^2 = 0.95) than AQUA (32,391 pairs) (R^2 = 0.94), with TERRA RMSE (0.05) smaller than AQUA (0.07). Independent temporal validation on 2023–2024 data confirmed that the trained model generalizes beyond the training period, with stable performance metrics. This temporal robustness, combined with the LCL spatial transferability, demonstrates that the proposed machine-learning approach can substantially reduce AOD data gaps in complex terrain environments with diverse land cover. The methodology holds potential for application in other regions with similar geographical diversity and limited ground-based monitoring networks, particularly in Southeast Asian island systems.

Funding:

This work was financially supported by the Taiwan National Science and Technology Council (NSTC), grant 114-2111-M-008-033 and NSTC 113-2111-M-008-029.

Materials availability

The data and associated products used in this study are available from: https://aeronet.gsfc.nasa.gov/ (accessed on 4 January 2024). The MODIS aerosol products retrieved from https://modis.gsfc.nasa.gov/ (accessed on 2 February 2024), the Enhance Vegetation Index freely downloaded from https://modis.gsfc.nasa.gov/data/dataprod/mod13.php (accessed on 4 February 2024). The Land Use Land Cover can be accessed at https://modis.gsfc.nasa.gov/data/dataprod/mod12.php (accessed on 19 April 2024), and the meteorology indicators are available at https://e-service.cwa.gov.tw/HistoryDataQuery/index.jsp (accessed on 6 Mei 2024).

Author Informations:

Authors and Affiliations

International Program of Environmental Science and Technology, National Central University, Taoyuan, Taiwan R.O.C.
Jelita Citrawati Jihan

Center for Space and Remote Sensing Research, National Central University, Taoyuan, Taiwan R.O.C.
Jelita Citrawati Jihan

Center for Space and Remote Sensing Research, National Central University, Taoyuan, Taiwan R.O.C.
Tang-Huang Lin

Center for Astronautical Physics and Engineering, National Central University, Taoyuan, Taiwan R.O.C.
Tang-Huang Lin

Geoinformation Research Center, Thai Nguyen University of Agriculture and Forestry, Thai Nguyen, Vietnam
Nguyen Van Hieu

Author Contributions:

Conceptualization, J.C.J. and T.-H.L.; methodology, J.C.J.; software, J.C.J. and N.V.H.; validation, J.C.J. and T.-H.L.; formal analysis, J.C.J.; investigation, J.C.J.;

resources, T.-H.L.; data curation, J.C.J.; writing original draft preparation, J.C.J.; writing review and editing, T.-H.L.; visualization, J.C.J.; supervision, T.-H.L.; project administration, T.-H.L. All authors have read and agreed to the published version of the manuscript.

Acknowledgements. The authors gratefully acknowledge the National Aeronautics and Space Administration (NASA) for providing MODIS aerosol, enhanced vegetation index (EVI), and land use/land cover products, as well as AERONET aerosol optical depth and extinction profile measurements, and further thank the Climate Observation Data Inquire Service (CODiS)-Central Weather Administration (CWA) of Taiwan database for meteorological data, in addition to the editors and reviewers for their constructive comments and valuable efforts in improving this manuscript.

Declarations

Conflict of Interest

The authors declare no competing interest.

Appendix A Section title of first appendix

Table A1. The amount are from a image available inventories in Taiwan study periods 2018 to 2024 for satellite studies AOD retrieval TERRA at 01.45-03.35 and AQUA 04.10-06.10 UTC+8

TERRA Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
2018	47	43	50	46	45	45	47	47	41	42	45	45	543
2019	47	42	46	44	49	47	46	49	47	50	51	52	570
2020	34	29	30	26	23	29	29	24	25	33	26	28	336
2021	52	51	55	53	60	59	57	54	50	51	50	55	647
2022	54	55	48	45	50	54	56	51	43	31	51	51	589
2023	51	45	53	49	54	49	53	51	52	52	50	54	613
2024	51	46	54	48	51	48	48	51	46	54	47	51	595
Total													3893
AQUA Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
2018	53	46	52	54	53	50	49	53	51	52	46	52	611
2019	52	42	49	50	51	50	49	51	52	53	53	53	605
2020	31	31	31	26	23	28	31	14	29	28	29	23	324
2021	55	48	54	53	54	51	54	54	51	54	52	55	635
2022	54	51	52	27	58	56	52	54	52	53	52	53	614

2023	54	49	54	51	57	50	53	52	52	55	49	54	630
2024	52	50	48	50	49	49	53	52	52	53	50	54	612
Total													4031

References

Boucher, O. D., Randall, P., Artaxo, C., Bretherton, G., Feingold, P., Forster, V. M., & Zhang, X. Y. (2021). Climate Change 2021—The Physical Science Basis. In Chemistry International (Vol. 43). https://doi.org/10.1515/ci-2021-0407

Basheer, I. A., & Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. Journal of Microbiological Methods, 43(1), 3–31. https://doi.org/10.1016/S0167-7012(00)00201-3

Cohen, W. B., & Justice, C. O. (1999). Validating MODIS terrestrial ecology products: Linking in situ and satellite measurements. Remote Sensing of Environment, 70(1), 1–3. https://doi.org/10.1016/S0034-4257(99)00053-X

Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A., O'Neill, N. T., … Kinne, S. (1999). Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols. Journal of Geophysical Research Atmospheres, 104(D24), 31333–31349. https://doi.org/10.1029/1999JD900923

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sensing of Environment, 114(1), 168–182. https://doi.org/https://doi.org/10.1016/j.rse.2009.08.016

Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., … Lyapustin, A. I. (2019). Advancements in the Aerosol Robotic Network (AERONET) Version 3 database - Automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements. Atmospheric Measurement Techniques, 12(1), 169–209. https://doi.org/10.5194/amt-12-169-2019

P. Gupta, S. A. Christopher, J. Wang, R. Gehrig, Y. Lee, and N. Kumar. (2006). Satellite remote sensing of particulate matter and air quality assessment over global cities. Atmos. Environ., vol. 40, no. 30, pp. 5880–5892. https://doi.org/10.1016/j.atmosenv.2006.03.016

Hsu, N. C., Jeong, M. J., Bettenhausen, C., Sayer, A. M., Hansell, R., Seftor, C. S., Tsay, S. C. (2013). Enhanced Deep Blue aerosol retrieval algorithm: The second generation. Journal of Geophysical Research Atmospheres, 118(16), 9296–9315. https://doi.org/10.1002/jgrd.50712

Holben, B. N., Eck, T. F., Slutsker, I., Smirnov, A., Sinyuk, A., Schafer, J., … Dubovik, O. (2006). Aeronet's Version 2.0 quality assurance criteria. In S.-C. Tsay, T. Nakajima, R. P. Singh, & R. Sridharan (Eds.), Remote Sensing of the Atmosphere and Clouds (Vol. 6408, p. 64080Q). https://doi.org/10.1117/12.706524

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment, 83(1), 195–213.

https://doi.org/10.1016/S0034-4257(02)00096-2

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. Neural Networks, 2(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Ichoku, C., Chu, D. A., Mattoo, S., Kaufman, Y. J., Remer, L. A., Tanré, D., Holben, B. N. (2002). A spatio-temporal approach for global validation and analysis of MODIS aerosol products. Geophysical Research Letters, 29(12), MOD1-1-MOD1-4. https://doi.org/10.1029/2001GL013206

J. H. Seinfeld and S. N. Pandis, (1999). Atmospheric Chemistry and Physics: From Air Pollution to Climate Change By John H. Seinfeld (California Institute of Technology) and Spyros N. Pandis (Carnegie Mellon University). Wiley-VCH: New York. 1997. xxvii + 1326 pp. ISBN 0-471-17815-2., J. Am. Chem. Soc., vol. 121, no. 6, p. 1423.

IPCC, (2022). Land-climate interactions. In Climate Change and Land: An IPCC Special Report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems. https://doi.org/10.1017/9781009157988.004

Kang, Y., Kim, M., Kang, E., Cho, D., & Im, J. (2022). Improved retrievals of aerosol optical depth and fine mode fraction from GOCI geostationary satellite data using machine learning over East Asia. ISPRS Journal of Photogrammetry and Remote Sensing, 183, 253–268. https://doi.org/10.1016/j.isprsjprs.2021.11.016

Kaufman, Yoram J, Holben, B. N., Tanré, D., Slutsker, I., Smirnov, A., & Eck, T. F. (2000). Will aerosol measurements from Terra and Aqua Polar Orbiting satellites represent the daily aerosol abundance and properties? Geophysical Research Letters, 27(23), 3861–3864. https://doi.org/10.1029/2000GL011968

Kaufman, Y. J., Tanré, D., Remer, L. A., Vermote, E. F., Chu, A., & Holben, B. N. (1997). Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer. Journal of Geophysical Research Atmospheres, 102(14), 17051–17067. https://doi.org/10.1029/96jd03988

King, M. D., Kaufman, Y. J., Menzel, W. P., & Tanre, D. (1992). Remote sensing of cloud, aerosol, and water vapor properties from the moderate resolution imaging spectrometer (MODIS). IEEE Transactions on Geoscience and Remote Sensing, 30(1), 2–27. https://doi.org/10.1109/36.124212

Kaufman, Y J, & Tanre, D. (1992). Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. IEEE Transactions on Geoscience and Remote Sensing, 30(2), 261–270. https://doi.org/10.1109/36.134076

Levy, R. C., Remer, L. A., Shi, Y., & Kleidman, R. G. (2023). Characterizing Aerosol from Space with the MODerate-resolution Imaging Spectroradiometer (MODIS) on the Terra and Aqua Satellites. In H. Akimoto & H. Tanimoto (Eds.), Handbook of Air Quality and Climate Change (pp. 271–294). https://doi.org/10.1007/978-

981-15-2760-9_60

Lanzaco, B. L., Olcese, L. E., Palancar, G. G., & Toselli, B. M. (2017). An improved aerosol optical depth map based on machine-learning and modis data: Development and application in South America. Aerosol and Air Quality Research, 17(6), 1523–1536. https://doi.org/10.4209/aaqr.2016.11.0484

Lanzaco, B. L., Olcese, L. E., Palancar, G. G., & Toselli, B. M. (2016). A method to improve MODIS AOD values: Application to South America. Aerosol and Air Quality Research, 16(6), 1509–1522. https://doi.org/10.4209/aaqr.2015.05.0375

Lary, D. J., Remer, L. A., MacNeill, D., Roscoe, B., & Paradise, S. (2009). Machine learning and bias correction of MODIS aerosol optical depth. IEEE Geoscience and Remote Sensing Letters, 6(4), 694–698. https://doi.org/10.1109/LGRS.2009.2023605

Levy, R. C., Remer, L. A., & Dubovik, O. (2007). Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land. Journal of Geophysical Research Atmospheres, 112(13), 1–15. https://doi.org/10.1029/2006JD007815

Lin, C.-Y., Wang, Z., Chen, W.-N., Chang, S.-Y., Chou, C. C. K., Sugimoto, N., & Zhao, X. (2007). Long-range transport of Asian dust and air pollutants to Taiwan: observed evidence and model simulation. Atmospheric Chemistry and Physics, 7(2), 423–434. https://doi.org/10.5194/acp-7-423-2007

Liang, L., Daniels, J., Biancardi, M., & Zhou, Y. (2023a). Reconstructing aerosol optical depth using spatiotemporal Long Short-Term Memory convolutional autoencoder. Scientific Data, 10(1), 842. https://doi.org/10.1038/s41597-023-02696-w

Liang, L., Daniels, J., Biancardi, M., & Zhou, Y. (2023b). Reconstructing aerosol optical depth using spatiotemporal Long Short-Term Memory convolutional autoencoder. 1–12. https://doi.org/10.1038/s41597-023-02696-w

Liu, G.-R., Chen, A. J., Lin, T.-H., & Kuo, T.-H. (2002). Applying SPOT data to estimate the aerosol optical depth and air quality. Environmental Modelling & Software, 17(1), 3–9. https://doi.org/10.1016/S1364-8152(01)00047-0

NASA, 2014. URL http://earthobservatory.nasa.gov/.

Owili, P. O., Lien, W. H., Muga, M. A., & Lin, T. H. (2017). The associations between types of ambient PM2.5 and under-five and maternal mortality in Africa. International Journal of Environmental Research and Public Health, 14(4), 1–20. https://doi.org/10.3390/ijerph14040359

Pöschl, U. (2005). Atmospheric aerosols: Composition, transformation, climate and health effects. Angewandte Chemie - International Edition, 44(46), 7520–7540. https://doi.org/10.1002/anie.200501122

Remer, L. A., Kaufman, Y. J., Tanré, D., Mattoo, S., Chu, D. A., Martins, J. V., Holben, B. N. (2005). The MODIS aerosol algorithm, products, and validation. Journal of the Atmospheric Sciences, 62(4), 947–973. https://doi.org/10.1175/JAS3385.1

Running, Steven W, Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., & Hashimoto, H. (2004). A Continuous Satellite-Derived Measure of Global Terrestrial Primary Production. BioScience, 54(6), 547–560. https://doi.org/10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.CO;2

Running, S. W., Baldocchi, D. D., Turner, D. P., Gower, S. T., Bakwin, P. S., & Hibbard, K. A. (1999). A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and EOS satellite data. Remote Sensing of Environment, 70(1), 108–127. https://doi.org/10.1016/S0034-4257(99)00061-9

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533–536. https://doi.org/10.1038/323533a0

Sulla-Menashe, D., Gray, J. M., Abercrombie, S. P., & Friedl, M. A. (2019). Hierarchical mapping of annual global land cover 2001 to present: The MODIS Collection 6 Land Cover product. Remote Sensing of Environment, 222, 183–194. https://doi.org/10.1016/j.rse.2018.12.013

Sayer, A. M., Hsu, N. C., Bettenhausen, C., Jeong, J., & Meister, G. (2015). Effect of MODIS Terra radiometric calibration improvements on Collection 6 Deep Blue aerosol products: Validation and Terra/Aqua consistency. Journal of Geophysical Research: Atmospheres, 12(120(23)), 157–174. https://doi.org/10.1002/2015JD023878

Sayer, A. M., Munchak, L., Hsu, N. C., Levy, R. C., Bettenhausen, C., & Jeong, M. J. (2014). MODIS Collection 6 aerosol products Comparison between Aqua s e-Deep Blue Dark Target. Journal of Geophysical Research: Atmospheres, 119. https://doi.org/10.1002/2014JD022453

Sayer, A. M., Hsu, N. C., Bettenhausen, C., & Jeong, M. J. (2013). Validation and uncertainty estimates for MODIS Collection 6 "deep Blue" aerosol data. Journal of Geophysical Research Atmospheres, 118(14), 7864–7872. https://doi.org/10.1002/jgrd.50600

Yang, J., & Hu, M. (2018). Filling the missing data gaps of daily MODIS AOD using spatiotemporal interpolation. Science of The Total Environment, 633, 677–683. https://doi.org/10.1016/j.scitotenv.2018.03.202