https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

A blue-band surface residual correction method for land aerosol optical depth retrieval from DQ-1 WSI

Jingxuan Zhao[1], Yujing Wang[4,5], Hui Chen[2], Mengjie Zhang[1], Lili Wang[3], Xingchuan Yang[1], Shuo Dong[2], Wenji Zhao[1] and Pengfei Ma[2]

[1]College of Resource Environment and Tourism, Capital Normal University, Beijing 100048, China
[2]Ministry of Ecology and Environment Center for Satellite Application on Ecology and Environment State Environmental Protection Key Laboratory of Satellite Remote Sensing, Beijing, 100094, China
[3]School of Geology and Geomatics, Tianjin Chengjian University, Tianjin, 300384, China
[4]Henan Institute of Geo-Environment Exploration Co., Ltd. Zhengzhou 450000, China
[5]Henan Institute of Geo-Environment Planning & Design Co., Ltd. Zhengzhou 450000, China

Correspondence to: Wenji Zhao (4973@cnu.edu.cn) and Pengfei Ma (mapengfei@secmep.cn)

Abstract. Surface reflectance prior uncertainty remains a major source of error in land aerosol optical depth (AOD) retrieval, particularly over complex surfaces and for newly launched satellite sensors lacking long-term observations for constructing stable surface priors. Here, we develop a hybrid AOD retrieval framework for the DQ-1 Wide Swath Imager (WSI) that integrates a static surface reflectance prior, machine-learning-based blue-band surface residual correction, and a physically based lookup-table (LUT) inversion. A multi-band static surface prior is first constructed from quality-controlled clear-sky observations under low-aerosol-loading conditions. Rather than directly predicting AOD, a random forest model is used to estimate the residual between the static and reference surface terms at 0.443 μm from the static blue-band prior, top-of-atmosphere reflectance, NDVI, viewing geometry, and elevation. The corrected blue-band surface term is then propagated through the conventional LUT retrieval, thereby retaining the physical aerosol inversion framework while using machine learning only to mitigate surface-prior mismatch. MCD19A2 AOD is used exclusively in the offline stage to support the construction of LUT-derived reference surface terms and for independent cross-product spatial consistency assessment, whereas AERONET observations provide the primary ground-based validation. Relative to the static-prior LUT retrieval, residual correction reduces the AERONET-based RMSE from 0.256 to 0.107 and mean bias error from 0.187 to 0.059, while increasing the fraction of retrievals within the expected error envelope from 14.8% to 55.6%. Evaluation on a held-out scene further demonstrates that the model generalizes the relationship between static-prior error and the LUT-derived reference surface term, reducing the blue-band surface-term MAE from 0.0088 to 0.0044. On dates excluded from model training and internal validation, the retrieved DQ-1 WSI AOD also exhibits strong regional spatial consistency with MCD19A2 across 200,507 collocated pixels (R = 0.863, RMSE = 0.131, MAE = 0.093, MBE = 0.012, and regression slope = 0.99). These results indicate that the performance improvement primarily arises from alleviating blue-band surface-prior mismatch within the physical LUT inversion rather than from direct empirical fitting of AOD. The proposed framework therefore provides a physically interpretable strategy for improving land AOD retrieval from DQ-1 WSI over springtime complex surfaces and

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

offers a transferable solution for new wide-swath sensors for which robust long-term surface reflectance priors are not yet available.

Keywords: DaQi-1 satellite; Wide Swath Imager (WSI); aerosol optical depth (AOD); surface reflectance; random forest; residual correction

1 Introduction

Aerosol optical depth (AOD) is an important parameter for characterizing atmospheric aerosol loading and its climatic and environmental effects (Zou et al., 2025; Luo et al., 2024a), and it has been widely used in air-quality monitoring and regional climate studies (Intergovernmental Panel On Climate Change (Ipcc), 2023; Wei et al., 2021). Owing to their wide spatial coverage and continuous observation capability, satellite remote sensing techniques have become a primary means of obtaining AOD at large scales (Wei et al., 2020; Yan et al., 2026). However, over land surfaces, especially in urban and urban–rural transitional areas, strong spectral heterogeneity and spatial non-uniformity make it difficult to effectively separate surface signals from aerosol scattering signals, which remains one of the major challenges for high-accuracy AOD retrieval.

Although physically based retrieval algorithms represented by the MODIS Dark Target (DT) (Kaufman et al., 1997), Deep Blue (DB) (Hsu et al., 2013; Levy et al., 2013), and MAIAC (Lyapustin et al., 2018) products have become relatively mature, errors in surface reflectance estimation are still a key factor limiting retrieval accuracy over complex surfaces (Bilal et al., 2013). This is particularly true under low-AOD conditions, where even a small bias in the blue-band surface prior may propagate through the radiative transfer chain and be amplified into a substantial AOD retrieval error. In addition, bidirectional reflectance distribution function (BRDF) effects and seasonal phenological changes may further weaken the applicability of static surface priors (Lucht et al., 2000). Therefore, for new remote-sensing sensors with limited temporal accumulation, it remains challenging to construct surface background constraints that are well matched to actual observation conditions (Yan et al., 2025).

In recent years, machine-learning methods have been increasingly used for AOD retrieval and related error correction and have shown strong capability in fitting nonlinear relationships in complex scenes (Luo et al., 2024b; Tian et al., 2022; Wei et al., 2024; Chen et al., 2025). However, end-to-end data-driven methods that directly target AOD usually incorporate radiative transfer constraints only implicitly, and their physical interpretability and spatiotemporal generalizability still require further examination (Yan et al., 2024; Peng et al., 2026; Yang et al., 2025). In contrast, physics-guided hybrid methods use machine learning to correct selected parameters or error terms while retaining the physical retrieval framework, thereby improving retrieval accuracy while better preserving physical consistency (Dong et al., 2024; Lemmouchi et al., 2023; Man et al., 2024; Yan et al., 2022; Yan et al., 2023). Among these methods, tree-based models such as random forest (RF) are well suited to handling multi-source heterogeneous features and representing nonlinear relationships, and have been shown to be effective in various remote-sensing retrieval tasks (Liang et al., 2022). For new wide-swath sensors, constraining machine learning to

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

surface-term error compensation, rather than using it to directly replace the physics-based AOD retrieval, provides a feasible way to improve retrieval stability over complex surfaces.

The DaQi-1 (DQ-1) atmospheric environment monitoring satellite carries the Wide Swath Imager (WSI), which provides relatively high spatial resolution and wide-swath coverage, offering new observational capability for regional aerosol monitoring (Dai et al., 2024; Wang et al., 2021). Previous studies have shown that DQ-1 WSI has the potential for land AOD retrieval (Wang et al., 2024b). However, as a new sensor with still-limited temporal accumulation, WSI faces several challenges over complex land surfaces. First, long-term and stable surface reflectance priors are not yet sufficiently available, and surface assumptions developed for mature sensors cannot be directly transferred. Second, clouds, thin clouds, warm haze, water, snow, and mixed pixels over complex land surfaces may affect the quality of analysis-ready data. In addition, the Beijing–Tianjin–Hebei region is characterized by highly heterogeneous urban–cropland–bare-land mosaics, which further increases the difficulty of surface reflectance characterization and stable AOD retrieval (Zang et al., 2023). Therefore, it is important to develop a retrieval-oriented workflow for DQ-1 WSI that integrates quality control, surface-prior constraints, physical retrieval, and error correction.

To address these challenges, this study develops a hybrid land AOD retrieval framework for DQ-1 WSI that combines machine-learning-based surface correction with a physically based lookup-table (LUT) inversion. A static multi-band surface reflectance prior is first constructed from quality-controlled low-AOD clear-sky observations. A random forest model is then used to correct the residual of the 0.443 μm surface term, with MCD19A2 AOD used only in the offline stage to constrain the LUT-derived reference surface term. The corrected blue-band surface term is subsequently introduced into the LUT retrieval to obtain AOD. Unlike direct machine-learning AOD retrieval, the proposed framework uses machine learning only to mitigate surface-prior mismatch while preserving the physical retrieval chain and its interpretability. The main contributions are threefold: (1) an integrated DQ-1 WSI land AOD retrieval workflow combining quality control, static surface-prior construction, surface residual correction, and LUT inversion; (2) a targeted correction of blue-band surface-prior errors to reduce retrieval bias over complex surfaces; and (3) a multi-level evaluation linking surface-term correction to final AOD performance using AERONET validation, held-out-scene tests, and independent consistency comparisons with MCD19A2. These analyses are designed to determine whether the retrieval improvement arises from alleviating surface-prior mismatch rather than from direct empirical fitting of AOD.

2. Study Area and Data Preprocessing

2.1 Study Area

The Beijing–Tianjin–Hebei (BTH) region was selected as the study area in this work. This region contains a mixture of urban areas, croplands, and bare land, resulting in strong surface heterogeneity, and therefore provides a representative test bed for examining the impact of surface reflectance estimation errors on satellite AOD retrieval. In addition, under the combined influence of anthropogenic activities, topography, and meteorological conditions, aerosol pollution in the BTH region exhibits

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

pronounced regional and episodic characteristics (Che et al., 2015; Tao et al., 2012). The wide range of AOD variability in this region provides representative scenes for evaluating the applicability of the DQ-1 WSI AOD retrieval algorithm under complex surface and different pollution conditions.

2.2 DQ-1 WSI Data

The DQ-1 atmospheric environment monitoring satellite was launched on 16 April 2022 and operates in a Sun-synchronous orbit at an altitude of approximately 700 km, with an overpass time of about 13:30 Beijing time over the study region (Wang et al., 2024b). The Wide Swath Imager (WSI) onboard DQ-1 provides wide-swath multispectral observations from the visible to the long-wave infrared region. The WSI has a swath width of approximately 2300 km and provides observations over approximately 0.415–12 μm, with nominal spatial resolutions of 75, 150, 300, and 600 m depending on the channel (Cao, 2023; Wang et al., 2024b). Pre-launch radiometric calibration showed uncertainties of approximately 3.2% for the visible and near-infrared channels and 4.6% for the shortwave-infrared channels, while the relative radiometric calibration accuracy of individual channels was better than 1.0% (Wang et al., 2024a). At 30% solar albedo, the signal-to-noise ratios are greater than 500 for the 0.443, 0.490, and 0.681 μm bands, greater than 200 for the 0.865 μm band, and greater than 110 for the 2.13 μm band (Wang et al., 2024a).

For the AOD retrieval developed in this study, the 0.443, 0.490, 0.681, 0.865, and 2.13 μm bands at 600 m spatial resolution were used to construct the main retrieval inputs. In addition, the 10.8 μm thermal infrared brightness temperature at 300 m resolution, together with solar/viewing geometry and elevation information from the navigation data, was used to support quality control and retrieval. The main bands used in this study and their applications are listed in Table 1.

Table 1. Main DQ-1 WSI bands used in this study and their applications.

Central Wavelength / μm	Spatial Resolution / m	Application
0.443	600	Core retrieval band; blue-band residual learning
0.490	600	Construction of the static surface reflectance prior
0.681	600	Surface-state characterization; NDVI calculation
0.865	600	Surface-state characterization; NDVI calculation
2.13	600	Water and complex-surface identification; quality control
10.8	300	Brightness-temperature retrieval; cloud contamination screening

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

2.3 Ground-Based Validation and Satellite Reference Products

AERONET (Aerosol Robotic Network) sun-photometer observations were used as the independent ground-based reference for validating the DQ-1 WSI AOD retrievals. Observations from stations in the Beijing–Tianjin–Hebei region were collected for the DQ-1 WSI overpass dates using the Version 3 Level 1.5 instantaneous product (Giles et al., 2019; Holben et al., 1998). The AOD at 550 nm was interpolated from the 500 nm AOD and the 440–675 nm Ångström exponent. To ensure a direct comparison between the static-prior LUT retrieval and the residual-corrected retrieval, the AERONET matchup analysis was conducted using only station–date samples for which both DQ-1 WSI retrievals were valid. Specifically, AERONET observations within the satellite overpass time window were averaged at each station to obtain the ground-based reference AOD. The satellite matchup value was then extracted as the mean AOD within a 3 × 3 pixel window centered on the station location for both retrievals. A station–date pair was retained only when both DQ-1 WSI retrievals were valid. This strategy ensures consistent sampling between the two retrievals, although it also constrains the final number of matchups.

The MODIS MAIAC MCD19A2 AOD product was also used to support low-AOD surface-prior construction and regional cross-product comparison. Previous studies have reported its good retrieval performance and spatial characterization over complex surfaces (Wei et al., 2019b). For dates corresponding to DQ-1 WSI observations, the Optical_Depth_055 layer was extracted, and high-quality pixels were selected according to the AOD_QA flag. The MCD19A2 data were then reprojected and clipped to the study domain to obtain same-day 550 nm AOD reference images. In this study, MCD19A2 was used mainly (1) to provide an external AOD constraint for static surface-prior construction under low-AOD conditions, (2) to provide a reference for regional spatial-structure comparison, and (3) to conduct hold-out-date cross-product consistency analysis. The independent accuracy assessment of the proposed retrieval remains based on AERONET observations. Level 1.5 data were used because fully quality-assured Level 2.0 observations were not consistently available for all DQ-1 WSI overpass dates at the time of this analysis.

2.4 DQ-1 WSI Preprocessing and Quality Control

To ensure observational consistency, the digital numbers (DNs) of the WSI visible-to-shortwave infrared bands were first converted to top-of-atmosphere (TOA) reflectance, and the thermal infrared band was converted to brightness temperature, using the official radiometric calibration coefficients provided by the China Centre for Resources Satellite Data and Application (CRESDA). Solar zenith angle, viewing zenith angle, and relative azimuth angle were extracted from the navigation data. TOA reflectance, geometric parameters, elevation, and other continuous variables were then resampled to a common spatial resolution of 600 m to generate spatially co-registered analysis-ready data.

A unified quality-control procedure was applied to reduce the influence of cloud contamination and anomalous surface targets on AOD retrieval. Water pixels were identified mainly using MNDWI (Xu, 2006), NDVI, and related spectral constraints, while snow pixels were detected using an NDSI threshold. For cloud screening, a conservative cloud–haze masking strategy was adopted for AOD retrieval, considering that thin clouds, cloud edges, and warm haze are easily missed over complex

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

surfaces. High-reflectance thick clouds were first identified using visible-band whiteness, while thin clouds and light haze were detected using blue-band enhancement. The 10.8 μm brightness temperature was further introduced to separate cold-cloud and warm-haze branches, and visible-band gradient and texture features were used to improve the detection of cloud edges and locally semi-transparent contaminated pixels. NDVI- and NDSI-based constraints were also applied to suppress false detections over bright vegetation and snow-like non-cloud targets. This conservative strategy was designed to remove suspicious contaminated pixels slightly more aggressively rather than allowing residual cloud or haze contamination to enter the AOD retrieval, thereby improving retrieval stability and physical consistency.

After the initial cloud–haze mask was obtained, morphological operations, including dilation, hole filling, and removal of small patches, were applied following existing cloud-masking approaches for optical satellite imagery (Qiu et al., 2019; Zhang et al., 2024). These operations improved the spatial completeness and stability of the mask. A unified QA mask was then generated using the priority order of snow, water, and cloud, and only clear-sky valid pixels were retained for subsequent surface reflectance estimation and AOD retrieval. Through these preprocessing and quality-control steps, DQ-1 WSI analysis-ready data with consistent radiometric units, unified spatial registration, and rigorous screening were obtained, providing reliable inputs for the subsequent LUT-based physical retrieval and residual correction. The specific thresholds, spectral tests, thermal constraints, and morphological operations used in the unified QA mask are summarized in Table S1 in the Supplement. The QA mask was encoded as 0 for clear sky, 1 for cloud or haze contamination, 2 for water, 3 for snow, and 255 for no data.

3. Methods

3.1 Overall Framework

The proposed framework consists of an offline preparation stage and an online AOD retrieval stage. The offline stage is used to construct the static surface prior and train the blue-band residual model, whereas the online stage performs AOD retrieval without requiring any external AOD product.

In the offline preparation stage, low-AOD clear-sky DQ-1 WSI reference scenes are first selected after quality control. For these scenes, collocated MCD19A2 AOD is used only as an external atmospheric constraint in the LUT inversion, together with WSI TOA reflectance, observation geometry, and elevation, to derive LUT-based reference surface terms. These reference surface terms are then used to construct the static multi-band surface reflectance prior. The difference between the LUT-derived reference blue-band surface term and the corresponding static blue-band prior is defined as the residual-learning target, and an RF model is trained to predict this residual from the static blue-band prior, blue-band TOA reflectance, NDVI, observation geometry, and elevation.

In the online retrieval stage, no MCD19A2 or other external AOD product is required. For each valid WSI pixel, the trained RF model predicts the blue-band surface residual, which is added to the static blue-band prior to obtain the residual-corrected blue-band surface term. The corrected surface term is then reintroduced into the LUT-based physical retrieval to derive AOD. Thus, machine learning is restricted to the surface-term correction step and does not directly estimate AOD.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

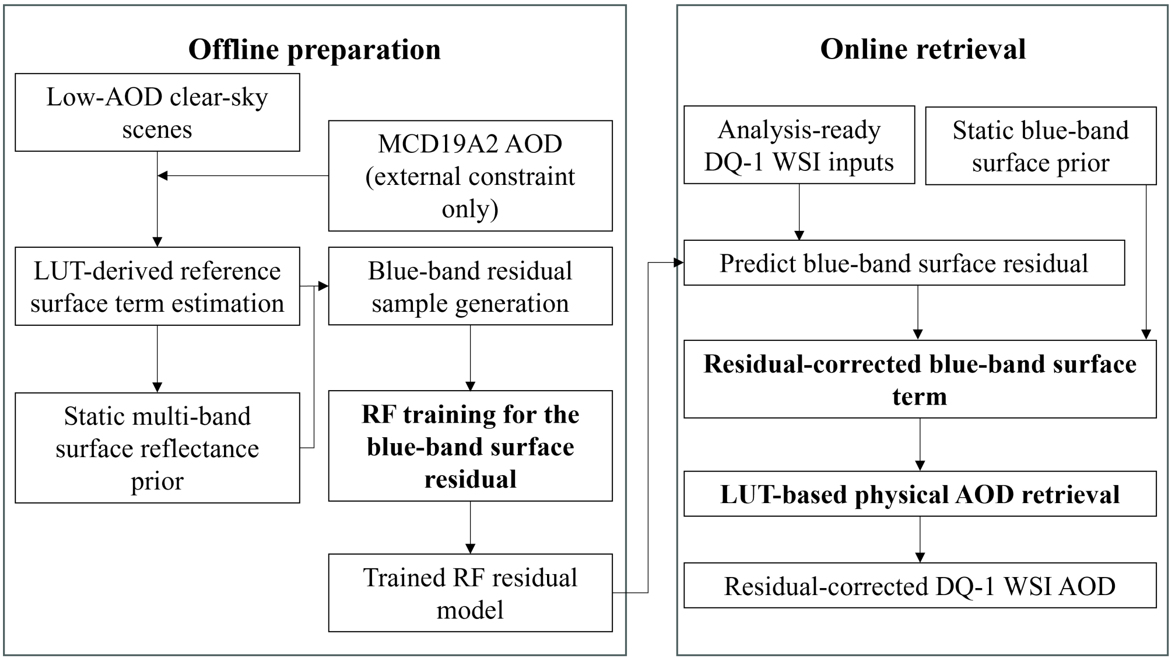

Figure 1. Framework of the proposed blue-band residual-correction method for DQ-1 WSI AOD retrieval. The left panel shows the offline preparation stage, in which MCD19A2 AOD is used as an external constraint for LUT-derived reference surface-term estimation under low-AOD clear-sky conditions, followed by static multi-band surface-prior construction, blue-band residual-sample generation, and RF model training. The right panel shows the online retrieval stage, in which no external AOD product is required; the trained RF model predicts the blue-band surface residual, and the resulting residual-corrected blue-band surface term is used for LUT-based physical AOD retrieval.

3.2 LUT Construction

Under the Lambertian surface assumption, this study used the 6S radiative transfer model (Kotchenova et al., 2006; Vermote et al., 1997) to simulate top-of-atmosphere (TOA) reflectance, and AOD was retrieved by minimizing the residual between observed and simulated reflectance. To improve the retrieval efficiency for large-area imagery, a lookup table (LUT) covering multidimensional observation geometry, elevation, and representative aerosol models was preconstructed. The key discretization settings of the LUT are listed in Table 2.

Table 2. Key parameter settings of the LUT.

LUT Parameter	Number of Levels	Parameter Values
Solar zenith angle	5	20–60 (step of 10)
Satellite zenith angle	5	0–80 (step of 20)
Relative azimuth angle	7	0–180 (step of 30)

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

550 nm AOD	13	0.01, 0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.80, 1.00, 1.50, 2.00, 2.50
Aerosol model	2	Continental and urban
Sensor band	5	0.443, 0.490, 0.681, 0.865, 2.130 µm

The continental and urban aerosol models adopted in this study are predefined aerosol types implemented in 6S rather than aerosol types retrieved or fitted from the DQ-1 WSI observations. Considering the influence of anthropogenic emissions and regional aerosol transport over the BTH region, the continental and urban models were adopted as two predefined candidate aerosol assumptions to provide distinct optical conditions within the LUT retrieval. The continental and urban models were evaluated separately without performing explicit scene-level aerosol-type classification. These two predefined models have different spectral scattering, absorption, and extinction characteristics; representative values of single-scattering albedo and normalized spectral aerosol optical depth are provided in Table S5 of the Supplement. They are used here as simplified candidate representations rather than as a complete description of aerosol-type variability over the BTH region.

Considering that gas absorption in the broad WSI spectral bands may introduce additional uncertainty into transmittance estimation, the transmittance parameters in the LUT were further corrected for gas absorption so as to obtain LUT parameters representing total atmospheric transmittance. During online retrieval, the corresponding radiative parameters were interpolated from the LUT according to pixel-wise observation geometry, elevation, and candidate AOD, and TOA reflectance was then simulated using the given surface reflectance. AOD was subsequently solved by minimizing the matching residual between observed and simulated reflectance. It should be noted that although the LUT was preconstructed for multiple spectral bands, the final AOD matching objective in the present retrieval uses the 0.443 µm band. The continental and urban aerosol models are evaluated separately, and the optimal model is selected according to the matching criterion. This process constitutes the static-prior LUT retrieval used as the control case in this study, while the subsequent residual learning is used only to correct the surface term and does not alter the main LUT retrieval chain.

3.3 Construction of the Static Multi-Band Surface Reflectance Library

A static surface reflectance prior was constructed from low-AOD clear-sky reference scenes. This clear-sky compositing and surface-background inversion strategy has been widely used to obtain reliable surface reflectance estimates when multi-angle observations are unavailable (Knapp et al., 2005). Specifically, a small number of representative low-AOD clear-sky scenes were first selected. After invalid pixels affected by clouds, water, and snow were removed, the corresponding MCD19A2 AOD was used as an external constraint. Together with DQ-1 WSI TOA reflectance, viewing geometry, and elevation, the LUT was interpolated to obtain the path reflectance, downward and upward transmittance, and atmospheric spherical albedo. Candidate surface reflectance values for the blue, green, red, and near-infrared bands were then inversely estimated on a pixel-by-pixel basis. To reduce the influence of aerosol contamination on surface-term estimation, only clear-sky pixels within a low-AOD range were used, with the AOD constraint set to 0.01–0.50.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

To maintain reasonable spectral consistency among bands in the static prior, reference dates were not selected independently for each band. Instead, a pixel-wise reference-date selection strategy was adopted. Among the candidate reference scenes, a given date was allowed to participate in the selection only when valid inversions were available for at least three bands. A composite score was then calculated for each candidate date, and the date with the highest score was selected as the final reference date for that pixel.

Here, Brightness is defined as the mean candidate surface reflectance of the blue, green, red, and near-infrared bands on the candidate date:

(1) $$Brightness = \frac{\rho_{blue} + \rho_{Green} + \rho_{Red} + \rho_{NIR}}{4}$$

where ρ_{Blue}, ρ_{Green}, ρ_{Red}, and ρ_{NIR} denote the candidate surface reflectance values of the four bands on that date. *Brightness* represents the overall brightness level of the pixel for a given candidate date. A multi-band mean brightness was used rather than a single-band brightness to reduce the influence of band-specific outliers, local noise, or individual-band biases on reference-date selection, thereby improving the overall stability of the multi-band static prior. Meanwhile, the deviation between the NDVI of the candidate date and the median NDVI of all candidate dates for the same pixel was used to represent vegetation-state consistency. The composite score was defined as

(2) $$Score = Brightness - \alpha \cdot |NDVI - NDVI_{median}|$$

Here, $\alpha = 0.06$ is an empirical parameter used to balance the brightness constraint and the vegetation-state consistency constraint. To examine the reasonableness of this empirical setting, a simple sensitivity analysis was conducted, as shown in Fig. 2. Within the tested range, increasing α gradually reduced the selected-scene NDVI deviation, while the mean brightness remained nearly stable. This indicates that changes in α mainly adjusted the relative weight of the vegetation-state consistency constraint, with little effect on the overall brightness characteristics. In addition, the reference-date switch ratio reached its minimum near $\alpha = 0.06$, suggesting that this value provides a reasonable compromise among brightness constraint, vegetation-state consistency, and reference-date selection stability. Therefore, $\alpha = 0.06$ was adopted as the empirical parameter for constructing the static surface reflectance prior.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

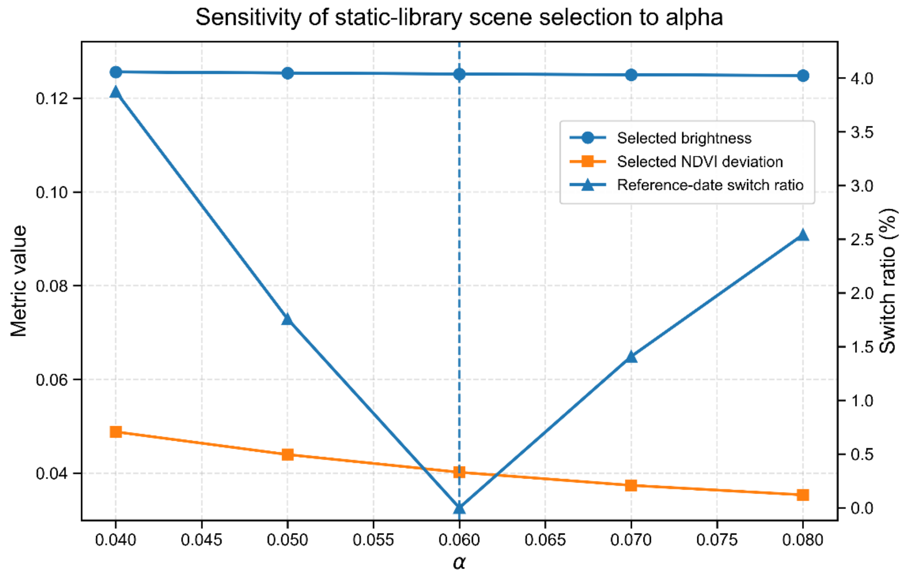

Figure 2. Sensitivity analysis of the empirical parameter α used in the pixel-wise reference-date selection for the static surface reflectance prior. The x-axis denotes the value of α. The left y-axis shows the mean selected brightness and the mean selected NDVI deviation, and the right y-axis shows the reference-date switch ratio relative to the default setting ($\alpha = 0.06$). The results indicate that $\alpha = 0.06$ provides a reasonable compromise among brightness constraint, vegetation-state consistency, and reference-date selection stability.

After the optimal reference date was determined, the blue, green, red, and near-infrared surface reflectance values from that same date were extracted simultaneously as the static prior values. This avoided spectral inconsistency caused by using different reference dates for different bands. For pixels that still had missing values, spatial nearest-neighbor filling was further applied to improve the spatial completeness of the static prior.

3.4 Surface Correction and AOD Retrieval Coupled with Random-Forest Residual Learning

Based on the static multi-band surface reflectance prior, a residual-learning model was constructed for the 0.443 μm blue band to compensate for scene-dependent mismatch between the static blue-band prior and the LUT-derived reference surface term. The blue band was selected as the correction target because it serves as the primary aerosol-sensitive retrieval channel in the current framework and is particularly sensitive to errors in surface characterization over complex land surfaces. Surface-prior mismatch may also occur in the other spectral bands; the present study does not assume these bands to be error-free. Rather, the correction is deliberately focused on the blue band to evaluate whether targeted correction of the surface term in the primary AOD-sensitive channel can improve the subsequent LUT-based AOD retrieval.

Instead of directly fitting AOD with machine learning, this study restricts the learning target to the residual between the static surface prior and the LUT-derived reference blue-band surface term. For a given pixel, the blue-band surface residual is defined as

(3) $\Delta\rho_{blue} = \rho_{surf,blue}^{ref} - \rho_{surf,blue}^{lib}$.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

where $\rho_{surf,blue}^{ref}$ denotes the reference blue-band surface reflectance inversely estimated under low-AOD conditions using external AOD constraints and the LUT, and $\rho_{surf,blue}^{lib}$ denotes the blue-band surface reflectance from the static prior library.

To ensure the reliability of the residual samples, only pixels satisfying all of the following conditions were retained: valid quality-control flags, valid external AOD within the low-AOD range (0.01–0.50), and reasonable blue-band TOA reflectance and static surface reflectance. In addition, the LUT-derived reference blue-band surface term was constrained to the range of 0–0.8, and samples with absolute residuals larger than 0.12 were excluded as outliers during residual-sample construction. The residual threshold was used only for offline sample screening to remove extreme discrepancies between the reference surface term and the static prior and does not directly enter the subsequent online AOD inversion.

A random forest (RF) model was used to learn the nonlinear mapping between the blue-band surface reflectance residual and the input features (Breiman, 2001; Wei et al., 2019). The input features included blue-band surface reflectance from the static library, blue-band TOA reflectance, NDVI derived from red and near-infrared TOA reflectance, solar zenith angle, viewing zenith angle, relative azimuth angle, and elevation. It should be noted that although external MODIS AOD was used to assist the construction of reference surface reflectance samples under low-AOD conditions, it was not used as an input feature for RF training. Therefore, no external AOD product is required during online inference.

The study period corresponded to spring 2025 (March–May). Three scenes acquired on 22 March, 22 April, and 19 May 2025 were used for model training, and one independent scene acquired on 28 April 2025 was used for internal validation. Up to 80,000 pixels were randomly sampled from each training scene, and up to 60,000 pixels were sampled from the validation scene. Following previous studies on random-forest hyperparameters and tuning strategies (Probst et al., 2019), the main RF settings were as follows: 500 trees, maximum depth of 22, minimum leaf size of 40, minimum split size of 80, and maximum feature fraction of 0.7, while other parameters were kept at their default values.

The detailed residual-sample screening criteria and RF training configuration are summarized in Table S2 in the Supplement. Because neighbouring pixels are not fully independent, the pixel-level sample size should not be interpreted as the number of independent atmospheric cases; therefore, the residual module was additionally evaluated using a held-out scene, and the final retrieval performance was assessed using independent AERONET station–date matchups.

In the online stage, the blue-band residual was predicted pixel by pixel from WSI observations, observation geometry, elevation, and the static surface prior, and the static prior was then corrected as

(4) $\rho_{,blue}^{corr} = \rho_{blue}^{lib} + \widehat{\Delta\rho_{blue}}$

where $\widehat{\Delta\rho_{blue}}$ is the RF-predicted blue-band residual and $\rho_{,blue}^{corr}$ is the corrected blue-band surface reflectance. To prevent unusually large surface corrections from propagating into the subsequent LUT inversion, range constraints were imposed on both the predicted residual and the corrected surface term. Specifically, the RF-predicted residual was limited to [−0.10, 0.10] as a safeguard against extreme model extrapolation, and the corrected blue-band surface term was constrained to [0, 0.8], consistent with the admissible range used for the LUT-derived reference surface term.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

The corrected blue-band surface reflectance was then reintroduced into the LUT retrieval framework described in Sect. 3.2, and AOD was retrieved by matching the observed and simulated TOA reflectance. The discrete AOD nodes were set to 0.01, 0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.80, 1.00, 1.50, 2.00, and 2.50. Retrieval was performed separately under the continental and urban aerosol models, and the optimal aerosol model was selected according to the minimum median matching cost at the scene scale. For solutions falling between discrete AOD nodes, local parabolic interpolation was further applied to obtain continuous AOD estimates. Finally, a quality flag was constructed using the pixel-level matching cost, defined as the minimum squared difference between the observed and LUT-simulated 0.443 μm TOA reflectance over the discrete AOD nodes. A matching-cost threshold of 0.03 was used as a conservative gross-mismatch screening criterion, and pixels satisfying this criterion were flagged as high-quality for subsequent analysis. A sensitivity check showed that thresholds of 0.01, 0.03, and 0.05 retained 99.86 %, 99.89 %, and 99.94 % of valid inversion pixels, respectively, indicating that retrieval coverage is only weakly sensitive to this criterion. Small isolated high-quality patches were removed, and a 3 × 3 median filter was further applied as a spatial post-processing step to improve the stability of the final AOD product.

In this way, a DQ-1 WSI AOD retrieval workflow consisting of "static surface prior–blue-band residual correction–LUT-based physical retrieval" was established. It should be emphasized that the external AOD product was used only to assist the construction of reference surface reflectance samples under low-AOD conditions, rather than as a direct learning target of the model. Therefore, the proposed method is essentially a constrained correction of surface-term errors within the LUT-based physical retrieval framework.

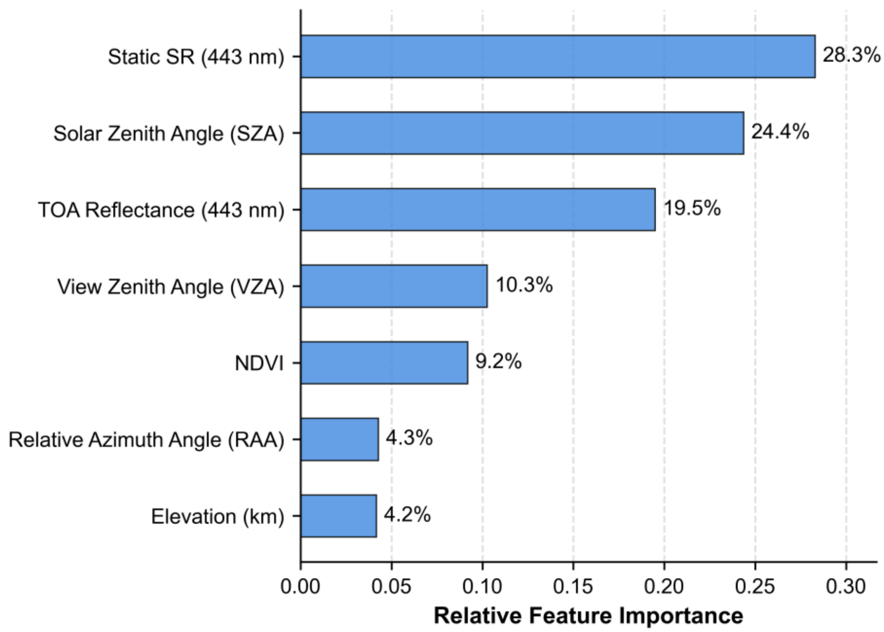

Figure 3. Relative importance of the input features used in the blue-band surface-residual learning model. The x-axis shows the relative feature importance in the random forest model, and the y-axis lists the input variables. Static blue-band surface prior, solar zenith angle, and blue-band TOA reflectance show the highest contributions, indicating that the residual correction mainly relies on the blue-band prior, observational information, and observation geometry.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

To further illustrate how the residual-learning model uses the input information, the relative feature importance of the random forest model was analyzed, as shown in Fig. 3. The results indicate that static blue-band surface prior, solar zenith angle, and blue-band TOA reflectance are the three most important variables, with relative contributions of 28.3%, 24.4%, and 19.5%, respectively. Viewing zenith angle and NDVI also contribute to the prediction, whereas the relative contributions of relative azimuth angle and elevation are lower. Previous studies have shown that feature-importance analysis can help examine whether machine-learning models mainly rely on physically meaningful input variables, thereby improving interpretability in Earth-science applications (Roscher et al., 2020). Overall, the feature-importance results indicate that the RF prediction relies primarily on the static blue-band surface prior, blue-band TOA reflectance, and observation geometry, which are physically relevant to the surface–atmosphere separation problem addressed in this study. These results support the physical interpretability of the residual-learning model, while the effect of the surface-term correction on the intermediate surface term and final AOD retrieval is examined explicitly in Sects. 4.2 and 4.3.

4. Results and Validation

4.1 Validation against AERONET

The AERONET matchup strategy is described in Sect. 2.3 and was used to validate the DQ-1 WSI AOD retrievals.

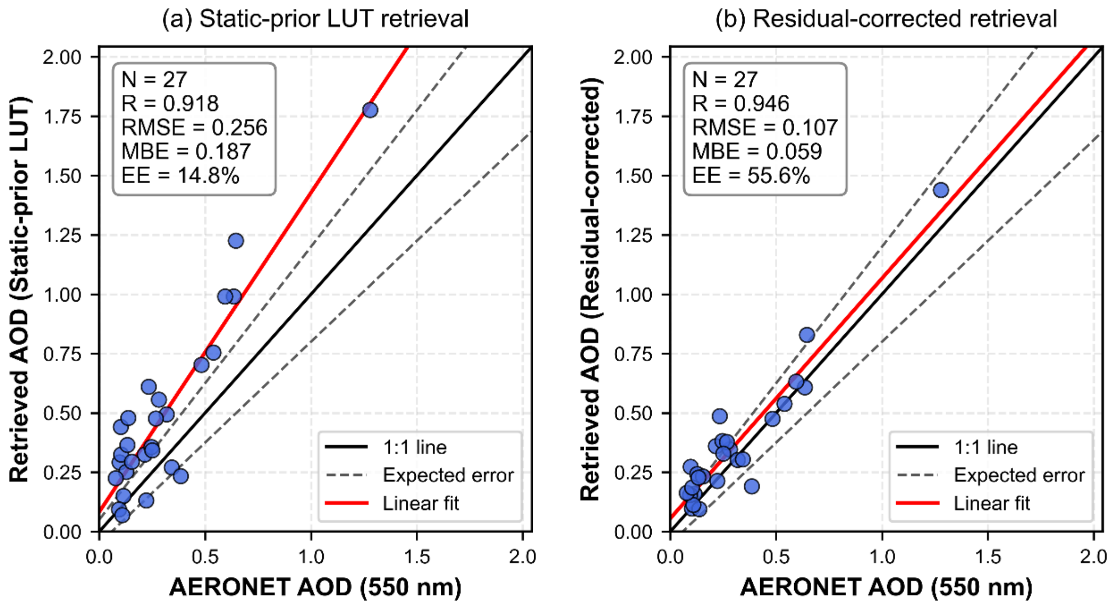

Figure 4. AERONET validation of the static-prior LUT retrieval and the residual-corrected retrieval. (a) Static-prior LUT retrieval versus AERONET AOD; (b) residual-corrected retrieval versus AERONET AOD. The black solid line denotes the 1:1 line, the grey dashed lines denote the expected error envelope, and the red line denotes the linear fit.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

Figure 4 shows the AERONET validation results for the static-prior LUT retrieval and the residual-corrected retrieval. To ensure direct comparability between the two retrievals, the statistics were calculated using only station–date matchups for which both retrievals were valid. Each station–date pair was treated as an independent sample unit, resulting in 27 independent matchups. This sample size should be regarded as a conservative result obtained under a unified temporal window, a consistent spatial averaging window, and the requirement that both DQ-1 WSI retrievals were valid, rather than as a simple accumulation of all available AERONET observations.

The static-prior LUT retrieval shows a relatively high correlation with AERONET, with R = 0.918, but it also exhibits a clear positive bias, with an RMSE of 0.256, an MBE of 0.187, and only 14.8 % of the samples falling within the expected error envelope. Here, the expected error envelope is used as a commonly adopted metric in satellite AOD validation, following previous MODIS aerosol validation studies (Sayer et al., 2013). In contrast, the residual-corrected retrieval shows a scatter distribution closer to the 1:1 line, with R increasing to 0.946, RMSE decreasing to 0.107, MBE decreasing to 0.059, and the fraction within the expected error envelope increasing to 55.6 %. These results indicate that introducing the blue-band surface residual correction substantially reduces the positive bias and overall retrieval error relative to the static-prior LUT retrieval. The relationship between the surface-term correction and the resulting AOD changes is examined further in Sects. 4.2 and 4.3. The current AERONET matchups mainly cover observations in spring 2025, with limited station coverage and temporal sampling. Therefore, these results should be interpreted as preliminary and conservative independent validation of the proposed method under springtime complex-surface conditions, rather than as sufficient evidence for year-round robustness. The following sections further examine the stability and physical consistency of the method through held-out-scene evaluation of the residual correction module, AOD response analysis, and regional spatial-structure and hold-out-date cross-product consistency analyses with MCD19A2.

4.2 Held-out-scene evaluation of the blue-band surface residual correction module

To further examine whether the blue-band residual correction learns physically meaningful surface-term adjustments, the residual correction module was evaluated using a held-out scene. It should be noted that the reference blue-band surface term used here does not represent the true surface reflectance. Instead, it was inversely estimated from DQ-1 WSI TOA reflectance using the LUT constrained by MCD19A2 AOD. Therefore, this reference is used only to assess the ability of the residual correction module to learn the mismatch between the static surface prior and the LUT-derived reference surface term.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

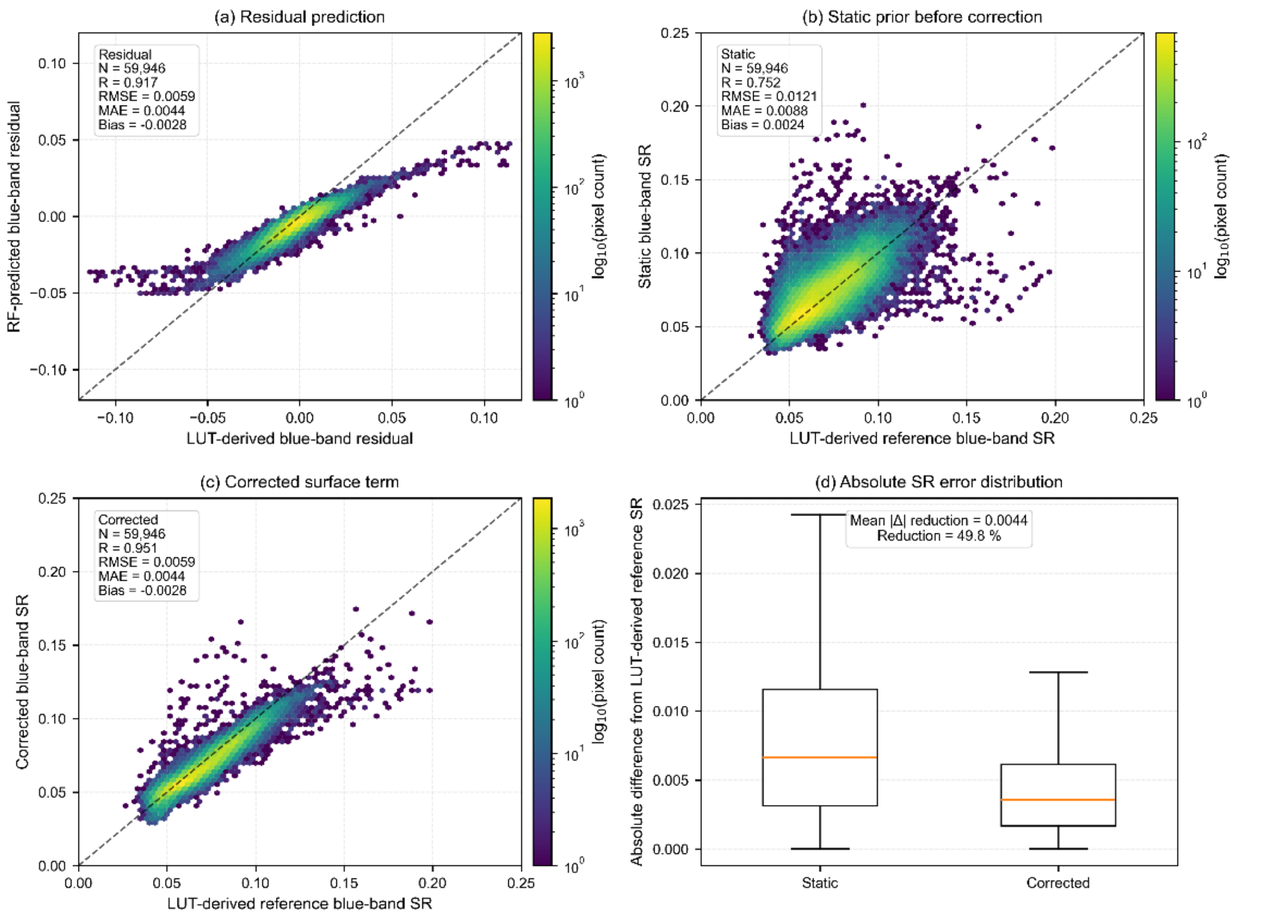

Figure 5. Evaluation of the blue-band surface residual correction module on a held-out scene. (a) Relationship between the LUT-derived blue-band surface residual and the RF-predicted residual; (b) relationship between the LUT-derived reference blue-band surface term and the original static blue-band surface prior; (c) relationship between the LUT-derived reference blue-band surface term and the residual-corrected blue-band surface term; (d) distributions of the absolute differences before and after correction relative to the LUT-derived reference surface term. The LUT-derived reference surface term was obtained through LUT inversion constrained by MCD19A2 AOD and is used only to evaluate the correction of surface-prior mismatch; it should not be interpreted as true surface reflectance. The dashed lines denote the 1:1 line.

As shown in Fig. 5, the RF-predicted residual agrees well with the LUT-derived residual, with a correlation coefficient of 0.917 and RMSE and MAE values of 0.0059 and 0.0044, respectively. This indicates that the model can capture both the direction and magnitude of the blue-band surface-term residual. Before correction, the static blue-band surface prior shows only moderate agreement with the reference surface term, with a correlation coefficient of 0.752 and an MAE of 0.0088. After residual correction, the agreement is substantially improved, with the correlation coefficient increasing to 0.951 and the MAE decreasing to 0.0044. The error distribution further shows that the mean absolute difference of the blue-band surface term is reduced by approximately 49.8 % after correction.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

These results indicate that the proposed method does not empirically shift the AOD retrieval results. Instead, it corrects the blue-band surface-term mismatch before the final LUT-based AOD retrieval. This intermediate-layer evidence suggests that static surface-prior error is an important source of instability in the static-prior LUT retrieval, while the residual correction improves the surface-term input to the LUT retrieval chain. This provides mechanistic support for the subsequent improvement in AOD accuracy and regional spatial continuity.

4.3 AOD response to blue-band surface-term correction and stratified analysis

Based on the evaluation of the blue-band surface residual correction module, we further examined how the surface-term correction propagates to the final AOD retrieval. For this purpose, the relationship between the change in blue-band surface reflectance and the change in AOD was analyzed by comparing the static-prior LUT retrieval with the residual-corrected retrieval. The surface reflectance change was defined as $\Delta SR = SR_{\text{residual-corrected}} - SR_{\text{static-prior}}$, and the AOD change was defined as $\Delta AOD = AOD_{\text{residual-corrected}} - AOD_{\text{static-prior}}$. The results are shown in Fig. 6.

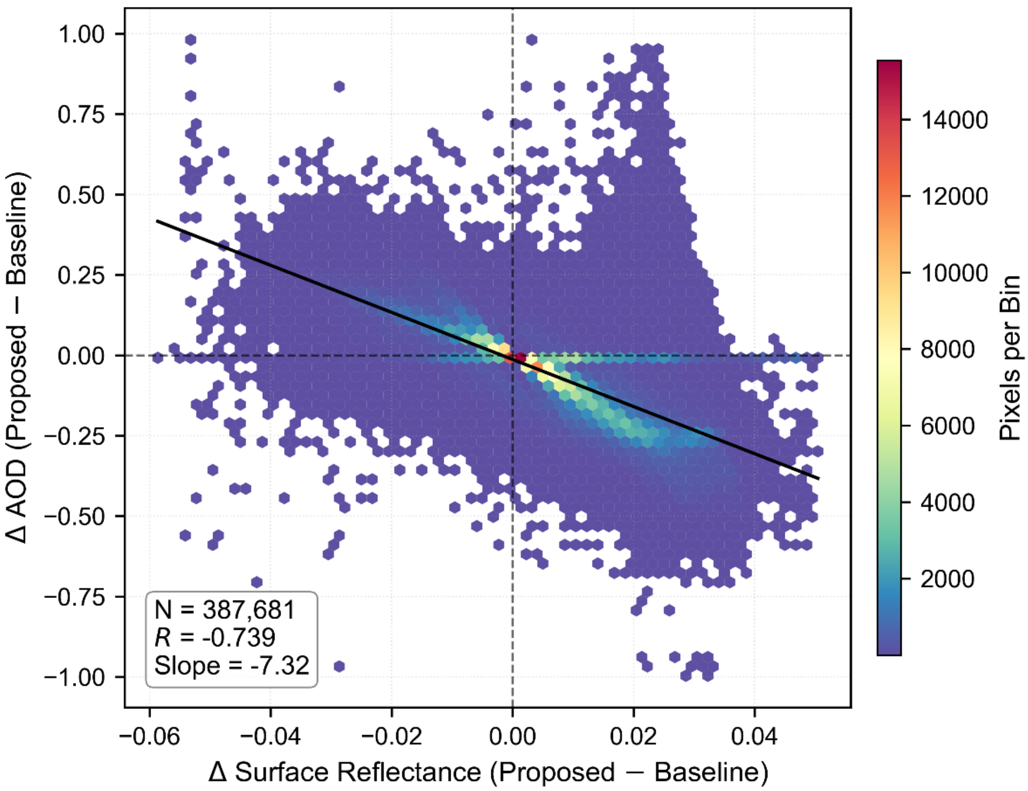

Figure 6. Overall relationship between blue-band surface reflectance correction and AOD change. The x-axis denotes the change in surface reflectance between the residual-corrected retrieval and the static-prior LUT retrieval, and the y-axis denotes the corresponding change in AOD. The color indicates the number of pixels in each hexagonal bin, and the black line denotes the linear fit.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

For the full sample, ΔSR and ΔAOD show a clear negative relationship. In other words, when the corrected surface reflectance increases, the retrieved AOD generally decreases; conversely, when the corrected surface reflectance decreases, the retrieved AOD tends to increase. This response direction is consistent with the trade-off between surface-reflected contribution and atmospheric scattering contribution in radiative-transfer-based retrievals. It indicates that correction of the blue-band surface term directly affects the partitioning of the TOA signal between the surface and aerosol components, thereby changing the final AOD estimate. This behavior is also consistent with the critical surface albedo theory, which suggests that surface reflectance errors can be amplified and propagated into aerosol retrievals under specific observational and surface conditions (Seidel and Popp, 2012). Therefore, rather than applying a uniform empirical shift to the AOD results, the residual correction produces AOD changes whose direction is consistent with radiative-transfer expectations. This response supports the interpretation that modifying the blue-band surface term changes the partitioning of the TOA signal between surface reflection and aerosol scattering before the final LUT solution.

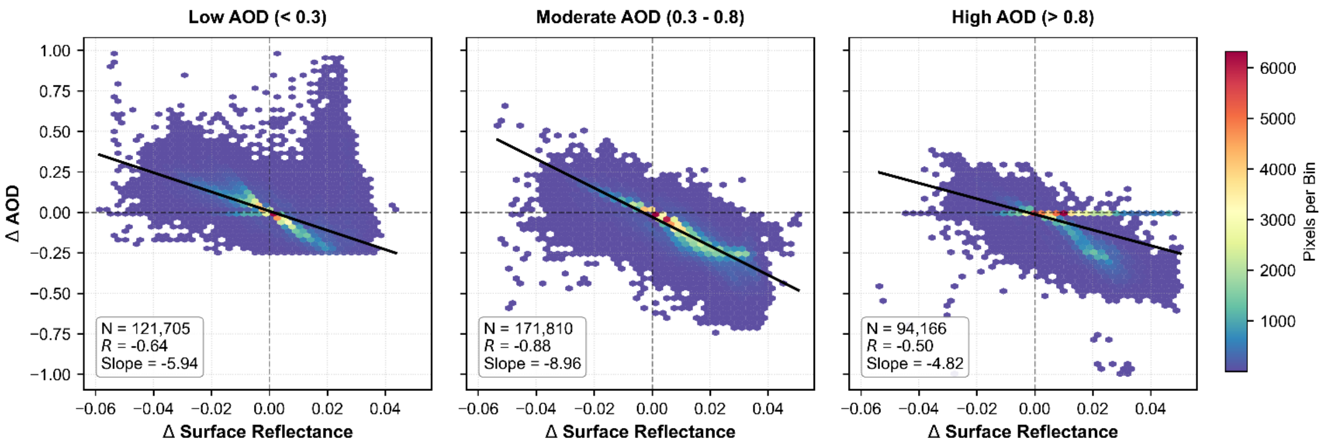

Figure 7. Stratified relationship between blue-band surface reflectance correction and AOD change under different AOD loading levels. (a) Low AOD (< 0.3), (b) moderate AOD (0.3–0.8), and (c) high AOD (> 0.8). The x-axis denotes the change in surface reflectance between the residual-corrected retrieval and the static-prior LUT retrieval, and the y-axis denotes the corresponding change in AOD. The color indicates the number of pixels in each hexagonal bin, and the black line denotes the linear fit.

The strength of the AOD response to blue-band surface-term correction varies with aerosol loading. Under low-to-moderate AOD conditions, the negative relationship between ΔSR and ΔAOD is more pronounced. This suggests that, within this range, the TOA signal is still jointly affected by surface reflection and aerosol scattering, so errors in the static surface prior are more likely to propagate into AOD biases. As a result, surface residual correction has a stronger influence on the retrieval. In contrast, under high-AOD conditions, the atmospheric scattering contribution becomes more dominant, and the relative influence of surface-term error on the final solution decreases. Consequently, the response amplitude of ΔAOD to ΔSR is reduced.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

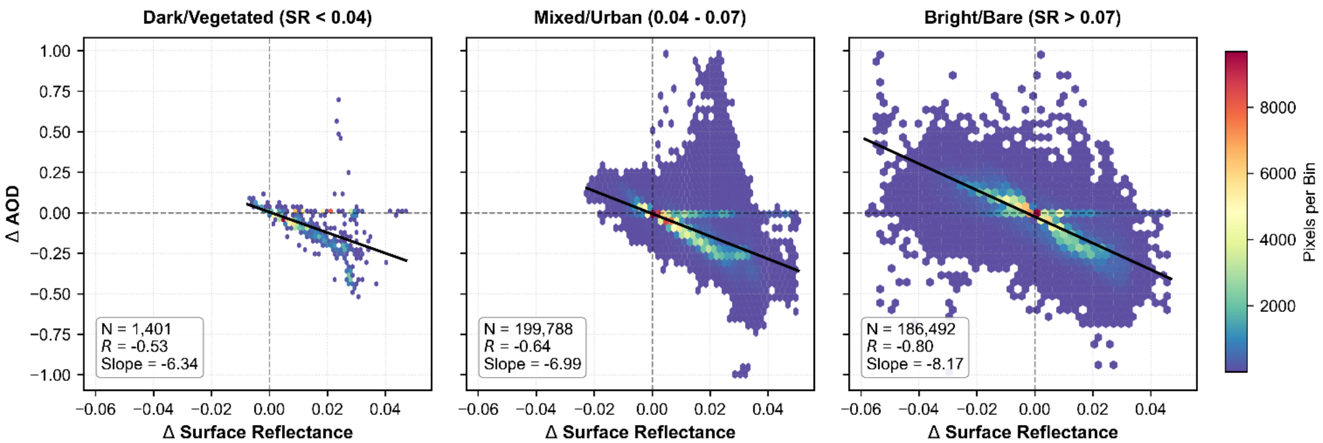

Figure 8. Sensitivity of the AOD response to surface reflectance correction under different surface brightness conditions. (a) Dark or vegetated surfaces (SR < 0.04), (b) mixed or urban surfaces (0.04–0.07), and (c) bright or bare surfaces (SR > 0.07). The x-axis denotes the change in surface reflectance between the residual-corrected retrieval and the static-prior LUT retrieval, and the y-axis denotes the corresponding change in AOD. The color indicates the number of pixels in each hexagonal bin, and the black line denotes the linear fit.

The stratification by surface brightness shows that mixed or urban surfaces and bright surfaces exhibit a more pronounced AOD response. In these regions, surface reflectance varies rapidly in space and mixed pixels are more common, making the static surface prior more likely to be mismatched with the actual observation scene. This mismatch can introduce local anomalous high values or spatial speckle into the static-prior LUT retrieval. After blue-band residual correction, the surface contribution is redistributed, and the corresponding AOD estimate is adjusted accordingly. By comparison, over dark or vegetated surfaces, the surface reflectance is lower and the static-prior LUT retrieval is less affected by surface-term mismatch; therefore, the AOD response induced by residual correction is weaker. Previous studies have shown that aerosol retrievals over bright and highly reflective source regions are generally more sensitive to errors in surface reflectance assumptions, further highlighting the importance of accurate surface-term characterization under complex surface conditions (Hsu et al., 2004). Overall, the stratified results show that the AOD response to blue-band surface-term correction is more pronounced over mixed/urban and bright surfaces, where static-prior mismatch is more likely to occur. Together with the $\Delta SR - \Delta AOD$ relationship, these results support the interpretation that the observed AOD changes are consistent with alleviation of blue-band surface-term mismatch rather than with a uniform empirical adjustment of the AOD results.

4.4 Regional spatial consistency and hold-out-date cross-product consistency with MCD19A2

To evaluate the regional-scale spatial performance of the proposed method and to reduce the effects of resolution differences and grid misalignment, the DQ-1 WSI AOD retrievals and the MCD19A2 product were first reprojected, spatially aligned, and scale-matched before comparison. Block averaging at an approximate 3 km scale was further applied to improve the spatial comparability between the two products. Pixels with very low AOD values (< 0.05 in either product) and evident residual cloud-contamination outliers were excluded before calculating the cross-product statistics.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

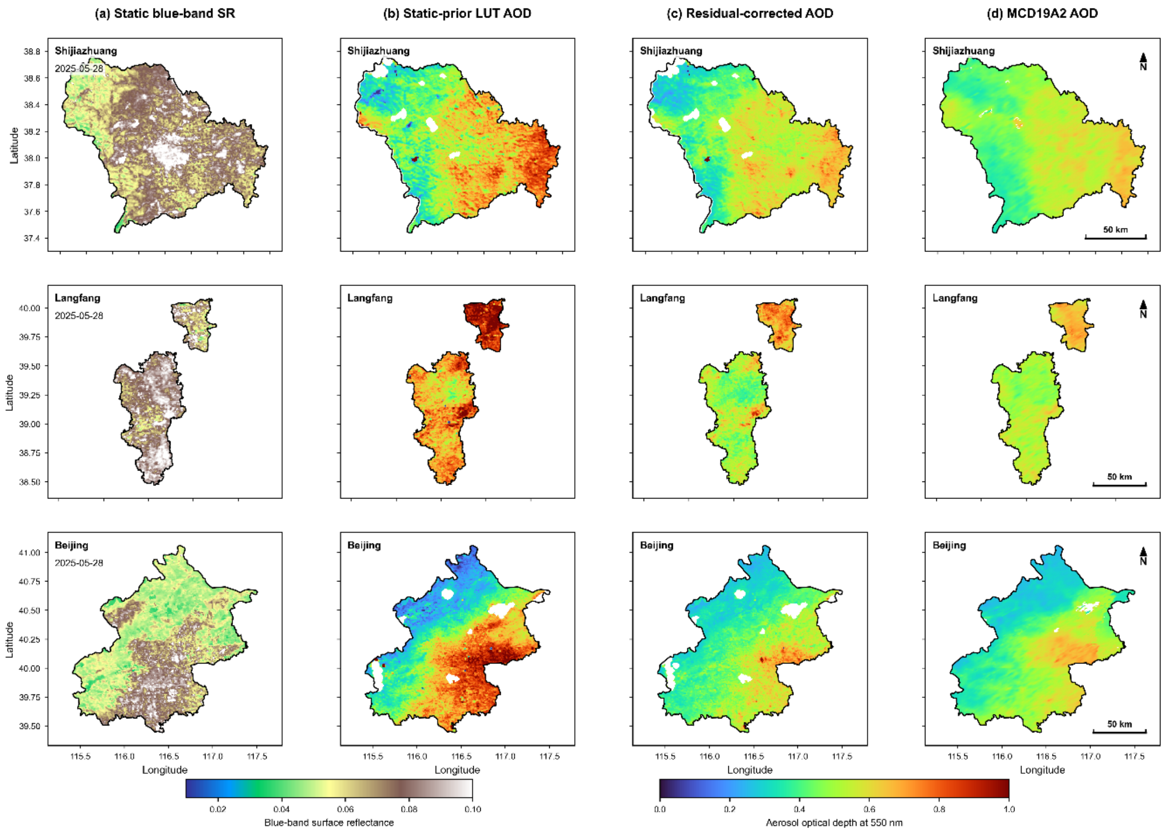

Figure 9. Spatial comparison of the static-prior LUT retrieval, residual-corrected retrieval, and MCD19A2 AOD over representative urban regions on 28 May 2025. (a) Static blue-band surface reflectance; (b) static-prior LUT AOD; (c) residual-corrected AOD; (d) MCD19A2 AOD. The three rows correspond to Shijiazhuang, Langfang, and Beijing, respectively. MCD19A2 is used here as a reference for regional spatial structure in the cross-product consistency analysis, rather than as an independent accuracy validation standard.

To visually examine the effect of blue-band surface residual correction on the regional AOD spatial structure, three representative urban regions in the Beijing–Tianjin–Hebei region, namely Shijiazhuang, Langfang, and Beijing, were selected on 28 May 2025. The spatial distributions of the static-prior LUT retrieval, the residual-corrected retrieval, and MCD19A2 were compared. This date corresponds to the later stage of a regional pollution episode over the Beijing–Tianjin–Hebei region, during which the inter-city AOD distribution was influenced by regional transport, boundary-layer development, and complex surfaces, leading to a clear spatial gradient (Li et al., 2017).

As shown in Fig. 9, the static-prior LUT retrieval tends to produce locally anomalous high values and spatial speckle over urban and urban–rural transitional areas, particularly over Langfang and southern Beijing. This suggests that, under complex surface conditions, mismatch between the static surface prior and the actual observation scene can propagate into the final AOD retrieval, resulting in spatial discontinuities or local overestimation. After blue-band surface residual correction, the anomalous high values are reduced in all three urban regions, the AOD spatial transitions become smoother, and the regional gradient structure becomes more consistent with the large-scale spatial pattern represented by MCD19A2.

460 It should be noted that MCD19A2 and DQ-1 WSI differ in spatial resolution, viewing geometry, retrieval strategy, and quality-control procedure. Therefore, exact agreement in local texture or fine-scale intensity between the two products is not expected. The comparison in Fig. 9 is intended mainly to illustrate the regional-scale spatial plausibility of the residual-corrected DQ-1 WSI AOD, rather than to provide an independent accuracy validation. The independent assessment of retrieval effectiveness in this study remains based on the AERONET station matchups. Overall, this representative regional episode indicates that blue-band surface residual correction can effectively reduce local anomalous high values and speckle in the static-prior LUT retrieval and improve the spatial continuity and stability of AOD distributions over complex surfaces.

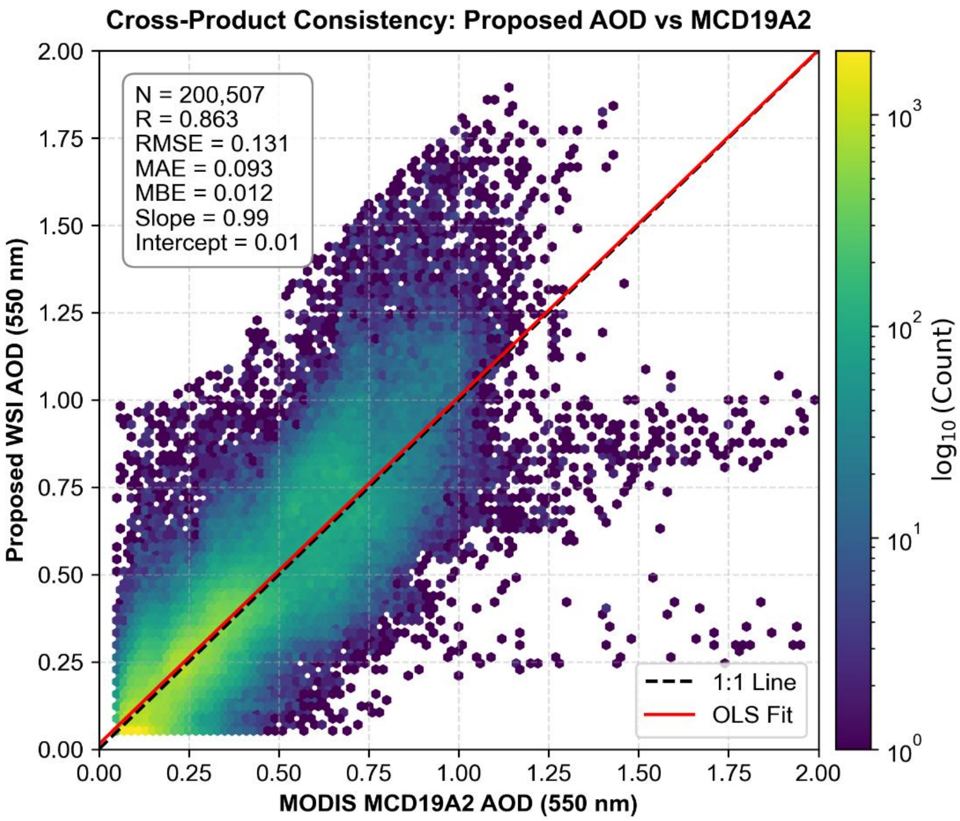

Figure 10. Cross-product consistency between the residual-corrected DQ-1 WSI AOD and MCD19A2 after excluding the training and internal validation dates. Dates used for residual-model training and internal validation were excluded from the statistics. The x-axis denotes MCD19A2 AOD, and the y-axis denotes residual-corrected DQ-1 WSI AOD. The color indicates the logarithmic pixel count, the black dashed line denotes the 1:1 line, and the red line denotes the linear fit.

To provide additional regional-scale evidence beyond the limited ground-based matchups, a large-sample hold-out-date cross-product consistency analysis was conducted. After excluding the scenes used for residual-model training and internal validation, 200,507 collocated valid pixels were extracted to compare the residual-corrected DQ-1 WSI AOD with MCD19A2 on a pixel-wise basis. As shown in Fig. 10, when complex-surface pixels were retained, the residual-corrected retrieval showed close regional-scale consistency with MCD19A2, with a correlation coefficient of 0.863 and RMSE and MAE values of 0.131 and

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

0.093, respectively. The mean bias error (MBE) was 0.012, and the fitted slope and intercept were 0.99 and 0.01, respectively, indicating that the method does not show an evident overall overestimation or underestimation on dates not used for training or internal validation.

The hold-out-date comparison therefore supports the regional-scale consistency of the residual-corrected retrieval within the spring study period. It should be emphasized that MCD19A2 is used here mainly as a reference for regional spatial consistency and hold-out-date cross-product consistency. This analysis does not constitute strict independent accuracy validation, and the independent evaluation of the proposed retrieval remains based on AERONET station matchups.

5. Discussion

5.1 Methodological positioning of the DQ-1 WSI AOD retrieval workflow

The main objective of this study is not to replace physical retrieval with machine learning, but to compensate for key surface-term errors that affect the stability of DQ-1 WSI AOD retrieval over complex surfaces. For new sensors with limited temporal accumulation, long-term and stable surface reflectance priors are often not yet available, and clouds, haze, water, snow, and mixed pixels over complex land surfaces further increase the difficulty of generating analysis-ready data and characterizing the surface term. Therefore, this study develops a retrieval-oriented AOD workflow for DQ-1 WSI that integrates quality control, static surface-prior construction, blue-band surface residual correction, and LUT-based physical retrieval.

Within this framework, the unified quality-control procedure provides relatively reliable clear-sky inputs for static surface-prior construction and subsequent retrieval. The multi-band static surface prior inverted from low-AOD clear-sky reference scenes provides the basic surface constraint for LUT retrieval. The blue-band residual correction further compensates for the remaining mismatch between the static prior and the actual observation scene. In this way, machine learning is embedded only in the surface-term error correction step, rather than being used to directly fit AOD, thereby improving retrieval stability over complex scenes while preserving radiative-transfer interpretability.

The results in Sect. 4 support this methodological positioning from several perspectives. The AERONET validation shows that the residual-corrected retrieval substantially reduces AOD error and positive bias compared with the static-prior LUT retrieval. The held-out-scene evaluation indicates that the RF model can effectively learn the residual relationship between the static blue-band surface prior and the LUT-derived reference surface term. The response relationship between ΔSR and ΔAOD further shows that surface-term correction affects the final AOD estimate in a direction consistent with radiative-transfer expectations, rather than simply shifting the AOD results empirically. The regional spatial comparison and hold-out-date cross-product consistency analysis also show that the method can reduce, to some extent, local anomalous high values and spatial discontinuities in the static-prior LUT retrieval.

In terms of applicability, the proposed method shows more pronounced improvements over complex surfaces, such as urban areas, urban–rural transition zones, and bare land, as well as under low-to-moderate AOD conditions. In these regions, the

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

static surface prior is more likely to be mismatched with the actual observation scene, making the blue-band surface residual correction more effective. By contrast, over dark surfaces or relatively stable vegetated surfaces, the static-prior LUT retrieval is less affected by surface-term mismatch. Under high-AOD conditions, the atmospheric scattering contribution becomes more dominant, reducing the relative influence of surface-term errors on the final retrieval solution; consequently, the improvement is generally weaker than under low-to-moderate AOD conditions.

Overall, the proposed "static surface prior–blue-band residual correction–LUT physical retrieval" framework provides an initial demonstration of the feasibility of DQ-1 WSI for regional AOD retrieval over springtime complex surfaces. It should be emphasized that this study focuses on establishing and evaluating an interpretable retrieval workflow for this sensor, rather than specifically assessing the realism of sub-kilometre pollution structures at the 600 m scale.

5.2 Applicability limits and uncertainties

Although the proposed method leads to clear improvements in DQ-1 WSI AOD retrieval, several applicability limits and uncertainties remain. First, the number of independent AERONET matchups is still limited by the current availability of DQ-1 WSI observations, the distribution of AERONET stations, and the requirement of commonly valid retrieval samples. The 27 matchups used in this study were obtained under a unified temporal window, a consistent spatial averaging window, and the condition that both DQ-1 WSI retrievals were valid. Therefore, these matchups provide preliminary independent support for the effectiveness of the method, but they are not sufficient to fully characterize its performance over broader regions, more seasons, and more complex pollution conditions.

Second, the spatial consistency analysis and large-sample cross-product comparison with MCD19A2 should not be interpreted as strict independent accuracy validation. On the one hand, MCD19A2 is a well-established satellite AOD product and provides a useful reference for regional spatial structure and hold-out-date stability. On the other hand, MCD19A2 was used under low-AOD clear-sky conditions to support the construction of LUT-derived reference surface-term samples, so the two datasets are not fully independent. Potential biases in MCD19A2 AOD may propagate through the LUT inversion into the derived reference surface term and consequently partly enter the residual-learning target. Such effects may be more relevant over bright or heterogeneous urban surfaces, where the separation of atmospheric and surface contributions is more sensitive to the assumed AOD. Therefore, the LUT-derived surface term is treated here as a pseudo-reference for residual learning rather than as a direct observation of true surface reflectance. For this reason, AERONET matchups are used as the primary independent basis for evaluating retrieval effectiveness, while comparisons with MCD19A2 are interpreted as regional spatial-structure comparison and cross-product consistency analysis.

Third, the current static surface prior and blue-band residual model were mainly constructed from springtime scenes, so the conclusions of this study are currently most applicable to springtime complex-surface conditions. For summer scenes with dense vegetation, surface spectral states, BRDF characteristics, and vegetation phenology may differ substantially from those represented in the current reference library, while winter snow or other bright surfaces may introduce additional surface-characterization challenges. Therefore, the existing static prior and RF residual model should not be assumed to transfer

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

unchanged across seasons. Extension to other seasons would likely require updating the static surface prior and augmenting or retraining the residual model using representative reference scenes from the target season. The hold-out-date cross-product consistency analysis indicates that the method maintains a certain degree of stability on dates not used for training or internal validation, but this mainly supports preliminary temporal robustness within the spring study period and cannot replace a systematic year-round evaluation. Similarly, regional transfer may require additional reference scenes and reconsideration of the candidate aerosol models when the dominant surface or aerosol regimes differ substantially from those in the BTH region. The present study focuses primarily on reducing uncertainty associated with the land-surface term, and the current implementation therefore uses two predefined 6S aerosol models as candidate aerosol assumptions. This simplified representation does not fully capture the temporal and compositional variability of aerosols over the BTH region, and uncertainty associated with the assumed aerosol optical model may become relatively more important under higher aerosol loading. Future extensions should therefore consider a broader set of aerosol models or additional aerosol-type constraints when applying the framework under more diverse aerosol conditions.

In addition, very thin clouds, semi-transparent cloud edges, warm-haze mixed pixels, and strongly heterogeneous surfaces may still affect blue-band AOD retrieval. Although a relatively strict QA masking procedure was applied, residual contamination may still introduce additional high-reflectance signals in the blue band, thereby interfering with surface residual learning and the final AOD retrieval (Christensen et al., 2017; Norgren et al., 2022). Meanwhile, the current framework still adopts an approximate Lambertian surface assumption. Under large viewing-geometry variations, high-contrast urban surfaces, and bare-land conditions, BRDF effects may become an additional source of uncertainty.

5.3 Future improvements and extension potential

Future work can further improve the proposed framework in several directions. First, as DQ-1 WSI in-orbit observations continue to accumulate, multi-seasonal surface reflectance priors that account for seasonal surface evolution can be constructed, following surface-reconstruction strategies developed for new Chinese satellite sensors (Shi et al., 2023). Long-term evaluations over broader climate zones, surface types, and pollution conditions will also be needed. In particular, for summer scenes with dense vegetation, dynamic surface priors capable of representing phenological changes should be introduced to reduce seasonal mismatch in AOD retrieval.

Second, the current residual correction deliberately focuses on the 0.443 μm blue band as the primary aerosol-sensitive retrieval channel. Surface-prior mismatch may also occur in other bands, including 0.490 and 0.681 μm. Future work could therefore extend the residual-correction framework to multiple visible and near-infrared bands while preserving the LUT-based physical retrieval chain. Such an extension would require consideration of band-dependent aerosol sensitivity, surface anisotropy, and cross-band error characteristics. Combined with statistical optimization or hybrid physical–statistical retrieval concepts (Litvinov et al., 2025), multi-band joint constraints may further improve spectral consistency and surface-term stability over complex surfaces.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

Third, explicit BRDF information could be introduced to replace or correct the current Lambertian approximation (Zhou et al., 2024). For wide-swath sensors, variations in viewing geometry can substantially affect the directional reflectance of the surface, especially over urban areas, bare land, and other highly reflective or heterogeneous surfaces. Accounting for BRDF explicitly would help reduce additional surface-term errors caused by viewing-angle variations and improve the stability of cross-date and cross-region applications.

Overall, this study does not aim to provide an operational AOD product applicable under all-season conditions. Instead, it provides an interpretable and extensible technical pathway for AOD retrieval from DQ-1 WSI and similar new wide-swath sensors over springtime complex surfaces. More systematic external validation using additional AERONET stations, ground-based observation networks, and independent satellite products will still be needed in future work.

6. Conclusions

This study developed a retrieval-oriented land AOD workflow for DQ-1 WSI, a new wide-swath sensor, over springtime complex surfaces. The workflow integrates quality control, a static surface reflectance prior, blue-band surface residual correction, and LUT-based physical retrieval. To address scene-dependent mismatch in the static surface prior over complex surfaces, a multi-band static surface reflectance prior was first constructed from low-AOD clear-sky reference scenes. A random forest model was then used to correct the 0.443 µm blue-band surface-term residual. In this way, the proposed method improves AOD retrieval accuracy and spatial stability while preserving the interpretability of the LUT-based radiative-transfer retrieval chain.

AERONET-based independent validation shows that the residual-corrected retrieval substantially reduces AOD retrieval error and positive bias compared with the static-prior LUT retrieval. Under a consistent spatiotemporal matching strategy and the requirement that both DQ-1 WSI retrievals were valid, the residual-corrected retrieval achieved R = 0.946, RMSE = 0.107, MBE = 0.059, and 55.6 % of the samples within the expected error envelope. In contrast, the static-prior LUT retrieval had RMSE = 0.256, MBE = 0.187, and only 14.8 % of the samples within the expected error envelope. These results demonstrate a clear improvement in retrieval accuracy and a substantial reduction in positive bias relative to the static-prior LUT retrieval under the current validation conditions.

The held-out-scene evaluation further shows that the RF model can capture the residual relationship between the static blue-band surface prior and the LUT-derived reference surface term. After correction, the correlation coefficient increased from 0.752 to 0.951, the MAE decreased from 0.0088 to 0.0044, and the mean absolute difference was reduced by approximately 49.8 %. In addition, the relationship between ΔSR and ΔAOD shows that the resulting AOD changes occur in a direction consistent with radiative-transfer expectations rather than as a uniform empirical shift. Together, these intermediate-layer and AOD-response results support the interpretation that the retrieval improvement is associated with alleviation of blue-band surface-prior mismatch within the LUT retrieval chain.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

At the regional scale, representative spatial comparisons show that the proposed method suppresses local anomalous high values, speckle, and unrealistic discontinuities in the static-prior LUT retrieval, producing more continuous AOD distributions that are more consistent with the large-scale regional gradients represented by MCD19A2. In the hold-out-date cross-product consistency analysis, after excluding scenes used for training and internal validation, the residual-corrected retrieval remained consistent with MCD19A2 at the regional scale, with 200,507 collocated pixels, R = 0.863, RMSE = 0.131, MAE = 0.093, MBE = 0.012, and a fitted slope of 0.99. This suggests that the method does not show an evident overall overestimation or underestimation on the hold-out dates.

In summary, the proposed framework provides an initial physically interpretable approach for improving DQ-1 WSI AOD retrieval over springtime complex surfaces. The combined surface-term, AOD-response, ground-based validation, and regional consistency analyses support the interpretation that the observed improvement is associated with alleviation of blue-band surface-prior mismatch within the LUT retrieval chain rather than direct empirical fitting of AOD. However, owing to the current limitations in observation period, static-prior construction scenes, and independent AERONET matchup samples, the conclusions of this study are mainly applicable to springtime complex-surface conditions and cannot yet fully represent performance during dense-vegetation seasons or throughout the year. Future work should incorporate multi-seasonal surface priors, longer time-series observations, explicit BRDF constraints, and additional independent ground-based samples to further evaluate and extend the cross-seasonal and cross-regional robustness of the method.

Code and data availability

The code used for QA-mask generation, blue-band surface-residual learning, and LUT-based AOD retrieval is available from the corresponding author upon reasonable request. AERONET observations are publicly available from the AERONET data portal. MODIS MAIAC MCD19A2 aerosol products are publicly available from NASA Earthdata. DQ-1 WSI data and the associated radiometric calibration coefficients were obtained from the China Centre for Resources Satellite Data and Application (CRESDA) and are subject to its data access policy. The processed matchup data, validation statistics, and figure data generated in this study are available from the corresponding author upon reasonable request.

Supplement link

The supplement related to this article is provided as a separate file.

Author contributions

JXZ designed the study, developed the retrieval algorithm, processed the DQ-1 WSI, AERONET, and MCD19A2 data, performed the experiments, analyzed the results, prepared the figures, and wrote the original manuscript. Yujing Wang contributed to manuscript formatting, editorial checking, and discussion of the research direction. MJZ contributed to data preprocessing and validation. LLW contributed to the organization of satellite and ground-based reference data. SD contributed to figure preparation and statistical analysis. XCY contributed to manuscript revision, formatting, and editorial checking. PFM supervised the study, contributed to the interpretation of the results, and revised the manuscript. HC contributed to methodological discussion and validation design, provided funding support, and revised the manuscript. WJZ supervised the study, contributed to project administration, and critically revised the manuscript. All authors reviewed and approved the final manuscript.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Acknowledgements

The authors thank the AERONET team and site principal investigators for providing the ground-based aerosol observations. The authors also acknowledge the MODIS MAIAC team for providing the MCD19A2 aerosol product. The authors thank the China Centre for Resources Satellite Data and Application (CRESDA) for providing the DQ-1 WSI data and associated radiometric calibration information.

Financial support

This research has been supported by Beijing Natural Science Foundation (L241082).

References

Bilal, M., Qiu, Z., and Lu, X.: A simplified high resolution MODIS Aerosol Retrieval Algorithm (SARA) for use over mixed surfaces, Remote Sens. Environ., 136, 135–145, https://doi.org/10.1016/j.rse.2013.04.014, 2013.

Breiman, L.: Random forests, Mach. Learn., 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Cao, J.: Geometric calibration of the whiskbroom DaQi-1 WSI based on an equivalent linear-array CCD, IEEE Trans. Geosci. Remote Sens., 61, 5616713, https://doi.org/10.1109/TGRS.2023.3300013, 2023.

Che, H., Xia, X., Zhu, J., Wang, H., Wang, Y., Sun, J., Zhang, X., and Shi, G.: Aerosol optical properties under the condition of heavy haze over an urban site of Beijing, China, Environ. Sci. Pollut. Res., 22, 1043–1053, https://doi.org/10.1007/s11356-014-3415-5, 2015.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

Chen, J., Chen, H. W., Li, Z., Wang, Q., Wang, G., Jia, K., and Yan, X.: Divergent radiative forcing of fine-mode aerosols across tree genera during wildfires in North America and Europe, Journal of Hazardous Materials, 495, 138881, https://doi.org/10.1016/j.jhazmat.2025.138881, 2025.

Christensen, M. W., Neubauer, D., Poulsen, C. A., Thomas, G. E., McGarragh, G. R., Povey, A. C., Proud, S. R., and Grainger, R. G.: Unveiling aerosol–cloud interactions – Part 1: Cloud contamination in satellite products enhances the aerosol indirect forcing estimate, Atmos. Chem. Phys., 17, 13151–13164, https://doi.org/10.5194/acp-17-13151-2017, 2017.

Dai, G., Wu, S., Long, W., Liu, J., Xie, Y., Sun, K., Meng, F., Song, X., Huang, Z., and Chen, W.: Aerosol and cloud data processing and optical property retrieval algorithms for the spaceborne ACDL/DQ-1, Atmos. Meas. Tech., 17, 1879–1890, https://doi.org/10.5194/amt-17-1879-2024, 2024.

Dong, Y., Li, J., Zhang, Z., Zheng, Y., Zhang, C., and Li, Z.: Machine learning-based retrieval of aerosol and surface properties over land from the Gaofen-5 Directional Polarimetric Camera measurements, IEEE Trans. Geosci. Remote Sens., 62, 1–14, https://doi.org/10.1109/TGRS.2024.3419169, 2024.

Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., Eck, T. F., Holben, B. N., Lewis, J. R., Campbell, J. R., et al.: 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, Atmos. Meas. Tech., 12, 169–209, https://doi.org/10.5194/amt-12-169-2019, 2019.

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., et al.: AERONET – A federated instrument network and data archive for aerosol characterization, Remote Sens. Environ., 66, 1–16, https://doi.org/10.1016/S0034-4257(98)00031-5, 1998.

Hsu, N. C., Jeong, M. J., Bettenhausen, C., Sayer, A. M., Hansell, R., Seftor, C. S., Huang, J., and Tsay, S. C.: Enhanced Deep Blue aerosol retrieval algorithm: the second generation, J. Geophys. Res. Atmos., 118, 9296–9315, https://doi.org/10.1002/jgrd.50712, 2013.

Hsu, N. C., Tsay, S. C., King, M. D., and Herman, J. R.: Aerosol properties over bright-reflecting source regions, IEEE Trans. Geosci. Remote Sens., 42, 557–569, https://doi.org/10.1109/TGRS.2004.824067, 2004.

Intergovernmental Panel on Climate Change (IPCC): Short-lived climate forcers, in: Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 817–922, https://doi.org/10.1017/9781009157896.008, 2023.

Kaufman, Y. J., Wald, A. E., Remer, L. A., Gao, B. C., Li, R. R., and Flynn, L. E.: The MODIS 2.1 μm channel – correlation with visible reflectance for use in remote sensing of aerosol, IEEE Trans. Geosci. Remote Sens., 35, 1286–1298, https://doi.org/10.1109/36.628795, 1997.

Knapp, K. R., Frouin, R., Kondragunta, S., and Prados, A.: Toward aerosol optical depth retrievals over land from GOES visible radiances: determining surface reflectance, Int. J. Remote Sens., 26, 4097–4116, https://doi.org/10.1080/01431160500099329, 2005.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

Kotchenova, S. Y., Vermote, E. F., Matarrese, R., and Klemm, F. J.: Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part I: Path radiance, Appl. Opt., 45, 6762–6774, https://doi.org/10.1364/AO.45.006762, 2006.

Lemmouchi, F., Cuesta, J., Lachatre, M., Brajard, J., Coman, A., Beekmann, M., and Derognat, C.: Machine learning-based improvement of aerosol optical depth from CHIMERE simulations using MODIS satellite observations, Remote Sens., 15, 1510, https://doi.org/10.3390/rs15061510, 2023.

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu, N. C.: The Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas. Tech., 6, 2989–3034, https://doi.org/10.5194/amt-6-2989-2013, 2013.

Li, Z., Guo, J., Ding, A., Liao, H., Liu, J., Sun, Y., Wang, T., Xue, H., Zhang, H., and Zhu, B.: Aerosol and boundary-layer interactions and impact on air quality, Natl. Sci. Rev., 4, 810–833, https://doi.org/10.1093/nsr/nwx117, 2017.

Liang, T., Liang, S., Zou, L., Sun, L., Li, B., Lin, H., He, T., and Tian, F.: Estimation of aerosol optical depth at 30 m resolution using Landsat imagery and machine learning, Remote Sens., 14, 1053, https://doi.org/10.3390/rs14051053, 2022.

Litvinov, P., Chen, C., Dubovik, O., Zhai, S., Matar, C., Li, C., Lopatin, A., Fuertes, D., Lapyonok, T., Bindreiter, L., et al.: Synergetic retrieval from multi-mission spaceborne measurements for enhanced aerosol and surface characterization, Atmos. Meas. Tech., 18, 7679–7716, https://doi.org/10.5194/amt-18-7679-2025, 2025.

Lucht, W., Schaaf, C. B., and Strahler, A. H.: An algorithm for the retrieval of albedo from space using semiempirical BRDF models, IEEE Trans. Geosci. Remote Sens., 38, 977–998, https://doi.org/10.1109/36.841980, 2000.

Luo, N., Zhang, Y., Jiang, Y., Zuo, C., Chen, J., Zhao, W., Shi, W., and Yan, X.: Unveiling global land fine- and coarse-mode aerosol dynamics from 2005 to 2020 using enhanced satellite-based monthly inversion data, Environmental Pollution, 348, 123838, https://doi.org/10.1016/j.envpol.2024.123838, 2024a.

Luo, N., Zou, J., Zhang, Z., Chen, T., and Yan, X.: Wide and deep learning model for satellite-based real-time aerosol retrievals in China, Atmosphere, 15, 564, https://doi.org/10.3390/atmos15050564, 2024b.

Lyapustin, A., Wang, Y., Korkin, S., and Huang, D.: MODIS Collection 6 MAIAC algorithm, Atmos. Meas. Tech., 11, 5741–5765, https://doi.org/10.5194/amt-11-5741-2018, 2018.

Man, W., Tao, M., Xu, L., Xu, X., Jiang, J., Wang, J., Wang, L., Wang, Y., Fan, M., and Chen, L.: Improving aerosol retrieval from MISR with a physics-informed deep learning method, IEEE Trans. Geosci. Remote Sens., 62, 1–11, https://doi.org/10.1109/TGRS.2024.3376598, 2024.

Norgren, M. S., Wood, J., Schmidt, K. S., van Diedenhoven, B., Stamnes, S. A., Ziemba, L. D., Crosbie, E. C., Shook, M. A., Kittelman, A. S., LeBlanc, S. E., et al.: Above-aircraft cirrus cloud and aerosol optical depth from hyperspectral irradiances measured by a total-diffuse radiometer, Atmos. Meas. Tech., 15, 1373–1394, https://doi.org/10.5194/amt-15-1373-2022, 2022.

Peng, R., Wang, Q., Jia, K., Cao, B., Dong, W., Chen, Z., Chen, J., and Yan, X.: Satellite On-Orbit Chip-Level Deep Learning Model for Real-Time Dust Storm Monitoring, Environmental Science & Technology, 60(16), 12280–12291, https://doi.org/10.1021/acs.est.5c14697, 2026.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

Probst, P., Wright, M. N., and Boulesteix, A. L.: Hyperparameters and tuning strategies for random forest, Wiley Interdiscip. Rev. Data Min. Knowl. Discov., 9, e1301, https://doi.org/10.1002/widm.1301, 2019.

Qiu, S., Zhu, Z., and He, B.: Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4–8 and Sentinel-2 imagery, Remote Sens. Environ., 231, 111205, https://doi.org/10.1016/j.rse.2019.05.024, 2019.

Roscher, R., Bohn, B., Duarte, M. F., and Garcke, J.: Explainable machine learning for scientific insights and discoveries, IEEE Access, 8, 42200–42216, https://doi.org/10.1109/ACCESS.2020.2976199, 2020.

Sayer, A. M., Hsu, N. C., Bettenhausen, C., and Jeong, M. J.: Validation and uncertainty estimates for MODIS Collection 6 "Deep Blue" aerosol data, J. Geophys. Res. Atmos., 118, 7864–7872, https://doi.org/10.1002/jgrd.50600, 2013.

Seidel, F. C. and Popp, C.: Critical surface albedo and its implications to aerosol remote sensing, Atmos. Meas. Tech., 5, 1653–1665, https://doi.org/10.5194/amt-5-1653-2012, 2012.

Shi, Z., Xie, Y., Li, Z., Zhang, Y., Chen, C., Mei, L., Xu, H., Wang, H., Zheng, Y., Liu, Z., et al.: A generalized land surface reflectance reconstruction method for aerosol retrieval: application to the Particulate Observing Scanning Polarimeter (POSP) onboard GaoFen-5 (02) satellite, Remote Sens. Environ., 295, 113683, https://doi.org/10.1016/j.rse.2023.113683, 2023.

Tao, M., Chen, L., Su, L., and Tao, J.: Satellite observation of regional haze pollution over the North China Plain, J. Geophys. Res. Atmos., 117, D12203, https://doi.org/10.1029/2012JD017915, 2012.

Tian, X., Gao, L., Li, J., Chen, L., Ren, J., and Li, C.: Retrieval of atmospheric aerosol optical depth from AVHRR over land with global coverage using machine learning method, IEEE Trans. Geosci. Remote Sens., 60, 1–13, https://doi.org/10.1109/TGRS.2021.3129853, 2022.

Vermote, E. F., Tanré, D., Deuzé, J. L., Herman, M., and Morcrette, J. J.: Second simulation of the satellite signal in the solar spectrum, 6S: an overview, IEEE Trans. Geosci. Remote Sens., 35, 675–686, https://doi.org/10.1109/36.581987, 1997.

Wang, Y., Zhang, D., Zhang, E., Xiong, Q., Niu, X., and Chen, S.: Pre-launch calibration technology of DQ-1 Wide-Field Imaging Spectrometer, Aerospace Shanghai (Chinese & English), 41, 137–145, https://doi.org/10.19328/j.cnki.2096-8655.2024.01.018, 2024a (in Chinese).

Wang, Z., Zhang, R., Chen, R., and Chen, H.: The aerosol optical depth retrieval from Wide-Swath Imaging of DaQi-1 over Beijing, Atmosphere, 15, 1476, https://doi.org/10.3390/atmos15121476, 2024b.

Wei, J., Huang, W., Li, Z., Xue, W., Peng, Y., Sun, L., and Cribb, M.: Estimating 1 km resolution PM2.5 concentrations across China using the space-time random forest approach, Remote Sens. Environ., 231, 111221, https://doi.org/10.1016/j.rse.2019.111221, 2019a.

Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., and Cribb, M.: Reconstructing 1 km resolution high-quality PM2.5 data records from 2000 to 2018 in China: spatiotemporal variations and policy implications, Remote Sens. Environ., 252, 112136, https://doi.org/10.1016/j.rse.2020.112136, 2021.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

Wei, J., Peng, Y., Mahmood, R., Sun, L., and Guo, J.: Intercomparison in spatial distributions and temporal trends derived from multi-source satellite aerosol products, Atmos. Chem. Phys., 19, 7183–7207, https://doi.org/10.5194/acp-19-7183-2019, 2019b.

Wei, J., Wang, Z., Li, Z., Li, Z., Pang, S., Xi, X., Cribb, M., and Sun, L.: Global aerosol retrieval over land from Landsat imagery integrating Transformer and Google Earth Engine, Remote Sens. Environ., 315, 114404, https://doi.org/10.1016/j.rse.2024.114404, 2024.

Wei, X., Chang, N.-B., Bai, K., and Gao, W.: Satellite remote sensing of aerosol optical depth: advances, challenges, and perspectives, Crit. Rev. Environ. Sci. Technol., 50, 1640–1725, https://doi.org/10.1080/10643389.2019.1665944, 2020.

Xu, H.: Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery, Int. J. Remote Sens., 27, 3025–3033, https://doi.org/10.1080/01431160600589179, 2006.

Yan, X., Chen, J., Chen, H. W., Zhang, Y., Jiang, Y., Wang, Q., Wang, G., Jia, K., Chen, Z., Dong, W., Fan, J., and Zhao, C.: Exploring Global Aerosol Size Dynamics from 2001 to 2024 Using the Pretrained Remote sensing pIxel-based Spatial-teMporal Deep Neural Network, Environ. Sci. Technol., 60, 7146–7158, https://doi.org/10.1021/acs.est.5c13469, 2026.

Yan, X., Chen, J., Zuo, C., Jiang, Y., and He, B.: Spectral deconvolution algorithm for global fine-mode aerosol retrieval in the 1990s using dual-angle satellite aerosol data, IEEE Transactions on Geoscience and Remote Sensing, 61, 4107811, https://doi.org/10.1109/TGRS.2023.3332316, 2023.

Yan, X., Lyu, H., Chen, Z., Peng, R., Chen, J., Zou, J., Dong, W., and Wang, Q.: Real-Time Remote sensing for Sudden surface Anomalies: A Review of principles and challenges, Space Science & Technology, 5, https://doi.org/10.34133/space.0362, 2025.

Yan, X., Zang, Z., Li, Z., Chen, H. W., Chen, J., Jiang, Y., Chen, Y., He, B., Zuo, C., Nakajima, T., and Kim, J.: Deep Learning with Pretrained Framework Unleashes the Power of Satellite-Based Global Fine-Mode Aerosol Retrieval, Environmental Science & Technology, 58(32), 14260-14270, https://doi.org/10.1021/acs.est.4c03456, 2024.

Yan, X., Zang, Z., Li, Z., Luo, N., Zuo, C., Jiang, Y., Li, D., Guo, Y., Zhao, W., Shi, W., and Cribb, M.: A global land aerosol fine-mode fraction dataset (2001–2020) retrieved from MODIS using hybrid physical and deep learning approaches, Earth System Science Data, 14(3), 1193–1213, https://doi.org/10.5194/essd-14-1193-2022, 2022.

Yang, Y., Li, Z., Guo, J., Wang, Y., Wu, H., Shang, Y., Wang, Y., Zhu, L., and Yan, X.: Revolutionizing Clear-Sky Humidity Profile Retrieval with Multi-Angle-Aware Networks for Ground-Based Microwave Radiometers, Journal of Remote Sensing, 5, https://doi.org/10.34133/remotesensing.0736, 2025.

Zang, Z., Zhang, Y., Zuo, C., Chen, J., He, B., Luo, N., Zou, J., Zhao, W., Shi, W., and Yan, X.: Exploring Global Land Coarse-Mode Aerosol Changes from 2001-2021 Using a New Spatiotemporal Coaction Deep-Learning Model, Environmental Science & Technology, 57(48), 19881-19890, https://doi.org/10.1021/acs.est.3c06097, 2023.

Zhang, H. K., Luo, D., and Roy, D. P.: Improved Landsat Operational Land Imager (OLI) cloud and shadow detection with the Learning Attention Network Algorithm (LANA), Remote Sens., 16, 1321, https://doi.org/10.3390/rs16081321, 2024.

https://doi.org/10.5194/egusphere-2026-2510
Preprint. Discussion started: 18 August 2026

Zhou, D., Wang, Q., Li, S., and Yang, J.: Preliminary retrieval and validation of aerosol optical depths from FY-4B Advanced Geostationary Radiation Imager images, Remote Sens., 16, 372, https://doi.org/10.3390/rs16020372, 2024.

Zou, J., Chen, H. W., Li, H., Wang, Q., Wang, G., Jia, K., Chen, Z., Zhao, C., Shi, W., Yang, Y., Tang, Y., Chen, J., Zhang, Y., Xu, T., Wang, Y., Liu, G., and Yan, X.: Amplified urban heat island effect in southern China's old towns following atmospheric regulation policies, Sustainable Cities and Society, 106675, https://doi.org/10.1016/j.scs.2025.106675, 2025.