International Journal of Applied Earth Observation and Geoinformation 152 (2026) 105393

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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

Hybrid physics–AI aerosol property retrieval algorithm for AMI/GK-2A with a deep learning radiative transfer emulator

Minseok Kim [a,b,c], Jhoon Kim [c,*], Hyunkwang Lim [d], Seoyoung Lee [e,f], Hyeji Cha [c], Yujin Chai [c], Sang Seo Park [a], Jun Wang [g,h]

[a] Department of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, the Republic of Korea
[b] AI-Transformed Aerospace InnoCORE Research Center, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, the Republic of Korea
[c] Department of Atmospheric Sciences, Yonsei University, Seoul 03722, the Republic of Korea
[d] National Institute for Environmental Studies (NIES), Tsukuba 305-0053, Japan
[e] Goddard Earth Sciences Technology and Research (GESTAR) II, University of Maryland Baltimore County (UMBC), Baltimore, MD 21250, USA
[f] Climate and Radiation Laboratory, NASA Goddard Space Flight Center (GSFC), Greenbelt, MD 20771, USA
[g] Department of Chemical and Biochemical Engineering, The University of Iowa, Iowa City, IA 52242, USA
[h] Center for Global and Regional Environmental Research, The University of Iowa, Iowa City, IA 52242, USA

A R T I C L E I N F O

Keywords:
Satellite
Remote sensing
Deep learning
Radiative transfer
Aerosol optical depth

A B S T R A C T

Computational advances have enabled the evolution and refinement of satellite aerosol retrieval algorithms to mitigate issues in previous versions and improve their results. This study addressed a fundamental limitation in operational satellite remote sensing which is the trade-off between computational efficiency and retrieval accuracy in conventional look-up table (LUT)-based algorithms. The hybrid-Yonsei Aerosol Retrieval (hybrid-YAER) algorithm for the Advanced Meteorological Imager (AMI) incorporates a flexible aerosol model input, which became available by using a deep learning (DL) radiative transfer model (RTM) that replaces the interpolation-based inversion with LUTs. The DLRTM, trained with 100,000 Vector LInearized Discrete Ordinate Radiative Transfer (VLIDORT) code simulations using a Set Transformer architecture, accurately reproduces top-of-atmosphere reflectance (the root-mean-square error is about 0.002 for Mie calculations) at computation times about 0.1% of those of a conventional RTM. Additionally, a dust call-back procedure and an extended surface reflectance database based on multi-year minimum reflectance are introduced in the hybrid-YAER algorithm. Validation against AERONET and comparison with Visible Infrared Imaging Radiometer Suite (VIIRS) Dark Target (DT) and Deep Blue (DB) AOD products demonstrate that the new algorithm mitigates underestimation of low values of aerosol optical depth (AOD) and discontinuities of AOD values between adjacent land and ocean, while improving overall correlation (R increased from 0.757 to 0.767 for land pixels and from 0.856 to 0.881 for ocean pixels) and reducing bias. Over ocean, bias patterns become more linear, and the retrieval stability is improved (R^2 increased from 0.88 to 0.94) compared to the original version. Uncertainty analysis shows that retrieval uncertainty scales linearly with AOD, confirming the robustness of the hybrid algorithm. The concept of advancement in aerosol retrieval via hybrid algorithm can be applied to any other aerosol remote sensing algorithms.

1. Introduction

Various algorithms for the retrieval of aerosol optical properties (AOPs) from satellite data have been developed since the introduction of the aerosol optical depth (AOD) retrieval algorithm for data from the Advanced Very High Resolution Radiometer aboard National Oceanic and Atmospheric Administration (NOAA) satellites (Kaufman et al., 1990; Higurashi and Nakajima, 1999; King et al., 1999). Computational advances and accumulating observations have facilitated continuous modification of these algorithms to improve their speed and accuracy. NASA introduced the Dark Target (DT) and Deep Blue (DB) algorithms with its Moderate Resolution Imaging Spectroradiometer (MODIS) (Remer et al., 2005; Hsu et al., 2006; Levy et al., 2007). Improvements for MODIS Collection 6 included enhanced methodologies for surface

* Corresponding author.
E-mail address: jkim2@yonsei.ac.kr (J. Kim).

https://doi.org/10.1016/j.jag.2026.105393
Received 5 March 2026; Received in revised form 18 May 2026; Accepted 2 June 2026
Available online 19 June 2026

reflectance estimation and aerosol model selection (Levy et al., 2013). The DT and DB algorithms were implemented in the Visible Infrared Imaging Radiometer Suite (VIIRS) to continue the MODIS aerosol record (Sayer et al., 2020; Hsu et al., 2019; Lee et al., 2024). The DB algorithm was also applied to data from the Sea-viewing Wide Field-of-view Sensor (Hsu et al., 2013). Meanwhile, the OMAERUV algorithm, based on Ozone Monitoring Instrument observations, was introduced by Torres et al. (2007) and updated by Torres et al. (2013) using additional A-Train constellation datasets. Torres et al. (2018) further enhanced the OMAERUV algorithm by incorporating realistic spheroidal shape distributions for dust particles and accounting for the effects of angular scattering by clouds in the ultraviolet (UV) aerosol index.

The Generalized Retrieval of Aerosol and Surface Properties (GRASP) algorithm is one of the most widely used aerosol retrieval frameworks. It retrieves aerosol properties from spectral, multi-angular, and polarimetric observations, and is sufficiently versatile that diverse satellite observation datasets can be used as inputs (Dubovik et al., 2021). Consequently, both low Earth orbit (LEO) and geostationary Earth orbit (GEO) instruments have adopted GRASP for retrieving AOPs. For example, Chen et al. (2020) used GRASP to obtain aerosol products using Polarization and Directionality of the Earth's Reflectances (POLDER) observations. The POLDER/GRASP AOD product proved to be comparable to the MODIS AOD product, but provided more information on other aerosol properties such as spectral fine-mode AOD, coarse-mode AOD, and Ångström exponent utilizing multi-angular polarimetry. Chen et al. (2022) then used the GRASP algorithm to process Ocean and Land Colour Imager (OLCI)/Sentinel-3A observations.

Recent studies have increasingly incorporated machine learning into aerosol retrievals (Remer et al., 2024). Early applications primarily involved the post-processing of existing AOD products (Lipponen et al., 2022; Cho et al., 2024; Kim et al., 2024), whereby models were trained to learn either reference AODs or errors in AODs using ground-based reference data (e.g., AErosol RObotic NETwork; AERONET). Kim et al. (2025) analytically derived aerosol fine-mode fractions using post-processed spectral AODs. Other studies trained machine learning models directly on Level 1B observations to retrieve AODs without an explicit physical inversion process (She et al., 2020; Yeom et al., 2022; Chen et al., 2022; Kang et al., 2022). Retrievals based on machine learning were comparable to those from physical algorithms. However, concerns such as the lack of physical interpretability of these "black box" models which can complicate uncertainty estimation of the AOD product remains. Moreover, when AOD products from these machine learning inversion models are used to retrieve other atmospheric variables, such as aerosol effective height (Park et al., 2025) or trace gases (Leitão et al., 2010), the neglect of radiative transfer processes can lead to inconsistencies between those variables and the AOD. This is because the AOD from machine learning model cannot physically restore the original observation which has to be used for another retrieval process. Shortly, previous studies reveal an important trade-off: purely data-driven inversion can improve efficiency and empirical skill, whereas physical interpretability, transferability outside the training distribution, and uncertainty characterization become more difficult. In addition, improvements in aerosol products for these data-driven algorithms rely mainly on general strategies, such as improving training datasets or adopting new machine learning frameworks. However, maintaining a physical retrieval framework allows targeted updates for specific error sources.

A more advanced approach to the use of machine learning for aerosol retrieval is to employ a deep learning (DL) model to emulate a radiative transfer model (RTM) (Man et al., 2024, Jiang et al., 2023; Tao et al., 2023). Traditional aerosol retrieval algorithms rely on pre-calculated RTM results in look-up tables (LUTs), because running an RTM online during retrieval is computationally expensive. To overcome this limitation, hybrid approaches have been proposed that maintain a physical retrieval framework while accelerating RTM calculations with machine learning. For example, Gao et al. (2021) developed the fast Multi-Angular Polarimetric Ocean coLor (fastMAPOL) algorithm, which retrieves AOPs using a deep learning radiative transfer model (DLRTM) to simulate top-of-atmosphere (TOA) reflectance and the degree of linear polarization. Wang et al. (2025) has also constructed machine learning-based radiative transfer model using Back Propagation Neural Network to retrieve single scattering albedo from Medium Resolution Spectral Imager-II (MERSI-II) aboard Fengyun-3D satellite. Yao et al. (2026)** has also used neural network to overcome the 'curse of dimensionality' problem in the LUT-based retrieval algorithm. However, such an advanced approach has yet to be introduced to an existing aerosol retrieval algorithm. When it comes to the DL-enhancement of aerosol retrieval by emulating RTMs, the ability of the emulator to accurately represent the nonlinearity is important. For the sake of the nonlinearity explanation, various DL frameworks developed and used in various other fields need to be considered. In addition to the neural network (Jin and Xu, 2025), Gaussian process regression (Jin and Xu, 2024), ensemble/composite method (Xu, 2020), regression trees (Zhang and Xu, 2022), least square boosting (Zhang and Xu, 2021), support vector regression (Alade et al., 2021), and Transformer-based architectures (Vaswani et al., 2017; Lee et al., 2023) were shown to have excessive ability to represent nonlinearities from variety of researches.

The Yonsei AErosol Retrieval (YAER) algorithms for GEO satellites have also been progressively updated. The original YAER algorithm was developed for the Geostationary Ocean Color Imager (GOCI) by Lee et al. (2010), based on the methodology of Kim et al. (2008). Choi et al. (2018) updated the GOCI YAER algorithm. GOCI YAER algorithm version 2 included an improved surface reflectance database for near-real-time processing and a refined cloud detection process; it outperformed version 1 over both land and ocean. The YAER algorithm was then applied to the second generation of GOCI (GOCI-II) to enhance the long-term aerosol dataset (Lee et al., 2023). The YAER algorithm was also implemented to retrieve AOD from the Advanced Himawari Imager (AHI), which has fewer visible (VIS) spectral bands than GOCI, but includes infrared (IR) channels (Lim et al., 2018). Aerosol products from meteorological imagers such as AHI provide broader spatial coverage over Asia than GOCI, which is limited to East Asia. The AHI YAER algorithm was subsequently ported to the Advanced Meteorological Imager (AMI) and the Advanced Geostationary Radiation Imager (AGRI) with minimal adjustments (Kim et al., 2024).

In this study, a DLRTM emulator is introduced for a hybrid aerosol inversion process to the existing AMI YAER algorithm. The AMI hybrid-YAER algorithm also mitigates known issues in the AMI original-YAER algorithm such as discontinuities between land and ocean owing to a positive bias over ocean pixels, and cloud detection procedure sometimes misclassifying thick dust as cloud (Kim et al., 2024). The remainder of this paper is organized as follows. Section 2 describes the datasets used, and section 3 provides details of the algorithm updates. Section 4 presents validation results through comparison with the AMI original-YAER and VIIRS DT/DB products. Finally, section 5 provides a summary of the findings and outlines prospects for future algorithm development.

2. Data

2.1. GEO-KOMPSAT-2A (GK-2A)/AMI

AMI is a meteorological payload aboard the GEO-KOMPSAT-2A (GK-2A) satellite. AMI follows the mission of the previous Meteorological Imager, continuously monitoring the Asia–Pacific region with advanced optical performance and high temporal resolution. It provides 52 official products, including surface and atmospheric variables. Two independent scanning mirrors pivot in the east–west and north–south directions to acquire full-disk images of the Earth. Within the optical system, two beam splitters divide incoming light into three wavelength ranges: six channels at VIS–near IR (NIR), and five each at mid-wave IR and longwave IR. Focal plane modules detect incoming radiance and convert it to electrical signals. AMI observes solar-reflected Earth radiance across the 16 VIS–IR bands. Full-disk observations are performed every 10 min (Huh and Jin, 2022). The red channel has a spatial resolution of 0.5 km, while the remaining VIS–NIR channels have a resolution of 1.0 km. The other IR channels are observed at 2.0 km resolution. The instrument features on-board black-body and space-view calibration systems to ensure radiometric stability across all IR channels, while the VIS–NIR bands are routinely cross-calibrated with NOAA-20 and SNPP/VIIRS. AMI was launched on 5 December 2018 and began operational service on 25 July 2019. Table 1 provides a summary of the instrument specifications.

2.2. NOAA-20/VIIRS DT/DB AOD

NASA introduced the DT algorithm with MODIS Collection 4 (Remer et al., 2005). The algorithm was updated for MODIS Collection 5, with improvements to its surface reflectance estimation, aerosol model optical properties, and LUTs (Levy et al., 2007). The DB algorithm, which employs different surface reflectance estimation and aerosol model assumptions, was also introduced in MODIS Collection 5 (Hsu et al., 2006). In particular, the DB algorithm utilizes spectral bands centered near 412 nm to aid the quantification of aerosol loading over bright surfaces. Both algorithms were further refined in MODIS Collection 6, with enhanced surface reflectance estimation and aerosol model selection methodologies (Levy et al., 2013). They were implemented in VIIRS to ensure continuity of the long-term MODIS aerosol record. This study compares VIIRS AOD from NASA's DT and DB algorithms with AMI AOD products. While VIIRS DT/DB AOD has been updated to version 2 (Levy et al., 2024; Lee et al., 2024), this study uses version 1 AOD because the version 2 product was made available only recently. NASA's Atmosphere Archive & Distribution System Distributed Active Archive Center has offered DT since November 2023 and DB since June 2023. Collocation between VIIRS and AMI is performed by remapping the geolocation of VIIRS to that of AMI, selecting the closest VIIRS pixel for each AMI pixel.

2.3. AERONET AOD

AERONET is a network of ground-based aerosol remote sensing instruments, consisting of numerous sun photometer stations operating worldwide. AERONET version 3 (V3), level 2.0 AOD products are used as the validation reference for AMI YAER AOD. The uncertainty of AERONET AOD is 0.01–0.02 (Holben et al., 1998; Eck et al., 1999). All AERONET data used in this study are cloud-screened and quality assured following the standard AERONET V3 protocol. During the validation period (2022–2023), a total of 71 stations reported ground-based level 2.0 AOD. To match the AERONET and spaceborne wavelengths, AOD at 550 nm was derived through quadratic interpolation from AODs measured at 340, 380, 440, 500, 675, 870, and 1020 nm. For spatiotemporal collocation, satellite AODs within a 25 km radius of each AERONET site were averaged, and AERONET measurements within ±30 min of each exact hour was also averaged. Fig. S1 shows the spatial distribution of the AERONET stations used for validation of the AMI YAER AOD. The AERONET sites cover diverse surface types and aerosol conditions across East and South Asia, ensuring robust regional representativeness of the validation results.

Table 1
AMI measurement specifications (Kim et al., 2021).

Satellite	GeoKompsat-2A
Launch	December 2018
Lifetime	10 years
Location	128°E
Wavelength range	0.4–13.3 μm
Channels	4 visible, 2 shortwave infrared, and 10 infrared bands
Spatial resolution	0.5 km (red), 1.0 km (blue, green, near-infrared), and 2.0 km (infrared)
Temporal resolution	10 min
Full-width half maximum (FWHM)	10–20 nm

3. AMI hybrid-YAER algorithm

3.1. Overview of the AMI YAER algorithm

Fig. 1 presents the workflow of the AMI hybrid-YAER algorithm, based on the AHI (Lim et al., 2018) and AMI (Kim et al., 2024) algorithms. The retrieval process consists of three primary steps: (i) masking of bright pixels including clouds; (ii) pixel aggregation; and (iii) inversion with aerosol-type selection. Inputs to the inversion include surface reflectance, model AOPs, and TOA reflectances simulated by an RTM. Compared with the original-YAER AMI algorithm of Kim et al. (2024), this study introduces four major updates: (1) a dust call-back procedure, (2) improved surface-reflectance estimation, (3) revised aerosol model assumptions, and (4) a hybrid inversion process with DLRTM online simulation.

3.2. Pixel masking and aggregation

As aerosol retrieval algorithms rely primarily on VIS spectral observations, highly reflective clouds can block the underlying aerosol signal. Consequently, most aerosol algorithms exclude cloud-affected pixels from retrieval. Critical reflectance is defined as the surface reflectance at which an increase in aerosol loading no longer affects the radiance observed by passive satellite sensors (von Hoyningen-Huene et al., 2011; Zhu et al., 2011; Kim et al., 2014). As this critical reflectance is generally higher than the reflectance of most land surfaces on Earth, aerosol retrieval algorithms typically exclude bright-surface pixels from processing. In the AMI YAER algorithm, bright surfaces over land include snow, inland water, and deserts, whereas over the ocean they include sun-glint and turbid-water regions. AMI provides IR spectral bands that enable effective separation of cloud signals from those of the surface and atmosphere. After identifying the pixels eligible for AOD retrieval, the AMI YAER algorithm aggregates arrays of 6 × 6 pixels into single retrieval units. Within each aggregation window, the brightest 40% and the darkest 20% of valid pixels are discarded, and the remaining pixels are averaged. The removal of the brightest pixels mitigates residual cloud contamination, while excluding the darkest pixels minimizes the influence of cloud shadows.

3.2.1. Dust call-back

During the cloud masking procedure, thick dust plumes can occasionally be misclassified as clouds and erroneously removed. To recover such cases, AMI hybrid-YAER implements a dust call-back procedure, allowing AOD retrievals to be retained for thick dust events. Detection of these plumes follows the brightness-temperature criteria recommended by Hansell et al. (2007). Table 2 summarizes the specific pixel-masking thresholds used in the AMI hybrid-YAER algorithm. The thresholds were selected because they are physically tied to the distinct thermal-infrared signature of mineral dust relative to water and ice clouds, and they performed robustly for the East Asian dust cases examined in this study. Nevertheless, their robustness is not universal: the separability of dust and cloud can vary with surface temperature, atmospheric humidity, dust altitude, and mixed dust–cloud scenes. Accordingly, the current call-back is intended as a conservative recovery step for optically thick dust, and some false positives remain possible under spectrally ambiguous conditions.

3.3. Surface reflectance estimation

3.3.1. Minimum reflectance method update

The original AMI YAER algorithm considers a minimum reflectance

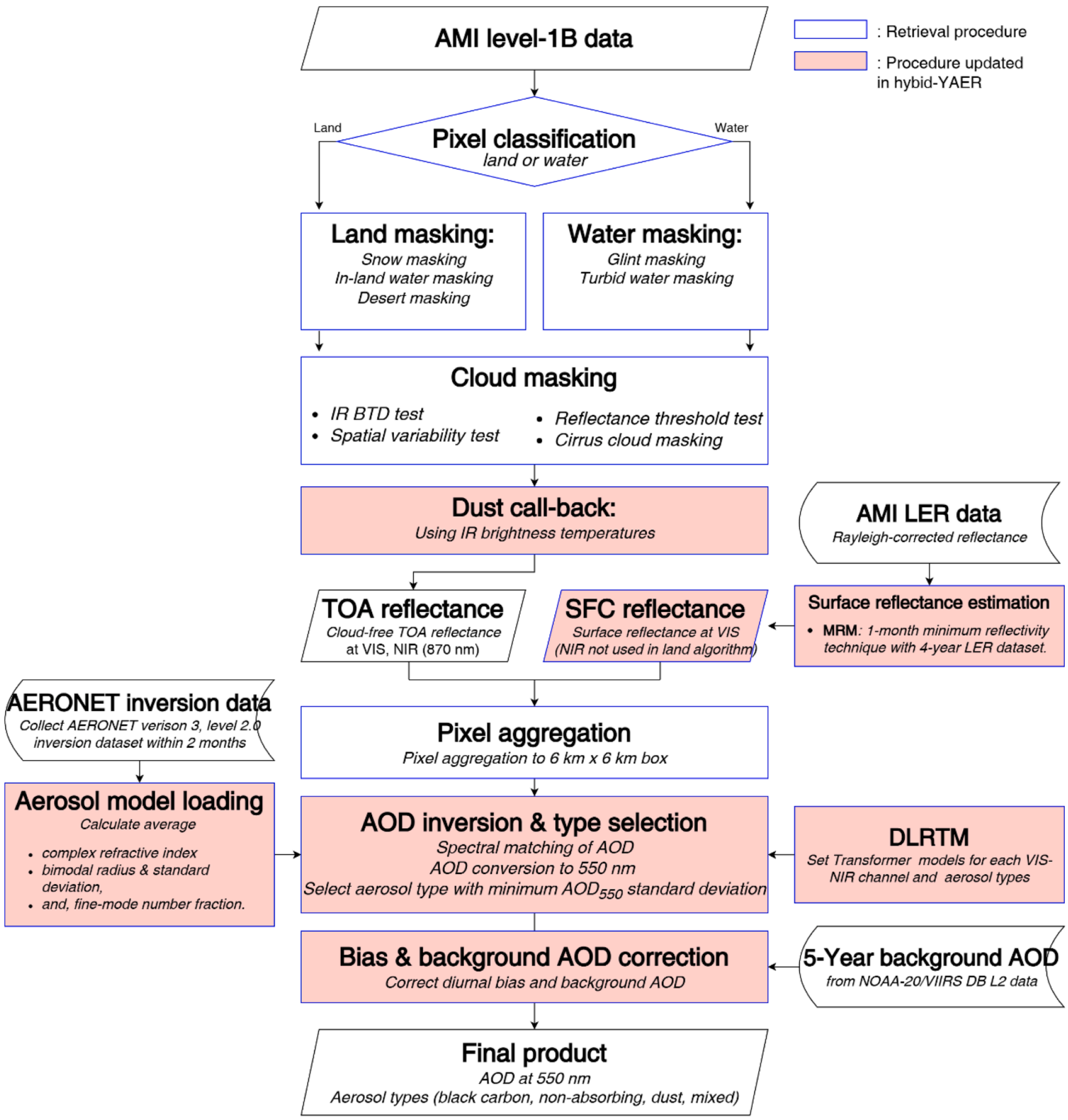

Fig. 1. Flowchart for AMI hybrid-YAER AOD algorithm. Updated procedures in hybrid-YAER are indicated with a red background. BTD, LER, and MRM stand for bright temperature difference, Lambertian equivalent reflectance, and minimum reflectance method. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

method (MRM) database constructed using land surface reflectance (LER) data accumulated in a period of only 2 years during the early mission. Too short a period of LER accumulation can allow persistent background aerosol loading to bias the derived surface reflectance toward higher values. Furthermore, regions or seasons with frequent cloud cover can introduce cloud contamination into the LER database, again leading to the overestimation of surface reflectance. The propagation of these effects into the retrieval process leads to AOD underestimation, particularly under conditions of low aerosol loading when surface contributions dominate the observed signal. Choi et al. (2018) improved the accuracy of GOCI AOD by extending the minimum-reflectance compositing period from 1 to 5 years, while Su et al. (2022) demonstrated similar improvements in AOD retrieval by using 11 years of MODIS observations to derive band-wise surface-reflectance relationships.

3.3.2. *Land–ocean AOD discontinuity mitigation*

The AMI original-YAER simulated ocean surface reflectance using the model of Cox and Munk (1954), incorporating the effects of wind speed and whitecaps. While this physically based approach is robust, systematic bias between Level 1B calibration and RTM simulations can lead to mismatches between the model and observations. AMI original-YAER had Level 1B TOA reflectances in its VIS channels systematically higher than those simulated using a linearized pseudo-spherical vector discrete ordinate radiative transfer code, (VLIDORT; Spurr, 2006) with the Cox and Munk model. Consequently, Kim et al. (2024) reported a systematic positive AOD bias over ocean in the AMI original-YAER AOD product. However, the land AOD did not show such a positive bias, given the different surface estimation schemes for land and ocean. As a result, AMI original-YAER generated distinct discontinuities in the AOD over land and ocean. The MRM used for land surface estimation constructs its database from long-term observations, which helps compensate for calibration-related biases. Therefore, this study adopts MRM-based ocean surface reflectance to mitigate the positive bias over ocean and alleviate the discontinuity in AOD over land and ocean.

Table 2
Cloud masking and dust-call back criteria used in this study.

	Cloud masking criteria	Classification
Land masking	$NDSI = \frac{R0.5 - R1.6}{R0.5 + R1.6} > 0.14$	Snow
	$NDVI = \frac{R0.8 - R0.6}{R0.8 + R0.6} < 0.01$	In-land water
	$R1.6 > 0.2$ and $BT11 - BT14 < -5$ K, or $R1.6 > 0.2$ and $R0.8 / R1.6 < 0.77$	Desert
Water masking	Glint angle $= cos^{-1}(cos(SZA)cos(VZA) + sin(SZA)sin(VZA)cos(RAA)) < 25°$	Sun glint
	$R0.6 - R0.4 + \frac{R1.6 - R0.4}{(R1.6 - R0.4)(R0.6 - R0.4)} > 0.01$	Turbid water
Cloud masking	$R0.4 > 0.35$	
	$R1.3 > 0.025$	Cirrus
	$BT15 - BT16 < 10$ K	High cloud
	$BT13 - BT16 < 11$ K	Low cloud
	Standard deviation of 3×3 pixels in R0.5 or R0.8 > 0.0025	
	Average weighted standard deviation of 3×3 pixels in $R0.4 > 0.0018$	
	BT09 (BT14) − Maximum BT09 (BT14) in previous 10 days < −28 K	
	Dust call-back criteria	
Dust call-back	$BT038 - BT105 > 10$ K	
	$BT105 - BT123 < -1.0$ K	
	$BT086 - BT105 > -0.5$ K	
	Dust index $= \frac{(BT112 - BT123 + 0.5\,K)}{(BT086 - BT112 - 15\,K)} > 0.94$	

3.4. AOD inversion scheme

The principal update in AMI hybrid-YAER is the replacement of the conventional LUT-based inversion with an online RTM driven by deep learning. As executing a full RTM online for all pixels is computationally prohibitive, traditional algorithms precompute a LUT at discrete nodes for each variable and then interpolate TOA reflectances to retrieve AOD. This interpolation linearizes a fundamentally nonlinear inversion across multiple nodes, so an insufficient node density will lead to interpolation errors. Constructing the nodes at fine resolution alleviates such errors but has enormous computational cost. Moreover, AOPs such as the complex refractive index, effective radius and its standard deviation, and fine-mode fraction were prescribed for each aerosol type (black carbon, non-absorbing, dust and mixed) and fixed once the LUT was generated. In contrast, the speed at which the DLRTM reproduces RTM outputs enables online calculations with flexible and scene-specific aerosol property inputs. This greatly reduces computational cost while allowing AOPs to be dynamically adjusted to match actual observation conditions. Accordingly, AMI hybrid-YAER employs DLRTM for efficient online RTM calculations.

For each pixel, given its geometry, terrain height, and surface reflectance conditions, the DLRTM computes TOA reflectances at AOD nodes of 0, 0.5, 1.0, 2.0, and 3.0 for each of the four aerosol types. For each band and aerosol type, the observed TOA reflectance is matched to these DLRTM-simulated values using piecewise linear interpolation between the two bracketing AOD nodes; when the observation lies outside the simulated range, the nearest endpoint is used. Thus, the interpolation in hybrid-YAER is one-dimensional in AOD after the forward simulation, rather than multidimensional interpolation across a precalculated multi-dimensional LUT space. All spectral AODs are subsequently converted to 550 nm using type-specific climatological AE, and the standard deviation of the resulting 550-nm AODs across channels is calculated. The two aerosol types with the smallest standard deviations are then selected, and an inverse-variance weighted average of their 550-nm AODs is taken to obtain the final retrieval. By replacing LUT interpolation with DLRTM-based online calculations, the algorithm avoids the interpolation errors inherent to the five-dimensional LUT space defined by three geometric dimensions, along with one dimension each for terrain height and surface reflectance.

The deep learning model used for the DLRTM is the Set Transformer, an attention-based neural network (Lee et al., 2019). Attention mechanisms allow the model to focus selectively on the most informative parts of the input. Transformer architecture processes sequences using attention alone, and has been widely adopted in natural language processing (Vaswani et al., 2017). The Set Transformer extends this framework to permutation-invariant set inputs, ensuring that output remains independent of input ordering; its compressed-attention formulation reduces computational cost and accelerates training.

3.4.1. Replacement of LUT interpolation with DLRTM

We used VLIDORT version 2.7 (Spurr, 2006) as the RTM to generate training targets for the deep learning model. Random samples were drawn within the parameter ranges listed in Table 3 to produce 100,000 input combinations, and VLIDORT was used to compute TOA reflectances at 470, 510, 640, and 865 nm. Two particle-shape configurations were considered: spherical aerosols simulated using Mie theory, and non-spherical aerosols simulated using T-matrix calculations (Dubovik et al., 2006). The T-matrix setup modeled aerosols as ellipsoids with an aspect ratio of 1.4, consistent with the aspect ratios of approximately 1.2–1.5 reported at desert AERONET sites (Nie and Mao, 2023).

The Set Transformer is an attention-based neural network designed to process unordered inputs while preserving permutation invariance (Lee et al., 2023). Unlike conventional feed-forward architectures, it explicitly models interactions among input variables through self-attention, allowing the network to learn complex nonlinear dependencies that arise from coupled physical parameters. We tested performances of widely used machine learning frameworks such as multi-layer perceptron (MLP), residual MLP, and XGBoost for an RTM simulation of spherical aerosols at 470 nm (Fig. S1 and Table S1). The Set Transformer outperformed the other machine learning models in both training efficiency and test dataset evaluation. Therefore, the Set Transformer using an encoder–decoder architecture is adopted in this study. As shown in Fig. 2, the encoder consists of two stacked Induced Set Attention Blocks (ISAB), which project the input features into a latent space while capturing global relationships among variables. The use of inducing points enables efficient approximation of full self-attention, substantially reducing computational cost while retaining expressive power. The resulting latent representations are subsequently aggregated by Pooling by Multihead Attention (PMA), which performs learnable attention-based pooling to produce a compact summary of the input set. This pooled representation is further refined using successive Self-Attention Blocks (SAB), and a final linear layer maps the output to the target variable. Model training is performed using supervised learning with the mean absolute error as the loss function and the Adam optimizer. During each iteration, predictions are obtained through forward propagation, and parameter updates are carried out via back-propagation through all attention and linear layers. The hidden dimension is 128, the number of attention heads is eight, the optimizer is

Table 3
Input variables and their sampling ranges for training the DLRTM.

Variable (unit)	Value range
Solar zenith angle (°)	0–80
Viewing zenith angle (°)	0–80
Relative azimuth angle (°)	0–180
Terrain height (km)	0–10
Surface reflectance	0–0.3
Aerosol optical depth	0–6
Fine-mode mean radius (μm)	0.06–0.1
Coarse-mode mean radius (μm)	0.15–2
Fine-mode radius standard deviation (μm)	1.25–2
Coarse-mode radius standard deviation (μm)	1.5–2.3
Fine-mode number fraction	0.99–1.00
Real part of refractive index	1.35–1.6
Imaginary part of refractive index	0.0005–0.02
Aerosol peak height (km)	0.01–6

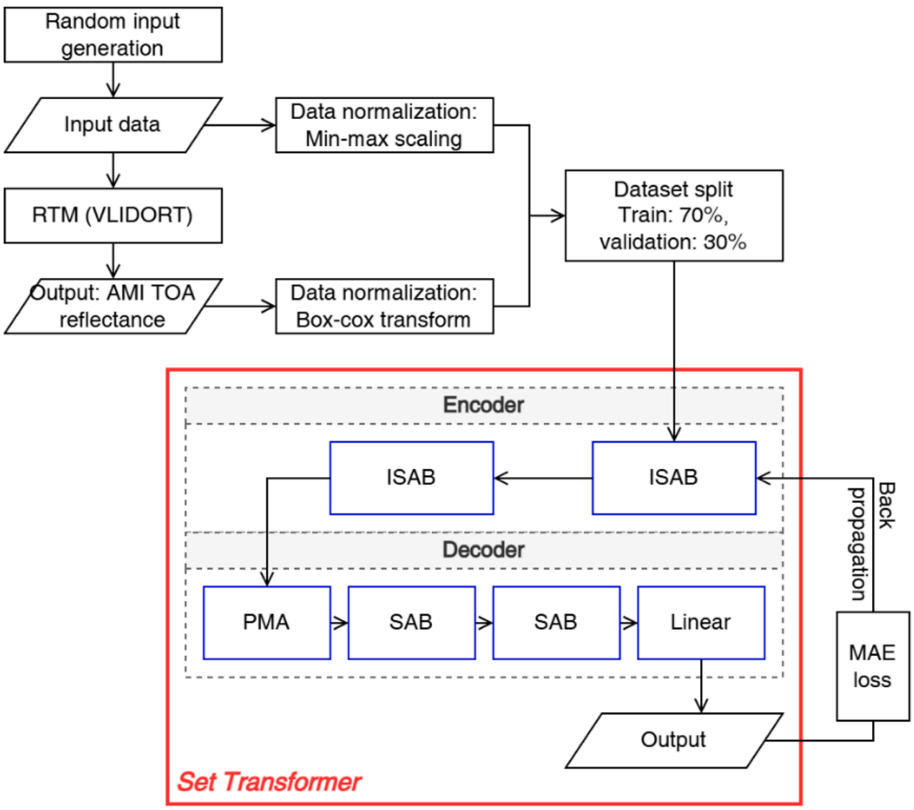

Fig. 2. Flowchart of DLRTM training. ISAB, PMA, SAB, and MAE refers to Induced Set Attention Blocks, Pooling by Multihead Attention, Self-Attention Blocks, and Mean Absolute Error.

Adam, the learning rate is 1 × 10^{-5}, the batch size is 1024, and each model is trained for 5000 epochs using mean absolute error as the loss function.

Eight DLRTMs based on the Set Transformer were trained: each of the two particle shapes (Mie: spherical; T-matrix: ellipsoidal) was modeled at each of the four wavelengths. All eight models share the same architecture so that differences in performance can be attributed primarily to the complexity of the underlying radiative transfer mapping rather than to architectural changes. Fig. 3 shows the validation results for each of the eight DLRTMs. VLIDORT simulated an additional 30,000 random samples not included in training, which were used to evaluate model performance. The DLRTMs reproduced Mie calculations more accurately than T-matrix calculations, with root-mean-square errors (RMSEs) of ~0.002 for Mie and >0.005 for T-matrix cases. The higher RMSEs for the T-matrix cases indicate that the angular dependence of non-spherical scattering introduces additional complexity that cannot be fully captured by the current deep learning architecture. This does not directly translate one-to-one into AOD retrieval error, because the

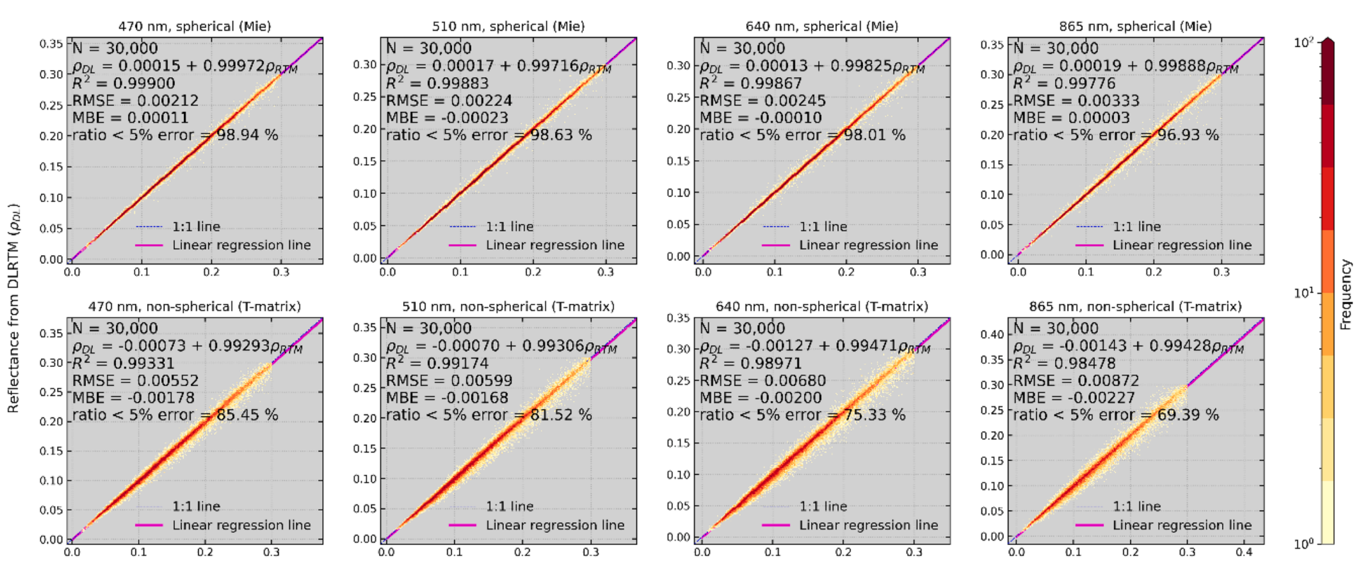

Fig. 3. Two-dimensional histograms for the validation of DLRTMs against reference VLIDORT output for Mie calculation (upper row) and T-matrix calculation (lower row). Columns show results for each band (470, 510, 640, and 865 nm).

inversion still uses multi-band consistency tests and aerosol-type selection, which partly suppresses forward-model noise. Nevertheless, retrievals dominated by the dust model should be interpreted as having higher forward-model uncertainty than corresponding Mie-based cases. The error propagation from the T-matrix DLRTM is assessed in Section 4.2, with validation of pixels which showed dust as a primary aerosol type. Using a single Center Processing Unit (CPU), calculating 100 samples takes ~143 s for VLIDORT and ~0.1 s for the DLRTM, which is ~0.1% computation time. These timings refer to CPU-only inference; thus, the reported speedup does not rely on GPU acceleration. The CPU model is Intel(R) Xeon(R) Gold 6254 CPU @ 3.10 GHz. The DL framework used for this study is PyTorch version 2.4.0. Although generating the offline VLIDORT training set is computationally expensive, it is a one-time cost that is amortized during operational processing, where the much lower online latency is the relevant factor. The maximum memory footprint and total processing time of the AMI hybrid-YAER algorithm is approximately 35.1 GB and 30 min, which is sufficiently small for deployment in near-real-time processing.

3.4.2. *Flexible aerosol model assumption*

LUT-interpolation algorithms cannot revise AOPs once the LUTs are generated and prescribed for each aerosol type, as updating them requires recomputing an extensive RTM simulation. For AMI original-YAER, we compiled AERONET inversion level 2.0 data for all sites between 10°S–50°N and 70°E–150°E from the start of observations through 2022 to generate AOP inputs for the retrieval. Following Lee et al. (2010), four aerosol types (black carbon, non-absorbing, dust, and mixed) were classified, and type-averaged AOPs were derived for LUT construction. In contrast, the inversion scheme of AMI hybrid-YAER allows type-specific AOPs to be readily updated by supplying new parameter values to the DLRTM, without the need to regenerate an entire LUT. Fig. S2 presents yearly average AOPs for Asian AERONET sites that reported at least 100 observations per year over a decade for each aerosol type. Restricting the analysis to long-term stations ensures the representativeness of the Asian AOP climatology and minimizes sampling artifacts associated with site turnover. The optical properties of black carbon aerosols in Asia have changed little since 2010. The imaginary part of the refractive index decreased between 2010 and 2013, and then remained nearly constant through 2023. Non-absorbing aerosols form mainly through secondary aerosol production, which is a temporally independent process that leads to no distinct long-term trend. For both black carbon and non-absorbing aerosols, the coarse-mode radius exhibits year-to-year variability, driven mainly by fluctuations in dust transport intensity. The long-term trend of the imaginary refractive index of the mixed aerosols closely follows that of black carbon, suggesting that the mixed category includes a substantial black carbon component. For dust aerosols, the fine-mode number fraction has increased since 2010, implying a gradual mixing of mineral dust with anthropogenic fine particles. The flexible aerosol model framework implemented in AMI hybrid-YAER can accommodate such temporal variations in AOPs.

3.5. *Background AOD and diurnal bias correction*

The AMI original-YAER AOD product exhibited a systematic low bias at several AERONET stations in heavily polluted regions (e.g., #6: Kanpur and #8: Gandhi_College in Fig. S1). The VIIRS DB AOD product is used to calculate background AOD over Asia, because the DB algorithm (applied with the Satellite Ocean Aerosol Retrieval algorithm; Sayer et al., 2012) provides a more complete global view of aerosol distribution using deep-blue wavelength observations compared to the DT algorithm to retrieve aerosols over bright surfaces such as deserts (Hsu et al., 2006, 2013). Therefore, a background AOD database was constructed using 5 years of VIIRS DB AOD data (2019–2023). After collocating the VIIRS geolocation to that of AMI, the lowest 1% of VIIRS DB AOD values with the highest quality flags were averaged for each pixel. If the number of valid samples for a given pixel was <3% of the total 1826 days, the background AOD was set to zero to ensure reliability. Over ocean, a non-zero natural background associated with sea-salt aerosol can exist, particularly under high-wind conditions. However, for the present correction scheme we set ocean background AOD to zero because the target of the correction is the persistent low bias over land, not open-ocean aerosol climatology. In addition, the VIIRS DB algorithm can produce errors over turbid water due to difficulties in estimating surface reflectance (Sayer et al., 2020). Therefore, the background AOD over any ocean pixel was set to zero, and no correction was applied to AMI AODs over the ocean. This simplification may introduce a slight negative bias in remote marine regions with enhanced natural aerosol, and it should be revisited in future versions using dedicated marine aerosol constraints. Fig. 4 shows the spatial distribution of the derived background AOD. Pixels with insufficient valid samples are shaded gray, indicating no application of the correction. Regions such as equatorial Southeast Asia and Sichuan in China exhibited limited sampling due to persistent cloud cover, while northwest China showed sparse valid data owing to unfavorable surface conditions that prevented reliable DB retrievals. These gray areas mark where the correction is intentionally withheld to preserve regional consistency rather than applying a poorly sampled adjustment. Over most of East Asia, background AOD values were below 0.05, supporting the assumption that at least one clear-sky day is available per month within a 3-year period for the MRM-based surface reflectance estimation. In contrast, South Asia, particularly Nepal and Bangladesh, showed elevated background AODs, challenging the MRM's assumption. The background AOD correction is applied after AOD inversion, as illustrated in Fig. 1.

4. Results and discussion

4.1. *Qualitative assessment of AMI hybrid-YAER*

4.1.1. *Dust call-back*

Fig. 5 shows a dust transport event on 4 March 2022. A dense dust plume was transported eastward from the Taklamakan Desert. The RGB image in Fig. 5a shows yellow dust aerosols over the northern Yellow Sea, the Shandong Peninsula, and the Bohai Sea. However, the AMI original-YAER AOD (Fig. 5b) masked the thick core of the dust plume by misclassifying it as cloud. In contrast, the dust call-back routine of AMI hybrid-YAER successfully recovered the entire plume (Fig. 5c). A few pixels over the southwestern portion of the region (32°N–35°N, 114°E–120°E) appear to be contaminated by clouds and were inadvertently reintroduced by the call-back process. This behavior confirms that the fixed thermal thresholds are effective for recovering thick dust but are not fully sufficient to reject all spectrally similar cloud scenes. In future versions, this false-positive dust pixels could be reduced by combining the current thermal tests with additional spatial-coherence screening, neighborhood texture tests, or cloud-shadow geometry constraints before inversion. NOAA's Advanced Baseline Imager (ABI) aerosol detection product (ADP) utilized multiple spectral channels spanning from VIS to IR (e.g., 0.47, 0.64, 8.9, 11, 12 μm; ABI ADP ATBD, 2018) for the accurate detection of dust pixels. Since the AMI has similar channels for dust detection, upcoming updates will utilize those channels for the removal of cloud-contaminated pixels. As accurate retrieval of dust aerosols is essential for monitoring frequent long-range transport events over East Asia, the dust call-back procedure was retained in AMI hybrid-YAER despite this remaining trade-off. Future updates on the dust call-back using IR channels will also have to consider latitudinal and temporal dependence of the threshold.

4.1.2. *Land–ocean continuity*

Fig. 6 illustrates the mitigation of the land–ocean discontinuity achieved by revising the ocean surface reflectance estimation scheme and extending the land surface compositing period. This update

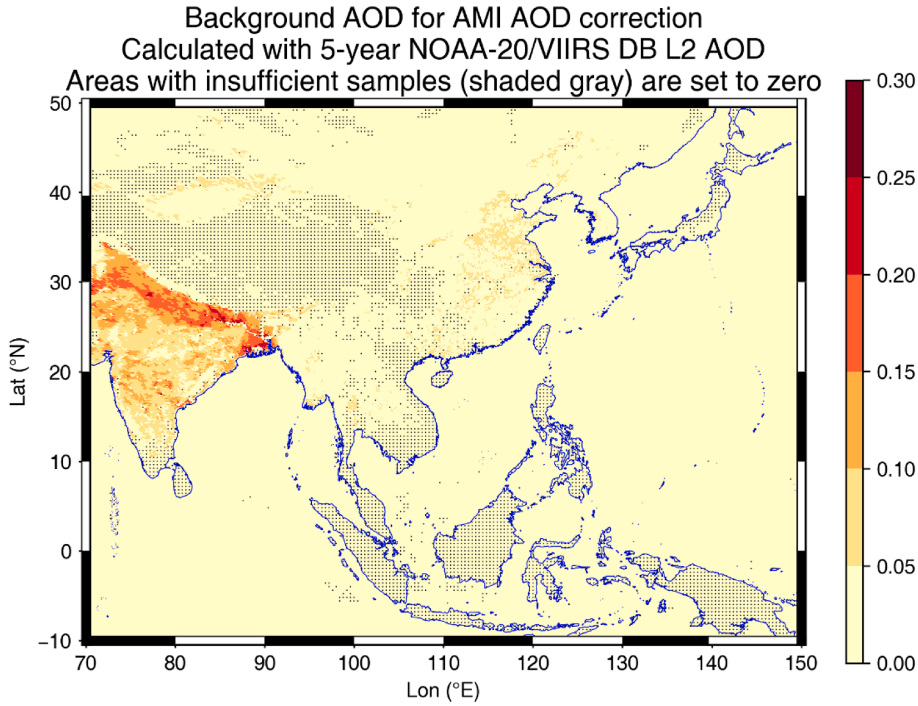

Fig. 4. Background AOD calculated with VIIRS DB AOD with high quality flags. Areas with an insufficient number of samples are shaded gray and excluded from the background AOD correction. Ocean pixels are also filled with 0 and excluded from the background AOD correction.

increased land AOD values and decreased ocean values, thus smoothing the transition between land and ocean pixels. In addition, previously invalid pixels in the northwestern portion of the image (21°N–22°N, 104°E–105°E) were successfully reintroduced in the AMI hybrid-YAER AOD retrievals. This expansion of valid retrievals resulted from the hybrid-YAER algorithm implementing a more stable surface reflectance database.

4.2. Validation and error analysis against AERONET

Fig. 7 shows the results of the validation of the AMI original- and hybrid-YAER AOD products over land and ocean against AERONET AOD; Table 4 lists the validation metrics. The two-dimensional histograms of land AOD (Fig. 7a and c) show a notable reduction in the frequency of underestimation within the AOD range of 0.0–0.5 for the hybrid-YAER. The hybrid-YAER algorithm shows a slightly improved Pearson correlation coefficient (R) of 0.767 compared with 0.757 for original-YAER. However, the RMSE and mean bias error (MBE) increased because the number of overestimated cases is not correspondingly improved, resulting in overall errors being dominated by overestimation. The RMSE increased from 0.211 to 0.217, while the RRMSE has decreased from 0.482 to 0.448. This implies that the hybrid-YAER algorithm performed better at higher aerosol loading, but worse at lower aerosol loading. Considering that the accuracy at higher aerosol loading is more affected by aerosol model assumption, the flexible aerosol model assumption which was available by using DLRTM enhanced the accuracy at high aerosol loading. Meanwhile, increased error at lower AOD seems to be induced by increased cloud contamination, because of the increased frequency where AERONET AOD is extremely low (AERONET AOD < 0.1), and AMI AOD is moderately low (AMI AOD ~ 0.3). Although the surface mitigation in the new algorithm resulted in reduced underestimation of the AMI AOD (AERONET AOD ~ 0.2 and AMI AOD ~ 0.1), the error magnitude of the cloud contamination was more dominant on the validation metric of RMSE. Compared with the AOD results over land, those over the ocean (Fig. 7b and d) show more apparent improvements by AMI hybrid-YAER. The systematic overestimation at low AODs (0.0–0.5) is mitigated, but underestimation at high AODs (>1.0) is also reduced. Consequently, the scatter plot of AMI YAER AOD vs. AERONET results becomes more linear after the version update. The validation metrics of R, RMSE, and MBE improve by $+0.025$, -0.027, and -0.045 (in absolute terms), respectively. The percentage of retrievals over ocean within the expected error (EE) requirement increase by 27.3%p. Throughout this work, EE $= \pm(0.05 + 0.15 \times \text{AOD})$ (Levy et al., 2013). The number of AMI hybrid-YAER–AERONET collocations also increases over both land and ocean compared with original-YAER owing to the dust call-back procedure and the more stable surface reflectance database.

Ablation experiments was conducted by comparing the fully updated configuration (“All included” in Table 4) with sensitivity configurations in which one algorithmic update was removed. The validation is based on one-day validation on the 15th day of each month from 2022 to 2023. The fully updated configuration showed balanced validation statistics against AERONET, with R = 0.783, RMSE = 0.206, MBE = 0.020, 48.7% within EE, and a GCOS fraction of 25.5%. These values are used here as the reference for evaluating the effect of each sensitivity configuration. By using shorter surface reflectance composite period (“No surface update” in Table 4) produced the clearest shift toward underestimation. The MBE decreased markedly from 0.020 to − 0.064, and the regression intercept decreased from 0.145 to 0.075. RMSE also increased from 0.206 to 0.215, while the number of collocations decreased from 1414 to 1293. These changes indicate that the longer surface reflectance database in the fully updated configuration improves retrieval stability and reduces negative AOD bias. The result is consistent with the interpretation that an insufficient compositing period can retain residual aerosol or cloud-contamination effects in the surface reflectance database, which can then propagate into AOD underestimation. When the dust call-back procedure was removed (“No dust call-back” in Table 4), the number of collocations decreased substantially from 1414 to 1253. The remaining validation statistics were similar to those of the fully updated configuration, with R = 0.790, RMSE = 0.208, MBE = 0.021, and 48.8% within EE. This does not imply that removing the dust call-back improves the retrieval. Rather, it indicates that the dust call-back primarily affects data availability and the retention of thick dust cases, some of which are more difficult retrieval conditions. Therefore, aggregate

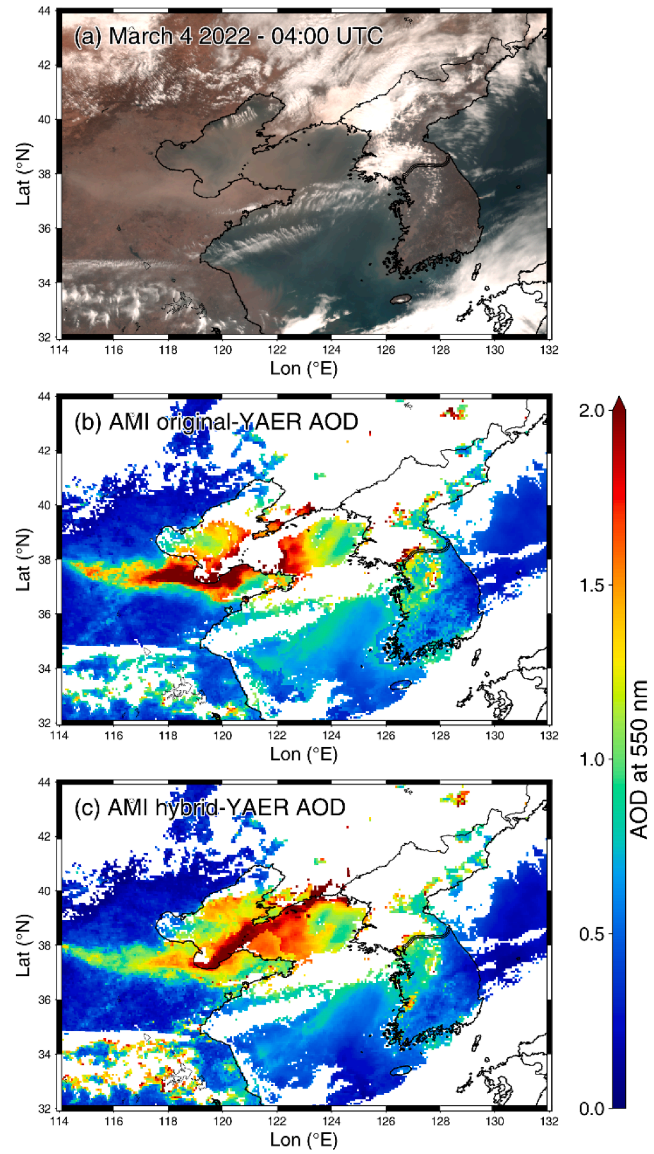

Fig. 5. Observations for March 4, 2022 demonstrating dust call-back: (a) an RGB image, (b) AMI original-YAER AOD, and (c) AMI hybrid-YAER AOD.

validation metrics based only on the remaining collocations can mask the physical benefit of recovering thick dust plumes that would otherwise be misclassified as cloud. The LUT-based configuration ("No DLRTM" in Table 4), in which the DLRTM-based online RTM calculation was replaced by the conventional LUT approach, produced similar R and RMSE values to the fully updated configuration. However, its MBE increased from 0.020 to 0.036, the fraction within EE decreased from 48.7% to 46.6%, and the GCOS fraction decreased from 25.5% to 24.6%. Therefore, the main advantage of the DLRTM-based inversion is not expressed as a large improvement in the simple correlation coefficient, but rather as better bias control and a more favorable error distribution. This is consistent with the role of the DLRTM in reducing multidimensional LUT interpolation constraints and enabling more flexible aerosol optical property inputs. When the background AOD correction was removed ("No background AOD correction" in Table 4), the validation performance degraded in most metrics. The correlation decreased from 0.783 to 0.764, RMSE increased from 0.206 to 0.213, and the fraction within EE decreased from 48.7% to 46.3%. The MBE shifted from a small positive value of 0.020 to a negative value of −0.013, while the regression slope decreased from 0.653 to 0.616. This indicates that the background AOD correction mainly acts to compensate for the persistent low bias, especially under conditions where the minimum-reflectance assumption is weakened by non-negligible background aerosol loading. However, the positive MBE in the fully updated configuration also suggests that this correction can shift the retrieval slightly upward. Overall, the ablation results suggest that the background AOD correction and extended surface reflectance compositing mainly regulate the sign and magnitude of the bias, the DLRTM-based inversion improves the retrieval error distribution and bias stability, and the dust call-back expands valid retrievals for optically thick dust events. However, these results should be interpreted with caution. Because the algorithmic updates are mutually coupled through cloud masking, surface reflectance estimation, aerosol-type selection, and AOD inversion, removing one component changes not only that component but also the retrieval population and the operating conditions of the remaining components. Therefore, the difference between each sensitivity configuration and the fully updated configuration should not be interpreted as the absolute or independent performance contribution of a single update.

The validation against AERONET further reveals that overestimation of low AOD becomes more apparent after the mitigation of underestimation of low AOD in AMI hybrid-YAER. Site-specific analyses provide further insight into this behavior. From a total of 77 validation plots for each station, we select stations showing overestimation of low AOD, as listed in Fig. S3. The selected AERONET sites include one desert site (Dalanzadgad; Fig. S3a), several stations in Taiwan (Chen-Kung_Univ, Xitun, Lulin, and EPA-NCU; Fig. S3b–e), and additional sites in Japan (Osaka, DRAGON_Matsumoto, and DRAGON_Kofu; Fig. S3f–h). The cause of the overestimation at the desert site is relatively straightforward. The low NDVI near Dalanzadgad (Fig. S3a) corresponds to high surface reflectance. As AOD retrieval over a bright surface is less sensitive to changes in aerosol loading (von Hoyningen-Huene et al., 2011) than that over a dark surface, an increased TOA reflectance from a bright surface can lead to positively biased AODs even under low aerosol conditions. In contrast, the overestimation observed at the Taiwanese and Japanese sites is likely related to frequent cloud cover in these island regions (Fig. 8a). Cloud can affect retrieval in two ways: cloud-affected pixels with high TOA reflectance may enter during pixel aggregation, and cloud shadows may decrease surface reflectance in the MRM-based surface database. Both result in AOD overestimation. Additionally, complex surface conditions in Taiwan and central–western Japan may complicate retrievals when aggregating 6 × 6 pixels prior to inversion. As shown in Fig. 8a, regions such as the Himalaya–Sichuan Basin and equatorial Southeast Asia also experience frequent cloud/snow cover; however, such validation errors were not observed in these areas. This is likely because Taiwan exhibits a sharp gradient in cloud occurrence over a very small spatial scale due to its complex topography, making accurate detection more challenging. The high spatial contrast in surface characteristics across relatively small islands may lead to underestimation of the aggregated surface reflectance (Fig. 8b and c). Future updates to the algorithm should seek to improve the handling of cloud shadow and cloud masking. Trees et al. (2024) effectively mitigated cloud shadow contamination in the TROPOMI absorbing aerosol retrieval by geometrically estimating shadow locations using cloud height information.

Fig. 9 presents error analyses of the AMI original- and hybrid-YAER AODs with respect to AERONET AOD, observation time, solar zenith angle, and NDVI. Over land, as AMI hybrid-YAER mitigates underestimation at low AOD, the overall bias shifts toward positive values. Both versions tend to underestimate AOD at higher aerosol loadings, but the bias of hybrid-YAER is of lower magnitude. Over the ocean, original-YAER systematically overestimates at AOD < 0.5 but underestimates at higher AODs, resulting in a nonlinear bias pattern that hinders uncertainty characterization. In comparison, the bias of hybrid-YAER is of smaller magnitude (slightly negative), and the shape of the bias becomes more linear with respect to aerosol loading. Fig. 9c and d shows two peaks at 0100 and 0500 UTC in the hourly bias patterns for both versions; however, hybrid-YAER over ocean maintains a relatively stable

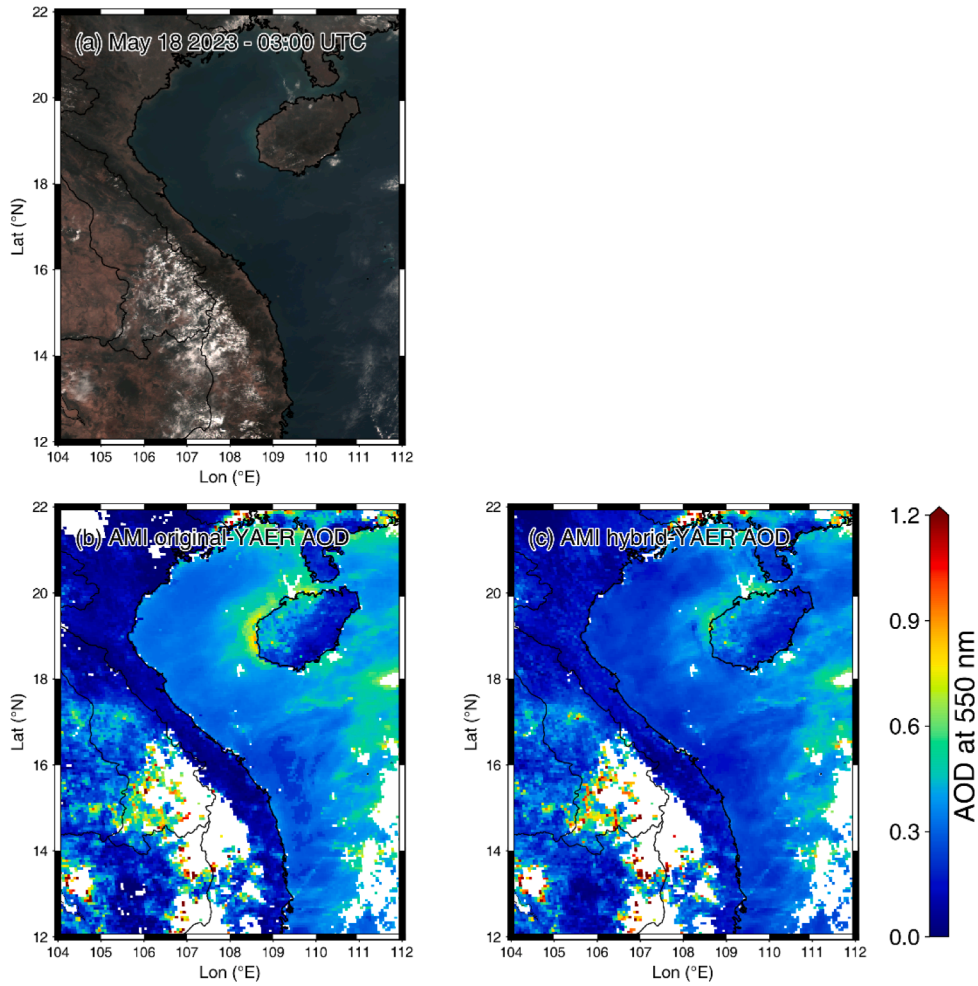

Fig. 6. Representative example of land–ocean AOD discontinuity mitigation (18 May 2023): (a) RGB image, (b) AMI original-YAER AOD, and (c) AMI hybrid-YAER AOD.

bias during daytime. Over land, although the AMI hybrid-YAER AOD still shows fluctuations, its interquartile range (25th–75th percentile) and overall amplitude of bias are lower than those of original-YAER, indicating improved stability. The number of AERONET collocations drops around 06 UTC, because AMI underwent daily wheel offloading at that time from 22 August to 20 December. Fig. 9e–f suggests that the version update did not substantially mitigate bias patterns associated with solar geometry, as the update did not include geometric corrections. Biases with respect to NDVI over land decrease with increasing vegetation fraction for both versions (Fig. 9g), likely because vegetation generally appears relatively dark in the VIS spectrum, thus providing favorable conditions for AOD retrieval. The interquartile range of bias values narrows after the update to hybrid-YAER, particularly at NDVI ~ 0.2. The extended compositing period used for MRM-based surface reflectance estimation contributes to hybrid-YAER showing more stable AOD retrievals over relatively bright arid regions in comparison to original-YAER.

To investigate the possible error propagation induced from the DLRTM for T-matrix, we performed error analysis for pixels that was selected to be dust as the primary aerosol type (Fig. 10). When collocated with AERONET, more than half of the valid AMI pixels within 25 km from an AERONET station had to be dust for the analysis. Meanwhile, the AOD $<$ 0.3 are excluded to focus on the effect of aerosol model calculation, since the accuracy of low AOD are more affected by surface conditions. Compared to the original-YAER AOD, the hybrid-YAER AOD exhibits more random and widely dispersed errors. This indicates that uncertainties in the T-matrix calculations propagate into the AOD retrieval, leading to a broader error distribution. As shown in Fig. 3, the errors in the T-matrix are randomly distributed, resulting in errors that spread in a non-directional (random) manner rather than introducing a systematic bias. In contrast, the original YAER AOD shows an overall negative bias, which becomes more pronounced as AOD increases. On the other hand, the errors in the hybrid-YAER AOD are more symmetrically distributed around zero, indicating a reduction in systematic bias. Although the relatively large errors from the DLRTM-based T-matrix calculations are propagated into the AOD, the improvements enabled by DLRTM, such as online RTM calculation and more flexible aerosol model inputs, lead to better performance. As a result, for dust AOD greater than 0.3, the RMSE decreases from 0.263 to 0.213. Therefore, despite the uncertainties in dust T-matrix calculations, the introduction of DLRTM ultimately improves the accuracy of dust AOD retrievals.

4.3. Comparison with VIIRS DT/DB AOD

To evaluate the spatial robustness of the algorithmic improvements in AMI hybrid-YAER, both original- and hybrid-YAER are compared with the VIIRS AOD DT and DB products from low Earth orbit. Table 5 lists the evaluation metrics, and Fig. 11 shows the results. With respect to the DB product, hybrid-YAER shows less AOD underestimation than original-YAER at low AOD values (Fig. 11a and c). Scatter plots with frequencies of 26 to 193 show that hybrid-YAER agrees with VIIRS AOD moderately better than does original-YAER at higher AODs, likely due to the updated aerosol model. As a result, the slope of the linear regression increases from 0.802 (for original-YAER) to 0.926 (for hybrid-YAER),

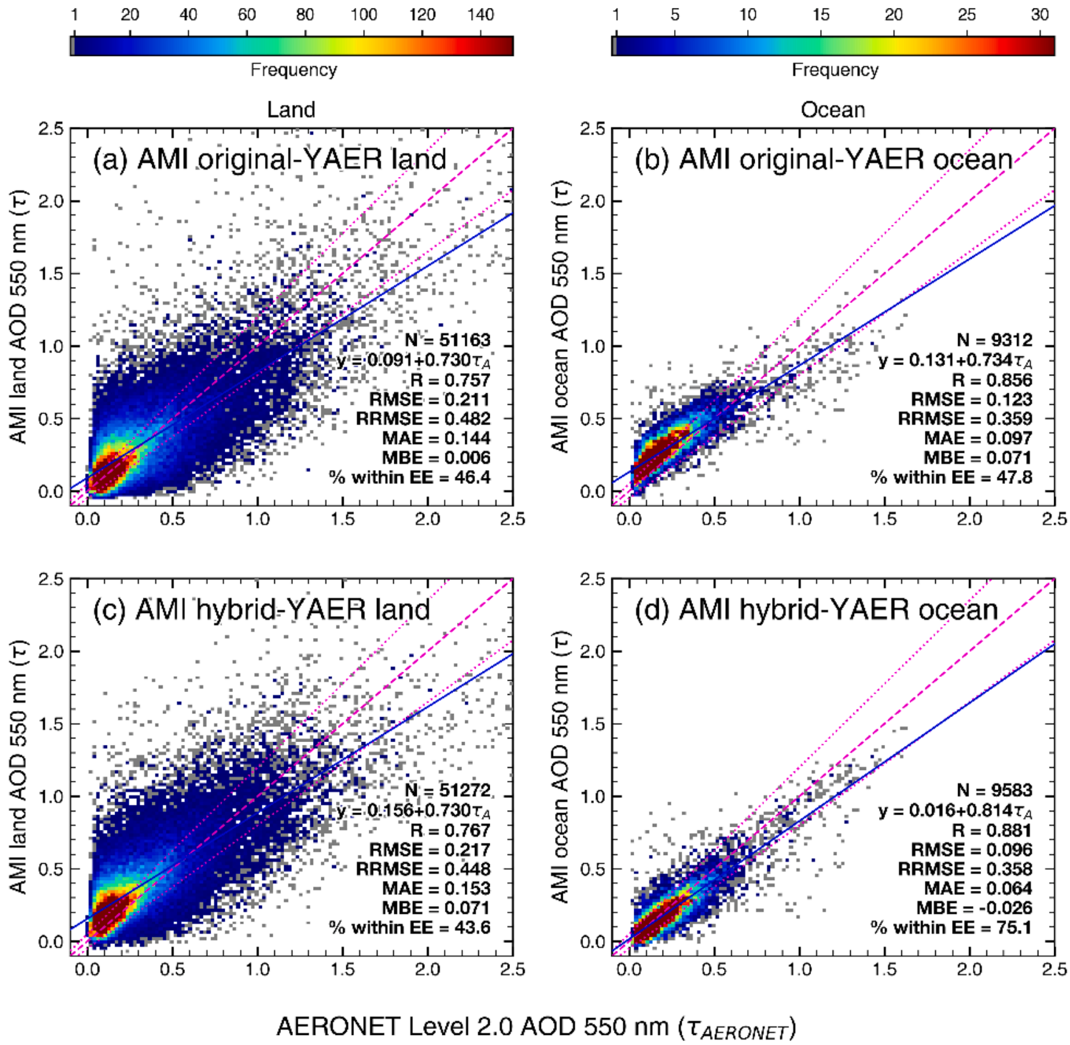

Fig. 7. Two-dimensional histograms of AMI YAER AOD validation over land (left) and ocean (right). The upper row shows results for original-YAER, and the lower row shows results for hybrid-YAER. Dotted magenta lines indicate the expected error envelop of $\pm(0.05 + 0.15 \times \text{AOD})$, and dashed magenta lines indicate the 1:1 trend. Blue lines indicate the linear regression line between AERONET AOD and AMI YAER AOD. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
AERONET validation statistics for the ablation experiments, based on one-day validation on the 15th day of each month from 2022 to 2023. “All included” denotes the fully updated algorithm, and the other cases remove one update at a time. The results represent coupled sensitivity responses, not independent contributions from each update.

Case	N	Regression Equation	R	RMSE	MBE	% within EE	GCOS
No surface update	1293	(y = 0.075 + 0.630x)	0.787	0.215	−0.064	48.6	24.9
No dust call-back	1253	(y = 0.147 + 0.664x)	0.790	0.208	0.021	48.8	25.0
No DLRTM (use LUT)	1417	(y = 0.156 + 0.666x)	0.788	0.206	0.036	46.6	24.6
No background AOD correction	1414	(y = 0.125 + 0.616x)	0.764	0.213	−0.013	46.3	25.2
All included	1414	(y = 0.145 + 0.653x)	0.783	0.206	0.020	48.7	25.5

and the fraction of points within the EE rises by 1.6 percentage points. Other validation metrics either slightly deteriorate or show negligible improvement, which can be attributed to the overestimation of AMI hybrid-YAER AOD in the 0.5–2.0 range of VIIRS DB AOD. This overestimation likely results from the dust call-back procedure, which enables retrievals over bright desert surfaces where the spectral sensitivity of AMI is limited.

In contrast, such overestimation is not observed in the comparison with the VIIRS DT product (Fig. 11b and d), as the DB algorithm already exploits deep-blue spectral channels to retrieve aerosols over bright surfaces. Comparison of the results from original- and hybrid-YAER with the VIIRS DT product reveals a clearer overall improvement of hybrid-YAER over original-YAER than does their comparison with the VIIRS DB product. Going from original-YAER to hybrid-YAER, the indicators show the following changes: the regression slope increases from 0.726 to 0.786, R increases from 0.857 to 0.870, RMSE decreases from 0.157 to 0.148, MBE improves by decreasing in magnitude from 0.009 to −0.001, and the percentage of points within the EE increases from 53.7% to 58.4%. Similar to the comparisons against VIIRS DB AOD, the scatter plots with respect to VIIRS DT AOD with frequencies between 26 and 193 show a more consistent distribution for hybrid-YAER than do those for original-YAER.

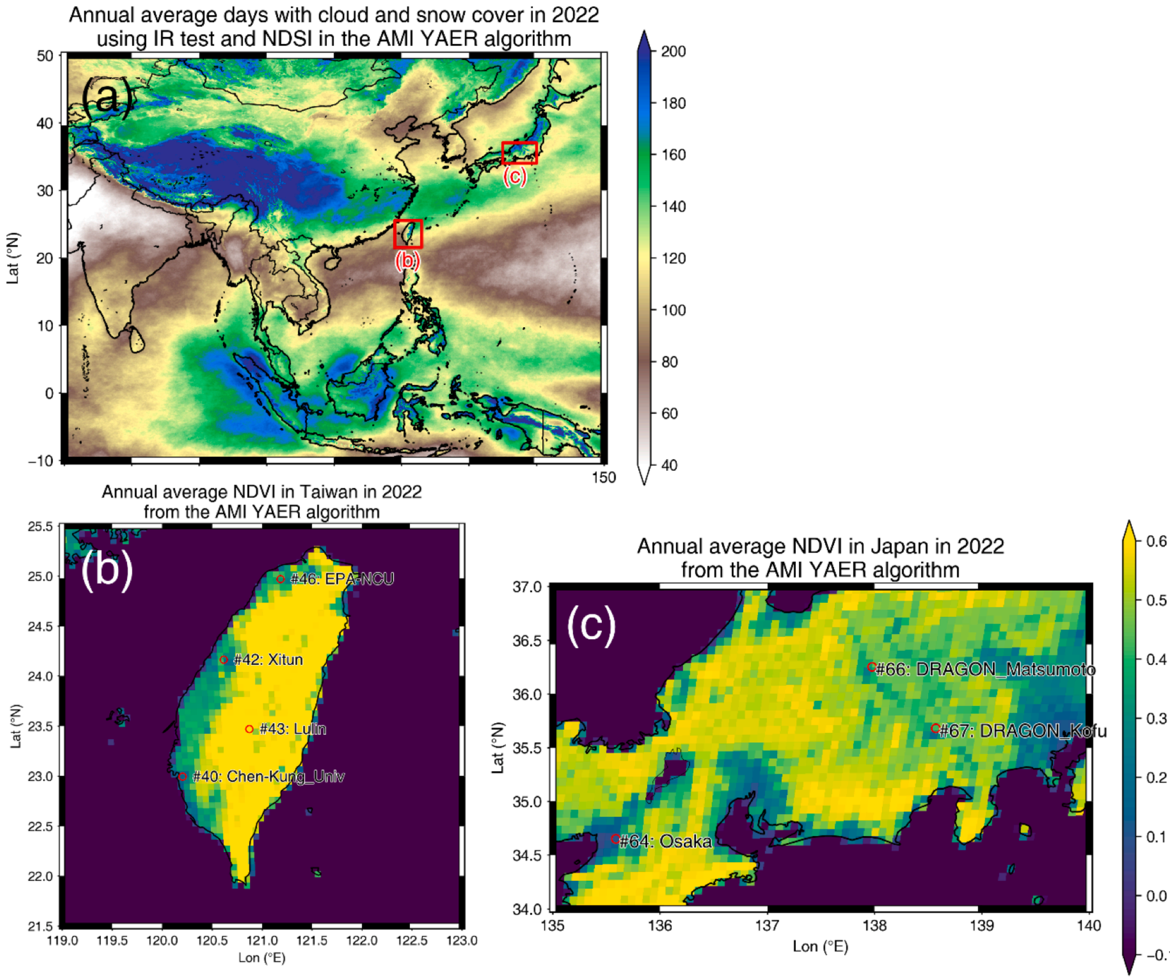

Fig. 8. (a) Average cloud and snow covered days in 2022 in Asia. Cloud and snow cover are detected with IR tests and NDSI in the AMI YAER algorithm. Red boxes indicate the areas shown in (b) and (c). (b, c) Average NDVI in Taiwan and Japan in 2022. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.4. Uncertainty analysis of the AOD from AMI hybrid-YAER

Based on the prognostic error estimation framework introduced by Sayer et al. (2020), we quantify the uncertainty of each AMI AOD product as follows. First, retrieval errors of AMI AOD (τ_{AMI}) are defined as ε and are sorted with respect to τ_{AMI}:

$$\varepsilon_i = (\tau_{\mathrm{AMI}} - \tau_{\mathrm{AERONET}})_s, \text{where}, s = \mathrm{argsort}(\tau_{AMI}) \tag{1}$$

The total number of ε_i is N, and they are all binned into K bins, with each bin numbered M:

$$\mathscr{B}_k = \{j | j \in [k-1]M + 1, \min(kM, N)]\}, k = 1, 2, \cdots, K \tag{2}$$

where B_k is the index range corresponding to the k-th bin. Then, we define E_k as including the retrieval errors within the index range of B_k:

$$E_k = \{\varepsilon_j | j \in \mathscr{B}_k\} \tag{3}$$

Parameters σ and μ of the error distribution in E_k are calculated as follows:

$$\sigma = \mathrm{Quantile}(\mathrm{abs}(E_k); 0.68) - \mathrm{Quantile}(\mathrm{abs}(E_k); 0.50) \tag{4}$$

$$\mu = \mathrm{Quantile}(E_k; 0.50) \tag{5}$$

Assuming a Gaussian error distribution for AMI AOD, the quantities σ and μ correspond respectively to the standard deviation and average of the Gaussian distribution.

Fig. 12 shows the calculated σ and μ for the AMI AOD values, where σ represents retrieval uncertainty and μ denotes bias. The differences of the biases between Fig. 9a and b and Fig. 12 arise from the different references (x-axes): AERONET AOD in Fig. 9 and AMI AOD in Fig. 12. Over land (Fig. 12a), σ for both AMI original-YAER and hybrid-YAER increases with aerosol loading. Both versions show similar linear regression equations for the uncertainty listed in **Table 6 (i.e., expected error or prognostic error): $0.02 + 0.11 \times \mathrm{AOD}$ for original-YAER, and $0.02 + 0.12 \times \mathrm{AOD}$ for hybrid-YAER. Their respective coefficients of determination (R^2) are 0.95 and 0.96. However, as discussed earlier, the μ value for AMI hybrid-YAER shifts toward positive values due to the mitigation of underestimation at low AOD. Over ocean, the regression slope of the σ value of hybrid-YAER decreases slightly by 0.02 compared with that of original-YAER, while μ is of smaller magnitude after the update. The coefficient of determination also improves from 0.88 to 0.94. This improvement indicates that the expected error more accurately represents the actual retrieval error, suggesting enhanced stability of the AMI hybrid-YAER AOD over oceanic regions.

4.5. Limitation and future study

Several limitations of the present hybrid-YAER framework should be noted. First, the DLRTM error is larger for non-spherical T-matrix cases than for Mie cases, so dust-dominated retrievals remain the least certain component of the product. Second, the land validation indicates that the current combination of cloud screening, pixel aggregation, and MRM-

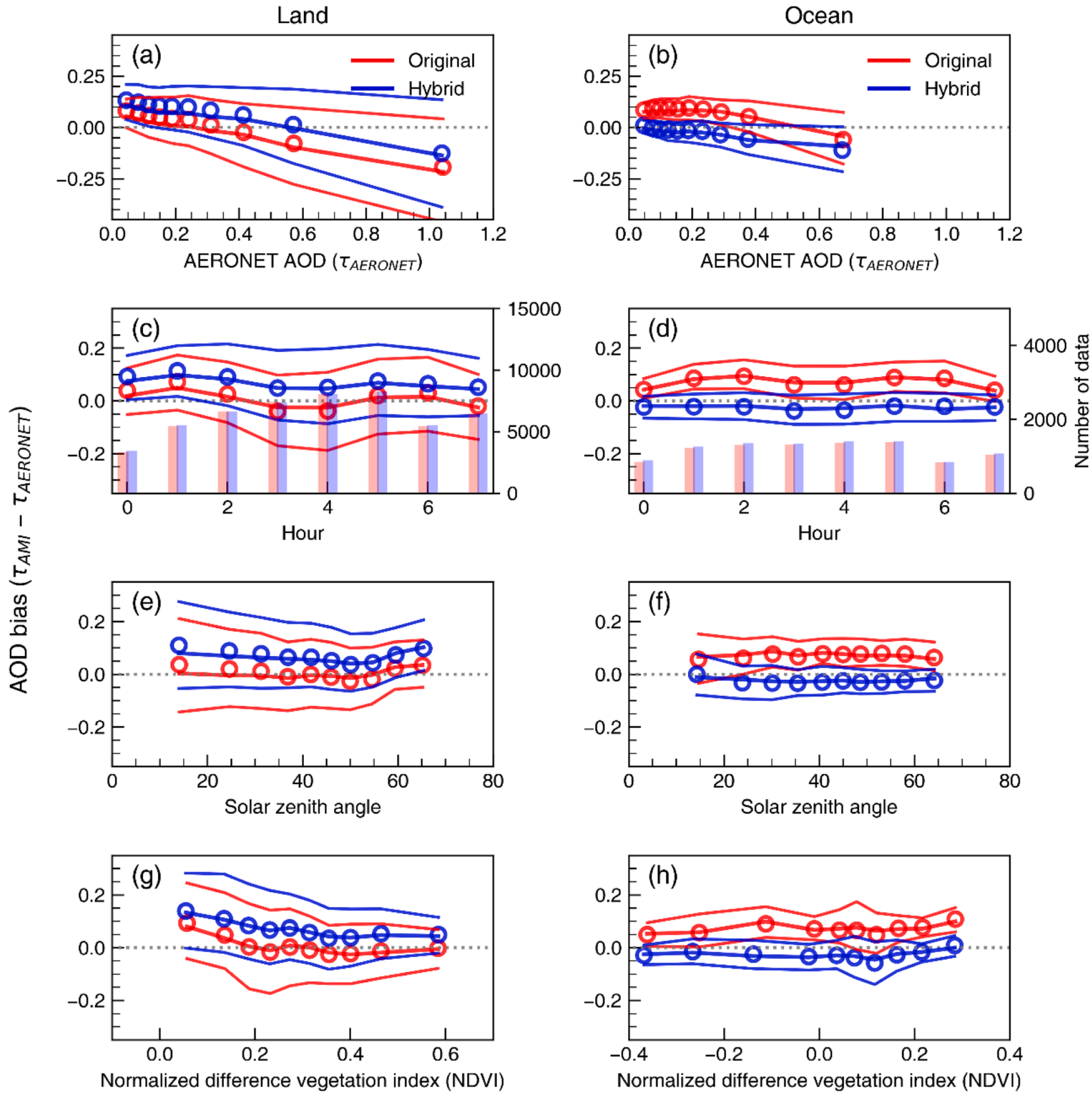

Fig. 9. Error analysis of AMI YAER AOD products with respect to (a, b) AERONET AOD, (c, d) observation time, (e, f) solar zenith angle, and (g, h) NDVI. Red indicates the 5,116 land and 931 ocean data points from original-YAER, and blue indicates the 5,127 land and 958 ocean data from hybrid-YAER. The bar charts in (c and d) give the number of data at each observation time. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

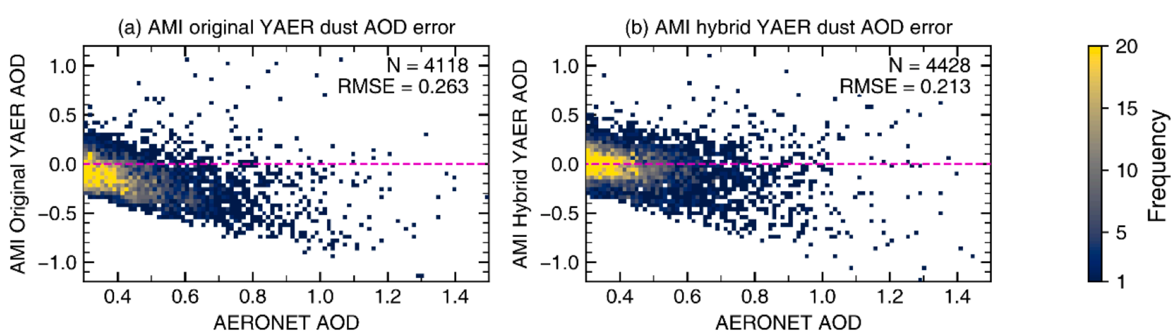

Fig. 10. Error analysis as a function of AERONET AOD for pixels where the aerosol type is identified as dust and AOD > 0.3. (a) AMI original YAER AOD error, and (b) AMI hybrid YAER AOD error.

based surface reflectance can still induce positive biases over bright or heterogeneous surfaces, particularly in Taiwan, Japan, and desert regions. Third, because multiple updates were introduced simultaneously, the present manuscript does not provide a formal ablation analysis of the individual contributions from the DLRTM, revised aerosol models, extended surface database, dust call-back, and background correction. Future work will therefore focus on dust-specific error propagation, ablation experiments, improved cloud-shadow and heterogeneous-

Table 5
Metrics comparing each of AMI original-YAER and hybrid-YAER with each of VIIRS DB and DT.

	AMI original-YAER	AMI hybrid-YAER
vs. VIIRS DB	N: 2,015,024 Linear regression: y = 0.060 + 0.802x R = 0.849 RMSE = 0.149 MBE = 0.011 % within EE = 58.9	N: 2,107,548 Linear regression: y = 0.018 + 0.926x R = 0.845 RMSE = 0.164 MBE = 0.000 % within EE = 60.5
vs. VIIRS DT	N: 1,830,134 Linear regression: y = 0.076 + 0.726x R = 0.857 RMSE = 0.157 MBE = 0.009 % within EE = 53.7	N: 1,901,758 Linear regression: y = 0.050 + 0.786x R = 0.870 RMSE = 0.148 MBE = 0.001 % within EE = 58.4

surface handling.

5. Conclusion

The AMI YAER algorithm was updated with improvements to dust call-back, surface reflectance estimation, aerosol model update, inversion method, background AOD correction, and hourly bias correction. The updated adopts a hybrid inversion method that DLRTMs replace LUT interpolation with online RTM calculation. Maintaining a physics-based inversion has several advantages. For example, the retrieved AOD can be used to reconstruct the observed L1B through physical calculations. This enables greater physical consistency when the AOD product is used as an input for retrieving other variables from the same satellite observations. In addition, for future algorithm updates, the physical framework allows for targeted improvements by addressing specific sources of error. In contrast, data-driven approaches are typically limited to more general strategies, such as extending the training dataset or adopting new machine learning architectures. By preserving the physical basis, more systematic and refined version control based on identified error sources becomes possible. Eight DLRTMs for four VIS–NIR channels and two particle geometries (Mie and T-matrix) were trained to emulate VLIDORT. Validation of the DLRTMs showed T-matrix emulation to be more erroneous than Mie calculation. The background AOD for the correction was calculated using 5 years of VIIRS DB AOD data. Across the study area, India was the only region with distinct changes after the background AOD correction. Case studies demonstrated the efficacy of the dust call-back process, and the surface estimation update mitigated land–ocean AOD discontinuity. Validation of AMI YAER AOD against ground-based AERONET AOD revealed that the algorithm update greatly improved AOD over ocean. The underestimation by original-YAER of low AOD over land was improved in hybrid-YAER. However, some overestimation remained unaffected, resulting

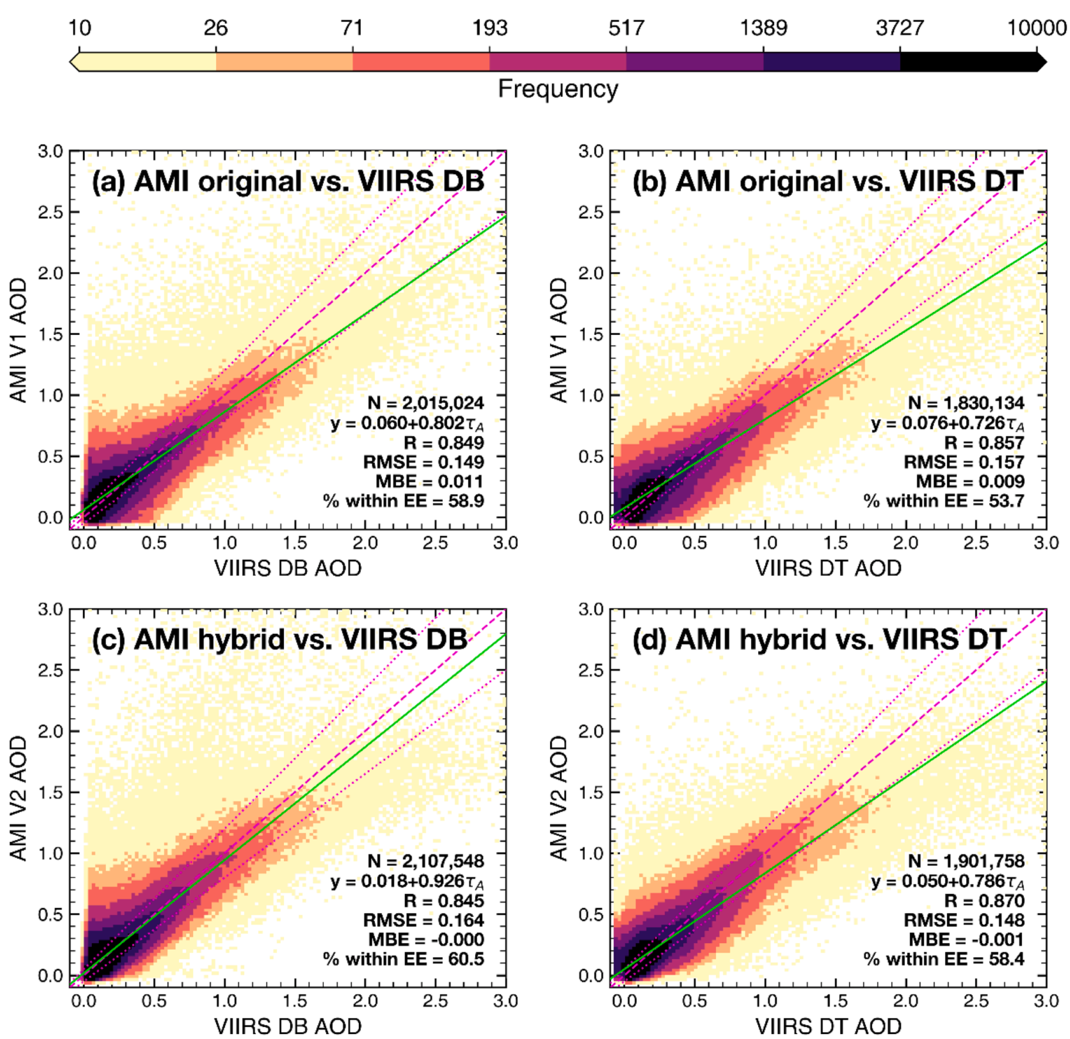

Fig. 11. Two-dimensional histograms comparing AODs from AMI YAER (a and b) original-YAER and (c and d) hybrid-YAER with AODs from VIIRS (a and c) DB and (b and d) DT.

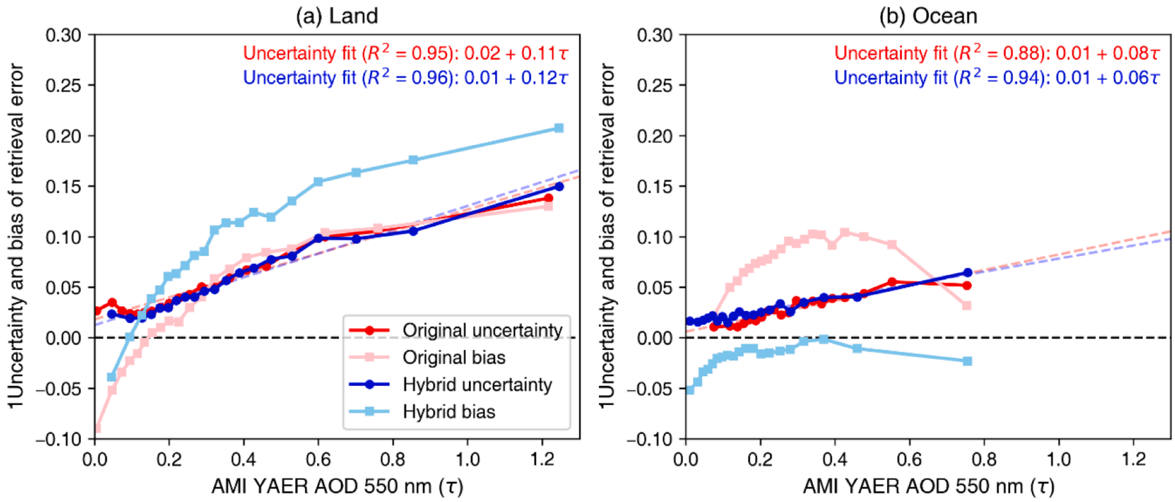

Fig. 12. Uncertainty analysis of the AMI original-YAER and hybrid-YAER AOD products. Red and blue plots are the σ of the retrieval error distribution as retrieval uncertainty for original- and hybrid-YAER, respectively. Dashed lines are linear regression lines of σ. Pink and sky-blue plots are the μ of the retrieval error distribution as bias for original- and hybrid-YAER, respectively. Retrieval error distributions are assumed to be Gaussian, so that the median is equivalent to μ, and the difference between the 68th percentile and the median is equivalent to σ.

Table 6
Linear uncertainty (σ) fit of AMI original-YAER and hybrid-YAER algorithms.

	Land	Ocean
AMI original-YAER	0.02 + 0.11 AOD	0.01 + 0.08 AOD
AMI hybrid-YAER	0.01 + 0.12 AOD	0.01 + 0.06 AOD

in a positive bias at low aerosol loading in the AMI hybrid-YAER AOD data. The overestimation at low aerosol loading was distinct at AERONET stations near desert regions and stations in Taiwan and Japan. It was associated with cloud contaminated pixels and/or low surface reflectance due to cloud shadow. Error analysis revealed that the bias patterns with respect to AERONET AOD, observation time, and solar zenith angles were not changed over land, while the magnitudes of the bias became positive. Over ocean, the bias patterns were reduced in magnitude and became more robust. For both land and ocean, the version update stabilized the bias pattern with respect to observation time. Comparisons with VIIRS DT/DB AOD showed that AMI hybrid-YAER AOD agrees better with LEO AOD products than does AMI original-YAER AOD. However, comparison with VIIRS DB AOD revealed that the dust call-back procedure of AMI hybrid-YAER retained data for bright desert surfaces with moderate AOD, resulting in AOD overestimation for VIIRS DB AOD of 0.5–2.0. Several limitations remain in the hybrid-YAER framework. DLRTM errors are larger for non-spherical cases, and positive biases persist over bright or heterogeneous land surfaces. In addition, the lack of a formal ablation analysis limits the attribution of individual component improvements, which will be addressed in future work.

Data availability

The training code for the DLRTM using the Set Transformer is available via GitHub (https://github.com/KIMminseok-github/JAG-D-26-01071R1). The AMI original-YAER and hybrid-YAER AOD products are available upon request to the corresponding author (Jhoon Kim; email: jkim2@yonsei.ac.kr).

Declaration of competing interest

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

Acknowledgements

This work was supported by the InnoCORE program [grant number N10250155(00)] of the Korean Ministry of Science and ICT (MSIT), the National Research Foundation of Korea funded by MSIT [grant number RS-2024-00346149], and the Korea Institute of Science and Technology [grant number Global-24-004]. Jhoon Kim's work was supported by the Yonsei Fellow Program, funded by Lee Youn Jae. The authors gratefully acknowledge all principal investigators and their staff for establishing and maintaining the AERONET sites used in this study. The authors also acknowledge the NASA VIIRS aerosol teams for providing the Dark Target and Deep Blue products. The National Meteorological Satellite Center of the Korea Meteorological Administration is acknowledged for providing the GK-2A/AMI satellite data used in this research. We acknowledge helpful discussions with Dr. Edward J. Hyer.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2026.105393.

References

Alade, I.O., Zhang, Y., Xu, X., 2021. Modeling and prediction of lattice parameters of binary spinel compounds (AM2X4) using support vector regression with Bayesian optimization. New J. of Chem. 45 (34), 15255–15266. https://doi.org/10.1039/D1NJ01523K.

Chen, C., Dubovik, O., Fuertes, D., Litvinov, P., Lapyonok, T., Lopatin, A., Ducos, F., Derimian, Y., Herman, M., Tanré, D., Remer, L.A., Lyapustin, A., Sayer, A.M., Levy, R.C., Hsu, N.C., Descloitres, J., Li, L., Torres, B., Karol, Y., Herrera, M., Herreras, M., Aspetsberger, M., Wanzenboeck, M., Bindreiter, L., Marth, D., Hangler, A., Federspiel, C., 2020. Validation of GRASP algorithm product from POLDER/PARASOL data and assessment of multi-angular polarimetry potential for aerosol monitoring. Earth Syst. Sci. Data 12, 3573–3620. https://doi.org/10.5194/essd-12-3573-2020.

Chen, C., Dubovik, O., Litvinov, P., Fuertes, D., Lopatin, A., Lapyonok, T., Matar, C., Karol, Y., Fischer, J., Preusker, R., Hangler, A., Aspetsberger, M., Bindreiter, L., Marth, D., Chimot, J., Fougnie, B., Marbach, T., Bojkov, B., 2022a. Properties of aerosol and surface derived from OLCI/Sentinel-3A using GRASP approach: Retrieval development and preliminary validation. Remote Sens. Environ. 280, 113142. https://doi.org/10.1016/j.rse.2022.113142.

Chen, X., Zhao, L., Zheng, F., Li, J., Li, L., Ding, H., Zhang, K., Liu, S., Li, D., de Leeuw, G., 2022b. Neural Network AEROsol Retrieval for Geostationary Satellite (NNAeroG) based on temporal, spatial and spectral measurements. Remote Sens. 14, 980. https://doi.org/10.3390/rs14040980.

Cho, Y., Kim, J., Go, S., Kim, M., Lee, S., Kim, M., Chong, H., Lee, W.J., Lee, D.W., Torres, O., Park, S.S., 2024. First atmospheric aerosol-monitoring results from the geostationary environment monitoring spectrometer (GEMS) over Asia. Atmos. Meas. Tech. 17, 4369–4390. https://doi.org/10.5194/amt-17-4369-2024.

Hansell, R.A., Ou, S.C., Liou, K.N., Roskovensky, J.K., Tsay, S.C., Hsu, C., Ji, Q., 2007. Simultaneous detection/separation of mineral dust and cirrus clouds using MODIS thermal infrared window data. Geophys. Res. Lett. 34, L11808. https://doi.org/10.1029/2007GL029388.

Choi, M., Kim, J., Lee, J., Kim, M., Park, Y.J., Holben, B., Eck, T.F., Li, Z., Song, C.H., 2018. GOCI Yonsei aerosol retrieval version 2 products: an improved algorithm and error analysis with uncertainty estimation from 5-year validation over East Asia. Atmos. Meas. Tech. 11, 385–408. https://doi.org/10.5194/amt-11-385-2018.

Cox, C., Munk, W., 1954. Measurement of the roughness of the sea surface from photographs of the Sun's glitter. J. Opt. Soc. Am. 44, 838–850. https://doi.org/10.1364/JOSA.44.000838.

Dubovik, O., Fuertes, D., Litvinov, P., Lopatin, A., Lapyonok, T., Doubovik, I., Xu, F., Ducos, F., Chen, C., Torres, B., Derimian, Y., Li, L., Herreras-Giralda, M., Herrera, M., Karol, Y., Matar, C., Schuster, G.L., Espinosa, R., Puthukkudy, A., Li, Z., Fischer, J., Preusker, R., Cuesta, J., Kreuter, A., Cede, A., Aspetsberger, M., Marth, D., Bindreiter, L., Hangler, A., Lanzinger, V., Holter, C., Federspiel, C., 2021. A comprehensive description of multi-term LSM for applying multiple a priori constraints in problems of atmospheric remote sensing: GRASP algorithm, concept, and applications. Front. Remote Sens. 2. https://doi.org/10.3389/frsen.2021.706851.

Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B.N., Mishchenko, M., Yang, P., Eck, T.F., Volten, H., Muñoz, O., Veihelmann, B., van der Zande, W.J., Leon, J.-F., Sorokin, M., Slutsker, I., 2006. Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust. J. Geophys. Res. 111, D11208. https://doi.org/10.1029/2005JD006619.

Eck, T.F., Holben, B.N., Reid, J.S., Dubovik, O., Smirnov, A., O'Neill, N.T., Slutsker, I., Kinne, S., 1999. Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols. J. Geophys. Res. 104, 31333–31349. https://doi.org/10.1029/1999JD900923.

Gao, M., Franz, B.A., Knobelspiesse, K., Zhai, P.W., Martins, V., Burton, S., Cairns, B., Ferrare, R., Gales, J., Hasekamp, O., Hu, Y., Ibrahim, A., McBride, B., Puthukkudy, A., Werdell, P.J., Xu, X., 2021. Efficient multi-angle polarimetric inversion of aerosols and ocean color powered by a deep neural network forward model. Atmos. Meas. Tech. 14, 4083–4110. https://doi.org/10.5194/amt-14-4083-2021.

Higurashi, A., Nakajima, T., 1999. Development of a two-channel aerosol retrieval algorithm on a global scale using NOAA AVHRR. J. Atmos. Sci. 56, 924–941. https://doi.org/10.1175/1520-0469(1999)056<0924:DOATCA>2.0.CO;2.

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., Lavenu, F., Jankowiak, I., Smirnov, A., 1998. 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.

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

Hsu, N.C., Lee, J., Sayer, A.M., Kim, W., Bettenhausen, C., Tsay, S.-C., 2019. VIIRS Deep Blue aerosol products over land: extending the EOS long-term aerosol data records. J. Geophys. Res. Atmos. 124, 4026–4053. https://doi.org/10.1029/2018JD029688.

Hsu, N.C., Tsay, S.C., King, M.D., Herman, J.R., 2006. Deep Blue retrievals of Asian aerosol properties during ACE-Asia. IEEE Trans. Geosci. Remote Sens. 44, 3180–3195. https://doi.org/10.1109/TGRS.2006.879540.

Huh, S., Jin, K.W., 2022. GEO-KOMPSAT-2A/2B AMI/GOCI-II/GEMS data & products. Geo Data 4, 39–49. https://doi.org/10.22761/DJ2022.4.4.005.

Jiang, J., Tao, M., Xu, X., Jiang, Z., Man, W., Wang, J., Wang, L., Wang, Y., Zheng, Y., Tao, J., Chen, L., 2023. A generalized aerosol algorithm for multi-spectral satellite measurement with physics-informed deep learning method. Geophys. Res. Lett. 50, e2023GL106806. https://doi.org/10.1029/2023GL106806.

Jin, B., Xu, X., 2024. Machine learning coffee price predictions. J. Uncert. Systems 17 (04), 2450023. https://doi.org/10.1142/s1752890924500235.

Jin, B., Xu, X., 2025. High-frequency CSI300 spot and futures price predictions via the neural network. J. Uncert. Systems 18 (04), 2550008. https://doi.org/10.1142/s1752890925500084.

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 J. Photogramm. Remote Sens. 183, 253–268. https://doi.org/10.1016/j.isprsjprs.2021.11.016.

Kaufman, Y.J., Fraser, R.S., Ferrare, R.A., 1990. Satellite measurements of large-scale air pollution: methods. J. Geophys. Res. 95, 9895–9909. https://doi.org/10.1029/JD095iD07p09895.

Kim, D., Gu, M., Oh, T.-H., Kim, E.-K., Yang, H.-J., 2021. Introduction of the Advanced Meteorological Imager of Geo-Kompsat-2a: in-orbit tests and performance validation. Remote Sens. 13, 1303. https://doi.org/10.3390/rs13071303.

Kim, J., Yoon, J.M., Ahn, M.H., Sohn, B.J., Lim, H.S., 2008. Retrieving aerosol optical depth using visible and mid-IR channels from geostationary satellite MTSAT-1R. Int. J. Remote Sens. 29, 6181–6192. https://doi.org/10.1080/01431160802175553.

Kim, M., Kim, J., Lee, S., Lim, H., Cho, Y., Lee, H.C., Kuk, S.K., 2025. Aerosol fine-mode-fraction retrieval from GEO-KOMPSAT-2A/AMI using a deep neural network and spectral deconvolution algorithm. IEEE Trans. Geosci. Remote Sens. 63, 1–12. https://doi.org/10.1109/TGRS.2025.3591177.

Kim, M., Kim, J., Lim, H., Lee, S., Cho, Y., Chan, P.W., 2024a. Implementation of the Yonsei Aerosol retrieval algorithm in the GK-2A/AMI and FY-4A/AGRI remote-sensing systems. AIP Conf. Proc. 2988, 050002. https://doi.org/10.1063/5.0183243.

Kim, M., Kim, J., Lim, H., Lee, S., Cho, Y., Lee, Y.G., Go, S., Lee, K., 2024b. Aerosol optical depth data fusion with Geostationary Korea Multi-Purpose Satellite (GEO-KOMPSAT-2) instruments GEMS, AMI, and GOCI-II: statistical and deep neural network methods. Atmos. Meas. Tech. 17, 4317–4335. https://doi.org/10.5194/amt-17-4317-2024.

Kim, M., Kim, J., Wong, M.S., Yoon, J., Lee, J., Wu, D., Chan, P.W., Nichol, J.E., Chung, C.-Y., Ou, M.-L., 2014. Improvement of aerosol optical depth retrieval over Hong Kong from a geostationary meteorological satellite using critical reflectance with background optical depth correction. Remote Sens. Environ. 142, 176–187. https://doi.org/10.1016/j.rse.2013.12.003.

King, M.D., Kaufman, Y.J., Tanré, D., Nakajima, T., 1999. Remote sensing of tropospheric aerosols from space: past, present, and future. Bull. Am. Meteorol. Soc. 80, 2229–2260. https://doi.org/10.1175/1520-0477(1999)080<2229:RSOTAF>2.0.CO;2.

Lee, J., Hsu, N.C., Kim, W.V., Sayer, A.M., Tsay, S.-C., 2024. VIIRS Version 2 deep blue aerosol products. J. Geophys. Res. 129, e2023JD040082. https://doi.org/10.1029/2023JD040082.

Lee, J., Kim, J., Song, C.H., Ryu, J.-H., Ahn, Y.-H., Song, C.K., 2010. Algorithm for retrieval of aerosol optical properties over the ocean from the Geostationary Ocean Color Imager. Remote Sens. Environ. 114, 1077–1088. https://doi.org/10.1016/j.rse.2009.12.021.

Lee, J., Lee, Y., Kim, J., Kosiorek, A., Choi, S., Teh, Y.W., 2019. Set transformer: a framework for attention-based permutation-invariant neural networks. Proc. Mach. Learn. Res. 97, 3744–3753. https://doi.org/10.48550/arXiv.1810.00825.

Lee, S., Choi, M., Kim, J., Park, Y.-J., Choi, J.-K., Lim, H., Lee, J., Kim, M., Cho, Y., 2023. Retrieval of aerosol optical properties from GOCI-II observations: Continuation of long-term geostationary aerosol monitoring over East Asia. Sci. Total Environ. 903, 166504. https://doi.org/10.1016/j.scitotenv.2023.166504.

Leitão, J., Richter, A., Vrekoussis, M., Kokhanovsky, A., Zhang, Q.J., Beekmann, M., Burrows, J.P., 2010. On the improvement of NO_2 satellite retrievals – aerosol impact on the airmass factors. Atmos. Meas. Tech. 3, 475–493. https://doi.org/10.5194/amt-3-475-2010.

Levy, R.C., Mattoo, S., Munchak, L.A., Remer, L.A., Sayer, A.M., Patadia, F., Hsu, N.C., 2013. 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.

Levy, R.C., Mattoo, S., Sawyer, V., Kiliyanpilakkil, P., Shi, Y., Gupta, P., Remer, L., Kim, M., Zhou, Y., Kleidman, R., Tanré, D., Kaufman, Y. J., 2024. Algorithm for remote sensing of tropospheric aerosol over dark targets, NASA Goddard Space Flight Center. Available at: https://darktarget.gsfc.nasa.gov/sites/default/files/users/user9/ATBD_DarkTarget_April3.pdf.

Levy, R.C., Remer, L.A., Mattoo, S., Vermote, E.F., Kaufman, Y.J., 2007. Second-generation operational algorithm: retrieval of aerosol properties over land from inversion of moderate resolution imaging spectroradiometer spectral reflectance. J. Geophys. Res. 112. https://doi.org/10.1029/2006JD007811.

Lim, H., Choi, M., Kim, J., Kasai, Y., Chan, P.W., 2018. AHI/Himawari-8 Yonsei Aerosol Retrieval (YAER): algorithm, validation and merged products. Remote Sens. 10, 699. https://doi.org/10.3390/rs10050699.

Lipponen, A., Reinvall, J., Väisänen, A., Taskinen, H., Lähivaara, T., Sogacheva, L., Kolmonen, P., Lehtinen, K., Arola, A., Kolehmainen, V., 2022. Deep-learning-based post-process correction of the aerosol parameters in the high-resolution Sentinel-3 Level-2 Synergy product. Atmos. Meas. Tech. 15, 895–914. https://doi.org/10.5194/amt-15-895-2022.

Man, W., Tao, M., Xu, L., Xu, X., Jiang, J., Wang, J., Wang, L., Wang, Y., Fan, M., Chen, L., 2024. Improving aerosol retrieval from MISR with a physics-informed deep learning method. IEEE Trans. Geosci. Remote Sens. 62, 4102911. https://doi.org/10.1109/TGRS.2024.3376598.

Nie, X., Mao, Q., 2023. Study on inversion of atmospheric aerosol nonsphericity based on satellite and ground observations. Atmos. Res. 283, 106582. https://doi.org/10.1016/j.atmosres.2022.106582.

Park, S.S., Kim, J., Cho, Y., Lee, H., Park, J., Lee, D.W., Lee, W.J., Kim, D.R., 2025. Retrieval algorithm for aerosol effective height from the Geostationary Environment Monitoring Spectrometer (GEMS). Atmos. Meas. Tech. 18, 2241–2259. https://doi.org/10.5194/amt-18-2241-2025.

Remer, L.A., Kaufman, Y.J., Tanré, D., Mattoo, S., Chu, D.A., Martins, J.V., Li, R.-R., Ichoku, C., Levy, R.C., Kleidman, R.G., Eck, T.F., Vermote, E., Holben, B.N., 2005. The MODIS aerosol algorithm, products, and validation. J. Atmos. Sci. 62, 947–973. https://doi.org/10.1175/JAS3385.1.

Remer, L.A., Levy, R.C., Martins, J.V., 2024. Opinion: Aerosol remote sensing over the next 20 years. Atmos. Chem. Phys. 24, 2113–2127. https://doi.org/10.5194/acp-24-2113-2024.

Sayer, A.M., Govaerts, Y., Kolmonen, P., Lipponen, A., Luffarelli, M., Mielonen, T., Patadia, F., Popp, T., Povey, A.C., Stebel, K., Witek, M.L., 2020. A review and framework for the evaluation of pixel-level uncertainty estimates in satellite aerosol remote sensing. Atmos. Meas. Tech. 13, 373–404. https://doi.org/10.5194/amt-13-373-2020.

Sayer, A.M., Hsu, N.C., Bettenhausen, C., Ahmad, Z., Holben, B.N., Smirnov, A., Thomas, G.E., Zhang, J., 2012. SeaWiFS Ocean Aerosol Retrieval (SOAR): Algorithm, validation, and comparison with other data sets. J. Geophys. Res. 117. https://doi.org/10.1029/2011JD016599.

She, L., Zhang, H.K., Li, Z., de Leeuw, G., Huang, B., 2020. Himawari-8 aerosol optical depth (AOD) retrieval using a deep neural network trained using AERONET observations. Remote Sens. 12, 4125. https://doi.org/10.3390/rs12244125.

Spurr, R.J.D., 2006. VLIDORT: a linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media. J. Quant. Spectrosc. Radiat. Transf. 102, 316–342. https://doi.org/10.1016/j.jqsrt.2006.05.005.

Su, X., Wei, Y., Wang, L., Zhang, M., Jiang, D., Feng, L., 2022. Accuracy, stability, and continuity of AVHRR, SeaWiFS, MODIS, and VIIRS deep blue long-term land aerosol retrieval in Asia. Sci. Total Environ. 832, 155048. https://doi.org/10.1016/j.scitotenv.2022.155048.

Tao, M., Chen, J., Xu, X., Man, W., Xu, L., Wang, L., Wang, Y., Wang, J., Fan, M., Shahzad, M.I., Chen, L., 2023. A robust and flexible satellite aerosol retrieval algorithm for multi-angle polarimetric measurements with physics-informed deep learning method. Remote Sens. Environ. 297, 113763. https://doi.org/10.1016/j.rse.2023.113763.

Torres, O., Ahn, C., Chen, Z., 2013. Improvements to the OMI near-UV aerosol algorithm using A-train CALIOP and AIRS observations. Atmos. Meas. Tech. 6, 3257–3270. https://doi.org/10.5194/amt-6-3257-2013.

Torres, O., Bhartia, P.K., Jethva, H., Ahn, C., 2018. Impact of the ozone monitoring instrument row anomaly on the long-term record of aerosol products. Atmos. Meas. Tech. 11, 2701–2715. https://doi.org/10.5194/amt-11-2701-2018.

Torres, O., Tanskanen, A., Veihelmann, B., Ahn, C., Braak, R., Bhartia, P.K., Veefkind, P., Levelt, P., 2007. Aerosols and surface UV products from ozone monitoring Instrument observations: an overview. J. Geophys. Res. 112, D24S47. https://doi.org/10.1029/2007JD008809.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. Adv. Neural Inf. Process. Syst. 30. https://doi.org/10.48550/arXiv.1706.03762.

von Hoyningen-Huene, W., Yoon, J., Vountas, M., Istomina, L.G., Rohen, G., Dinter, T., Kokhanovsky, A.A., Burrows, J.P., 2011. Retrieval of spectral aerosol optical thickness over land using ocean color sensors MERIS and SeaWiFS. Atmos. Meas. Tech. 4, 151–171. https://doi.org/10.5194/amt-4-151-2011.

Xu, X., 2020. Corn cash price forecasting. Amer. J. Agr. Econ. 102, 1297–1320. https://doi.org/10.1002/ajae.12041.

Yeom, J.M., Jeong, S., Ha, J.S., Lee, K.H., Lee, C.S., Park, S., 2022. Estimation of the hourly aerosol optical depth from GOCI geostationary satellite data: deep neural network, machine learning, and physical models. IEEE Trans. Geosci. Remote Sens. 60, 1–12. https://doi.org/10.1109/TGRS.2021.3107542.

Zhang, Y., Xu, X., 2021. Solid particle erosion rate predictions through LSBoost. Powder Tech 388, 517–525. https://doi.org/10.1016/j.powtec.2021.04.072.

Zhang, Y., Xu, X., 2022. Disordered MgB2 superconductor critical temperature modeling through regression trees. Physica C: Supercond. Appl. 597, 1354062. https://doi.org/10.1016/j.physc.2022.1354062.

Zhu, L., Martins, J.V., Remer, L.A., 2011. Biomass burning aerosol absorption measurements with MODIS using the critical reflectance method. J. Geophys. Res. 116, D07202. https://doi.org/10.1029/2010JD015187.