First Atmospheric Aerosol Monitoring Results from Geostationary Environment Monitoring Spectrometer (GEMS) over Asia

Yeseul Cho[1], Jhoon Kim[1], Sujung Go[2,3], Mijin Kim[4], Seoyoung Lee[2,3], Minseok Kim[1], Heesung Chong[5], Won-Jin Lee[6], Dong-Won Lee[6], Omar Torres[3], Sang Seo Park[7]

[1]Department of Atmospheric Sciences, Yonsei University, Seoul, Republic of Korea

[2]Goddard Earth Sciences Technology and Research (GESTAR) II, University of Maryland, Baltimore County, Baltimore, MD 21250, USA

[3]NASA Goddard Space Flight Center, Greenbelt, MD, USA

[4]Goddard Earth Sciences Technology and Research (GESTAR) II, Morgan state university, Baltimore, MD 21251, USA

[5]Center for Astrophysics | Harvard & Smithsonian, Cambridge, MA 02138, USA

[6]National Institute of Environmental Research, Incheon, Republic of Korea

[7]Department of Civil, Urban, Earth and Environmental Engineering, Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea

Correspondence to: Jhoon Kim (jkim2@yonsei.ac.kr)

Abstract. Aerosol optical properties have been provided from the Geostationary Environment Monitoring Spectrometer (GEMS), the world's first geostationary earth orbit (GEO) satellite instrument designed for air quality monitoring. This study describes improvements to the GEMS aerosol retrieval algorithm (AERAOD), which include spectral binning, surface reflectance estimation, cloud masking, and post-processing, along with validation results. These enhancements are aimed at providing more accurate and reliable aerosol-monitoring results for Asia. The adoption of spectral binning in the lookup table (LUT) approach reduces random errors and enhances the stability of satellite measurements. In addition, we introduced a new high-resolution database for surface reflectance estimation based on the minimum reflectance method adapted to the GEMS pixel resolution. Monthly background aerosol optical depth (BAOD) values were used to estimate hourly GEMS surface reflectance consistently. Advanced cloud removal techniques have been implemented to significantly improve the effectiveness of cloud detection and enhance aerosol retrieval quality. An innovative post-processing correction method based on machine learning is introduced to address artificial diurnal biases in aerosol optical depth (AOD) observations. In this study, we investigated selected aerosol events. This highlights the capability of GEMS to monitor and provide insights into hourly aerosol optical properties during various atmospheric events. The performance of the GEMS AERAOD products was validated against the Aerosol Robotic Network (AERONET) and Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) data for the period from November 2021 to October 2022. The GEMS AOD at 443 nm demonstrates a strong correlation with the AERONET AOD at 443 nm ($R = 0.792$). However, it exhibited biased patterns, including underestimation of high AOD values and overestimation of low AOD conditions. Different aerosol types (highly absorbing fine, dust, and non-absorbing) exhibited distinct validation results. The GEMS single scattering albedo (SSA) at 443 nm retrievals agreed well with the AERONET SSA at 440 nm within reasonable error ranges, with variations observed among aerosol types. For GEMS AOD at 443 nm exceeding 0.4 (1.0), 42.76% (56.61%) and 67.25% (85.70%) of GEMS SSA data points fall within the ±0.03 and ±0.05 error bounds, respectively. Model-enforced post-processing correction improved the GEMS AOD and SSA performance, thereby reducing the diurnal variation in the biases. The validation of the GEMS aerosol layer height (ALH) retrievals against the CALIOP data demonstrates a good agreement, with a mean bias of -0.225 km and 55.29% (71.70%) of data within ±1 km (1.5 km).

1 Introduction

The regional and global monitoring of aerosol optical properties (AOPs) was conducted using satellite measurements. Low earth orbit (LEO) instruments such as the Advanced Very High-Resolution Radiometer (AVHRR), Moderate Resolution Imaging Spectroradiometer (MODIS), Multiangle Imaging Spectro Radiometer (MISR), Visible Infrared Imaging Radiometer Suite (VIIRS), and Sea-viewing Wide Field-of-view Sensor (SeaWiFS), can provide daily aerosol properties for the global domain (Hsu et al.,2004,2006,2017,2019; Jackson et al., 2013; Jethva et al.,2007; Levy et al., 2013; Lyapustin et al., 2018; Lee et al., 2012; Martonchik et al., 2009; Remer et al., 2005). Although significant diurnal variations in AOPs have been observed at daily and local scales (Kassianov et al., 2013; Kuang et al., 2015), emphasizing the importance of geostationary satellite measurements for both air quality and climate studies, the temporal resolution of LEO satellites (typically one per day) has limitations in investigating the diurnal variation and transboundary transportation of aerosols (Lennartson et al., 2018; Zhang et al., 2018). Geostationary earth orbits (GEO) instruments such as the Advanced Baseline Imager (ABI), Geostationary Ocean Color Imager (GOCI), GOCI-II, Meteorological Imager (MI), and Advanced Himawari Imager (AHI), have contributed to the operational monitoring of the continuous spatio-temporal variations in AOPs at continental spatial scales with temporal resolutions of minutes to hours using the visible and near-infrared channel (Choi et al., 2018; Kim et at., 2016; Kim et al., 2014; Kondragunta et al., 2020; Lee et al., 2023; Yoshida et al., 2018).

In addition to spatial and temporal resolutions, channel specification is another critical consideration for satellite aerosol retrieval. All instruments except GOCI-II used only visible (Vis) and near-infrared channels. However, the near-ultraviolet (UV) spectral region uniquely leverages the sensitivity to aerosol absorption. Therefore, this study provides valuable insights into the optical properties of aerosols. A significant advantage of near-UV measurements is that surface reflectance in the near-UV region is darker than that in the visible region. This enables the derivation of AOPs over a bright surface, typically aerosol source regions. In addition, observations in the UV region are sensitive to aerosol radiative absorption and aerosol layer height (ALH) information. The contribution of Rayleigh scattering to the total Top of the Atmosphere (TOA) reflectance enhancement is reduced below the aerosol layer owing to aerosol attenuation (Kayetha et al., 2022; Torres et al., 2005).

The Ozone Monitoring Instrument (OMI) serves as an example of an LEO sensor that utilizes UV wavelengths for aerosol retrievals. It has measured radiances in the 270–500 nm spectral range and offered global coverage at a spatial resolution of 13 × 24 km at nadir since 2004 (Levelt et al., 2018). OMI employs two aerosol algorithms. The first one, OMAERO (Curier et al., 2008), developed and maintained by the Royal Netherlands Meteorological Institute (KNMI), is a multiwavelength algorithm that relies on spectral fitting procedures to derive aerosol properties. The other is the OMI near-UV aerosol retrieval algorithm (OMAERUV). It focuses on retrieving key atmospheric aerosol properties, including aerosol optical depth (AOD), single scattering albedo (SSA), and absorbing aerosol index (AI) (Torres et al., 2007).

The OMAERUV algorithm has its heritage in the Total Ozone Mapping Spectrometer (TOMS) aerosol retrieval algorithm. It uses reflectance measurements at 354 and 388 nm to determine AOD and single scattering albedo (SSA) using the two-channel inversion method (Torres et al., 2002; Torres et al., 2007). Global statistics reported by Ahn et al. (2014) indicate a correlation coefficient (R) of 0.81. However, OMAERUV provided a lower R (0.63) over Central and East Asia (Zhang et al., 2015). In addition, the Tropospheric Monitoring Instrument (TROPOMI) aerosol algorithm (TropOMAER) was developed as an adaptation of the OMAERUV. A comparison between Aerosol Robotic Network (AERONET) and TropOMAER AOD at 12 locations yielded an R of 0.82 and a root mean square error (RMSE) of 0.19 (Torres et al., 2020).

The Geostationary Environment Monitoring Spectrometer (GEMS) is the first UV-Vis hyperspectral satellite instrument in GEO. It is onboard Geostationary Korea Multi-Purpose Satellite-2B (GEO-KOMPSAT-2B or GK-2B), launched on February 19, 2020 (Kim et al., 2020). The objective of the GEMS mission is to monitor hourly air quality in Asia (5°–45°N, 75°–145°E) with a fine spatial resolution (3.5 × 7.7 km^2 in Seoul, South Korea). GEMS provides hyperspectral measurements covering 300–500 nm at 0.2 nm spectral sampling and 0.6 nm full width at half maximum (FWHM) spectral resolution. The GEMS retrieval domain coverage changes with time because of the varying GEMS scan patterns with the solar zenith angle (SZA). The GEMS aerosol retrieval (AERAOD) algorithm is based on the OMAERUV algorithm and the optimal estimation (OE) method by determining the optimized values of AOD, SSA, and ALH from GEMS measurements at six wavelengths (354, 388, 412, 443, 477, and 490 nm). This algorithm employs the two-channel inversion method used in the OMAERUV algorithm to retrieve the AOD and SSA to overcome the challenge posed by the limited degrees of freedom for signals in the GEMS wavelength range. Subsequently, these retrievals were used as first estimates for the OE method (Kim et al., 2018). The six wavelengths in the UV-Vis region contained information regarding aerosol absorption in the UV region and the absorption bands of the oxygen dimer (O_2-O_2) at 477 nm. Before the GEMS was launched, this method was first tested using OMI Level 1 data and used to derive key aerosol parameters, including AOD, SSA, ALH, UV, and VisAI (Jeong et al., 2016; Kim et al., 2018; Go et al., 2020a, 2020b). Kim et al. (2018) reported that a comparison between AERONET and GEMS AOD at 26 locations in Asia yielded an R of 0.71 and an RMSE of 0.46. The percentage of GEMS SSA within the expected error range of the AERONET inversion data (±0.03) was denoted by 27.54%. Spectral variations in aerosol absorption in the UV-Vis region, as investigated by Go et al. (2020a), were applied to the GEMS aerosol algorithm to achieve improved AOPs retrieval. This adjustment accounts for the spectral dependence of aerosol absorption, which was previously treated as independent of wavelength. The GEMS AOD demonstrated a strong correlation with the AERONET AOD (R = 0.847 and RMSE = 0.285), and the percentage of GEMS SSA within the expected error of ±0.03 increased to 41.64% (Go et al., 2020a). To improve the accuracy of the GEMS aerosol retrieval, Go et al.. al. (2020b) tested the use of cloud mask information from MODIS IR channels to remove cirrus and sub-pixel cloud contamination, as well as the total dust confidence index for the classification of aerosol type. The limitations associated with the UV-Vis regions of GEMS were overcome using the IR channels of other satellites, leading to research on the synergistic use of hyperspectral satellite instruments and broadband meteorological imagers.

However, because the testbed for the GEMS algorithm was on the LEO platform, a time-dependent retrieval bias was not previously observed. The diurnal variations in the satellite-retrieved AOPs may differ from the actual diurnal variations. This discrepancy can be explained by the different patterns of bias observed over time among different GEO satellites and retrieval algorithms (Choi et al., 2018; Lennartson et al., 2018; Wei et al., 2019; Zhang et al., 2020). This diurnal bias in AOP measurements can originate from various factors, such as errors in the surface reflectance assumption used in the retrieval algorithm, calibration issues in the Level 1 data, or the presence of short light paths at noon (Ceamanos et al., 2023).

To address this, Zhang et al.. al. (2020) developed an empirical AOD bias-correction algorithm that utilized the lowest AOD values observed within 30 days in conjunction with the background AOD to obtain a smoothed bias curve for each pixel of the ABI AOD data. This approach helps mitigate the impact of diurnal bias in satellite AOD retrievals to improve accuracy by removing artifacts from the retrieval. By applying bias-correction methods, more reliable diurnal variations in AOD can be explained. In addition to traditional statistical methods, bias correction methods based on machine learning have also been proposed. Model-enforced post-processing correction involves the use of a machine-learning-based model to predict errors in conventional aerosol retrievals (Lipponen et al. 2021, 2022a, 2022b). This method was trained to learn the relationship between the input parameters of satellite measurements and the associated retrieval errors. This approach provides a practical and effective method for enhancing the accuracy of aerosol retrieval without extensive modifications to existing retrieval algorithms. It leverages machine-learning capabilities to improve the reliability and precision of hourly aerosol measurements obtained from GEO satellite observations.

In this study, we report on AOPs, including AOD, SSA, and ALH, derived from GEMS operational observations using the GEMS aerosol retrieval algorithm. The remainder of this paper is organized as follows: Section 2 describes the GEMS data and aerosol retrieval algorithm. It also highlights the algorithm updates after the GEMS (IOT) period. Section 3 discusses the post-process correction for near-real-time retrieval. Section 4 discusses the GEMS aerosol monitoring results for dust, biomass burning, and absorbing aerosol events over Asia. Section 5 presents an evaluation of the GEMS AOD, SSA, and ALH retrievals against AERONET and Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) data and directions for future work. Finally, Section 6 presents a summary.

2 Data and GEMS aerosol algorithm

2.1 Data description

2.1.1 GEMS normalized radiance

The National Institute of Environmental Research (NIER) of Korea provides the GEMS Level-1C (L1C) dataset, which includes various auxiliary variables necessary for retrieval to improve the efficiency of the Level 2 algorithm. In this study, the aerosol retrieval algorithm used radiances only with the quality flags of 0 (Good) or 2 (interpolated radiances), determined by the "bad_pixel_mask" variable. Rather than the GEMS irradiance, we used the KNMI solar reference spectrum to calculate the GEMS-normalized radiance (Dobber et al., 2008). The GEMS irradiance is within the range of -5% to -20% compared with the KNMI solar reference spectrum. Further improvements in L1 processing are ongoing. The KNMI solar reference spectrum was convolved with a GEMS spectral response function (Kang et al., 2020). GEMS-measured irradiances will be employed when the NIER releases an improved version of the Sun L1C product.

Normalized radiances are defined in the following equation:

$$N_{\lambda} = \frac{I_{\lambda}}{ESD \times E_{\lambda}} \quad (1)$$

where *I*, *E*, *ESD*, and λ are the GEMS radiance, KNMI solar reference spectrum, earth–sun distance correction factor, and wavelength (354, 388, 412, 443, 477, and 490 nm), respectively. Spectral radiance and irradiance were spectrally binned and averaged within ±2.2 nm of each wavelength to enhance the measurement signals. Additionally, Earth-Sun distance correction was used to calculate the normalized radiance.

2.1.2 AERONET

AERONET is a global ground-based remote-sensing network that measures aerosol optical, microphysical, and radiative properties (Giles et al., 2019; Holben et al., 1998; Sinyuk et al., 2020). The measurement systems used Cimel sun photometers to measure solar irradiance at eight wavelengths ranging from 340 to 1020 nm and sky radiances at four wavelengths ranging from 440 to 1020 nm. The AERONET data provide global aerosol information, including spectral AOD and inversion products, such as the SSA, aerosol size distribution, and refractive index. The uncertainties in AODs are wavelength-dependent. It is approximately 0.01 (Vis) to 0.02 (Near-UV) in direct sun measurements (Dubovik et al., 2002). The uncertainties of SSA are ±0.03 when AOD exceeds 0.4 at 440 nm (Dubovik et al., 2002). For the evaluation of GEMS AOD and SSA data from November 2021 to October 2022, we used AERONET V3 Level 2.0 data for AOD and AERONET V3 Level 2.0 hybrid inversion data for SSA from all sites within the entire GEMS domain, ensuring higher quality compared to Level 1.5. However, we used AERONET V3 Level 1.5, data for AOD, and AERONET V3 Level 1.5 hybrid inversion data for SSA for post-process correction to ensure a sufficient volume of data during the modelling and near-real-time processing.

2.1.3 CALIOP

The CALIOP instrument is a two-wavelength polarization-sensitive lidar on the cloud aerosol lidar and infrared pathfinder satellite observation (CALIPSO) satellite. It was launched on April 28, 2006 (Winker et al., 2009). CALIOP monitors the global vertical profiles of aerosols and clouds by measuring three signals: backscatter intensity at 1064 nm and the orthogonally polarized components of the backscattered signal at 532 nm.

Quantitative scattering information from the CALIOP instruments was used as reference data to validate the ALH obtained from passive sensors (Xu et al., 2017; Xu et al., 2019; Nanda et al., 2020; Park et al., 2023). We used the CALIPSO Lidar Level 2 Aerosol Profile V3-41 data to validate the GEMS ALH. CALIOP profiles of the extinction coefficient () at the 532 nm channel were utilized to calculate the CALIOP ALH using the following equation:

$$Z_{aer} = \sum_{i-1}^{n} H(i) \left[\frac{\beta_{ext}(i)}{\sum_{i=1}^{n} \beta_{ext}(i)} \right] \quad (2)$$

where (i) is the CALIOP profile of the 532 nm extinction coefficient at height $H(i)$, and n is the number of layers.

2.2 GEMS AERAOD retrieval algorithm

2.2.1 Aerosol optical properties retrieval algorithm for GEMS

The GEMS AERAOD algorithm produces AOD, SSA, and ALH data using the OE method. An early version of the GEMS AERAOD was developed using OMI L1B normalized radiance (Kim et al., 2018; Go et al., 2020a, 2020b). After the launch, the algorithm was tested using GEMS observations during the IOT period, and several parts of the algorithm were updated. This section briefly describes the GEMS AERAOD algorithm, AERAOD L2 data, and updates, including the lookup table (LUT), cloud-masking procedure, surface reflectance estimation, and post-processing after the IOT period. The general flow of the GEMS AERAOD retrieval algorithm is illustrated in Figure 1.

The GEMS algorithm adopts an LUT approach to optimize computation efficiency. The LUT was calculated assuming the AOPs of three aerosol types using a radiative transfer model (RTM), the Vector Linearized Discrete Ordinate Radiative Transfer code (VLIDORT) (Spurr, 2006). The highly absorbing fine (HAF), Dust, and Non-absorbing (NA) are integrated from the AERONET inversion data and applied for the RTM simulation. The details of the updated LUT are described in Section 2.1.2. The preliminary algorithm used the OMI climatology Lambertian equivalent reflectance (OMLER v003) datasets as surface reflectance (Kleipool et al., 2008). However, for the GEMS AERAOD algorithm, GEMS L2 surface reflectances at 354, 388, 412, 443, 477, and 490 nm were obtained using the minimum reflectance method. The details of the surface reflectance estimation are described in Section 2.1.3.

GEMS AERAOD provides UV and visible (Vis) AI to indicate qualitative radiative absorptivity and particle size information, respectively (Torres et al., 2002). The GEMS UVAI and VisAI were calculated using the following equations:

$$\mathrm{AI} = -100 \left[\log\left(\frac{N_{\lambda_1}}{N_{\lambda_2}}\right)_{meas} - log\left(\frac{N_{\lambda_1}(LER_{\lambda_1})}{N_{\lambda_2}(LER_{\lambda_2})}\right)_{calc} \right] \quad (3)$$

where and are the normalized radiances at the 354/388 (477/490) nm wavelength pair for UVAI and VisAI, respectively. Subscripts *meas* and *calc* represent the measured and calculated normalized radiances, respectively.

The aerosol types HAF, dust, and NA were selected using UVAI and VisAI. A negative UVAI value was detected for the NA type. The dust and HAF types were distinguished using the VisAI. HAF type was selected when UVAIs are positive and VISAI is negative. The dust type was selected when both AIs were positive. Sun glint and cloud masking leave only pixels appropriate for aerosol retrieval. The glint mask was set for glint angles less than 35°. The details of the cloud-masking procedure are described in Section 2.1.4. The *a priori* states of AOD and SSA at 443 nm were obtained by two-channel inversion with neighboring wavelengths (354 and 388 nm) over both land and ocean, with *a priori* states of ALH based on the climatology of CALIOP ALH. The *a priori* states of AOD and SSA were supplied to solve the Levenberg–Marquardt equation (Rodgers, 2000). The optimal ALH was determined by fitting the normalized radiance between the measured and calculated values for the OE routine. Details of the GEMS aerosol inversion procedure are described in Kim et al. (2018).

To improve the accuracy of near real-time GEMS AOD retrieval, a model-enforced post-processing correction step was implemented using a random forest (RF) model. By combining GEMS aerosol retrieval with this post-processing correction model, more reliable and accurate near-real-time AOD estimates can be obtained.

2.1.2 LUT calculation

In this study, the AOPs were determined as described by Kim et al. (2018) and Go et al. (2020a). However, the dimensions of the LUT varied (Table 1), which is different from Kim et al. (2018). The nodes for the 412 nm SSA node for the NA were added. In addition, the nodes for the AOD in the LUT were extended to include the values at 5.0° and 10.0°, enabling the retrieval of exceptionally severe aerosol events during GEMS observations. The early version of the GEMS AERAOD retrieval algorithm utilized normalized radiance at six specific monochromatic wavelengths (354, 388, 412, 443, 477, and 490 nm). However, satellite measurements averaged over a specific wavelength range produce more stable values than measurements obtained at individual monochromatic wavelengths, owing to the averaging of random errors (i.e., instrument noise).

Consequently, a spectral-binning LUT approach was employed to reduce random errors and improve measurement stability. This enabled more reliable and consistent observations. Compared to monochromatic wavelengths, the spectral binning method is computationally intensive. Therefore, the calculations were performed using a forward RTM coupled with the Mie theory. The aerosol parameters, including the mean radii and standard deviations of the fine and coarse modes, respectively, of the aerosol bimodal number size distribution, fine mode particle fraction with respect to the total number concentration, and real part of the refractive index, were used to generate the LUT (Kim et al., 2018).

The process of spectral binning LUT in the GEMS aerosol algorithm involves three steps: 1) A reference spectrum is generated using an RTM, which provides a spectral interval of 0.1 nm. 2) The calculated spectrum was convolved with the GEMS spectral response function and resampled to the target spectral grids with a resolution of 0.2 nm (Kang et al., 2020). 3) The resampled spectrum is averaged at intervals of ±2.2 nm at six central wavelengths (354, 388, 412, 443, 477, and 490 nm) and saved in the LUT. Intervals of ±2.2 nm were selected to account for calculation capacity and reduce the impact of random errors. During the retrieval process, the GEMS L1C normalized radiances, after being averaged at intervals of ±2.2 nm at six central wavelengths, are compared with the calculated spectrum in the LUT. Through these steps, the spectral-binning LUT aims to generate more stable retrieval results for aerosol properties.

2.1.3 Surface reflectance estimation

Several improvements were introduced in this study. These include an updated GEMS surface reflectance estimation method. The early version of the GEMS AERAOD retrieval algorithm used the OMI surface reflectance climatology-data product (OMLER v003) (Kleipool et al. 2008), with a spatial resolution of 0.5 × 0.5°, which is too coarse compared with GEMS pixel size, therefore, resulting in discontinuities in the GEMS AOPs. The updated GEMS surface reflectance had a finer spatial resolution (0.1 × 0.1°) to address this limitation. This aligns closely with the pixel resolution of the GEMS. This enhancement enabled more accurate aerosol retrieval at the pixel level. The compiled hourly surface reflectance indirectly reflects the bidirectional reflectance distribution function (BRDF) effect. In addition, a new hourly surface reflectance database was generated using the minimum reflectance method based on GEMS data (Herman and Celarier, 1997; Hsu et al., 2004). The algorithm adopts the climatological minimum reflectance method for each pixel over a ±15-day window spanning two years. Several tests were performed to evaluate different time windows and methods for constructing an accurate surface reflectance. These tests evaluated the effectiveness of using a ±15-day window as well as alternative options such as a previous 30-day window. In addition, different methods, including the minimum and second minimum reflectance approaches, were evaluated to determine the most suitable method for generating appropriate surface reflectance values (not included in this study).

The background AOD (BAOD) was considered in the retrieval algorithm. The BAOD represents the baseline level of AOD that is consistently present in a region. This was then used to derive the surface reflectance dataset. The Rayleigh scattering, gaseous absorption, and BAOD were corrected during the atmospheric correction process to create a surface reflectance dataset. Recent studies have shown that incorporating BAOD into an algorithm can reduce the uncertainty associated with satellite-based AOD retrieval (Kim et al., 2014, 2021). Zhang et al. (2016) estimated the BAOD as the lowest fifth percentile of the AERONET AOD over two years and improved the performance of the VIIRS aerosol algorithm. It has been observed that Asia experiences relatively high BAOD values with seasonal variation. For example, at the Dhaka University site, the monthly BAOD over the past two years varied from a minimum of 0.124 in August to a maximum of 0.685 in April. Therefore, considering seasonal variations in BAOD for atmospheric correction can help mitigate the uncertainty in satellite-derived AOD retrieval, particularly over Asia. The monthly BAODs were calculated using the following equation for each 0.1 × 0.1° box from November 2020 to October 2021:

$$\tau_{grid,b,m}(lat, lon) = \sum_i \; W_i \, \tau_{b,m,i} / \sum_i \; W_i \qquad (4)$$

where is the interpolated BAOD 443 nm at *(lat, lon)* in month *m*. is the inverse distance weighting function, which is defined as . (lat, lon) is the distance between the AERONET site and the GEMS pixel, and is a constant. is the lowest fifth percentile of AERONET AOD over two years at AERONET site *i* in month *m*.

Figure S1 shows the monthly BAOD obtained based on AERONET AOD data. Additionally, the fifth percentiles of the AERONET AOD 443 nm values at each AERONET site are plotted as circles for reference. It is evident that regions such as India exhibit a high BAOD of approximately 0.15 throughout the year, regardless of the month. However, seasonal variations in BAOD occurred over the Indo-Chinese Peninsula, Korea, and China. These areas experience heavy pollution from biomass burning during the dry season and dust events from the deserts. Both these factors contribute to increased atmospheric aerosol concentrations. These enhancements, including the use of hourly GEMS surface reflectance and the incorporation of monthly BAOD, can improve aerosol retrieval.

2.1.4 Cloud masking procedure

The GEMS aerosol algorithm retrieved AOPs only for cloud-free pixels. Clouds exhibit spatial inhomogeneity and a higher brightness than aerosols. This study aimed to enhance the cloud-masking process in the GEMS aerosol algorithm by addressing the limitations of previous simple cloud-masking techniques. The previous method relied on a (Step 1) fixed threshold for reflectance at 477 nm and (Step 2) standard deviation test of reflectance at 477 nm within a 3 × 3 pixel area. An additional cloud-removal technique was introduced in this study to improve the cloud masking performance. These tests included the following: (Step 3) a 470/477 nm normalized radiance ratio test. This involved a threshold test for the ratio of the normalized radiance values at 470 and 477 nm. This contrasts with the presence of clouds using the absorption bands of O_2-O_2 due to the decrease in absorption of O_2-O_2 at 477 nm in the presence of clouds (Kim et al., 2024) (Step 4). The difference between the hourly surface reflectance database and calculated scene reflectivity at 412 nm indicates the presence of clouds (Torres et al., 2013). (Step 5) Standard deviation at 477 nm within 3 × 3 pixels> *f(latitude)*: The threshold for this test can vary based on the latitude, considering regional differences in cloud characteristics. (Step 6) Standard deviation at 477 nm within 3 × 3 pixels> *f(latitude, number of cloud pixels):* The threshold for this test can vary based on the latitude and number of pixels detected as clouds from Steps 1 to 4. A final cloud mask was applied after aerosol retrieval. (Step 7) Filter out the high AOD over the ocean > *f(number of cloud pixels):* a threshold that is a function of the number of cloud pixels detected as clouds from steps 1 to 6 in an 11 × 11-pixel window (Lyapustin et al., 2021). This helps remove residual clouds over the ocean. By implementing these new methods, the proposed algorithm aims to improve the effectiveness of cloud detection and removal in GEMS pixels.

Quantitative analysis was performed to assess the impact of the improvements in each section on the retrieval results of GEMS AOD at 443 nm (Table S1). We analyzed the influence of each update factor on the AOD validation results. The validation periods were January, April, and July 2022. The statistics included R, RMSE, mean bias error (MBE), slope, y-offset, Q value indicating the percentage of AOD retrievals falling within the uncertainty envelope of 0.1 or ±30% of AOD error range, and the Global Climate Observing System (GCOS) requirement is defined as the percentage of AOD retrieval falling within the uncertainty envelope of 0.03 or ± 10% for AOD error range. The early version of the GEMS AERAOD had an MBE of 0.36, indicating an overestimation of the GEMS AOD. When using KNMI irradiance instead of GEMS irradiance and changing to spectral binning LUT, Set1 resulted in a closer MBE of -0.074 to zero and an increased Q-value of 50.63%, approximately 30% higher than the results of the early version of GEMS AERAOD. Set 2 was the result of the analysis using the GEMS surface reflectance instead of the OMI climatology values as the surface reflectance (Section 2.1.3). Set 2 showed a slight decrease in the R-value but an improvement in the Q-value by over 7%. Finally, introducing a new cloud removal method (Set3) increased R and decreased RMSE, leading to an increase in the Q value compared to Set2.

.

3 GEMS post-process correction for the near-real-time retrieval

The GEMS AOD exhibits a diurnal bias pattern that fluctuates throughout the day. It formed a U-shape, with a minimum at 03:00 UTC (as will be demonstrated in Section 5.1). A model-enforced post-processing correction step was implemented using the random forest (RF) model proposed by Lipponen et al. (2021) to improve the accuracy of near real-time GEMS AOD retrieval.

This concept was trained to learn the relationship between hourly GEMS data and AOD errors (GEMS-AERONET AOD) and to predict AOD errors at the target time. The proposed method consists of two main parts: modelling and prediction to enable near-real-time retrieval. In the modelling part, the input data for the RF model included GEMS data (normalized radiances at six wavelengths, scattering angle, viewing zenith angle (VZA), relative azimuth angle (RAA), SZA UV and VisAI, aerosol type, AOD, and a clear fraction (ClearFrac) (which is the ratio of clear-sky pixels to the total pixels within a 0.25° radius from the pixel center)). The data also include auxiliary information, such as time, land-sea mask, and elevation. The target data for training were the AOD errors, which were calculated as the difference between the GEMS AOD and AERONET AOD at the corresponding single GEMS pixel where the AERONET site was located. AERONET data are temporally matched within a ±10 min window of the GEMS measurement time. Data from three AERONET sites (Sorong, Jambi, and BMKG_GAW_PALU) with severe subpixel cloud contamination were excluded from the modeling to exclude cloud-contaminated pixels during the modelling process. The predictors and target variables were collected for a time window ranging from N days to one day before the target time. After conducting several tests, N was determined to be 30 days. In the prediction part, the input variables, including the GEMS data and auxiliary information at the target time, were used for the pre-trained RF model. Using these inputs, the model predicts the error in the GEMS AOD in real-time. This predicted error value was then applied to the first GEMS AOD retrieved using the retrieval algorithm. This resulted in the production of postprocessed GEMS AOD.

To investigate the performance in areas without AERONET data, we conducted a Leave-One-Site-Out Cross-Validation. This principle involves excluding data from one site and training the model using data from all other sites. The performance of the model was evaluated using data from excluded sites. The station selected for evaluation was excluded from the model-fitting process. For the period ranging from 30 days prior to the current day to 1 h before the target day, modelling was conducted, excluding data from one site. The predictive accuracy of the model was evaluated for one site on the target day. Figure S2 shows the statistical maps illustrating the results of the Leave-One-Site-Out Cross-Validation for post-process-corrected GEMS AOD for the one year from November 1, 2021, to October 31, 2022. In Northeast Asia, there was a notably high R, indicating a strong relationship in the AERONET data. However, sites closer to the equator tended to exhibit lower R values, around 0.5. The RMSE followed a similar pattern, with lower values in densely populated Northeast Asia, reflecting a better fit between the predicted and AERONET values in this region. The MBE in Northeast Asia tended to be close to zero, suggesting minimal bias in the predictions. In contrast, the Indian region shows negative MBE values, indicating an underestimation, whereas Southeast Asia shows positive values, signifying an overestimation.

A variable importance analysis for post-processing correction of the GEMS AOD was conducted (Figure S3). The GEMS AOD was the most crucial variable, emphasizing its direct influence on the correction process. The VZA and elevation are highly important. However, their significance can be attributed not only to their inherent properties but also to their role in conveying AERONET location-related information. The aerosol type appeared to be less significant in the RF models. This result contrasts the notable importance of the GEMS UVAI and VisAI. This discrepancy can be attributed to the inaccurate aerosol-type classification in the GEMS aerosol algorithm.

In addition, the diurnal bias pattern in the GEMS SSA exhibited fluctuations throughout the day, forming a bell shape with a minimum at 03:45 UTC. This is discussed in Section 5.2. The post-processing method adopted was similar to that used for the AOD. This method was trained to determine the relationship between hourly GEMS data and SSA errors (the difference between GEMS at 443 nm and AERONET SSA at 440 nm) and to predict SSA errors for the target time. The key difference between the RF model predicting the AOD error and predicting the SSA error is as follows. The second model includes the GEMS SSA as an input variable as well. A variable importance analysis for the post-processing correction of the GEMS SSA was conducted (Figure S4). The GEMS SSA was the most critical variable in the correction process. The GEMS AOD also emerged as a highly influential variable in the RF models for GEMS SSA post-process correction. In addition, the aerosol types appeared to have relatively low significance within the RF models for SSA correction.

Unlike AOD and SSA, the post-processing of ALH using an RF model is inherently limited by the fact that CALIOP is an LEO satellite, and pixels co-located with GEMS ALH data are available only from 03:45 to 07:45 UTC. This rendered it inaccessible as a reference hourly dataset covering 22:45–02:45 UTC. Unlike AEROENT, the use of data from ground-based lidar is severely constrained by the limited number of observation stations and restricted geographical areas in which the lidars are deployed.

4 Aerosol events

4.1 Dust aerosol event (2022.04.08)

Figure 2 presents an example of hourly maps of the GEMS aerosol product, including AOD, SSA, and ALH, for April 8, 2022. Note that these results are the GEMS AOD, SSA, and ALH before post-processing. The GEMS false RGB is shown using R (477 nm), G (412 nm), and B (354 nm) bands similar to those of the OMI false RGB method (Levelt et al., 2006).

As shown in Figure 2, the GEMS retrieval domain coverage changed with time owing to the varying GEMS scan patterns with the SZA. Overall, the GEMS AOD showed a significantly good agreement with the AERONET AOD measurements. It captures higher values in the Beijing–Hebei–Tianjin (BTH) region and lower values in South Korea and Japan. High GEMS AOD values were evident along the dust plume, reaching 2.0 at 06:45 UTC. In the case of SSA, the retrieval results demonstrated a relatively lower accuracy (notably in the BTH region) compared with AOD. In general, from 22:45 to 05:45 UTC, the SSA values displayed good concordance with both the AERONET and GEMS SSA. However, from 06:45–07:45 UTC, the SSA numbers did not match those for Beijing. Compared with the Beijing region, the results were more consistent for the dust plume. The SSA values remained relatively stable at approximately 0.92–0.96 over time. However, the GEMS SSA tended to have a positive bias compared with the AERONET values. This is discussed in Section 5.2. The GEMS ALHs were ~3–4 km for the dust plume over the Taklamakan Desert and ~1.0 km over the Beijing region. The GEMS ALH exhibited continuous spatial and temporal patterns.

4.2 Biomass burning event (2022.03.19)

Figure 3 shows maps of the GEMS aerosol product at 06:45 UTC on March 19, 2022. This represents a biomass-burning event in mainland Southeast Asia. These results were obtained for the GEMS AOD and SSA before post-processing. Fine pollution particles are prevalent in this region during the dry season (Yin et al., 2019). The GEMS AOD > 1.6. This indicated significant aerosol loading and enhancement during the event. The GEMS SSA is approximately 0.88. This indicates aerosol absorption during this event. The ALH ranged from 2 to 3 km within the biomass-burning plume. The GEMS ALH was not retrieved along the east-to-west straight line at ~22.5°N, which is a bad pixel in the CCD. The GEMS UVAI revealed hotspots and fine features associated with this event. Thus, it captures the aerosol absorption in the ultraviolet spectrum. GEMS VisAI did not clearly show signals from small particles caused by biomass burning, indicating that signals from the surface were not completely removed. There may be limitations in considering aerosol size information using GEMS VisAI (Go et al., 2020b). This case study demonstrates that GEMS provides valuable insights into aerosol properties during specific events such as biomass burning, and can capture temporal and spatial variations in AOD, SSA, ALH, UVAI, and VisAI.

Figure 3g shows a comparison of the CALIOP extinction coefficients at 532 nm, the CALIOP ALH, and the GEMS ALH over the CALIOP path (green line on the GEMS false RGB image in Figure 3a). Figure 3g illustrates the precise relationship between the GEMS AOD and the accuracy of the GEMS ALH. Accurate retrieval of ALH requires the presence of a sufficient amount of aerosols in the atmosphere. The GEMS ALH closely followed the latitudinal variation in the CALIOP ALH. As the latitude increased from 18° to 21°, the GEMS ALH followed the CALIOP ALH and exhibited an increase in altitude. In the latitude range of 24°–28°, the GEMS AOD decreased, and the GEMS ALH exhibited scattered variations owing to weaker signals. In the scatter plot comparing CALIOP ALH and GEMS ALH (Figure 3h), 39.88% of the pixels are within the expected error range of 0.5 km, and 68.10% of the pixels are within the expected error range of 1 km. As the GEMS AOD values decreased, the GEMS ALH pixels were more likely to fall outside the expected error range.

4.3 Absorbing aerosol event (2021.12.04, 2021.12.23)

Figure 4 shows an example of the GEMS AOD before and after post-processing for an absorbing aerosol case over the Indo-Gangatic Plane (IGP) at 04:45 UTC on December 4, 2021. Atmospheric haze is prevalent in this region during the winter (Ram et al., 2012). Recent studies have shown that primary aerosols and precursors of secondary aerosols emitted from fossil fuel combustion and biomass burning are released into the atmosphere (Singh et al., 2021). Figure 4a shows a GEMS false RGB image with the AERONET stations represented by circles. Colors indicate AERONET AOD. Two distinct aerosol plumes were observed. The northwest showed an AOD of approximately 0.8, whereas the southeast showed a value of approximately 1.3. Figure 4b shows the GEMS AOD data. The spatial distribution of GEMS AOD was similar to that of AERONET AOD, as shown in Figure 4a. However, the values were marginally lower than those of the AERONET AOD. However, the post-processed AOD showed an elevated value, particularly in the moderate original AOD range (~0.7), bringing the GEMS AOD closer to the AERONET AOD (Figure 4c). Specifically, at the Gandhi-College site (25.871°N, 84.128°E) and Lahore (31.480°N, 74.264°E), post-processing resulted in more reasonable values.

Figure 5 shows the maps of the GEMS SSA and the GEMS SSA after post-processing for an absorbing aerosol case over India, Bangladesh, and mainland Southeast Asia at 03:45 UTC on December 23, 2021. The GEMS false-color RGB with AERONET stations, represented by circles, is shown in Figure 5a. The color indicates the AERONET SSA at 440 nm. AERONET SSA values are ~0.9 in India and Bangladesh and ~0.93 in Thailand. Before post-processing, the GEMS SSAs exhibited values of ~0.96 in the Indian region and ~1.0 in other areas. However, after post-processing, the GEMS SSA values converged and became more similar to the AERONET SSA values. Nonetheless, a marginal tendency toward overestimation remained.

5 Validation in GEMS AERAOD product

This section evaluates the GEMS AOD and SSA at 443 nm according to aerosol type and measurement time using AERONET data in the entire GEMS domain. We used AERONET version 3 level 2.0 data to validate both the AOD and SSA, as it is quality-assured. Figure 6 illustrates a map of the AERONET sites used for GEMS AOD and SSA validation in conjunction with site-specific data counts. The AERONET AOD data generally showed higher counts in South Korea and Taiwan. Sites in South and Southeast Asia typically have fewer data points. The number of AERONET SSA data points showed a distribution similar to that of the AOD. In addition, we retrieved the GEMS ALH and compared it with the CALIPSO level 2 extinction coefficient profiles at 532 nm, as well as with the CALIOP ALH defined by Equation (2).

5.1 Aerosol optical depth

In this section, the GEMS AOD at 443 nm was validated against AERONET data across the entire GEMS domain from November 1, 2021, to October 31, 2022. The GEMS AOD data were spatially collocated within a 0.25° radius of the AERONET stations and temporally within a 30-minute window of the GEMS measurement time. When a specific aerosol type in the GEMS was present in more than 90% of the pixels within the validation radius, aerosol-type validation was conducted.

Figure 7 presents the results for all the pixels and each aerosol type (HAF, dust, and NA). The total GEMS AOD demonstrated a good correlation with the AERONET AOD, with R = 0.781, RMSE = 0.221, and MBE = 0.047 (Figure 7a). The Q value was calculated as 52.93%, with 18.17% of the AOD satisfying the GCOS requirements. However, the slope and y-intercept are 0.572 and 0.202, respectively. This indicated an overestimation of low AERONET AOD and an underestimation of high AERONET AOD. There is evidence of cloud contamination effects in the case of low AERONET AOD. This results in an overestimation of the retrieved GEMS AOD.

The validation showed differences according to aerosol type. The HAF type showed the highest R and Q values compared with the other aerosol types (Figure 7b). Pixels that deviated beyond the error range owing to GEMS AOD underestimation were notably observed in two main categories: sites in the Indian region (which still showed bias notwithstanding the consideration of BAOD) and sites located in Beijing with an AERONET AOD of approximately 2.0, and a GEMS AOD of approximately 1.0. Among the three aerosol types, the dust type had the fewest samples, accounting for 1 / 15 of the NA

(Figure 7c). The R-value was 0.821, and the slope was the highest among the three types. Pixels that deviated beyond the error range due to GEMS AOD underestimation were primarily observed in the Indian region. In contrast, pixels exceeding the error range owing to GEMS AOD overestimation were located in Northeast Asia. Currently, the GEMS uses the same aerosol model (number-size distribution parameters and refractive index) over the entire domain for each aerosol type. However, given the varying bias patterns observed in the dust types, it is necessary to consider regional variations in the GEMS aerosol model (and, thus, the LUT) in future studies. NA was selected most frequently among the three aerosol types (Figure 7d). Figure 7d shows that a significant number of pixels were influenced by cloud contamination, which was particularly evident in regions with low NA AOD values. The GEMS aerosol cloud masking process requires further improvement, particularly over the ocean. The current cloud-masking process may not effectively distinguish small clouds (i.e., broken clouds) near the equatorial regions. This resulted in an overestimation of the AOD owing to cloud contamination. This phenomenon has frequently been observed at AERONET stations located near the equator. The underestimation of high AOD values by the GEMS aerosol algorithm can be attributed to the effects of the current aerosol model assumptions used in the algorithm. This emphasizes the importance of understanding AOPs to better characterize them in the atmosphere, particularly in the UV region.

Figure S5 and Table 2 present the hourly AOD validation results and statistical metrics, including N, R, slope, y-intercept, RMSE, MBE, Q value, and GCOS. It is important to note that the E–W scan profile of the GEMS varied depending on the SZA. Therefore, the sites used for the validation may not have remained consistent over time. For example, the AERONET stations around 22:45 and 23:45 UTC were mainly used for validation in the eastern region of GEMS, whereas those around 06:45 and 07:45 UTC were expected to be located in the western region of GEMS. A systematic error analysis will be planned in future studies. Nevertheless, the hourly validation results of the GEMS AOD provide significant insights. The hourly slopes of the GEMS AOD exhibited diurnal variations, starting at 0.725 at 22:45 UTC, decreasing to 0.490 and 0.533 at 1:45 UTC and 2:45 UTC, respectively, and subsequently increasing to 0.606 and 0.632 at 06:45 and 7:45 UTC, respectively. However, the R-values remained relatively stable over time. Most time intervals exhibited R values of approximately 0.7 or higher except at 00:45. Figure S5 and Table 2 show that the diurnal variation in GEMS AOD does not precisely reflect the actual diurnal AOD variation. Thus, it is necessary to correct and produce a consistent dataset over time to investigate diurnal variations in aerosol properties. A machine-learning model using RF was used to train the hourly dependent error characteristics, remove artifacts in the retrieval processes, and maintain the physical signals.

Figure 8a shows the comparison results for GEMS AOD after model-enforced post-processing correction with AERONET data. Figure 8a shows that all statistical metrics improved. In particular, the slope was closer to one at 0.857, and the y-intercept was closer to zero at 0.049. Additionally, the R, RMSE, and MBE were 0.920, 0.135, and -0.001, respectively. The Q value and GCOS requirements were improved by 82.17% and 37.29%, respectively. The bias near low AOD values of approximately zero was significantly reduced. Furthermore, high AOD values were closer to the 1:1 line. Figure 8b shows the bias of the GEMS AODs before and after post-processing correction with respect to time for all AOD pixels. After applying the model-enforced post-processing correction to the GEMS AOD data, significant improvements in bias were observed over the diurnal cycle. The original GEMS AOD exhibited an hourly-dependent bias. It formed a U-shape, with a minimum value near noon at 03:45 UTC. However, with the implementation of a model-enforced post-processing correction, the diurnal bias was effectively mitigated. This resulted in a bias value close to zero throughout the day and decreased standard deviation. Figure 8c illustrates the diurnal variation in the bias of low AOD (AERONET AOD < 0.4). The GEMS AOD (red circles) exhibits a positive bias of ~0.1. It was mainly corrected to values close to zero after postprocessing (blue circles). However, a positive bias was observed at approximately 22:45 and 23:45 UTC and 06:45 and 07:45 UTC. Figure 8d shows the diurnal variation in the bias of high AOD (AERONET AOD > 0.4). The diurnal variation in GEMS AOD (red circles) shows a clear U-shaped pattern with a maximum negative bias of approximately -0.2 at 0.3 UTC. However, after post-processing, the bias was still negative but less than -0.1, which is significantly closer to zero. By incorporating the predicted error, we obtain an improved GEMS AOD that considers the uncertainties and biases inherent in the retrieval process. This approach helps reduce these biases, including low AOD overestimation, high AOD underestimation, and artificial diurnal bias in near-real-time AOD retrievals. A reduction in artifactual diurnal bias is crucial for ensuring the reliability of hourly GEMS AOD data. This eliminates time-dependent discrepancies and provides a more representative hourly aerosol distribution. Users can now rely on corrected GEMS AOD data for various applications without the influence of diurnal variations in the original measurements.

5.2 Single-scattering albedo

This section presents a comparison of the GEMS SSA at 443 nm with the AERONET SSA at 440 nm over the entire GEMS domain. The validation period and collocation criteria for the AERONET sites were identical to those for GEMS AOD. Similar to the AOD, when a particular aerosol type in the GEMS was detected in over 90% of the pixels within a 0.25° radius, we performed aerosol-type validation. Figure 9 and Table 3 show the validation results for all pixels and each aerosol type. Statistics, including N values and percentages, were within the expected error ranges (0.03 and 0.05). The uncertainty of SSA is ±0.03 when AERONET AOD 440 nm is over 0.4 (Dubovik et al., 2002). The gray dashed lines indicate an uncertainty envelope of ±0.03 in SSA, whereas the black dashed lines indicate an uncertainty envelope of ±0.05 in SSA. These reference lines help to assess the agreement between the GEMS SSA and AERONET data within a reasonable error range. Capturing SSA signals from satellite observations is challenging when atmospheric aerosols are not abundant. Therefore, for validation, separate analyses were conducted for cases where the GEMS AOD was > 0.4 (indicated by the red open circles) and the GEMS AOD was > 1.0 (as indicated by the blue open circles). Notwithstanding the significant uncertainties associated with the satellite measurements, the GEMS aerosol product showed good overall agreement with the AERONET SSA. When GEMS AOD exceeds 0.4, the percentage of GEMS SSA within the expected error range of ±0.03 is denoted by 34.22%, and that within the expected error range of ±0.05 is denoted by 61.38%. When the aerosol signal is strong (when GEMS AOD exceeds 1.0), the percentage of GEMS SSA within the expected error of ±0.03 (0.05) increases to 48.85% (84.48%). However, the percentages within the expected error range and scatter plots varied depending on the aerosol type. For the HAF type, the SSAs exhibited the largest spread. This indicates a lower accuracy. This is likely a result of ineffective aerosol-type selection (red circles). However, when the AOD exceeded 1.0, it tended to approach the 1:1 line (blue circles). Moreover, the percentage falling within the expected error range of ±0.03 increases significantly. For the dust type, the GEMS SSA exhibited a positive bias of approximately 0.04 compared with the AERONET SSA (red circles). Similarly, when the AOD exceeded 1.0, these biases decreased, approaching the 1:1 line (blue circles). However, the systematic bias observed in the GEMS SSA for dust type indicates the need to refine the assumed dust AOPs in the LUT. The NA type in the GEMS was observed to have significantly lower variability than the AERONET SSA. The GEMS SSAs showed values close to one compared with the AERONET data. According to Lee et al. (2010), the NA type is identified when the SSA is above 0.95. However, many NA-type pixels were observed with AERONET SSA values below 0.95 in the NA type. This indicates potential inaccuracies in the classification of the absorbing and NA GEMS aerosol types. However, when the AOD was high (blue circles), these classification errors tended to decrease. This resulted in values closer to those of the AERONET SSA.

Figure S6 and Table 4 present the hourly SSA validation results and statistic metrics, including the N and percentage within the expected error range of ±0.03 (±0.05). The GEMS and AERONET SSA exhibited varying distributions over time. The difference between the GEMS and AERONET SSA was most significant at 03:45 UTC and 04:45 UTC, with a positive bias. This difference decreased at 22:45 and 23:45 UTC and at 05:45 and 06:45 UTC (Figure S6). Similar to the GEMS AOD, the GEMS SSA exhibited diurnal variations. These values are also reflected in the EE% values shown in Table 4. At 22:45 and

23:45 UTC, the percentage within the expected error range of ±0.03 exceeded 64. However, it decreased to less than 19% at 03:45 UTC and 23% at 04:45 UTC before increasing again. Further studies are required to understand the bias and accuracy variations in the SSA and improve the retrieval results. This can also be attributed to the shorter path length in the observation geometry when the influence of the surface reflectance increases, similar to that in AODs.

Figure 10a presents the comparison results for the GEMS SSA after post-process correction and the AERONET data. The validation period was from November 1, 2021, to October 31, 2022. Notably, all statistical metrics demonstrated improvements. Specifically, the percentage of GEMS SSA falling within the expected error range of ±0.03 was 68.54%, whereas the percentage within the range of ±0.05 was 88.95%. Furthermore, the SSA exhibited a closer alignment with the 1:1 line. Figure 10b depicts the difference between the GEMS and AERONET SSA over the measurement time. Notably, the bias pattern observed in the GEMS SSA exhibited artifactual characteristics, thereby forming a bell-shaped curve. In particular, during the time interval from 01:45 to 05:45 UTC, the mean bias of GEMS SSA consistently surpassed the expected error range of ±0.03. However, the implementation of model-enforced post-processing correction was demonstrated to be highly effective in mitigating this artificial diurnal bias. This correction methodology significantly improved the GEMS SSA values within the expected error range. Therefore, it enhanced the overall accuracy of the SSA retrieval.

5.3 Aerosol layer height

From November 1, 2021, to October 31, 2022, the GEMS and CALIOP data were co-located for comparison. In this section, level-2 aerosol extinction coefficients at 532 nm were used to calculate the CALIOP ALH. This is expressed as (2). GEMS ALH pixels within a 0.05° radius surrounding each CALIOP pixel were averaged and compared with the CALIOP ALHs within a time window of 1 h of the GEMS observation time. Validation was conducted when the GEMS AOD values were greater than 0.2. This is because the error in ALH retrieval increased when the presence of aerosols in the atmosphere was insufficient. Figure 11a shows a histogram of the differences between the GEMS and CALIOP ALH. The total co-located number of data is 77,318, and the mean difference is -0.225 km. The median difference was -0.167 km. This indicates that the histogram of the differences follows a Gaussian distribution, although it is marginally skewed in the positive direction. Figure 11b shows a comparison between GEMS and CALIOP ALH. These were distributed predominantly at altitudes of less than 2 km. The percentage of data falling within the expected error of ±1 km was 55.3%, and the percentage falling within the expected error of ±1.5 km was 71.7%. The variability in the GEMS ALH was comparable to that of the CALIOP ALH.

5.4 Limitations of the current GEMS AOPs and future work

Figure S7 shows seasonal and regional variations as a function of UTC for each of the following four regions: Korea (33° N–39° N and 124° E–132° E), North China (33° N–34° N and 110° E–124° E), South China (21° N–33° N and 110° E–122° E), Indochina Peninsula (8° N–22° N and 92° E–110° E). The Indian region was excluded from the regional analysis because the observable area within the entire region of India varied significantly depending on the GEMS scan profiles. After gridding the GEMS AOPs into a 0.1° × 0.1° grid box, monthly averages were calculated. After monthly averaging, seasonal averages were calculated for each pixel only when all three months within a season had data available for the given pixel. Regional averages were calculated when more than 50% of the available values were within the domain. For the AOD, U-shaped or flat diurnal variations were observed in all four regions. In the case of SSA, higher values were observed during June, July, and August (JJA) in Korea, North China, and South China, which are considered to be influenced by aerosol hygroscopic growth owing to relatively high atmospheric humidity. However, the Indochina Peninsula showed the highest SSA values in SON (September, October, and November) and the lowest values in DJF (December, January, and February), which is consistent with the relatively low SSA values observed at the Chiang Mai AERONET site from 2011 to 2016 during DJF (Liang et al., 2019). However, there are limitations to the investigation of diurnal variations in ALH. The diurnal variations in the ALH were not consistent with the diurnal variations in the mixing layer height. One reason for the uncertainty in the ALH is that it is retrieved from the OE depending on the uncertainty of the *a priori* AOD, SSA, and ALH. Before post-processing, GEMS AOD and SSA exhibited diurnal bias patterns compared to the AERONET data (details in Sections 5.1 and 5.2). These uncertainties affect the uncertainty in the diurnal variation of ALH. Because the GEMS ALH cannot be post-corrected using CALIOP data (details in Section 3), we are considering post-process-corrected ALH using ground-based lidar observation networks (i.e., the Korea Aerosol Lidar Observation Network and the Asian dust and aerosol lidar observation network) in future studies. Therefore, one of the limitations of this study is that the GEMS ALH has limitations in the detailed investigation of diurnal variations in ALH.

Several methods can be employed to improve the results of the GEMS aerosol algorithm. First, additional satellite data could be integrated for cloud detection. Incorporating data from other satellite sensors with IR channels, such as the AMI, can provide complementary information for cloud masking. Secondly, it is necessary to consider the AOPs used in the LUT to improve the GEMS aerosol algorithm. It is essential to incorporate additional ground-based observations in the UV region, such as those from the Pandora Instrument and SKYNET. Collecting ground-based observations in the UV region and incorporating them into the LUT can enhance the algorithm's performance. Finally, regional LUTs with data from diverse regions that consider variability in AOPs based on regional characteristics are crucial.

6 Summary

In this study, we present the first atmospheric aerosol monitoring results from GEMS over Asia. Given that the GEMS AERAOD algorithm was developed using OMI as the input data before the GEMS launch, modifications were made to consider GEO observation characteristics during the IOT period. A new hourly surface reflectance database was created using the minimum reflectance method with fine spatial resolution aligned with the GEMS pixel resolution. In addition, monthly BAOD maps were incorporated to estimate hourly GEMS surface reflectance. New cloud-removal techniques have significantly improved the effectiveness of cloud detection and enhanced the quality of aerosol retrieval. To avoid discrepancies between the observed and simulated radiances that may arise because of the monochromatic assumption of the LUT calculation, we applied a spectral binning approach to the LUT calculation. Finally, post-processing correction methods based on machine learning were used to remove non-physical diurnal biases in the AOD and SSA retrieval. This reduced the biases over time and provided more reliable hourly GEMS aerosol products in near-real time.

The GEMS aerosol products were investigated for three specific events: dust events over Northeast Asia, biomass burning in Southeast Asia, and absorbing aerosol in India. These events highlight the capability of the GEMS to monitor and provide insights into aerosol properties during various atmospheric events while also emphasizing the importance of post-processing for data accuracy and agreement with ground measurements.

The GEMS aerosol products were validated against AERONET and CALIOP data for the entire GEMS domain for one year (from November 2021 to October 2022). The performance of the GEMS aerosol algorithm was validated to verify its applicability in studying the distribution of AOPs across Asia. The validation results for each product are summarized below.

The GEMS AOD showed a good correlation with the AERONET AOD (R = 0.792). However, specific biased patterns were observed. Notably, the underestimation of AOD in high AERONET AOD and the overestimation of AOD in low AERONET AOD occurred because of cloud contamination. Different aerosol types exhibited varying validation results: the HAF type with the highest R and Q values; the dust type with underestimation in India but overestimation in Northeast Asia; and the

NA type with cloud contamination issues, particularly for low AOD. This indicates the need for improvement of the cloud-masking process, particularly over the ocean. Certain deviations beyond the error range of GEMS AOD were observed in India and Beijing. The underestimation of the high AOD values can be attributed to the aerosol model. Diurnal variations in the retrieval performance were evident, with varying slopes and other comparison statistics throughout the day. Because the testbed for the GEMS algorithm was on the LEO platform, a time-dependent retrieval bias was not previously observed. Therefore, we adopted a model-enforced post-process correction and found that this enhanced GEMS AOD performance reduced the overall biases. These corrected data ensure the reliability of various applications.

The GEMS SSA at 443 nm was validated against AERONET SSA at 440 nm over the entire GEMS region. The GEMS SSA's agreement with the AERONET data was evaluated within a reasonable error range of ±0.03 (±0.05). For GEMS AOD exceeding 0.4, 42.76 (67.25)% of GEMS SSA is within ±0.03 (0.05) error. This increases to 56.61 (85.70%) for the strong aerosol signals (GEMS AOD > 1.0). However, the accuracy varied among the aerosol types. The HAF type exhibited higher variability and lower accuracy. The dust type had a marginal positive bias, mainly when the AOD was high. Similar to the AOD, the post-processing correction for the GEMS SSA data yielded significant enhancements in the statistical metrics.

The GEMS and CALIOP data were compared. GEMS ALH was compared with CALIOP ALH when the GEMS AOD exceeded 0.2. The results showed a mean difference of -0.225 km, with 55.29% of data being within ±1 km and 71.70% being within ±1.5 km. GEMS ALH exhibited variability similar to that of CALIOP ALH.

Overall, improvements in the GEMS aerosol algorithm have contributed to advancing our understanding of aerosol properties and their effects on the environment. Therefore, it provides valuable information for diverse applications, including air quality monitoring, air quality data assimilation, and health impact assessments in Asia.

Code availability. The GEMS L2 AERAOD algorithm is not available publicly.

Data availability. GEMS L2 AERAOD data were downloaded from the Environmental Satellite Center website (https://nesc.nier.go.kr/en/html/datasvc/index.do).

Author Contribution. YC, JK, SG, and MK designed the experiments. WL and DL provided support for data collection. SL, MK, HC, OT, and SP contributed to algorithm development. YC wrote the manuscript with contributions from all co-authors. JK reviewed and edited the manuscript. JK provided support and supervision. All authors analyzed the measurement data and prepared the manuscript.

Competing Interests. At least one of the (co-)authors is a member of the editorial board of Atmospheric Measurement Techniques.

Acknowledgements. We thank all principal investigators and their staff for establishing and maintaining the AERONET sites used in this investigation. The CALIOP V3.41 data were obtained from the NASA Langley Research Center Atmospheric Science Data Center at https://asdc.larc.nasa.gov/project/CALIPSO. The authors acknowledge the National Institute of Environmental Research Korea Aerospace Research Institute for providing satellite data and Professor Myoung-Hwan Ahn and his research group at Ewha Womans University for providing information regarding GEMS specifications and Level 1 data.

Financial Support. This work was supported by a grant from the National Institute of Environment Research (NIER), funded by the Ministry of Environment (MOE) of the Republic of Korea (NIER-2024-04-02-028). The GEMS program is supported by the National Institute of Environmental Research (NIER), the Ministry of Environment, South Korea. This project is supported by the Korea Ministry of Environment (MOE) as Public Technology Program based on Environmental Policy (2017000160001).

References

Ahn, C., Torres, O., and Jethva, H.: Assessment of OMI near-UV aerosol optical depth over land, J. Geophys. Res. Atmos., 119, 2457–2473, https://doi.org/10.1002/2013jd020188, 2014.

Ceamanos, X., Six, B., Moparthy, S., Carrer, D., Georgeot, A., Gasteiger, J., Riedi, J., Attié, J.-L., Lyapustin, A., and Katsev, I.: Instantaneous aerosol and surface retrieval using satellites in geostationary orbit (iAERUS-GEO) – estimation of 15 min aerosol optical depth from MSG/SEVIRI and evaluation with reference data, Atmos. Meas. Tech., 16, 2575–2599, https://doi.org/10.5194/amt-16-2575-2023, 2023.

Choi, M., Kim, J., Lee, J., Kim, M., Park, Y.-J., Holben, B., Eck, T. F., Li, Z., and Song, C. H.: 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, 2018.

Curier, R. L., Veefkind, J. P., Braak, R., Veihelmann, B., Torres, O., and de Leeuw, G.: Retrieval of aerosol optical properties from OMI radiances using a multiwavelength algorithm: Application to western Europe. J. Geophys. Res. Atmos., 113(D17), 2008.

Dobber, M., Kleipool, Q., Dirksen, R., Levelt, P., Jaross, G., Taylor, S., Kelly, T., Flynn, L., Leppelmeier, G., and Rozemeijer, N.: Validation of Ozone Monitoring Instrument level 1b data products, J. Geophys. Res., 113, https://doi.org/10.1029/2007jd008665, 2008.

Dubovik, O. and King, M. D.: A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements, J. Geophys. Res., 105, 20673–20696, https://doi.org/10.1029/2000JD900282, 2000.

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., Welton, E. J., Korkin, S. V., and Lyapustin, A. I.: 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.

Go, S., Kim, J., Mok, J., Irie, H., Yoon, J., Torres, O., Krotkov, N. A., Labow, G., Kim, M., Koo, J.-H., Choi, M., and Lim, H.: Ground-based retrievals of aerosol column absorption in the UV spectral region and their implications for GEMS measurements, Remote Sens. Environ., 245, https://doi.org/10.1016/j.rse.2020.111759, 2020.

Go, S., Kim, J., Park, S. S., Kim, M., Lim, H., Kim, J.-Y., Lee, D.-W., and Im, J.: Synergistic Use of Hyperspectral UV-Visible OMI and Broadband Meteorological Imager MODIS Data for a Merged Aerosol Product, Remote Sens., 12, https://doi.org/10.3390/rs12233987, 2020.

Herman, J. R. and Celarier, E. A.: Earth surface reflectivity climatology at 340–380 nm from TOMS data, J. Geophys. Res.-Atmos., 102, 28003–28011, 1997.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y., Nakajima, T., Lavenue, F., Jankowiak, I., and Smirnov, A.: 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., Lee, J., Sayer, A. M., Kim, W., Bettenhausen, C., and Tsay, S. C.: 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, 2019.

Hsu, N. C., Si-Chee, T., 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.

Hsu, N. C., Tsay, S. C., King, M. D., and Herman, J. R.: 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, 2006.

Jackson, J. M., Liu, H., Laszlo, I., Kondragunta, S., Remer, L. A., Huang, J., and Huang, H.-C.: Suomi-NPP VIIRS aerosol algorithms and data products, J. Geophys. Res. Atmos., 118, 12,673–612,689, https://doi.org/10.1002/2013jd020449, 2013.

Jeong, U., Kim, J., Ahn, C., Torres, O., Liu, X., Bhartia, P. K., Spurr, R. J. D., Haffner, D., Chance, K., and Holben, B. N.: An optimal-estimation-based aerosol retrieval algorithm using OMI near-UV observations, Atmos. Chem. Phys., 16, 177–193, https://doi.org/10.5194/acp-16-177-2016, 2016.

Jethva, H., Satheesh, S. K., and Srinivasan, J.: Assessment of second-generation MODIS aerosol retrieval (Collection 005) at Kanpur, India, Geophys. Res. Lett., 34, https://doi.org/10.1029/2007gl029647, 2007.

Kalashnikova, O. V., Garay, M. J., Martonchik, J. V., and Diner, D. J.: MISR Dark Water aerosol retrievals: operational algorithm sensitivity to particle non-sphericity, Atmos. Meas. Tech., 6, 2131–2154, https://doi.org/10.5194/amt-6-2131–2013, 2013.

Kang, M., Ahn, M.-H., Liu, X., Jeong, U., and Kim, J.: Spectral Calibration Algorithm for the Geostationary Environment Monitoring Spectrometer (GEMS), Remote Sens., 12, https://doi.org/10.3390/rs12172846, 2020.

Kassianov, E., Barnard, J., Pekour, M., Berg, L. K., Michalsky, J., Lantz, K., and Hodges, G.: Do diurnal aerosol changes affect daily average radiative forcing? Geophys. Res. Lett., 40, 3265–3269, 2013.

Kuang, Y., Zhao, C. S., Tao, J. C., and Ma, N.: Diurnal variations of aerosol optical properties in the North China Plain and their influences on the estimates of direct aerosol radiative effect, Atmos. Chem. Phys., 15, 5761–5772, https://doi.org/10.5194/acp-15-5761-2015, 2015.

Kayetha, V., Torres, O., and Jethva, H.: Retrieval of UV–visible aerosol absorption using AERONET and OMI–MODIS synergy: spatial and temporal variability across major aerosol environments, Atmos. Meas. Tech., 15, 845–877, https://doi.org/10.5194/amt-15-845-2022, 2022.

Kim, M., Kim, J., Wong, M. S., Yoon, J., Lee, J., Wu, D., Chan, P. W., Nichol, J. E., Chung, C. Y., and Ou, M. L.: 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, 2014.

Kim, J., Jeong, U., Ahn, M.-H., Kim, J. H., Park, R. J., Lee, H., Song, C. H., Choi, Y.-S., Lee, K.-H., Yoo, J.-M., Jeong, M.-J., Park, S. K., Lee, K.-M., Song, C.-K., Kim, S.-W., Kim, Y. J., Kim, S.-W., Kim, M., Go, S., Liu, X., Chance, K., Chan Miller, C., Al-Saadi, J., Veihelmann, B., Bhartia, P. K., Torres, O., Abad, G. G., Haffner, D. P., Ko, D. H., Lee, S. H., Woo, J.-H., Chong, H., Park, S. S., Nicks, D., Choi, W. J., Moon, K.-J., Cho, A., Yoon, J., Kim, S.-k., Hong, H., Lee, K., Lee, H., Lee, S., Choi, M., Veefkind, P., Levelt, P. F., Edwards, D. P., Kang, M., Eo, M., Bak, J., Baek, K., Kwon, H.-A., Yang, J., Park, J., Han, K. M., Kim, B.-R., Shin, H.-W., Choi, H., Lee, E., Chong, J., Cha, Y., Koo, J.-H., Irie, H., Hayashida, S., Kasai, Y., Kanaya, Y., Liu, C., Lin, J., Crawford, J. H., Carmichael, G. R., Newchurch, M. J., Lefer, B. L., Herman, J. R., Swap, R. J., Lau, A. K. H., Kurosu, T. P., Jaross, G., Ahlers, B., Dobber, M., McElroy, C. T., and Choi, Y.: New Era of Air Quality Monitoring from Space: Geostationary Environment Monitoring Spectrometer (GEMS), Bull. Am. Meteorol. Soc., 101, E1–E22, https://doi.org/10.1175/bams-d-18-0013.1, 2020.

Kim, M., Kim, J., Jeong, U., Kim, W., Hong, H., Holben, B., Eck, T. F., Lim, J. H., Song, C. K., Lee, S., and Chung, C. Y.:

Aerosol optical properties derived from the DRAGON-NE Asia campaign, and implications for a single-channel algorithm to retrieve aerosol optical depth in spring from Meteorological Imager (MI) on-board the Communication, Ocean, and Meteorological Satellite (COMS), Atmos. Chem. Phys., 16, 1789–1808, https://doi.org/10.5194/acp-16-1789-2016, 2016.

Kim, M., Kim, J., Torres, O., Ahn, C., Kim, W., Jeong, U., Go, S., Liu, X., Moon, K., and Kim, D.-R.: Optimal Estimation-Based Algorithm to Retrieve Aerosol Optical Properties for GEMS Measurements over Asia, Remote Sens., 10, https://doi.org/10.3390/rs10020162, 2018.

Kim, M., Kim, S. H., Kim, W. V., Lee, Y. G., Kim, J., and Kafatos, M. C.: Assessment of Aerosol optical depth under background and polluted conditions using AERONET and VIIRS datasets, Atmos. Environ., 245, https://doi.org/10.1016/j.atmosenv.2020.117994, 2021.

Kleipool, Q. L., Dobber, M. R., de Haan, J. F., and Levelt, P. F.: Earth surface reflectance climatology from 3 years of OMI data, J. Geophys. Res., 113, https://doi.org/10.1029/2008jd010290, 2008.

Kondragunta, S., Laszlo, I., Zhang, H., Ciren, P., and Huff, A.: Air Quality Applications of ABI Aerosol Products from the GOES-R Series, in The GOES-R Series: A New Generation of Geostationary Environmental Satellites, Elsevier, Amsterdam, the Netherlands, Oxford, UK, Cambridge MA, USA, 203–217, 2020.

Lee, J., Kim, J., Yang, P., and Hsu, N. C.: Improvement of aerosol optical depth retrieval from MODIS spectral reflectance over the global ocean using new aerosol models archived from AERONET inversion data and tri-axial ellipsoidal dust database, Atmos. Chem. Phys., 12, 7087–7102, https://doi.org/10.5194/acp-12-7087-2012, 2012.

Lee, S., Choi, M., Kim, J., Park, Y. J., Choi, J. K., Lim, H., Lee, J., Kim, M., and Cho, Y.: 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, 2023.

Lee, Y., Ahn, M.-H., Kang, M., and Eo, M.: Spectral replacement using machine learning methods for continuous mapping of the Geostationary Environment Monitoring Spectrometer (GEMS), Atmos. Meas. Tech., 16, 153–168, https://doi.org/10.5194/amt-16-153-2023, 2023.

Lennartson, E. M., Wang, J., Gu, J., Castro Garcia, L., Ge, C., Gao, M., Choi, M., Saide, P. E., Carmichael, G. R., Kim, J., and Janz, S. J.: Diurnal variation of aerosol optical depth and PM2.5 in South Korea: a synthesis from AERONET, satellite (GOCI), KORUS-AQ observation, and the WRF-Chem model, Atmos. Chem. Phys., 18, 15125–15144, https://doi.org/10.5194/acp-18-15125-2018, 2018.

Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., Malkki, A., Huib, V., Johan de, V., Stammes, P., Lundell, J. O. V., and Saari, H.: The ozone monitoring instrument, IEEE Trans. Geosci. Remote Sens., 44, 1093–1101, https://doi.org/10.1109/tgrs.2006.872333, 2006.

Levelt, P. F., Joiner, J., Tamminen, J., Veefkind, J. P., Bhartia, P. K., Stein Zweers, D. C., Duncan, B. N., Streets, D. G., Eskes, H., van der A, R., McLinden, C., Fioletov, V., Carn, S., de Laat, J., DeLand, M., Marchenko, S., McPeters, R., Ziemke, J., Fu, D., Liu, X., Pickering, K., Apituley, A., González Abad, G., Arola, A., Boersma, F., Chan Miller, C., Chance, K., de Graaf, M., Hakkarainen, J., Hassinen, S., Ialongo, I., Kleipool, Q., Krotkov, N., Li, C., Lamsal, L., Newman, P., Nowlan, C., Suleiman, R., Tilstra, L. G., Torres, O., Wang, H., and Wargan, K.: The Ozone Monitoring Instrument: overview of 14 years in space, Atmos. Chem. Phys., 18, 5699–5745, https://doi.org/10.5194/acp-18-5699-2018, 2018.

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.

Liang, Y., Che, H., Gui, K., Zheng, Y., Yang, X., Li, X., Liu, C., Sheng, Z., Sun, T., and Zhang, X.: Impact of Biomass Burning in South and Southeast Asia on Background Aerosol in Southwest China, Aerosol Air Qual. Res., 19, 1188–1204, https://doi.org/10.4209/AAQR.2018.08.0324, 2019.

Lipponen, A., Kolehmainen, V., Kolmonen, P., Kukkurainen, A., Mielonen, T., Sabater, N., Sogacheva, L., Virtanen, T. H., and Arola, A.: Model-enforced post-process correction of satellite aerosol retrievals, Atmos. Meas. Tech., 14, 2981–2992, https://doi.org/10.5194/amt-14-2981-2021, 2021.

Lipponen, A., Reinvall, J., Väisänen, A., Taskinen, H., Lähivaara, T., Sogacheva, L., Kolmonen, P., Lehtinen, K., Arola, A., and Kolehmainen, V.: 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, 2022.

Lyapustin, A., Wang, Y., Go, S., Choi, M., Korkin, S., Huang, D., Knyazikhin, Y., Blank, K., and Marshak, A.: Atmospheric Correction of DSCOVR EPIC: Version 2 MAIAC Algorithm, Front. Remote Sens., 2, https://doi.org/10.3389/frsen.2021.748362, 2021.

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.

Nanda, S., de Graaf, M., Veefkind, J. P., Sneep, M., ter Linden, M., Sun, J., and Levelt, P. F.: A first comparison of TROPOMI aerosol layer height (ALH) to CALIOP data, Atmos. Meas. Tech., 13, 3043–3059, https://doi.org/10.5194/amt-13-3043-2020, 2020.

Park, S. S., Kim, J., Cho, Y., Lee, H., Park, J., Lee, D.-W., Lee, W.-J., and Kim, D.-R.: Retrieval Algorithm for Aerosol Effective Height from the Geostationary Environment Monitoring Spectrometer (GEMS), Atmos. Meas. Tech. Discuss., https://doi.org/ 10.5194/amt-223-136.

Powell, K. A., Hu, Y., Omar, A., Vaughan, M. A., Winker, D. M., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms, J. Atmos. Ocean. Tech., 26, 2310–2323, https://doi.org/10.1175/2009jtecha1281.1, 2009.

Ram, K., Sarin, M.M., Sudheer, A.K., and Rengarajan, R.: Carbonaceous and Secondary Inorganic Aerosols during Wintertime Fog and Haze over Urban Sites in the Indo-Gangetic Plain, Aerosol Air Qual. Res., 41, 359–370, https://doi.org/10.4209/aaqr.2011.07.0105, 2012.

Rodgers, C. D.: Inverse method for atmospheric sounding: theory and practice. World Scientific Publishing Co. Pte. Ltd., Singapore, 2000.

Sayer, A. M., Hsu, N. C., Lee, J., Carletta, N., Chen, S. H., and Smirnov, A.: Evaluation of NASA Deep Blue/SOAR aerosol retrieval algorithms applied to AVHRR measurements, J. Geophys. Res. Atmos., 122, 9945–9967, https://doi.org/10.1002/2017JD026934, 2017.

Singh, G. K., Choudhary, V., Rajeev, P., Paul, D., and Gupta, T.: Understanding the origin of carbonaceous aerosols during periods of extensive biomass burning in northern India. Environ. Pollut., 270, 116082, https://doi.org/10.1016/j.envpol.2020.116082, 2021.

Sinyuk, A., Holben, B. N., Eck, T. F., Giles, D. M., Slutsker, I., Korkin, S., Schafer, J. S., Smirnov, A., Sorokin, M., and Lyapustin, A.: The AERONET Version 3 aerosol retrieval algorithm, associated uncertainties and comparisons to Version 2, Atmos. Meas. Tech., 13, 3375–3411, https://doi.org/10.5194/amt-13-3375-2020, 2020.

Spurr, R. J. D.: 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, 2006.

Taskinen, H., Väisänen, A., Hatakka, L., Virtanen, T. H., Lähivaara, T., Arola, A., Kolehmainen, V., and Lipponen, A.: High-Resolution Post-Process Corrected Satellite AOD, Geophys. Res. Lett., 49, https://doi.org/10.1029/2022gl099733, 2022.

Torres, O., Bhartia, P. K., Herman, J. R., Syniuk, A., Ginoux, P., and Holben, B.: A long term record of aerosol optical depth from TOMS observations D and comparison to AERONET measurements, J. Atm. Sci., 59, 398–413, 2002.

Torres, O., Ahn, C., and Chen, Z.: 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, 2013.

Torres, O., Bhartia, P. K., Sinyuk, A., Welton, E. J., and Holben, B.: Total Ozone Mapping Spectrometer measurements of aerosol absorption from space: Comparison to SAFARI 2000 ground-based observations, J. Geophys. Res., 110, D10S18, https://doi.org/10.1029/2004JD004611, 2005.

Torres, O., Tanskanen, A., Veihelmann, B., Ahn, C., Braak, R., Bhartia, P. K., Veefkind, P., and Levelt, P.: Aerosols and surface UV products from Ozone Monitoring Instrument observations: An overview, J. Geophys. Res., 112, https://doi.org/10.1029/2007jd008809, 2007.

Torres, O., Jethva, H., Ahn, C., Jaross, G., and Loyola, D. G.: TROPOMI aerosol products: evaluation and observations of synoptic-scale carbonaceous aerosol plumes during 2018–2020, Atmos. Meas. Tech., 13, 6789–6806, https://doi.org/10.5194/amt-13-6789–2020, 2020.

Wei, J., Li, Z., Sun, L., Peng, Y., Zhang, Z., Li, Z., Su, T., Feng, L., Cai, Z., and Wu, H.: Evaluation and uncertainty estimate of next-generation geostationary meteorological Himawari-8/AHI aerosol products, Sci. Total Environ., 692, 879–891, https://doi.org/10.1016/j.scitotenv.2019.07.326, 2019.

Xu, X., Wang, J., Wang, Y., Zeng, J., Torres, O., Reid, J. S., Miller, S. D., Martins, J. V., and Remer, L. A.: Detecting layer height of smoke aerosols over vegetated land and water surfaces via oxygen absorption bands: hourly results from EPIC/DSCOVR in deep space, Atmos. Meas. Tech., 12, 3269–3288, https://doi.org/10.5194/amt-12-3269-2019, 2019.

Xu, X., Wang, J., Wang, Y., Zeng, J., Torres, O., Yang, Y., Marshak, A., Reid, J., and Miller, S.: Passive remote sensing of altitude and optical depth of dust plumes using the oxygen A and B bands: first results from EPIC/DSCOVR at Lagrange-1 point, Geophys. Res. Lett., 44, 7544–7554, https://doi.org/10.1002/2017gl073939, 2017.

Yin, S., Wang, X., Zhang, X., Guo, M., Miura, M., and Xiao, Y.: Influence of biomass burning on local air pollution in mainland Southeast Asia from 2001 to 2016, Environ. Pollut., 254, 112949, https://doi.org/10.1016/j.envpol.2019.07.117, 2019.

Yoshida, M., Kikuchi, M., Nagao, T. M., Murakami, H., Nomaki, T., and Higurashi, A.: Common Retrieval of Aerosol Properties for Imaging Satellite Sensors, J. Meteorol. Soc. Jpn. Ser. II, 96B, 193–209, https://doi.org/10.2151/jmsj.2018-039, 2018.

Zhang, H., Kondragunta, S., Laszlo, I., and Zhou, M.: Improving GOES Advanced Baseline Imager (ABI) aerosol optical depth (AOD) retrievals using an empirical bias correction algorithm, Atmos. Meas. Tech., 13, 5955–5975, https://doi.org/10.5194/amt-13-5955–2020, 2020.

Zhang, H., Kondragunta, S., Laszlo, I., Liu, H., Remer, L. A., Huang, J., Superczynski, S., and Ciren, P.: An enhanced VIIRS aerosol optical thickness (AOT) retrieval algorithm over land using a global surface reflectance ratio database, J. Geophys. Res. Atmos., 121, https://doi.org/10.1002/2016jd024859, 2016.

Zhang, W., Gu, X., Xu, H., Yu, T., and Zheng, F.: Assessment of OMI near-UV aerosol optical depth over Central and East Asia, J. Geophys. Res. Atmos., 121, 382–398, https://doi.org/10.1002/2015jd024103, 2016.

Zhang, Z., Wu, W., Fan, M., Tao, M., Wei, J., Jin J., Tan, Y. and Wang, Q.: Validation of Himawari-8 aerosol optical depth retrievals over China, Atmos. Environ., 199, https://doi.org/10.1016/j.atmosenv.2018.11.024, 2019.

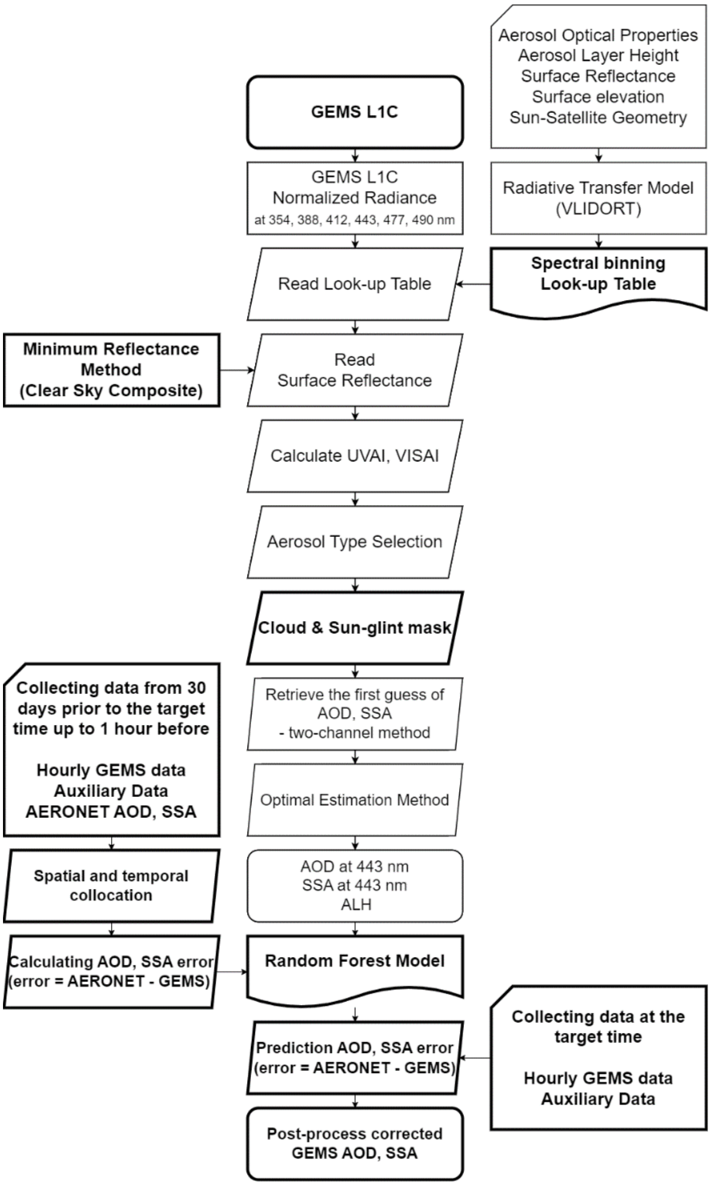

Figure 1: The flowchart of the GEMS AERAOD retrieval algorithm and the modifications in the study (in bold boxes)

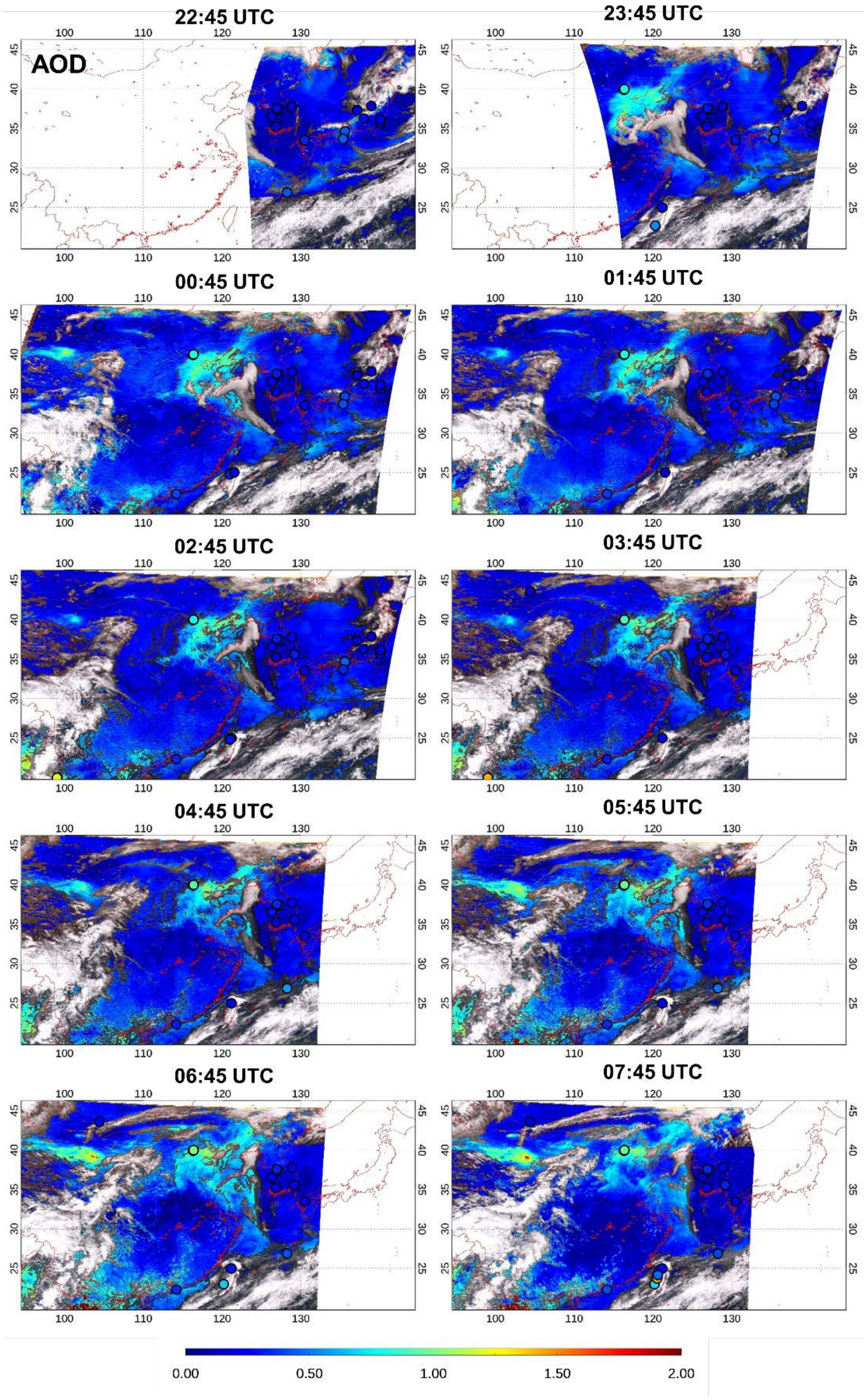

Figure 2: Hourly GEMS aerosol products for the dust case on April 8, 2022 over northwestern China. Time-series maps of AOD at 443 nm, SSA at 443 nm, and ALH (km) from 22:45 UTC to 07:45 UTC. The circle denotes an AERONET station, and the filled color indicates the AERONET AOD and SSA at 443 nm in the AOD and SSA columns. GEMS SSA and ALH are displayed only when GEMS AOD > 0.2.

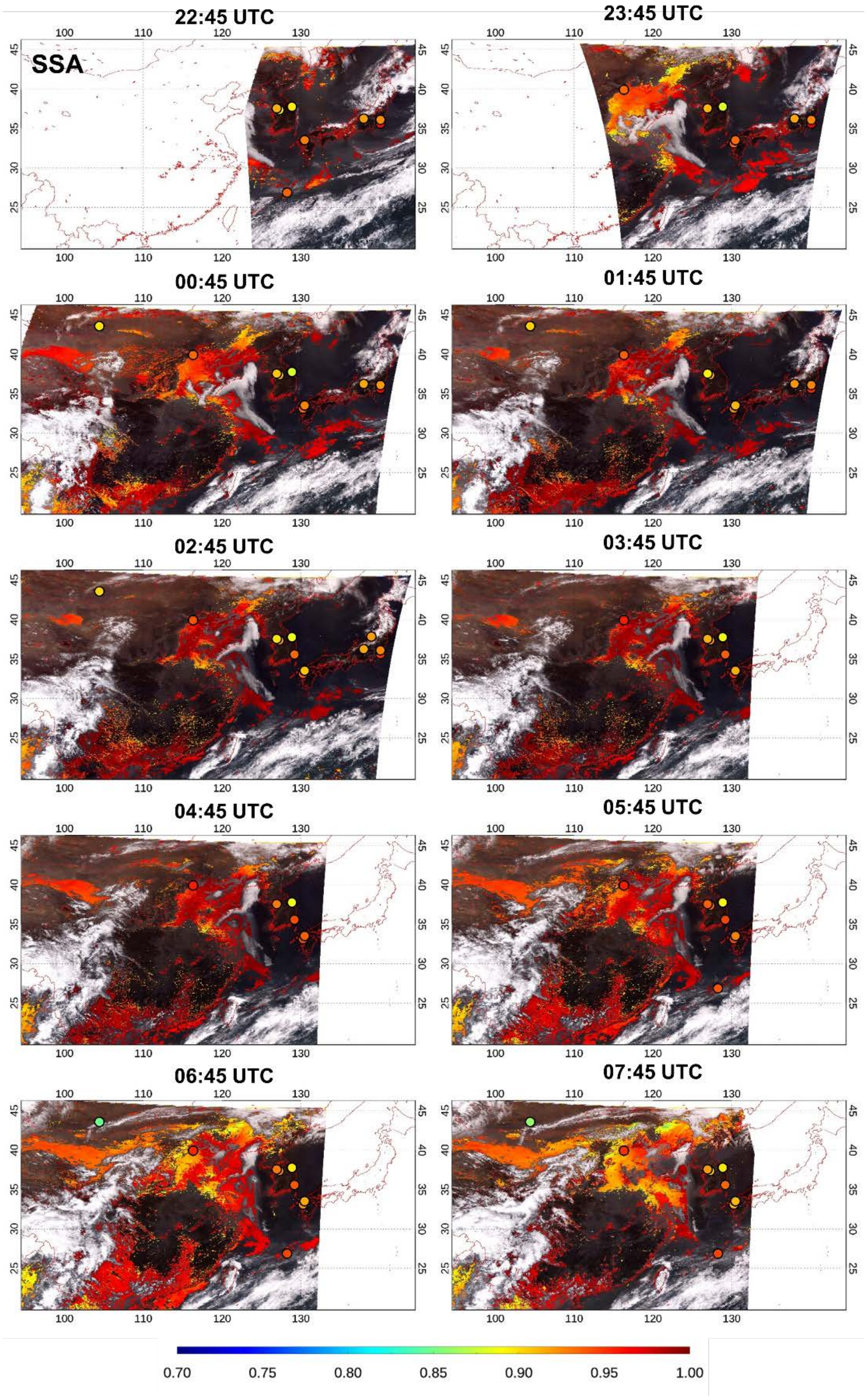

Figure 2: Continue

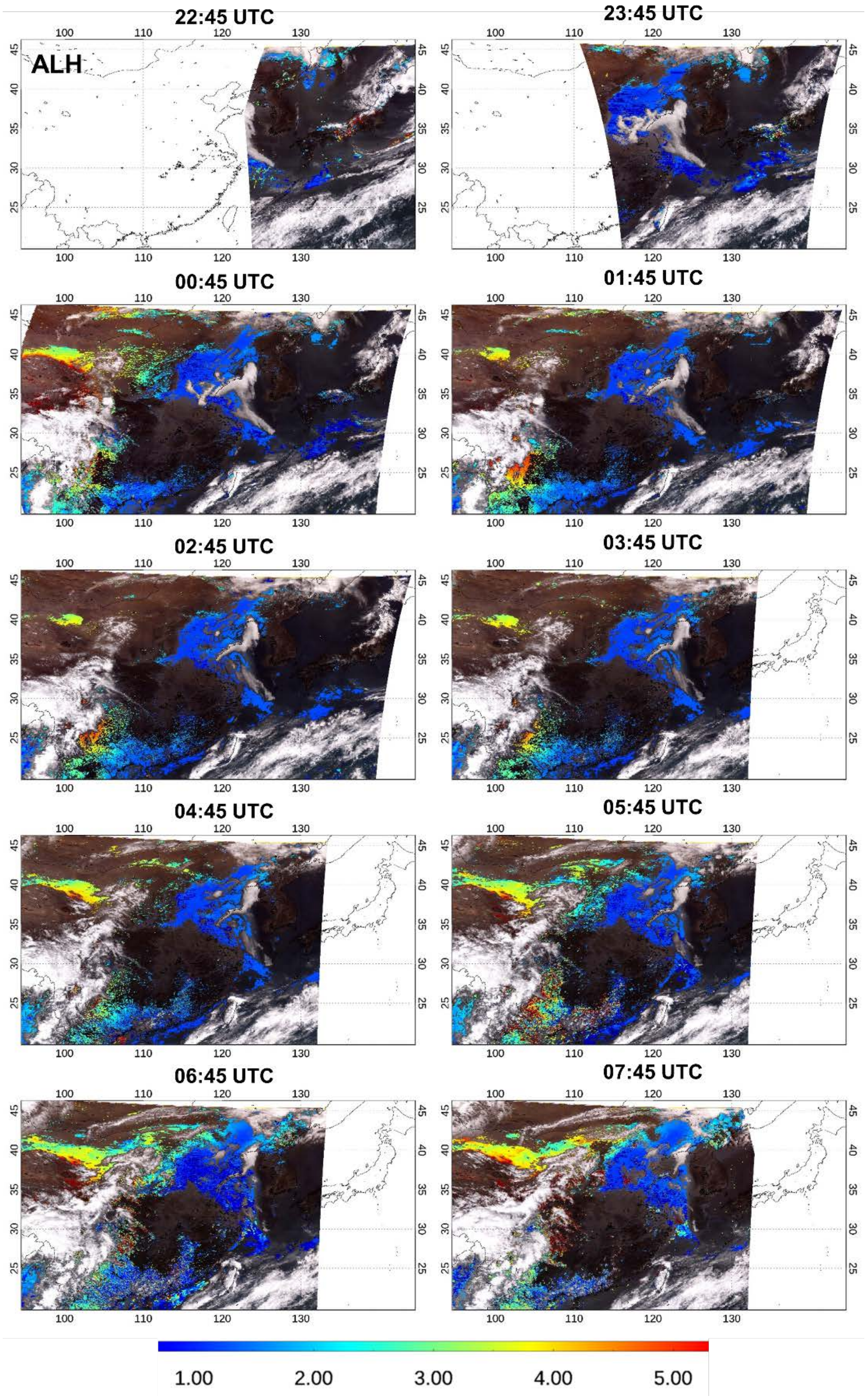

Figure 2: Continue

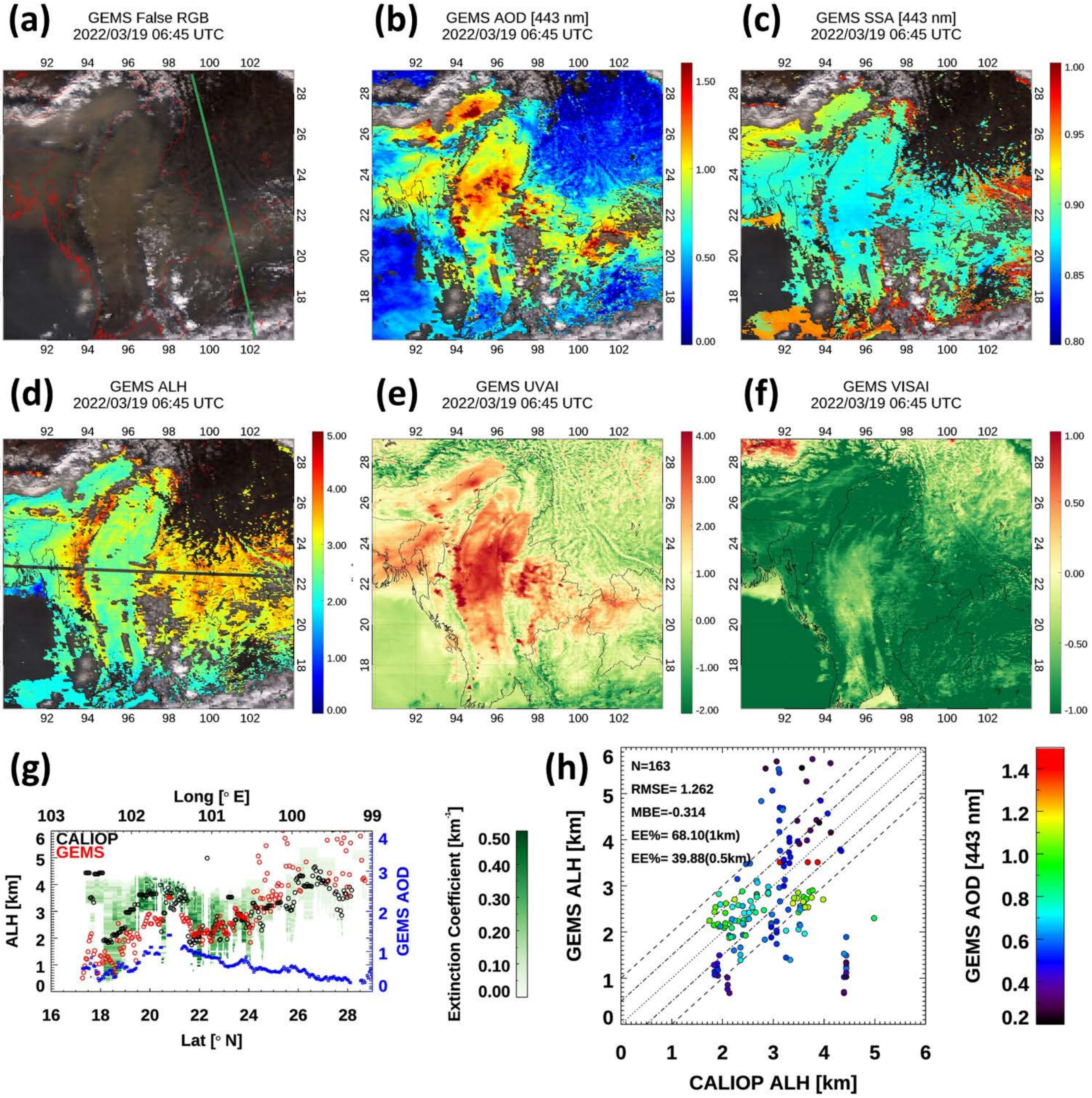

Figure 3: The example of GEMS aerosol products for biomass burning over mainland Southeast Asia. The maps of (a) GEMS False RGB, (b) AOD, (c) SSA, (d) ALH, (e) UVAI, and (f) VisAI. The green line in GEMS False RGB indicates the overpass path of CALIOP. The GEMS SSA and ALH are displayed only when the GEMS AOD is over 0.2. (g) GEMS ALH compared with CALIOP extinction coefficient in the domain. The background color represents the CALIOP extinction coefficient. The black open circles denote the CALIOP ALH, whereas the red open circles represent the GEMS ALH. The blue squares represent the GEMS AOD. (h) Comparison of GEMS and CALIOP ALH when GEMS AOD > 0.2. The dashed and dash-dotted lines indicate an uncertainty envelope of ±1 km and ±0.5 km in ALH, respectively. The dotted lines represent the 1:1 line. The color in the circles represents the GEMS AOD.

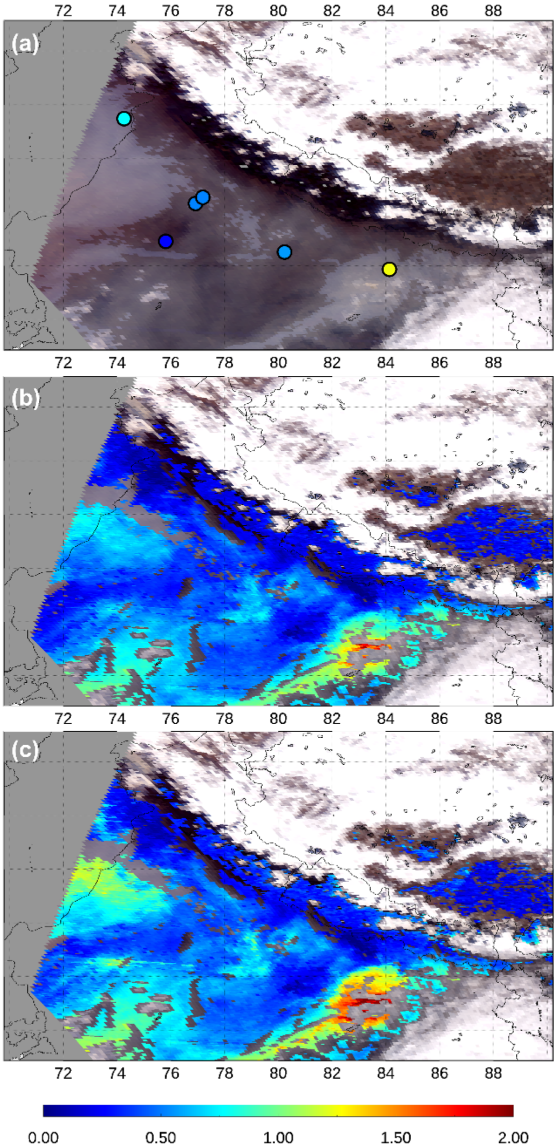

Figure 4: The example of the GEMS AOD before and after post-processing for an absorbing aerosol case over Indo-Gangatic Plane at 04:45 UTC on December 4, 2021. (a) GEMS false RGB. The circle denotes an AERONET station, and the filled color indicates the AERONET AOD at 443 nm, (b) GEMS AOD, and (c) GEMS AOD after post-process correction.

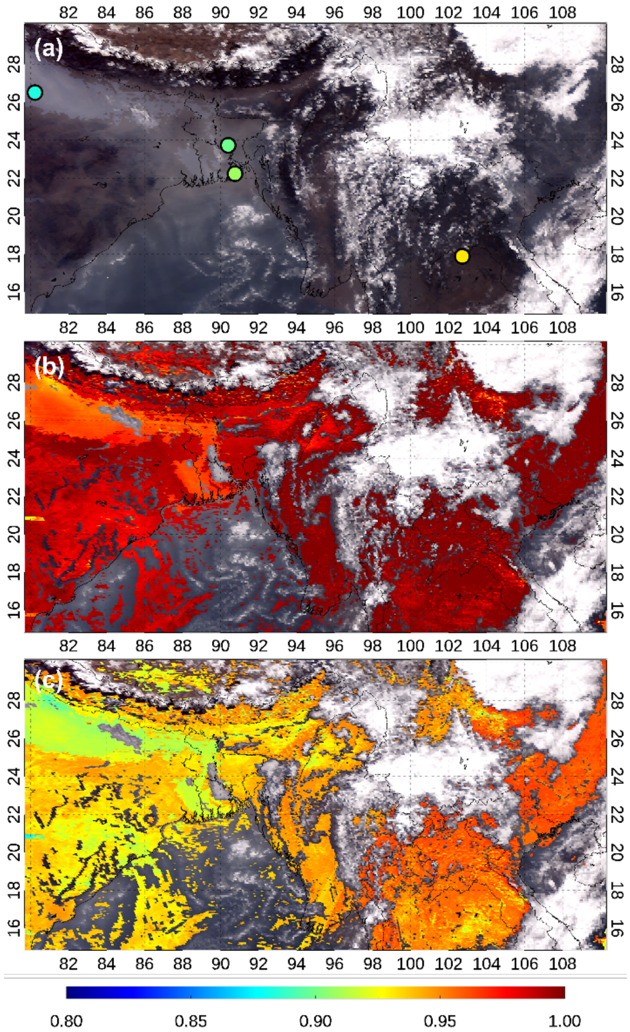

Figure 5: The example of GEMS SSA and the GEMS SSA after post-processing for an absorbing aerosol case over India, Bangladesh, and mainland Southeast Asia at 03:45 UTC on December 23, 2021. (a) GEMS false RGB. The circle denotes an AERONET station, and the filled color indicates the AERONET SSA at 440 nm, (b) GEMS SSA, and (c) GEMS SSA after post-process correction.

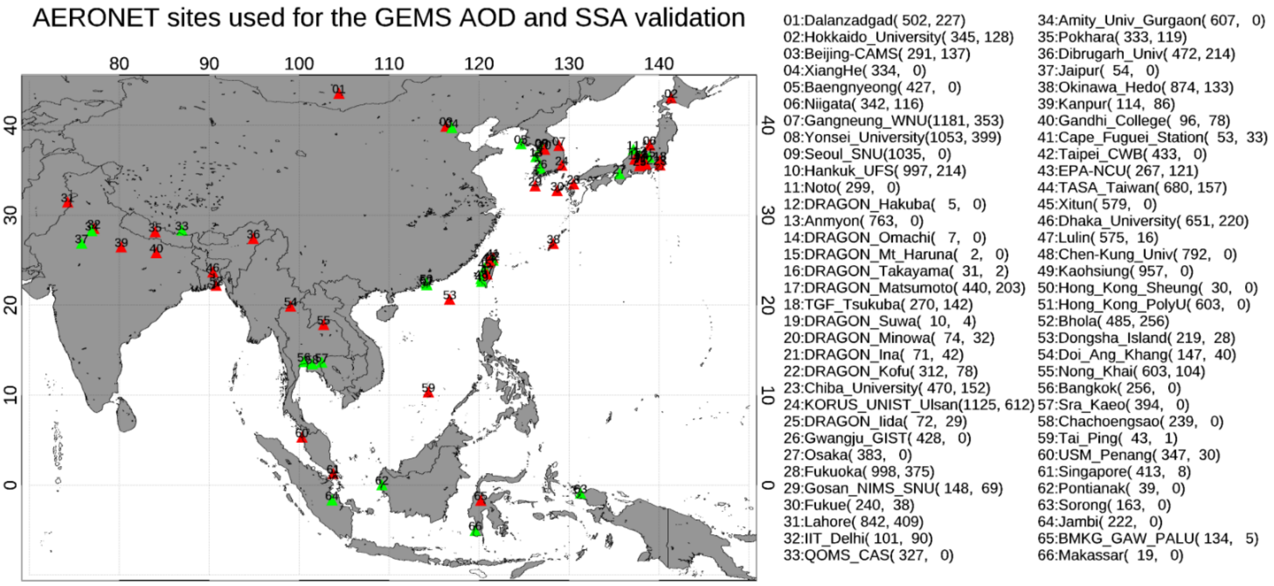

Figure 6: AERONET sites used for the GEMS AOD and SSA validation. The red color indicates the site where validation points exist for both AOD and SSA. The green color indicates the site where validation points exist only for AOD. The list of station names in conjunction with the number of AERONET AOD and SSA data points for validation at each station.

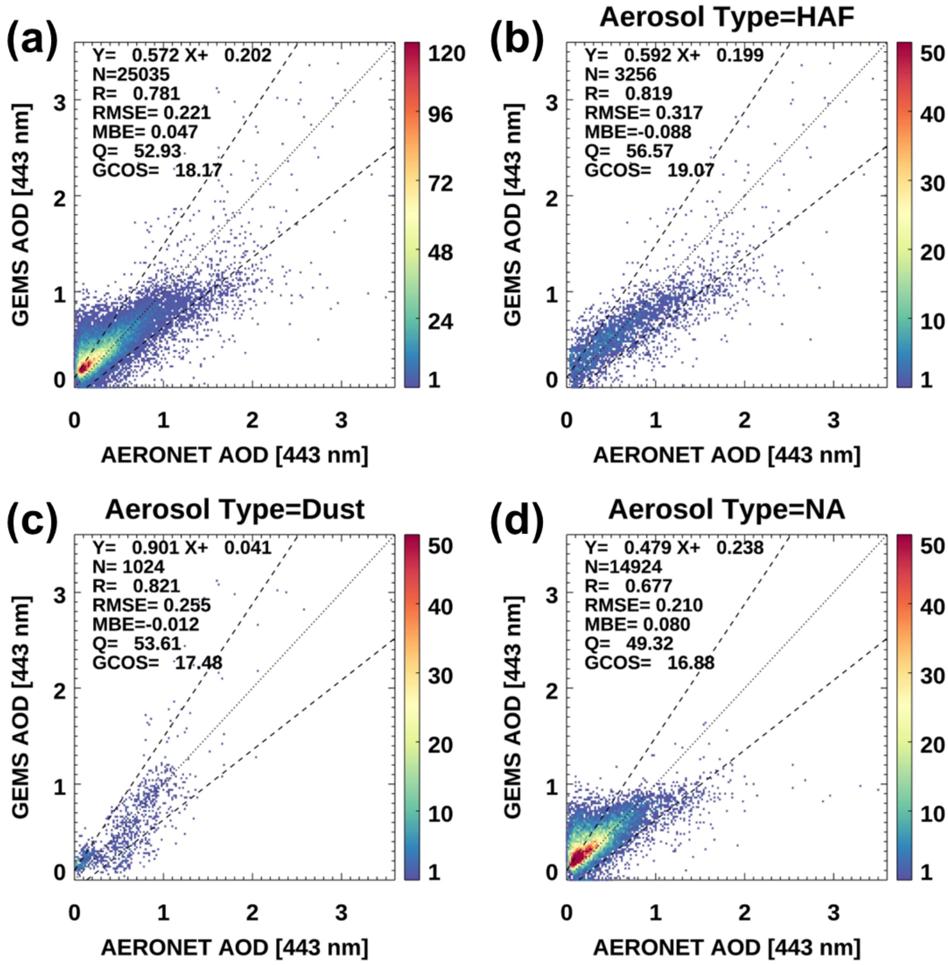

Figure 7: Comparison of GEMS and AERONET AOD for (a) total and individual aerosol types: (b) HAF, (c) dust, and (d) NA. The dashed lines indicate an uncertainty envelope of maximum (0.1 or 30%) in AOD. The dotted lines represent the 1:1 line. Data from November 1, 2021, to October 31, 2022, are used for comparison.

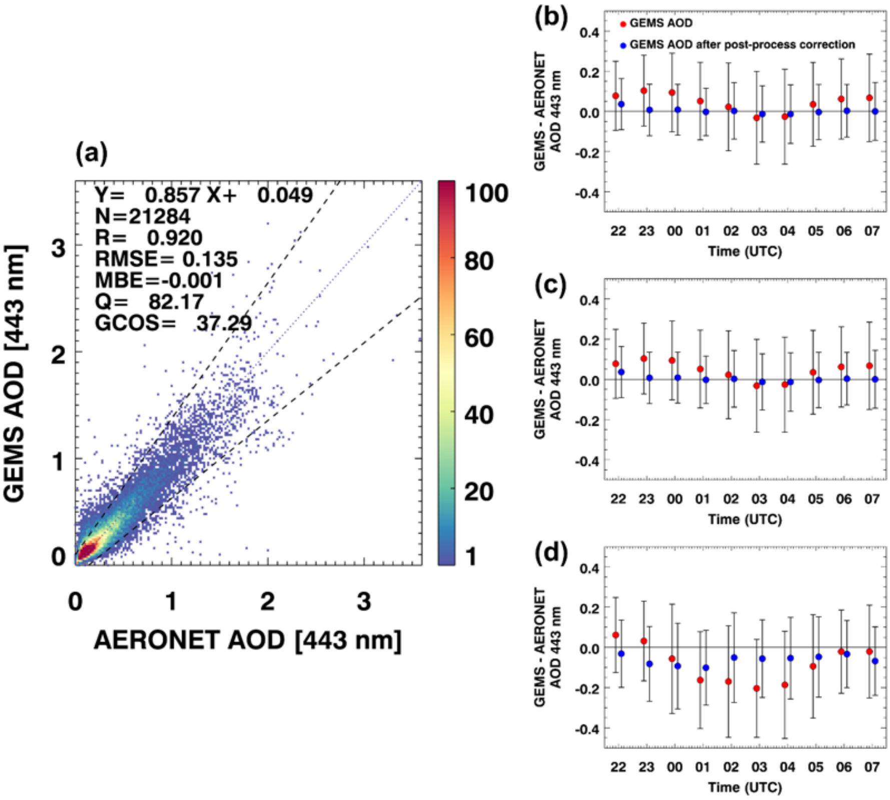

Figure 8: (a) Comparison of GEMS AOD after post-process correction by machine learning and AERONET AOD. The dashed lines indicate an uncertainty envelope of a larger 0.1 or ±30% in AOD. The dotted lines represent the 1:1 line. The difference between GEMS AOD and AERONET AOD in terms of time. (b) All pixels, (c) pixels when AERONET AOD < 0.4, and (d) pixels when AERONET AOD > 0.4. The red circles represent the GEMS AOD, and the blue circles represent the GEMS AOD after post-process correction. The error bars correspond to the standard deviation. Data from November 1, 2021, to October 31, 2022, are used for comparison.

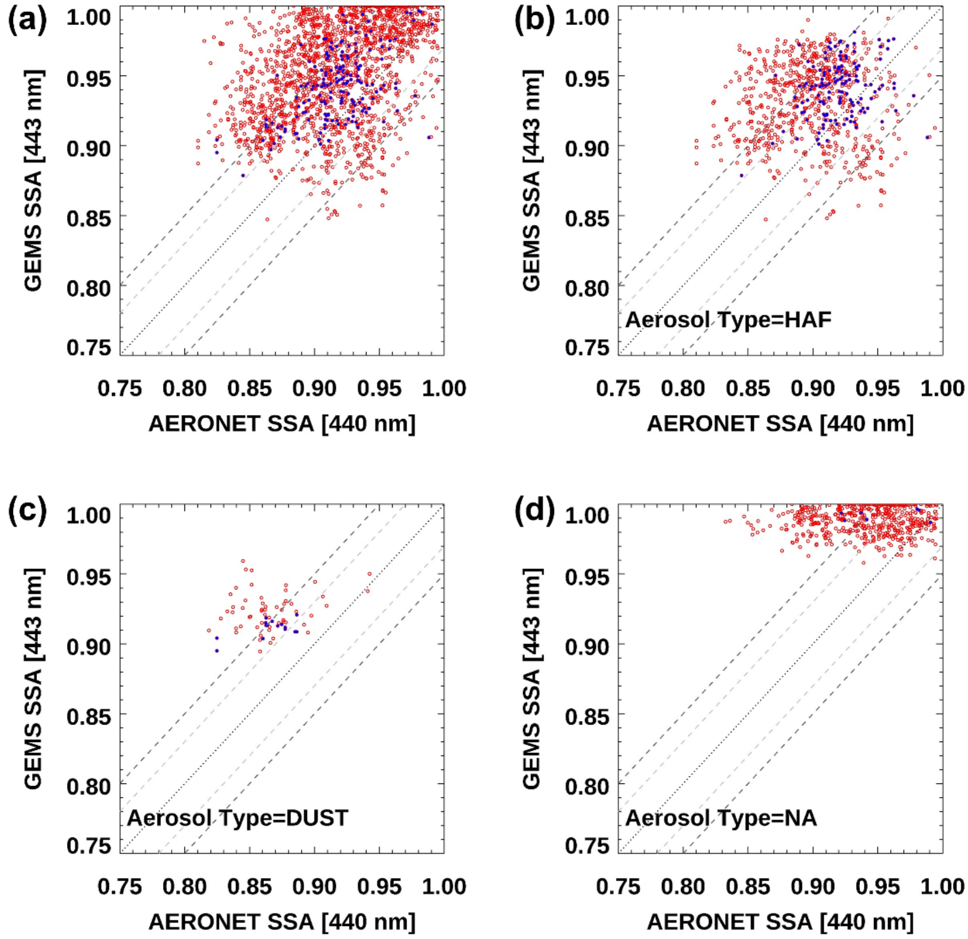

Figure 9: Comparison of GEMS and AERONET SSA for (a) total and individual aerosol types: (b) HAF, (c) dust, and (d) NA. The red circles represent the pixels when AOD > 0.4, and the blue circles represent the pixels when AOD > 1.0. The gray dashed lines indicate an uncertainty envelope of ±0.03 in SSA, the black dashed lines indicate an uncertainty envelope of ±0.05 in SSA, and the dotted lines represent the 1:1 line. Data from November 1, 2021, to October 31, 2022, are used for comparison.

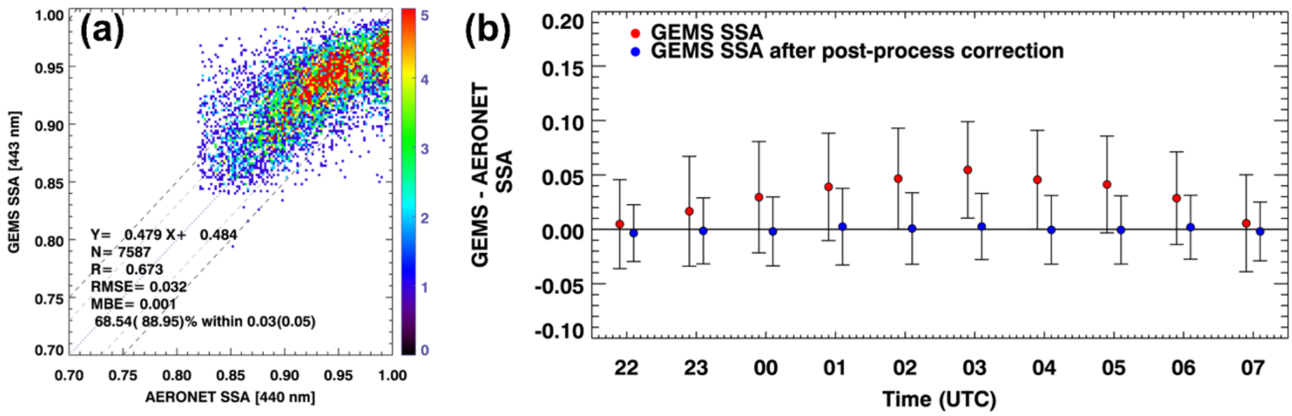

Figure 10: (a) Comparison of GEMS SSA after post-process correction and AERONET SSA. The gray dashed lines indicate an uncertainty envelope of ±0.03 in SSA, the black dashed lines indicate an uncertainty envelope of ±0.05 in SSA, and the dotted lines represent the 1:1 line. (b) The difference between GEMS and AERONET SSA in terms of time. Data from November 1, 2021, to October 31, 2022, are used for comparison.

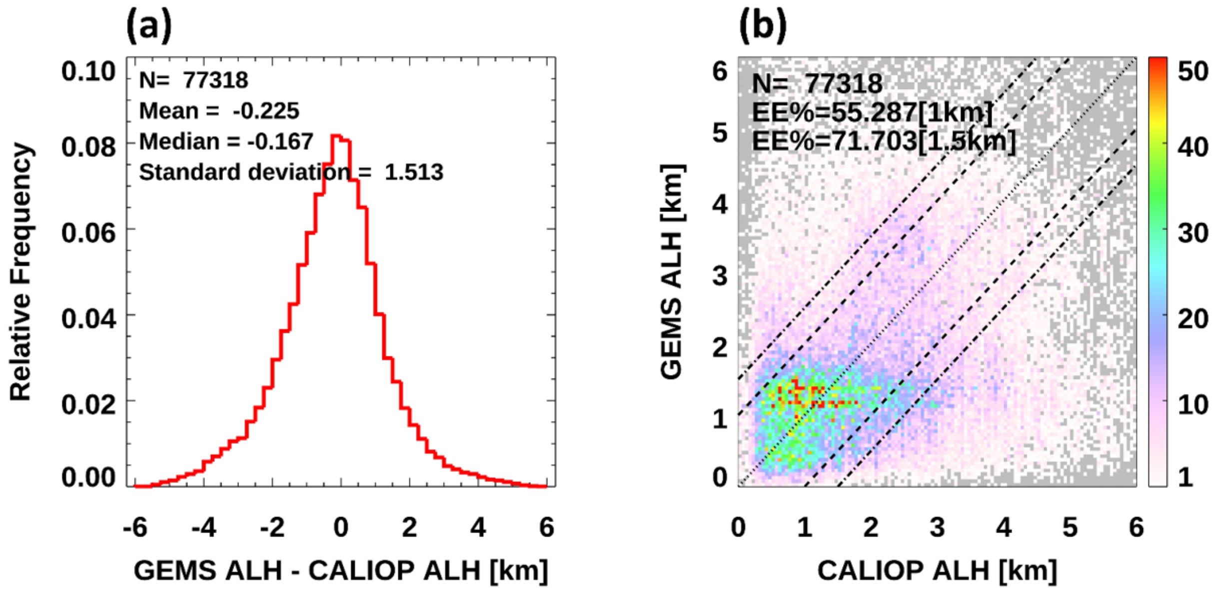

Figure 11: (a) Histogram of difference between GEMS and CALIOP ALH and (b) comparison of GEMS and CALIOP ALH. The dashed lines indicate an uncertainty envelope of ±1 km in ALH. The dash-dotted lines indicate an uncertainty envelope of ±1.5 km in ALH. The dotted lines represent the 1:1 line. Data from November 1, 2021, to October 31, 2022, are used for comparison.

Table 1: Dimension of LUT in GEMS Aerosol algorithm.

Variable Name [Unit]	Number of Entries	Entries
Wavelength [nm]	6	354, 388, 412, 443, 477, 490
SZA [°]	12	0.01, 5, 10, 15, 20, 27, 34, 41, 48, 55, 62, 69
VZA [°]	12	0.01, 5, 10, 15, 20, 27, 34, 41, 48, 55, 62, 69
RAA [°]	11	0.01, 15, 30, 45, 60, 80, 100, 120, 140, 160, 180
Surface reflectance [-]	4	0.0, 0.05, 0.1, 0.2
AOD at 443 nm [-]	8	0.0, 0.1, 0.4, 0.8, 1.5, 2.0, 2.8, 3.6, 5.0, 10.0
SSA at 443 nm [-]	8	1.0, 0.98, 0.96, 0.94, 0.91, 0.88, 0.85, 0.82 for HAF and Dust 1.0, 0.99, 0.98, 0.97, 0.96, 0.94, 0.92, 0.90 for NA
ALH above the surface [km]	5	0.5, 1.5, 3.0, 4.5, 6.0
Elevation [km]	3	0, 3, 6

Table 2: Statistic of hourly comparison of GEMS and AERONET AOD in Figure S5.

Time	N	Slope	y-intercept	R	RMSE	MBE	Q (%)	GCOS (%)
22:45	801	0.725	0.177	0.738	0.181	0.094	60.42	24.97
23:45	1413	0.728	0.193	0.752	0.187	0.115	53.93	19.89
00:45	2879	0.600	0.221	0.698	0.218	0.112	48.32	15.56
01:45	3345	0.490	0.211	0.715	0.209	0.063	52.68	16.95
02:45	3718	0.533	0.193	0.780	0.214	0.039	52.66	17.86
03:45	3504	0.577	0.171	0.830	0.238	-0.011	53.48	16.67
04:45	3556	0.592	0.176	0.824	0.238	-0.001	53.12	17.97
05:45	3186	0.518	0.233	0.725	0.043	0.043	50.00	18.33
06:45	2117	0.606	0.241	0.766	0.239	0.069	52.01	19.79
07:45	1299	0.632	0.227	0.754	0.245	0.063	54.89	19.86

Table 3: Comparison of GEMS and AERONET SSA for different aerosol types in Figure 9. N represents the number of data, and EE% denotes the percentage within the expected error range of ±0.03 (±0.05).

	GEMS AOD > 0.4		GEMS AOD > 1.0	
Aerosol Type	N	EE% ±0.03 (±0.05)	N	EE% ±0.03 (±0.05)
All	1841	34.22(61.38)	174	48.85(84.48)
HAF	764	31.68(62.43)	136	54.41(89.71)
Dust	71	12.68(45.07)	15	13.33(66.67)
NA	536	32.46(56.72)	7	42.86(57.14)

Table 4: Statistic of comparison of GEMS and AERONET SSA in Figure S6.

	GEMS AOD > 0.4		GEMS AOD > 1.0	
Time	N	EE% ±0.03 (±0.05)	N	EE% ±0.03 (±0.05)
22:45	49	67.35(89.80)	13	61.54(92.31)
23:45	76	64.47(82.89)	18	77.78(94.44)
00:45	100	62.00(87.00)	21	90.48(100.00)
01:45	138	57.25(81.16)	29	72.41(96.55)
02:45	190	31.58(56.84)	72	31.94(56.94)
03:45	391	18.67(44.76)	206	15.05(46.60)
04:45	406	22.41(52.46	209	23.44(58.85)
05:45	223	30.49(61.88)	94	28.72(65.96)
06:45	175	37.14(69.71)	83	40.96(75.90)
07:45	93	53.76(73.12)	46	54.35(76.09)