Evaluation and uncertainty estimate of next-generation geostationary meteorological Himawari-8/AHI aerosol products

Jing Wei a, Zhanqing Li b,⁎, Lin Sun c, Yiran Peng d, Zhaoyang Zhang e, Zhengqiang Li f, Tianning Su b, Lan Feng g, Zhaoxin Cai a, Hao Wu h

State Key Laboratory of Remote Sensing Science, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China
b Department of Atmospheric and Oceanic Science, Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA
c College of Geomatics, Shandong University of Science and Technology, Qingdao, China
d Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, China
e College of Geography and Environmental Sciences, Zhejiang Normal University, Jinhua, China
f State Environmental Protection Key Laboratory of Satellite Remote Sensing, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China
g School of Geography and Information Engineering, China University of Geosciences, Wuhan, China

HIGHLIGHTS

• Himawari-8 10-min aerosol products are evaluated over land and ocean.
• AOD retrievals show large uncertainties at the regional scale.
• AE retrievals perform poorly and are always underestimated.
• The Expected Errors for Himawari-8 AOD products are determined.

ABSTRACT

The next-generation geostationary meteorological Himawari-8 satellite carrying the Advanced Himawari Imager (AHI) allows frequent observations of the atmosphere, the surface, and oceans every 10 min. With its retrieval algorithms recently updated, Himawari-8/AHI Version 2 Level 2 aerosol products are now available. However, these retrievals have not yet undergone a quality assessment. This study aims to comprehensively validate the official aerosol optical properties derived from Himawari-8/AHI over land and ocean. Aerosol Robotic Network and Sun-Sky Radiometer Observation Network ground-based measurements at 98 stations in the Himawari-domain region are used to validate aerosol optical depth (AOD, or τ) retrievals at 500 nm and Ångström exponent (AE) retrievals at 440–675 nm from the year 2016. The AOD retrievals agree well with surface observations (i.e., from linear regression, slope = 0.876, intercept = 0.076, and correlation coefficient = 0.756) with a mean absolute error and a root-mean-square error of 0.168 and 0.293, respectively. On site and regional scales, large uncertainties are seen, especially in Australia (significant overestimation) and South Asia (significant underestimation). The AOD retrievals can correctly capture daily variations and show the best (worst) performance in summer (spring). The AE performance is poorer on all scales, showing overall underestimations, especially in Australia, Southeast Asia, and China. The data quality of AOD retrievals improves as the vegetation coverage and...
the AE increases. This suggests that the official aerosol retrieval algorithm still faces great challenges over bright surfaces and under coarse-particle-dominated conditions. In general, approximately 61% and 64% of the AOD matchups meet the newly defined expected errors of $[0.330 \times \tau + 0.024; -0.132 \times \tau - 0.125]$ and $[0.519 \times \tau + 0.005; -0.007 \times \tau - 0.194]$ determined by ground measurements and aerosol retrievals, respectively. The highly variable accuracy of aerosol retrievals raises a concern about the reliability of the current product under different environmental conditions and underlying surfaces. It also sheds light on what future improvements need implementing to the aerosol retrieval algorithm.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

The atmospheric aerosol is a stable mixing system consisting of solid particles and liquid particles dispersed uniformly in the atmosphere. Aerosols affect weather and the climate system through their direct and indirect radiative forcings (Charlson et al., 1992; Feng et al., 2018; Haywood and Boucher, 2000; Pöschl, 2005; Ramanathan and Carmichael, 2017; Wang et al., 2019). Closely associated with air pollution, aerosols also have serious adverse effects on human health, especially fine particulate matter with aerodynamic diameters ≤ 2.5 (PM$_{2.5}$) and 1 (PM$_1$) microns (Gupta et al., 2006; Sun et al., 2016a, 2016b; Viana et al., 2014). PM$_{2.5}$ and PM$_1$ can do much damage to human health (Bartell et al., 2013; Crouse et al., 2012; Peng et al., 2009; Pöschl, 2005). Aerosol optical depth (AOD) and the Ångström exponent (AE) are two fundamental aerosol optical properties used to quantify the columnar amount and size of aerosol particles. Based on atmospheric physics theories, these parameters can be extracted from remotely sensed images over wide spatial scales and long timescales (Kaufman et al., 1997). With the development of satellite remote sensing techniques, their spatial, spectral, and temporal resolutions have improved continuously.

Himawari-8, successfully launched on 7 October 2014 by the Japan Meteorological Agency (JMA), is the next-generation geostationary meteorological satellite carrying the Advanced Himawari Imager (AHI). Himawari-8/AHI data are used in meteorological applications and became freely available to the public on 7 July 2015. It has similar spectral and spatial characteristics as those of the Advanced Baseline Imager planned for the U.S. Geostationary Operational Environmental Satellites-R series of satellites. It acquires data at 16 independent spectral channels from visible to infrared wavelengths covering the main Asia-Pacific region (80°E–160°W, 60°S–60°N, Fig. 1). Images are recorded every 10 min at a spatial resolution of 0.5–1 km in the visible channels and 1–2 km in the infrared channels (Table 1). Himawari-8/AHI has produced various operational products, including atmospheric motion vectors, clear-sky radiative fluxes, and high-resolution cloud analysis information. The high spatial resolution and frequent imaging time allow more accurate and detailed observations of atmospheric, surface, and ocean targets.

Like instruments on other Sun-synchronous satellites, e.g., the Advanced Very High Resolution Radiometer (Che et al., 2018), the MODerate-resolution Imaging Spectrometer (MODIS, Levy et al., 2013; Wei et al., 2019a, 2019b, 2019c), and the Visible Infrared Imaging

Fig. 1. Full scene of a next-generation geostationary meteorological Himawari-8/AHI image where red and purple dots show the locations of Aerosol Robotic Network and Sun-Sky Radiometer Observation Network sites, respectively. The background map shows MODIS land cover data from 2016. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Radiometer Suite (Jackson et al., 2013), Himawari-8/AHI also provides aerosol products at a spatial resolution of 5 km with four temporal resolutions: 10 min (Level 2), 1 h (Level 3), 1 day (Level 3), and 1 month (Level 3). Multi-source satellite-derived aerosol products have been widely used to estimate near-surface particulate matter concentrations and monitor air quality (Chen et al., 2018; Lv et al., 2017; Ma et al., 2014; Yao et al., 2018; Wei et al., 2019d). Wang et al. (2017) developed an improved linear mixed-effect model to monitor surface hourly PM$_{2.5}$ concentrations from Himawari-8/AHI AOD products over the Beijing–Tianjin–Hebei (BTH) region in China. Zang et al. (2018) estimated hourly PM$_{10}$ concentrations from Himawari-8/AHI AOD products matched with coincident ground-based PM$_{10}$ measurements using a principal component analysis-general regression neural network method they developed for application in China.

These studies have presented preliminary and limited evaluations of the official aerosol products. The reliability of the aerosol products is closely related to its applications in these atmospheric pollution studies. With continuous improvements made to the aerosol algorithm, the data version was updated from Version 1 to Version 2 on 10 August 2018. The official Himawari-8 aerosol products, however, have not yet been comprehensively evaluated. This study thus aims to provide a thorough evaluation of the full-scene Himawari-8 aerosol products over land and ocean.

For this purpose, the latest Version 2 Level 2 10-min aerosol datasets, including AOD retrievals at 500 nm and the AE at 440–675 nm, at a spatial resolution of 5 km from 2016 are collected and validated against the newest releases of Aerosol Robotic Network (AERONET) and Chinese Sun–Sky Radiometer Observation Network (SONET) ground-based aerosol optical property measurements. The systematic validation of the official JMA aerosol product is expected to be an important reference for its future applications in a wide range of studies.

Section 2 provides descriptions of the Himawari-8/AHI aerosol products and ground measurements. Section 3 validates the aerosol retrievals at Himawari-domain, regional, and individual-site scales and their temporal performances at the diurnal, daily, and seasonal levels. Section 4 analyzes and discusses the retrieval errors due to different underlying surfaces and aerosol particle sizes and calculates their expected errors. Section 5 provides a summary and conclusions.

2. Datasets and method

2.1. Himawari-8/AHI aerosol products

The Himawari-8/AHI aerosol products (including AOD and AE) are generated using the newly developed aerosol retrieval algorithm based on the Lambertian assumption over land and ocean (Fukuda et al., 2013; Higurashi and Nakajima, 1999; Yoshida et al., 2018). This algorithm introduces the weight to the objective function for each channel then automatically selects the optimum channels for aerosol retrievals. Surface reflectances are estimated based on second-minimum reflectance synthetic technology within one month where both the sensor (solar) geometry is almost identical for each hour (month) to reduce the effects of cloud shadows and surface bidirectional characteristics. Common fine (including rural, sea-spray, and yellow sand), coarse marine, and coarse dust aerosol models with monomodal lognormal volume size distributions are assumed through a cluster analysis with AERONET measurements (Omar et al., 2005).

Cloud pixels are screened using the previously developed common cloud detection algorithm (Ishida and Nakajima, 2009; Ishida et al., 2011). The gas absorption for ozone and water vapor are corrected based on MODIS empirical approximations using the total ozone and water vapor columns from the Ozone Monitoring Instrument and JMA global analysis data (Levy et al., 2007). Lastly, the aerosol retrieval is performed based on the look-up-table approach using the University of Tokyo radiative transfer code (Nakajima and Tanaka, 1986, 1988; Stamnes et al., 1988).

The Himawari-8/AHI Version 2 Level 2 aerosol dataset includes a quality assurance (QA) flag with several confidence tests concerning land cover type (i.e., land, water, snow/ice), clouds, sunglint, high solar and satellite zenith angles, surface reflectance, AOD, and AE [https://www.eorc.jaxa.jp/ptree/documents/Himawari_Monitor_Aerosol_Product_v5.pdf]. Four levels of confidence are assigned to AOD and AE retrievals: (1) very good, (2) good, (3) marginal, and (4) no confidence (or no retrieval). This study uses only those retrievals passing the recommended QA with low uncertainties for validation purposes.

2.2. Ground-based measurements

2.2.1. AERONET ground measurements

AERONET provides freely available spectral-independent AOD measurements at numerous globally distributed monitoring sites every 15 min with a low uncertainty of approximately 1–2% under cloud-free conditions (Holben et al., 2001; Smirnov et al., 2000). The newly updated AERONET Version 3 Level 2.0 (cloud screened and quality assured) aerosol property measurements, which have undergone further cloud screening and quality control, are selected for validation purposes (Giles et al., 2019; Wei et al., 2019a, 2019b, 2019e). The study uses measurements from 84 AERONET sites located within the area covered by Himawari-8 images.

2.2.2. SONET ground measurements

SONET is a ground-based Cimel radiometer network that provides long-term columnar atmospheric aerosol properties at multiple wave-lengths in China. The sun-sky radiometer CE318-DP manufactured by Cimel Électronique in France provides the measurements used to do the retrievals. The cloud screening and QA for SONET AOD-related products mainly follow AERONET protocols. The data records are reported every 15 min with a high precision of 0.25–0.5% at visible and near-infrared wavelengths (Li et al., 2018). The study uses SONET Level 1.5 (cloud screened) AOD measurements at 14 long-term sites (purple dots in Fig. 1) for validation purposes.

2.3. Data processing and validation approaches

This study uses Himawari-8/AHI AOD retrievals at 500 nm and AE retrievals at 440–675 nm with recommended QA at a 5-km resolution from 2016. The average value within a sampling window of 3 × 3 pixels (with at least 3 out of 9 pixels available) centered on a ground monitoring site defines the retrievals. The average of at least two ground AOD measurements within 5 min of the Himawari-8/AHI overpass time is taken as the ground truth AOD (Sun et al., 2016c; Wei et al., 2018a, 2019a, 2019b). Statistical tools and techniques used to quantify the accuracy and uncertainty of aerosol retrievals against ground measurements include the Pearson product-moment correlation coefficient (R), the mean absolute error (MAE), the root-mean-square error (RMSE), and the relative mean bias (RMB). RMSE values > 1.0 or < 1.0 represent the overestimation or underestimation uncertainty, respectively (Wei and Sun, 2017; Wei et al., 2017, 2019d, 2019f).
3. Validation against surface measurements

3.1. Spatial validation results

3.1.1. Himawari-domain-scale performance

Fig. 2 presents density scatterplots of Himawari-8/AHI AOD and AE retrievals against ground measurements at all available sites on land and ocean in 2016. A total of 141,071 AOD retrievals were collected from 98 sites (including 84 AERONET and 14 SONET sites). They agree well with ground measurements ($R = 0.756$, Fig. 2a). The matched data points are evenly distributed on both sides of the 1:1 line with a slope of 0.876 and a $y$-intercept of 0.076. The MAE and RMSE values are 0.168 and 0.293, respectively, and the largest concentration of data points are below AOD = 0.8. However, aerosol loading is generally overestimated ($RMB = 1.403$) on the Himawari-domain scale. A large number of matched data points are found near the ordinate, suggesting contamination by thin and fragmentary clouds in the geostationary remote sensing images. The number of matched AE data points collected from 84 AERONET sites is 121,156. They are poorly related to AERONET AE measurements ($R = 0.191$) with a slope and a $y$-intercept of 0.243 and 0.803, respectively (Fig. 2b). The range from 1.1 to 1.7 contains the bulk of AE retrievals. In general, AE retrievals have large estimation uncertainties with MAE and RMSE values of 0.403 and 0.511, respectively. On the Himawari-domain scale, AE is underestimated.

3.1.2. Regional-scale performance

Examined next is the performance of the next-generation geostationary meteorological Himawari-8/AHI aerosol products at national/regional scales. The regions include the BTH and Yangtze River Delta (YRD) regions of China, Australia, Japan, Korea, South Asia (SAA), and Southeast Asia (SEA) (Table 2). Korea shows the best performance in terms of aerosol loading with a mean AOD (0.399) close to the ground truth (0.368) and the best evaluation metrics ($R = 0.737$, $MAE = 0.131$, $RMSE = 0.196$). A similar performance ($R = 0.773$, $MAE = 0.131$, $RMSE = 0.196$) but with larger overestimations ($RMB = 1.447$) is found in Japan. By contrast, Australia and South Asia have poor agreements between retrieved and measured AODs and large estimation uncertainties. In particular, aerosol loadings are significantly overestimated in Australia ($RMB = 2.065$) but significantly underestimated in South Asia ($R = 0.767$). In China, the AOD retrievals agree well with ground measurements ($R = 0.750$) with an overall MAE of 0.207 and RMSE of 0.386, and with -17% overestimations.

Table 2

<table>
<thead>
<tr>
<th>Region</th>
<th>Property</th>
<th>Number</th>
<th>True Retrieval</th>
<th>R</th>
<th>MAE</th>
<th>RMSE</th>
<th>RMB</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>AOD</td>
<td>44,556</td>
<td>0.420</td>
<td>0.492</td>
<td>0.750</td>
<td>0.207</td>
<td>0.386</td>
</tr>
<tr>
<td>BTH</td>
<td>AOD</td>
<td>14,538</td>
<td>0.525</td>
<td>0.614</td>
<td>0.865</td>
<td>0.203</td>
<td>0.357</td>
</tr>
<tr>
<td>YRD</td>
<td>AOD</td>
<td>14,466</td>
<td>1.139</td>
<td>1.049</td>
<td>0.219</td>
<td>0.423</td>
<td>0.521</td>
</tr>
<tr>
<td>AE</td>
<td></td>
<td>3111</td>
<td>1.219</td>
<td>1.251</td>
<td>0.241</td>
<td>0.308</td>
<td>0.930</td>
</tr>
<tr>
<td>Australia</td>
<td>AOD</td>
<td>13,552</td>
<td>0.069</td>
<td>0.143</td>
<td>0.508</td>
<td>0.085</td>
<td>0.156</td>
</tr>
<tr>
<td>Japan</td>
<td>AOD</td>
<td>14,520</td>
<td>0.217</td>
<td>0.314</td>
<td>0.737</td>
<td>0.131</td>
<td>0.196</td>
</tr>
<tr>
<td>Korea</td>
<td>AOD</td>
<td>37,194</td>
<td>0.368</td>
<td>0.399</td>
<td>0.828</td>
<td>0.132</td>
<td>0.190</td>
</tr>
<tr>
<td>AE</td>
<td></td>
<td>37,179</td>
<td>1.307</td>
<td>1.147</td>
<td>0.290</td>
<td>0.342</td>
<td>0.438</td>
</tr>
<tr>
<td>SEA</td>
<td>AOD</td>
<td>21,493</td>
<td>0.527</td>
<td>0.433</td>
<td>0.803</td>
<td>0.194</td>
<td>0.283</td>
</tr>
<tr>
<td>SAA</td>
<td>AOD</td>
<td>3111</td>
<td>1.258</td>
<td>1.181</td>
<td>0.422</td>
<td>0.304</td>
<td>0.378</td>
</tr>
</tbody>
</table>

Mean ground-truth AE values range from 1.139 to 1.470 over land where China’s BTH region and Southeast Asia have the minimum and maximum values, respectively. However, the satellite-retrieved AEs have a low correlation ($R < 0.5$) with surface measurements, with large mean AE differences and estimation uncertainties ($MAE = 0.31$, $RMSE > 0.39$) in most regions, e.g., Australia, Southeast Asia, and China. The AE retrievals are overall underestimated to varying degrees in all regions except the YRD region. In general, the retrievals show the best performance with the highest correlation of 0.422 and smallest estimation uncertainty ($MAE = 0.304$, $RMSE = 0.378$, $RMB = 0.939$) in South Asia. These results illustrate the great challenge in estimating AE at the regional scale.

3.1.3. Individual-site-scale performance

Aggregated global/regional analyses mainly provide the overall performance of aerosol products. Due to the inhomogeneity of ground-
based sites around the world and different data collections among the sites, additional site-scale validation is performed to avoid uneven-weighting issues in combined metrics (Wei et al., 2019a). Fig. 3 presents the validation of Himawari-8/AHI AOD retrievals against ground measurements for each AERONET and SONET site in 2016. Only data from sites with at least 20 matchups are used to ensure statistical significance. AOD retrievals agree well with surface observations with correlations > 0.7 at 63 out of 94 selected sites. These 63 sites are mainly located in South Korea, Japan, eastern and southern China, and some countries in Southeast Asia. Sites with low correlations are in Australia, South Asia, other Southeast Asian countries, and northwest China. MAE and RMSE have similar spatial distribution patterns where the combination of the two metrics represents the estimation uncertainty. Small estimation uncertainties are found at most sites. MAE and RMSE values are < 0.2 and 0.3 at approximately 72% and 71% of the sites, respectively. Sites in South Asia, and northwest and central China have large estimation uncertainties. Furthermore, retrievals are overestimated at 61% of the sites, especially in Australia, central-southeast Asia, and northwest and southwest China, with large RMB values > 1.4. By contrast, in central China, Southeast Asia, and South Asia, retrievals are underestimated.

Fig. 4 illustrates the validation of Himawari-8/AHI AE retrievals against AERONET AE measurements for each site in 2016. Retrievals of 440–675-nm AE are not made at SONET sites. Data from sites with at least 20 matchups are used here. AE retrievals do not agree well with ground measurements at most sites with > 61% of the sites showing low, even negative, correlation coefficients below 0.3. Some individual sites in South Korea and Japan have relatively better correlations. MAE and RMSE have similar spatial patterns with approximately 62% and 61% sites having MAE and RMSE values > 0.35 and 0.45, respectively. By contrast, some sites in eastern China and South Korea have small estimation uncertainties. Furthermore, retrievals are underestimated at ~82% of the sites, especially in South Korea and Southeast Asia.

3.1.4. Zonal-scale performance

Zonally averaged Himawari-8/AHI AOD and AE retrievals in different latitudinal bands from 60°S to 60°N over both land and ocean in 2016 are examined here to help study the radiative effects of aerosols on Earth’s climate (Fig. 5). Due to few observation sites in the Southern Hemisphere, there are relatively smaller numbers of AOD and AE matchups than in the Northern Hemisphere. The AOD retrievals show good performance with small MAE and RMSE values in the Southern
Hemisphere and at middle-high latitudes (> 30°) in the Northern Hemisphere. However, aerosol loadings are highly overestimated. Between 20°N and 30°N, estimate uncertainties (i.e., MAE and RMSE) are large, and aerosol loadings are always underestimated. The AE performance is relatively poor in the Southern Hemisphere and areas above 40°N. The estimate uncertainties (i.e., MAE, RMSE, and RMB) between 20°N and 30°N are generally larger than those in other regions.

Fig. 4. Same as Fig. 3 but for Himawari-8/AHI AE retrievals at 440–675 nm.

Fig. 5. Validation of Himawari-8/AHI (a) AOD at 500 nm and (b) 440–675-nm AE retrievals against ground measurements in different latitudinal bands.
uncertainties are generally low with MAE and RMSE values agree well with surface measurements. In general, daily mean AOD retrievals abnormally high values, the aerosol loading remains at a relatively low that frequently occur in spring in Asia may explain this. Despite some retrievals until around day 125 of the year. Severe dust and haze episodes at the beginning of the year, then increase with large estimation uncertainties (MAE = 0.167, RMSE = 0.314) and winter is 44,086, 29,387, 21,731, and 25,952, respectively. In all sea-

3.2. Temporal validation results

3.2.1. Validation at the diurnal level

The high temporal resolution of the Himawari-8/AHI sensor allows for analysis of AOD and AE retrievals on a diurnal scale. Fig. 6 shows the diurnal variations of Himawari-8/AHI 500-nm AOD and 440–675-nm AE retrievals against ground measurements. The effective aerosol observation time from the Himawari-8 satellite ranges from 0400 to 1800 Beijing Time (BT). Aerosol loadings are generally low in the morning and gradually increase, reaching a maximum value around 1100 BT then decreasing gradually. In general, hourly Himawari-8/AHI and ground-based AOD retrievals do not have the same distributions. The satellite retrieval has a bimodal distribution with maximum values seen around 0700 BT and 1300 BT. Ground-based AE retrievals gradually increase over time in a day. The reduction in coarse particulate matter with larger particle sizes through sedimentation over time may explain this. At the diurnal level, satellite- and ground-based AE retrievals are poorly correlated (R = 0.14). The satellite-based AE retrievals have a decreasing trend with increasing negative biases and estimation errors over time in a day.

3.2.2. Validation at the daily level

Fig. 7 shows the time series of daily-averaged Himawari-8/AHI AOD and AE retrievals and ground measurements. Aerosol loadings are low at the beginning of the year, then increase with large standard deviations until around day 125 of the year. Severe dust and haze episodes that frequently occur in spring in Asia may explain this. Despite some abnormally high values, the aerosol loading remains at a relatively low mean level for the rest of the year. In general, daily mean AOD retrievals agree well with surface measurements (R = 0.92), and the estimation uncertainties are generally low with MAE and RMSE values < 0.2 and 0.3 on 297 and 296 days of the year, respectively. Underestimated daily aerosol loading estimates with negative biases occur during the first third of the year. Overestimations with small positive biases occur during the rest of the year. These results illustrate that Himawari-8/AHI AOD retrievals can more accurately capture the daily aerosol variations in the year.

AERONET AE measurements are relatively small during the first 125 days, mainly because of the larger-sized particles present in the springtime when dust and haze episodes are frequent. Daily mean AE values increase during the rest of the year, indicating decreases in particle size. In general, the satellite- and ground-based AE measurements have similar daily variations during the year. However, the retrievals are poorly correlated (R = 0.27). Underestimated daily AE estimates with negative biases occur on most days of the year. These results illustrate that Himawari-8/AHI AE retrievals perform poorly on a daily level.

3.2.3. Validation at the seasonal level

Fig. 8 shows density scatterplots of the Himawari-8/AHI AOD retrievals against ground measurements for each season at all available sites. The number of matched AOD retrievals for spring, summer, autumn, and winter is 50,352, 35,014, 24,419, and 31,286, respectively. The Himawari-8/AHI AOD product performs worst in spring. The number of matched AE retrievals for spring, summer, autumn, and winter is 44,086, 29,387, 21,731, and 25,952, respectively. In all sea-

![Fig. 6](image-url) Diurnal variations of Himawari-8/AHI (a) 500-nm AOD and (b) 440–675-nm AE retrievals. Ground-based measurements are shown in red. The shaded parts in each panel represent the standard deviations of the data samples. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
4. Discussion and uncertainty analysis

4.1. Retrieval error associated with vegetation conditions

The normalized difference vegetation index (NDVI) describes the greenness of a target surface. Fig. 9 shows the bias boxplots and performance of Himawari-8/AHI 500-nm AOD retrievals as a function of NDVI over land. For sparsely vegetated areas (NDVI ≤ 0.2), the AOD retrievals agree poorly with ground-based measurements with large positive biases and large estimation uncertainties (MAE = 0.18, RMSE = 0.33). Negative correlations arise when NDVI ≤ 0.1. However, as NDVI increases, the performance of the Himawari-8/AHI aerosol algorithm improves with greater correlations, decreasing biases, and reduced MAE and RMSE values for moderately vegetated surfaces (0.2 < NDVI < 0.6). For densely vegetated surfaces (NDVI ≥ 0.6), the AOD retrievals agree well with surface-based measurements (R > 0.8) with small biases close to 0 and small standard deviations. The MAE and RMSE values are generally < 0.13 and 0.19, respectively. In general, as the surface vegetation coverage increases, the performance of the aerosol retrieval algorithm becomes more robust with decreasing estimation uncertainties.

4.2. Retrieval error associated with land use type

Fig. 10 shows density scatterplots of Himawari-8/AHI 500-nm AOD retrievals against ground measurements over six main land-use types (Wei et al., 2019a, 2019b). The Himawari-8/AHI aerosol algorithm performs well over water bodies with a high correlation coefficient of 0.828 and a highly linear regression relationship. Overall, the estimation uncertainties are small with low MAE and RMSE values of 0.146 and 0.205, but there are large aerosol overestimations (RMB = 1.451). Of the two typical bright surfaces, the aerosol algorithm performs the worse over bare land (R = 0.469, MAE = 0.205, and RMSE = 0.394) with the smallest sample size (N = 1130) and the highest overestimations (RMB = 3.032). The performance improves over urban surfaces (R = 0.819, MAE = 0.171, and RMSE = 0.260) but with ~24% overestimations (RMB = 1.243). The main reason is that the complex surfaces and intensive human activities increase the difficulties in estimating the surface reflectance and in assuming the aerosol type, reducing the quality of the aerosol retrieval (Wei et al., 2018b, 2019a, 2019e). Of the three vegetation types, the aerosol algorithm performs best over forests with almost all the highest evaluation metrics (i.e., R = 0.866, MAE = 0.105, and RMSE = 0.170), mainly because of the perennially high vegetation coverage. Retrievals over cropland have the largest estimation uncertainties. Seasonal variations in vegetation coverage, especially at the beginning or end of the vegetation growth period, may explain this. The performance over grassland is slightly worse than that over forests, but are much better than that over cropland.

4.3. Retrieval error associated with aerosol particle size

Fig. 11 shows the bias boxplots and performance of Himawari-8/AHI 500-nm AOD retrievals as a function of AERONET 440–675-nm AE retrievals. Table 3 summarizes the statistics. Three aerosol particle size scenarios are considered: (1) AE ≤ 0.8: Aerosols with large particle sizes mainly dominated by mineral dust; (2) 0.8 < AE ≤ 1.4: Aerosols with medium particle sizes considered a mixture of fine-mode and
coarse-mode aerosols; and (3) AE > 1.4: aerosols with small particle sizes likely dominated by small particles, e.g., smoke and industrial pollution (Huang et al., 2016).

In general, the performance of the aerosol retrieval algorithm improves as the aerosol particle size decreases. For dust-dominated aerosols, the Himawari-8/AHI AOD retrievals are poorly related to the surface-based measurements with large positive mean biases ranging from 0.13 to 0.29, standard deviations (STD) > 0.28, large estimation uncertainties (MAE > 0.18 and RMSE > 0.33), and significant overestimations for > 22% of the retrievals (RMB > 1.223). For mixed aerosols, the performance of the AOD retrieval algorithm gradually improves as AE increases with reduced mean biases closer to zero and decreasing
MAE and RMSE values. Overestimations are also significantly reduced to < 20% of the retrievals (RMB = 1.00–1.17). For fine-mode-dominated aerosols, the retrievals have small mean biases, which tend to be negative. The mean biases become positive as the aerosol particle size continues to decrease.

The number of matched retrievals for the three particle-size cases is 6359, 63,040, and 51,579 (Table 3). AOD retrievals for large-sized aerosol particles have a relatively large AOD-independent mean bias of 0.162 with a large MAE of 0.212 and RMSE of 0.402. Approximately 74% of the retrievals are significantly overestimated (RMB = 1.742). AOD retrievals for medium-sized aerosol particles have a greatly reduced mean bias of 0.033, an MAE of 0.163, and an RMSE of 0.263. Only ~8% of the retrievals are overestimated (RMB = 1.084). AOD retrievals for small-sized aerosol particles have the
The smallest negative mean bias of magnitude 0.015 and a standard deviation of 0.145. The MAE and RMSE values are 0.145 and 0.227, respectively.

4.4. Estimation of expected errors

The expected error (EE) is a widely used evaluation metric for describing the overall performance of, for example, satellite-derived aerosol retrievals against ground measurements and differs among different satellites. The objective of this section is to explore and estimate the accepted EE definition for Himawari-8/AHI AOD products. The EE for MODIS Dark Target (DT) aerosol products, based on the assumption that the AOD bias has a Gaussian distribution, is the envelope containing 68% of the data points that have the least bias between the DT retrievals and matched AERONET AOD measurements (Hsu et al., 2013; Levy et al., 2010). However, for MODIS Deep Blue aerosol products, EEs are defined using a similar approach but as a function of the satellite retrievals (Sayer et al., 2013). The following two approaches are thus selected to define EEs for the Himawari-8/AHI aerosol products: (1) expected accuracy (EA) and expected precision (EP), functions of the ground measurements, where EA represents the regressed mean bias line, and EP represents the regressed expected precision line using the one standard deviation of biases; and (2) EA and EP that are functions of the Himawari-8/AHI AOD retrievals.

Fig. 12a shows boxplots of the bias between Himawari-8/AHI AOD retrievals and ground measurements as a function of actual aerosol loading. In general, the retrievals worsen as the aerosol loading increases. Retrievals associated with AOD < 0.4 (lightly polluted) have small standard deviations with positive biases (overestimation). However, as the aerosol loading increases (0.4 < AOD < 1.0), the biases approach zero, gradually becoming more negative (increasing underestimations). This tendency intensifies when AOD ≥ 1.0 (heavily polluted). This suggests that retrieving AOD under heavily polluted conditions is still a challenge. As a function of ground measurements, the EA (red line in Fig. 12a) is estimated by fitting the mean biases for 25 bins using the least-squares method: $0.099 \times \tau + 0.055$. The EP is estimated by fitting the one standard deviation of the mean biases using the same least-squares method: $0.231 \times \tau + 0.180$. Therefore, the EE envelope for the AOD retrievals is $\text{EA} \pm \text{EP}$ (two blue lines in Fig. 12a): $[0.330 \times \tau + 0.024$ and $-0.132 \times \tau - 0.125]$. About 61% of data points fall within this EE envelope.

Fig. 12b shows boxplots of the bias between Himawari-8/AHI AOD retrievals and ground measurements as a function of the satellite-based AOD retrievals. In general, the retrievals worsen as the retrieved AOD increases. However, opposite to what Fig. 12a shows, there are large positive biases when the retrieved AOD < 0.35, and biases tend to be positive as the retrieved AOD increases in magnitude. When AOD > 1.0, the aerosol retrievals have large estimation uncertainties and large positive biases $>0.15$ (significant overestimations). As a function of Himawari-8/AHI AOD retrievals, the EA and EP are $0.256 \times \tau - 0.091$ and $0.263 \times \tau + 0.096$, respectively. Therefore, the EE envelope is $[0.519 \times \tau + 0.005$ and $-0.007 \times \tau - 0.194]$. About 64% of the data points fall within this EE envelope.

5. Summary and conclusions

The Himawari-8/AHI official Version 2 Level 2 aerosol products, including AOD retrievals at 500 nm and AE retrievals at 440–675 nm, from 2016 are obtained over land and ocean to gain knowledge of their performances at various spatial scales (site-specific, regional, and
Himawari-domain) and temporal (diurnal, daily, and seasonal) levels. These satellite-derived products are validated against AERONET and SONET ground-based measurements of corresponding aerosol optical properties at 98 sites. Discussed in detail are the retrievals errors associated with vegetation conditions, land-use type, and aerosol particle size. The expected errors for Himawari-8/AHI AOD products are also determined.

Overall, the satellite- and ground-based matched AOD retrievals are in good consistency with a correlation of 0.756, an MAE of 0.168, and an RMSE of 0.293 on the Himawari-domain scale. Observed are larger differences in performance on regional and site-specific scales. Australia and South Asia have significantly overestimated and underestimated, respectively, aerosol loadings. The AOD retrievals can more accurately capture daily variations than diurnal variations and perform better in summer and autumn than in spring and winter. In general, the satellite- and ground-based matched AE retrievals do not agree well on any spatial scale considered, showing overall underestimated uncertainties, especially over Australia, Southeast Asia, and China. The matched AE retrievals also do not agree well on any temporal scale considered.

The uncertainty analysis shows that the performance of Himawari-8/AHI AOD retrievals gradually improves, i.e., the estimation uncertainties decrease, as the vegetation coverage increases. They perform better over forests and water, but perform poorly over bare land and cropland, mainly because of the high surface brightness and significant seasonal variations in land cover. The Himawari-8/AHI AOD retrievals show large estimation errors for large-sized particles (e.g., mineral dust) and in general, the estimation errors decrease as the aerosol particle size decreases. The EEs are $[0.330 \times \tau + 0.024; -0.132 \times \tau - 0.125]$ and $[0.519 \times \tau + 0.005; -0.007 \times \tau - 0.194]$ as functions of surface- and satellite-based AODs, respectively. About 61% and 64% of the data points, respectively, fall within these EE envelopes.

Despite the overall sound performance of AOD retrievals over Himawari-domain region, this study illustrates that the accuracy of aerosol optical property retrievals from Himawari-8/AHI is highly variable, pending on different environmental conditions and underlying surfaces, especially for the AE. It raises concerns about the reliability of the current algorithm, especially under highly polluted conditions and for sparsely vegetated areas. The Himawari-8/AHI needs a more accurate aerosol optical property retrieval algorithm, which will require further exploration.

Declaration of Competing Interest

All authors made substantial contributions to this work. We declare no conflict of interest.

Acknowledgments

The National Key Research and Development Program of China (2017YFC1501702), the National Natural Science Foundation of China (91544217), the U.S. National Science Foundation (AGS1534670 and AGS1837811), and the BNU Interdisciplinary Research Foundation for the First-Year Doctoral Candidates (BNUXKJC1808) support this work. The Himawari-8/AHI aerosol products are available from the Himawari Monitor and P-Tree system (ftp.ptree.jaxa.jp), and AERONET data are available at https://aeronet.gsfc.nasa.

References
