Response to Reviewer1

This paper sets out a methodology and presents summary results for assimilating aerosol index measurements in to an aerosol forecasting model. This is relevant and interesting for the modelling community as it is effectively aerosol radiance assimilation. Radiance assimilation is common place in the NWP data assimilation community but has still to be explored for aerosol assimilation. For NWP it provides improved results compared to a level 2 retrieval and it has not yet been established whether the same may be true for aerosol assimilation. The article is very nicely written and provides a clear and precise overview of the work carried out. The detail of the forward model and assimilation procedure used is thoroughly covered but the clear structure of the article means the overall message of the paper is not lost in all the detail. The results of the assimilation experiment are succinctly presented in easy to understand figures without inflating the results or claiming more than is shown. This well written paper presents an advance to modelling science and deserves publication. I do, however, have a few minor comments that I list below

We thank the reviewer for his/her constructive comments

Minor comments

Question 1. It was not quite clear to me from the article whether the three models whose results are compared were the same version of the NAAPS model? I understand that the NAAPS reanalysis v1 was used to show the results with AOD assimilation (pg 8, paragraph 1) and that a free running version was used to provide the results without any aerosol assimilation at all (line 176). You also state that the assimilation system is based on variations of aerosol particles from NAAPS (line 106). Are all three the same version at the same resolution or are there differences between them? It would be beneficial to clarify this in the article as any differences will also impact on the results of the three experiments compared to Aeronet.

Response: *The same research version of the NAAPS model is used for all three experiments. For the natural runs, only the NAAPS forecast model was used, that is, without any form of data assimilation. For the NAAPS reanalysis version 1, NAAPS was run with additional assimilation using MODIS and MISR AOD data. For the OMI AI data assimilation as presented in the study, NAAPS was run with the newly developed OMI AI assimilation. All three runs are at the same spatial and temporal resolutions, and are driven by the same meteorology and model physics. We expect that differences among three model runs resulted from the different aerosol data assimilation schemes implemented versus the natural run. We have added the following sentence to clarify the issue:*

"Note that the same version of the NAAPS model with the same temporal and spatial resolutions, and driven by the same meteorological data, were used in constructing Figure 5 and thus the differences in Figures 5a, 5b and 5c only result from different aerosol data assimilation methods implemented (no data assimilation for the natural run)."

Question 2. Related to this, I'm slightly confused by your description of the post-processing system in lines 209-211. I would consider the construction of a new NAAPS analysis based on the background NAAPS aerosol concentrations and increments as derived from the assimilation system to be part of the assimilation process itself. In fact I would assume that this updated analysis state would be forecast forward in time to create the background state for the next cycle of the data assimilation process. Is this not the case?

Response: *Post-processing as mentioned in lines 209-211 is a part of the typical data assimilation process. In a typical data assimilation method, increments are constructed based on the differences between observations and modeled parameters (innovations), as well as error characteristics of both model and observations. These increments include new changes that need to be made for each model grid. At the last step of a typical data assimilation process, the modeled background is updated by adding those increments (or corrections) to construct a revised background state (analysis). The revised background state is then used as the initial state for the forecast for the next time cycle.*

In another word, analysis = background + increments. Note that a similar post-processing step is also included in the NAVDAS-AOD for MODIS and MISR AOD assimilation (Zhang et al., 2008).

Zhang, J. and J. S. Reid, D. Westphal, N. Baker, and E. Hyer, A System for Operational Aerosol Optical Depth Data Assimilation over Global Oceans, J. Geophys. Res., 113, D10208, doi:10.1029/2007JD009065, 2008.

Question: 3. Your Figure 7 is a comparison of the vertical profiles of the NAAPS natural and AI DA runs. Assuming that the AI DA runs are as described above, so an analysis model state that is used as the initial condition for a short forecast to create the background state for the next assimilation cycle, then I don't believe you can draw the conclusions that you do in lines 493-498. There is no guarantee that the profile before assimilation is the same as the nature run profile and so you can not disentangle what profile differences come from previous assimilation versus what is due to the assimilation of the AI data in the current cycle. To look at the impact of assimilating AI data in one specific cycle you would need to plot the background model state versus the analysis state, rather than the nature run.

Response: *Both natural and OMI AI DA runs were performed with the same version of the NAAPS model, at the same spatial and temporal resolutions, with the same initial conditions at the beginning of the study period (00Z, July 1, 2007). The only difference between the two-month natural and OMIAI DA runs is that OMI AI data assimilation was implanted in the OMI AI DA run, while OMI AI data assimilation was not implanted for the natural run. Therefore, the differences between the two model runs arise uniquely from the OM AI data assimilation process.*

Note that for a given cycle, once the model has begun integrating forward in time, the differences in vertical profiles between the natural and OMI AI DA runs will also be impacted by increments from previous cycles (after the starting date of the study period). So the differences between

OMI DA and natural runs as shown in Figure 7 can be considered as an integrated effect of OMI AI DA from 00Z, July 01 to 12 Z, July 28, 2007.

We added the following sentence to avoid confusion: "Note that the differences between OMI DA and natural runs as shown in Figure 7 are essentially an integrated effect of OMI AI DA from 00Z, July 01 to 12 Z, July 28, 2007."

Question 4. What do you think is the impact of using gridded OMI data (line 130-133) versus the higher resolution (I assume) AOD data of the reanalysis. Do you think that the results would change if you were able to use the AI data at its native resolution and that it would closer match the results of the reanalysis?

Response *: I assume the reviewer meant to say "high resolution (I assume) AI data" based on the second sentence. Changes are definitely expected with the use of AI at its native resolution. This is because each data point included/removed will introduce changes in the computed increments. Still, for a given grid, the gridded OMI data represents the averaged properties for that grid. Thus, we expect the difference between using gridded data or OMI data at the native resolution to be marginal.*

Question 5. It is interesting and useful to have an idea of the computational burden of the call to the radiative transfer model in Section 4.4, but it would add perspective if this could be compared to the equivalent computational burden for AOD assimilation.

Response: The time scale for running AOD assimilation for 1 month is at the hourly level, depending on the machines used. We have added the following remark:

"In comparison, the time scale for running AOD assimilation for 1 month is at the hourly level."

Typos

Question: Pg. 7, line 147: AERONET

Response: *done.*

Question: Pg. 8, line 169: precipitation data are used to constrain the wet removal process

Response: *done*

Question: Pg. 18, line 405-407: It is unclear to me which figures you are talking about in this sentence. I assume it is Figure 3c, but coming directly after discussion of a comparison of 3b to 3d it needs further clarification.

Response: *We added Figure 3c in the text.*

Responses to reviewer 2 comments

This paper develops a data assimilation scheme using the VLIDORT radiative transfer model and simulated aerosol information from the NAAPS model to assimilate OMI AI measurements into the NAAPS model. Including the OMI AI assimilation improves the NAAPS simulation compared to the OMI AI, and improves NAAPS simulated AOD compared to AERONET AOD, but it does not outperform the NAAPS reanalysis AOD compared to AERONET. Overall the paper is well written and their data assimilation approach is well explained. I do have some comments.

We thank the reviewer for his/her comments

Question : My main issue with the paper is that the authors state in the abstract: “Improvements in model simulations demonstrate the utility of OMI AI data assimilation for improving the accuracy of aerosol model analysis over cloudy regions and bright surfaces.” But this is not really shown anywhere in the paper.. On line 149 it is even stated: “As AERONET data require a cloud-free line of sight to the solar disk, the performance of OMI AI data assimilation over overcast regions is not evaluated.” Yes there are AI measurements over cloudy regions and bright surfaces, but nowhere in the paper have the authors specifically evaluated the performance of their analysis over bright or cloudy surfaces compared to, say, the NAAPS reanalysis AOD from MODIS and MISR. The authors even state that their assimilation does not improve the NAAPS AOD compared to the reanalysis AOD, so where is the evidence of improvement over bright and cloudy surfaces? It is not explicitly stated which products from MODIS and MISR go into the NAAPS reanalysis, but both MODIS deep blue and MISR retrieve AOD quite accurately over bright surfaces, especially deserts, so this statement really should be backed up somehow.

Response: *One of the advantages of OMI AI is its ability to detect UV- absorbing aerosols over cloudy skies as well bright surfaces such as over desert regions and snow/ice-covered regions. In this study, we examined the possibility of assimilating OMI AI data over cloudy regions as well as desert regions (bright surfaces). All quality-checked (excluding noisy data and data with row anomalies) OMI AI data over cloudy regions and desert regions were used in the assimilation process. In comparison, no reliable AOD retrievals are available over cloudy regions from traditional passive-based aerosol retrieval methods. Also, retrievals over the desert regions are also limited to select algorithms. Therefore, having the OMI AI data assimilation capability over cloudy regions and over bright surfaces is an advancement in aerosol data assimilation.*

We agree with the reviewer that it is hard to evaluate NAAPS performance over cloudy regions. We also agree that OMI AI is an indirect indicator of aerosol properties, and assimilating OMI AI typically cannot out-perform assimilating of MODIS/MISR AOD over cloud free regions. Nonetheless, the improvements in NAAPS analyses over cloudy regions or bright surfaces through OMI AI DA can be directly or indirectly illustrated from three aspects.

First, our study suggests, based on the AERONET evaluation, that over cloud-free regions, in comparing NAAPS natural runs (without aerosol assimilation), the accuracy of NAAPS analyses is improved with the assimilation of OMI AI data. This suggests OMI AI data can be used to improve NAAPS performance. Also, OMI AI has comparable capability to detect UV absorbing aerosols over cloud-free skies as well over cloudy skies, thus, benefits in NAAPS analysis over cloudy regions or bright surfaces are expected through assimilating quality- controlled OMI AI data over cloudy and bright surfaces. Note, no passive-based AOD data are currently available for assimilation over cloudy regions.

Secondly, as the reviewer mentioned, there are AI measurements over cloudy regions and bright regions for evaluation. We have performed this approach in the paper. One of the steps for a data assimilation system is to check the difference between observation and analysis (O-A), as well as the difference between observation and background (O-B). OMI AI can be considered as observations. NAAPS data includes aerosol concentrations, and thus to perform O-A or O-B, we used the forward model and computed simulated OMI AI using NAAPS data. The two-month (July and August 2007) mean O-A is shown in Figure 4d and the two-month mean O-B is shown in Figure 4h. While near zero O-A values are found for the study region as shown in Figure 4d, large O-B values can be found in Figure 4h over heavy smoke and dust aerosol polluted regions. Note to compute two-month mean O-A and O-B, both NAAPS and OMI AI data over both cloudy and cloud-free skies were used. At the instantaneous level, Figures 3b and 3f show the O and A for 12UTC, July 28, 2007. Figures3b and 3e show the O and B for 12 UTC, July 28, 2007 as well. Again, while observation and simulated AI using NAAPS analysis are similar over both cloudy regions and cloud free regions, large discrepancies can be found between OMI AI and simulated OMI AI using NAAPS natural run data. The O-A/O-B analyses at both two-month mean and instantaneous levels indicating NAAPS performance can be improved over cloudy regions.

Third, as a qualitative check, as highlighted in red ellipses in Figure 3, the NAAPS AOD patterns after OMI AI DA show a very similar spatial pattern to OMI AI over both cloudy and non-cloudy regions. This can be considered as an indirect indicator that NAAPS AOD patterns match OMI AI patterns after OMIAI DA, even over cloudy regions.

However, we have revised the sentence along the lines suggested by the reviewer: "Improvements in model simulations demonstrate the utility of OMI AI data assimilation for aerosol model analysis over cloudy regions and bright surfaces"

Other comments:

Question: - In section 4.3 Sensitivity Analysis, the authors discuss how varying smoke SSA affects the AI and conclude that there is a need for regionally varying SSA values for smoke to be included for future studies. However, the issue is not necessarily varying smoke SSA, it is the fact that the model used in this paper treats all "smoke" as one aerosol type with a single SSA value. In reality, "smoke" is composed of both black and organic (that is, brown) carbon, which have different SSA values, and different areas have different contributions of black and brown carbon to the overall "smoke". So really what the authors are showing is a major limitation in modelling absorbing aerosol with the particular model they chose.

Response: *Agreed. However, the problem we are encountering is very similar to that faced by the passive-based AOD retrieval community. Dust/smoke aerosol properties vary as a function of region and season, creating a problem not only for this study but for AOD retrievals using passive sensors. To deal with this problem, regional-based aerosol properties are used in some algorithms (e.g MODIS Dark Target). Similar methods may be also adopted for this study, as we have mentioned. However, this is outside the scope of our paper and is the subject for a future study.*

Question:- Also in section 4.3, the authors state: "Interestingly, although simulated AI values are significantly affected by perturbing SSA values as shown in Figure 6, less significant impacts are observed for NAAPS AOD." However, this is to be expected, because AOD is a measure of the total extinction due to the presence of aerosols, so changing the fraction that is either scattering or absorbing would not change the overall extinction.

Response: *NAAPS-modeled UV-absorbing aerosol (dust and smoke) concentrations are corrected based on OMI AI observations. We agree that dust and smoke aerosols are only a fraction of the total aerosol concentration.*

Question: - Lines 136-139: "Isolated high AI values are removed as follows. First, for a 4x4 pixel box, if the mean AI is less than 0.7 but an individual AI value is larger than 0.7, then that one value is removed. Second, if the standard deviation of AI values for a 3x3 pixel box surrounding a pixel is larger than 0.5, that individual AI value is likewise removed" It is not explained how the authors came up with this criteria, and it might be helpful for them to include a bit of an explanation.

Response: *Both approaches are essentially homogeneity tests that are used for identifying outliers. The thresholds are estimated empirically through visual inspection.*

We added this sentence: "Note that both approaches are essentially homogeneity tests that are used for identifying outlies. The thresholds are estimated empirically through visual inspection."

Technical comments:

Question: - Lines 80-86 are worded a little confusingly: "AI retrievals are currently computed using observations from sensors with ozone-sensitive channels. For example, the Ozone Monitoring Instrument (OMI), Ozone Mapping and Profiler Suite (OMPS), TROPOspheric Monitoring Instrument (TROPOMI) and the future Plankton, Aerosol, Cloud and ocean Ecosystem (PACE) mission can detect UV-absorbing aerosol particles, such as black carbon laden smoke or iron-bearing dust, over bright surfaces, such as desert, snow and ice covered regions, and aerosol plumes above clouds (e.g. Torres et al., 2012; Yu et al., 2012; Alfaro-Contreras et al., 2014; 2016).” At first it is being discussed how AI retrievals use ozone sensitive channels, then the “for example” is talking about detecting absorbing aerosols.

Response: *We revised the sentence to read:“For example, the Ozone Monitoring Instrument (OMI), Ozone Mapping and Profiler Suite (OMPS), TROPOspheric Monitoring Instrument (*TROPOMI*) and the future Plankton, Aerosol, Cloud and ocean Ecosystem (PACE) mission include ozone-sensitive channels that can detect UV-absorbing aerosol particles, such as black carbon laden smoke or iron-bearing dust, over bright surfaces, such as desert, snow and ice covered regions, and aerosol plumes above clouds (e.g. Torres et al., 2012; Yu et al., 2012; Alfaro-Contreras et al., 2014; 2016).”*

Question: - Line 276: dust “plums” should be “plumes”

Response: *Done.*

Question: - Line 453: “proving” should be “providing”

Response: *Done.*

Response to short comment

We thank the executive editor for the comment. We will add the version number in the revised version of the paper as suggested.

The OMI data assimilation scheme (V1.0) is constructed using VLIDORT and NAVDAS-AOD for NAAPS analyses and forecasts. The VLIDORT radiative transfer code is a property of RT Solutions Inc. The distribution of the full VLIDORT package is publicly available, and comes with a standard GNU public license, through direct contact with RT Solutions Inc. (http://www.rtslidort.com/mainprod_vlidort.html). Both NAAPS and NAVDAS-AOD are proprietary to the Naval Research Laboratory, United States Department of the Navy. Given their association with a defense system, they are not available publicly. This situation is similar to that in other major centers such as ECMWF, JMA, and UKMO. Nevertheless, both NAAPS and NAVDAS-AOD are well documented in past studies (e.g. Lynch et al., 2016; Zhang et al., 2008; 2011; 2014; Rubin et al., 2017) and we have made every effort to thoroughly report our methods so that they may be replicated. In addition, AOD fields from the NAAPS OMI AI DA runs and natural runs over the study region and for the study period will be shared in the supplement.

We have revised the code and data availability section to read:

Code and data availability: The OMI data assimilation scheme (V1.0) is constructed using VLIDORT and NAVDAS-AOD for NAAPS analyses and forecasts. The VLIDORT radiative transfer mode is a property of RT Solutions Inc. The VLIDORT code is publicly available, and comes with a standard GNU public license, through direct contact with RT Solutions Inc. (http://www.rtslidort.com/mainprod_vlidort.html). Both NAAPS and NAVDAS-AOD are proprietary to Naval Research Laboratory, United States Department of the Navy. Nevertheless, both NAAPS and NAVDAS-AOD are well documented in past studies (e.g. Lynch et al., 2016; Zhang et al., 2008; 2011; 2014; Rubin et al., 2017) and we have made every effort to thoroughly report our methods so that they may be replicated. AOD fields from the NAAPS OMI AI DA and natural runs over the study region and period are shared as the supplement to the paper for readers who are interested. The NAAPS reanalysis data are available from the USGODAE web site (https://nrlgodae1.nrlmry.navy.mil/cgi-bin/datalist.pl?dset=nrl_naaps_reanalysis&summary=Go**.** The OMI OMAERUV data are available from the NASA's Goddard Earth Sciences Data and Information Services Center (*GES DISC*; https://disc.gsfc.nasa.gov/datasets/OMAERUV_003/summary). AERONET data are obtained from the NASA AERONET webpage (https://aeronet.gsfc.nasa.gov/).

Development of an OMI AI data assimilation scheme for aerosol modeling over bright surfaces—a step toward direct radiance assimilation in the UV spectrum

Jianglong Zhang[1], Robert J. D. Spurr[2], Jeffrey S. Reid[3], Peng Xian[3], Peter R. Colarco[4], James R. Campbell[3], Edward J. Hyer[3], and Nancy Baker[3]

[1]Department of Atmospheric Sciences, University of North Dakota, Grand Forks, ND

[2]RT SOLUTIONS Inc., Cambridge MA

[3]Marine Meteorology Division, Naval Research Laboratory, Monterey, CA

[4]NASA Goddard Space Flight Center, Greenbelt, MD

Revised version s~~S~~ubmitted to

GMD

September~~June~~ 2020

Corresponding Author: jianglong.zhang@und.edu

Abstract

Using the Vector LInearized Discrete Ordinate Radiative Transfer (VLIDORT) code as the main driver for forward model simulations, a first-of-its-kind data assimilation scheme has been developed for assimilating Ozone Monitoring Instrument (OMI) aerosol index (AI) measurements into the Naval Aerosol Analysis and Predictive System (NAAPS). This study suggests both RMSE and absolute errors can be significantly reduced in NAAPS analyses with the use of OMI AI data assimilation, when compared to values from NAAPS natural runs. Improvements in model simulations demonstrate the utility of OMI AI data assimilation for ~~r improving the accuracy of~~ aerosol model analysis over cloudy regions and bright surfaces. However, the OMI AI data assimilation alone does not out-perform aerosol data assimilation that uses passive-based aerosol optical depth (AOD) products over cloud free skies and dark surfaces. Further, as AI assimilation requires the deployment of a fully-multiple-scatter-aware radiative transfer model in the forward simulations, computational burden is an issue. Nevertheless, the newly-developed modeling system contains the necessary ingredients for assimilation of radiances in the ultra-violet (UV) spectrum, and our study shows the potential of direct radiance assimilation at both UV and visible spectrums, possibly coupled with AOD assimilation, for aerosol applications in the future. Additional data streams can be added, including data from TROPOspheric Monitoring Instrument (TROPOMI), Ozone Mapping and Profiler Suite (OMPS) and eventually with the Plankton, Aerosol, Cloud and ocean Ecosystem (PACE) mission.

1.0 Introduction

Operational chemical transport modeling (CTM) of atmospheric aerosol particles, including simulation of sources and sinks and long-range transport of aerosol events such as biomass burning aerosols from fires and dust outbreaks, is now commonplace at global meteorology centers for air quality and visibility forecasts (e.g. Sessions et al, 2015; Lynch et al., 2016). Variational and ensemble-based assimilation of satellite derived aerosol products such as aerosol optical depth (AOD), lidar backscatter measurements, and surface aerosol properties, can substantially improve accuracies in CTM analyses and forecasts (Zhang et al., 2008; 2011; 2014; Yumimoto et al., 2008; Uno et al., 2008**;** Benedetti et al., 2009; Schutgens et al., 2010; Sekiyama et al., 2010; Saide et al. 2013; Schwartz, 2012; Li et al., 2013; Rubin et al., 2017; Lynch et al., 2016).

Currently, the main satellite inputs for operational aerosol modeling are AOD products derived from passive-based polar orbiting imagers, such as the Moderate Resolution Imaging Spectroradiometer (MODIS), the Visible Infrared Imaging Radiometer Suite (VIIRS), and the Advance Very High Resolution Radiometer (AVHRR). Experimentation is proceeding with the use of products from the multi-angle imaging spectroradiometer (MISR) (e.g., Lynch et al., 2016; Randles et al. 2017; Buchard et al. 2017) and from geostationary instruments such as Himawari and Geostationary Operational Environmental Satellite (GOES). A major advantage with such passive-based satellite sensors is that the AOD is retrieved with high spatial and temporal resolutions over relatively broad fields-of-view (e.g. Zhang et al., 2014). For example, MODIS and VIIRS provide near-global daily daytime coverage (e.g. Levy et al., 2013; Hsu et al., 2019) and GOES and Himawari are capable of retrieving AOD over North American and East Asia regions at sub-hourly temporal resolution (e.g. Bessho et al., 2016).

To date, these traditional passive-based satellite AOD retrievals have been limited to darker surfaces and relatively cloud-free conditions. The widely-used MODIS Dark Target aerosol data, for instance, are available globally over only oceans and dark land surfaces (e.g. Levy et al., 2013). The MISR and MODIS Deep Blue aerosol products are also available over some arid environments, but are not applicable to snow and ice covered regions (e.g. Kahn et al., 2010; Hsu et al., 2013). Also, none of the above-mentioned aerosol products are valid over cloudy regions.

In comparison to AOD, the semi-quantitative UV-based aerosol index (AI) has long been used to monitor major aerosol events such as smoke plumes and dust storms, starting with the Total Ozone Mapping Spectrometer (TOMS) from the late 1970s (Herman et al., 1997). AI is derived using the ratio of observed UV radiances to simulated ones assuming only a clear Rayleigh sky (e.g. Torres et al., 2007). AI retrievals are currently computed using observations from sensors with ozone-sensitive channels. For example, the Ozone Monitoring Instrument (OMI), Ozone Mapping and Profiler Suite (OMPS), TROPOspheric Monitoring Instrument (TROPOMI) and the future Plankton, Aerosol, Cloud and ocean Ecosystem (PACE) mission include ozone sensitive channels that can detect UV-absorbing aerosol particles, such as black carbon laden smoke or iron-bearing dust, over bright surfaces, such as desert, snow and ice covered regions, and aerosol plumes above clouds (e.g. Torres et al., 2012; Yu et al., 2012; Alfaro-Contreras et al., 2014; 2016).

To complement existing AOD assimilating systems, we have developed an AI data assimilation (AI-DA) system that is capable of assimilating OMI AI over bright surfaces and cloudy regions for aerosol analyses and forecasts. This study can be considered as one of the first attempts for direct radiance assimilation in the UV spectrum for aerosol applications, as AI can be directly computed from UV radiances and the developed OMI AI-DA system has all necessary components for a typical radiance assimilation package. In time we expect our assimilation model to merge with AOD or solar radiance assimilation to influence aerosol loading, height and absorption (e.g., VIIRS+OMPS product; such as Lee et al. 2015). Details of the developed OMI AI assimilation system are presented in the paper, which is organized as follows: Data sets used in the study are summarized in Section 2; Section 3 discusses the components of the AI-DA system. Section 4 provides an evaluation of the developed system; and Section 5 contains a summary discussion.

2.0 Datasets and Models

Three datasets are used in this study. These are: (i) the OMI level 2 UV aerosol product (OMAERUV; Torres et al., 2007), (ii) the Aerosol Robotic Network (AERONET; Holben et al., 1998) AOD product, and (iii) reanalysis data from the Naval Aerosol Analysis and Prediction System (NAAPS; Lynch et al., 2016), which was the first operational global aerosol mass transport model available to the community. The assimilation system is based on spatial and temporal variations of aerosol particles from NAAPS (Zhang et al., 2006; 2008), and the Vector LInearized Discrete Ordinate Radiative Transfer (VLIDORT; Spurr, 2006) code is used to construct a forward model for the AI-DA system.

2.1 OMI aerosol product

UV Aerosol Index data from the OMI level 2 version 3 UV aerosol products (OMAERUV) are used in this study. The OMI instrument is on board the Aura satellite (launched in 2004) and it observes the earth's atmosphere over the UV/visible spectrum with a pixel size of 13x24 km at nadir for the global scan mode, and a swath of ~2600 km (Levelt et al., 2018). The daytime equatorial crossing for the Aura platform is ~1:30 p.m. The dataset comprises the UV AI, viewing and solar geometries, spectrally-dependent surface albedos at the 354 and 388 nm spectral channels, terrain pressure, geolocations, x-track and algorithm quality flags, plus other aerosol and ancillary parameters. The UV AI is designed to detect UV-absorbing aerosol particles, and is based on radiance observations at 354 nm (I_{obs354}) and calculated radiance (I_{cal354}) at 354 nm for a Rayleigh (no aerosol) atmosphere (e.g. Torres et al., 2007) as defined as

$$AI = -100 \log_{10} \frac{I_{obs354}}{I_{cal354}}. \qquad (1)$$

Unbiased, noise-reduced, quality-assured AI data are necessary for AI data assimilation. This is especially important for OMI observations, due to this particular sensor suffering from the well-referenced "row anomalies" issues (Torres et al., 2018). To remove pixels with row anomalies, only retrievals with x-track flag values of 0 are retained. Also, abnormal AI values were identified over mountain regions. Thus, retrievals with terrain/surface pressure less than 850 hpa are excluded in the study. Finally, only retrievals with OMI AI values larger than -2 are used. Therefore, OMI observations over cloudy skies, which could have negative OMI AI values, are also included.

Both cloud-free and above-cloud AI data satisfying these quality checks are aggregated / averaged in 1x1° (Latitude/Longitude) bins. As a radiative transfer model run is applied for each observation, the gridded data are used in the assimilation process in order to reduce the computational burden. Averaged parameters for the gridded data include the solar and sensor zenith angles, the relative azimuth angles, the spectrally-dependent surface albedos at 354 and 388 nm, the cloud fraction, and the AI values themselves. Additional quality assurance steps are also applied during the spatial-averaging process. Isolated high AI values are removed as follows. First, for a 4x4 pixel box, if the mean AI is less than 0.7 but an individual AI value is larger than 0.7, then that one value is removed. Second, if the standard deviation of AI values for a 3x3 pixel box surrounding a pixel is larger than 0.5, that individual AI value is likewise removed. Note that both approaches are essentially homogeneity tests that are used for identifying outlies. The thresholds are estimated empirically through visual inspection.

2.2 AERONET data

Version 3 level 2 daytime, cloud-cleared and quality-assured AERONET data are used to evaluate the performance of the OMI AI data assimilation in our study (Holben et al., 1998; Giles et al., 2019). During daytime, AOD from AERONET instruments are derived by measuring the attenuated solar radiance typically at seven wavelengths ranging from 340 to 1020 nm. In this study, AERONET data are collocated with NAAPS analyses with and without OMI AI assimilation. In order to collocate AERONET and NAAPS AOD data, AERONET~~ENT~~ AOD values within ±30 minutes of a given NAAPS analysis time are averaged and used as ground-based AOD values for the NAAPS 1x1° (Latitude/Longitude) collocated bins. As AERONET data require a cloud-free line of sight to the solar disk, the performance of OMI AI data assimilation over overcast regions is not evaluated.

2.3 NAAPS and NAAPS reanalysis data

The NAAPS (http://www.nrlmry.navy.mil/aerosol/) model is a multi-species, three-dimensional, Eulerian global transport model using operational Navy Global Environmental Model (NAVGEM) as the meteorological driver (Hogan et al., 2014). NAAPS provides 6-day forecasts at a 3-hour interval with a spatial resolution of 1/3° (latitude/Longitude) and 42 vertical levels on a global scale. NAAPS predicts four aerosol particle classes: anthropogenic and biogenic fine particles (ABF, such as primary and secondary organic aerosols and sulfate aerosols); dust, biomass burning smoke; and sea salt (Lynch et al, 2016).

The 2003-2018 NAAPS reanalysis version 1 (v1) (Lynch et al., 2016) is a modified version of the operational NAAPS model. In this version, quality-controlled retrievals of AOD from MODIS and MISR (Zhang et al., 2006; Hyer et al., 2011; Shi et al., 2014) are assimilated into NAAPS through the Naval Research Laboratory Atmospheric Variation Data Assimilation System-AOD system (NAVDAS-AOD; e.g., Zhang et al., 2008; Zhang et al., 2011; Zhang et al., 2014). Aerosol source functions, including biomass burning, smoke and dust emissions, are tuned regionally based on the AERONET data. Other aerosol processes, including dry deposition over water, are also tuned based on AOD data assimilation correction fields. NOAA Climate Prediction Center (CPC) MORPHing (CMORPH) precipitation data are used to constrain~~t~~ the wet removal process within the tropics (Joyce et al., 2004). The usage of CMORPH avoids the ubiquitous precipitation bias that exists in all global atmospheric models (e.g. Dai, 2006) and is proven to improve aerosol wet deposition, therefore yielding better AOD (Xian et al., 2009). The reanalysis agrees reasonably well with AERONET data on a global scale (Lynch et al., 2016) and also reproduces AOD trends that are in a good agreement with satellite based analysis (e.g., Zhang and Reid, 2010; Hsu et al., 2012). In this study, we use a free running version of NAAPS reanalysis v1 without AOD assimilation to provide aerosol fields every 6 hours at 1°x1° (Latitude/Longitude) resolution.

2.4 VLIDORT radiative transfer code

VLIDORT is a linearized, multiple-scatter radiative transfer model for the simultaneous generation of Stokes 4-vectors and analytically-derived Jacobians (weighting functions) of these

4-vectors with respect to any atmospheric or surface property (Spurr, 2006). The model uses discrete-ordinate methods to solve the polarized plane-parallel RT equations in a multi-layer atmosphere, plus the solution of a boundary value problem and subsequent source-function integration to obtain radiation fields at any geometry and any atmospheric level. VLIDORT has a "pseudo-spherical" *ansatz*: the treatment of solar-beam attenuation in a spherical-shell atmosphere before scattering. Single-scattering in VLIDORT is accurate for both line-of-sight and solar-beam spherical geometry. The model has a full thermal emission capability. VLIDORT has two supplements, one dealing with bidirectional (non-Lambertian) reflection at the surface, and the other with the inclusion of surface light sources (SIF or water-leaving radiances). Full details on the VLIDORT model may be found in a recent review paper (Spurr and Christi, 2019, and references to VLIDORT therein).

VLIDORT is used to simulate the AI in this study. Simulations at 354 and 388 nm are performed both for Rayleigh atmospheres, and for scenarios with aerosol loadings (four mass-mixing profiles for different aerosol types) taken from the NAAPS model. In addition to the AI, Jacobian calculations are needed with respect to these aerosol profiles. Firstly, radiance Jacobians with respect to these four mass-mixing profiles are computed analytically using VLIDORT's linearization facility, and secondly the associated Jacobians of AI are further derived through a second VLIDORT linearization with respect to the Lambertian-equivalent reflectivity. The details of this process is given in the next section

3.0 OMI AI assimilation system

The OMI assimilation system has three components: a forward model, a 3-D variational assimilation system, and a post-processing system. Based on the background NAAPS 3-D aerosol concentrations for dust, smoke, ABF, and sea salt aerosols, the forward model not only computes the associated AI values, but also their Jacobians of AI with respect to the four aerosol mass-loading profiles. The 3-D variational assimilation system is a modified 3-D AOD system (Zhang et al., 2008; 2011; 2014) that computes increments for dust and smoke aerosol concentrations based on OMI AI data. The post-processing system constructs a new NAAPS analysis based on the background NAAPS aerosol concentrations and increments as derived from the 3-D variational assimilation system. Details of the forward model and the modified NAVDAS-AOD system are described in this section.

3.1 Forward model for simulating OMI AI

To construct an AI-DA system, a forward model is needed to simulate AI using aerosol concentrations from NAAPS. In this study, the forward model is built around the VLIDORT model, following a similar method to that suggested in Buchard et al. (2015). Here VLIDORT is configured to compute OMI radiances and Jacobians as functions of the observational conditions at 354 and 388 nm, using geolocation information from OMI data such as satellite zenith, solar zenith and relative azimuth angles, as well as ancillary OMI data (surface albedos at 354 and 388 nm).

To convert from NAAPS mass-loading concentrations to aerosol extinction and scattering profiles, we require aerosol optical properties for the four species at 354 and 388 nm, which are summarized in Table 1. The optical properties of ABF (assumed to be sulfate in this study), sea salt, dust and smoke aerosols, including mass extinction cross sections and single scattering albedos at 354 and 388 nm are adapted from NASA's Goddard Earth Observing System version 5 (GEOS-5) model (e.g. Colarco et al., 2014; Buchard et al., 2015). Note that the study period is

July and August of 2007 over Africa, coinciding with the early biomass burning season associated with lower single scattering albedo values (Eck et al., 2013). With that in mind, we choose a quite low value of 0.85 for the single-scattering albedo value at 354nm (e.g. Eck et al., 2013; Cochrane et al., 2019). A slightly higher single scattering albedo of 0.86 is assumed at 388 nm. The slight increase in single scattering albedo from 354 to 388 nm has also been observed from Solar Spectral Flux Radiometer (SSFR) observations during the recent NASA ObseRvations of CLouds above Aerosols and their intEractionS (ORACLES) Campaign **(**Pistone et al., 2019). Scattering matrices for dust, smoke, sea salt and sulfate (to represent ABF) aerosols are based on associated expansion coefficients (e.g. Colarco et al., 2014; Buchard et al., 2015) taken from NASA's GEOS-5 model. Also to reduce computational expenses, scalar radiative transfer calculations are performed.

To simulate OMI AI, the Lambertian Equivalent Reflectivity (LER) at 388 nm (R_{388}) is needed for estimating LER at 354 nm. The R_{388} is calculated from VLIDORT, based on equation 2 below, adapted from Buchard et al. (2015), or

$$R_{388} = \frac{I_{aer388}(\rho_{388}) - I_{ray388}(0)}{T + S_b(I_{aer388}(\rho_{388}) - I_{ray388}(0))} \quad . \qquad (2)$$

$I_{ray388}(0)$ is the calculated path radiance at 388 nm assuming a Rayleigh atmosphere with surface albedo 0. T and S_b are the calculated transmittance and spherical albedo at 388 nm. $I_{aer388}(\rho_{388})$ is the computed radiance including 3-D aerosol fields from NAAPS and the 388 nm surface albedo from OMI data. In Buchard et al. (2015), an adjusting factor is applied to R_{388} by adding the difference between climatological surface albedos at 354 and 388 nm. The similar approach is also adopted in this study, as shown in their Equation 3.

$$R'_{388} = R_{388} - (\rho_{388} - \rho_{354}) \quad . \qquad (3)$$

Here, R'_{388} is surface albedo adjusted Lambertian Equivalent Reflectivity at 388 nm. ρ_{388} and ρ_{354} are surface albedo values at 388 and 354 nm channels that are obtained from the OMI OMAERUV data. Finally, the simulated AI (AI_{naaps}) is given by

$$AI_{naaps} = -100 \log_{10} \frac{I_{aer354}(\rho_{354})}{I_{ray354}(R'_{388})} \quad . \qquad (4)$$

Here, $I_{aer354}(\rho_{354})$ is the calculated radiance at 354 nm using NAAPS aerosol fields as well as the OMI-reported surface albedo at 354 nm (ρ_{354}). $I_{ray354}(R'_{388})$ is the calculated radiance assuming a Rayleigh atmosphere and the derived value of R'_{388} as surface albedo (Buchard et al., 2015).

The forward model-simulated OMI AI values are inter-compared with OMI AI values as shown in Figure 1 for the study region. A total of one month (01-31 July 2007) of NAAPS reanalysis data and OMI AI data were used. Note that OMI AI data over both cloud-free and cloudy skies were used. Since surface albedos included in the OMI data represent reflectivities under clear-sky situations, the albedo under cloudy sky is then computed

$$\rho_{cld} = \rho_{clr} * (1 - f_c) + 0.8 * f_c \quad . \qquad (5)$$

Here, ρ_{clr} and f_c are the clear sky surface albedo (e.g. ρ_{354} or ρ_{388}) and the cloud fraction, both quantities obtained from the OMI dataset. Clouds are assumed to be tropospheric (close to the surface) with an UV albedo of 0.8, such that this equation applies to both the 354 and 388 nm channels.

Figure 1a shows the spatial distribution of NAAPS AOD over Central and North Africa, using collocated NAAPS and OMI AI datasets. OMI AI data are grid-averaged in 1°x1° (latitude/longitude) bins. Also, we focus over Africa in this paper as this area includes dust plumes over deserts and smoke plumes overlying stratus cloud decks. The Arctic is not included as additional efforts may be needed to fully understand properties of sea ice reflectivity; we leave this topic for a future paper. Only bins that have valid NAAPS and OMI AI data are used to generate

Figure 1. Dust plumes are visible over North Africa and the Persian Gulf, and a smoke plume from Central Africa is also evident. These UV-absorbing aerosol plumes are also captured by OMI AI, as seen in Figure 1c. Shown in Figure 1b are the simulated OMI AI using the NAAPS aerosol fields and viewing geometries and surface albedos from OMI. The simulated OMI AI shows similar patterns to those derived from OMI, especially for the dust plumes over North Africa and smoke plumes over Central Africa. An overall correlation of 0.785 is found between simulated and satellite-retrieved OMI AI values, as shown in Figure 1, suggesting the forward model is functioning reasonably as designed.

3.2~~1~~ Forward model for Jacobians of AI

Jacobians of OMI AI with respect to aerosol mass concentrations are needed for the OMI AI assimilation system. In this study, AI Jacobians (K) are calculated from radiance Jacobians with respect to aerosol mass concentrations for four aerosol species (smoke, dust, ABF/sulfate, sea-salt) at 354 nm ($K_{354,nk} = \frac{\partial I_{aer354}}{\partial M_{nk}}$) and 388 nm ($K_{388,nk} = \frac{\partial I_{aer388}}{\partial M_{nk}}$) wavelengths. Here M_{nk} is the mass concentration for aerosol type, k, and for vertical layer, n. I_{aer354} and I_{aer388} are radiances for the 354 and 388 nm channels, respectively. $K_{354,nk}$ and $K_{388,nk}$ are the corresponding radiance Jacobians at 354 and 388 nm, respectively. AI Jacobians can then be calculated by analytic differentiation of the basic formula in Equation (1), and, after some algebra, we find the following result:

$$\frac{\partial AI}{\partial M_{nk}} = \mathcal{A}_1 K_{354,nk}(\rho_{354}) + \mathcal{A}_2 K_{388,nk}(\rho_{388}) \quad . \qquad (6)$$

Here, $\mathcal{A}_1$ and $\mathcal{A}_2$ are given respectively by Equations (7) and (8), as

$$\mathcal{A}_1 = \left(-\frac{100}{I_{aer354}(\rho_{354}) \times \ln 10}\right) \quad \text{, and} \qquad (7)$$

$$\mathcal{A}_2 = \left(-\frac{100}{I_{ray354}(\cancel{R_{388}}R'_{388})\times \ln 10}\right)\frac{\partial I_{ray354}(\cancel{R_{388}}R'_{388})}{\partial R}\left[\frac{(1-S_{388}R_{388})^2}{T_{388}}\right] \quad . \qquad (8)$$

Based on these equations, radiance Jacobians with respect to aerosol particles, $K_{354,nk}$ and $K_{388,nk}$, are computed at 354 and 388 nm, respectively, using OMI-reported surface albedo values (ρ_{354} and ρ_{388}), followed by a calculation of the albedo Jacobian $\frac{\partial I_{aer354}(\cancel{R_{388}}R'_{388})}{\partial R}$ at 354 nm.

To check this analytic Jacobian calculation in Eqns. (6)-(8), we compute the aerosol AI Jacobians using a finite difference (FD) method. Here, the derivative of AI as a function of aerosol concentration of a species, k, in layer n, is computed using

$$\frac{\partial AI}{\partial M_{nk}} = \frac{(AI-AI')}{(C_{nk}-C'_{nk})} \quad . \qquad (9)$$

Here C_{nk} and C_{nk}' are the baseline and perturbed aerosol concentrations, respectively, and AI and AI' are computed using C_{nk} and C_{nk}', respectively.

Figure 2b shows the comparison of Jacobians of dust aerosols estimated from the analytic and the FD solutions. Dust, smoke, ABF and sea salt aerosol concentrations as a function of altitude are shown in Figure 2a. To compute FD Jacobians with respect to dust aerosols, a 10% perturbation is introduced in the dust profiles. A very close match is found between analytic and FD Jacobians. This validates the analytical solution used in the study. The analytic solution is of course much faster, as a single call to VLIDORT will deliver all necessary Jacobians at one wavelength, as compared to 97 separate calls to VLIDORT with the FD calculation (baseline; 4 species perturbations in the 24-layer atmosphere).

3.2 The variational OMI AI assimilation system

The OMI AI assimilation system is based on AI simulations (with Jacobians) from the forward model. Two principles underlay the assimilation procedure. First, we assume that OMI AI is sensitive to UV-absorbing aerosol particles, such as NAAPS smoke and dust, or that only smoke and dust are injected high enough into the troposphere to impact AI. Therefore, innovations are limited to modifications of dust and smoke aerosol properties. For classes that do not strongly project onto AI, such as sea salt and ABF aerosols, aerosol concentrations are not modified during the process. Second, contributions of smoke/dust aerosols to AI (AI_{smoke} / AI_{dust}) prior to assimilation are estimated by multiplying smoke/dust aerosol concentrations from NAAPS with Jacobians of AI respective of smoke/dust aerosols. The ratio of AI innovation from smoke aerosols (ΔAI_{smoke}) to total AI innovation (ΔAI or OMI AI - AI_{naaps}) is assumed to be the ratio of AI_{smoke} to $AI_{smoke} + AI_{dust}$. The same assumption holds for dust aerosols.

Given these two principles, the overall design concept for the OMI AI assimilation can be expressed as

$$C^a = C^b +$$

$$\frac{P_{dust}\boldsymbol{H}_{dust}{}^T}{\boldsymbol{H}_{dust}{}^T P_{dust}\boldsymbol{H}_{dust}+R}[\text{y-H}(C^b)]\times\frac{\boldsymbol{H}_{\boldsymbol{dust}}C_{dust}{}^b}{\boldsymbol{H}_{\boldsymbol{dust}}C_{dust}{}^b+\boldsymbol{H}_{smk}C_{smk}{}^b}+$$

$$\frac{P_{smk}\boldsymbol{H}_{smk}{}^T}{\boldsymbol{H}_{smk}{}^T P_{smk}\boldsymbol{H}_{smk}+R}[\text{y-H}(C^b)]\times\frac{\boldsymbol{H}_{smk}C_{smk}{}^b}{\boldsymbol{H}_{\boldsymbol{dust}}C_{dust}{}^b+\boldsymbol{H}_{smk}C_{smk}{}^b}, \tag{10}$$

where C^b and C^a are NAAPS aerosol concentrations for the analysis and background fields, respectively, $C_{dust}{}^b$ and $C_{smk}{}^b$ are background NAAPS particle mass concentrations for dust and smoke, H(C) is the NAAPS forward model that links NAAPS ~~partcle~~particle mass concentrations to AI, and ***H*** is defined as ∂H(C)/∂C, which is the Jacobian matrix of AI with respect to aerosol concentrations. Y is the observed OMI AI, and Y- H(C^b) is the innovation of AI, representing the difference between observed and modeled AI values.

The $\frac{H_{dust}C_{dust}{}^{b}}{\boldsymbol{H_{dust}}C_{dust}{}^{b}+\boldsymbol{H}_{smk}C_{smk}{}^{b}}$ and $\frac{\boldsymbol{H}_{smk}C_{smk}{}^{b}}{\boldsymbol{H_{dust}}C_{dust}{}^{b}+\boldsymbol{H}_{smk}C_{smk}{}^{b}}$ terms are the fractional contribution of innovation from dust and smoke aerosol, respectively. These terms are estimated using NAAPS aerosol concentrations for relatively high aerosol loading cases (AOD > 0.15). For low aerosol loading (AOD < 0.15) as reported from NAAPS, it is possible that NAAPS could underestimate aerosol concentrations. Thus, the fractional contribution of innovations is assigned to 1 for the dominant aerosol type based on a NAAPS aerosol climatology (Zhang et al., 2008). Note that the term [y-H(C^{b})]× $\frac{\boldsymbol{H}_{dust}C_{dust}{}^{b}}{\boldsymbol{H_{dust}}C_{dust}{}^{b}+\boldsymbol{H}_{smk}C_{smk}{}^{b}}$ is in observational space. P_{dust} and P_{smk} are model error ~~spatial~~ covariance matrices for dust and smoke ~~(model space)~~ aerosols (e.g. Zhang et al., 2008; 2011; 2014). R is the observation-based error covariance ~~in model space~~. The $\frac{P_{dust}\boldsymbol{H}_{dust}{}^{T}}{\boldsymbol{H}_{dust}{}^{T}P_{dust}\boldsymbol{H}_{dust}+R}$ [y-H(C^{b})]× $\frac{\boldsymbol{H_{dust}}C_{dust}{}^{b}}{\boldsymbol{H_{dust}}C_{dust}{}^{b}+\boldsymbol{H}_{smk}C_{smk}{}^{b}}$ and $\frac{P_{smk}\boldsymbol{H}_{smk}{}^{T}}{\boldsymbol{H}_{smk}{}^{T}P_{smk}\boldsymbol{H}_{smk}+R}$ [y-H(C^{b})]× $\frac{\boldsymbol{H}_{smk}C_{smk}{}^{b}}{\boldsymbol{H_{dust}}C_{dust}{}^{b}+\boldsymbol{H}_{smk}C_{smk}{}^{b}}$ terms represent the estimated increments in model space.

The background error covariance matrix is constructed from modeled error variances and error correlations, following the methodology in previous studies (Zhang et al., 2008; 2011). The horizontal background error covariance is generated using the second-order regressive function (SOAR), as shown in Equation 11 (Zhang et al., 2008), or

$$C(x,y) = (1 + R_{xy}/L)\exp(-\frac{R_{xy}}{L}) \quad . \qquad (11)$$

Here, x and y are two given locations, and R_{xy} is the great circle distance. L is the averaged error correlation length and is set to 200 km based on Zhang et al. (2008). Similarly, the vertical error correlation between two pressure levels p_1 and p_2 is also based on the SOAR function, this time in pressure space, based on Zhang et al., (2011), is

$$C(p_1,p_2) = [1 + \left|\int_{p_1}^{p_2} \frac{\mathrm{dln}\, p}{L}\right|]e^{-\left|\int_{p_1}^{p_2} \frac{\mathrm{dln}\, p}{L}\right|} \quad . \qquad (12)$$

Here, L is a unit-less number representing vertical correlation length and is set to 0.015.

The horizontal error variance is based on the RMS error of aerosol concentrations, which is arbitrarily set to 100 μg/m^3 for near-surface dust aerosols (ground to 700 hPa). The RMS error of dust aerosol mass is assumed to decrease as altitude increases, and is set to 50%, 25%, and 1% of the near-surface values for 500-700, 350-500 and 70-350 hPa respectively. Note that different aerosol species have different mass extinction values. Here we assume the modeled error in aerosol extinction is the same for different aerosol species and thus, the RMS error of smoke aerosol concentration is scaled by mass extinction cross section ratio between smoke and dust aerosols. The observational errors are assumed to be non-correlated in this study (e.g. Zhang et al., 2008). OMI AI values over cloud-free and cloudy skies are used in the study and therefore, RMS errors of AI are required for both these situations. Note, as suggested by Yu et al. (2012), for the same above cloud CALIOP AOD, variations in AI are found to be of the order of 1 for cloud optical depth changing from 2 to 20. Thus, we assume the RMS error of OMI AI is 0.5 for cloud-free skies, increasing linearly with cloud fraction up to a value of 1 for the 100% overcast.

Lastly, we assume that detectable UV absorbing aerosols have AI values larger than 0.8 (e.g. Torres et al., 2013). Therefore, for regions with OMI AI values larger than 0.8, UV absorbing aerosol particles can both be added or removed from air columns based on innovations, which are the differences between OMI reported and simulated AI values. For regions with OMI AI values less than 0.8, innovations are only used to remove UV absorbing aerosol particles from air columns.

4.0 System evaluation & discussion

4.1 Evaluating the performance of the AI assimilation system over Africa

Using two months of OMI data (July-August, 2007), the performance of OMI AI assimilation was evaluated around the Africa region (20°S-40°N; 6~~1~~0°W-5~~6~~0°E). The study region was chosen to examine the performance of OMI AI data assimilation over bright surfaces such as the deserts of North Africa, as well as study aerosol advection over clouds, in this case smoke off the west coast of Southern Africa. In this demonstration, two NAAPS runs were performed for the period of July 1 to August 31, 2007, one with and one without the use of OMI AI assimilation (AI-DA run). Both runs were initialized with the use of NAAPS reanalysis data at 0000 UTC 1 July and do not include any other form of aerosol assimilation.

Figure 3a shows the true color composite from Aqua MODIS for July 28, 2007 over the study region that is obtained from the NASA world view site (https://worldview.earthdata.nasa.gov/; last accessed June 2020). Visible in the image are the dust plumes from North Africa transported to the Atlantic Ocean, and smoke plumes from Central and Southern Africa transported to the west coast of South Africa. As indicated by the aggregated OMI AI data for 1200 UTC 28 July 2007 (Figure 3b), dust plumes from North Africa are transported to the North corner of the west coast of North Africa. Smoke plumes are also visible in the OMI AI plot in Southern Africa and are transported to the west coast and over the Atlantic. Comparing Figure 3a and Figure 3b, smoke plumes, as identified from OMI, are also found over cloudy regions as indicated from the MODIS visible imagery. Note that Figure 3b shows the OMI AI data used in the assimilation process and again, AI retrievals over both cloud free and cloudy conditions are included as suggested by Figure 3b.

Figure 3c is the 1200 UTC 28 July 2007 NAAPS AOD product from the natural run. In comparison, Figure 3d shows the same situation, this time with the use of OMI AI data assimilation. Comparing 3b with 3d, dust and smoke aerosol patterns as shown from OMI AI resemble more closely the NAAPS AOD fields after AI assimilation. Over the northeast coast of Africa, heavy aerosol plumes, as hinted at in NAAPS AOD from the natural run (Figure 3c), cover larger spatial areas than those inferred from OMI AI data. In comparison, NAAPS AOD patterns from the OMI AI data assimilation cycle closely resemble aerosol patterns as suggested from OMI AI data. Also shown in Figures 3e and 3f are the simulated AI using NAAPS data from the natural and OMI AI DA runs (data from Figures 3c and 3d) respectively. Clearly, with the use of NAAPS data from the natural run, simulated OMI AI are overestimated in comparison with OMI AI data (Figure 3b). Simulated AI patterns with the used of NAAPS data from the OMI AI DA run rather closely resemble AI patterns from the OMI data, again, indicating the OMI AI DA system is functioning reasonably as designed.

The performance of AI-DA is also evaluated using OMI AI for the whole study period, as shown in Figure 4. These data are constructed using collocated OMI AI and NAAPS data according to the conditions introduced in Sec. 3. Here, Figures 4a and 4e are spatial distributions of two-monthly averaged (July and August 2007) AODs for NAAPS AI-DA and natural runs, respectively. Figure 4b is the spatial distribution of the simulated AI using NAAPS data from AI-DA runs, and Figure 4c is the spatial distribution of OMI AI for the two-month period. Figures 4f and 4g show similar plots to those in Figures 4c and 4d, but this time for NAAPS natural runs. While simulated AI values from NAAPS natural runs (Figure 4f) are overestimated compared to OMI AI values (Figure 4g) for the study region, the patterns of simulated AI from NAAPS AI-DA runs (Figure 4b) are similar to patterns shown from OMI AI (Figure 4c). This is also seen from Figure 4d, which is the difference between simulated AI from NAAPS AI-DA runs and OMI AI. In contrast with the situation in Figure 4d, Figure 4h, which is the difference between simulated AI from NAAPS natural runs and OMI AI, shows much larger differences in AI values.

While it is not too difficult to make the model mimic the AI product, proof of real skill lies in any improvements to AOD calculations. To this end, the performance of OMI AI assimilation was evaluated with the use of AERONET data. Figure 5a shows the inter-comparison of NAAPS AOD versus AERONET AOD at 0.55 μm. A total of 1443~~50~~ collocated pairs of NAAPS and AERONET data were compiled for the study region over the two months test period. Comparing with AERONET data, NAAPS AOD from the natural run had a correlation of 0.68~~4~~, a mean absolute error in AOD of 0.154~~7~~, and an RMSE of 0.220~~5~~. In comparison, with AI assimilation, NAAPS AOD correlations to AERONET increased to 0.74~~2~~ (Figure 5b), the absolute error reduced to 0.104~~2~~, and RMSE reduced to 0.156~~8~~, both roughly a 30% reduction. Note that AERONET AOD values are only available for lines-of-sight that are free of cloud presence for the sun photometer instruments. Also, the slope of AERONET versus NAAPS AOD is 0.87~~93~~ for the NAAPS natural runs, and a similar slope of 0.84~~92~~ is found for the NAAPS AI-DA runs.

4.2 Inter-comparison with AOD data assimilation

Typically, NAAPS reanalyses are constructed through assimilation of MISR and MODIS aerosol products (NAAPS AOD assimilation). Thus, the performances of NAAPS AOD and AI-DA assimilations are compared against AERONET data. Figure 5c shows the comparison of AERONET AOD and NAAPS AOD after AOD assimilation, while Figure 5b shows a similar plot but using NAAPS data from AI-DA. Note that the same version of the NAAPS model with the same temporal and spatial resolutions, and driven by the same meteorological data, were used in constructing Figure 5 and thus the differences in Figures 5a, 5b and 5c only result from different aerosol data assimilation methods implemented (no data assimilation for the natural run). A better correlation between AERONET and NAAPS data of 0.79~~82~~ is~~and a slope of 1.01 are~~ found using AOD data assimilation. In comparison, the correlation is 0.74~~2~~ ~~and the slope is 0.92~~ for the AI-

DA runs. Slightly better RMSE (0.140~~5~~ versus 0.156~~8~~) and absolute error (0.095~~11~~ versus 0.104~~2~~) values are also found for the AOD data assimilation runs. This result is not surprising as OMI AI provides only a proxy for aerosol properties while passive-based AOD retrievals are often considered as a more reliable parameter for representing column-integrated aerosol properties. But still, the evaluation efforts are over cloud-free line-of-sight as detected from AERONET, AI DA may further assist traditional AOD data assimilation by ~~proving~~providing AI assimilation over cloudy regions.

4.3 Sensitivity test

As mentioned in Section 3, aerosol properties for non-smoke aerosol types were obtained from the NASA GEOS-5 model (e.g. Colarco et al., 2014; Buchard et al., 2015). Yet, different smoke aerosol SSA values are used in this study, as values for central Africa have a strong seasonal dependency (e.g. Eck et al., 2013). While SSA values of 0.85 and 0.86 are used for the 354 and 388 nm channels, respectively, in our study, we have also examined the sensitivity of simulated OMI AI with respect to differing SSA values (Figure 6). Figures 6a-c show the simulated AI at 1200 UTC 28 July 2007 using NAAPS reanalysis data (Lynch et al, 2016) for three scenarios: SSA values at 354 and 388 nm of 0.84 and 0.84 (Figure 6a), 0.85 and 0.85 (Figure 6b) and 0.86 and 0.86 (Figure 6c). Over the central Africa area, where smoke plumes are expected, simulated OMI AI patterns are similar for Figures 6a and 6b, but reduced values in AI are found when using higher SSA values of 0.86 at both 354 and 388 nm. This is further confirmed by the averaged AI for the smoke region over central Africa (14.5~~-0.5~~° to ~~-1~~4~~5~~0.5° S latitude and 10.5° to 30.5° E longitude; indicated using the black box in Figure 6f) of 0.96, 0.94 and 0.78 for Figures 6a, 6b and 6c respectively.

Figures 6d-f show the sensitivity for adjustments of the SSA values at 388nm while maintaining a fixed SSA value of 0.85 at 354 nm. Here the SSA values at 388 nm are set to 0.85, 0.855 and 0.86 for Figures 6d, 6e and 6f respectively. Interestingly, the spectral dependence of SSA seems to affect the simulated AI significantly, and this phenomenon has also been reported by previous studies (e.g. Hammer et al., 2017). The averaged AI values over central Africa (again, indicated by the black box in Figure 6f) are 0.94, 1.11 and 1.32 for 388 nm SSAs of 0.85, 0.855 and 0.86, respectively. This exercise suggests that simulated AI is a strong function of SSA, so that both the spectral dependence of SSA values at 354 and 388 nm and reliable SSA values are needed on a regional basis for future applications.

Interestingly, although simulated AI values are significantly affected by perturbing SSA values as shown in Figure 6, less significant impacts are observed for NAAPS AOD. This is found by running the OMI AI DA for 1200UTC, July 28, 2015 for SSA values used in generating Figure 6. For example, for the black box highlighted region in Figure 6f, the averaged values for the simulated OMI AI are 0.96, 0.94 and 0.78 for using SSA values at 354 / 388 nm channels of 0.84 / 0.84, 0.85 / 0.85 and 0.86 / 0.86, respectively. The corresponding NAAPS AODs are found to be 0.559, 0.560 and 0.585 after OMI AI DA, which is a change of less than 5%. Similar, by fixing the SSA value of the 354 nm channel as 0.85 and perturbing SSA values at 388 nm from 0.85 to 0.86, a ~30% change is found in simulated OMI AI (from 0.94 to 1.32), yet a ~10% change is found for the NAAPS AOD (from 0.560 to 0.504) after OMI AI DA.

It is also of interest to investigate the changes in aerosol vertical distributions due to the OMI AI DA. For this exercise, we selected the 1200 UTC 28 July 2007 case and compared vertical distributions of smoke and dust aerosols near the peak AI value of the smoke plume (9.5°S and 20.5°E) for the NAAPS natural and AI DA runs (Figure 7a). Note that the differences between

OMI DA and natural runs as shown in Figure 7 are essentially an integrated effect of OMI AI DA from 00Z, July 01 to 12 Z, July 28, 2007. As shown in Figure 7a, the corrections to dust and smoke aerosol concentrations from the AI DA system seem to be systematic changes across the majority of vertical layers, instead of moving dust or smoke aerosol plumes vertically. As dust aerosol concentrations are reduced at all layers and a systematic correction to smoke aerosol concentrations, although non-linear, is also observed. AI assimilation helps reduce the amount of upper troposphere dust (likely to be artifact) but does change the layer centroid slightly upwards. We have also evaluated NAAPS vertical distributions near a peak dust plume region (25.5°N and 12.5°W) for the 12Z 28 July 2007 case as shown in Figure 7b. Similar to Figure 7a, a non-linear correction to dust aerosol concentrations is also observed across the vertical domain.

4.4 Issues and discussions

The OMI AI data assimilation system is a proxy for all-sky, all-band modeling system radiance assimilation. It contains all the necessary components for such radiance assimilation, including a forward model for simulating radiances and AI values and their Jacobians, based on a full vector linearized radiative transfer model called for every observation. Therefore, the computational burden is a direct issue associated with the deployment of calls to a radiative transfer model for each observation. For the study area in this work, after binning OMI AI data into a 1°×1° (Latitude/Longitude) product, it still takes about ~1 CPU day for NAAPS to run for one month of model time. In comparison, the time scale for running AOD assimilation for 1 month is at the hourly ~~scale~~level. Clearly, there will be an unavoidable computational burden of some sort for OMI AI assimilation and by extension, for future radiance assimilation in the UV/visible spectrum for aerosol analyses. Performance enhancement methods, such as parallel processing (the VLIDORT software is thread-safe and can be used in parallel environments such as OpenMP), or fast look-up-table extraction based on neural-networks and trained data sets of forward simulation, must be explored in order to enable such assimilation applications in near real time on a global scale.

In contrast with the assimilation of retrieved aerosol properties, both aerosol absorption and scattering need to be accounted for when assimilating radiance or OMI AI in the UV spectrum. This requires the inclusion of more dynamic aerosol optical properties into the data assimilation process, and properties that vary with region and season. As noted already, even for biomass burning aerosols over South Africa, lower single scattering albedo values were found at earlier stages of burning seasons (e.g. Eck et al., 2013). A look-up-table of aerosol optical properties as functions of region and season will be needed for global implications of OMI AI as well as future radiance assimilation for aerosol modeling.

OMI AI is sensitive to above-cloud UV-absorbing aerosols (e.g. Yu et al., 2012; Alfaro-Contreras et al., 2014), and therefore, OMI AI values over cloudy scenes were also used in this study. However, OMI AI cannot be used to infer aerosol properties for aerosol plumes beneath a cloud deck. For regions with high clouds, the use of OMI AI data assimilation will likely result in an underestimation of AOD as below-cloud aerosol plumes are not accounted for. Therefore, only OMI AI data over low cloud scenes are to be used for aerosol assimilation efforts. In addition, although some quality assurance steps were applied in this study for the OMI AI data, lower AI values were observed over glint regions near the west coast of Africa. Abnormally high OMI AI values are also seen near the Arctic region - this may be related to the presence of floating ice sheets. Thus, innovative and detailed data screening and quality assurance steps are needed to exclude potentially noisy OMI AI retrievals and for further application of OMI AI data assimilation on a global scale.

Even with these known issues, OMI AI assimilation as presented in the study illustrates a new method for assimilating non-conventional aerosol products. Bearing in mind that OMI AI assimilation is essentially radiance assimilation in the UV spectrum, this study demonstrates the potential of directly assimilating satellite radiance in the UV/visible spectrum for aerosol modeling and analyses.

5.0 Conclusions

The OMI aerosol index (AI), which measures the differences between simulated radiances over Rayleigh sky and observed radiances at 354 nm, has been used to detect the presence of absorbing aerosols over both dark and bright surfaces. We have constructed a new assimilation system, based on the VLIDORT radiative transfer code as the major component of the forward model, for the direct assimilation of OMI AI. The aim is to improve accuracies of aerosol analyses over bright surfaces such as cloudy regions and deserts.

The performance of the OMI AI data assimilation system was evaluated over South-Central and Northern Africa regions for the period of 01 July -31 August 2007. This evaluation was done through inter-comparing NAAPS analyses with and without the inclusion of OMI AI data assimilation. Besides cloud-free AI retrievals over dark surfaces, OMI AI retrievals over desert regions and over areas were also considered. When compared against AERONET data, a total of ~29~~8~~% reduction in Root-Mean-Square-Error (RMSE) with a ~32% reduction in absolute error were found for NAAPS analyses with the use of OMI AI assimilation. Also, NAAPS analyses with the inclusion of OMI AI data assimilation show similar aerosol patterns to those in the OMI AI data sets, showing that our OMI AI data assimilation system works as expected.

This study also suggests that NAAPS analyses with OMI AI data assimilation cannot out-perform NAAPS reanalyses data that were incorporated with MODIS and MISR AOD assimilation, and validated against AERONET data. This is not surprising, as OMI AI is only a proxy for the AOD and is sensitive to other factors such as surface albedo and aerosol vertical distribution. Also, AERONET data are only available over cloud-free field of views, so the performance of our OMI AI data assimilation system over cloudy regions has not been evaluated.

There are a number of issues arising from our study. For example, aerosol optical properties are needed for the OMI AI-DA system - these have strong regional and temporal signatures that need to be carefully quantified before applying them to the AI-DA on a global scale. Also, OMI AI retrievals are rather noisy and contain known and unknown biases. Abnormally high OMI AI values are found over mountain regions as well the polar regions. Sporadic high AI values are also known to occur, for reasons that are still not properly understood. Even though quality assurance steps were proposed in this study, detailed analysis of OMI AI data are needed for future implementation of OMI AI data assimilation for aerosol studies.

Lastly, AI values are derived from radiances and thus, the AI-DA system presented in the study can be thought of as a radiance assimilation system for the UV spectrum. This is because the AI-DA system contains all necessary components for radiance assimilation, based on a forward model for calculating not only simulated satellite radiances, but also the aerosol-profile Jacobians of these radiance, both quantities as functions of observation conditions. This study is among the first attempts at radiance assimilation at the UV spectrum and indicates the future potential for direct radiance assimilation at the UV and visible spectra for aerosol analyses and forecasts.

Author contributions. All authors contributed to the overall design of the study. Authors JZ and RS coded the system. Author JSR provided valuable suggestions though the study. Author PX assist with the evaluation of the system.

Code and data availability: The OMI data assimilation scheme (V1.0) is constructed using VLIDORT and NAVDAS-AOD for NAAPS analyses and forecasts. The VLIDORT radiative transfer mode is a property of RT Solutions Inc. The VLIDORT code is publicly available, and comes with a standard GNU public license, through direct contact with RT Solutions Inc. (http://www.rtslidort.com/mainprod_vlidort.html). Both NAAPS and NAVDAS-AOD are proprietary to Naval Research Laboratory, United States Department of the Navy. Nevertheless, both NAAPS and NAVDAS-AOD are well documented in past studies (e.g. Lynch et al., 2016; Zhang et al., 2008; 2011; 2014; Rubin et al., 2017) and we have made every effort to thoroughly report our methods so that they may be replicated. AOD fields from the NAAPS OMI AI DA and natural runs over the study region and period are shared as the supplement to the paper for readers who are interested. The NAAPS reanalysis data are available from the USGODAE web site (https://nrlgodae1.nrlmry.navy.mil/cgi-bin/datalist.pl?dset=nrl_naaps_reanalysis&summary=Go**.** The OMI OMAERUV data are available from the NASA's Goddard Earth Sciences *Data* and Information Services Center (*GES DISC*; https://disc.gsfc.nasa.gov/datasets/OMAERUV_003/summary). AERONET data are obtained from the NASA AERONET webpage (https://aeronet.gsfc.nasa.gov/).

~~**Code and data availability:** The VLIDORT radiative transfer model is available to the public through contacting RT solutions Inc. (http://www.rtslidort.com/mainprod_vlidort.html). The NAAPS model belongs to the Naval Research Laboratory and is not publically available. The~~

~~NAAPS reanalysis data are available from the USGODAE web site (https://nrlgodae1.nrlmry.navy.mil/cgi-bin/datalist.pl?dset=nrl_naaps_reanalysis&summary=Go. The OMI OMAERUV data are available from the NASA's Goddard Earth Sciences Data and Information Services Center (GES DISC; https://disc.gsfc.nasa.gov/datasets/OMAERUV_003/summary). AERONET data are obtained from the NASA AERONET webpage (https://aeronet.gsfc.nasa.gov/).).~~

Competing interests. The authors claim no competing interests.

Acknowledgements. We thank the NASA AERONET group for the sun-photometer data used in the study. “We acknowledge the use of imagery from the NASA Worldview application (https://worldview.earthdata.nasa.gov), part of the NASA Earth Observing System Data and Information System (EOSDIS).”

Financial support. This study is supported by the NASA grant NNX17AG52G. Co-Author JSR was supported by the Office of Naval Research Code 322.

References:

Alfaro-Contreras, R., Zhang, J., Campbell, J. R., and Reid, J. S.: Investigating the frequency and trends in global above-cloud aerosol characteristics with CALIOP and OMI, Atmos. Chem. Phys., 16, 47-69, doi:10.5194/acp-16-47-2016, 2016.

Alfaro-Contreras R., Zhang, J., Campbell, J. R., Holz, R. E., and Reid, J. S.: Evaluating the Impact of Aerosol Particles above Cloud on Cloud Optical Depth Retrievals from MODIS, J. Geophys. Res. Atmos., 119, 5410–5423, doi:10.1002/2013JD021270, 2014.

Benedetti, A., et al., Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation, J. Geophys. Res., 114, D13205, doi:10.1029/2008JD011115, 2009.

Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., Kumagai, Y., Miyakawa, T., Murata, H., Ohno, T., Okuyama, A., Oyama, R., Sasaki, Y., Shimazu, Y., Shimoji, K., Sumida, Y., Suzuki, M., Taniguchi, H., Tsuchiyama, H., Uesawa, D., Yokota, H., & Yoshida, R.: An Introduction to Himawari-8/9- Japan's New-Generation Geostationary Meteorological Satellites. Journal of the Meteorological Society of Japan. Ser. II, 94*(*2*)*, 151–183. https://doi.org/10.2151/jmsj.2016-009, 2016.

Buchard, V., Randles, C., Silva, A., Darmenov, A., Colarco, P., Govindaraju, R., Ferrare, R., Hair, J., Beyersdorf, A., Ziemba, L., Yu, H.: The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part II: Evaluation and Case Studies Journal of Climate https://dx.doi.org/10.1175/jcli-d-16-0613.1, 2017.

Buchard, V., da Silva, A. M., Colarco, P. R., Darmenov, A., Randles, C. A., Govindaraju, R., Torres, O., Campbell, J., and Spurr, R.: Using the OMI aerosol index and absorption aerosol optical depth to evaluate the NASA MERRA Aerosol Reanalysis, Atmos. Chem. Phys., 15, 5743–5760, https://doi.org/10.5194/acp-15-5743-2015, 2015.

Cochrane, S. P., Schmidt, K. S., Chen, H., Pilewskie, P., Kittelman, S., Redemann, J., LeBlanc, S., Pistone, K., Kacenelenbogen, M., Segal Rozenhaimer, M., Shinozuka, Y., Flynn, C., Platnick, S., Meyer, K., Ferrare, R., Burton, S., Hostetler, C., Howell, S., Freitag, S., Dobracki, A., and Doherty, S.: Above-cloud aerosol radiative effects based on ORACLES 2016 and ORACLES 2017 aircraft experiments, Atmos. Meas. Tech., 12, 6505–6528, https://doi.org/10.5194/amt-12-6505-2019, 2019.

Colarco, P. R., Nowottnick, E. P., Randles, C. A., Yi, B., Yang, P., Kim, K.-M., Smith, J. A. and Bardeen, C. G.: Impact of radiatively interactive dust aerosols in the NASA GEOS-5 climate model: sensitivity to dust particle shape and refractive index. Journal of Geophysical Research: Atmospheres, 119(2), 753– 786. https://doi.org/10.1002/2013JD020046, 2014.

Dai, A.: Precipitation characteristics in eighteen coupled climate models. Journal of Climate, 19, 4605– 4630, 2006.

Eck, T. F., Holben, B. N., Reid, J. S., Mukelabai, M. M., Piketh, S. J., Torres, O., Jethva, H. T., Hyer, E. J., Ward, D. E., Dubovik, O., and Sinyuk, A.: A seasonal trend of single scattering albedo in southern African biomass-burning particles: Implications for satellite products and estimates of emissions for the world's largest biomass-burning source, J. Geophys. Res.-Atmos., 118, 6414–6432, 2013.

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.

Hammer, M. S., Martin, R. V., van Donkelaar, A., Buchard, V., Torres, O., Ridley, D. A., and Spurr, R. J. D.: Interpreting the ultraviolet aerosol index observed with the OMI satellite instrument to understand absorption by organic aerosols: implications for atmospheric oxidation and direct radiative effects, Atmos. Chem. Phys., 16, 2507–2523, https://doi.org/10.5194/acp-16-2507-2016, 2016.

Herman, J. R., Bhartia, P. K., Torres, O., Hsu, C., Seftor, C., and Celarier, E.: Global distribution of UV-absorbing aerosols from Nimbus 7/TOMS data, J. Geophys. Res., 102(D14), 16911–16922, doi:10.1029/96JD03680., 1997.

Holben, B. N., and coauthors: AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization. Remote Sensing of Environment, 66(1), 1–16. https://doi.org/10.1016/S0034-4257(98)00031-5, 1998.

Hogan, T.F., Liu, M., Ridout, J. A., Peng, M. S., Whitcomb, T. R., Ruston, B. C., Reynolds, C. A., Eckermann, S. D., Moskaitis, J. R., Baker, N. L., McCormack, J. P., Viner, K. C., McLay, J. G., Flatau, M. K., Xu, L., Chen, C., and Chang, S. W.: The Navy Global Environmental Model, Oceanography, 27(3):116–125, https://doi.org/10.5670/oceanog.2014.73., 2014.

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

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

Hsu, N. C., Gautam, R., Sayer, A. M., Bettenhausen, C., Li, C., Jeong, M. J., Tsay, S.-C., and Holben, B. N.: Global and regional trends of aerosol optical depth over land and ocean using SeaWiFS measurements from 1997 to 2010, Atmos. Chem. Phys., 12, 8037–8053, https://doi.org/10.5194/acp-12-8037-2012, 2012.

Hyer, E. J., Reid, J. S., and Zhang, J.: An over-land aerosol optical depth data set for data assimilation by filtering, correction, and aggregation of MODIS Collection 5 optical depth retrievals, Atmos. Meas. Tech., 4, 379-408, doi:10.5194/amt-4-379-2011, 2011.

Joyce, J. R., Janowiak, E. J., Arkin, P. A., and Xie, P.: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution, J. Hydrometeor., 5, 487-503, 2004.

Kahn, R. A., Gaitley, B. J., Garay, M. J., Diner, D. J., Eck, T. F., Smirnov, A. and Holben B. N.: Multiangle Imaging SpectroRadiometer global aerosol product assessment by comparison with the Aerosol Robotic Network, J. Geophys. Res., 115, *D23209, doi:*10.1029/2010JD014601., 2010.

Lee, J., Hsu, N.C., Bettenhausen, C., Sayer, A. M., Seftor, C. J., and Jeong, M.-J.: Retrieving the height of smoke and dust aerosols by synergistic use of VIIRS, OMPS, and CALIOP observations, J. Geophys. Res. Atmos.,120, 8372–8388,doi:10.1002/2015JD023567, 2015.

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, doi:10.5194/amt-6-2989-2013, 2013.

Li, Z., Zang, Z., Li, Q. B., Chao, Y., Chen, D., Ye, Z., Liu, Y., and Liou, K. N.: A three-dimensional variational data assimilation system for multiple aerosol species with WRF/Chem and an application to $PM_{2.5}$ prediction, Atmos. Chem. Phys., 13, 4265–4278, https://doi.org/10.5194/acp-13-4265-2013, 2013.

Lynch, P., Reid, J. S., Westphal, D. L., Zhang, J., Hogan, T. F., Hyer, E. J., Curtis, C. A., Hegg, D. A., Shi, Y., Campbell, J. R., Rubin, J. I., Sessions, W. R., Turk, F. J., and Walker, A. L.: An 11-year global gridded aerosol optical thickness reanalysis (v1.0) for atmospheric and climate sciences, Geosci. Model Dev., 9, 1489-1522, https://doi.org/10.5194/gmd-9-1489-2016, 2016.

Pistone, K., Redemann, J., Doherty, S., Zuidema, P., Burton, S., Cairns, B., Cochrane, S., Ferrare, R., Flynn, C., Freitag, S., Howell, S. G., Kacenelenbogen, M., LeBlanc, S., Liu, X., Schmidt, K. S., Sedlacek III, A. J., Segal-Rozenhaimer, M., Shinozuka, Y., Stamnes, S., van Diedenhoven, B., Van Harten, G., and Xu, F.: Intercomparison of biomass burning aerosol optical properties from in situ and remote-sensing instruments in ORACLES-2016, Atmos. Chem. Phys., 19, 9181–9208, https://doi.org/10.5194/acp-19-9181-2019, 2019.

Randles, C., Silva, A., Buchard, V., Colarco, P., Darmenov, A., Govindaraju, R., Smirnov, A., Holben, B., Ferrare, R., Hair, J., Shinozuka, Y., Flynn, C.: The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part I: System Description and Data Assimilation Evaluation, Journal of Climate https://dx.doi.org/10.1175/jcli-d-16-0609.1, 2017.

Reid, J. S., Hyer, E. J., Prins, E. M., Westphal, D. L., Zhang, J., Wang, J., Christopher, S. A., Curtis, C. A., Schmidt, C. C., Eleuterio, D. P., Richardson, K. A., and Hoffman, J. P.: Global monitoring and forecasting of biomass-burning smoke: Description and lessons from the Fire Locating and Modeling of Burning Emissions (FLAMBE) program, IEEE J. Sel. Top. Appl., 2, 144–162, 2009.

Rubin J. I., Reid, J. S., Hansen, J. A., Anderson, J. L., Holben, B. N., Lynch, P., Westphal, D. L., and Zhang, J.: Assimilation of AERONET and MODIS AOT observations using Variational and Ensemble Data Assimilation Methods and Its Impact on Aerosol Forecasting Skill, J. Geophys. Res., DOI: 10.1002/2016JD026067, 2017.

Saide, P. E., Carmichael, G. R., Liu, Z., Schwartz, C. S., Lin, H. C., da Silva, A. M., and Hyer, E.: Aerosol optical depth assimilation for a size-resolved sectional model: impacts of observationally constrained, multi-wavelength and fine mode retrievals on regional scale analyses and forecasts, Atmos. Chem. Phys., 13, 10425–10444, https://doi.org/10.5194/acp-13-10425-2013, 2013.

Schutgens, N. A. J., Miyoshi, T., Takemura, T., and Nakajima, T.: Applying an ensemble Kalman filter to the assimilation of AERONET observations in a global aerosol transport model, Atmos. Chem. Phys., 10, 2561–2576, https://doi.org/10.5194/acp-10-2561-2010, 2010.

Schwartz, C. S., Liu, Z., Lin, H.-C., and McKeen, S. A.: Simultaneous three-dimensional variational assimilation of surface fine particulate matter and MODIS aerosol optical depth, J. Geophys. Res., 117, D13202, doi:10.1029/2011JD017383, 2012.

Sekiyama, T. T., Tanaka, T. Y., Shimizu, A., and Miyoshi, T.: Data assimilation of CALIPSO aerosol observations, Atmos. Chem. Phys., 10, 39–49, https://doi.org/10.5194/acp-10-39-2010, 2010.

Sessions, W. R., Reid, J. S., Benedetti, A., Colarco, P. R., da Silva, A., Lu, S., Sekiyama, T., Tanaka, T. Y., Baldasano, J. M., Basart, S., Brooks, M. E., Eck, T. F., Iredell, M., Hansen, J. A., Jorba, O. C., Juang, H.-M. H., Lynch, P., Morcrette, J.-J., Moorthi, S., Mulcahy, J., Pradhan, Y., Razinger, M., Sampson, C. B., Wang, J., and Westphal, D. L.: Development towards a global operational aerosol consensus: basic climatological characteristics of the International Cooperative for Aerosol Prediction Multi-Model Ensemble (ICAP-MME), Atmos. Chem. Phys., 15, 335–362, https://doi.org/10.5194/acp-15-335-2015, 2015.

Shi Y., Zhang J., J.S. Reid, B. Liu, E.J. Hyer, Critical evaluation of cloud contamination in the MISR aerosol products using MODIS cloud masking products, Atmos. Meas. Tech., 7, 1791-1801, doi:10.5194/amt-7-1791-2014, 2014.

Spurr, R. J. D., and Christi. M.: The LIDORT and VLIDORT Linearized Scalar and Vector Discrete Ordinate Radiative Transfer Models: An Update for the last 10 Years, Light Scattering Reviews, Volume 12, ed. A. Kokhanovsky, Springer, 2019.

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. Transfer, 102(2), 316-342, doi:10.1016/j.jqsrt.2006.05.005, 2006.

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

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., Jethva, H., and Bhartia,P. K.: Retrieval of Aerosol Optical Depth above Clouds from OMI Observations: Sensitivity Analysis and Case Studies. *J. Atmos. Sci.,* 69, 1037–1053, https://doi.org/10.1175/JAS-D-11-0130.1, 2012.

Torres, O., Tanskanen, A., Veihelmann, B., Ahn, C., Braak, R., Bhartia, P. K., Veefkind, P., Levelt, P.: Aerosols and surface UV products from Ozone Monitoring Instrument observations: An overview. Journal of Geophysical Research, 112, D24S47. https://doi.org/10.1029/2007JD008809, 2007.

Uno, I., Yumimoto, K., Shimizu, A., Hara, Y., Sugimoto, N., Wang, Z., Liu, Z., and Winker D. M.: 3D structure of Asian dust transport revealed by CALIPSO lidar and a 4DVAR dust model, Geophys. Res. Lett., *35*, L06803, doi:10.1029/2007GL032329, 2008.

Westphal, D. L., Toon, O. B., and Carlson, T. N.: A case study of mobilization and transport of Saharan dust. J. Atmos. Sci., 45, 2145-2175, 1988.

Witek M. L., Flatau, P. J., Quinn, P. K., Westphal, D. L., Global sea-salt modeling: results and validation against multicampaign shipboard measurements, J. Geophys. Res., 112, p. D08215, 10.1029/2006JD007779, 2007.

Xian, P., Reid, J.S., Turk, J.F., Hyer, E.J. and Westphal, D.L.: Impact of modeled versus satellite measured tropical precipitation on regional smoke optical thickness in an aerosol transport model, Geophysical Research Letters, 36, L16805, https://doi.org/10.1029/2009GL038823, 2009.

Yu, H., Zhang. Y., Chin M., Liu Z., Omar A., Remer L. A. Yang Y., Yuan T. and Zhang J.: An Integrated Analysis of Aerosol above Clouds from A-Train Multi-sensor Measurements, Remote Sens. Environ., 121, 125–131, https://doi.org/10.1016/j.rse.2012.01.011, 2012.

Yumimoto, K., Uno, I., Sugimoto, N., Shimizu, A., Liu, Z., and Winker, D. M.: Adjoint inversion modeling of Asian dust emission using lidar observations, Atmos. Chem. Phys., 8, 2869–2884, https://doi.org/10.5194/acp-8-2869-2008, 2008.

Zhang J., Reid, J. S., Campbell, J. R., Hyer, E. J., and Westphal, D. L.: Evaluating the Impact of Multi-Sensor Data Assimilation on A Global Aerosol Particle Transport Model. J. Geophys. Res. Atmos., 119, 4674–4689, doi:10.1002/2013JD020975, 2014.

Zhang, J., Campbell, J. R., Reid, J. S., Westphal, D. L., Baker, N. L., Campbell, W. F., and Hyer, E. J.: Evaluating the impact of assimilating CALIOP-derived aerosol extinction profiles on a global mass transport model, Geophys. Res. Lett., Vo. 38, No. 14, L14801, doi: 10.1029/2011GL04773, 2011.

Zhang, J. and Reid, J. S.: A decadal regional and global trend analysis of the aerosol optical depth using a data-assimilation grade over-water MODIS and Level 2 MISR aerosol products, Atmos. Chem. Phys. Discuss., 10, 18879-18917, doi:10.5194/acpd-10-18879-2010, 2010.

Zhang, J., Reid, J. S., Westphal, D. L., Baker, N. L., and Hyer, E. J.: A system for operational aerosol optical depth data assimilation over global oceans, J. Geophys. Res., 113, No. D10, D10208, doi: 10.1029/2007JD009065, 2008.

Zhang, J. and Reid., J.S.: MODIS Aerosol Product Analysis for Data Assimilation: Assessment of Level 2 Aerosol Optical Thickness Retrievals, J. Geophysical Research-Atmospheres, VOL. 111, D22207, doi:10.1029/2005JD006898, 2006.

Table 1. Mass extinction cross-sections (σ, m^2/g) and single scattering albedos (ω_o) used in this study.

	ABF	Dust	Smoke	Sea Salt
σ (354 nm)	7.81	0.56	6.91	0.52
ω_o (354 nm)	1.0	0.88	0.85	1.0
σ (388 nm)	6.96	0.58	6.07	0.52
ω_o (388 nm)	1.0	0.91	0.86	1.0

Figure Captions

Figure 1. (a) Spatial distribution of NAAPS AODs, using NAAPS reanalysis data from the collocated OMI and NAAPS dataset for July 2007. (b). Simulated AI using NAAPS reanalysis data as shown in (a). (c). Spatial distribution of OMI AI using gridded OMI data from the collocated OMI and NAAPS dataset for July 2007. Grey color highlights those 1x1° (Latitude/Longitude) bins that have less than thre~~wo~~ collocated NAAPS and OMI AI data for the study period.

Figure 2. (a). Vertical distributions of smoke, dust, anthropogenic and sea salt aerosols for the test case as shown in (b). (b) Scatter plot of Jacobians of AI as a function of dust concentration: analytic versus finite difference solutions.

Figure 3. (a). Aqua MODIS true-color image over Central and North Africa for July 28, 2007. This composite was obtained from the NASA worldview site (https://worldview.earthdata.nasa.gov/). (b). Spatial distribution of Gridded OMI AI for 12 UTC, July 28, 2007. (c). Spatial distribution of NAAPS AOD from the NAAPS natural run for 12 UTC, July 28, 2007. (d). Similar to (c) but using NAAPS AOD from the AI-DA run. (e). Simulated AI using data from (c). (f). Simulated AI using data from (d).

Figure 4. (a). Spatial distribution of NAAPS AOD using NAAPS data from the AI-DA runs for July and August 2007. Only NAAPS data that have collocated OMI AI data are used. (b). Spatial distribution of simulated AI for July and August 2007 using NAAPS data from the AI-DA runs. (c). Spatial distribution of gridded OMI AI for July and August 2007. (d). Differences between Figures 4(b) and 4(c). (e-h) Similar to Figures 4(a)-4(d) but using NAAPS natural runs. Grey color highlights those 1x1° (Latitude/Longitude) bins that have less than thre~~wo~~ collocated NAAPS and OMI AI data for the study period.

Figure 5. (a). Scatter plot of AERONET and NAAPS AOD (0.55 μm) using NAAPS data from the natural runs for July-August 2007 over the study region. (b). Similar to Figure 5(a) but using NAAPS data from the AI-DA runs. (c). Similar to Figure 5(a) but with AODs taken from the NAAPS reanalysis.

Figure 6. Spatial distributions of simulated AI at 12 Z on July 28, 2007 using NAAPS reanalysis data, with single scattering albedos of smoke aerosol at 354 and 388 nm taken to be: (a) 0.84 and 0.84; (b) 0.85 and 0.85; (c) 0.86 and 0.86; (d) 0.85 and 0.85; (e) 0.85, 0.855; (f) 0.85 and 0.86.

Figure 7. (a). Vertical distributions of smoke and dust aerosol concentrations over 9.5°S and 10.5°E at 12 Z on July 28, 2007 for both natural and AI DA runs. (b). Similar as (a) but over 25.5°N and 12.5°W.

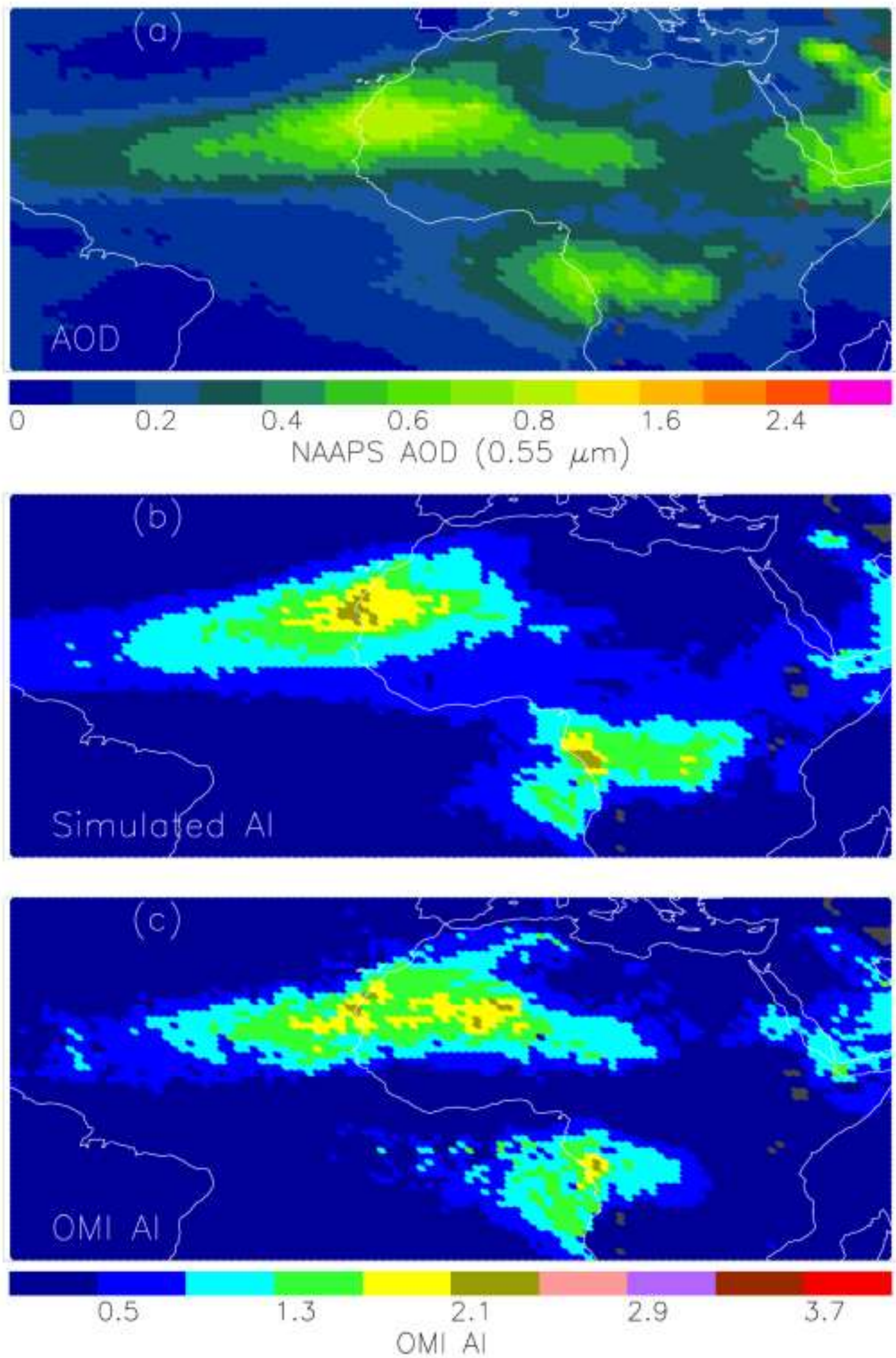


Figure 1. (a) Spatial distribution of NAAPS AODs, using NAAPS reanalysis data from the collocated OMI and NAAPS dataset for July 2007. (b). Simulated AI using NAAPS reanalysis data as shown in (a). (c). Spatial distribution of OMI AI using gridded OMI data from the collocated OMI and NAAPS dataset for July 2007. Grey color highlights those 1x1° (Latitude/Longitude) bins that have less than three~~wo~~ collocated NAAPS and OMI AI data for the study period.

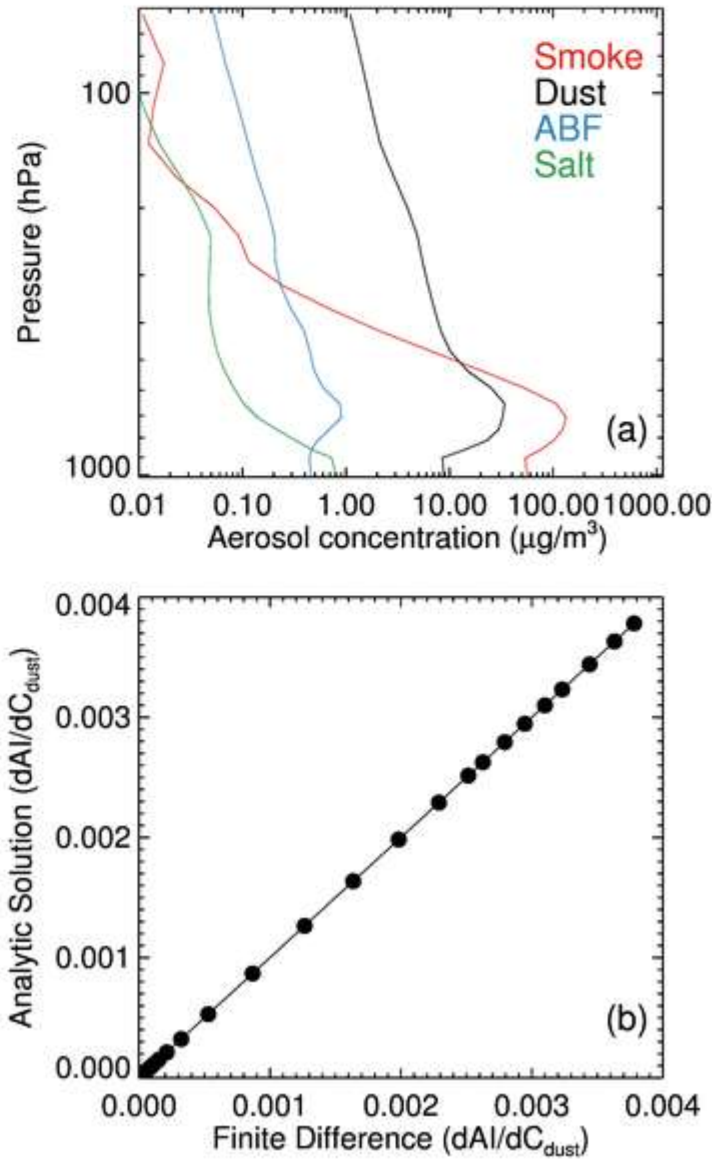


Figure 2. (a). Vertical distributions of smoke, dust, anthropogenic and sea salt aerosols for the test case as shown in (b). (b) Scatter plot of Jacobians of AI as a function of dust concentration: analytic versus finite difference solutions.

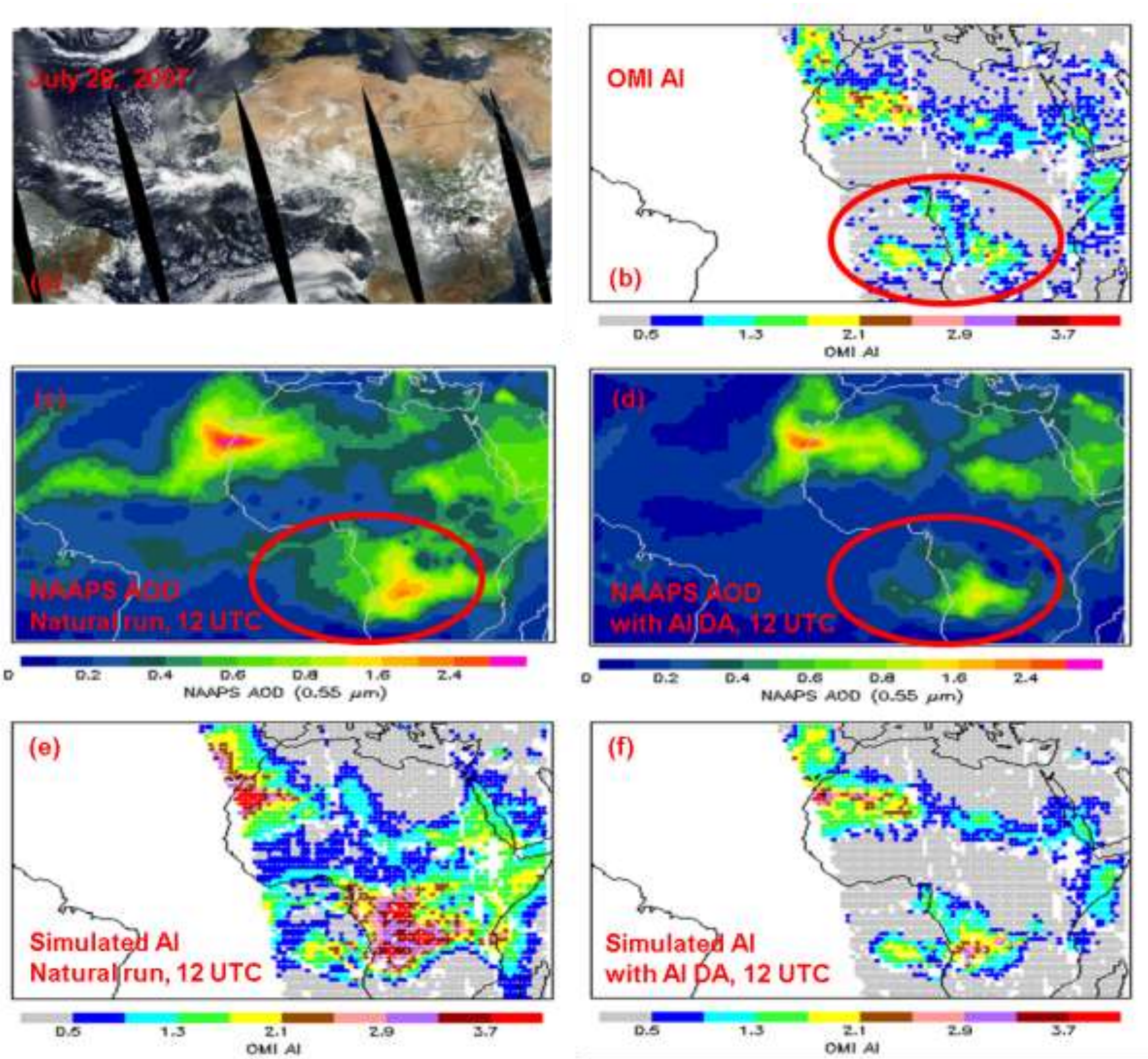


Figure 3. (a). Aqua MODIS true-color image over Central and North Africa for July 28, 2007. This composite was obtained from the NASA worldview site (https://worldview.earthdata.nasa.gov/). (b). Spatial distribution of Gridded OMI AI for 12 UTC, July 28, 2007. (c). Spatial distribution of NAAPS AOD from the NAAPS natural run for 12 UTC, July 28, 2007. (d). Similar to (c) but using NAAPS AOD from the AI-DA run. (e). Simulated AI using data from (c). (f). Simulated AI using data from (d).

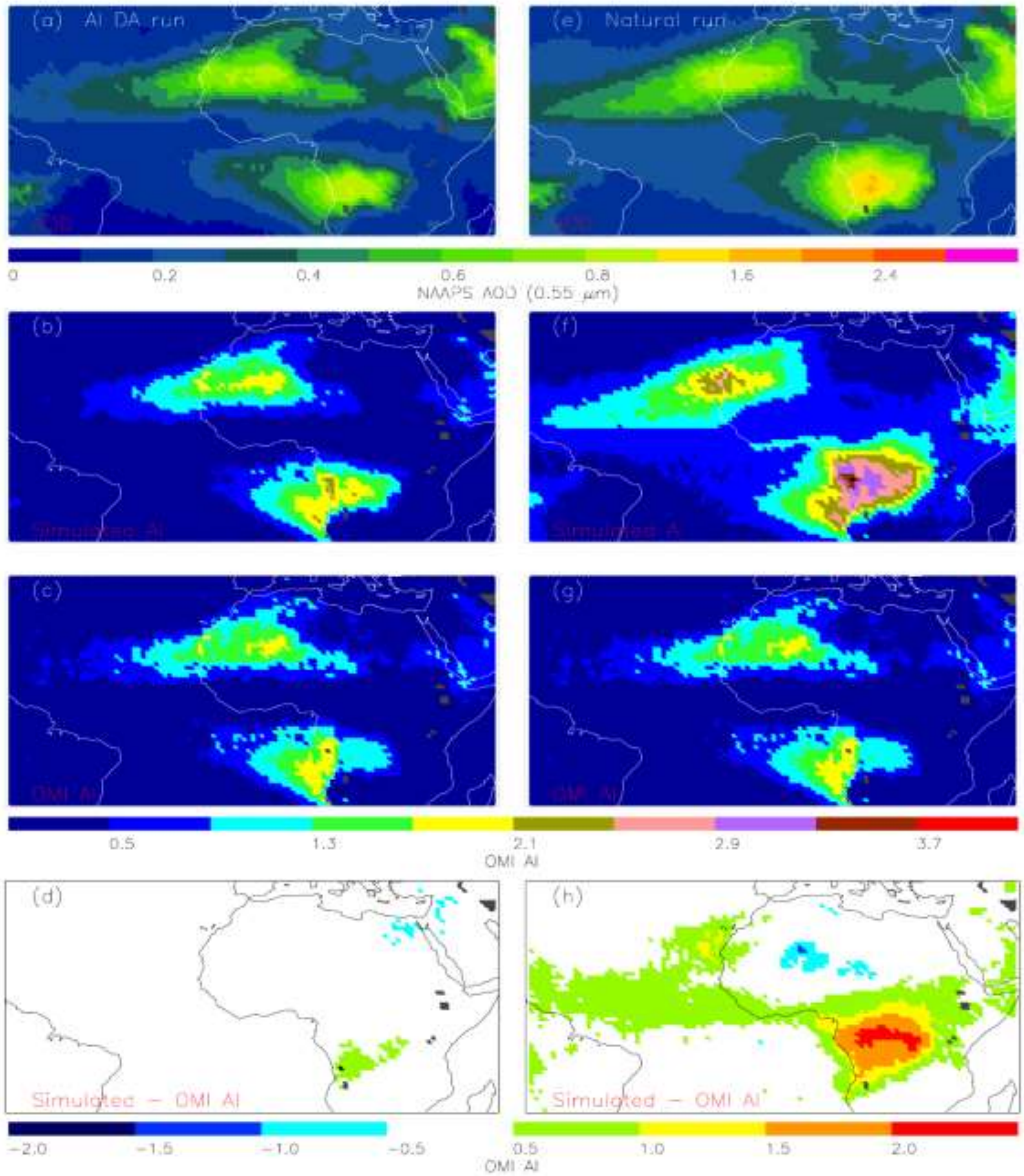
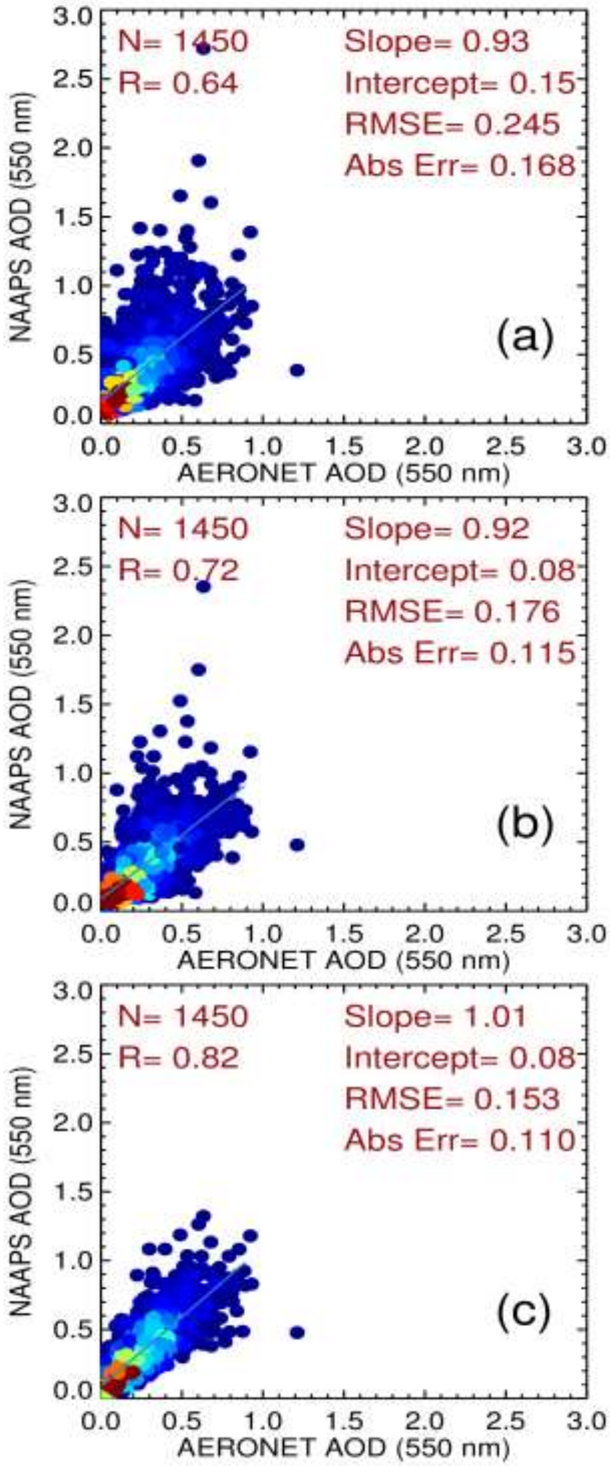


Figure 4. (a). Spatial distribution of NAAPS AOD using NAAPS data from the AI-DA runs for July and August 2007. Only NAAPS data that have collocated OMI AI data are used. (b). Spatial distribution of simulated AI for July and August 2007 using NAAPS data from the AI-DA runs. (c). Spatial distribution of gridded OMI AI for July and August 2007. (d). Differences between Figures 4(b) and 4(c). (e-h) Similar to Figures 4(a)-4(d) but using NAAPS natural runs. Grey color highlights those 1x1° (Latitude/Longitude) bins that have less than three~~wo~~ collocated NAAPS and OMI AI data for the study period.


N= 1450
R= 0.64
Slope= 0.93
Intercept= 0.15
RMSE= 0.245
Abs Err= 0.168
(a)
N= 1450
R= 0.72
Slope= 0.92
Intercept= 0.08
RMSE= 0.176
Abs Err= 0.115
(b)
N= 1450
R= 0.82
Slope= 1.01
Intercept= 0.08
RMSE= 0.153
Abs Err= 0.110
(c)
NAAPS AOD (550 nm)
AERONET AOD (550 nm)

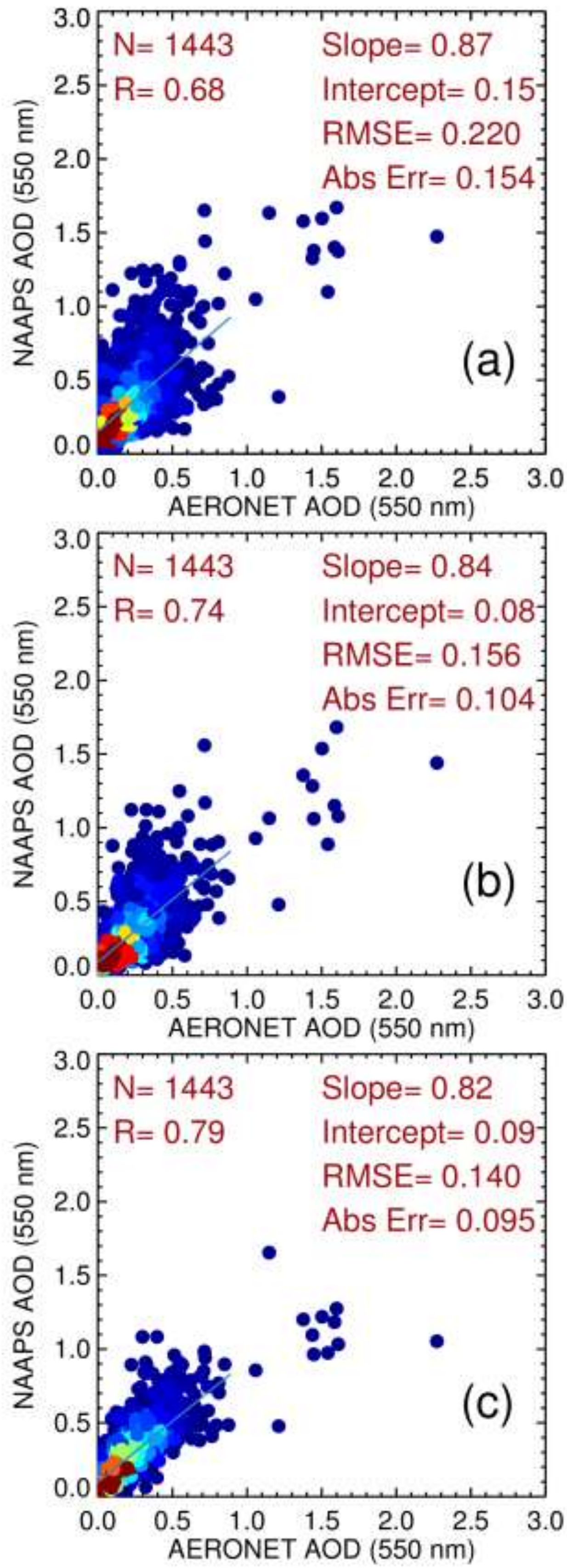

N= 1443
R= 0.68
Slope= 0.87
Intercept= 0.15
RMSE= 0.220
Abs Err= 0.154
(a)
NAAPS AOD (550 nm)
AERONET AOD (550 nm)
N= 1443
R= 0.74
Slope= 0.84
Intercept= 0.08
RMSE= 0.156
Abs Err= 0.104
(b)
NAAPS AOD (550 nm)
AERONET AOD (550 nm)
N= 1443
R= 0.79
Slope= 0.82
Intercept= 0.09
RMSE= 0.140
Abs Err= 0.095
(c)
NAAPS AOD (550 nm)
AERONET AOD (550 nm)

Figure 5. (a). Scatter plot of AERONET and NAAPS AOD (0.55 μm) using NAAPS data from the natural runs for July-August 2007 over the study region. (b). Similar to Figure 5(a) but using NAAPS data from the AI-DA runs. (c). Similar to Figure 5(a) but with AODs taken from the NAAPS reanalysis.

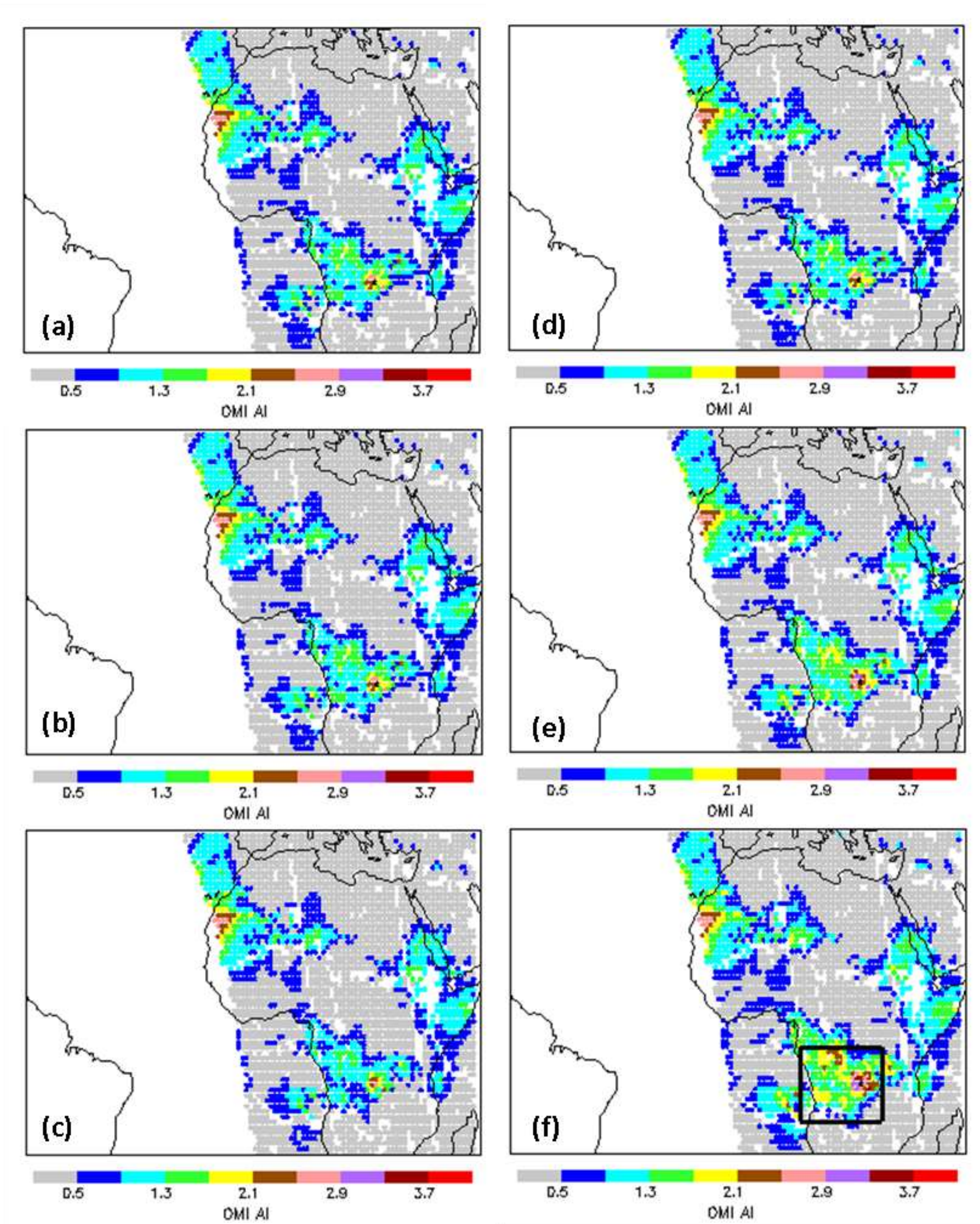
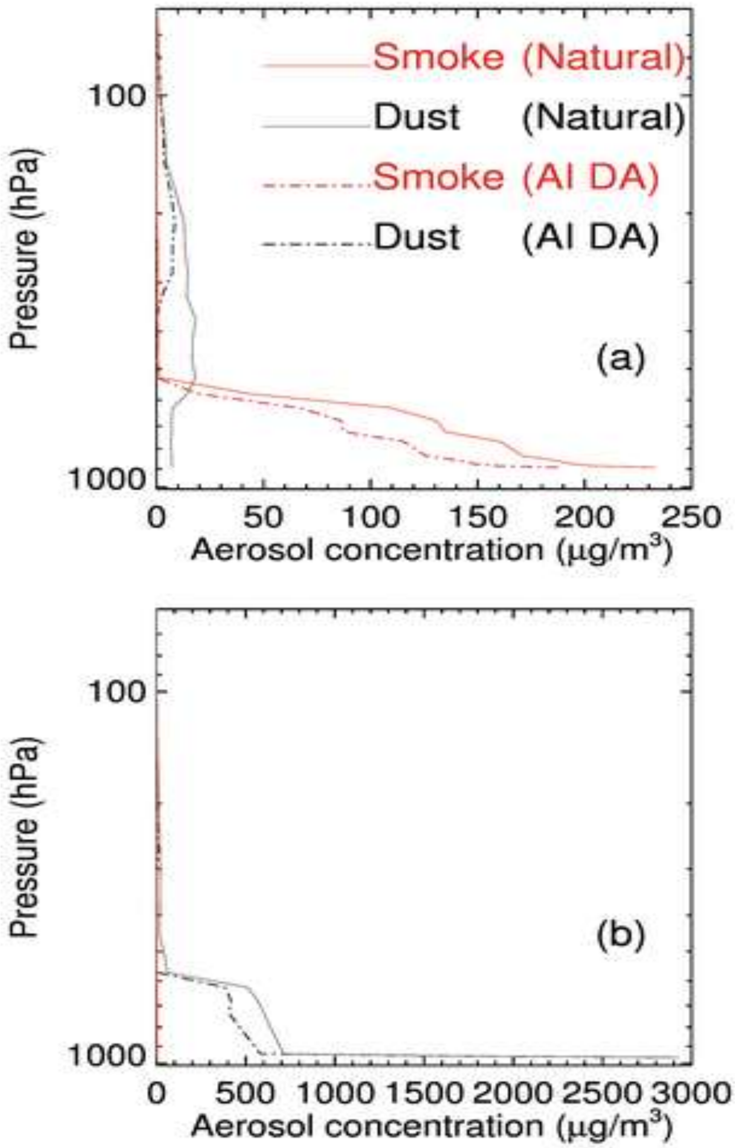


Figure 6. Spatial distributions of simulated AI at 12 Z on July 28, 2007 using NAAPS reanalysis data, with single scattering albedos of smoke aerosol at 354 and 388 nm taken to be: (a) 0.84 and 0.84; (b) 0.85 and 0.85; (c) 0.86 and 0.86; (d) 0.85 and 0.85; (e) 0.85, 0.855; (f) 0.85 and 0.86.



926

927 **Figure 7.** (a). Vertical distributions of smoke and dust aerosol concentrations over 9.5°S and
928 10.5°E at 12 Z on July 28, 2007 for both natural and AI DA runs. (b). Similar as (a) but over
929 25.5°N and 12.5°W.
930

931