

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

**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<sup>1</sup>, Robert J. D. Spurr<sup>2</sup>, Jeffrey S. Reid<sup>3</sup>, Peng Xian<sup>3</sup>, Peter R. Colarco<sup>4</sup>, James R. Campbell<sup>3</sup>, Edward J. Hyer<sup>3</sup>, and Nancy Baker<sup>3</sup>

<sup>1</sup>Department of Atmospheric Sciences, University of North Dakota, Grand Forks, ND

<sup>2</sup>RT SOLUTIONS Inc., Cambridge MA

<sup>3</sup>Marine Meteorology Division, Naval Research Laboratory, Monterey, CA

<sup>4</sup>NASA Goddard Space Flight Center, Greenbelt, MD

Revised version submitted to

GMD

September 2020

Corresponding Author: [jianglong.zhang@und.edu](mailto:jianglong.zhang@und.edu)

26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

## 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 aerosol model analysis over cloudy regions and bright surfaces. However, the OMI AI data assimilation alone does not outperform 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.

## 47 **1.0 Introduction**

48 Operational chemical transport modeling (CTM) of atmospheric aerosol particles,  
49 including simulation of sources and sinks and long-range transport of aerosol events such as  
50 biomass burning aerosols from fires and dust outbreaks, is now commonplace at global  
51 meteorology centers for air quality and visibility forecasts (e.g. Sessions et al, 2015; Lynch et al.,  
52 2016). Variational and ensemble-based assimilation of satellite derived aerosol products such as  
53 aerosol optical depth (AOD), lidar backscatter measurements, and surface aerosol properties, can  
54 substantially improve accuracies in CTM analyses and forecasts (Zhang et al., 2008; 2011; 2014;  
55 Yumimoto et al., 2008; Uno et al., 2008; Benedetti et al., 2009; Schutgens et al., 2010; Sekiyama  
56 et al., 2010; Saide et al. 2013; Schwartz, 2012; Li et al., 2013; Rubin et al., 2017; Lynch et al.,  
57 2016).

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

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

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

87 To complement existing AOD assimilating systems, we have developed an AI data  
88 assimilation (AI-DA) system that is capable of assimilating OMI AI over bright surfaces and  
89 cloudy regions for aerosol analyses and forecasts. This study can be considered as one of the first  
90 attempts for direct radiance assimilation in the UV spectrum for aerosol applications, as AI can be  
91 directly computed from UV radiances and the developed OMI AI-DA system has all necessary  
92 components for a typical radiance assimilation package. In time we expect our assimilation model

93 to merge with AOD or solar radiance assimilation to influence aerosol loading, height and  
94 absorption (e.g., VIIRS+OMPS product; such as Lee et al. 2015). Details of the developed OMI  
95 AI assimilation system are presented in the paper, which is organized as follows: Data sets used  
96 in the study are summarized in Section 2; Section 3 discusses the components of the AI-DA  
97 system. Section 4 provides an evaluation of the developed system; and Section 5 contains a  
98 summary discussion.

99

## 100 **2.0 Datasets and Models**

101 Three datasets are used in this study. These are: (i) the OMI level 2 UV aerosol product  
102 (OMAERUV; Torres et al., 2007), (ii) the Aerosol Robotic Network (AERONET; Holben et al.,  
103 1998) AOD product, and (iii) reanalysis data from the Naval Aerosol Analysis and Prediction  
104 System (NAAPS; Lynch et al., 2016), which was the first operational global aerosol mass transport  
105 model available to the community. The assimilation system is based on spatial and temporal  
106 variations of aerosol particles from NAAPS (Zhang et al., 2006; 2008), and the Vector LLinearized  
107 Discrete Ordinate Radiative Transfer (VLIDORT; Spurr, 2006) code is used to construct a forward  
108 model for the AI-DA system.

109

### 110 **2.1 OMI aerosol product**

111 UV Aerosol Index data from the OMI level 2 version 3 UV aerosol products (OMAERUV)  
112 are used in this study. The OMI instrument is on board the Aura satellite (launched in 2004) and  
113 it observes the earth's atmosphere over the UV/visible spectrum with a pixel size of 13x24 km at  
114 nadir for the global scan mode, and a swath of ~2600 km (Levelt et al., 2018). The daytime  
115 equatorial crossing for the Aura platform is ~1:30 p.m. The dataset comprises the UV AI, viewing

116 and solar geometries, spectrally-dependent surface albedos at the 354 and 388 nm spectral  
117 channels, terrain pressure, geolocations, x-track and algorithm quality flags, plus other aerosol and  
118 ancillary parameters. The UV AI is designed to detect UV-absorbing aerosol particles, and is  
119 based on radiance observations at 354 nm ( $I_{obs354}$ ) and calculated radiance ( $I_{cal354}$ ) at 354 nm for a  
120 Rayleigh (no aerosol) atmosphere (e.g. Torres et al., 2007) as defined as

$$121 \quad AI = -100 \log_{10} \frac{I_{obs354}}{I_{cal354}}. \quad (1)$$

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

130 Both cloud-free and above-cloud AI data satisfying these quality checks are aggregated /  
131 averaged in  $1 \times 1^\circ$  (Latitude/Longitude) bins. As a radiative transfer model run is applied for each  
132 observation, the gridded data are used in the assimilation process in order to reduce the  
133 computational burden. Averaged parameters for the gridded data include the solar and sensor  
134 zenith angles, the relative azimuth angles, the spectrally-dependent surface albedos at 354 and 388  
135 nm, the cloud fraction, and the AI values themselves. Additional quality assurance steps are also  
136 applied during the spatial-averaging process. Isolated high AI values are removed as follows.  
137 First, for a  $4 \times 4$  pixel box, if the mean AI is less than 0.7 but an individual AI value is larger than  
138 0.7, then that one value is removed. Second, if the standard deviation of AI values for a  $3 \times 3$  pixel

139 box surrounding a pixel is larger than 0.5, that individual AI value is likewise removed. Note that  
140 both approaches are essentially homogeneity tests that are used for identifying outliers. The  
141 thresholds are estimated empirically through visual inspection.

142

## 143 **2.2 AERONET data**

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

154

## 155 **2.3 NAAPS and NAAPS reanalysis data**

156 The NAAPS (<http://www.nrlmry.navy.mil/aerosol/>) model is a multi-species, three-  
157 dimensional, Eulerian global transport model using operational Navy Global Environmental  
158 Model (NAVGEM) as the meteorological driver (Hogan et al., 2014). NAAPS provides 6-day  
159 forecasts at a 3-hour interval with a spatial resolution of  $1/3^\circ$  (latitude/Longitude) and 42 vertical  
160 levels on a global scale. NAAPS predicts four aerosol particle classes: anthropogenic and biogenic

161 fine particles (ABF, such as primary and secondary organic aerosols and sulfate aerosols); dust,  
162 biomass burning smoke; and sea salt (Lynch et al, 2016).

163         The 2003-2018 NAAPS reanalysis version 1 (v1) (Lynch et al., 2016) is a modified version  
164 of the operational NAAPS model. In this version, quality-controlled retrievals of AOD from  
165 MODIS and MISR (Zhang et al., 2006; Hyer et al., 2011; Shi et al., 2014) are assimilated into  
166 NAAPS through the Naval Research Laboratory Atmospheric Variation Data Assimilation  
167 System-AOD system (NAVDAS-AOD; e.g., Zhang et al., 2008; Zhang et al., 2011; Zhang et al.,  
168 2014). Aerosol source functions, including biomass burning, smoke and dust emissions, are tuned  
169 regionally based on the AERONET data. Other aerosol processes, including dry deposition over  
170 water, are also tuned based on AOD data assimilation correction fields. NOAA Climate Prediction  
171 Center (CPC) MORPHing (CMORPH) precipitation data are used to constrain the wet removal  
172 process within the tropics (Joyce et al., 2004). The usage of CMORPH avoids the ubiquitous  
173 precipitation bias that exists in all global atmospheric models (e.g. Dai, 2006) and is proven to  
174 improve aerosol wet deposition, therefore yielding better AOD (Xian et al., 2009). The reanalysis  
175 agrees reasonably well with AERONET data on a global scale (Lynch et al., 2016) and also  
176 reproduces AOD trends that are in a good agreement with satellite based analysis (e.g., Zhang and  
177 Reid, 2010; Hsu et al., 2012). In this study, we use a free running version of NAAPS reanalysis v1  
178 without AOD assimilation to provide aerosol fields every 6 hours at  $1^\circ \times 1^\circ$  (Latitude/Longitude)  
179 resolution.

180

## 181 **2.4 VLIDORT radiative transfer code**

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

184 4-vectors with respect to any atmospheric or surface property (Spurr, 2006). The model uses  
185 discrete-ordinate methods to solve the polarized plane-parallel RT equations in a multi-layer  
186 atmosphere, plus the solution of a boundary value problem and subsequent source-function  
187 integration to obtain radiation fields at any geometry and any atmospheric level. VLIDORT has a  
188 “pseudo-spherical” *ansatz*: the treatment of solar-beam attenuation in a spherical-shell atmosphere  
189 before scattering. Single-scattering in VLIDORT is accurate for both line-of-sight and solar-beam  
190 spherical geometry. The model has a full thermal emission capability. VLIDORT has two  
191 supplements, one dealing with bidirectional (non-Lambertian) reflection at the surface, and the  
192 other with the inclusion of surface light sources (SIF or water-leaving radiances). Full details on  
193 the VLIDORT model may be found in a recent review paper (Spurr and Christi, 2019, and  
194 references to VLIDORT therein).

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

203

### 204 **3.0 OMI AI assimilation system**

205 The OMI assimilation system has three components: a forward model, a 3-D variational  
206 assimilation system, and a post-processing system. Based on the background NAAPS 3-D aerosol

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

215

### 216 **3.1 Forward model for simulating OMI AI**

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

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

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

240 To simulate OMI AI, the Lambertian Equivalent Reflectivity (LER) at 388 nm ( $R_{388}$ ) is  
 241 needed for estimating LER at 354 nm. The  $R_{388}$  is calculated from VLIDORT, based on equation  
 242 2 below, adapted from Buchard et al. (2015), or

$$243 \quad R_{388} = \frac{I_{aer388}(\rho_{388}) - I_{ray388}(0)}{T + S_b(I_{aer388}(\rho_{388}) - I_{ray388}(0))} \quad . \quad (2)$$

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

$$250 \quad R'_{388} = R_{388} - (\rho_{388} - \rho_{354}) \quad . \quad (3)$$

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

$$254 \quad AI_{naaps} = -100 \log_{10} \frac{I_{aer354}(\rho_{354})}{I_{ray354}(R'_{388})} \quad . \quad (4)$$

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

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

$$263 \quad \rho_{cld} = \rho_{ctr} * (1 - f_c) + 0.8 * f_c \quad . \quad (5)$$

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

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

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

282

### 283 3.2 Forward model for Jacobians of AI

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

$$293 \quad \frac{\partial AI}{\partial M_{nk}} = \mathcal{A}_1 K_{354,nk}(\rho_{354}) + \mathcal{A}_2 K_{388,nk}(\rho_{388}) \quad . \quad (6)$$

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

$$295 \quad \mathcal{A}_1 = \left( -\frac{100}{I_{aer354}(\rho_{354}) \times \ln 10} \right) \quad , \text{ and} \quad (7)$$

296 
$$\mathcal{A}_2 = \left( -\frac{100}{I_{ray354}(R'_{388}) \times \ln 10} \right) \frac{\partial I_{ray354}(R'_{388})}{\partial R} \left[ \frac{(1 - S_{388} R_{388})^2}{T_{388}} \right] \quad (8)$$

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

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

303 
$$\frac{\partial AI}{\partial M_{nk}} = \frac{(AI - AI')}{(C_{nk} - C'_{nk})} \quad (9)$$

304 Here  $C_{nk}$  and  $C'_{nk}$  are the baseline and perturbed aerosol concentrations, respectively, and AI and  
 305 AI' are computed using  $C_{nk}$  and  $C'_{nk}$ , respectively.

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

314

### 315 **3.2 The variational OMI AI assimilation system**

316 The OMI AI assimilation system is based on AI simulations (with Jacobians) from the  
 317 forward model. Two principles underlay the assimilation procedure. First, we assume that OMI AI

318 is sensitive to UV-absorbing aerosol particles, such as NAAPS smoke and dust, or that only smoke  
 319 and dust are injected high enough into the troposphere to impact AI. Therefore, innovations are  
 320 limited to modifications of dust and smoke aerosol properties. For classes that do not strongly  
 321 project onto AI, such as sea salt and ABF aerosols, aerosol concentrations are not modified during  
 322 the process. Second, contributions of smoke/dust aerosols to AI ( $AI_{\text{smoke}} / AI_{\text{dust}}$ ) prior to  
 323 assimilation are estimated by multiplying smoke/dust aerosol concentrations from NAAPS with  
 324 Jacobians of AI respective of smoke/dust aerosols. The ratio of AI innovation from smoke aerosols  
 325 ( $\Delta AI_{\text{smoke}}$ ) to total AI innovation ( $\Delta AI$  or  $OMI\ AI - AI_{\text{naaps}}$ ) is assumed to be the ratio of  $AI_{\text{smoke}}$  to  
 326  $AI_{\text{smoke}} + AI_{\text{dust}}$ . The same assumption holds for dust aerosols.

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

329

330  $C^a = C^b +$

331 
$$\frac{P_{\text{dust}} \mathbf{H}_{\text{dust}}^T}{\mathbf{H}_{\text{dust}}^T P_{\text{dust}} \mathbf{H}_{\text{dust}} + R} [y - H(C^b)] \times \frac{\mathbf{H}_{\text{dust}} C_{\text{dust}}^b}{\mathbf{H}_{\text{dust}} C_{\text{dust}}^b + \mathbf{H}_{\text{smk}} C_{\text{smk}}^b} +$$

332 
$$\frac{P_{\text{smk}} \mathbf{H}_{\text{smk}}^T}{\mathbf{H}_{\text{smk}}^T P_{\text{smk}} \mathbf{H}_{\text{smk}} + R} [y - H(C^b)] \times \frac{\mathbf{H}_{\text{smk}} C_{\text{smk}}^b}{\mathbf{H}_{\text{dust}} C_{\text{dust}}^b + \mathbf{H}_{\text{smk}} C_{\text{smk}}^b}, \quad (10)$$

333

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

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

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

$$355 \quad C(x, y) = (1 + R_{xy}/L)\exp(-\frac{R_{xy}}{L}) \quad . \quad (11)$$

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

$$360 \quad C(p_1, p_2) = [1 + \left| \int_{p_1}^{p_2} \frac{d \ln p}{L} \right|] e^{-\left| \int_{p_1}^{p_2} \frac{d \ln p}{L} \right|} \quad . \quad (12)$$

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

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

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

381

## 382 **4.0 System evaluation & discussion**

### 383 **4.1 Evaluating the performance of the AI assimilation system over Africa**

384 Using two months of OMI data (July-August, 2007), the performance of OMI AI  
385 assimilation was evaluated around the Africa region (20°S-40°N; 60°W-50°E). The study region  
386 was chosen to examine the performance of OMI AI data assimilation over bright surfaces such as  
387 the deserts of North Africa, as well as study aerosol advection over clouds, in this case smoke off  
388 the west coast of Southern Africa. In this demonstration, two NAAPS runs were performed for  
389 the period of July 1 to August 31, 2007, one with and one without the use of OMI AI assimilation  
390 (AI-DA run). Both runs were initialized with the use of NAAPS reanalysis data at 0000 UTC 1  
391 July and do not include any other form of aerosol assimilation.

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

404 Figure 3c is the 1200 UTC 28 July 2007 NAAPS AOD product from the natural run. In  
405 comparison, Figure 3d shows the same situation, this time with the use of OMI AI data  
406 assimilation. Comparing 3b with 3d, dust and smoke aerosol patterns as shown from OMI AI

407 resemble more closely the NAAPS AOD fields after AI assimilation. Over the northeast coast of  
408 Africa, heavy aerosol plumes, as hinted at in NAAPS AOD from the natural run (Figure 3c), cover  
409 larger spatial areas than those inferred from OMI AI data. In comparison, NAAPS AOD patterns  
410 from the OMI AI data assimilation cycle closely resemble aerosol patterns as suggested from OMI  
411 AI data. Also shown in Figures 3e and 3f are the simulated AI using NAAPS data from the natural  
412 and OMI AI DA runs (data from Figures 3c and 3d) respectively. Clearly, with the use of NAAPS  
413 data from the natural run, simulated OMI AI are overestimated in comparison with OMI AI data  
414 (Figure 3b). Simulated AI patterns with the used of NAAPS data from the OMI AI DA run rather  
415 closely resemble AI patterns from the OMI data, again, indicating the OMI AI DA system is  
416 functioning reasonably as designed.

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

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

442

#### 443 **4.2 Inter-comparison with AOD data assimilation**

444 Typically, NAAPS reanalyses are constructed through assimilation of MISR and MODIS  
445 aerosol products (NAAPS AOD assimilation). Thus, the performances of NAAPS AOD and AI-  
446 DA assimilations are compared against AERONET data. Figure 5c shows the comparison of  
447 AERONET AOD and NAAPS AOD after AOD assimilation, while Figure 5b shows a similar plot  
448 but using NAAPS data from AI-DA. Note that the same version of the NAAPS model with the  
449 same temporal and spatial resolutions, and driven by the same meteorological data, were used in  
450 constructing Figure 5 and thus the differences in Figures 5a, 5b and 5c only result from different  
451 aerosol data assimilation methods implemented (no data assimilation for the natural run). A better  
452 correlation between AERONET and NAAPS data of 0.79 is found using AOD data assimilation.  
453 In comparison, the correlation is 0.74 for the AI-DA runs. Slightly better RMSE (0.140 versus

454 0.156) and absolute error (0.095 versus 0.104) values are also found for the AOD data assimilation  
455 runs. This result is not surprising as OMI AI provides only a proxy for aerosol properties while  
456 passive-based AOD retrievals are often considered as a more reliable parameter for representing  
457 column-integrated aerosol properties. But still, the evaluation efforts are over cloud-free line-of-  
458 sight as detected from AERONET, AI DA may further assist traditional AOD data assimilation by  
459 providing AI assimilation over cloudy regions.

460

### 461 **4.3 Sensitivity test**

462 As mentioned in Section 3, aerosol properties for non-smoke aerosol types were obtained  
463 from the NASA GEOS-5 model (e.g. Colarco et al., 2014; Buchard et al., 2015). Yet, different  
464 smoke aerosol SSA values are used in this study, as values for central Africa have a strong seasonal  
465 dependency (e.g. Eck et al., 2013). While SSA values of 0.85 and 0.86 are used for the 354 and  
466 388 nm channels, respectively, in our study, we have also examined the sensitivity of simulated  
467 OMI AI with respect to differing SSA values (Figure 6). Figures 6a-c show the simulated AI at  
468 1200 UTC 28 July 2007 using NAAPS reanalysis data (Lynch et al, 2016) for three scenarios: SSA  
469 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  
470 0.86 (Figure 6c). Over the central Africa area, where smoke plumes are expected, simulated OMI  
471 AI patterns are similar for Figures 6a and 6b, but reduced values in AI are found when using higher  
472 SSA values of 0.86 at both 354 and 388 nm. This is further confirmed by the averaged AI for the  
473 smoke region over central Africa ( $14.5^{\circ}$  to  $40.5^{\circ}$  S latitude and  $10.5^{\circ}$  to  $30.5^{\circ}$  E longitude;  
474 indicated using the black box in Figure 6f) of 0.96, 0.94 and 0.78 for Figures 6a, 6b and 6c  
475 respectively.

476 Figures 6d-f show the sensitivity for adjustments of the SSA values at 388nm while  
477 maintaining a fixed SSA value of 0.85 at 354 nm. Here the SSA values at 388 nm are set to 0.85,  
478 0.855 and 0.86 for Figures 6d, 6e and 6f respectively. Interestingly, the spectral dependence of  
479 SSA seems to affect the simulated AI significantly, and this phenomenon has also been reported  
480 by previous studies (e.g. Hammer et al., 2017). The averaged AI values over central Africa (again,  
481 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  
482 and 0.86, respectively. This exercise suggests that simulated AI is a strong function of SSA, so  
483 that both the spectral dependence of SSA values at 354 and 388 nm and reliable SSA values are  
484 needed on a regional basis for future applications.

485 Interestingly, although simulated AI values are significantly affected by perturbing SSA  
486 values as shown in Figure 6, less significant impacts are observed for NAAPS AOD. This is found  
487 by running the OMI AI DA for 1200UTC, July 28, 2015 for SSA values used in generating Figure  
488 6. For example, for the black box highlighted region in Figure 6f, the averaged values for the  
489 simulated OMI AI are 0.96, 0.94 and 0.78 for using SSA values at 354 / 388 nm channels of 0.84  
490 / 0.84, 0.85 / 0.85 and 0.86 / 0.86, respectively. The corresponding NAAPS AODs are found to  
491 be 0.559, 0.560 and 0.585 after OMI AI DA, which is a change of less than 5%. Similar, by fixing  
492 the SSA value of the 354 nm channel as 0.85 and perturbing SSA values at 388 nm from 0.85 to  
493 0.86, a ~30% change is found in simulated OMI AI (from 0.94 to 1.32), yet a ~10% change is  
494 found for the NAAPS AOD (from 0.560 to 0.504) after OMI AI DA.

495 It is also of interest to investigate the changes in aerosol vertical distributions due to the  
496 OMI AI DA. For this exercise, we selected the 1200 UTC 28 July 2007 case and compared vertical  
497 distributions of smoke and dust aerosols near the peak AI value of the smoke plume (9.5°S and  
498 20.5°E) for the NAAPS natural and AI DA runs (Figure 7a). Note that the differences between

499 OMI DA and natural runs as shown in Figure 7 are essentially an integrated effect of OMI AI DA  
500 from 00Z, July 01 to 12 Z, July 28, 2007. As shown in Figure 7a, the corrections to dust and  
501 smoke aerosol concentrations from the AI DA system seem to be systematic changes across the  
502 majority of vertical layers, instead of moving dust or smoke aerosol plumes vertically. As dust  
503 aerosol concentrations are reduced at all layers and a systematic correction to smoke aerosol  
504 concentrations, although non-linear, is also observed. AI assimilation helps reduce the amount of  
505 upper troposphere dust (likely to be artifact) but does change the layer centroid slightly upwards.  
506 We have also evaluated NAAPS vertical distributions near a peak dust plume region (25.5°N and  
507 12.5°W) for the 12Z 28 July 2007 case as shown in Figure 7b. Similar to Figure 7a, a non-linear  
508 correction to dust aerosol concentrations is also observed across the vertical domain.

509

#### 510 **4.4 Issues and discussions**

511 The OMI AI data assimilation system is a proxy for all-sky, all-band modeling system  
512 radiance assimilation. It contains all the necessary components for such radiance assimilation,  
513 including a forward model for simulating radiances and AI values and their Jacobians, based on a  
514 full vector linearized radiative transfer model called for every observation. Therefore, the  
515 computational burden is a direct issue associated with the deployment of calls to a radiative transfer  
516 model for each observation. For the study area in this work, after binning OMI AI data into a  
517  $1^\circ \times 1^\circ$  (Latitude/Longitude) product, it still takes about ~1 CPU day for NAAPS to run for one  
518 month of model time. In comparison, the time scale for running AOD assimilation for 1 month is  
519 at the hourly level. Clearly, there will be an unavoidable computational burden of some sort for  
520 OMI AI assimilation and by extension, for future radiance assimilation in the UV/visible spectrum  
521 for aerosol analyses. Performance enhancement methods, such as parallel processing (the

522 VLIDORT software is thread-safe and can be used in parallel environments such as OpenMP), or  
523 fast look-up-table extraction based on neural-networks and trained data sets of forward simulation,  
524 must be explored in order to enable such assimilation applications in near real time on a global  
525 scale.

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

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

544 exclude potentially noisy OMI AI retrievals and for further application of OMI AI data  
545 assimilation on a global scale.

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

551

## 552 **5.0 Conclusions**

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

559 The performance of the OMI AI data assimilation system was evaluated over South-Central  
560 and Northern Africa regions for the period of 01 July -31 August 2007. This evaluation was done  
561 through inter-comparing NAAPS analyses with and without the inclusion of OMI AI data  
562 assimilation. Besides cloud-free AI retrievals over dark surfaces, OMI AI retrievals over desert  
563 regions and over areas were also considered. When compared against AERONET data, a total of  
564 ~29% reduction in Root-Mean-Square-Error (RMSE) with a ~32% reduction in absolute error  
565 were found for NAAPS analyses with the use of OMI AI assimilation. Also, NAAPS analyses  
566 with the inclusion of OMI AI data assimilation show similar aerosol patterns to those in the OMI  
567 AI data sets, showing that our OMI AI data assimilation system works as expected.

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

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

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

589

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

593  
594 **Code and data availability:** The OMI data assimilation scheme (V1.0) is constructed using  
595 VLIDORT and NAVDAS-AOD for NAAPS analyses and forecasts. The VLIDORT radiative  
596 transfer mode is a property of RT Solutions Inc. The VLIDORT code is publicly available, and  
597 comes with a standard GNU public license, through direct contact with RT Solutions Inc.  
598 ([http://www.rtslidort.com/mainprod\\_vlidort.html](http://www.rtslidort.com/mainprod_vlidort.html)). Both NAAPS and NAVDAS-AOD are  
599 proprietary to Naval Research Laboratory, United States Department of the Navy. Nevertheless,  
600 both NAAPS and NAVDAS-AOD are well documented in past studies (e.g. Lynch et al., 2016;  
601 Zhang et al., 2008; 2011; 2014; Rubin et al., 2017) and we have made every effort to thoroughly  
602 report our methods so that they may be replicated. AOD fields from the NAAPS OMI AI DA and  
603 natural runs over the study region and period are shared as the supplement to the paper for readers  
604 who are interested. The NAAPS reanalysis data are available from the USGODAE web site  
605 ([https://nrlgodae1.nrlmry.navy.mil/cgi-bin/datalist.pl?dset=nrl\\_naaps\\_reanalysis&summary=Go](https://nrlgodae1.nrlmry.navy.mil/cgi-bin/datalist.pl?dset=nrl_naaps_reanalysis&summary=Go)).  
606 The OMI OMAERUV data are available from the NASA's Goddard Earth Sciences *Data* and  
607 Information Services Center (*GES DISC*;  
608 [https://disc.gsfc.nasa.gov/datasets/OMAERUV\\_003/summary](https://disc.gsfc.nasa.gov/datasets/OMAERUV_003/summary)). AERONET data are obtained  
609 from the NASA AERONET webpage (<https://aeronet.gsfc.nasa.gov/>).

610  
611 **Competing interests.** The authors claim no competing interests.

612

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

617

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

620

621

622

623 **References:**

- 624 Alfaro-Contreras, R., Zhang, J., Campbell, J. R., and Reid, J. S.: Investigating the frequency and  
625 trends in global above-cloud aerosol characteristics with CALIOP and OMI, *Atmos. Chem.*  
626 *Phys.*, 16, 47-69, doi:10.5194/acp-16-47-2016, 2016.
- 627 Alfaro-Contreras R., Zhang, J., Campbell, J. R., Holz, R. E., and Reid, J. S.: Evaluating the Impact  
628 of Aerosol Particles above Cloud on Cloud Optical Depth Retrievals from MODIS, *J.*  
629 *Geophys. Res. Atmos.*, 119, 5410–5423, doi:10.1002/2013JD021270, 2014.
- 630 Benedetti, A., et al., Aerosol analysis and forecast in the European Centre for Medium-Range  
631 Weather Forecasts Integrated Forecast System: 2. Data assimilation, *J. Geophys. Res.*, 114,  
632 D13205, doi:10.1029/2008JD011115, 2009.
- 633 Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., Kumagai, Y., Miyakawa, T.,  
634 Murata, H., Ohno, T., Okuyama, A., Oyama, R., Sasaki, Y., Shimazu, Y., Shimoji, K.,  
635 Sumida, Y., Suzuki, M., Taniguchi, H., Tsuchiyama, H., Uesawa, D., Yokota, H., &  
636 Yoshida, R.: An Introduction to Himawari-8/9- Japan's New-Generation Geostationary  
637 Meteorological Satellites. *Journal of the Meteorological Society of Japan. Ser. II*, 94(2),  
638 151–183. <https://doi.org/10.2151/jmsj.2016-009>, 2016.
- 639 Buchard, V., Randles, C., Silva, A., Darmenov, A., Colarco, P., Govindaraju, R., Ferrare, R., Hair,  
640 J., Beyersdorf, A., Ziemba, L., Yu, H.: The MERRA-2 Aerosol Reanalysis, 1980 Onward.  
641 Part II: Evaluation and Case Studies *Journal of Climate* [https://dx.doi.org/10.1175/jcli-d-](https://dx.doi.org/10.1175/jcli-d-16-0613.1)  
642 [16-0613.1](https://dx.doi.org/10.1175/jcli-d-16-0613.1), 2017.
- 643 Buchard, V., da Silva, A. M., Colarco, P. R., Darmenov, A., Randles, C. A., Govindaraju, R.,  
644 Torres, O., Campbell, J., and Spurr, R.: Using the OMI aerosol index and absorption

645 aerosol optical depth to evaluate the NASA MERRA Aerosol Reanalysis, *Atmos. Chem.*  
646 *Phys.*, 15, 5743–5760, <https://doi.org/10.5194/acp-15-5743-2015>, 2015.

647 Cochrane, S. P., Schmidt, K. S., Chen, H., Pilewskie, P., Kittelman, S., Redemann, J., LeBlanc,  
648 S., Pistone, K., Kacenelenbogen, M., Segal Rozenhaimer, M., Shinozuka, Y., Flynn, C.,  
649 Platnick, S., Meyer, K., Ferrare, R., Burton, S., Hostetler, C., Howell, S., Freitag, S.,  
650 Dobracki, A., and Doherty, S.: Above-cloud aerosol radiative effects based on ORACLES  
651 2016 and ORACLES 2017 aircraft experiments, *Atmos. Meas. Tech.*, 12, 6505–6528,  
652 <https://doi.org/10.5194/amt-12-6505-2019>, 2019.

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

658 Dai, A.: Precipitation characteristics in eighteen coupled climate models. *Journal of Climate*, 19,  
659 4605– 4630, 2006.

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

665 Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., Eck, T. F.,  
666 Holben, B. N., Lewis, J. R., Campbell, J. R., Welton, E. J., Korkin, S. V., and Lyapustin,  
667 A. I.: Advancements in the Aerosol Robotic Network (AERONET) Version 3 database –

668 automated near-real-time quality control algorithm with improved cloud screening for Sun  
669 photometer aerosol optical depth (AOD) measurements, *Atmos. Meas. Tech.*, 12, 169–209,  
670 <https://doi.org/10.5194/amt-12-169-2019>, 2019.

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

676 Herman, J. R., Bhartia, P. K., Torres, O., Hsu, C., Seftor, C., and Celarier, E.: Global distribution  
677 of UV-absorbing aerosols from Nimbus 7/TOMS data, *J. Geophys. Res.*, 102(D14),  
678 16911–16922, doi:[10.1029/96JD03680](https://doi.org/10.1029/96JD03680)., 1997.

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

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

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

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

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

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

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

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

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

709 Levelt, P. F., Joiner, J., Tamminen, J., Veefkind, J. P., Bhartia, P. K., Stein Zweers, D. C., Duncan,  
710 B. N., Streets, D. G., Eskes, H., van der A, R., McLinden, C., Fioletov, V., Carn, S., de  
711 Laat, J., DeLand, M., Marchenko, S., McPeters, R., Ziemke, J., Fu, D., Liu, X., Pickering,

712 K., Apituley, A., González Abad, G., Arola, A., Boersma, F., Chan Miller, C., Chance, K.,  
713 de Graaf, M., Hakkarainen, J., Hassinen, S., Ialongo, I., Kleipool, Q., Krotkov, N., Li, C.,  
714 Lamsal, L., Newman, P., Nowlan, C., Suleiman, R., Tilstra, L. G., Torres, O., Wang, H.,  
715 and Wargan, K.: The Ozone Monitoring Instrument: overview of 14 years in space, *Atmos.*  
716 *Chem. Phys.*, 18, 5699–5745, <https://doi.org/10.5194/acp-18-5699-2018>, 2018.

717 Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu N. C.:  
718 The Collection 6 MODIS aerosol products over land and ocean, *Atmos. Meas. Tech.*, 6,  
719 2989–3034, [doi:10.5194/amt-6-2989-2013](https://doi.org/10.5194/amt-6-2989-2013), 2013.

720 Li, Z., Zang, Z., Li, Q. B., Chao, Y., Chen, D., Ye, Z., Liu, Y., and Liou, K. N.: A three-dimensional  
721 variational data assimilation system for multiple aerosol species with WRF/Chem and an  
722 application to PM<sub>2.5</sub> prediction, *Atmos. Chem. Phys.*, 13, 4265–4278,  
723 <https://doi.org/10.5194/acp-13-4265-2013>, 2013.

724 Lynch, P., Reid, J. S., Westphal, D. L., Zhang, J., Hogan, T. F., Hyer, E. J., Curtis, C. A., Hegg,  
725 D. A., Shi, Y., Campbell, J. R., Rubin, J. I., Sessions, W. R., Turk, F. J., and Walker, A.  
726 L.: An 11-year global gridded aerosol optical thickness reanalysis (v1.0) for atmospheric  
727 and climate sciences, *Geosci. Model Dev.*, 9, 1489–1522, [https://doi.org/10.5194/gmd-9-](https://doi.org/10.5194/gmd-9-1489-2016)  
728 [1489-2016](https://doi.org/10.5194/gmd-9-1489-2016), 2016.

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

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

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

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

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

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

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

759 Sekiyama, T. T., Tanaka, T. Y., Shimizu, A., and Miyoshi, T.: Data assimilation of CALIPSO  
760 aerosol observations, *Atmos. Chem. Phys.*, 10, 39–49, [https://doi.org/10.5194/acp-10-39-](https://doi.org/10.5194/acp-10-39-2010)  
761 2010, 2010.

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

769 Shi Y., Zhang J., J.S. Reid, B. Liu, E.J. Hyer, Critical evaluation of cloud contamination in the  
770 MISR aerosol products using MODIS cloud masking products, *Atmos. Meas. Tech.*, 7,  
771 1791-1801, doi:10.5194/amt-7-1791-2014, 2014.

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

775 Spurr, R. J. D.: VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative  
776 transfer code for forward model and retrieval studies in multilayer multiple scattering  
777 media, *J. Quant. Spectrosc. Radiat. Transfer*, 102(2), 316-342,  
778 doi:10.1016/j.jqsrt.2006.05.005, 2006.

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

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

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

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

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

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

797 Witek M. L., Flatau, P. J., Quinn, P. K., Westphal, D. L., Global sea-salt modeling: results and  
798 validation against multicampaign shipboard measurements, *J. Geophys. Res.*, 112, p.  
799 D08215, [10.1029/2006JD007779](https://doi.org/10.1029/2006JD007779), 2007.

800 Xian, P., Reid, J.S., Turk, J.F., Hyer, E.J. and Westphal, D.L.: Impact of modeled versus satellite  
801 measured tropical precipitation on regional smoke optical thickness in an aerosol transport

802 model, Geophysical Research Letters, 36, L16805,  
803 <https://doi.org/10.1029/2009GL038823>, 2009.

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

807 Yumimoto, K., Uno, I., Sugimoto, N., Shimizu, A., Liu, Z., and Winker, D. M.: Adjoint inversion  
808 modeling of Asian dust emission using lidar observations, Atmos. Chem. Phys., 8, 2869–  
809 2884, <https://doi.org/10.5194/acp-8-2869-2008>, 2008.

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

813 Zhang, J., Campbell, J. R., Reid, J. S., Westphal, D. L., Baker, N. L., Campbell, W. F., and Hyer,  
814 E. J.: Evaluating the impact of assimilating CALIOP-derived aerosol extinction profiles on  
815 a global mass transport model, Geophys. Res. Lett., Vo. 38, No. 14, L14801, doi:  
816 [10.1029/2011GL04773](https://doi.org/10.1029/2011GL04773), 2011.

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

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

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

826

827

828 Table 1. Mass extinction cross-sections ( $\sigma$ ,  $\text{m}^2/\text{g}$ ) and single scattering albedos ( $\omega_0$ ) used in  
 829 this study.

	ABF	Dust	Smoke	Sea Salt
$\sigma$ (354 nm)	7.81	0.56	6.91	0.52
$\omega_0$ (354 nm)	1.0	0.88	0.85	1.0
$\sigma$ (388 nm)	6.96	0.58	6.07	0.52
$\omega_0$ (388 nm)	1.0	0.91	0.86	1.0

830

831

832 **Figure Captions**

833  
834 **Figure 1.** (a) Spatial distribution of NAAPS AODs, using NAAPS reanalysis data from the  
835 collocated OMI and NAAPS dataset for July 2007. (b). Simulated AI using NAAPS reanalysis  
836 data as shown in (a). (c). Spatial distribution of OMI AI using gridded OMI data from the  
837 collocated OMI and NAAPS dataset for July 2007. Grey color highlights those  $1 \times 1^\circ$   
838 (Latitude/Longitude) bins that have less than three collocated NAAPS and OMI AI data for the  
839 study period.

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

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

849 **Figure 4.** (a). Spatial distribution of NAAPS AOD using NAAPS data from the AI-DA runs for  
850 July and August 2007. Only NAAPS data that have collocated OMI AI data are used. (b). Spatial  
851 distribution of simulated AI for July and August 2007 using NAAPS data from the AI-DA runs.  
852 (c). Spatial distribution of gridded OMI AI for July and August 2007. (d). Differences between  
853 Figures 4(b) and 4(c). (e-h) Similar to Figures 4(a)-4(d) but using NAAPS natural runs. Grey  
854 color highlights those  $1 \times 1^\circ$  (Latitude/Longitude) bins that have less than three collocated NAAPS  
855 and OMI AI data for the study period.

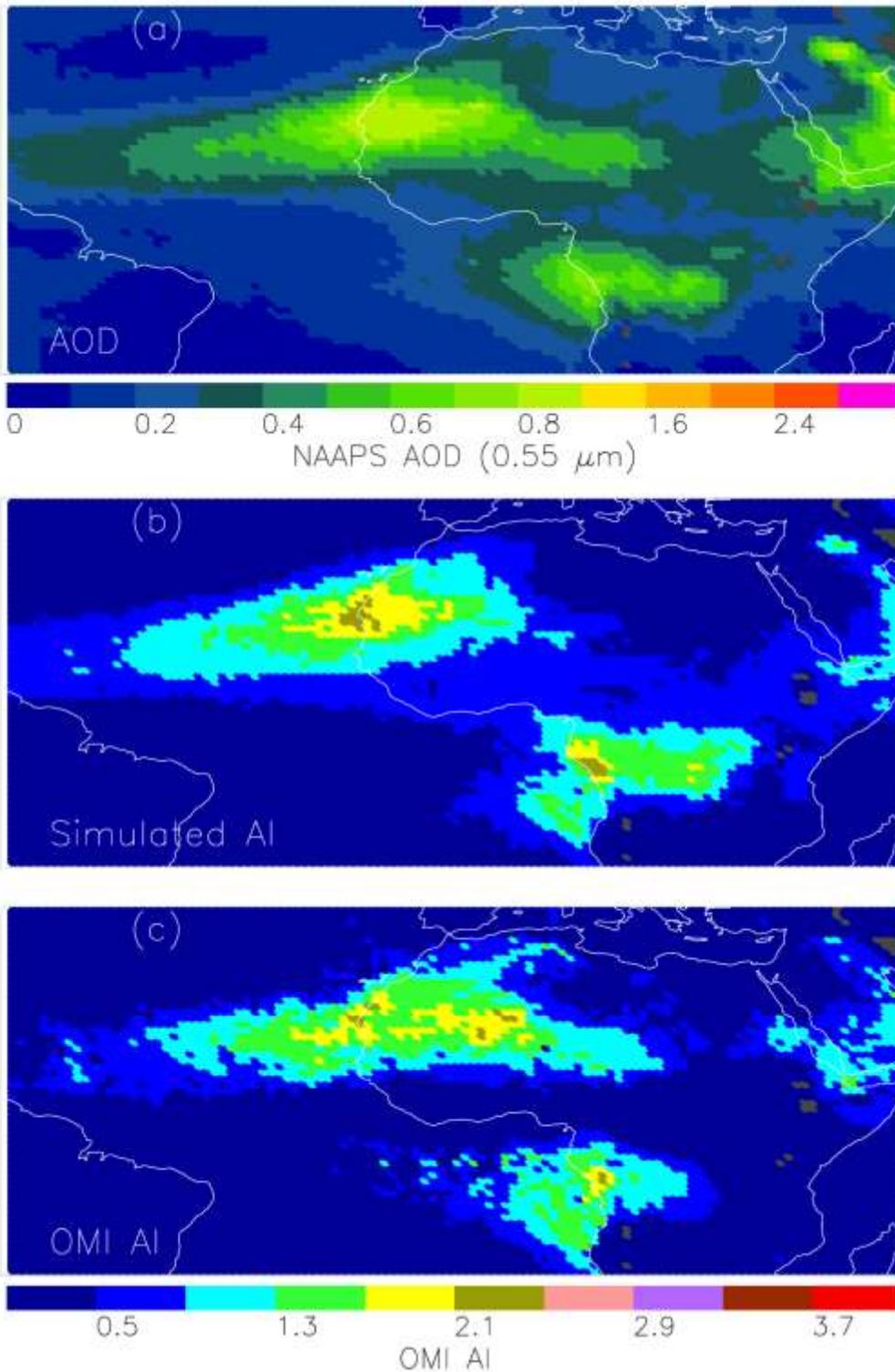
856 **Figure 5.** (a). Scatter plot of AERONET and NAAPS AOD ( $0.55 \mu\text{m}$ ) using NAAPS data from  
857 the natural runs for July-August 2007 over the study region. (b). Similar to Figure 5(a) but using  
858 NAAPS data from the AI-DA runs. (c). Similar to Figure 5(a) but with AODs taken from the  
859 NAAPS reanalysis.

860 **Figure 6.** Spatial distributions of simulated AI at 12 Z on July 28, 2007 using NAAPS reanalysis  
861 data, with single scattering albedos of smoke aerosol at 354 and 388 nm taken to be: (a) 0.84 and  
862 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.

863 **Figure 7.** (a). Vertical distributions of smoke and dust aerosol concentrations over  $9.5^{\circ}\text{S}$  and  
864  $10.5^{\circ}\text{E}$  at 12 Z on July 28, 2007 for both natural and AI DA runs. (b). Similar as (a) but over  
865  $25.5^{\circ}\text{N}$  and  $12.5^{\circ}\text{W}$ .

866

867



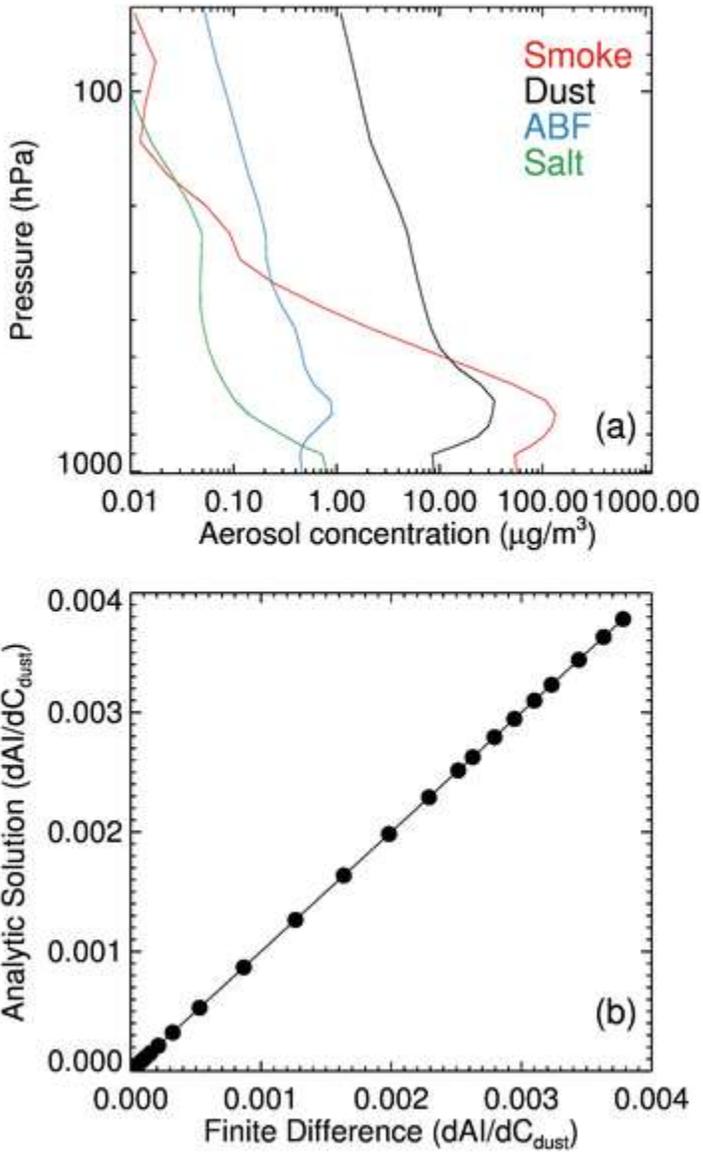
868

869

870

**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 collocated NAAPS and OMI AI data for the study period.

871



872

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

876

877

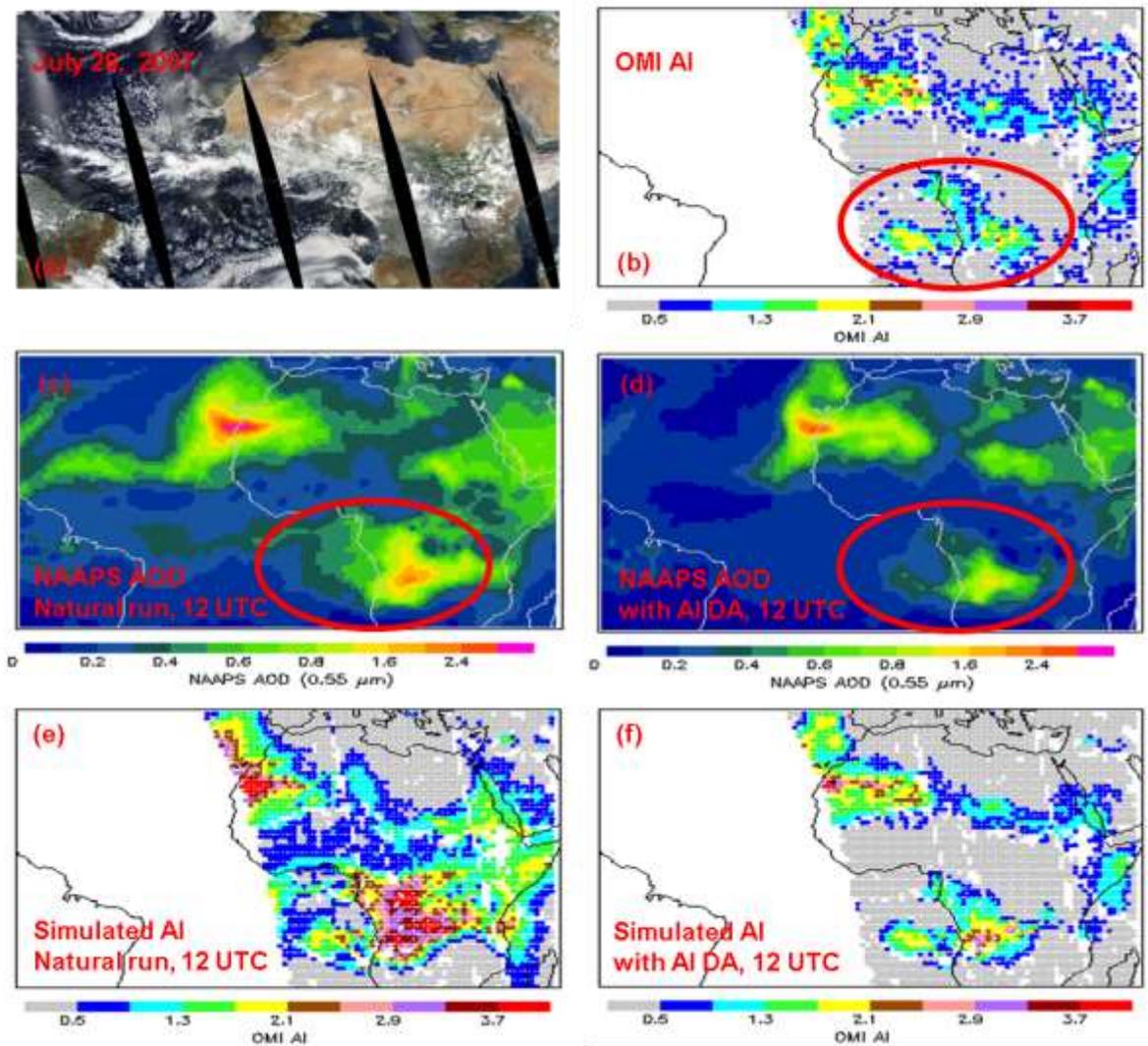
878

879

880

881

882



883

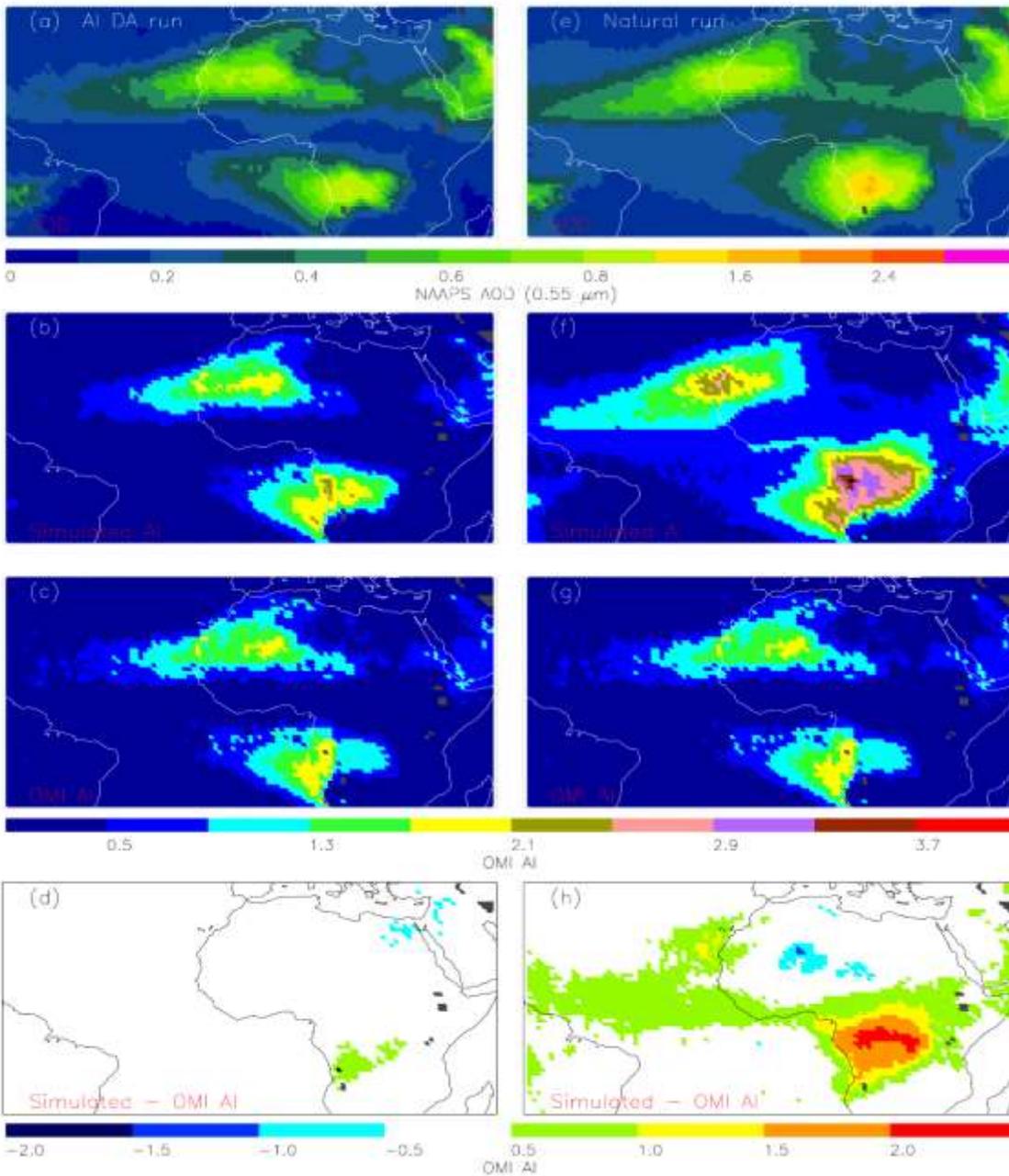
884 **Figure 3.** (a). Aqua MODIS true-color image over Central and North Africa for July 28, 2007.  
885 This composite was obtained from the NASA worldview site  
886 (<https://worldview.earthdata.nasa.gov/>). (b). Spatial distribution of Gridded OMI AI for 12 UTC,  
887 July 28, 2007. (c). Spatial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

890

891

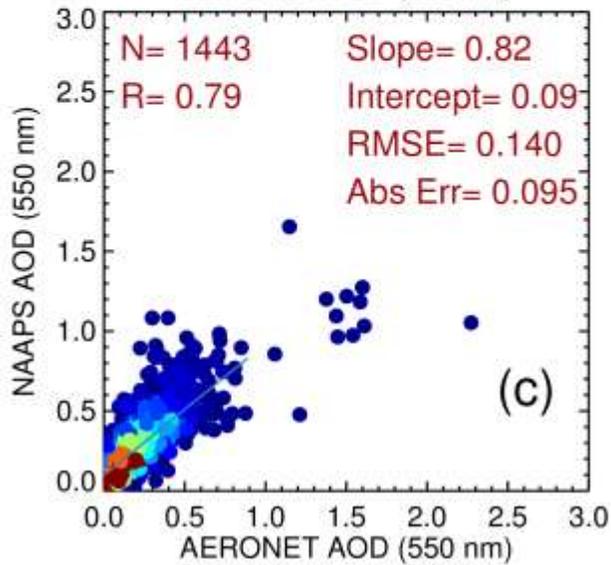
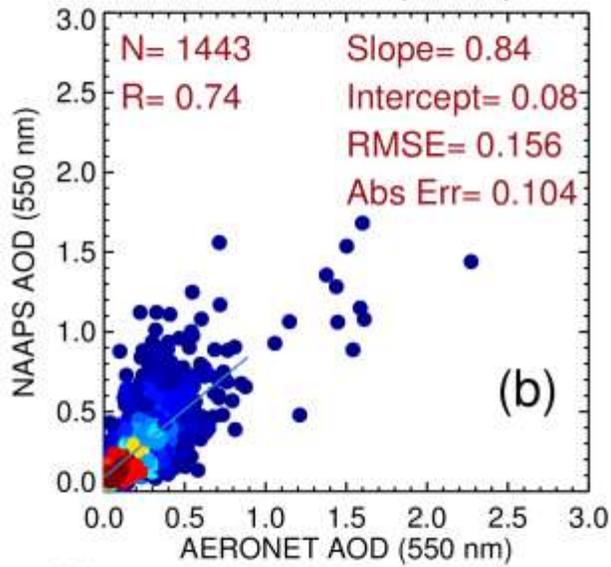
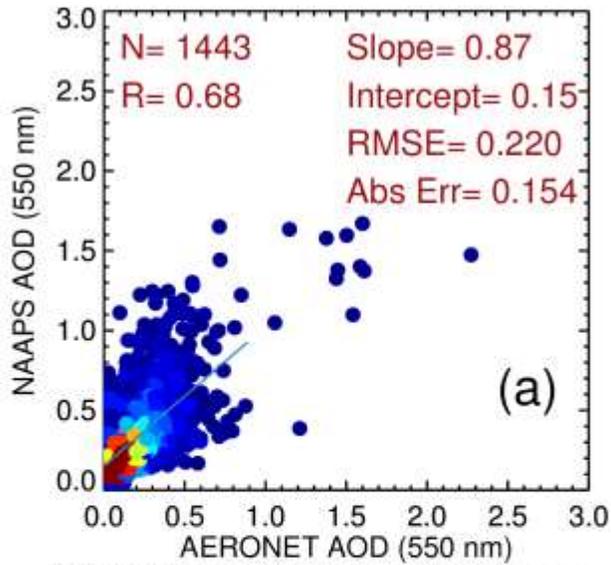
892

893

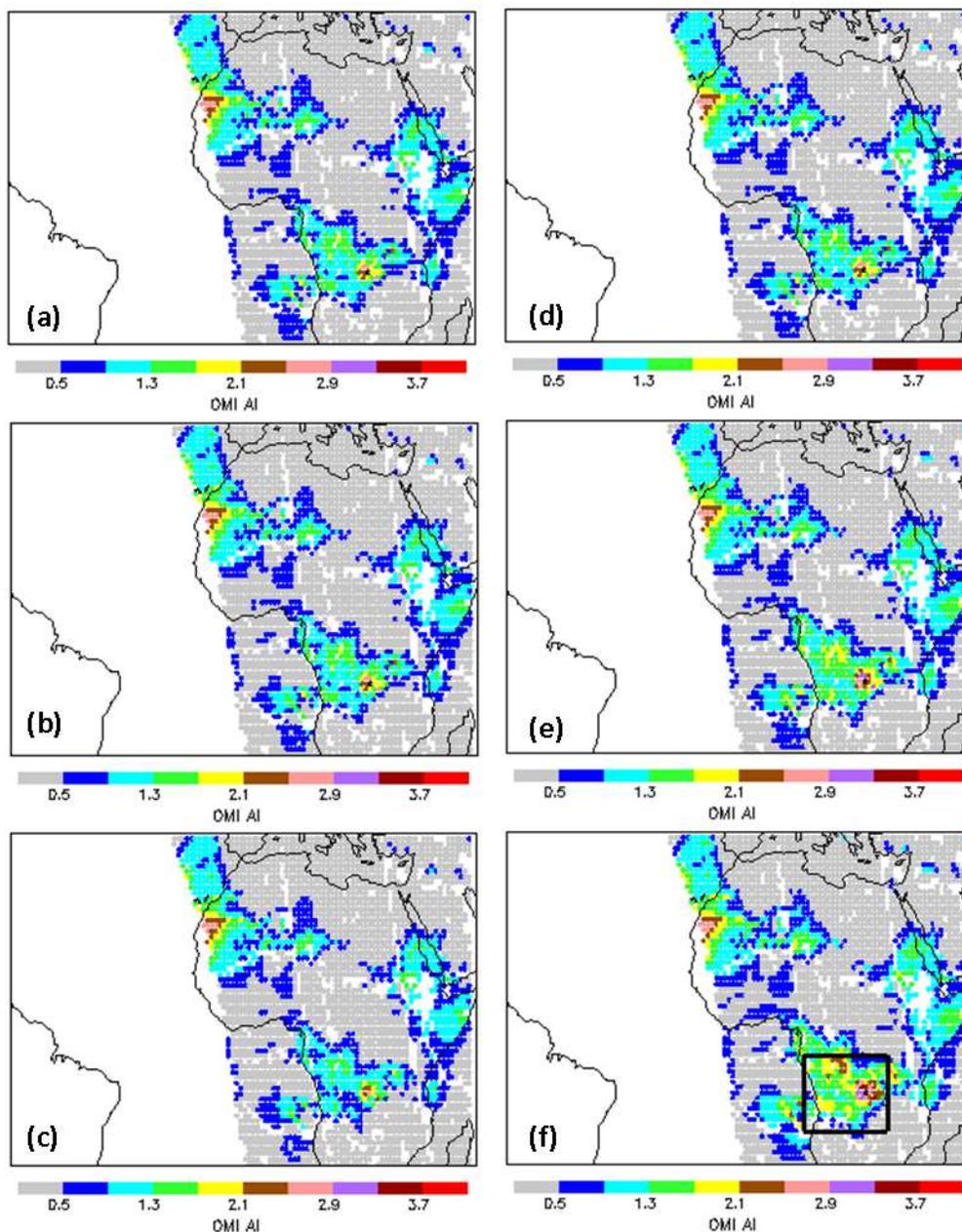


894

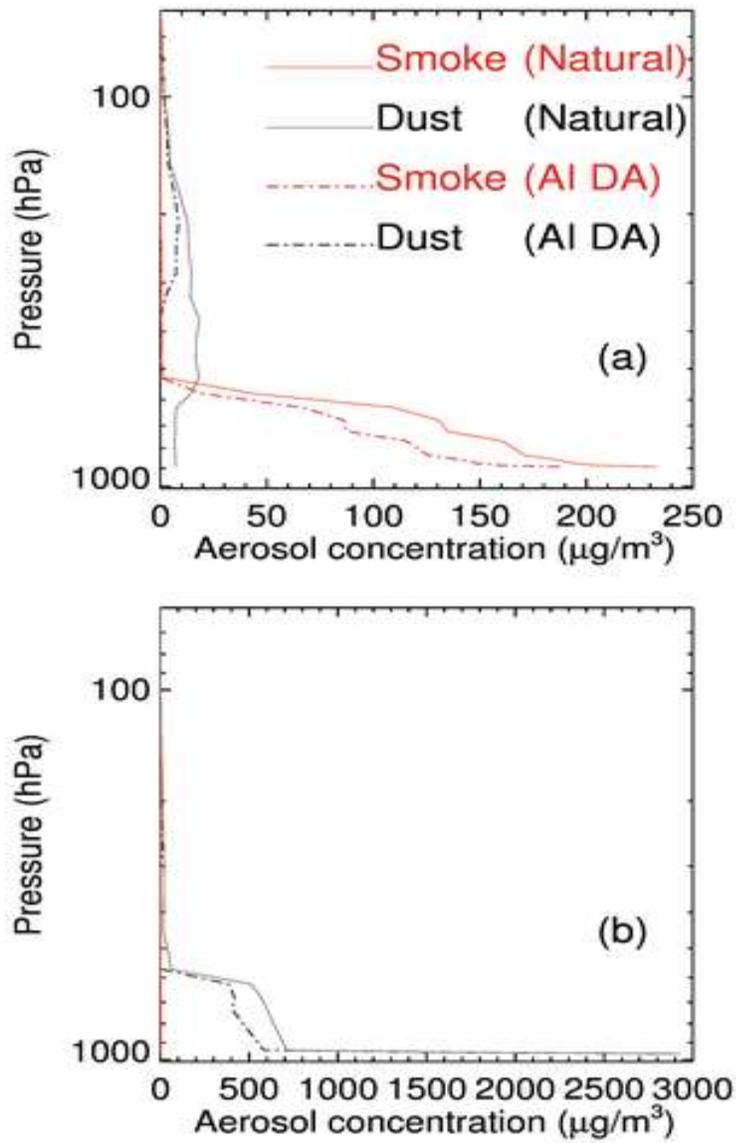
895 **Figure 4.** (a). Spatial distribution of NAAPS AOD using NAAPS data from the AI-DA runs for  
 896 July and August 2007. Only NAAPS data that have collocated OMI AI data are used. (b). Spatial  
 897 distribution of simulated AI for July and August 2007 using NAAPS data from the AI-DA runs.  
 898 (c). Spatial distribution of gridded OMI AI for July and August 2007. (d). Differences between  
 899 Figures 4(b) and 4(c). (e-h) Similar to Figures 4(a)-4(d) but using NAAPS natural runs. Grey  
 900 color highlights those  $1 \times 1^\circ$  (Latitude/Longitude) bins that have less than three collocated NAAPS  
 901 and OMI AI data for the study period.  
 902



904 **Figure 5.** (a). Scatter plot of AERONET and NAAPS AOD ( $0.55 \mu\text{m}$ ) using NAAPS data from  
 905 the natural runs for July-August 2007 over the study region. (b). Similar to Figure 5(a) but using  
 906 NAAPS data from the AI-DA runs. (c). Similar to Figure 5(a) but with AODs taken from the  
 907 NAAPS reanalysis.  
 908



909  
 910 **Figure 6.** Spatial distributions of simulated AI at 12 Z on July 28, 2007 using NAAPS reanalysis  
 911 data, with single scattering albedos of smoke aerosol at 354 and 388 nm taken to be: (a) 0.84 and  
 912 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.  
 913



914

915 **Figure 7.** (a). Vertical distributions of smoke and dust aerosol concentrations over 9.5°S and  
 916 10.5°E at 12 Z on July 28, 2007 for both natural and AI DA runs. (b). Similar as (a) but over  
 917 25.5°N and 12.5°W.

918

919