Analysis of the MODIS Above-Cloud Aerosol Retrieval Algorithm Using MCARS

Galina Wind[1,2], Arlindo M. da Silva[2], Kerry G. Meyer[2], Steven Platnick[2] and Peter M. Norris[3,2]

[1] SSAI, Inc. 10210 Greenbelt Road, Suite 600, Lanham, Maryland 20706, USA

[2] NASA Goddard Space Flight Center, 8800 Greenbelt Rd. Greenbelt, Maryland, 20771, USA

[3] Universities Space Research Association, 7178 Columbia Gateway Dr., Columbia, MD 21046, USA

Correspondence to: G. Wind (Gala.Wind@nasa.gov)

13 **Abstract**

14   The Multi-sensor Cloud and Aerosol Retrieval Simulator (MCARS) presently produces

15 synthetic radiance data from Goddard Earth Observing System version 5 (GEOS-5) model

16 output as if the Moderate Resolution Imaging Spectroradiometer (MODIS) was viewing a

17 combination of atmospheric column inclusive of clouds, aerosols and a variety of gases and

18 land/ocean surface at a specific location. In this paper we use MCARS to study the MODIS

19 Above-Cloud AEROsol retrieval algorithm (MOD06ACAERO). MOD06ACAERO is

20 presently a regional research algorithm able to retrieve aerosol optical thickness over clouds,

21 in particular absorbing biomass burning aerosols overlying marine boundary layer clouds in

22 the Southeastern Atlantic Ocean. The algorithm's ability to provide aerosol information in

23 cloudy conditions makes it a valuable source of information for modeling and climate studies

24 in an area where current clear sky-only operational MODIS aerosol retrievals effectively have

25 a data gap between the months of June and October. We use MCARS for a verification and

26 closure study of the MOD06ACAERO algorithm. The purpose of this study is to develop a set

27 of constraints a model developer might use during assimilation of MOD06ACAERO data.

28   Our simulations indicate that the MOD06ACAERO algorithm performs well for marine

29 boundary layer clouds in the SE Atlantic provided some specific screening rules are observed.

30 For the present study, a combination of five simulated MODIS data granules was used for a

31 dataset of 13.5 million samples with known input conditions. When pixel retrieval uncertainty

32 was less than 30%, optical thickness of the underlying cloud layer was greater than 4 and

33 scattering angle range within the cloud bow was excluded, MOD06ACAERO retrievals

34 agreed with the underlying ground truth (GEOS-5 cloud and aerosol profiles used to generate

35 the synthetic radiances) with a slope of 0.913, offset of 0.06, and RMSE=0.107. When only

36 near-nadir pixels were considered (view zenith angle within +/-20 degrees) the agreement

37 with source data further improved (0.977, 0.051 and 0.096 respectively). Algorithm closure

38   was examined using a single case out of the five used for verification. For closure, the

39   MOD06ACAERO code was modified to use GEOS-5 temperature and moisture profiles as

40   ancillary. Agreement of MOD06ACAERO retrievals with source data for the closure study

41   had a slope of 0.996 with offset -0.007 and RMSE of 0.097 at pixel uncertainty level of less

42   than 40%, illustrating the benefits of high-quality ancillary atmospheric data for such

43   retrievals.

1   Introduction

The MODerate resolution Imaging Spectroradiometer (MODIS) (Barnes et al., 1998) has proven to be an important sensor for aerosol data assimilation purposes for models such as the Goddard Earth Observing System Model, Version 5 (GEOS-5; Rienecker et al. 2008, Molod et al. 2012). There are two MODIS instruments on board NASA's Earth Observing System (EOS) *Terra* and *Aqua* spacecraft. There is a wide variety of data products available from these instruments for Land, Ocean and Atmosphere disciplines. Atmosphere discipline products include cloud mask, cloud top properties, cloud optical and microphysical properties and atmospheric aerosol properties. The MODIS data product files use a designation of MOD for Terra MODIS and MYD for Aqua MODIS. In this paper for brevity we will use "MOD" to refer to both instruments.

The largest contributor of biomass burning aerosols is Southern Africa (Reid et al, 2009, van der Werf et al, 2010, Chang et al, 2021). Biomass burning occurring from June through October creates thick smoke plumes that extend over the adjacent Atlantic Ocean. Prevailing winds in the area transport the smoke over the Southeast Atlantic Ocean (SEAO) and then as far as the Americas (Swap et al., 1996). The same time period coincides with a near-persistent layer of marine boundary-layer (MBL) stratus cloud that extends for several hundred miles westward from the Namibian coast (Devasthale and Thomas, 2011). The MODIS Dark Target aerosol retrieval algorithm (MOD04) that is used for ocean retrievals operates in clear sky conditions only. MOD04_DT retrievals are not provided for each individual MODIS pixel-level, but rather are performed over a 3x3 or 10x10 set of pixels. Moreover aerosol properties are not retrieved over sun glint regions (Kaufman et al, 1997, Levy et al, 2009, 2013). The SEAO region has both extensive seasonal cloud cover and a significant portion of MODIS

68 granules containing sun glint, leading to equally extensive loss of continuous observations

69 from the area.

70   Figure 1 illustrates these conditions using Terra MODIS data from 2006 through 2013.

71 Panel a) shows the percentage of ocean gridboxes in the SEAO area that had daily mean cloud

72 fraction greater than 50% in the MODIS Daily Level-3 gridded product (Hubanks, et al. 2019)

73 stored at 1x1 degree resolution. Here, the SEAO area is defined the same way as in Meyer et

74 al (2015), specifically between -20 and +20 degrees longitude and +4 to -20 degrees

75 longitude. As much as 60% of all ocean gridboxes have cloud fraction greater than 50% in

76 June (day 152) and only increase to the end of September (day 304). A 1-degree resolution

77 gridbox will contain some clear sky and thus at least some aerosol retrievals are possible. As

78 shown in Figure 1 b), in June between 70-80% of all ocean gridboxes contain some aerosol

79 retrievals, though by September that number drops to between 30-50% year over year.

80   Due to aforementioned limitations of the standard dark-target MODIS aerosol algorithm,

81 a model that assimilates aerosol data from SEAO would have very few aerosol retrievals over

82 the ocean available to it. Most of the transport mechanism in the model would be thus

83 governed by the model physical processes (e.g., advection, sedimentation and wet removal

84 and vertical transport) instead of being constrained by observations.

85   The MOD06ACAERO algorithm (Meyer et al. 2015) fills in the aerosol data gap in

86 SEAO as it is able to perform retrievals of aerosol properties above MBL clouds. The

87 algorithm has been evaluated against observations from the Cloud-Aerosol Lidar and Infrared

88 Pathfinder Satellite Observation (CALIPSO) (Winker et al, 2009), but CALIPSO only

89 provides data at nadir and with a very limited spatial coverage. Recent improvements in

90 CALIPSO version 4 aerosol products (Kim et al, 2018) indicate that the comparisons shown

91 of the MOD06ACAERO algorithm with CALIPSO in Meyer et al (2015) would improve

5

92   somewhat as significant work had been done to remedy the low bias that CALIPSO retrievals

93   have. However, Kim et al (2018) state that the remaining SEA low bias in CALIPSO

94   retrievals of AOD with respect to AERONET and MODIS makes CALIPSO retrievals

95   somewhat problematic as means of aerosol algorithm evaluation for SEAO area. (e.g., Meyer

96   et al, 2013, 2015, Jethva et al, 2014). Observations collected during the ObseRvations of

97   Aerosols above CLouds and their IntEractionS (ORACLES) (Redemann et al, 2021) are

98   currently being used to evaluate the MOD06ACAERO algorithm. Additional descriptions of

99   ORACLES aerosol data can be found in LeBlanc et al (2020) and Pistone et al (2019).

100   In this study we applied an Observing System Simulation Experiment (OSSE) framework

101   to gain insight on the performance of the MOD06ACAERO algorithm. Rather than using the

102   classic analysis/forecast error metric common in Numerical Weather Prediction OSSE studies

103   (e.g., Hoffman and Atlas 2016) we adopt here a "Retrieval OSSE" perspective where the

104   quality of the retrieval is used as the verification metric (Wind et al. 2013, 2016). A radiative

105   transfer code is applied to the model quantities combined with sensor geometry to simulate

106   how a model scene appears to a specific instrument. A retrieval algorithm designed for that

107   instrument can be executed on the simulated measurements. Physical quantities retrieved by

108   the algorithm can be compared to the known simulation input. The algorithm can be examined

109   for closure over a large spatial domain and thus any areas or conditions that may be

110   problematic for the algorithm could be examined, and the strengths and limitations of the

111   algorithm can be extensively documented.

112   The Multi-sensor Cloud and Aerosol Retrieval Simulator (MCARS) is a tool that

113   combines model output with a radiative transfer code in order to simulate radiances that may

114   be measured by a remote sensing instrument if it were passing over the model fields (Wind et

115   al, 2013, 2016). In this paper, MCARS continues to use the combination of the GEOS-5

116   model, correlated-k models of atmospheric transmittance due to various gaseous absorbers for

117   MODIS channels as per Kratz (1995), inline Rayleigh scattering and the Discrete Ordinate

118   Radiative Transfer (DISORT) code (Stamnes et al. 1988) to simulate MODIS radiances. Two

119   improvements have been made to the MCARS code since last publication. The computational

120   resolution has been increased to 32 streams, up from 16. Additionally, for this study the

121   higher resolution 7 km GEOS-5 Nature Run (G5NR) was used in place of the standard 25 km

122   resolution GEOS-5 output (Gelaro et al. 2015, da Silva et al. 2015, Putman et al. 2015).

123   G5NR is a 2-year global, non-hydrostatic mesoscale model dataset for the period 2005-2006

124   produced with the GEOS-5 Atmospheric GCM. The model run is performed at a horizontal

125   resolution of 7 km using a cubed-sphere horizontal grid with 72 vertical levels, extending up

126   to 0.01 hPa (~ 80 km). In addition to standard meteorological parameters (wind, temperature,

127   moisture, surface pressure), this GCM includes 15 aerosol tracers (dust, sea-salt, sulfate, black

128   and organic carbon), O_3 and CO_2. The GEOS-5 NR is driven by prescribed sea-surface

129   temperature and sea-ice, daily volcanic and biomass burning emissions, as well as high-

130   resolution inventories of anthropogenic sources. A description of the GEOS-5 model

131   configuration used for the Nature Run can be found in Putman et al. (2014), while results

132   from a validation exercise appear in Gelaro et al. (2015) and Castellanos et al. (2019).

133   In a previous study of the MOD04_DT code (Wind et al, 2016), we had the advantage of

134   having simultaneous in situ aerosol property measurements from AErosol RObotic NETwork

135   (AERONET) (Holben et al., 1998). AERONET has very limited data available over ocean,

136   mainly from islands and ship transits. Even in places where AERONET is established, no

137   measurements can be obtained in presence of clouds. Therefore, no ground-based in-situ

138   measurements can be included in our analysis of the MOD06ACAERO product and so the

139   analysis is necessarily limited to verification and closure.

140   In sections that follow we will describe the application of MCARS to study the

141   MOD06ACAERO algorithm. Section 2 very briefly describes the MCARS code and the

142   experiment setup. Section 3 describes the MODIS MOD06ACAERO product of Meyer et al.

143   (2015). Section 4 shows the details of the study and study conclusions. Finally, section 5

144   discusses the next steps in MCARS development.

145   **2   MCARS description**

146   The MCARS code was previously described in detail in Wind et al (2013, 2016).

147   Therefore, only a brief description will be given here. Global aerosol, cloud, surface and

148   atmospheric column fields from the G5NR simulation as described above serve as the starting

149   point for radiance simulations. The GOCART bulk aerosol scheme currently used in the

150   G5NR is used for the simulations reported in this paper, with corresponding optical properties

151   as described in Randles et al. (2017), Hess et al (1998) and references within. The simulation

152   input data was produced in accordance with the methods outlined in Wind et al. (2016). The

153   G5NR model output was split into 1-km subcolumns (MODIS pixel resolution) using the

154   independent column approximation method as described in detail in Wind et al. (2013). Here

155   a brief summary of the model data preparation methodology is given.

156   MODIS pixels for each GEOS-5 gridbox were collected and the same number of pixel-

157   like sub-columns was generated using a statistical model of sub-gridcolumn moisture

158   variability. The sub-column generation used a parameterized probability density function

159   (PDF) of total water content for each model layer and a Gaussian copula to correlate these

160   PDFs in the vertical (Norris et al, 2008, Norris and da Silva 2016a,b).

161   The subcolumns generated in this way were subsequently rearranged, to give horizontal

162   spatial coherence, by using a horizontal Gaussian copula applied to condensed water path.

163   This arrangement had to be applied in order to create spatially coherent cloud-like structures.

164 The subcolumns themselves were not altered in any way during this process. If this step is

165 skipped and the subcolumns are placed randomly within each gridbox the MODIS Cloud

166 Optical and Microphysical Properties (MOD06) product (Platnick et al, 2017) would restore

167 many of the pixels to clear sky unless the initial gridbox had close to 100% cloud fraction

168 (Zhang and Platnick 2011; Pincus et al. 2012). The MOD06 product is a necessary input for

169 MOD06ACAERO and must be produced prior to MOD06ACAERO execution. The need for

170 this subcolumn rearrangement is significantly lessened when G5NR is used because the

171 smaller gridboxes are often close to 100% cloudy especially in MBL regimes, but removing

172 the method from the model preparation step was not practical due to its small impact on

173 execution time and possibility of introducing errors.

174   The layer aerosol properties were obtained using the independent column approximation

175 with the same PDF of total water content as used for clouds. A GEOS-5 aerosol species

176 output file was used in conjunction with aerosol optical properties as in Randles et al. (2017).

177 The aerosol phase functions for each of the 15 species output by GEOS-5 were produced and

178 combined on the fly to create a single bulk set of scattering properties and Legendre

179 coefficients. (Wind et al, 2016)

180   Model parameters such as profiles of temperature, pressure, ozone and water vapor

181 together with layer information about clouds and aerosols are combined with solar and view

182 geometry of the MODIS instrument. Surface information is also a combination of GEOS-5

183 information of surface temperature, snow and sea ice cover and MODIS-derived spectral

184 surface albedo (Moody et al. 2007, 2008). All of these parameters are transferred to the

185 DISORT-5 radiative transfer code and reflectances and radiances in 22 MODIS channels

186 between 470nm and 14.2μm are produced. The default computational resolution of DISORT-

187 5 has also been increased to 32 streams up from 16 used in the two previous studies.

188 Additionally some of the simulations in this study were executed at 64 streams. Final MCARS

189 output is packaged in a format identical to the standard MODIS Level-1B radiometric files

190 and is thus completely transparent to any operational or research-level retrieval algorithm

191 code.

192   These simulations were produced at the NASA Center for Climate Simulations (NCCS)

193 supercomputer. Each complete simulation of a MODIS-like granule requires 5.5 hours of wall

194 clock time on 300 processors. Computational throughput can be increased by limiting the

195 scope of the simulation to fit a particular investigation. For this study, however, we retain the

196 full set of channels needed for both cloud and aerosol research.

197

198 **3   MODIS above-cloud aerosol properties product**

199

200   The MODIS above-cloud aerosol properties product (MOD06ACAERO) (Meyer et al.

201 2015) is a regional algorithm able to simultaneously retrieve MBL cloud optical thickness

202 (COT), cloud effective radius, and aerosol optical depth (AOD) above-cloud in the SEAO

203 region. It uses six MODIS channels (bands 1-5 and 7) having central wavelengths of 0.47,

204 0.55, 0.66, 0.86, 1.24 and 2.1μm. The MOD06ACAERO algorithm takes advantage of the

205 strong biomass burning aerosol absorption gradient in the visible (VIS) to near-infrared (NIR)

206 spectrum that, when the aerosol layer overlies a bright cloud, yields differential attenuation

207 (stronger at shorter wavelengths) of the otherwise nearly spectrally invarient top-of-

208 atmosphere cloud reflectance across the VIS/NIR. Sensitivity to cloud optical thickness is

209 localized in the spectral range between 0.47 and 1.24μm and is directly related to the

210 magnitude of reflectance, while sensitivity to above-cloud aerosol optical depth is related to

211 the spectral slope of the reflectance. The MOD06ACAERO algorithm uses 2.1μm channel for

212 cloud effective radius information. That is also consistent with the principal retrieval

213 contained in the MOD06 product (Platnick, et al, 2017)

214   The MOD06ACAERO retrieval inversion uses an optimal estimation-like approach

215   (Rodgers, 1976) that attempts to minimize the difference (cost function) between the six

216   MODIS reflectance observations and forward-modeled reflectance that is a function of cloud

217   optical thickness, effective radius, and above-cloud AOD. However, rather than in-line

218   radiative transfer calculations, MOD06ACAERO relies on a set of pre-computed lookup

219   tables (LUTs) of coupled cloud and above-cloud aerosol reflectance. These LUTs are

220   generated using the same cloud microphysics models used by MOD06 (Platnick et al, 2017)

221   and the absorbing aerosol model used by MOD04_DT over land surfaces (Levy et al, 2013).

222   Retrievals using a second aerosol property model, one based on field campaign data from

223   SAFARI 2000 (Haywood et al, 2003), are also available in MOD06ACAERO output. While

224   these Haywood et al. model retrievals were recommended in Meyer et al (2015), evaluation

225   during the ORACLES campaign revealed deficiencies at certain scattering angle ranges (K.

226   Meyer, private communication). Thus, for this study we use the MOD06ACAERO results

227   based on the MOD04_DT aerosol models.

228   The MOD06ACAERO retrieval operates at 1km resolution, compared to the 10km and

229   3km MOD04_DT resolutions, and simultaneously provides pixel-level estimates of retrieval

230   uncertainty accounting for known and quantifiable error sources (e.g., radiometry,

231   atmospheric profile errors, cloud and aerosol forward model errors) consistent with the

232   MOD06 cloud product methodology (Platnick et al, 2020). Figure 2 shows an example

233   retrieval result from MOD06ACAERO compared to MOD04_DT standard 10km output. The

234   Terra MODIS granule shown here, from 2006 day 224 at 10:05 UTC, has extensive cloud

235   cover over the ocean, typical for this season. MOD04_DT provides a very limited amount of

236   data, localized to the few areas of clear sky, while MOD06ACAERO fills in the above-cloud

237   area. Shinozuka, et al (2020) suggest that above-cloud aerosol retrievals are similar to

238   adjacent clear-sky retrievals and so clear-sky retrievals could be used as an above-cloud

11

239   proxy. However conditions shown in Figure 2 are common during the SEAO burning season.

240   There are no clear-sky retrievals of aerosol over most of the area due to near uniform

241   coverage by marine stratus, with cloud fraction approaching 80%. Nearest successful clear-

242   sky retrievals are hundreds of miles away. Therefore an above-cloud aerosol retrieval

243   algorithm such as MOD06ACAERO is very much so necessary.

244   MOD06ACAERO uses National Center for Environmental Prediction (NCEP)

245   atmospheric profile products (Derber et al, 1991) for atmospheric correction. As part of our

246   investigation we will look at impact of discrepancies between NCEP and G5NR on retrieved

247   aerosol properties.

248

249   **4   Analysis**

250

251   To create the data used for the MOD06ACAERO verification study, we examined the

252   G5NR dataset for cases that were similar to conditions commonly encountered during the

253   burning season over SEAO. August 2006 was selected because it was a very active smoke

254   season and a significant amount of MBL clouds were present in the model output. Models

255   often have difficulties forming MBL clouds as higher than usual grid and vertical resolution is

256   needed in order to accurately represent the processes that lead to MBL formation in nature.

257   As real Terra and Aqua overpasses are needed in order to define the sun-satellite

258   geometry for the MCARS simulations, satellite orbital tracks had to be considered. Because

259   orbital gaps are prominent in the MODIS data over the SEAO MBL region, care must be

260   taken in selecting specific days and times having adequate sensor geometry. Technically

261   because MCARS is a simulation, orbital gaps have no meaning. But because of the need of

262   actual sensor geometry to start the simulation, it is most expedient to simply browse available

263   MODIS data for a suitable track. Even though G5NR does not perform any data assimilation,

12

264   the model code is identical to the standard GEOS-5 model. MCARS normally runs on

265   standard GEOS-5 output. In Wind et al (2013) we showed MCARS as a model output

266   verification tool. It is always very desirable to match date/time/orbit when model performance

267   may be compared to real concurrent sensor measurements. Even though no orbital match is

268   required in this study, a decision was made to not alter the standard MCARS operation in

269   order to avoid accidental introduction of software issues. Five cases were selected under these

270   considerations. Three came from Terra MODIS overpasses and two from Aqua MODIS. The

271   times and dates were as follows. Terra MODIS: 2006 day 224, 10:05 UTC, 2006 day 225

272   09:10 UTC, 2006 day 228 09:40 UTC. Aqua MODIS: 2006 day 224 12:55 UTC and 2006 day

273   226 12:40 UTC. This simulated radiance dataset comprises 13.5 million points where the

274   atmospheric column and surface conditions are explicitly known. MOD06ACAERO retrievals

275   were attempted over those points, but of course that does not mean that each attempt produced

276   a successful aerosol retrieval.

277   Figure 3 a) shows simulated RGB images for the 5 MCARS MODIS granules listed

278   above. Also shown in b) are the same simulated granules where the aerosols have been

279   removed from the radiative transfer simulations. This ability to remove clouds, aerosols or

280   gases from the simulation offers extensive control evaluating the performance of retrieval

281   algorithms and diagnosing algorithm deficiencies.

282   There is a significant similarity between the real Terra MODIS granule of Figure 2 and

283   the simulated granule for the same date and time. The G5NR is a free running model and does

284   not perform any data assimilation, and therefore it is not synoptically locked to the particular

285   day depicted in Figure 2. The apparent similarities between Figures 2 and 3 merely reflect the

286   persistent patterns of MBL clouds and smoke in the region. There is no expectation of a match

287   with any real data in this study. It is not a statement to G5NR performance as in other cases

288   the cloud amount/distribution had no match to any real data. It is merely an interesting

289   coincidence. Some granules were selected to include a significant portion of land surface for a

290   later examination of the MOD04_DT retrievals, repeating the study in Wind et al (2016) in a

291   different region (not reported here).

292   This dataset, both the complete and the clean (aerosol-free) versions, was fed through the

293   standard operational MODIS Data Collection 6 cloud product processing chain to produce

294   cloud mask, MOD06 cloud top and optical properties, and finally the MOD06ACAERO

295   output for each case. Results from all granules were then combined and only retrievals for

296   cloudy pixels were examined. The MOD06ACAERO aerosol retrievals were compared to

297   source aerosol optical depth provided by GEOS-5 (Wind et al, 2016). Figure 4 shows results

298   of this comparison. The only constraint on this comparison was that the algorithm-reported

299   pixel-level retrieval uncertainty had to be less than 40% for panel a) and less than 30% for

300   panel b). One of the motivations of this study was to characterize errors in the

301   MOD06ACAERO algorithm for subsequent aerosol data assimilation into GEOS-5. Pixels

302   with higher uncertainties could be considered in the analysis, but assimilating data where the

303   retrieval error is 50% or greater could negatively impact the assimilated fields. As depicted in

304   Figure 4, filtering retrievals at the reported algorithm uncertainty at 40% is very effective to

305   produce a good match between MOD06ACAERO and the G5NR output variables, with the

306   exception of very low AODs. G5NR uses aerosol models described in detail in Randles et al

307   (2017). It is a set of 15 absorbers, properties of which are a function of column relative

308   humidity. MOD06ACAERO in this study uses the MOD04_DT aerosol models, which are

309   distinct in composition and additionally computed at a constant 80% column relative humidity

310   (Levy et al, 2013). Because G5NR mixes aerosols on-the-fly to create bulk layer properties

311   and MOD06ACAERO has a constant regional mixture, there is a natural source of uncertainty

312   in any comparison of MOD06ACAERO retrievals with G5NR. However the regional mixture

313   of MOD04_DT had been used extensively to train the GOCART model used by both GEOS-5

314   and G5NR. Thus we expect the uncertainty due to aerosol model mismatch to be fairly

315   minimal. Same exact situation of aerosol mixture mismatch exists in real data and is most

316   likely greater than the one existing in this simulation. Detailed comparison of GOCART and

317   MOD04_DT aerosol models for biomass burning aerosols has been performed in Wind et al

318   (2016).

319   Meyer et al. (2015) suggest that additionally MOD06ACAERO retrievals should be

320   screened by retrieved cloud optical thickness and that they should be discarded if COT is less

321   than 4.0. We applied this additional constraint onto the retrieval comparison and the result is

322   shown in Figure 5. Discarding the AOD retrievals when cloud is thin improved the match-up

323   against GEOS-5, but there still appears to be an issue when GEOS-5 AOD is very close to

324   zero.

325   The power of MCARS lies in being able to tightly control simulation parameters. The

326   MOD06ACAERO algorithm appears to run into a difficulty at low source AOD. In order to

327   examine the causes for this discrepancy in more detail, we turn our attention to the clean

328   MCARS case shown in figure 3b) by setting the AOD precisely to zero and examining the

329   retrieval performance in such situation. Ideally MOD06ACAERO should retrieve a zero AOD

330   throughout. With an exception of a narrow range of scattering angles between 135 and 145

331   degrees, which corresponds to the cloud bow direction, the algorithm indeed retrieved AOD

332   that was extremely close to zero. Figure 6 depicts the difference between retrieval and source

333   as a function of scattering angle. Retrievals where MOD06ACAERO matched GEOS-5

334   precisely were discarded for clarity. Within the cloud bow MOD06ACAERO tends to return a

335   small positive AOD of about 0.15.

336   The liquid water phase function is very complex in the cloudbow region and is very

337   difficult to model accurately. That particular region has consistently caused difficulties to the

338   standard MOD06 product retrievals of MBL clouds. Both MOD06 and MOD06ACAERO

339   LUTs are computed at 64 DISORT streams. We performed some investigation of this area by

340   running a special simulation for a single case from Terra 2006 day 224 10:05 UTC. This case

341   was selected because the cloudbow is especially noticeable in both real and simulated data.

342   The simulation was also executed using 64 DISORT streams in order to reduce uncertainties

343   associated with the simulation being performed at half the resolution. In cloudbow region

344   more streams would potentially lead to a better model. Unfortunately the cloudbow persisted.

345   It thus may be the case that 64 streams are not sufficient to properly resolve the cloudbow in

346   either simulation or retrieval. Even higher resolution may be advisable. Increasing

347   computational resolution of MOD06 LUTs is presently considered for the upcoming MODIS

348   Data Collection 7. Depending on the results, same increase may occur for MOD06ACAERO.

349   At this time, for purpose of establishment of assimilation constraints, which is the focus of

350   this study, one might simply exclude the cloud bow scattering angle range from consideration

351   until more is known.

352   Figure 7 shows the results of MOD06ACAERO retrievals from Figure 5 where retrievals

353   within the cloud bow have been discarded. The comparison with source data is further

354   improved and the cluster of MOD06ACAERO retrievals present in Figure 5 when GEOS-5

355   AOD was near zero has disappeared.

356   Often better retrievals can be obtained when less oblique view geometry is considered in

357   real data. Pixel size, longer optical path length and 3D effects from clouds can all make

358   retrievals performed at oblique view angles less optimal. In the case of this study, another

359   consideration for imposition of a view zenith limit is that presently MCARS does not account

360   for pixel size growth at oblique view angles. The number of subcolumns generated does not

361   change with view zenith angle. Therefore, MCARS results when view angle is oblique may

362   not be an accurate measure of algorithm performance as only the effects of optical path length

363   are simulated.

364   The MOD06 cloud product outputs cloud top pressure, temperature and height limited to

365   near nadir in addition to full swath products. The "near nadir" is defined as viewing zenith

366   angle less than 32 degrees (Menzel et al, 2008). Figure 8 shows the MOD06ACAERO

367   retrievals of Figure 7 further limited by view zenith angle of less than 32 degrees. When view

368   zenith angle is limited to 32 degrees the comparison with GEOS-5 source data is again

369   improved. We can now show a slope of 0.866 for retrievals with less than 40% error and

370   0.913 for retrievals with error of less than 30%. Note that even though the data extent had

371   been limited, there are still over 600,000 data points left to be ingested into a model if data

372   assimilation were to be attempted in an area where previously the number of such data points

373   was close to 0.

374   We can constrain the view zenith angle range even further as shown in Figure 9, reducing

375   the threshold to 20 degrees. Whereas the comparison shows all around improvement with

376   slope of 0.931 and 0.977 for retrieval error of less than 40% and 30% respectively, the number

377   of points suitable for assimilation shrinks by half. It is not clear if this dataset size reduction

378   can be justified by the improvement in alignment with the source data.

379   With the 20 degree view angle constraint the algorithm results are very close to source

380   data and we could potentially state that we have closure against source GEOS-5 data even

381   though both MOD06 and MOD06ACAERO run under operational conditions used NCEP

382   GDAS data for atmospheric correction (implying a likely overestimation of the error in these

383   profiles). In order to assess the impact of using these GDAS-based profiles we consider a final

384   experiment where we use MCARS pixel-level input profiles for atmospheric correction. The

385   result is shown in Figure 10. When atmospheric profiles are removed as a source of

386   inconsistency, the agreement with source data improves to a slope of 0.996 with intercept of -

387   0.007 and RMSE of 0.097 for retrievals with less than 40% error and slope of 0.989, intercept

388   of 0.03 and RMSE of 0.085 for retrievals with less than 30% error. Small sample size for

389   retrievals with lower uncertainty is the reason for somewhat lesser agreement with source data

390   for this closure experiment. The remaining source of potential disagreement of

391   MOD06ACAERO retrieval with input GEOS-5 data is the difference between aerosol models

392   used by MCARS and MOD06ACAERO. Cloud models between MOD06ACAERO and

393   MCARS are identical in this study. The MOD06ACAERO model is fixed for the region,

394   while the GEOS-5 aerosols are fully dynamic as per Randles et al (2017). However, it is not

395   practical to change either MCARS or MOD06ACAERO code to use a different aerosol model

396   set, and with the agreement being as good as it presently is.

397

398   **5   Conclusions and future directions**

399   This paper is a direct evolution of work started in Wind et al, (2013) and continued in

400   Wind et al (2016). The Multi-sensor Cloud and Aerosol Retrieval Simulator (MCARS) has

401   now been applied as a verification tool for a research-level algorithm. The algorithm studied

402   was the MODIS above-cloud aerosol properties retrieval algorithm of Meyer et al (2015).

403   MCARS computational resolution has been doubled and for this study the high-resolution

404   (7km) GEOS-5 Nature Run model was utilized. The MCARS code produces radiances and

405   reflectances in a standard MODIS Level 1B format after sending the GEOS-5 data through

406   DISORT-5 radiative transfer code. The output can be directly ingested by any retrieval or

407   analysis code that reads data from the MODIS instrument.

408   We used the MCARS code to perform verification and closure study on the

409   MOD06ACAERO algorithm. In this study we generated a set of five MODIS granules located

410   in the Southeastern Atlantic Ocean off the coast of Namibia. We executed the

411   MOD06ACAERO code on this case set. In the verification part of the study the algorithm

412   performed very well. When pixels with less than 30% uncertainty were considered with

413   underlying cloud layer having optical thickness greater than 4 the algorithm matched the

414 source GEOS-5 aerosol optical depth with slope of 0.774 and offset of 0.076, RMSE = 0.131.

415 On further examination, executing the algorithm on the same case set with aerosols removed it

416 was determined that there might be data that is less useful around the scattering angle of 140

417 degrees, the cloud bow direction. When the cloud bow pixels were excluded the slope

418 improved to 0.913. The near-nadir slope with angle limit of 20 degrees improved the

419 agreement further to 0.977, RMSE=0.096.

420   To look at closure one of the five cases was selected. For closure both MOD06 and

421 MOD06ACAERO codes were modified to use MCARS input profiles as ancillary instead of

422 the NCEP analysis used in operations (Platnick et al, 2017). When the results were compared

423 to source GEOS-5 data a slope of 0.996 with offset of -0.007 and RMSE = 0.097 was reached

424 for pixels with less than 40% uncertainty. The agreement was slightly worse for uncertainties

425 less than 30% (slope 0.989, offset 0.03 and RMSE = 0.085) but that was mainly due to having

426 a smaller number of pixels in the set, only 130,000.

427   The results of this study suggest that retrievals produced by MOD06ACAERO are of

428 good initial quality and would be a valuable addition to model data assimilation streams with

429 the following constraints. MOD06ACAERO pixels should be assimilated if retrieval

430 uncertainly is less than 40%, if optical thickness of the underlying cloud layer is greater than

431 4.0 and if the pixel scattering angle is outside the cloud bow. Additionally, an even tighter

432 constraint can be added to only take pixels that are near nadir.

433   This study is yet another example of the capabilities of the MCARS framework. There

434 are many other potential applications of the MCARS code, including extending the simulator

435 to other sensors and examining the performance of fast retrieval simulators used in climate

436 modeling.

437 **6   Code and Data Availability**

438   The MCARS code is free of charge and can be downloaded here:

439   https://zenodo.org/record/5224964#.YR_DOdNKjaV

440

441

442

Author Contributions

GW is the development and experiment design lead on the MCARS project. She maintains the code, creates experiments and performs most of the analysis of experiment data.

AdS and PN assist with preparation, interpretation and integration of the GEOS-5 model data.

KM is the author of MODIS above-cloud aerosol retrieval algorithm, the subject of this simulation experiment. He assisted with interpretation of retrieval results and development of assimilation constraints for the above-cloud aerosol product.

SP assisted with analysis, evaluation and interpretation of all experiment data.

459   **References**

460   Barnes, W. L., T. S. Pagano, and V. V. Salomonson, 1998: Prelaunch characteristics of the
461   Moderate Resolution Imaging Spectroradiometer (MODIS) on EOS-AM1. IEEE Trans.
462   Geosci. Remote Sens., 36, 088–1100.

463   Castellanos, P., da Silva, A., Darmenov, A., Buchard, V., Govindaraju, R., Ciren, P., &
464   Kondragunta, S. (2019). A Geostationary Instrument Simulator for Aerosol Observing
465   System Simulation Experiments. *Atmosphere*, *10*(1), 2–36.
466   http://doi.org/10.3390/atmos10010002

467   Chang, I., Gao, L., Burton, S. P., Chen, H., Diamond, M. S., Ferrare, R. A., et al. (2021).
468   Spatiotemporal heterogeneity of aerosol and cloud properties over the southeast Atlantic:
469   An observational analysis. Geophysical Research Letters, 48, e2020GL091469.
470   https://doi.org/10.1029/2020GL091469
471

472   Chin, M., P. Ginoux, S. Kinne, O. Torres, B. N. Holben, B. N. Duncan, R. V. Martin, J. A.
473   Logan, A. Higurashi, and T. Nakajima, 2002: Tropospheric Aerosol Optical Thickness
474   from the GOCART Model and Comparisons with Satellite and Sun Photometer
475   Measurements. J. Atmos. Sci., 59, 461–483.
476

477   da Silva, A.M., W. Putman and J. Nattala, 2014: File Specification for the 7-km GEOS-5
478   Nature Run, Ganymed Release (Non-hydrostatic 7-km Global Mesoscale Simulation).
479   GMAO Office Note No. 6 (Version 1.0), 176 pp, available from
480   http://gmao.gsfc.nasa.gov/pubs/office_notes.
481

482   Derber J.C., D.F. Parrish and S.J. Lord, 1991: The new global operational analysis system at
483   the National Meteorological Center. Weath. Forec. 6, 538-547.

484   Devasthale, A., and M. A. Thomas 2011: A global survey of aerosol-liquid water cloud
485   overlap based on four years of CALIPSO-CALIOP data. Atmos. Chem. Phys., 11, 1143–
486   1154, doi:10.5194/acp-11-1143-2011

487   Gelaro, R., W. Putman, S. Pawson, C. Draper, A. Molod, P.M. Norris, L.E. Ott, N. Prive, O.
488   Reale, D. Achuthavarier, M. Bosilovich, V. Buchard, W. Chao, L. Coy, R. Cullather, A.M.
489   da Silva, A. Darmenov, R.M. Errico, M. Fuentes, M.J. Kim, R. Koster, W. McCarty, J.
490   Nattala, G. Partyka, S. Schubert, G. Vernieres, Y. Vikhliaev, K. Wargan, 2015: Evaluation
491   of the 7-km GEOS-5 Nature Run. Technical Report. NASA/TM-2014-104606, Goddard
492   Space Flight Center, National Aeronautics and Space Administration.

493   Haywood, J. M., S. R. Osborne, P. N. Francis, A. Keil, P. Formenti, M. O. Andreae, and P. H.
494   Kaye, 2003: The mean physical and optical properties of regional haze dominated by
495   biomass burning aerosol measured from the C-130 aircraft during SAFARI 2000, J.
496   Geophys. Res.,108(D13), 8473, doi:10.1029/2002JD002226.

497   Hess, M., P. Koepke, and I. Schult, 1998: Optical properties of aerosols and clouds: The
498   software package OPAC. B. Am. Meteorol. Soc., 79(5), 831–844.

499   Hill, C., C. DeLuca, V. Balaji, M. Suarez, A. da Silva, 2004: The architecture of the Earth
500   System Modeling Framework, Comp. Sci. Engr., 6(1), 18-28.

501   Hoffman, R. N. and R. Atlas, 2016: Future Observing System Simulation Experiments.
502   BAMS, 97(9), 1601–1616. http://doi.org/10.1175/BAMS-D-15-00200.1

503   Holben, B.N., T. F. Eck, I. Slutsker, D. Tanre, J.P. Buis, A. Setzer, E.F. Vermote, J. A.

504   Reagan, Y.J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, A. Smirnov, 1998:
505   AERONET – A federated instrument network and data archive for aerosol characterization.
506   Rem. Sens. Env., v.66, n1, p1-16.

507   Hsu, N.C., M-J. Jeong, C. Bettenhausen, A.M. Sayer, R. Hansell, C.S. Seftor, J. Huang, S-C.
508   Tsay 2013: Enhanced Deep Blue aerosol retrieval algorithm: The second generation. J.
509   Geophys. Res.-Atmos., 118(16), p 9296-9315.

510   Hubanks, P., S. Platnick, M.D. King, W. Ridgway, 2019: MODIS Level-3 daily gridded
511   product. https://modis-atmos.gsfc.nasa.gov/images/l3-daily-browse

512   Jethva, H., O. Torres, F. Waquet, D. Chand, D., & Y. Hu, 2014: How do A-train sensors
513   intercompare in the retrieval of above-cloud aerosol optical depth? A case study-based
514   assessment. Geophysical Research Letters. http://doi.org/10.1002/(ISSN)1944-8007

515   Kaufman, Y. J., A. E. Wald, L. A. Remer, B. C. Gao, R. R. Li, L. Flynn, 1997: The MODIS
516   2.1µm channel - Correlation with visible reflectance for use in remote sensing of aerosol
517   IEEE Trans. Geosci. Remote Sens., Vol. 35, 1286-1298.

518   Kim, M.-H., Omar, A. H., Tackett, J. L., Vaughan, M. A., Winker, D. M., Trepte, C. R., Hu,
519   Y., Liu, Z., Poole, L. R., Pitts, M. C., Kar, J., and Magill, B. E.: The CALIPSO version 4
520   automated aerosol classification and lidar ratio selection algorithm, Atmos. Meas. Tech.,
521   11, 6107–6135, https://doi.org/10.5194/amt-11-6107-2018, 2018.

522   Kratz, D.P. 1995: The correlated-k distribution technique as applied to the AVHRR channels.
523   J. Quant. Spectrosc. Rad. Trans. 53, 501-517.

524   LeBlanc, S. E., Redemann, J., Flynn, C., Pistone, K., Kacenelenbogen, M., Segal-
525   rosenheimer, M., Shinozuka, Y., Dunagan, S., Dahlgren, R. P., Meyer, K., Podolske, J.,
526   Howell, S. G., Freitag, S., Small-griswold, J., Holben, B., Diamond, M., Wood, R.,
527   Formenti, P., Piketh, S., Maggs-Kölling, G., Gerber, M. and Namwoonde, A.: Above-
528   cloud aerosol optical depth from airborne observations in the southeast Atlantic, Atmos.
529   Chem. Phys., 20, 1565–1590, doi:10.5194/acp-20-1565-2020, 2020.

530   Levy, R. C., S. Mattoo, L. A. Munchak, L. A. Remer, A. M. Sayer, F. Patadia, and N. C.
531   Hsu, 2013: The Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas.
532   Tech., 6, 2989-3034, doi:10.5194/amt-6-2989-2013.

533   Levy, R.C., L. A. Remer, D. Tanre, S. Mattoo, Y.J. Kaufman, 2009: Algorithm for remote
534   sensing of tropospheric aerosol over dark targers from MODIS Collections 005 and 051,
535   revision 2. ATBD Reference Number: ATBD-MOD-04. http://modis-
536   atmos.gsfc.nasa.gov/reference_atbd.html

537   Menzel, W. P., R. Frey, H. Zhang, D. Wylie, C. Moeller, R. Holz, B. Maddux, B. A. Baum,
538   K. Strabala, and L. Gumley, 2008: MODIS global cloud-top pressure and amount
539   estimation: Algorithm description and results. J. Appl. Meteor. Climatol., 47, 1175–1198.

540   Meyer, K. G., S. Platnick. 2015: Simultaneously inferring above-cloud absorbing aerosol
541   optical thickness and underlying liquid phase cloud optical and microphysical properties
542   using MODIS. *J. Geophys. Res. Atmos*, 120 (11): 5524–5547

543   Meyer, K. G., S. Platnick, L. Oreopoulos, and D. Lee. 2013: Estimating the direct radiative
544   effect of absorbing aerosols overlying marine boundary layer clouds in the southeast
545   Atlantic using MODIS and CALIOP. *J. Geophys. Res. Atmos.,* 118 (10): 4801-4815

546   Molod, A., L. Takacs, M. Suarez, J. Bacmeister, I.-S. Song, and A. Eichmann, 2012: The

547   GEOS-5 Atmospheric General Circulation Model: Mean Climate and Development from
548   MERRA to Fortuna. Tech. Rep. S. Gl. Mod. Data Assim., 28

549 Moody, E. G., M. D. King, C. B. Schaaf, D. K. Hall, and S. Platnick, 2007: Northern
550   Hemisphere five-year average (2000-2004) spectral albedos of surfaces in the presence of
551   snow: Statistics computed from Terra MODIS land products. Remote Sens. Environ., 111,
552   337–345.

553 Moody, E. G., M. D. King, C. B. Schaaf and S. Platnick, 2008: MODIS-derived spatially
554   complete surface albedo products: Spatial and temporal pixel distribution and zonal
555   averages. J. Appl. Meteor. Climatol., 47, 2879–2894.

556 Norris, P. M., L. Oreopoulos, A. Y. Hou, W.-K. Tao, X. Zeng, 2008: Representation of 3D
557   heterogeneous cloud fields using copulas: Theory for water clouds. J. Q. R. Meteorol. Soc.
558   134: 1843–1864. doi:10.1002/qj.321.

559 Norris, P. M., and A. M. da Silva. 2016a: Monte Carlo Bayesian inference on a statistical
560   model of sub-gridcolumn moisture variability using high-resolution cloud observations.
561   Part 1: Method. Q.J.R. Meteorol. Soc., 142 (699): 2505-2527.

562 Norris, P. M., and A. M. da Silva. 2016b: Monte Carlo Bayesian inference on a statistical
563   model of sub-gridcolumn moisture variability using high-resolution cloud observations.
564   Part 2: Sensitivity tests and results. Q.J.R. Meteorol. Soc., 142 (699): 2528-2540.

565 Notarnicola, C., D. Di Rosa, F. Posa, 2011: Cross-Comparison of MODIS and CloudSat Data
566   as a Tool to Validate Local Cloud Cover Masks. Atmos., 2, 242-255,
567   doi:10.3390/atmos2030242

568 Pincus, R., S. Platnick, S. A. Ackerman, R. S. Hemler, and R. J. P. Hofmann, 2012:
569   Reconciling simulated and observed views of clouds: MODIS, ISCCP, and the limits of
570   instrument simulators. J. Climate, 25, 4699-4720. doi:10.1175/JCLI-D-11-00267.1.

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

577 Platnick, S., K.G. Meyer, M.D. King, G. Wind, N. Amarasinghe, B. Marchant, G.T. Arnold, Z.
578   Zhang, P.A. Hubanks, R.E. Holz, P. Yang, W. Ridgway and J. Riedi, 2017: The MODIS
579   Cloud Optical and Microphysical Products: Collection 6 Updates and Examples from Terra
580   and Aqua. IEEE Trans. Geo. Rem. Sens. 55(1), 502-525, doi:10.1109/TGRS.2016.2610522

581 Platnick, S., K.G. Meyer, N. Amarasinghe, G. Wind, P.A. Hubanks and R.E. Holz 2020:
582   Sensitivity of Multispectral Imager Liquid Water Cloud Microphysical Retrievals to the
583   Index of Refraction. Rem. Sens. 2020, 12, 4165; doi:10.3390/rs12244165

584 Putman W., A.M. da Silva, L.E. Ott and A. Darmenov, 2014: Model Configuration for the 7-
585   km GEOS-5 Nature Run, Ganymed Release (Non-hydrostatic 7 km Global Mesoscale
586   Simulation). GMAO Office Note No. 5 (Version 1.0), 18 pp, available from
587   http://gmao.gsfc.nasa.gov/pubs/office_notes.

588 Randles, C. A., A.M. da Silva, and V Buchard, 2017: The MERRA-2 Aerosol Reanalysis,
589   1980–onward, Part I: System Description and Data Assimilation Evaluation. *Journal of*
590   *Climate*. http://doi.org/10.1175/JCLI-D-16-0609.s1

591

592 Redemann, J., Wood, R., Zuidema, P., Doherty, S. J., Luna, B., LeBlanc, S. E., Diamond, M.
593   S., Shinozuka, Y., Chang, I. Y., Ueyama, R., Pfister, L., Ryoo, J.-M., Dobracki, A. N., da
594   Silva, A. M., Longo, K. M., Kacenelenbogen, M. S., Flynn, C. J., Pistone, K., Knox, N. M.,
595   Piketh, S. J., Haywood, J. M., Formenti, P., Mallet, M., Stier, P., Ackerman, A. S., Bauer,
596   S. E., Fridlind, A. M., Carmichael, G. R., Saide, P. E., Ferrada, G. A., Howell, S. G.,
597   Freitag, S., Cairns, B., Holben, B. N., Knobelspiesse, K. D., Tanelli, S., L'Ecuyer, T. S.,
598   Dzambo, A. M., Sy, O. O., McFarquhar, G. M., Poellot, M. R., Gupta, S., O'Brien, J. R.,
599   Nenes, A., Kacarab, M., Wong, J. P. S., Small-Griswold, J. D., Thornhill, K. L., Noone, D.,
600   Podolske, J. R., Schmidt, K. S., Pilewskie, P., Chen, H., Cochrane, S. P., Sedlacek, A. J.,
601   Lang, T. J., Stith, E., Segal-Rozenhaimer, M., Ferrare, R. A., Burton, S. P., Hostetler, C.
602   A., Diner, D. J., Seidel, F. C., Platnick, S. E., Myers, J. S., Meyer, K. G., Spangenberg, D.
603   A., Maring, H., and Gao, L.: An overview of the ORACLES (ObseRvations of Aerosols
604   above CLouds and their intEractionS) project: aerosol–cloud–radiation interactions in the
605   southeast Atlantic basin, Atmos. Chem. Phys., 21, 1507–1563, https://doi.org/10.5194/acp-
606   21-1507-2021, 2021.

607 Reid, J. S., et al., 2009: Global monitoring and forecasting of biomass-burning smoke:
608   Description of and lessons from the Fire Locating and Modeling of Burning Emissions
609   (FLAMBE) program, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 2, 144–162,
610   doi:10.1109/JSTARS.2009.2027443.

611 Rienecker, M.M., M. J. Suarez, R. Todling, J. Bacmeister, L. Takacs, H.-C. Liu, W. Gu, M.
612   Sienkiewicz, R. D. Koster, R. Gelaro, I. Stajner, and J. E. Nielsen, 2008: The GEOS-5 Data
613   Assimilation System - Documentation of Versions 5.0.1, 5.1.0, and 5.2.0. Tech. Rep. S. Gl.
614   Mod. Data Assim., 27

615 Rodgers, C. D. (1976), Retrieval of atmospheric temperature and composition from remote
616   measurements of thermal radiation, Rev. Geophys. Space Phys., 14, 609–624.

617 Shinozuka, Y., Kacenelenbogen, M. S., Burton, S. P., Howell, S. G., Zuidema, P., Ferrare, R.
618   A., LeBlanc, S. E., Pistone, K., Broccardo, S., Redemann, J., Schmidt, K. S., Cochrane, S.
619   P., Fenn, M., Freitag, S., Dobracki, A., Segal-Rosenheimer, M. and Flynn, C. J.: Daytime
620   aerosol optical depth above low-level clouds is similar to that in adjacent clear skies at the
621   same heights: airborne observation above the southeast Atlantic, Atmos. Chem. Phys.,
622   20(19), 11275–11285, doi:10.5194/acp-20-11275-2020, 2020.

623 Swap, R., M. Garstang, S. A. Macko, P. D. Tyson, W. Maenhaut, P. Artaxo, P. Kallberg, and
624   R. Talbot (1996), The long-range transport of southern African aerosols to the tropical
625   South Atlantic, J. Geophys. Res., 101, 23,777–23,791, doi:10.1029/95JD01049.

626 van der Werf, G. R., et al., 2010: Global fire emissions and the contribution of deforestation,
627   savanna, forest, agriculture, and peat fires (1997–2009), Atmos. Chem. Phys., 10, 11,707–
628   11,735, doi:10.5194/acp-10-11707-2010.

629 Wind, G., A.M. da Silva, P. M. Norris, and S. Platnick, 2013: Multi-sensor cloud retrieval
630   simulator and remote sensing from model parameters – Part 1: Synthetic sensor radiance
631   formulation, Geosci. Model Dev., 6, 2049-2062, doi:10.5194/gmd-6-2049-2013.

632 Wind, G., A.M. da Silva, P.M. Norris, S. Platnick, S. Mattoo and R. C. Levy, 2016: Multi-
633   sensor cloud retrieval simulator and remote sensing from model parameters – Part 2:
634   Aerosols, Geosci. Model Dev., 9, 2377-2389, doi:10.5194/gmd-9-2377-2016.

635 Zhang, Z., and S. Platnick, 2011: An assessment of differences between cloud effective

636   particle radius for marine water clouds from three MODIS spectral bands. J. Geophys.
637   Res., 116, D20215, doi:10.1029/2011JD016216.

638
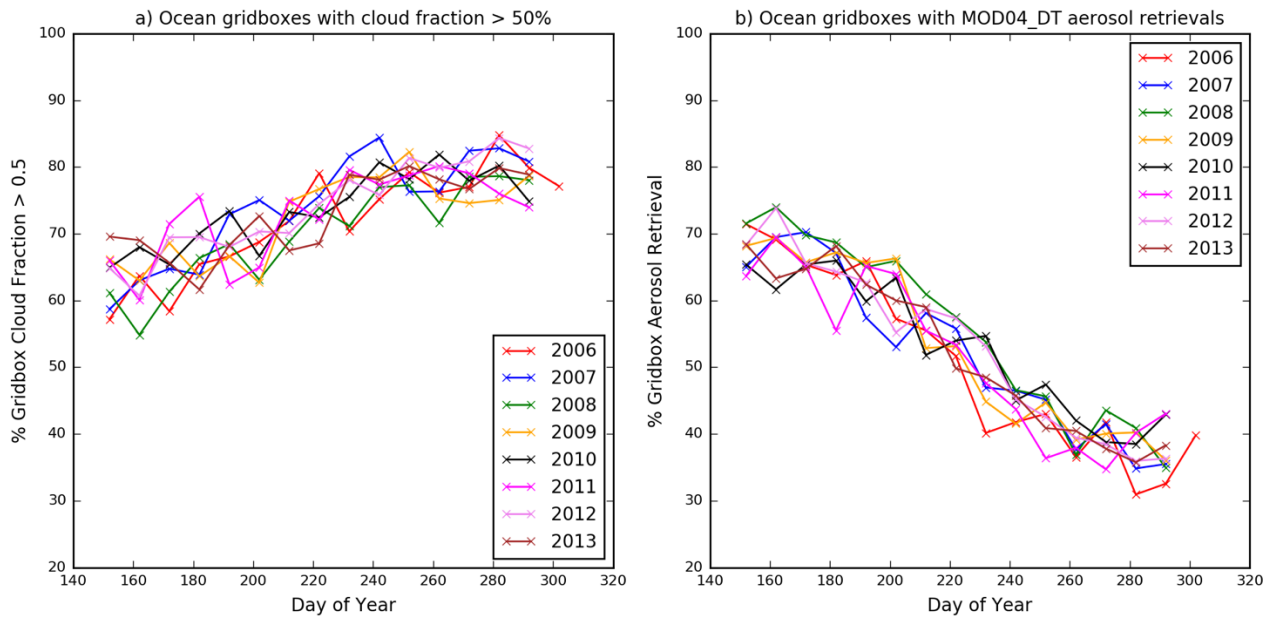
639
640 Figure 1. Terra MODIS Level-3 Daily 1-degree gridded product for SEAO area for years
641 2006-2013. Panel a) shows the percentage of SEAO ocean gridboxes that had cloud fraction
642 greater than 50%. Panel b) shows the percentage of SEAO ocean gridboxes that had any
643 successful MOD04DT aerosol property retrievals of any quality.
644

645
646
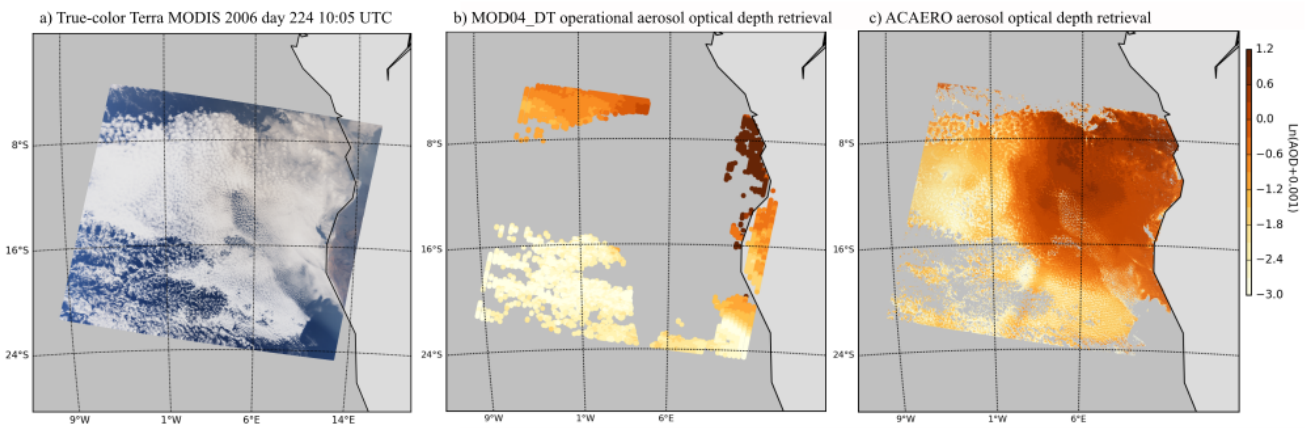
647
648
649   Figure 2. Real-data example of MOD06ACAERO retrieval. Terra MODIS 2006 day 224
650   10:05 UTC. Panel a) shows the true-color MODIS granule. There is extensive aerosol layer
651   above the equally extensive MBL cloud layer. Panel b) shows the MODIS Data Collection 6
652   operational Dark Target aerosol retrieval. It is a 10km resolution product with retrievals
653   available only in clear sky conditions and outside glint. Panel c) shows the MOD06ACAERO
654   above-cloud aerosol retrieval that is able to fill the data gap created by presence of MBL
655   layer.
656
657

a) MCARS study cases with aerosols included
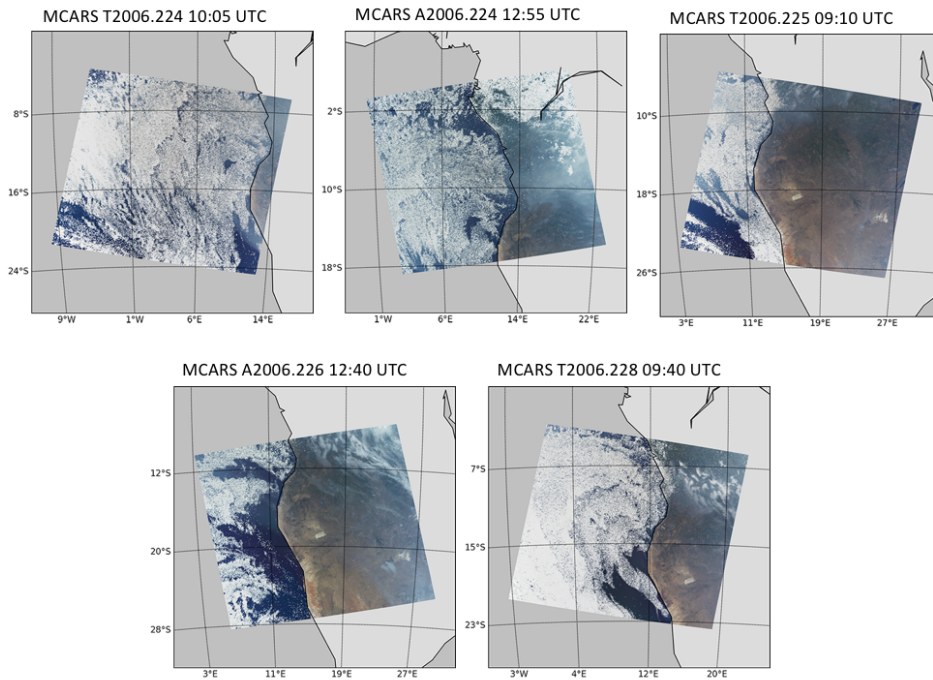
b) MCARS study cases with aerosols removed
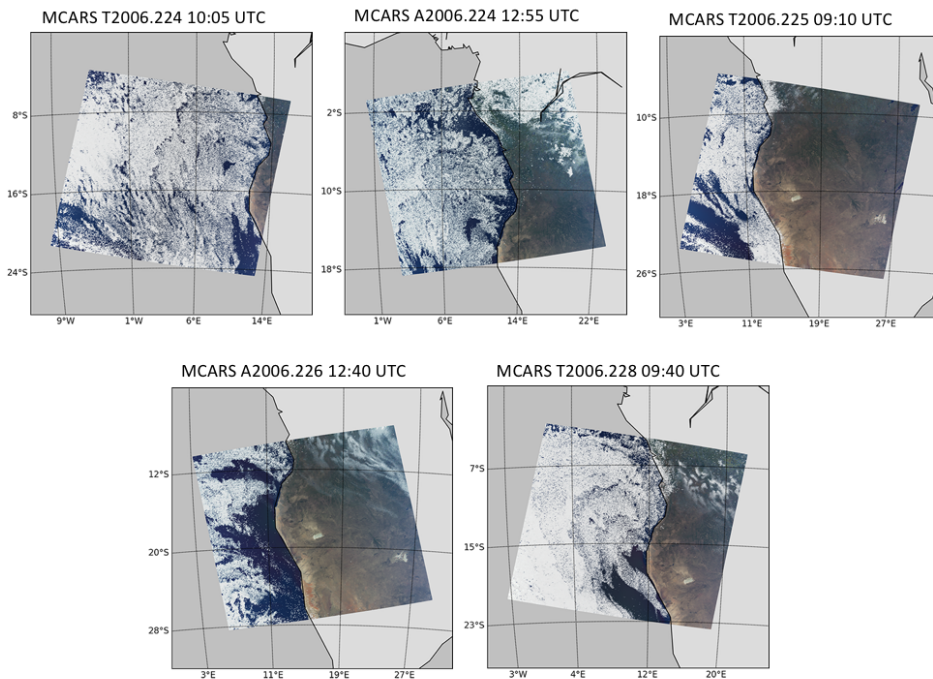
658
659 Figure 3. Scenes generated by MCARS from G5NR used in analysis of the MOD06ACAERO
660 product. There are three cases based on Terra MODIS, designated with a T next to the year.
661 There are two cases based on Aqua MODIS, designated with an A next to the year. Panel a)
662 shows the case set simulated with aerosols present. Panel b) shows the same case set but
663 simulated with aerosols removed.
664

a) Combined dataset, error < 40%   b) Combined dataset, error < 30%
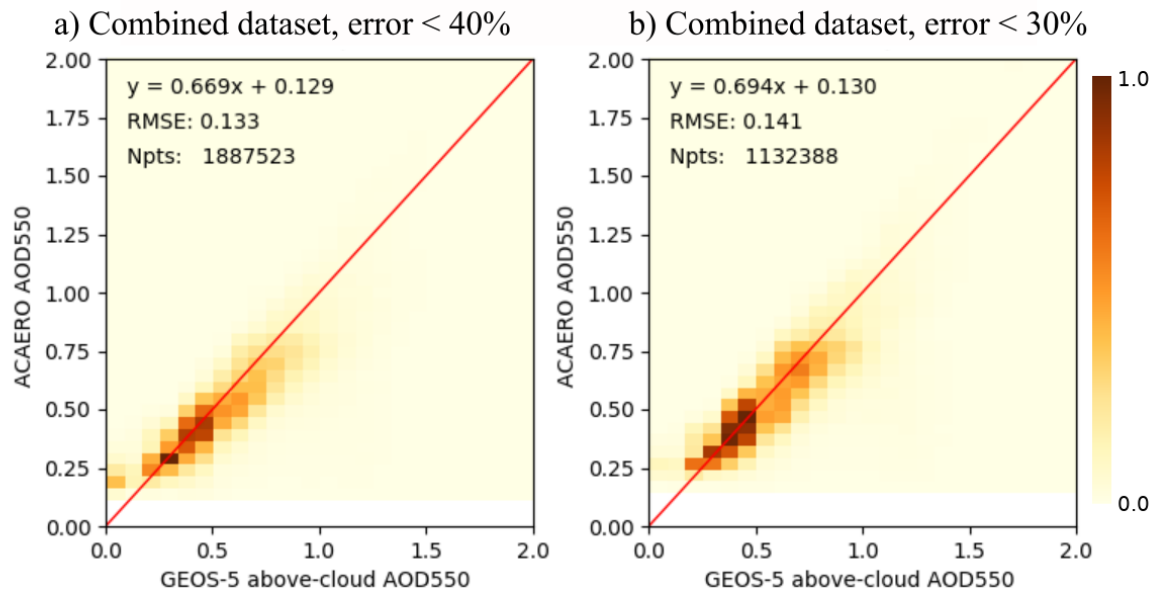
665
666 Figure 4. MOD06ACAERO retrieval results from the combined dataset of Figure 3a
667 compared to source GEOS-5 aerosol optical depth as a normalized density plot. No screening
668 of retrievals had been performed except for pixel-level uncertainty. Panel a) shows
669 MOD06ACAERO retrievals with uncertainty of less than 40% and panel b) shows same with
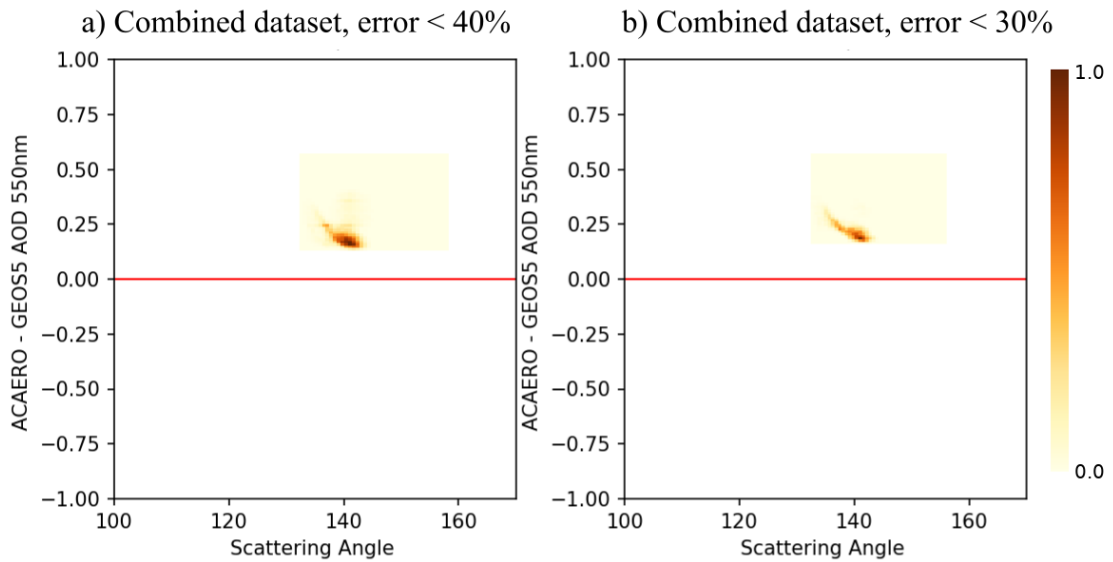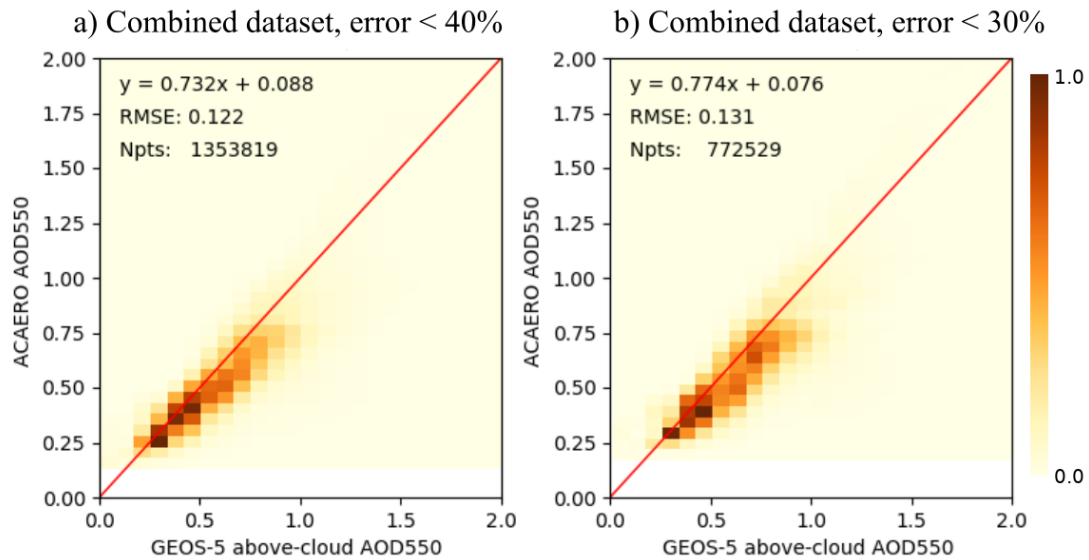670 uncertainty less than 30%.
671
672

a) Combined dataset, error < 40%   b) Combined dataset, error < 30%
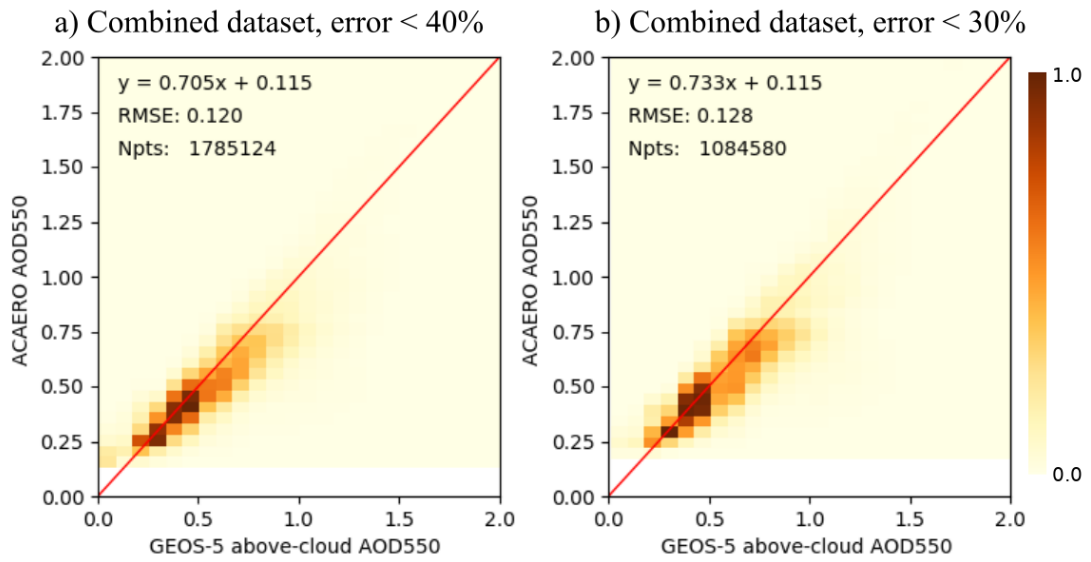
673
674 Figure 5. MOD06ACAERO retrieval results from the combined dataset of Figure 3a
675 compared to source GEOS-5 aerosol optical depth as a normalized density plot. AOD
676 retrievals where COT was less than 4 are now discarded. Panel a) shows MOD06ACAERO
677 retrievals with uncertainty of less than 40% and panel b) shows same with uncertainty less
678 than 30%.
679

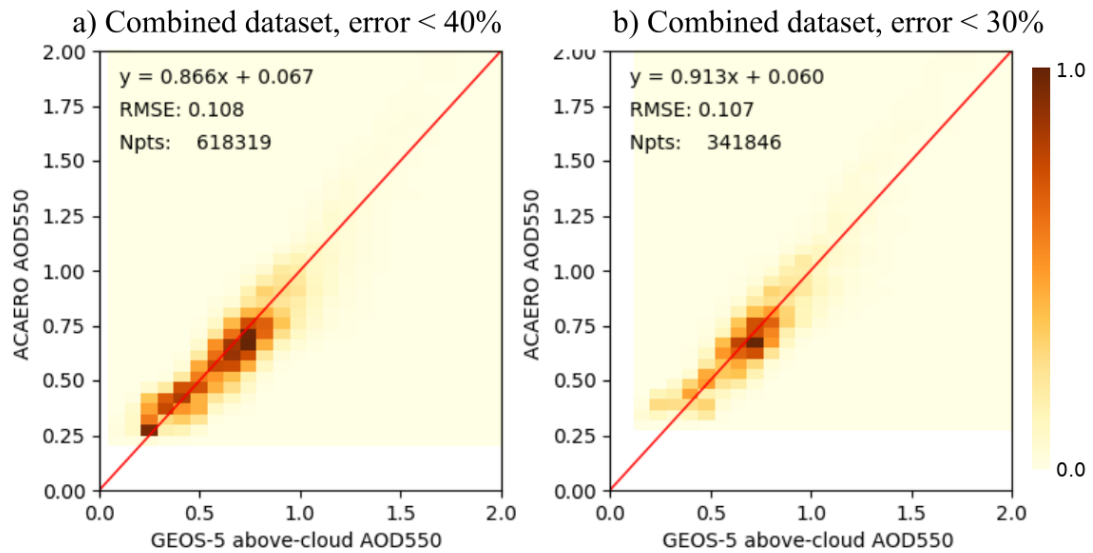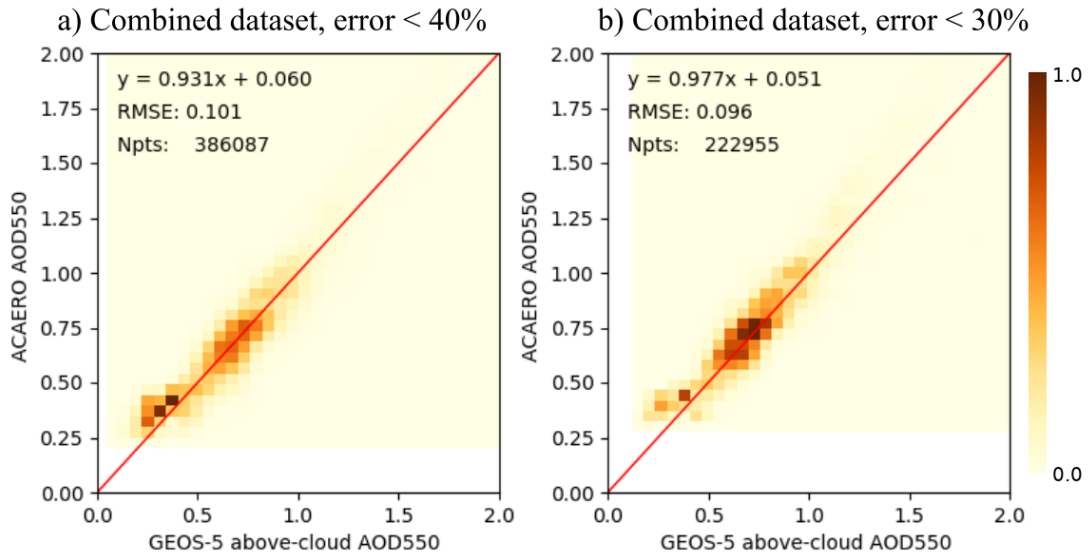a) Combined dataset, error < 40%   b) Combined dataset, error < 30%

680
681 Figure 6. MOD06ACAERO retrieval results from the combined dataset of Figure 3b, where
682 aerosols had been removed. The results are displayed as difference from GEOS-5 AOD,
683 which in this case was zero, as a function of scattering angle as a normalized density plot. All
684 retrievals where MOD06ACAERO result was also zero had been removed for clarity. All
685 non-zero MOD06ACAERO retrievals appear to be concentrated in a narrow angle range
686 between 135 and 145 degrees which corresponds to the cloud bow. Panel a) shows
687 MOD06ACAERO retrievals with uncertainty of less than 40% and panel b) shows same with
688 uncertainty less than 30%.
689
690

a) Combined dataset, error < 40%

b) Combined dataset, error < 30%

Figure 7. MOD06ACAERO retrieval results from the combined dataset of Figure 3a compared to source GEOS-5 aerosol optical depth as a normalized density plot. AOD retrievals where COT was less than 4 are now discarded. Additionally retrievals in the cloud bow region are also removed. It appears they were indeed the source of a cluster of higher MOD06ACAERO retrievals when GEOS-5 AOD was near zero and the match up with GEOS-5 source AOD is further improved. Panel a) shows MOD06ACAERO retrievals with uncertainty of less than 40% and panel b) shows same with uncertainty less than 30%.

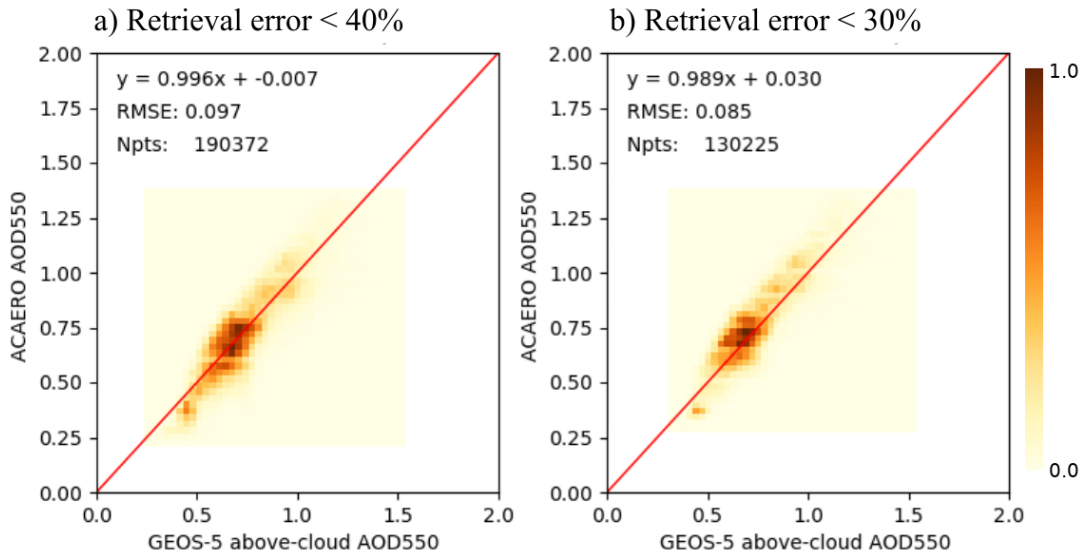a) Combined dataset, error < 40%   b) Combined dataset, error < 30%

701
702   Figure 8. MOD06ACAERO retrieval results from the combined dataset of Figure 3a
703   compared to source GEOS-5 aerosol optical depth as a normalized density plot. AOD
704   retrievals where COT was less than 4 and where the scattering angle was in the cloud bow are
705   now discarded. Additionally the data extent had been limited to only include pixels with view
706   zenith angle of less than 32 degrees. Retrieval comparison shows further improvement. Panel
707   a) shows MOD06ACAERO retrievals with uncertainty of less than 40% and panel b) shows
708   same with uncertainty less than 30%.

a) Combined dataset, error < 40%   b) Combined dataset, error < 30%

709
710 Figure 9. MOD06ACAERO retrieval results from the combined dataset of Figure 3a
711 compared to source GEOS-5 aerosol optical depth as a normalized density plot. AOD
712 retrievals where COT was less than 4 and where the scattering angle was in the cloud bow are
713 now discarded. Additionally the data extent had been limited to only include pixels with view
714 zenith angle of less than 20 degrees. Retrieval comparison shows further improvement
715 however it is not clear if the reduction in dataset size is worth the gain in accuracy. Panel a)
716 shows MOD06ACAERO retrievals with uncertainty of less than 40% and panel b) shows
717 same with uncertainty less than 30%.
718
719

MCARS Terra MODIS 2006 day 224, 10:05 UTC

a) Retrieval error < 40%

y = 0.996x + -0.007
RMSE: 0.097
Npts: 190372

b) Retrieval error < 30%

y = 0.989x + 0.030
RMSE: 0.085
Npts: 130225

Figure 10. MOD06ACAERO retrieval results from simulated MCARS granule based on Terra MODIS 2006 day 224 10:05 UTC compared to source GEOS-5 aerosol optical depth as a normalized density plot. In this experiment both MOD06 and MOD06ACAERO were modified to use MCARS pixel-level atmospheric profiles to perform atmospheric correction. AOD retrievals where COT was less than 4 and where the scattering angle was in the cloud bow are now discarded. Additionally the data extent had been limited to only include pixels with view zenith angle of less than 20 degrees. This experiment shows excellent agreement with source data. Panel a) shows MOD06ACAERO retrievals with uncertainty of less than 40% and panel b) shows same with uncertainty less than 30%. The small dataset size in panel b) is the reason for slightly lower agreement with source compared to panel a)