

1 Evaluation of NASA Deep Blue/SOAR aerosol
2 retrieval algorithms applied to AVHRR
3 measurements

A. M. Sayer^{1,2}, N. C. Hsu², J. Lee^{2,3}, N. Carletta^{2,4}, S.-H. Chen^{2,4}, and A.
Smirnov^{2,4}

AUTHOR'S ACCEPTED DRAFT OF PAPER
PRIOR TO COPY-EDITING AND TYPESETTING
PLEASE DO NOT DISTRIBUTE

REFER TO FINAL VERSION IN JGR WHEN AVAILABLE

N. Carletta, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.

S.-H. Chen, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.

N. C. Hsu, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.

J. Lee, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.

A. M. Sayer, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. (an-
drew.sayer@nasa.gov)

A. Smirnov, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.

¹Goddard Earth Sciences Technology and

4 **Abstract.** The Deep Blue (DB) and Satellite Ocean Aerosol Retrieval
5 (SOAR) algorithms have previously been applied to observations from sen-
6 sors like the Moderate Resolution Imaging Spectroradiometers (MODIS) and
7 Sea-viewing Wide Field-of-view Sensor (SeaWiFS) to provide records of mid-
8 visible aerosol optical depth (AOD) and related quantities over land and ocean
9 surfaces respectively. Recently, DB and SOAR have also been applied to Ad-
10 vanced Very High Resolution Radiometer (AVHRR) observations from sev-
11 eral platforms (NOAA11, NOAA14, and NOAA18), to demonstrate the po-
12 tential for extending the DB and SOAR AOD records. This study provides
13 an evaluation of the initial version (V001) of the resulting AVHRR-based AOD
14 data set, including validation against Aerosol Robotic Network (AERONET)
15 and ship-borne observations, and comparison against both other AVHRR AOD

Research (GESTAR), Universities Space
Research Association.

²NASA Goddard Space Flight Center,
Greenbelt, MD, USA.

³Earth Systems Science Interdisciplinary
Center (ESSIC), University of Maryland,
College Park, MD, USA.

⁴Science Systems and Applications, Inc.,
Lanham, MD, USA.

16 records and MODIS/SeaWiFS products at select long-term AERONET sites.
17 Although it is difficult to distil error characteristics into a simple expression,
18 the results suggest that one standard deviation confidence intervals on re-
19 trieved AOD of $\pm(0.03+15\%)$ over water and $\pm(0.05+25\%)$ over land rep-
20 resent the typical level of uncertainty, with a tendency towards negative bi-
21 ases in high-AOD conditions, caused by a combination of algorithmic assump-
22 tions and sensor calibration issues. Most of the available validation data are
23 for NOAA18 AVHRR, although performance appears to be similar for the
24 NOAA11 and NOAA14 sensors as well.

1. Introduction

25 Remote sensing of aerosol optical depth (AOD) from space has been performed using a
26 wide variety of sensor types. Passive polar-orbiting single-view imaging radiometers such
27 as the Advanced Very High Resolution Radiometer (AVHRR), Sea-viewing Wide Field-
28 of-view Sensor (SeaWiFS), Moderate Resolution Imaging Spectroradiometer (MODIS),
29 Medium Resolution Imaging Spectroradiometer (MERIS), and Visible Infrared Imaging
30 Radiometer Suite (VIIRS) include several important features suited for this task. Specif-
31 ically, they typically have moderate spatial pixel sizes (sub-km to several km), broad
32 swaths (providing views of a given location on the Earth approximately daily), and make
33 measurements in bands at solar (and often thermal) wavelengths sensitive to the atmo-
34 spheric aerosol loading. Thus they have been widely used via various techniques for such
35 applications over both land (e.g. *Hsu et al.*, 2004, *Levy et al.*, 2007, *von Hoyningen-Huene*
36 *et al.*, 2011, *Lyapustin et al.*, 2011) and water (e.g. *Stowe et al.*, 1997, *Tanré et al.*, 1997,
37 *Mishchenko et al.*, 1999, *Ahmad et al.*, 2010, *Sayer et al.*, 2012a, 2017) surfaces.

38 Similarities in observation characteristics between sensors of this type facilitate the ap-
39 plication of similar retrieval techniques, moving towards the goal of a long-term consistent
40 AOD record. AVHRR is particularly advantageous for long-term monitoring as the first
41 was launched in 1978 and AVHRRs are still flying at the present time. Even with a com-
42 mon algorithm, however, this goal is beset by numerous challenges as no two sensors are
43 identical; issues such as precise measurement capabilities, cloud screening, and calibra-
44 tion, among others, can influence sensor-to-sensor data consistency (e.g. *Jeong and Li*,
45 2005, *Li et al.*, 2009, *Kahn et al.*, 2011, *Mishchenko et al.*, 2012). Other instrument types

46 offer important capabilities of their own relevant to aerosol retrieval (e.g. multi-angle,
47 polarimetry, UV wavelengths, hyperspectral, lidar). These features provide additional
48 or alternative information content (e.g. *Hasekamp and Landgraf, 2007*), although such
49 sensors often lack comparatively in some combination of pixel size, swath width, or data
50 record length.

51 The long time series of the AVHRRs motivated recent efforts to apply versions of the
52 over-land Deep Blue (DB, *Hsu et al., 2004, 2006, 2013, Sayer et al., 2012b*) and over-water
53 Satellite Ocean Aerosol Retrieval (SOAR, *Sayer et al., 2012a, 2017*) algorithms, which
54 have previously been applied to AOD retrieval from SeaWiFS, MODIS, and VIIRS, to
55 the AVHRRs. An initial version (V001) of an AVHRR Deep Blue data set, combining
56 DB and SOAR retrievals, has been created for a subset of the AVHRR sensor records
57 (those flying on NOAA11 from 1989-1990; NOAA14 from 1995-1999; NOAA18 from 2006-
58 2011). Although the individual instruments were operational for longer, the satellites were
59 launched with nominal daytime Equatorial local solar crossing times around 1:30 pm and
60 drifted later while in orbit, which has consequences for sampling and time series analysis.
61 Thus the time periods processed to date were chosen to cover the periods where these
62 sensors had Equatorial crossing times between 1:30 pm and 3 pm, most comparable with
63 other early-afternoon platforms (e.g. the A-Train).

64 The new data set is freely available to download, along with a user guide, from
65 <https://portal.nccs.nasa.gov/datashare/AVHRRDeepBlue>. Additional information and
66 documentation is provided at <https://deepblue.gsfc.nasa.gov>. The specific implementa-
67 tion of these algorithms to the AVHRRs is described in a companion paper, *Hsu et al.*

68 [2017]. Note the data set is referred to as ‘AVHRR Deep Blue’ although it is composed
69 of both the distinct DB and SOAR algorithms.

70 The goal of this study is to evaluate these new data products, thereby providing guidance
71 for data users and suggest directions for refinement for a future processing of the whole
72 multi-satellite AVHRR data record. This is accomplished through several sets of compar-
73 isons. First, ground-truth reference data from the Aerosol Robotic Network (AERONET,
74 *Holben et al.*, 1998), Maritime Aerosol Network (MAN, *Smirnov et al.*, 2009), and earlier
75 ship-based AOD observations provide a validation. Second, comparing to existing related
76 satellite-based AOD records provides broader-scale context. Section 2 describes the data
77 products used, and the general validation methodology. The following Sections 3 and 4
78 provide a validation of the SOAR over-water and DB over-land AOD retrievals respec-
79 tively, while Section 5 is a comparison of the new AVHRR Deep Blue data set against
80 other satellite products. Section 6 provides a brief summary.

2. Data set descriptions

2.1. AVHRR Deep Blue/SOAR AOD retrievals

81 The adaptation of the DB and SOAR algorithms to the AVHRRs is described by *Hsu*
82 *et al.* [2017]. The physical principles behind the AVHRR application of the algorithms
83 are the same as those behind the SeaWiFS, MODIS, and VIIRS versions. However, as
84 only two solar bands are available for most of the AVHRRs and there is no on-board
85 solar band calibration, various algorithmic constraints and cloud screening tests must be
86 tightened to avoid unstable or unphysical results. Brief descriptions of some key features
87 of the AVHRR implementation follow.

88 The primary data products are the AOD at wavelengths of 550 nm (due to its com-
89 mon use as a reference wavelength in the scientific community), AOD at AVHRR band
90 1, and (over water only) AOD at band 2. In general, mentions of AOD without a specific
91 wavelength should be taken to refer to 550 nm. The exact central wavelengths of these
92 bands vary slightly between the different AVHRRs, and are referred to herein at 630 nm
93 and 830 nm respectively in the general discussion for simplicity. All calculations, however,
94 use exact sensor-specific wavelengths. Specifically, central wavelengths are 636, 636, and
95 633 nm for band 1, and 810, 820, and 848 nm for band 2, for NOAA11, 14, and 18 respec-
96 tively. Multiple AVHRR solar band calibrations have been derived; this initial version of
97 the data set uses that of *Vermote and Kaufman* [1995], which is also used for NASA's
98 long-term normalized difference vegetation index (NDVI) data sets, although the use of
99 other calibrations will be investigated for future versions.

100 Over land, DB has two methods of estimating surface reflectance for a given pixel,
101 depending on whether the location has a bright (e.g. barren ground, urban areas) or
102 vegetated surface. For bright surfaces, a global seasonally-varying data base of surface
103 reflectance is constructed using a similar method to the minimum reflectance technique,
104 applied to the whole sensor record. For the other applications of Deep Blue (cf. *Hsu et al.*,
105 2013) the primary wavelength is 412 nm, at which the surface reflectance is fairly dark,
106 even for deserts. AVHRR lacks this channel so band 1 near 630 nm is used instead. As the
107 surface is typically somewhat brighter at 630 nm than 412 nm, however, the aerosol signal
108 is somewhat reduced, and the resulting AOD uncertainty is larger. Over the brightest
109 surfaces (e.g. snow, salt pans, some deserts) the surface is too bright and no retrieval
110 is performed to due a lack of sensitivity to AOD variations. Over vegetated surfaces,

111 reflectance is estimated dynamically, as it often varies more rapidly in time than over
112 arid surfaces. As AVHRR lacks shortwave infrared (SWIR) bands which are useful to
113 track these changes, the surface reflectance is modelled as an empirical function of NDVI.
114 A similar approach was previously developed for SeaWiFS DB (*Hsu et al.*, 2013), as
115 SeaWiFS also lacks SWIR channels, and was found to perform well (*Sayer et al.*, 2012b).
116 Full details of both approaches are provided by *Hsu et al.* [2017]. Note that separate
117 surface data bases and NDVI relationships are constructed for each sensor, as they each
118 have slightly different spectral response functions.

119 For both land surface types, the aerosol optical model is assumed on a regional and
120 seasonal basis, due to the aforementioned limited information content of AVHRR. These
121 models are drawn from the same sets of models used for other DB applications, adapted
122 to AVHRR wavelengths. Once the surface reflectance has been obtained, band 1 AOD is
123 retrieved directly from the AVHRR measurement. The AOD at 550 nm is extrapolated
124 from this using an assumed (regionally and seasonally-dependent) Ångström exponent
125 (AE) based on AERONET climatologies (*Hsu et al.*, 2017). Thus, both 550 nm and band
126 1 AOD are provided within the data set, even though AVHRR has no band near 550 nm.

127 Over water, SOAR uses both bands 1 and 2 in a simultaneous inversion to determine
128 AOD and the best-fitting aerosol optical model from a choice of dust, fine-mode dom-
129 inated, and maritime optical models. Surface reflectance includes contributions from
130 wind-speed-dependent foam and Sun glint, as well as ‘underlight’ from suspended pig-
131 ments, although this latter term is small for AVHRR bands. This is essentially the same
132 basic approach as in the SeaWiFS retrieval (*Sayer et al.*, 2012a), although the AVHRR
133 algorithm makes use of improvements to the surface reflectance model and aerosol optical

134 models (e.g. nonspherical dust) which were developed during the VIIRS implementation
135 of SOAR (*Sayer et al.*, 2017). Unlike these other applications, for AVHRR the fine mode
136 fractional contribution to AOD is fixed (one different value for each aerosol type), rather
137 than retrieved directly, again due to the limited spectral information provided by AVHRR.
138 The AOD at 550 nm is then obtained in a self-consistent approach using the retrieved
139 aerosol loading and best-fitting aerosol optical model. This model and its associated AE
140 are also reported in the data set. As over land, separate lookup tables are created for
141 each AVHRR sensor.

142 Each retrieval also has an associated quality assurance (QA) flag between 1 and 3.
143 QA=1 ('poor') indicates internal tests (*Hsu et al.*, 2017) suggest some potential problem,
144 such as cloud-contamination or an improper surface model, so the retrieval is likely to
145 be quantitatively less reliable. These retrievals should not be used for most applications.
146 QA=3 ('good') pass all checks, and are therefore least likely to suffer from these issues.
147 QA=2 ('moderate') retrievals are an intermediate category. Most retrievals are assigned
148 either QA=1 or QA=3. In this analysis, only retrievals with QA=2 or 3 are used, which
149 is the general recommendation for almost all data users.

150 The resulting data Level 2 (L2) products are provided at approximately $8.8 \times 8.8 \text{ km}^2$
151 horizontal pixel size at the sub-satellite point (2×2 Global Area Coverage AVHRR pixels)
152 for daytime (solar zenith angle $< 84^\circ$) land and ocean pixels free from cloud, snow/ice, or
153 Sun glint. Level 3 (L3) daily/monthly composites are also available, created from $\text{QA} \geq 2$
154 retrievals gridded to 1° resolution. Consecutive orbits from AVHRR overlap, particularly
155 at high latitudes, and so some L3 daily grid cells contain contributions from multiple
156 orbits, spaced approximately 90 minutes apart. Note that the AVHRR daily L3 data

157 product requires at least 5 retrievals for a grid cell to be valid, and the monthly mean at
158 least 3 days with sufficient data within a month.

159 As with other Deep Blue data products, and indeed many other satellite data sets, the
160 uncertainty on retrieved AOD is a function of the true AOD. This is somewhat unavoidable
161 given the nature of the measurements and required retrieval assumptions. An expected
162 error (EE) envelope is defined, intended to represent a one-standard-deviation confidence
163 envelope around the retrieved AOD (e.g. *Sayer et al.*, 2013), such that one standard
164 deviation of retrievals (i.e. about 68 %) match the ground truth AOD to within this level
165 (and, following Gaussian statistics, approximately 95 % within twice the EE envelope,
166 etc.). For the initial AVHRR Deep Blue data set, the EE is taken (*Hsu et al.*, 2017)
167 as $\pm(0.03+15\%)$ over water, and $\pm(0.05+25\%)$ over land (with AOD defined relative
168 to the Sun photometer values, i.e. a diagnostic rather than prognostic measure), for all
169 wavelengths considered. These may be refined further in the future.

170 The validation analysis includes discussions of the fraction f of points where the
171 AVHRR-AERONET difference is smaller than the EE. By the definition of the EE, the
172 target in the ideal case for a useful uncertainty metric is thus $f \approx 0.68$, with substan-
173 tially lower values indicating retrieval errors are on average larger than this envelope, and
174 substantially higher values indicating that performance under this circumstance appears
175 better than anticipated.

2.2. Other satellite AOD products used

176 2.2.1. AVHRR

177 Two other over-ocean AVHRR AOD retrieval algorithms are also examined in this
178 work. The first is from the Global Aerosol Climatology Project (GACP), described most

179 recently by *Geogdzhayez et al.* [2015], which provides monthly AOD at 550 nm and AE
180 over ocean on a 1° grid. The second is the NOAA aerosol climate data record (CDR)
181 version 3 (*Zhao, 2016, Zhao et al., 2016*), which provides AOD at band 1 (but not 550
182 nm); daily and monthly data (level 3) are provided on a 0.1° grid. Various approaches
183 to retrieve AOD over land from the AVHRRs have been proposed and demonstrated on
184 local or regional scales (e.g. *Knapp and Stowe, 2002, Riffler et al., 2010, Mei et al., 2014,*
185 *Gao et al., 2016*). However these have not been applied to the full AVHRR records to
186 produce global publicly-available data products. The one exception is the Polar Multi-
187 sensor Aerosol product (PMAp, *EUMETSAT, 2016*), although this has only been applied
188 to the AVHRR sensors on the MetOp platforms in forward-processing, and does not
189 overlap with the current DB/SOAR AVHRR record. Therefore the comparison with
190 other AVHRR products is restricted to the aforementioned over-ocean data records only.

191 2.2.2. SeaWiFS

192 The SeaWiFS mission covered the time period September 1997-December 2010, with a
193 small number of temporary outages, and the SeaWiFS Deep Blue data set includes DB
194 retrievals over land (*Sayer et al., 2012b, Hsu et al., 2013*) and the initial application of
195 SOAR over water (*Sayer et al., 2012a*). The current version 4 is used; specifically, the
196 monthly mean 550 nm AOD product at 1° spatial resolution. As many 1° grid cells contain
197 both land and water pixels, the product provides both a ‘combined’ land/ocean data set
198 as well as results from the DB/SOAR algorithms individually. The latter are used here.
199 A level 3 monthly grid cell from this product is only filled if it contains data from at least
200 3 different days within a given month, and a grid cell is valid on a particular day if it
201 contains at least 3 retrievals passing QA checks.

202 **2.2.3. MODIS**

203 This study makes use of MODIS AOD from the Aqua platform (launched in May 2002
204 and still operational), as it has a similar local solar crossing time (1:30 pm at the Equator
205 for the daytime nodes) to the nominal orbit times of the platforms hosting the specific
206 AVHRR sensors considered in this study. Over land, data from the DB algorithm are
207 used (*Hsu et al.*, 2013, *Sayer et al.*, 2013). Over water, since there has not yet been an
208 application of SOAR to MODIS, the standard MODIS ocean retrieval product (*Tanré*
209 *et al.*, 1997, *Levy et al.*, 2013) is used as a point of reference. This shares similar physical
210 principles to SOAR (multispectral inversion based on matching observed reflectances to
211 results from radiative transfer models) but numerous algorithmic specifics are different.
212 Both DB and the MODIS ocean algorithm provide 550 nm AOD, used herein.

213 The results in this work are taken from the current Collection 6 level 3 monthly product
214 (identifier MYD08_M3). Note that this product as standard does not have any thresholds
215 applied to determine whether a grid cell is sufficiently well-sampled or not to be meaningful
216 (i.e. one retrieval passing QA checks in a whole month results in a populated monthly
217 mean AOD). In practice most populated grid cells contain several hundred retrievals from
218 multiple days, but a small fraction contain only a dozen or so. There is no way within the
219 MYD08_M3 data product to identify how many individual days contributed to a specific
220 cell within a given month. As a result, to mitigate the influence of a small number of
221 sparsely-populated cells, an additional filtering step is applied herein to remove grid cells
222 with fewer than 30 retrievals within a month. The specific threshold chosen does not
223 strongly affect the results presented herein.

2.3. AERONET

224 Starting from a few sites in the early 1990s, AERONET has expanded to provide sev-
225 eral hundred sites with multi-year (in some cases decadal or longer) aerosol observations
226 (*Holben et al.*, 1998, 2001), as well as dedicated deployments during intensive field cam-
227 paigns (*Holben et al.*, 2017). The Cimel Sun photometers used by AERONET provide
228 observations of columnar spectral AOD and water vapor from direct-Sun observations
229 with a temporal frequency of approximately 3-15 minutes (dependent on site, and limited
230 to daytime cloud-free periods), as well as a range of products from the spectral deconvolu-
231 tion of the AOD (*O'Neill et al.*, 2003), and from inversions of almucantur scans (*Dubovik
232 and King*, 2000). The direct-Sun products have become a standard for satellite/model
233 AOD validation, due to the low level of uncertainty (~ 0.01 at midvisible and longer wave-
234 lengths; *Eck et al.*, 1999) and consistency in instrument calibration and data processing
235 between sites and in time.

236 This study uses the AERONET direct-Sun version 2 level 2 (cloud-screened and quality-
237 assured; *Smirnov et al.*, 2000a) data products. All instruments provide a standard set of
238 wavelengths (440, 675, 870, and 1020 nm for AOD), and some include additional wave-
239 lengths. In this analysis, AERONET AOD are interpolated spectrally to 550 nm as well
240 as band 1 and 2 central wavelengths for the individual AVHRR sensor in question. This
241 interpolation is performed using the closest available AERONET wavelength and the AE,
242 and adds negligible additional uncertainty.

243 AVHRR and AERONET data are compared by averaging satellite data within 25 km
244 of the AERONET site and AERONET data within ± 30 minutes of the satellite overpass,
245 which has been the standard approach (cf. previously-cited satellite AOD papers). This
246 approach is designed to mitigate the influence of spatiotemporal variability on the com-

247 parison, although it cannot eliminate sampling differences entirely (see discussion by *Hyer*
248 *et al.*, 2011 and *Kahn et al.*, 2011). When comparing DB land retrievals only AVHRR
249 land pixels are considered, and when comparing SOAR ocean retrievals only AVHRR wa-
250 ter pixels are considered; as noted previously, the AVHRR data are QA-filtered prior to
251 this averaging process. A matchup is valid if there is at least one AVHRR retrieval in the
252 spatial window and at least one AERONET observation in the temporal window.

253 Note that no AERONET matchups are available for the NOAA11 part of the analysis,
254 because no sites were active during the time period available in the V001 AVHRR Deep
255 Blue data set (1989-1990).

2.4. Ship-borne AOD measurements

256 The Maritime Aerosol Network (MAN; *Smirnov et al.*, 2009, 2011) is a complement
257 to AERONET, consisting of ship-based AOD measurements made from hand-held Mi-
258 crotops II Sun photometers. These can be used to determine spectral AOD with an
259 approximate uncertainty of 0.02 (*Knobelspiesse et al.*, 2004), i.e. slightly greater than
260 that of the stationary Cimel instruments used in AERONET but still sufficient for a val-
261 idation of satellite/model data sets. Measurements are made on cruises where equipment
262 and personnel are available, thus enabling Sun photometer-based AOD validation in open
263 ocean regions. With the exception of two pilot cruises in 2004 and 2005, the MAN data
264 base includes cruises from 2006 onwards (and are most frequent in more recent years).
265 Thus, of the satellites considered in this study, MAN data are only available for NOAA18.
266 Here, the ‘series average’ (data acquired with a gap of <2 minutes between observations)
267 level 2 MAN product is used, with the same matchup methodology as applied over land
268 (Section 2.3).

269 Ship-based measurements of this type were also made prior to the formal establishment
270 of MAN. *Smirnov et al.* [2002] provide a discussion of some. As they were collected
271 by a variety of researchers and not formalised into a consistent data base, the available
272 information (both in terms of AOD wavelengths and precision of spatial/temporal location
273 data) for these earlier cruises is more variable. These measurements are used to provide
274 validation for the NOAA11 and NOAA14 data. To increase the available NOAA11 data
275 record, data from the year 1991 were also included rather than just the 1989-1990 period
276 provided in the initial AVHRR Deep Blue data release. It is possible, however, that 1991
277 results will be subject to different error characteristics as the June 1991 eruption of Mt.
278 Pinatubo injected a sizeable amount of aerosol into the atmosphere, which spread to cover
279 much of the globe, persisted for several years, with both different optical properties and
280 vertical distribution from tropospheric aerosols found in periods free from strong eruptions
281 (*Lambert et al.*, 1992, *Mishchenko and Geogdzhayev*, 2007).

282 The same spectral interpolation technique is applied throughout. For NOAA14
283 matchups, data content and format were similar to those of MAN and so the same matchup
284 criteria are used. NOAA11 matchups typically provide latitude/longitude information to
285 the nearest degree, and data reported as ‘morning’ or ‘afternoon’ averages (note NOAA11,
286 14, and 18 had early-afternoon overpass times), which is somewhat less precise than the
287 MAN data. Where this is the case, all available AVHRR retrievals over a 200 km radius
288 from this point on a given day are averaged and compared with the reported ship-based
289 temporal average. This inevitably increases sampling-related uncertainty, which should
290 be borne in mind in the interpretation, although given the limitations of the available

291 data, it is the best that can be done and the only option to provide a validation for these
292 earlier satellite missions.

3. AOD validation over ocean

3.1. AERONET island/coastal sites

293 AERONET has expanded significantly through the past few decades. As a result,
294 matchups over water were obtained at 40 island/coastal sites during the NOAA18 period,
295 but only 20 provided data for NOAA14, and none for NOAA11. Summary statistics
296 for the comparisons for these two satellites are provided in Table 1. The focus of this
297 discussion is on NOAA18 data, since it has the greatest available comparison volume,
298 for brevity; conclusions concerning error characteristics over ocean, where not discussed
299 explicitly, are qualitatively and quantitatively similar for NOAA14.

300 The overall tendencies of AOD retrieval error over ocean for NOAA18 are shown in
301 Figure 1, which splits the data according to AERONET AOD and AE, thus providing a
302 simple categorisation into background (low-AOD), elevated-AOD coarse-mode dominated
303 (i.e. dust, typically), and elevated-AOD fine-mode dominated (i.e. smoke/continental)
304 conditions. The 550 nm and band 1 results show similar behavior; for these bands there
305 is a tendency for a slight positive AOD bias of order 0.02 in the cleanest scenes, gradually
306 decreasing as AOD increases, with a negative bias of approximately 10% in high-AOD
307 conditions. The neutral point of AOD bias around 0 is about 0.15-0.2. On the other
308 end, the ability to examine the statistics of extreme conditions is limited, since the 95th
309 percentile of AERONET AOD is only 0.48 for these matchups. Note that in a statistical
310 sense a tendency for positive offset is expected for the cleanest conditions, due to the
311 simple fact that AOD is positive definite, so in comparison to a ‘truth’ reference even

312 with a random distribution of errors the aggregate is likely to be biased positive (i.e.
313 negative AOD is unphysical so an underestimate of AOD in conditions close to zero AOD
314 is impossible). For band 2 (near 848 nm for NOAA18), the bias is more small and positive
315 throughout, and only becomes negative, and to a lesser extent, for dust-like conditions.

316 The median and central 68% of retrieval errors fall within or are similar to the EE
317 lines in Figure 1 in most conditions, suggesting that this metric provides a reasonable
318 approximation of retrieval uncertainty on aggregate. It would be desirable in future
319 versions to address biases so that binned statistics of this type fall closer to the zero line.
320 These bias tendencies are indicative of a bias in some combination of sensor calibration or
321 radiative transfer assumptions (most likely aerosol optical model or trace gas absorption,
322 which is not negligible, especially for AVHRR band 2; *Tanré et al.*, 1992).

323 Figure 2 is analogous to Figure 1, except for NOAA14; the general tendencies between
324 the two are similar, except that (particularly for band 2) the biases are more positive,
325 by around 0.015-0.02 dependent on wavelength (Table 1). Since the two apply the same
326 algorithm, it is likely that calibration differences are the major reason for the discrepancy
327 here. It should also be noted that the data volume is smaller for NOAA14 by about an
328 order of magnitude (in terms of number of matchups) and a factor of two (in terms of
329 sites), since AERONET was less widespread during this period.

330 Returning to NOAA18, Figure 3 shows site-by-site statistics at 550 nm for the over-
331 ocean comparison. Spatial patterns are similar for NOAA14 data, as well as for data at
332 other wavelengths (not shown). Correlation coefficients tend to be high (0.8-1) for sites
333 with a large dynamic range of AOD (largely continental outflow regions), and smaller for
334 low-AOD regions, where the range of AOD becomes more comparable to the retrieval EE.

335 Biases tend to be small (magnitude <0.015 at most sites), with the sign dependent on
336 whether it is a predominantly clean or high-AOD region, consistent with Figure 1. Note
337 a few areas with high positive AOD bias are sites in complex coastal areas, particularly
338 ICIPE Mbita (on the shores of Lake Victoria), Hong Kong, Taihu (a large lake near
339 Shanghai), and Darwin (northern Australia). In these areas it is possible that either the
340 turbid water mask is not working effectively, or some pixels identified as ocean are in fact
341 mixed land and ocean, thereby providing a brighter signal than would be expected for an
342 open-ocean scene. These sites are also the ones at which the fraction of points matching
343 AERONET within the EE are significantly lower than the target of 68%.

344 The data were also examined for possible biases with respect to changing near-surface
345 wind speeds or total column water vapor amount (omitted for brevity), although these
346 were small (less than 0.02 change in median bias across the range of the variables). Over-
347 all, this analysis suggests that caution should be taken in analysis of retrievals in complex
348 coastal environments, and particularly lakeshores, but otherwise typical AOD retrieval
349 biases are close to zero and uncertainty is of order $\pm(0.03+15\%)$. Over ocean, the root
350 mean square error (RMSE) at individual sites is typically in the range 0.05-0.075 (Figure
351 3). Since the biases are in most cases significantly smaller than this, it is unlikely that de-
352 creasing the bias, whether through improvements to calibration or baseline aerosol optical
353 models, will significantly decrease the RMSE over ocean or shrink the EE envelope on a
354 global basis. This is a fundamental consequence of AVHRR's limited spectral information
355 and band digitisation. Consequently, improving the correlation over low-AOD ocean sites
356 may be difficult. This suggests that the best path forward for improvement to the ocean
357 retrieval may be to focus on improvement to QA tests in turbid or coastal waters, as these

358 aforementioned sites are those with highest RMSE and lowest compliance with the EE
359 metric.

360 AOD bias characteristics over ocean are similar to the SeaWiFS application of SOAR
361 reported by *Sayer et al.* [2012a], i.e. a small positive bias in low-AOD conditions but
362 $\sim 10\%$ low bias in high-AOD conditions. The optical models in both cases are based on
363 AERONET version 2 inversions, and are common to both sensors (except for the case of
364 dust, where AVHRR adopts a nonspherical model which had not been developed at the
365 time the SeaWiFS data set was created). The similar bias characteristics may plausibly
366 indicate systematic biases in the aerosol optical models (e.g. insufficient absorption), al-
367 though other causes such as sensor calibration cannot be discounted at the present time.
368 A version 3 AERONET inversion product is expected to become available within the next
369 year or so, at which point the two versions will be compared to see if there is any system-
370 atic shift in retrieved size distribution or absorption. If so, updated optical models can
371 be derived and implemented in future SeaWiFS/AVHRR reprocessings.

372 Finally, although aerosol type should not be considered a primary retrieval data product
373 here, Figure 4 shows histograms of the AERONET AE split according to whether SOAR
374 identified each matchup as predominantly dust-dominated, fine-mode dominated, or clean
375 marine (in terms of retrieved best-fit aerosol optical model). As a reminder (Section 2.1),
376 SOAR sequentially performs the retrieval for each aerosol optical model and reports the
377 best-fitting. So, over ocean, it is instructive to see to what extent SOAR's judgement of
378 likely aerosol optical model compares so the AE (which is related to aerosol fine/coarse-
379 mode optical dominance) derived from AERONET. Over land, the AVHRR application

380 of DB uses a fixed aerosol optical model dependent on location and season and so such a
381 comparison is not possible.

382 For both the ‘all points’ and ‘AERONET AOD ≥ 0.2 ’ cases, the general picture is
383 reasonable in that the most common AE when SOAR picks the dust model are low (0-
384 0.5), while the most common AE when SOAR picks the fine-dominated model are higher
385 (1.3-1.8). The distributions do however have fairly long tails, indicating cases where the
386 inferred likely aerosol type from AVHRR is probably incorrect. Therefore, while they
387 may often be reasonable, the best-fitting optical model should not be taken alone as a
388 definitive indicator of likely type or origin of the observed aerosols in the column.

389 The ‘marine’ histograms are both broader, reflecting both the more potentially more
390 mixed nature of clean scenes, but also the fact that AERONET AE is somewhat uncertain
391 in low-AOD conditions (e.g. *Wagner and Silva, 2008*). Note also that the marine AE his-
392 tograms skew to more positive values than expected for typical remote ocean conditions,
393 as reported by *Smirnov et al. [2011]* based on extensive ship-borne observations, which is
394 probably related to the fact that the available AERONET sites are, by their nature, situ-
395 ated in island/coastal areas (i.e. additional continental influence) which may be expected
396 to have a different fine/coarse aerosol partition from the open ocean. This points to the
397 need for validation in both coastal and remote regions.

3.2. Ship-based observations

398 Table 2 presents statistics of the comparison between AVHRR and ship-based AOD
399 measurements. The results are in general agreement with those obtained in Section 3.1
400 for coastal/island AERONET sites. Figure 5 shows the locations of each matchup for
401 each sensor, colored to show the aerosol optical model chosen by SOAR in each case.

402 Although the data volume is small and this is a series of instantaneous snapshots rather
403 than a climatology, it does match intuitive expectations (i.e. open-ocean conditions tend
404 to be have chosen the optical model for clean marine aerosols, and dust/fine-dominated
405 aerosols are chosen largely downwind of expected source locations typical for these aerosol
406 types). This is broadly in agreement with the histograms shown previously in Figure 4.

407 The 80 matchups with NOAA11 come from two distinct sources. The first is mea-
408 surements made by Y. Villevalde in the Pacific and North Atlantic oceans, reported in
409 *Villevalde et al.* [1994] and *Smirnov et al.* [1995a]. These both predominantly sampled
410 low-AOD conditions represented of the clean marine atmosphere; for the cruises as a
411 whole, *Villevalde et al.*, 1994 report mean 551 nm AOD of 0.13 and 0.11, and AE of 0.56
412 and 0.99, for the Pacific and Atlantic legs respectively. The NOAA11 data are in good
413 agreement with these cases, and indeed SOAR chose the ‘clean marine’ optical model
414 (*Sayer et al.*, 2012c) in almost all these cases. The second set of measurements were led
415 by O. Yershov and took place on several cruises in the North Atlantic, Mediterranean,
416 and Black Sea, and are described by *Smirnov et al.* [1995b]. These sampled both open-
417 ocean and continentally-influenced air masses. One outlying case from a Mediterranean
418 leg of these cruises is responsible for the lower correlation and higher RMSE of these
419 data compared to the NOAA14/NOAA18 observations in Table 2. Manual examination
420 of this case reveals a dust plume near the reported ship location; since the geolocation
421 information of these early ship-borne data were less precise than in later records (Section
422 2.4), this is likely attributable to sampling differences rather than true retrieval error.
423 AVHRR retrieved in the dust plume with AOD around 0.65, but the ship, potentially
424 up to 50 km in space and several hours distant in time, may have sampled outside the

425 plume, reporting AOD around 0.15. Overall, however, the matchups with NOAA11 are
426 consistent with comparable performance to the later AVHRR sensors, and as noted, the
427 available validation data for this time period are very limited.

428 All of the 20 NOAA14 matchups come from measurements between the US East Coast
429 and Bermuda during the summer 1996 TARFOX campaign, described in *Smirnov et al.*
430 [2000b]. This cruise sampled a mixture of clean marine and continentally-influenced air
431 masses; the matchups with NOAA14 were all low to moderate AOD (0.1-0.35). All are
432 in excellent agreement with the ship-based data (correlation 0.98 or higher, and RMSE
433 0.03 or lower, depending on wavelength, and 100% matching within the EE). While
434 encouraging, it is important to emphasise that this is a small number of measurements
435 from a small region and a limited time period, so should not be taken to imply that the
436 performance of the NOAA14 data set is superior to the others.

437 The NOAA18 matchups are from a broader set of cruises (see *Smirnov et al.*, 2009, 2011)
438 and cover many different regions (Figure 5). Comparison statistics are broadly similar to
439 those for the earlier AVHRR sensors (Table 2) and the AERONET island/coastal sites
440 (Table 1). In particular, the AOD bias tends to become more positive (or less negative),
441 and RMSE to decrease, as wavelength increases. The increased uncertainty at 550 nm is
442 expected since this AOD represents a slight extrapolation beyond the wavelength range of
443 AVHRR measurements, so is not as well-constrained (i.e. it is quite sensitive to the AE,
444 which is assumed rather than retrieved). Nevertheless, all data in Table 2 have $f > 0.68$,
445 suggesting the EE may be smaller than the assumed $\pm(0.03+15\%)$ over ocean, consistent
446 with results in Table 1 for island/coastal AERONET locations.

4. AOD validation over land

447 A total of 427 AERONET sites, shown in Figure 6, provided matchups with NOAA18
448 over land. Due to the larger variety of aerosol sources and sinks over land compared
449 to ocean, as well as the increased heterogeneity of terrain, unevenness of distribution of
450 AERONET sites, and regional rather than global nature of many analyses, regional as
451 well as global statistics are provided in Table 3. The boundaries of these regions are also
452 shown in Figure 6. Their definition is a balance between trying to keep areas with similar
453 aerosol/surface conditions together, and regions frequently used in analyses, balanced by
454 the distribution of the AERONET sites. As such it is inherently somewhat subjective
455 but provides a balance between level of detail, conciseness, and data volume. Figure 7
456 shows (for the 304 sites providing at least 25 matchups) site-by-site correlation, bias, and
457 fraction matching within the over-land EE of $\pm(0.05+25\%)$, and Figure 8 an examination
458 of retrieval error characteristics as a function of AOD and AE.

459 From Table 3, globally, 69% of matchups agree with AERONET within the EE at
460 550 nm, and 74% for band 1 (633 nm for NOAA18). Globally and regionally, the RMSE
461 tends to be 10-20% larger at 550 nm compared to band 1, and the AOD bias is less positive
462 (or more negative) at 550 nm than band 1. The AOD bias at both wavelengths (Figure
463 8) also tends to be small and positive in low-AOD conditions, but more negative (relative
464 bias around -20%) at high AODs, meaning it is on the lower end of the EE envelope; on
465 average it is small and negative at most sites (between 0 and -0.05; Figure 7). These bias
466 characteristics share similarities with those found over ocean (Section 3.1). Further, the
467 DB algorithm has two methods for modelling land surface reflectance (*Hsu et al.*, 2013,
468 2017): a method based on NDVI used over vegetated regions, and a surface data base for

469 brighter surfaces (deserts, mountains, urban areas), and similar bias characteristics are
470 found in both (Figure 8). This makes it likely that sensor calibration is a contributing
471 factor, since similar biases are found in both land and ocean algorithms, and for the two
472 different over-land surface reflectance determination methods.

473 Table 3 and Figure 7 also indicate that there is regional variability in performance.
474 Around half of global matchups are in North America or Europe, due to the density of
475 the AERONET network in these areas, and about 8 % are in the ‘boreal’ region (mostly
476 tundra or forested regions at high Northern latitudes). At sites in these regions, the DB
477 algorithm tends to perform well, with biases often smaller than 0.025 and more than 68 %
478 of retrievals matching within the EE. These are regions where the NDVI-based surface
479 reflectance determination method predominates. A fairly high quality of performance is
480 also seen in the South America, South Africa, and Oceania regions, although there is some
481 tendency to underestimate AOD in high-AOD conditions. Lower correlations at sites in
482 some of these regions (particularly Oceania) again reflect that the dynamic range of AOD
483 is fairly small compared to the magnitude of retrieval uncertainty.

484 Performance at tropical sites, particularly in the Sahel, Arabian Peninsula, Indian sub-
485 continent, and eastern Asia is poorer. This is likely due to a combination of the brighter
486 surface (less sensitivity to the aerosol signal, and potentially being near the critical albedo
487 where the TOA signal is invariant with AOD, e.g. *Seidel and Popp, 2012*), high variabil-
488 ity in aerosol composition (i.e. single aerosol models and the assumed AE for conversion
489 of band 1 AOD to 550 nm is less appropriate), and higher frequency of cirrus clouds
490 (which are harder to detect in AVHRR than sensors which have bands around 1.37 μm
491 like MODIS/VIIRS). Tropical cirrus cloud contamination is particularly problematic in

492 south-eastern Asia, and can affect Sun photometer data as well as satellite retrievals.
493 *Chew et al.* [2011] examined collocated Sun photometer and lidar data at Singapore and
494 found residual cirrus contamination present in around a third of the Sun photometer; the
495 resulting AOD bias for these cases was around 0.034, which is somewhat larger than the
496 instruments' nominal uncertainty. They also found that the bias induced in the Sun pho-
497 tometer data was larger than the typical bias introduced into satellite AOD from cirrus
498 contamination, so it is possible that the negative biases are in part due to this effect.
499 These regions often perform more poorly than others in over-land AOD retrieval algo-
500 rithms, so the difficulty is not limited to AVHRR or DB (*Levy et al.*, 2010, *Kahn et al.*,
501 2010, *Sayer et al.*, 2012b, 2013, 2014, *Reid et al.*, 2013, *Popp et al.*, 2016). This is also
502 reflected in Figure 8, in that uncertainties tend to be slightly larger for bright regions
503 where the data base method was used to estimate surface reflectance.

504 The available data volume for NOAA14 is an order of magnitude smaller (6,668 matches
505 from 123 sites, 58 of which provided at least 25 matches). As over ocean, this is due to
506 the more limited extent of the AERONET network during the 1995-1999 period. No
507 AERONET sites were active over the OCE region at this time. Figures 9 and 10 char-
508 acterize the AOD- and site- dependence of validation statistics for NOAA14 respectively,
509 and show the same tendencies as were observed for NOAA18 data in Figures 7 and 8.
510 Table 4 summarises statistics globally and regionally. The regional dependence of these
511 statistics is in general similar to that of NOAA18 (cf. Table 3), although the very limited
512 data volume in some regions makes it more difficult to assess how representative some
513 of these statistics are, particularly in Asia. Figure 10 suggests similar AOD- and type-

514 dependence of retrieval errors, thus it appears as though the over-land data from the two
515 sensors share similar error characteristics on the whole.

516 For both AVHRR sensors, a summary is that the retrieval tends to perform well in areas
517 with darker (more vegetated) surfaces, and where the aerosol type is not too variable in
518 time. In these cases the biases are small and the retrieval uncertainty is probably better
519 than $\pm(0.05+25\%)$, tracking the temporal variability of AOD well but with a tendency to
520 underestimate the AOD of high-AOD events. In more complicated tropical environments,
521 the data should be used with more caution, as there is a greater tendency to underestimate
522 AOD. However the correlation often remains high, suggesting the ability to identify high-
523 AOD events, despite this underestimation. Development of future versions of the AVHRR
524 DB products will therefore focus on better QA-filtering of data in these regions, whether
525 more appropriate aerosol optical models can be found, and development of separate error
526 models for the two different surface reflectance determination methods (NDVI vs. data
527 base) and/or geographic regions. The AOD biases relative to RMSE are larger over land
528 (e.g. Figure 7) than ocean, suggesting that decreasing the bias (which over land could be
529 achieved with improved radiometric calibration and/or surface reflectance determination)
530 could lead to non-negligible decreases in RMSE and shrinking the EE envelope.

531 The error characteristics for the DB AVHRR data over land also share some common
532 features with validation results from DB applied to SeaWiFS (*Sayer et al.*, 2012b) and
533 MODIS (*Sayer et al.*, 2013, 2014). All show better performance over vegetated than bright
534 land surfaces. This is a consequence of the fact that the aerosol signal is in general com-
535 paratively stronger over a vegetated (darker) surface, and the dynamic surface reflectance
536 model employed by DB over such surfaces helps in tracking temporal/directional varia-

537 tions. The similarity in bias characteristics between instruments, however, is harder to
538 explain. As AVHRR lacks bands in the blue spectral region which are key for the Sea-
539 WiFS/MODIS applications of DB, errors caused by aerosol optical model assumptions in
540 SeaWiFS/MODIS would not necessarily be expected to be the same. This similarity may
541 therefore be in part coincidental. The fact that AVHRR DB (land) and SOAR (water)
542 AOD biases show similar behavior, despite being independent algorithms, suggests that
543 sensor calibration plays some role in AVHRR's biases.

5. Comparison with other satellite products

5.1. AVHRR over ocean

544 As discussed in Section 2.2.1, the main other AVHRR data sets available at present
545 are the over-ocean NASA GACP and NOAA CDR products. It is difficult to do a direct
546 three-way comparison between these and SOAR, as there are differences in the available
547 wavelengths (i.e. GACP provides only 550 nm, CDR only band 1, SOAR both) and aggre-
548 gation levels in both time (GACP provides only monthly, CDR daily and monthly, neither
549 orbit-level) and space (CDR is on a 0.1° grid while GACP and SOAR are at 1°) between
550 the data sets. As a balance, this analysis provides a comparison of seasonal composites for
551 the year 2006 from NOAA18. This year was notable for aerosol events including a strong
552 dust storm in March, intense fires in north-eastern Russia and China in May, and a strong
553 El Niño leading to an intense biomass burning season in Indonesia, peaking in Septem-
554 ber/October (*Carboni et al.*, 2012, *Marlier et al.*, 2013, *Field et al.*, 2016). A seasonal
555 comparison means that the effects of calibration, sampling, and retrieval algorithm cannot
556 be directly separated, but it allows for a big-picture comparison which is more akin to the

557 way many data users approach these products (i.e. monthly or longer composites), and,
558 as noted, the types of comparison possible are constrained by the available data products.

559 The CDR product is aggregated to 1° to match the others (cf. Section 2.2.1), and
560 differences in monthly means are calculated before averaging to provide seasonal means
561 and differences. Figure 11 shows the resulting seasonal 550 nm AOD maps from SOAR, as
562 well as difference maps between SOAR and GACP/CDR (comparing SOAR and GACP
563 550 nm AOD, and then SOAR and CDR band 1 AOD, i.e. comparing common wavelengths
564 in both cases).

565 It is immediately apparent that the differences between SOAR and GACP, and SOAR
566 and CDR, show contrasting behaviour in several regions (e.g. SOAR is somewhat lower
567 than GACP in the Southern Ocean but somewhat higher than CDR in this region). Part
568 of the difference between the SOAR/GACP and SOAR/CDR comparisons is due to the
569 different wavelengths between the two comparisons (the former pair is 550 nm and the
570 latter 630 nm), although this should be a small effect (<0.02) in most cases since this
571 wavelength difference is not that large. Thus, where the SOAR/GACP and SOAR/CDR
572 comparisons show offsets of opposing signs, it is likely that the difference is dominated by
573 some combination of calibration, algorithm, and sampling, rather than this wavelength
574 difference.

575 The SOAR AOD is higher than both GACP and CDR in many high-AOD continental
576 outflow regions (e.g. the Saharan dust belt, central African biomass burning, north-eastern
577 Asia). Differences in such regions are expected to be particularly large, because of the
578 limited information content of the sensor and so need to make considerable simplifying
579 assumptions about aerosol optical model (size distribution and refractive index). SOAR

580 picks from one of several bimodal optical models, while both GACP and CDR assume
581 the same aerosol properties for fine and coarse aerosol modes globally. This will lead to
582 larger (systematic) errors in high-AOD conditions (as scattering/absorption properties of
583 marine and dust, smoke, continental, or other aerosol types differ). As GACP and CDR
584 assume a spherical coarse mode (SOAR includes nonspherical dust), further, errors will
585 exhibit a larger angular-dependence in the case of nonspherical dust, and such errors will
586 not necessarily cancel out through averaging to a longer time scale (e.g. *Zhao et al.*, 2004,
587 *Lee et al.*, 2017). Some further analysis of the implications of the aerosol optical model
588 assumptions, as it pertains to the GACP product and long-term trends in particular,
589 is provided by *Mishchenko et al.* [2012]. Such differences in these areas are therefore
590 expected, although it is interesting that SOAR AOD is higher than the others in these
591 cases given that Figure 1 indicates a tendency to underestimate the AOD in high-AOD
592 conditions. Therefore it is possible that GACP/CDR are biased more negatively.

593 Validation of the GACP product was performed by *Liu et al.* [2004], although this was on
594 a monthly 1° basis as opposed to an instantaneous basis (as performed for SOAR herein),
595 and predated NOAA18's launch. *Geogdzhayez et al.* [2015] did not present additional
596 validation for NOAA18, although noted that there did not appear to be sensor-to-sensor
597 discontinuities between the GACP record from different sensors, by using years where
598 data from multiple overlapping sensors were available. Hence, it is plausible that the bias
599 tendencies of NOAA18 are similar to those found for the earlier sensors by *Liu et al.* [2004],
600 which were ship-based measurements indicating a random error of 0.04 and positive bias
601 around 11%. In this sense the fact that SOAR AOD (which appears to have a small
602 bias with respect to AERONET/MAN in clean conditions) is higher than GACP is also

603 unusual if the NOAA18 GACP record really does have a positive bias. The version 3 CDR
604 product has not been validated extensively, particularly for NOAA18, although available
605 analyses (*Zhao et al.*, 2004, *Zhao*, 2016) suggest a systematic error at 630 nm in open-
606 ocean condition of order 0.03, and random errors of order 0.11. CDR also allows retrieval
607 of negative AOD (down to -0.2), although unphysical, in an attempt to stop retrieval
608 errors in low-AOD conditions being positively skewed (*Zhao*, 2016). It is not clear how
609 error characteristics are likely to change in areas of high aerosol loading. Due to the
610 small data volume, and monthly rather than instantaneous comparison, it is difficult to
611 disentangle how algorithm and sampling may be combining to cause the observed offsets
612 in low- and high-AOD regions.

613 Smaller differences in open-ocean conditions may arise from factors such as the relative
614 aggressiveness of cloud screening, both in terms of the risk of cloud contamination, which
615 typically causes high-AOD artifacts, and relative sampling of near-cloud vs. far-from-cloud
616 pixels, the former of which may have real higher AODs due to e.g. aerosol humidification
617 (*Twohy et al.*, 2009, *Várnai et al.*, 2013). *Zhao et al.* [2013] found differences in zonal or
618 monthly mean AVHRR-derived AOD at 630 nm of up to 0.04 dependent upon strictness
619 of cloud masking. Detection of optically-thin cirrus clouds is particularly difficult for the
620 AVHRRs compared to e.g. MODIS as they lack a band near $1.37 \mu\text{m}$, which is sensitive to
621 high clouds. Additional regional offsets can be explained by the fact that the GACP algo-
622 rithm assumes a globally-constant near-surface wind speed of 7 ms^{-1} (*Mishchenko et al.*,
623 1999) while SOAR uses ancillary meteorological information to calculate the influence of
624 wind speed on surface reflectance for each retrieval. This constant-wind assumption is
625 known to lead to regional offsets in AOD of either sign of order 0.01-0.02, dependent on

626 typical local wind speeds, and also mean that Sun glint can be under- or overscreened
627 (*Zhang and Reid, 2006, Sayer et al., 2010*). The CDR product uses a constant Lambertian
628 albedo (*Zhao, 2016*), which is more-or-less equivalent to a constant wind speed, although
629 does include a Sun-glnt contribution as well.

630 Another notable offset is that SOAR-GACP is quite negative in the Southern Ocean
631 while SOAR-CDR is positive (and somewhat smaller). The phenomenon of high Southern
632 Ocean AOD is found in several satellite data sets (including GACP but not CDR), but
633 not seen in AERONET or MAN and so thought to be partially an artifact. The feature
634 is also seen in Northern storm tracks, but is less prominent due to cloud and land cover.
635 The causes were investigated by *Toth et al. [2013]*, with a main focus on MODIS data,
636 who concluded that cloud contamination was responsible for up to 30-40 % but other
637 assumptions (such as a fixed assumed surface wind speed) were responsible for the rest.
638 It therefore seems likely that this conclusion is applicable to the AVHRR products as well.
639 The fact that this is not seen in CDR may suggest that cloud contamination is the larger
640 factor relevant for the AVHRRs, and its absence in CDR an indication of more aggressive
641 cloud masking; following *Zhao et al. [2013]*, a fairly strict cloud mask was adopted in the
642 version 3 CDR product.

5.2. MODIS over land and ocean

643 As many research applications take monthly AOD products as a basis, rather than L2
644 data, it is instructive to see how similar such composites are between the new AVHRR
645 data set and other commonly-used products such as MODIS. Figure 12 provides such a
646 comparison between NOAA18 AVHRR and MODIS Aqua monthly data (Section 2.2.3),
647 constructed from the overlapping time period of the two sensors (2006-2011). To increase

648 the robustness of statistics only grid cells containing data from at least 24 months are
649 considered, which removes points in areas of high cloud cover (e.g. tropical rainforests)
650 and high latitudes where clouds and polar night strongly limit coverage in some months.
651 Figure 12(a) shows the main global features of AOD are represented in the AVHRR data.
652 Note that as this is a multiannual mean composite, the strength of seasonal features can
653 be attenuated. The other panels provide important contextual information.

654 Over the open ocean, AVHRR AOD is often lower than MODIS by 0-0.03. This is
655 consistent with AVHRR having a near-zero AOD bias in such conditions (Section 3),
656 and MODIS having a positive bias of order 0.015 on average (*Sayer et al.*, 2012d, *Levy*
657 *et al.*, 2013). For some grid cells near the Equator a positive offset is seen instead, which
658 may be due to the aforementioned greater difficulty of thin cirrus cloud detection in
659 AVHRR than MODIS. In general over the remote ocean the correlation coefficient varies
660 from 0-0.8, dependent on the precise region. This is because the seasonal variation in
661 AOD is small relative to the retrieval uncertainties (which tend to have a non-negligible
662 systematic component), such that a large correlation is only found in areas with seasonal or
663 periodic continental aerosol transport. AVHRR has a slightly more negative offset at high
664 latitudes, which is consistent with *Toth et al.* [2013] who identified cloud contamination
665 as a probable contributing cause to a high band of AOD in MODIS. The RMS difference
666 is small (0-0.03 over the cleanest ocean regions, 0.03-0.06 over other open oceans), but
667 higher in these storm tracks, likely due again to cloud contamination in MODIS.

668 AOD is also lower in AVHRR than MODIS over dust aerosol outflow regions of Africa
669 and Asia, consistent both with a slight low bias in AVHRR, and a positive (on average) bias
670 in MODIS due to its lack of nonspherical dust aerosol models (*Levy et al.*, 2003, *Zhang*

671 *and Reid, 2006, Banks et al., 2017*). The correlation in these outflow regions remains
672 high (0.8-1), indicating that both track the same seasonal and interannual variability in
673 dust transport. Over smoke outflow regions East of southern Africa the correlation is
674 similarly high and offset/RMS difference small. In contrast, the RMS difference over
675 the smoke outflow region from southern Africa into the southern Atlantic is larger and
676 correlation lower. Closer examination reveals that this is due to some of this smoke being
677 masked as cloud in the AVHRR data, resulting in it being underrepresented in the monthly
678 composite. This doesn't show in Figure 12(c) because this smoke transport only occurs
679 in a few months of the year.

680 As over ocean, the regions over land with low correlation between AVHRR and MODIS
681 monthly composites tend to be those with fairly persistent low AODs such as large parts
682 of Australia and mountainous areas of North and South America. In most of these areas
683 the offset and RMS difference between the two sensors tends to be 0-0.03, confirming that
684 the two are consistent in this lack of temporal variation. In contrast the two are highly
685 correlated, with fairly low bias, in smoke source regions in South America and Africa.
686 Intermediate regions (i.e. fairly low AOD but moderate seasonality), such as much of the
687 Americas and Europe, have intermediate correlation and small biases (0-0.03, of either
688 sign). Improving correlation or decreasing RMS in these areas may be difficult as both
689 sensors show fairly small biases with respect to AERONET in these regions, although
690 those of AVHRR are slightly larger (Section 4 and *Sayer et al., 2013*). Larger offsets
691 and/or RMS differences are found in three land regions:

692 1. Near the limits of AVHRR spatial coverage around bright deserts. These differences
693 are likely dominated by a combination of AVHRR retrieval error, and differences in spatial
694 coverage.

695 2. Over high-AOD regions of China. These often have limited sampling due to high
696 cloud cover; available validation suggests MODIS DB has less bias than AVHRR DB.
697 Refinement of seasonal aerosol optical model assumptions may help, although this region
698 has very high spatiotemporal variation in aerosol sources.

699 3. In central Asia, most notably around Iran and surrounding countries. This difference
700 has been traced to a limitation of the MODIS C6 DB product in this area, which has been
701 fixed for the upcoming Collection 6.1 reprocessing. Future data versions should show a
702 higher level of consistency.

5.3. Time series comparison at AERONET sites

703 A goal of the Deep Blue aerosol project is to move towards consistency in AOD de-
704 rived from multiple satellite sensors using similar measurement types and retrieval tech-
705 niques. As such, this Section examines the AOD time series obtained at selected long-term
706 AERONET sites covering the era to which DB/SOAR have been applied. Although there
707 are several hundred AERONET sites in operation, very few have operated continuously
708 or with few gaps since the mid-1990s, which limits the extent of the comparison. A total
709 of five sites are considered in this analysis. Over ocean these are Capo Verde (Atlantic
710 dust outflow) and Wallops (US East coast continental outflow); there is unfortunately no
711 well-sampled long-term ‘clean marine’ site covering both the NOAA14 and NOAA18 eras.
712 Over land they are Alta Floresta (Brazilian rainforest with seasonal biomass burning),
713 NASA Goddard Space Flight Center (GSFC, suburban Eastern US) and Mongu (situ-

714 ated in a semi-arid part of Zambia with seasonal biomass burning). Four of these five
715 AERONET sites were identified by *Li et al.* [2016] as providing a moderate or high level of
716 representivity of their surrounding regions on these 1° spatial scales; the other (Wallops)
717 was not evaluated by *Li et al.* [2016]. Thus, although the choice of AERONET sites is
718 strongly constrained by the limited number which have been in operation for much of the
719 period from the mid-1990s until 2011, it is fortunate that these sites appear to sample air
720 masses representative of the spatial scales of satellite level 3 products.

721 In addition to AERONET and AVHRR, the time series analysis uses the monthly mean
722 MODIS and SeaWiFS data sets described in Section 2.2. The AERONET daily mean
723 product is used to calculate both the monthly mean AOD (for those months with at least
724 3 days with observations) as well as the central one standard deviation (68%) range of
725 daily mean AOD, to provide an indication of day-to-day AOD variability within the month
726 and thus indicate those periods where sampling issues are most likely to be important.
727 For all the satellite products, the grid cell in which the AERONET site lies was used to
728 extract the time series.

729 The resulting mean AOD time series, with the AERONET variability providing a shaded
730 background, are shown in Figures 13 and 14 for the ocean and land sites respectively. The
731 correlation and median bias between AVHRR and the other monthly mean AOD data sets
732 are given in Table 5. The comparison against AERONET here provides an additional look
733 at the validation. Even though SeaWiFS and MODIS are retrievals and not a ground truth
734 like AERONET, the rationale for providing statistics comparing AVHRR to each of these
735 is to assess the level of consistency between the satellite products, which is subtly different
736 than assessing the error in the AVHRR data. Thus these analyses provide different but

737 complementary information. Note that as the NOAA14 time series processed ends in
738 1999, prior to the launch of the Aqua platform in 2002, there is no comparison between
739 this pair. Both NOAA14 and NOAA18 AVHRR are represented by black lines, but as
740 there is no temporal overlap between the two the provenance of each part of the time
741 series is unambiguous.

742 The time series all provide similar AOD magnitude and seasonality, and monthly mean
743 values typically lie within the central 68% range of daily means observed by AERONET
744 for the month in question, which is encouraging. Correlation coefficients range between
745 0.72 and 0.99 (Table 5), confirming that the seasonal and interannual variability are
746 broadly consistent between AVHRR and the other data sets. Biases are often of similar
747 magnitude between NOAA14 and NOAA18 AVHRR, and in terms of correlation coefficient,
748 there does not appear to be a systematic pattern whereby NOAA14 or NOAA18 is
749 systematically more strongly-correlated with AERONET or the other data sets. However
750 it is important to note that (particularly for NOAA14) the number of overlapping months
751 between data sets is small. Hence, there is inherently likely to be larger uncertainty on
752 these statistics compared with, for example, the instantaneous matchup statistics obtained
753 in validation with AERONET direct-Sun data (see e.g. *Schonbrödt and Perugini, 2013* for
754 discussions of uncertainties in the estimation of correlation coefficients). Despite the small
755 data volume, the data do suggest the future potential for combining multi-sensor data sets
756 like DB and SOAR to produce a consistent long-term data record, possibly after further
757 bias-correction steps such as have been developed for data assimilation applications (e.g.
758 *Zhang and Reid, 2006, Hyer et al., 2011, Schutgens et al., 2013*).

6. Conclusions

759 A primary goal of the Deep Blue aerosol project is to be able to create a long-term
760 aerosol data record with broadly consistent error characteristics that is based on the
761 use of satellite sensors with similar measurement capabilities. The approach is to apply
762 similar retrieval algorithms that account for the particular characteristics of each sensor.
763 The feasibility of using the AVHRRs for AOD retrieval over ocean has been established for
764 decades, but existing over-land AOD retrievals proposed for AVHRR have been limited
765 in scope.

766 This study has established that the DB and SOAR algorithms can be adapted for use
767 with the AVHRR sensors to retrieve AOD over land (aside from snow-covered or very
768 bright desert) and ocean surfaces. As well as providing an over-ocean record with compa-
769 rable heritage to the other SOAR algorithms, this opens up (for the first time for AVHRR)
770 near-global over-land AOD products on both an instantaneous (i.e. Level 2 orbit-level)
771 and aggregated (Level 3 daily/monthly) basis. The bulk of the available validation data is
772 for NOAA18, although the results indicate a similar quality of performance, to the extent
773 that can be diagnosed, from the earlier NOAA11 and NOAA14 AVHRR instruments as
774 well. This is encouraging in terms of being able to extend these data records back in
775 time, particularly for the new over-land capability. The sparsity of available validation
776 data prior to the mid-1990s will, however, present a challenge for evaluation when the
777 algorithms are applied to the earlier AVHRRs.

778 The typical level of uncertainty on instantaneous AOD retrieved, which appears to be
779 around $\pm(0.05+25\%)$ over land and $\pm(0.03+15\%)$ over water, is a little higher than the
780 application of DB/SOAR or similar algorithms to more advanced similar sensors such as

781 SeaWiFS, MODIS, and VIIRS. This is due to the well-known more limited capabilities
782 of the AVHRR sensors (only two broad reflective solar bands, without on-board calibra-
783 tion). However this should still be sufficient for many quantitative scientific applications,
784 and may be able to be reduced further by refinement of retrieval algorithm and sensor
785 calibration. In particular, AOD time series at long-term AERONET sites examined are
786 well-correlated and typically exhibit small biases with respect to both AERONET and
787 other satellite products. Differences between AVHRR and MODIS AOD data are gen-
788 erally consistent with their known error characteristics, and can hopefully be decreased
789 in future versions. This suggests that the future goal of creating a harmonized data set
790 from multiple sensors, which would be a great advantage for the study of multi-decadal
791 variations in aerosol loading, is achievable. To assess and improve upon the sensor cali-
792 bration used in the creation of the data set, to further aerosol optical models refine, and
793 to extend DB/SOAR processing to the whole AVHRR record, making use of available
794 validation data and periods of overlap from multiple sensors, are therefore the next steps
795 toward this goal.

Acknowledgements

796 Further information about Deep Blue, including file formats, documentation, and data
797 download locations for the various data sets, is available at <https://deepblue.gsfc.nasa.gov>.
798 Data hosting resources were provided by the NASA High-End Computing (HEC)
799 Program through the NASA Center for Climate Simulation (NCCS) at Goddard
800 Space Flight Center; the AVHRR Deep Blue data products are freely available
801 from <https://portal.nccs.nasa.gov/datashare/AVHRRDeepBlue>. The MERRA data
802 used have been provided by the Global Modeling and Assimilation Office (GMAO)

803 at NASA Goddard Space Flight Center (<https://gmao.gsfc.nasa.gov>). AERONET
804 and MAN data are available from <https://aeronet.gsfc.nasa.gov>. The responsible
805 investigators for AERONET sites and MAN cruises are thanked for the creation
806 and stewardship of the Sun photometer data records. GACP data were obtained
807 from https://gacp.giss.nasa.gov/data/time_ser. NOAA AVHRR ocean AOD data were
808 obtained from <https://www.ncei.noaa.gov/data/avhrr-aerosol-optical-thickness>, and X.
809 Zhao (NOAA) is thanked for discussions about the CDR product. Data processing was
810 facilitated by use of the GNU Parallel utility by *Tange* [2011]. The Editor and three
811 reviewers are thanked for numerous useful comments, which improved the content and
812 clarity of the manuscript.

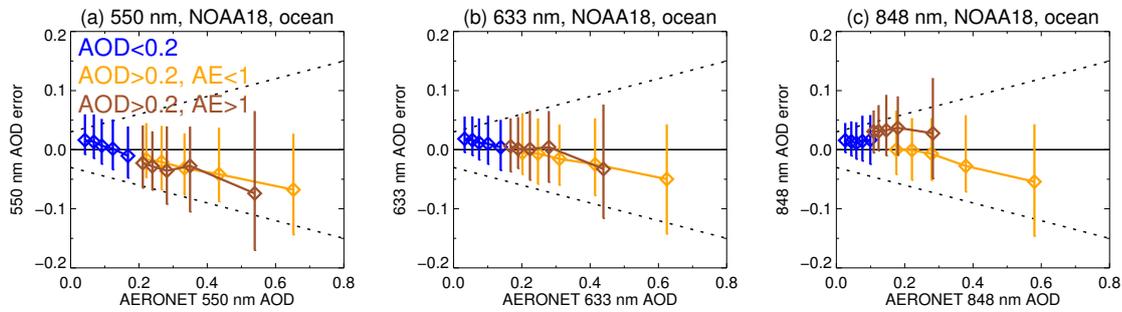


Figure 1. Binned median (points) and central 68% (lines) over-ocean AOD retrieval error (AVHRR-AERONET) for NOAA18, for (a) 550 nm, (b) band 1, and (c) band 2. Data are split into (blue) AERONET AOD at 550 nm < 0.2 , (orange) AERONET AOD at 550 nm ≥ 0.2 and $AE < 1$, and (brown) AERONET AOD at 550 nm ≥ 0.2 and $AE \geq 1$. Matchups within each category are divided into five equally-populated bins. Dashed black lines indicate the EE, $\pm(0.03+15\%)$.

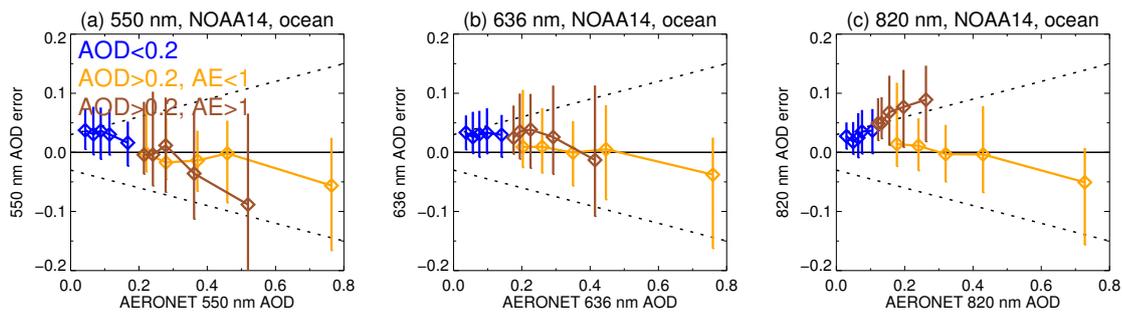


Figure 2. As Figure 1, except for NOAA14.

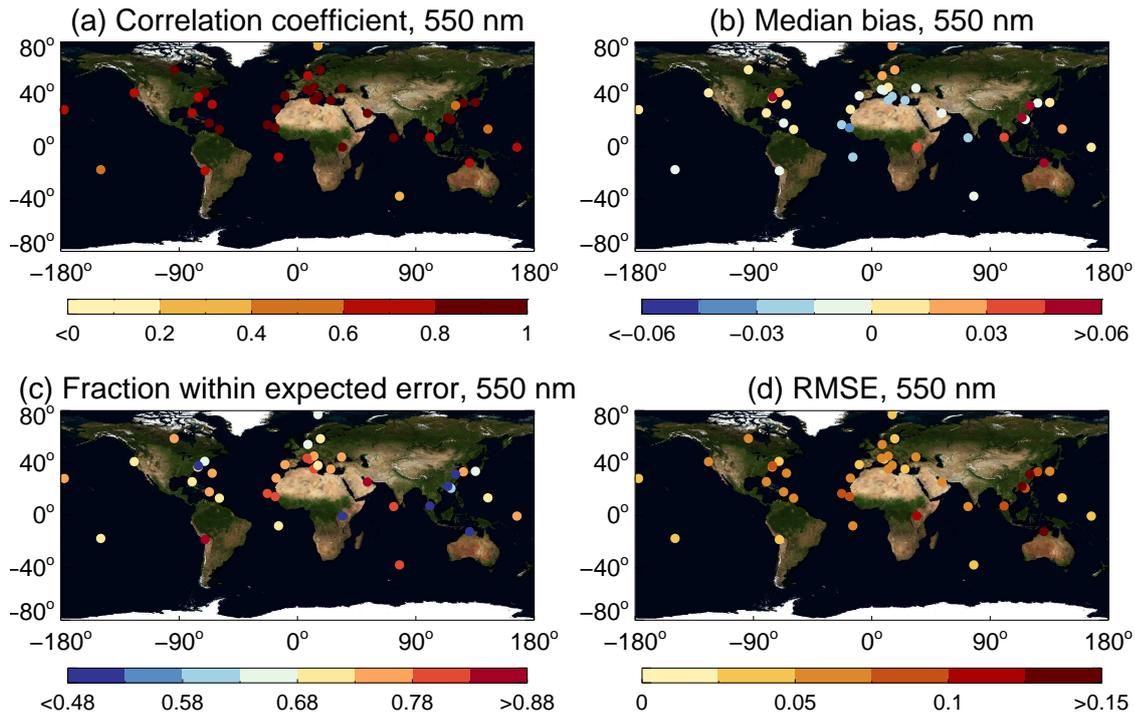


Figure 3. Site-by-site (a) correlation coefficient, (b) median bias, (c) fraction agreeing within the EE, and (d) root mean square error for over-ocean NOAA18 and AERONET matchups at 550 nm.

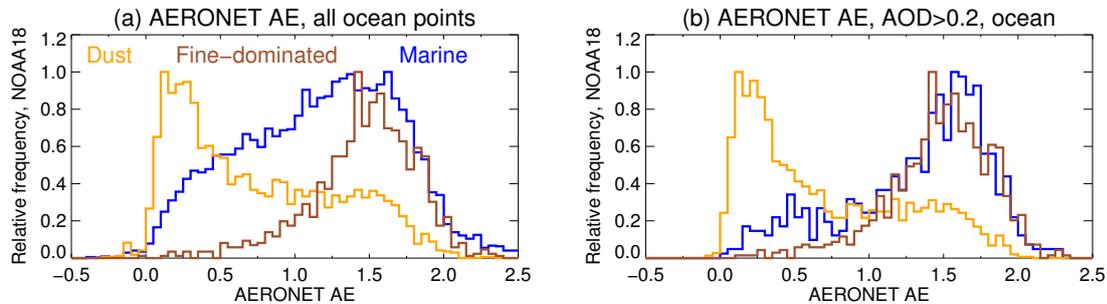


Figure 4. Histograms of AERONET AE, for (a) all ocean matchups with NOAA18, and (b) only NOAA18 matchups where AERONET AOD at 550 nm ≥ 0.2 . Points split to show cases where SOAR chose (blue) maritime, (orange) dust, and (brown) fine-dominated aerosol optical models.

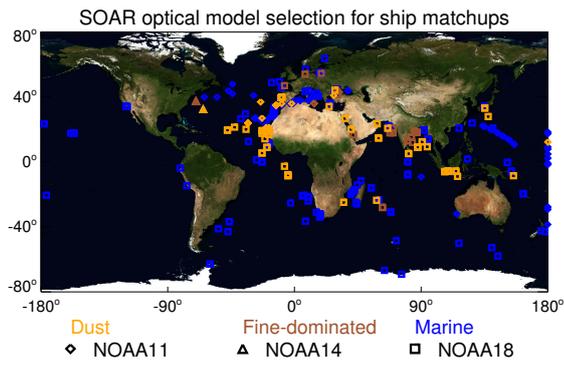


Figure 5. Optical models chosen by SOAR for the AVHRR/ship matchups. Orange indicates matchups where the dust model was chosen, brown the fine-dominated model, and blue the maritime model. Diamonds, triangles, and squares indicate NOAA11, NOAA14, and NOAA18 respectively.

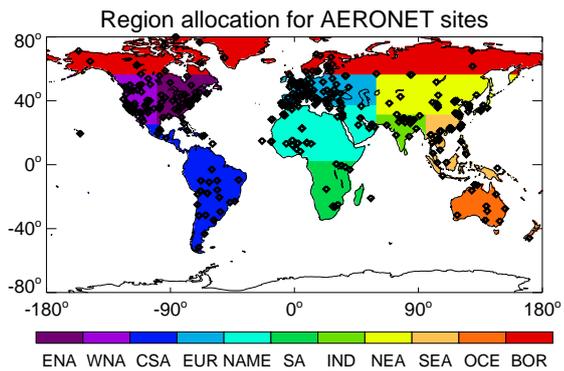


Figure 6. Site locations (black diamonds) and region assignment for over-land NOAA18 DB and AERONET matchups. Regions are boreal (BOR), Eastern North America (ENA), Western North America (WNA), Central/South America (CSA), Europe (EUR), North Africa/Middle East (NAME), Southern Africa (SA), Indian subcontinent (IND), North-Eastern Asia (NEA), South-Eastern Asia (SEA), and Oceania (OCE).

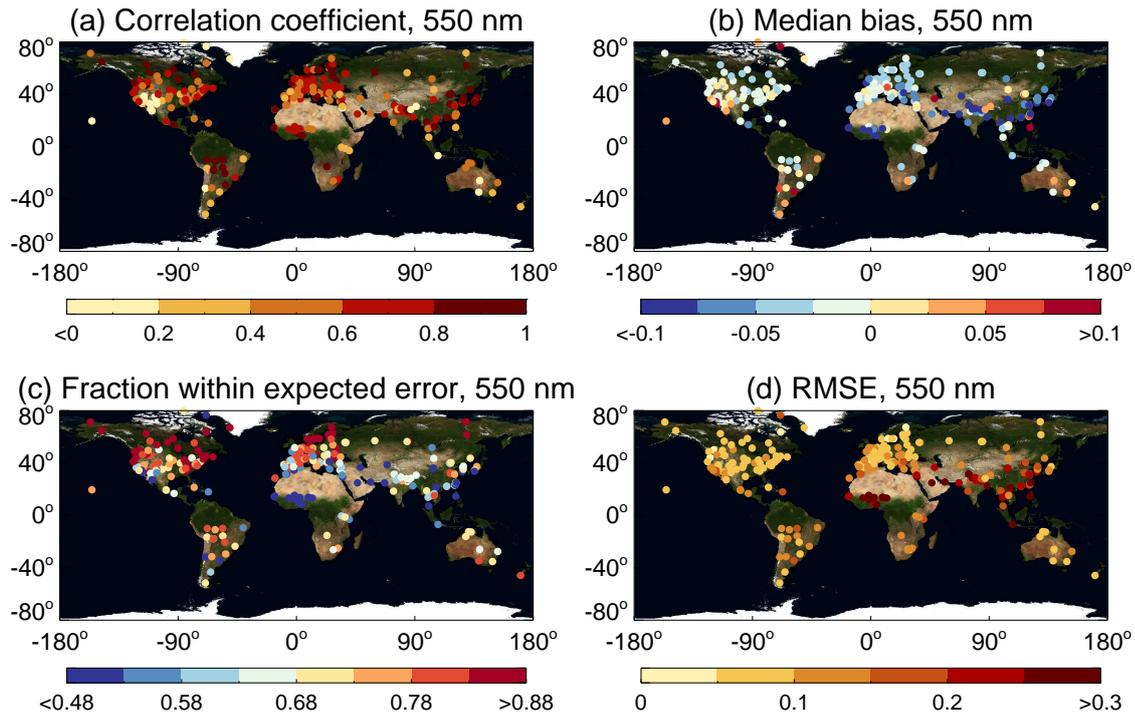


Figure 7. As Figure 3, except for the comparison between NOAA18 and AERONET sites over land, and note different color scale range in panels (b) and (d). Data shown only for sites with at least 25 matchups.

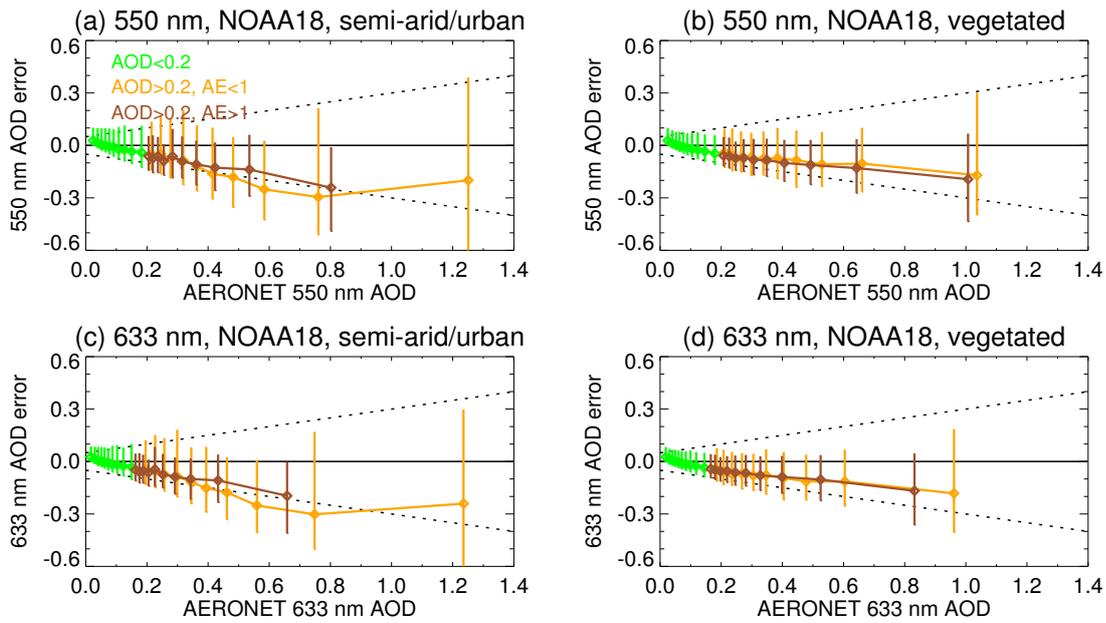


Figure 8. As Figure 1, except for NOAA18 DB matchups over land at 550 nm and band 1, and matchups within each category are divided into 10 equally-populated bins. Note that axis ranges are also different. Data shown separately for (a,c) matchups from semi-arid/urban pixels where the surface reflectance data base method was used and (b,d) vegetated pixels where the NDVI-based surface reflectance model was used. The low-AOD ‘background’ set are also indicated in green, rather than blue. Dashed black lines indicate the over-land EE, $\pm(0.05+25\%)$.

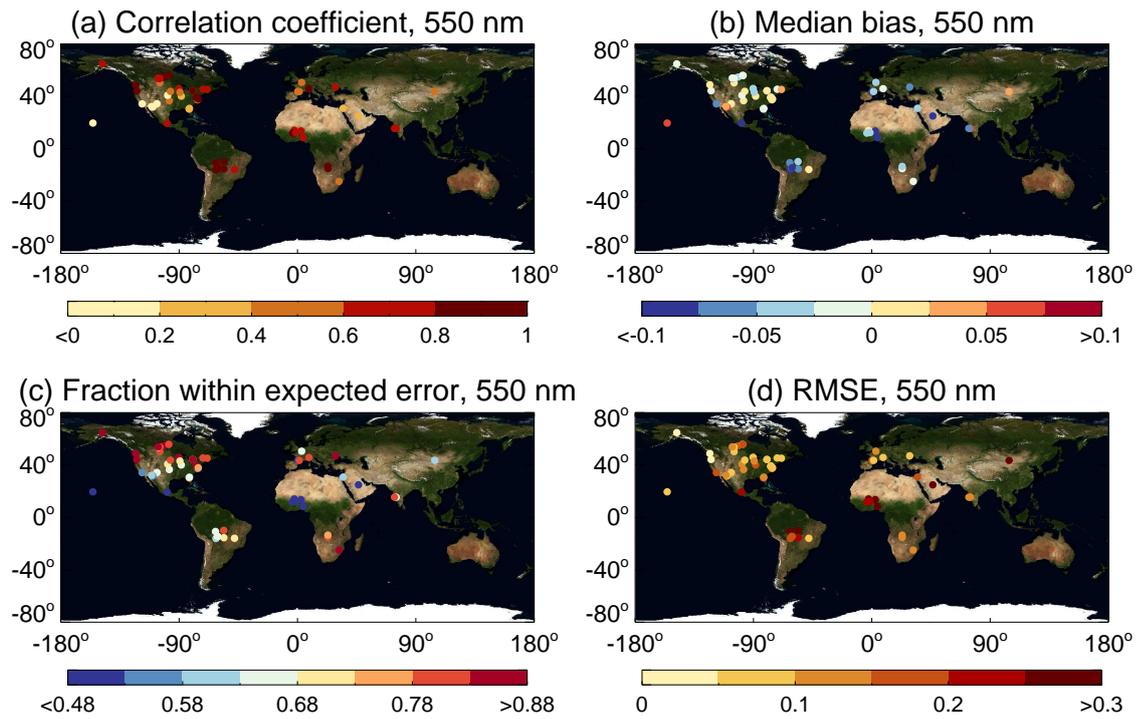


Figure 9. As Figure 7, except for the comparison between NOAA14 and AERONET sites over land.

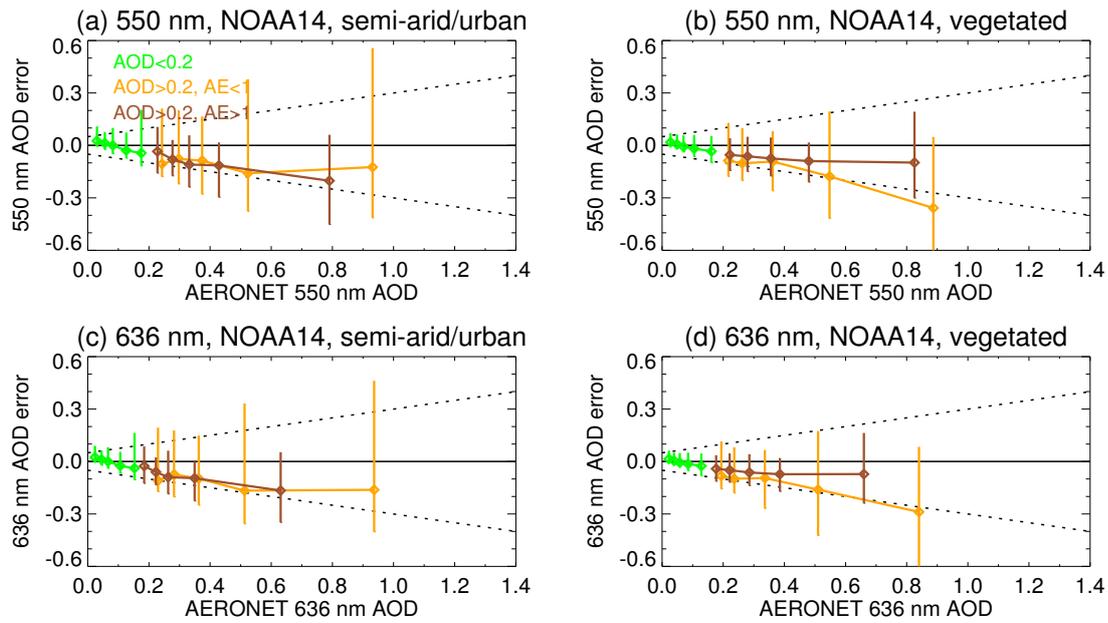


Figure 10. As Figure 8, except for NOAA14 matchups over land, and with half the number of bins in each category due to the smaller data volume.

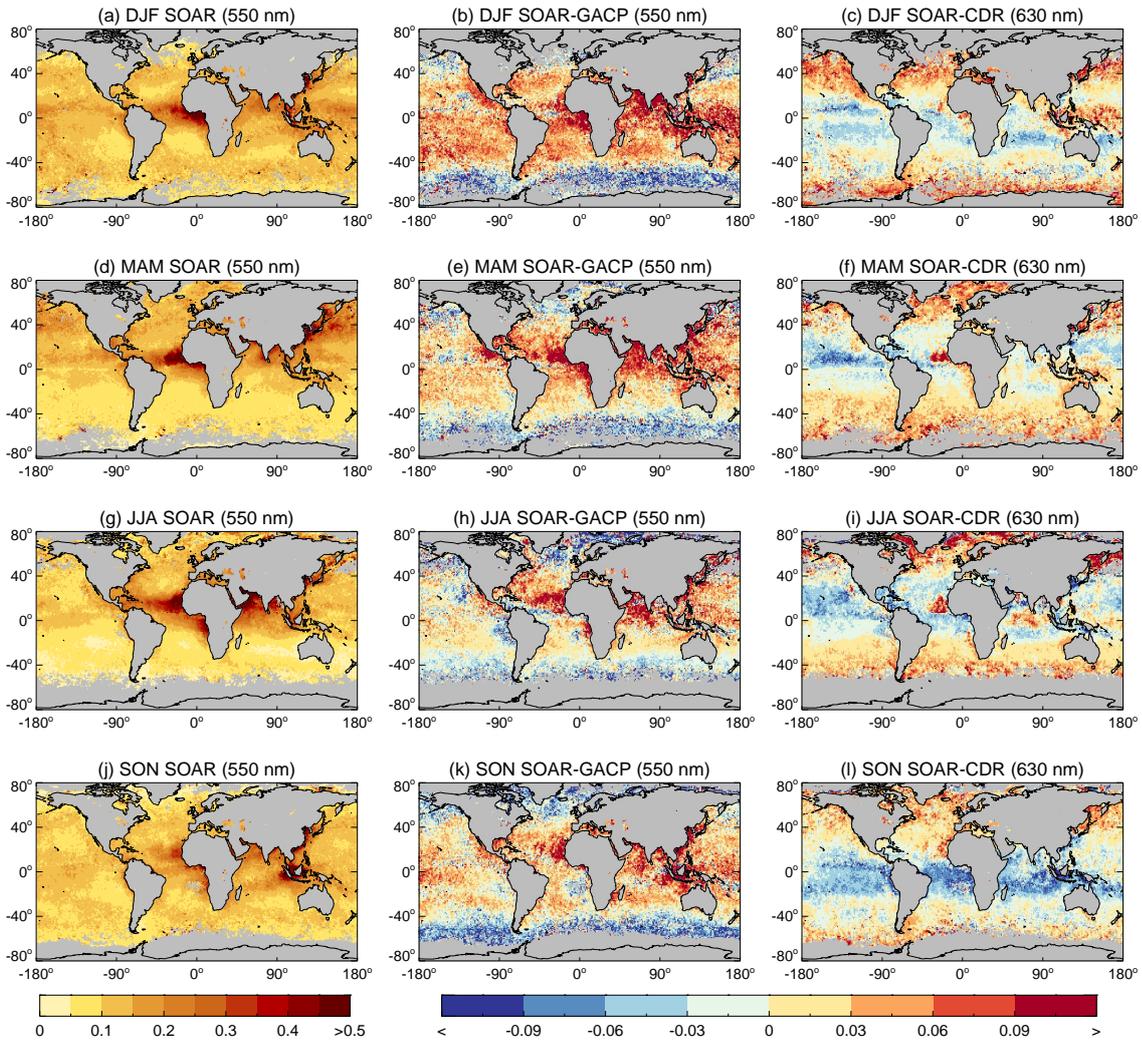


Figure 11. Seasonal composites from the year 2006, of NOAA18 (left column) SOAR 550 nm AOD, (center) SOAR-GACP 550 nm AOD, and (right) SOAR-CDR band 1 AOD. Grid cells without valid data are shaded in grey.

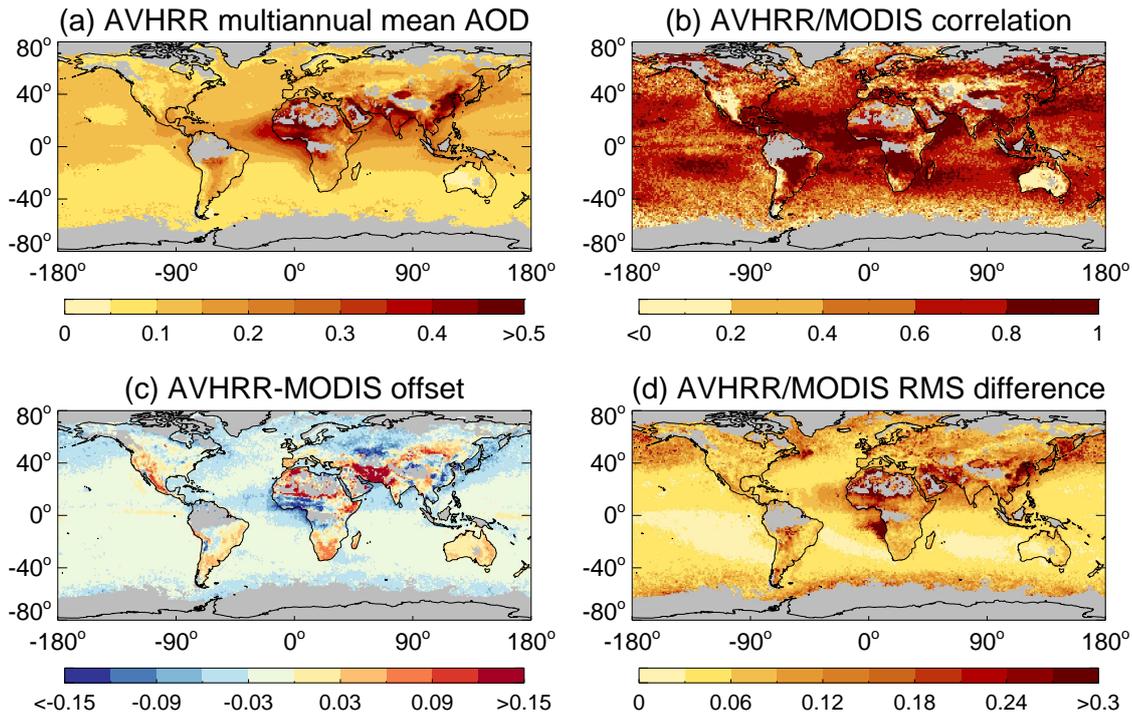


Figure 12. Global statistics of comparison between NOAA18 AVHRR and MODIS Aqua AOD at 500 nm. Panels show (a) multiannual mean AVHRR AOD from matched monthly points, (b) Pearson’s correlation coefficient, (c) the median AVHRR-MODIS offset, and (d) the RMS difference between the two. Grid cells without sufficient valid data are shaded in grey.

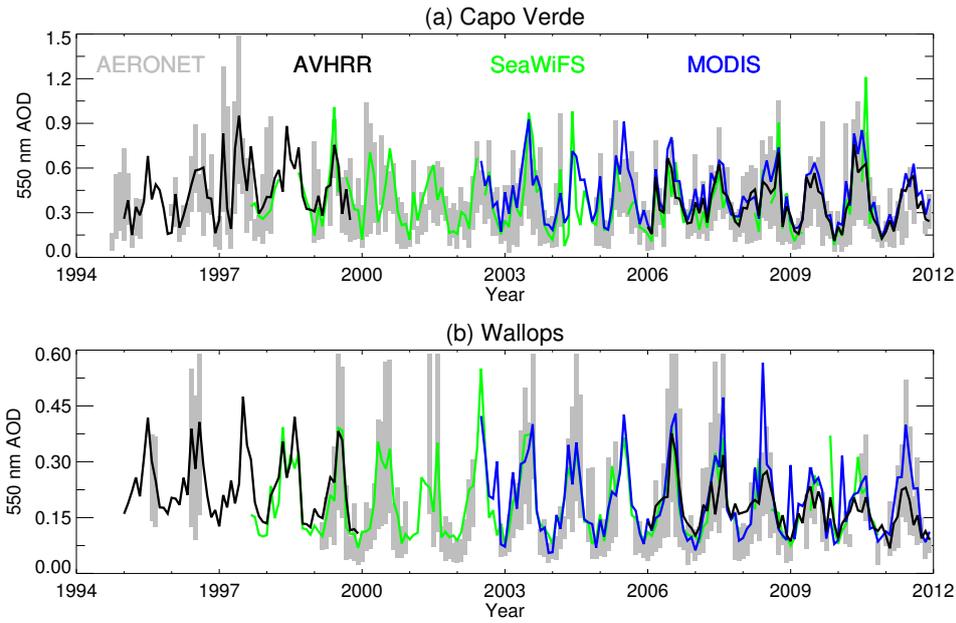


Figure 13. Time series of 550 nm AOD at two long-term coastal/island AERONET sites. The shaded grey area indicates the central 68% range of AERONET daily mean AOD within a month. Black, green, and blue lines indicate AVHRR (SOAR), SeaWiFS (SOAR), and MODIS Aqua (ocean) retrieved monthly mean AOD respectively.

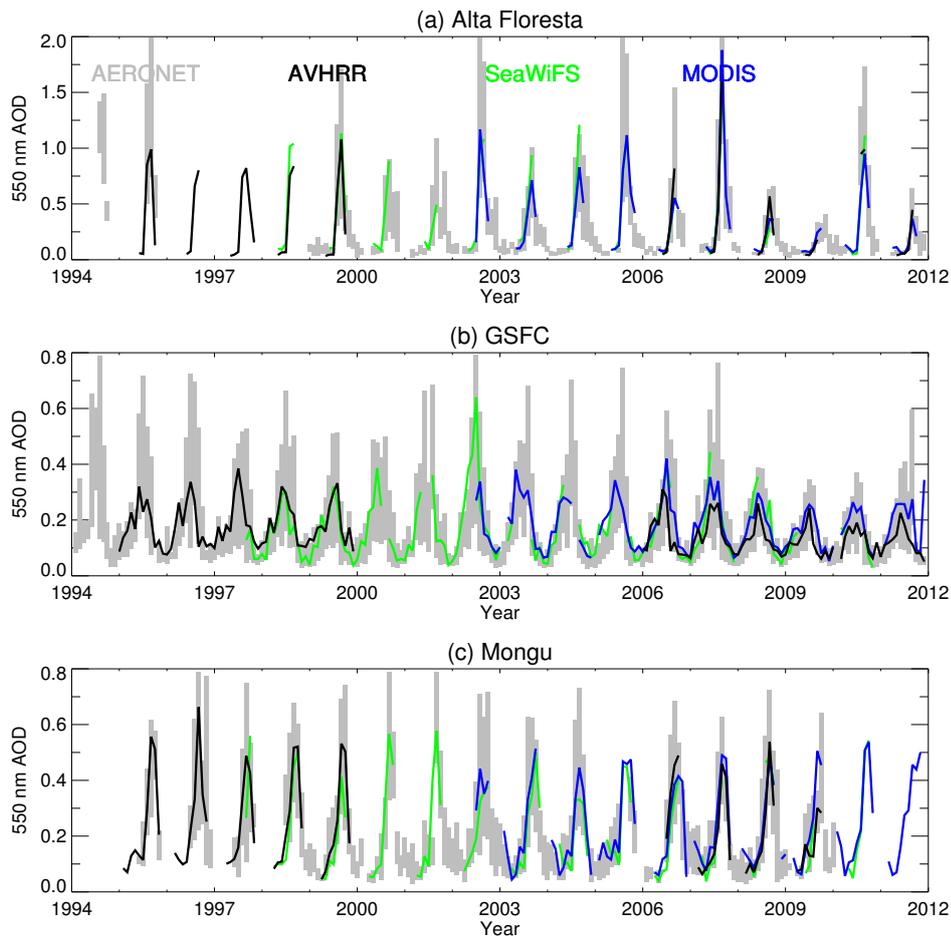


Figure 14. Time series of 550 nm AOD at three long-term land AERONET sites. The shaded grey area indicates the central 68 % range of AERONET daily mean AOD within a month. Black, green, and blue lines indicate AVHRR, SeaWiFS, and MODIS Aqua retrieved monthly mean AOD respectively, in all cases from the DB algorithm.

References

- 813 Ahmad, Z., B. A. Franz, C. R. McClain, E. J. Kwiatowska, J. Werdell, E. P. Shettle, and B. N.
 814 Holben (2010), New aerosol models for the retrieval of aerosol optical thickness and normalized
 815 water-leaving radiances from the SeaWiFS and MODIS sensors over coastal regions and open

- 816 oceans, *Appl. Opt.*, *49*(29), 5545–5560, doi:10.1364/AO.49.005545.
- 817 Banks, J. R., H. E. Brindley, G. Stenchikov, and K. Schepanski (2017), Satellite retrievals of
818 dust aerosol over the Red Sea and the Persian Gulf (2005-2015), *Atmos. Chem. Phys.*, *17*,
819 3987–4003, doi:10.5194/acp-17-3987-2017.
- 820 Carboni, E., G. E. Thomas, A. M. Sayer, R. Siddans, C. A. Poulsen, R. G. Grainger, C. Ahn,
821 D. Antoine, S. Bevan, R. Braak, H. Brindley, S. DeSouza-Machado, J. L. Deuzé, D. Diner,
822 F. Ducos, W. Grey, C. Hsu, O. V. Kalashnikova, R. Kahn, P. R. J. North, C. Salustro, A. Smith,
823 D. Tanré, O. Torres, and B. Veihelmann (2012), Intercomparison of desert dust optical depth
824 from satellite measurements, *Atmos. Meas. Tech.*, *5*, 1973–2002, doi:10.5194/amt-5-1973-2012.
- 825 Chew, B. N., J. R. Campbell, J. S. Reid, D. M. Giles, E. J. Welton, S. V. Salinas, and S. C.
826 Liew (2011), Tropical cirrus cloud contamination in sun photometer data, *Atm. Env.*, *45*(37),
827 6724–6731, doi:10.1016/j.atmosenv.2011.08.017.
- 828 Dubovik, O., and M. D. King (2000), A flexible inversion algorithm for retrieval of aerosol optical
829 properties from Sun and sky radiance measurements, *J. Geophys. Res.*, *105*.
- 830 Eck, T. F., B. N. Holben, J. S. Reid, O. Dubovik, A. Smirnov, N. T. O’Neill, I. Slutsker, and
831 S. Kinne (1999), Wavelength dependence of the optical depth of biomass burning, urban, and
832 desert dust aerosols, *J. Geophys. Res.*, *104*(D24), 31,333–31,349.
- 833 EUMETSAT (2016), Polar Multi-Sensor Aerosol Product: User guide, V2, *Tech. rep.*, EUMET-
834 SAT, Darmstadt, Germany, available online at www.eumetsat.int under Data > Technical
835 Documents > GDS Metop > PMAp [Accessed 27 Feb 2017].
- 836 Field, R. D., G. van der Werf, T. Fanin, E. J. Fetzer, R. Ruller, H. Jethva, R. Levy, N. J. Livesey,
837 M. Luo, O. Torres, and H. M. Wordern (2016), Indonesian fire activity and smoke pollution in
838 2015 show persistent nonlinear sensitivity to El Niño-induced drought, *Proc. Natl. Acad. Sci.*,

839 113(33), 9204–9209, doi:10.1073/pnas.1524888113.

840 Gao, L., J. Li, L. Chen, L. Zhang, and A. K. Heidinger (2016), Retrieval and validation of
841 atmospheric aerosol optical depth from AVHRR over China, *IEEE Trans. Geosci. Remote*
842 *Sens.*, 54(1), 6280–6291, doi:10.1109/TGRS.2016.2574756.

843 Geogdzhayez, I. V., M. I. Mishchenko, J. Li, W. B. Rossow, L. Liu, and B. Cairns (2015),
844 Extension and statistical analysis of the GACP aerosol optical thickness record, *Atmos. Res.*,
845 164-165, 268–277, doi:10.1016/j.atmosres.2015.05.013.

846 Hasekamp, O. P., and J. Landgraf (2007), Retrieval of aerosol properties over land surfaces: ca-
847 pabilities of multi-viewing-angle intensity and polarization measurements, *Appl. Opt.*, 46(16),
848 3332–3344, doi:10.1364/AO.46.003332.

849 Holben, B. N., T. F. Eck, I. Slutsker, D. Tanré, J. P. Buis, A. Setzer, E. Vermote, J. A. Reagan,
850 Y. J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov (1998), AERONET:
851 A federated instrument network and data archive for aerosol characterization, *Remote Sens.*
852 *Environ.*, 66, 1–16, doi:10.1016/S0034-4257(98)00031-5.

853 Holben, B. N., D. Tanré, A. Smirnov, T. F. Eck, I. Slutsker, N. Abuhassan, W. W. Newcomb,
854 J. S. Schafer, B. Chatenet, F. Lavenu, Y. J. Kaufman, J. Vande Castle, A. Setzer, B. Markham,
855 D. Clark, R. Frouin, R. Halthore, A. Karneli, N. T. O’Neill, C. Pietras, R. T. Pinker, K. Voss,
856 and G. Zibordi (2001), An emerging ground-based aerosol climatology: Aerosol optical depth
857 from AERONET, *J. Geophys. Res.*, 106(D11), 12,067–12,097, doi:10.1029/2001JD900014.

858 Holben, B. N., et al. (2017), An overview of meso-scale aerosol processes, comparison and vali-
859 dation studies from DRAGON networks, *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2016-
860 1182.

- 861 Hsu, N. C., S.-C. Tsay, M. D. King, and J. R. Herman (2004), Aerosol properties over
862 bright-reflecting source regions, *IEEE Trans. Geosci. Remote Sens.*, *42*(3), 557–569, doi:
863 10.1109/TGRS.2004.824067.
- 864 Hsu, N. C., S.-C. Tsay, M. D. King, and J. R. Herman (2006), Deep Blue retrievals of Asian
865 aerosol properties during ACE-Asia, *IEEE Trans. Geosci. Remote Sens.*, *44*(11), 3180 – 3195,
866 doi:10.1109/TGRS.2006.879540.
- 867 Hsu, N. C., M.-J. Jeong, C. Bettenhausen, A. M. Sayer, R. Hansell, C. S. Seftor, J. Huang, and
868 S.-C. Tsay (2013), Enhanced Deep Blue aerosol retrieval algorithm: the second generation, *J.*
869 *Geophys. Res.*, *118*, 9296–9315, doi:10.1002/jgrd.50712.
- 870 Hsu, N. C., J. Lee, A. M. Sayer, N. Carletta, S.-H. Chen, C. J. Tucker, and S.-C. Tsay (2017),
871 Retrieving near-global aerosol loading over land and ocean from AVHRR, *J. Geophys. Res.*,
872 submitted.
- 873 Hyer, E. H., J. S. Reid, and J. Zhang (2011), An over-land aerosol optical depth data set for
874 data assimilation by filtering, correction, and aggregation of MODIS Collection 5 optical depth
875 retrievals, *Atmos. Meas. Tech.*, *4*, 379–408, doi:10.5194/amt-4-379-2011.
- 876 Jeong, M. -J. and Z. Li (2005), Quality, compatibility, and synergy analyses of global aerosol
877 products derived from the advanced very high resolution radiometer and Total Ozone Mapping
878 Spectrometer, *J. Geophys. Res.*, *110*, D10S08, doi:10.1029/2004JD004647.
- 879 Kahn, R. A., B. J. Gaitley, M. J. Garay, D. J. Diner, T. F. Eck, A. Smirnov, and B. N.
880 Holben (2010), Multiangle Imaging SpectroRadiometer global aerosol product assessment
881 by comparison with the Aerosol Robotic Network, *J. Geophys. Res.*, *115*(D23209), doi:
882 10.1029/2010JD014601.

- 883 Kahn, R. A., M. J. Garay, D. L. Nelson, R. C. Levy, M. A. Bull, D. J. Diner, J. V. Martonchik,
884 E. G. Hansen, L. A. Remer, and D. Tanré (2011), Response to ‘Toward unified satellite clima-
885 tology of aerosol properties: 3. MODIS versus MISR versus AERONET’, *J. Quant. Spectrosc.*
886 *Radiative Trans.*, *112*(5), 901–909, doi:10.1016/j.jqsrt.2010.11.001.
- 887 Knapp, K. R., and L. L. Stowe (2002), Evaluating the potential for retrieving aerosol optical
888 depth over land from AVHRR Pathfinder atmosphere data, *J. Atmos. Sci.*, *59*, 279–293, doi:
889 10.1175/1520-0469(2002)059<0279:ETPFRA>2.0.CO;2.
- 890 Knobelspiesse, K. D., C. Pietras, G. S. Fargion, M. Wang, R. Frouin, M. A. Miller, A. Subrama-
891 niam, and W. M. Balch (2004), Maritime aerosol optical thickness measured by handheld Sun
892 photometers, *Remote Sens. Environ.*, *93*(1–2), 87–106, doi:10.1016/j.rse.2004.06.018.
- 893 Lambert, A., R. G. Grainger, J. J. Remedio, C. D. Rodgers, M. Corney, and F. W. Taylor (1992),
894 Measurements of the evolution of the Mt. Pinatubo aerosol cloud by ISAMS, *Geophys. Res.*
895 *Lett.*, *20*(12), 1287–1290, doi:10.1029/93GL00827.
- 896 Lee, J., N. C. Hsu, A. M. Sayer, C. Bettenhausen, and P. Yang (2017), AERONET-based non-
897 spherical dust optical models and effects on the VIIRS Deep Blue/SOAR over-water aerosol
898 product, *J. Geophys. Res.*, submitted.
- 899 Levy, R. C., L. A. Remer, D. Tanré, Y. J. Kaufman, C. Ichoku, B. N. Holben, J. M. Livingston,
900 P. B. Russell, and H. Maring (2003), Evaluation of the Moderate-Resolution Imaging Spec-
901 troradiometer (MODIS) retrievals of dust aerosol over the ocean during PRIDE, *J. Geophys.*
902 *Res.*, *108*(D19), doi:10.1029/2002JD002460.
- 903 Levy, R. C., L. A. Remer, S. Mattoo, E. F. Vermote, and Y. J. Kaufman (2007), Second-
904 generation operational algorithm: Retrieval of aerosol properties over land from inversion
905 of Moderate Resolution Imaging Spectroradiometer spectral reflectance, *J. Geophys. Res.*,

- 906 112(D13211), doi:10.1029/2006JD007811.
- 907 Levy, R. C., L. A. Remer, R. G. Kleidman, S. Mattoo, C. Ichoku, R. Kahn, and T. F. Eck (2010),
908 Global evaluation of the Collection 5 MODIS dark-target aerosol products over land, *Atmos.*
909 *Chem. Phys.*, *10*, 103,999–10,420, doi:10.5194/acp-10-10399-2010.
- 910 Levy, R. C., S. Mattoo, L. A. Munchak, L. A. Remer, A. M. Sayer, F. Patadia, and N. C. Hsu
911 (2013), The Collection 6 MODIS aerosol products over land and ocean, *Atmos. Meas. Tech.*,
912 *6*, 2989–3034, doi:10.5194/amt-6-2989-2013.
- 913 Li, J., X. Zhao, R. Kahn, M. Mishchenko, L. Remer, K. -H. Lee, M. Wang, I. Laszlo, T. Nakajima,
914 and H. Maring (2009), Uncertainties in satellite remote sensing of aerosols and impact on
915 monitoring its long-term trend: a review and perspective, *Ann. Geophys.*, *27*, 2755–2770, doi:
916 10.5194/angeo-27-2755-2009.
- 917 Li, J., X. Li, B. E. Carlson, R. A. Kahn, A. A. Lacis, O. Dubovik, and T. Nakajima (2016), Re-
918 ducing multi-sensor satellite monthly mean aerosol optical depth uncertainty part I: Objective
919 assessment of current AERONET locations, *J. Geophys. Res.*, *121*, doi:10.1002/2016JD025469.
- 920 Liu, L., M. I. Mishchenko, I. G. Geogdzhayev, A. Smirnov, S. M. Sakerin, D. M. Ka-
921 banov, and O. A. Ershov (2004), Global validation of two-channel AVHRR aerosol optical
922 thickness retrievals over the oceans, *J. Quant. Spect. Rad. Trans.*, *88*(1-3), 97–109, doi:
923 10.1016/j.jqsrt.2004.03.031.
- 924 Lyapustin, A., Y. Wang, I. Laszlo, R. Kahn, S. Korkin, L. Remer, R. Levy, and J. S. Reid
925 (2011), Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm,
926 *J. Geophys. Res.*, *116*(D03211), doi:10.1029/2010JD014986.
- 927 Marlier, M. E., R. S. DeFries, A. Voulgarakis, P. L. Kinney, J. T. Randerson, D. T. Shindell,
928 Y. Chen, and G. Faluvegi (2013), El Niño and health risks from landscape fire emissions in

- 929 southeast Asia, *Nature Clim. Change*, *3*, 131–136, doi:10.1038/nclimate1658.
- 930 Mei, L. L., Y. Xue, A. A. Kokhanovsky, W. von Hoyningen-Huene, G. de Leeuw, and J. P.
931 Burrows (2014), Retrieval of aerosol optical depth over land surfaces from AVHRR data, *Atmos.*
932 *Meas. Tech.*, *7*, 2411–2420, doi:10.5194/amt-7-2411-2014.
- 933 Mishchenko, M. I., and I. V. Geogdzhayev (2007), Satellite remote sensing reveals regional tro-
934 pospheric aerosol trends, *Opt. Express*, *15*(12), 7423–7438, doi:10.1364/OE.15.007423.
- 935 Mishchenko, M. I., I. V. Geogdzhayev, B. Cairns, W. B. Rossow, and A. A. Lacis (1999), Aerosol
936 retrievals over the ocean by use of channels 1 and 2 AVHRR data: sensitivity analysis and
937 preliminary results, *Appl. Opt.*, *38*(36), 7325–7341, doi:10.1364/AO.38.007325.
- 938 Mishchenko, M. I., L. Liu, I. V. Geogdzhayev, J. Li, B. E. Carlson, A. A. Lacis, B. Cairns,
939 and L. D. Travis (2012), Aerosol retrievals from channel-1 and -2 AVHRR radiances: Long-
940 term trends updated and revisited, *J. Quant. Spect. Rad. Trans.*, *113*(15), 1974–1980, doi:
941 10.1016/j.jqsrt.2012.05.006.
- 942 O’Neill, N. T., T. F. Eck, A. Smirnov, B. N. Holben, and S. Thulasiraman (2003), Spectral
943 discrimination of coarse and fine mode optical depth, *J. Geophys. Res.*, *108*(D17), 4559–4573,
944 doi:10.1029/2002JD002975.
- 945 Popp, T., et al. (2016), Development, production and evaluation of aerosol climate data records
946 from European satellite observations (Aerosol_cci), *Remote Sens.*, *8*(5), doi:10.3390/rs8050421.
- 947 Reid, J. S., E. J. Hyer, R. S. Johnson, B. N. Holben, R. J. Yokelson, J. Zhang, J. R. Campbell,
948 S. A. Christopher, D. G. L., L. Giglio, R. E. Holz, C. Kearney, J. Miettinen, E. A. Reid, F. J.
949 Turk, J. Wang, P. Xian, G. Zhao, R. Balasubramanian, B. N. Chew, S. Janjai, N. Lagrosas,
950 P. Lestari, N. H. Lin, M. Mahmud, A. Nguyen, B. Norris, N. T. K. Oanh, M. Oo, S. V. Salinas,
951 E. J. Welton, and S. C. Liew (2013), Observing and understanding the Southeast Asian aerosols

- 952 system by remote sensing: an initial review and analysis for the Seven Southeast Asian Studies
953 (7 SEAS) program, *Atmos. Res.*, *122*, 303–468, doi:10.1016/j.atmosres.2012.06.005.
- 954 Riffler, M., C. Popp, A. Hauser, F. Fontana, and S. Wunderle (2010), Validation of a modified
955 AVHRR aerosol optical depth retrieval algorithm over Central Europe, *Atmos. Meas. Tech.*, *3*,
956 1255–1270, doi:10.5194/amt-3-1255-2010.
- 957 Sayer, A. M., G. E. Thomas, and R. G. Grainger (2010), A sea surface reflectance model
958 for (A)ATSR, and application to aerosol retrievals, *Atmos. Meas. Tech.*, *3*, 813–838, doi:
959 10.5194/amt-3-813-2010.
- 960 Sayer, A. M., N. C. Hsu, C. Bettenhausen, Z. Ahmad, B. N. Holben, A. Smirnov, G. E. Thomas,
961 and J. Zhang (2012a), SeaWiFS Ocean Aerosol Retrieval (SOAR): Algorithm, validation, and
962 comparison with other data sets, *J. Geophys. Res.*, *117*(D03206), doi:10.1029/2011JD016599.
- 963 Sayer, A. M., N. C. Hsu, C. Bettenhausen, M.-J. Jeong, B. N. Holben, and J. Zhang (2012b),
964 Global and regional evaluation of over-land spectral aerosol optical depth retrievals from Sea-
965 WiFS, *Atmos. Meas. Tech.*, *5*, 1761–1778, doi:10.5194/amt-5-1761-2012.
- 966 Sayer, A. M., A. Smirnov, N. C. Hsu, and B. N. Holben (2012c), A pure marine aerosol model, for
967 use in remote sensing applications, *J. Geophys. Res.*, *117*(D05213), doi:10.1029/2011JD016689.
- 968 Sayer, A. M., A. Smirnov, N. C. Hsu, L. A. Munchak, and B. N. Holben (2012d), Estimating
969 marine aerosol particle volume and number from Maritime Aerosol Network data, *Atmos.*
970 *Chem. Phys.*, *12*, 8889–8909, doi:10.5194/acp-12-8889-2012.
- 971 Sayer, A. M., N. C. Hsu, C. Bettenhausen, and M.-J. Jeong (2013), Validation and uncertainty
972 estimates for MODIS Collection 6 "Deep Blue" aerosol data, *J. Geophys. Res.*, *118*, 7864–7872,
973 doi:10.1002/jgrd.50600.

- 974 Sayer, A. M., L. A. Munchak, N. C. Hsu, R. C. Levy, C. Bettenhausen, and M.-J. Jeong (2014),
975 MODIS Collection 6 aerosol products: Comparison between Aqua’s e-Deep Blue, Dark Target,
976 and merged data sets, and usage recommendations, *J. Geophys. Res.*, *119*(1396513989), doi:
977 10.1002/2014JD022453.
- 978 Sayer, A. M., N. C. Hsu, C. Bettenhausen, R. E. Holz, J. Lee, G. Quinn, and P. Veglio (2017),
979 Cross-calibration of S-NPP VIIRS moderate-resolution reflective solar bands against MODIS
980 Aqua over dark water scenes, *Atmos. Meas. Tech.*, *10*, 1425–1444, doi:10.5194/amt-2016-238.
- 981 Schonbrödt, F. D., and M. Perugini (2013), At what sample size do correlations stabilize?, *J.*
982 *Res. Person.*, *47*(5), 609–612, doi:10.1016/j.jrp.2013.05.009.
- 983 Schutgens, N. A. J., M. Nakata, and T. Nakajima (2013), Validation and empirical correction of
984 MODIS AOT and AE over ocean, *Atmos. Meas. Tech.*, *6*, 2455–2475, doi:10.5194/amt-6-2455-
985 2013.
- 986 Seidel, F. C., and C. Popp (2012), Critical surface albedo and its implications to aerosol remote
987 sensing, *Atmos. Meas. Tech.*, *5*, 1653–1665, doi:10.5194/amt-5-1653-2012.
- 988 Smirnov, A., Y. Villevalde, N. T. O’Neill, A. Royer, and A. Tarussov (1995a), Aerosol optical
989 depth over the oceans: analysis in terms of synoptic air mass types, *J. Geophys. Res.*, *100*(D8),
990 16,639–16,650.
- 991 Smirnov, A., O. Yershov, and Y. Villevalde (1995b), Measurement of aerosol optical depth in
992 the Atlantic Ocean and Mediterranean Sea, in *Proceedings of SPIE*, *2582*, pp. 203–214, doi:
993 10.1117/12.228530.
- 994 Smirnov, A., B. N. Holben, T. F. Eck, O. Dubovik, and I. Slutsker (2000a), Cloud-screening
995 and quality control algorithms for the AERONET database, *Remote Sens. Environ.*, *73*(3),
996 337–349.

- 997 Smirnov, A., B. N. Holben, O. Dubovik, N. T. O'Neill, L. A. Remer, T. F. Eck, I. Slutsker,
998 and D. Savoie (2000b), Measurement of atmospheric optical properties on U.S. Atlantic coast
999 sites, ships, and Bermuda during TARFOX, *J. Geophys. Res.*, *105*(D8), 9887–9901, doi:
1000 10.1029/1999JD901067.
- 1001 Smirnov, A., B. N. Holben, Y. J. Kaufman, O. Dubovik, T. F. Eck, I. Slutsker, C. Pietras, and
1002 R. H. Halthore (2002), Optical properties of atmospheric aerosol in maritime environments, *J.*
1003 *Atmos. Sci.*, *59*, 501–523.
- 1004 Smirnov, A., B. N. Holben, I. Slutsker, D. M. Giles, C. R. McClain, T. F. Eck, S. M. Sakerin,
1005 A. Macke, P. Croot, G. Zibordi, P. K. Quinn, J. Sciare, S. Kinne, M. Harvey, T. J. Smyth,
1006 S. Piketh, T. Zielinski, A. Proshuninsky, J. I. Goes, N. B. Nelson, P. Larouche, V. F. Radionov,
1007 P. Goloub, K. K. Moorthy, R. Matarresse, E. J. Robertson, and F. Jourdin (2009), Maritime
1008 aerosol network as a component of aerosol robotic network, *J. Geophys. Res.*, *112*(D06204),
1009 doi:10.1029/2008JD011257.
- 1010 Smirnov, A., B. N. Holben, D. M. Giles, I. Slutsker, N. T. O'Neill, T. F. Eck, A. Macke, P. Croot,
1011 Y. Courcoux, S. M. Sakerin, T. J. Smyth, T. Zielinski, G. Zibordi, J. I. Goes, M. J. Harvey,
1012 P. K. Quinn, N. B. Nelson, V. F. Radionov, C. M. Duarte, R. Losno, J. Sciare, K. J. Voss,
1013 S. Kinne, N. R. Nalli, E. Joseph, K. Krishna Moorthy, D. S. Covert, S. K. Gulev, G. Milinevsky,
1014 P. Larouche, S. Belanger, E. Horne, M. Chin, L. A. Remer, R. A. Kahn, J. S. Reid, M. Schulz,
1015 C. L. Heald, J. Zhang, K. Lapina, R. G. Kleidman, J. Griesfeller, B. J. Gaitley, Q. Tan, and
1016 T. L. Diehl (2011), Maritime Aerosol Network as a component of AERONET-first results and
1017 comparison with global aerosol models and satellite retrievals, *Atmos. Meas. Tech.*, *4*, 583–597,
1018 doi:10.5194/amt-4-583-2011.

- 1019 Stowe, L., A. Ignatov, and R. Singh (1997), Development, validation, and potential enhancements
1020 to the second-generation operational aerosol product at NOAA/NESDIS, *J. Geophys. Res.*,
1021 *102*(D14), 16,923–16,934.
- 1022 Tange, O. (2011), GNU Parallel - the command-line power tool, *login: The USENIX Magazine*,
1023 pp. 42–47.
- 1024 Tanré, D., B. N. Holben, and Y. J. Kaufman (1992), Atmospheric correction algorithm for
1025 NOAA-AVHRR products: theory and application, *IEEE Trans. Geosci. Remote Sens.*, *30*(2),
1026 231–248, doi:10.1109/36.134074.
- 1027 Tanré, D., Y. J. Kaufman, M. Herman, and S. Mattoo (1997), Remote sensing of aerosol prop-
1028 erties over oceans using the MODIS/EOS spectral radiances, *J. Geophys. Res.*, *102*(D14),
1029 16,971–16,988, doi:10.1029/96JD03437.
- 1030 Toth, T. D., J. Zhang, J. R. Campbell, J. S. Reid, Y. Shi, R. S. Johnson, A. Smirnov, M. A.
1031 Vaughan, and D. M. Winker (2013), Investigating enhanced Aqua MODIS aerosol optical
1032 depth retrievals over the mid-to-high latitude Southern Oceans through intercomparison with
1033 co-located CALIOP, MAN, and AERONET data sets, *J. Geophys. Res. Atmos.*, *118*, 4700–
1034 4714, doi:10.1002/jgrd.50311.
- 1035 Twohy, C. H., J. A. Coakley Jr., and W. R. Tahnk (2009), Effect of changes in relative humidity
1036 on aerosol scattering near clouds, *J. Geophys. Res.*, *114*(D0505), doi:10.1029/2008JD010991.
- 1037 Várnai, T., A. Marshak, and W. Yang (2013), Multi-satellite aerosol observations in the vicinity
1038 of clouds, *Atmos. Chem. Phys.*, *13*, 3899–3908, doi:10.5194/acp-13-3899-2013.
- 1039 Vermote, E. and Y. J. Kaufman (1995), Absolute calibration of AVHRR visible and near-807
1040 infrared channels using ocean and cloud views, *Int. J. Remote Sens.*, *16*(13), 2317–2340, doi:
1041 10.1080/01431169508954561.

- 1042 Villevalde, Y. V., A. Smirnov, N. T. O'Neill, S. P. Smyshlyaev, and V. V. Yakovlev (1994),
1043 Measurements of aerosol optical depth in the Pacific Ocean and the North Atlantic, *J. Geophys.*
1044 *Res.*, *99*(D10), 20,983–20,988, doi:10.1029/94JD01618.
- 1045 von Hoyningen-Huene, W., J. Yoon, M. Vountas, L. G. Istomina, G. Rohen, T. Dinter, A. A.
1046 Kokhanovsky, and J. P. Burrows (2011), Retrieval of spectral aerosol optical thickness over
1047 land using ocean color sensors MERIS and SeaWiFS, *Atmos. Meas. Tech.*, *4*, 151–171, doi:
1048 10.5194/amt-4-151-2011.
- 1049 Wagner, F., and A. M. Silva (2008), Some considerations about Ångström exponent distributions,
1050 *Atmos. Chem Phys.*, *8*, 481–489, doi:10.5194/acp-8-481-2008.
- 1051 Zhang, J., and J. S. Reid (2006), MODIS aerosol product analysis for data assimilation: Assess-
1052 ment of over-ocean level 2 aerosol optical thickness retrievals, *J. Geophys. Res.*, *111*(D22207),
1053 doi:10.1029/2005JD006898.
- 1054 Zhao, T. X.-P., O. Dubovik, A. Smirnov, B. N. Holben, J. Sapper, C. Pietras, K. J. Voss,
1055 and R. Frouin (2004), Regional evaluation of an advanced very high resolution radiome-
1056 ter (AVHRR) two-channel aerosol retrieval algorithm, *J. Geophys. Res.*, *109*(D02204), doi:
1057 10.1029/2003JD003817.
- 1058 Zhao, T. X.-P., P. K. Chan, and A. K. Heidinger (2013), A global survey of the effect of
1059 cloud contamination on the aerosol optical thickness and its long-term trend derived from
1060 operational AVHRR satellite observations, *J. Geophys. Res. Atmos.*, *118*, 2849–2857, doi:
1061 10.1002/jgrd.50278.
- 1062 Zhao, X. (2016), Climate Data Record (CDR) Program Climate Algorithm Theoretical
1063 cal Basis Document (C-ATBD) AVHRR Aerosol Optical Thickness (AOT), *Tech.*
1064 *rep.*, NOAA, report number CDRP-ATBD-0096 revision 3, available online from

1065 <https://www.ncei.noaa.gov/data/avhrr-aerosol-optical-thickness/access/doc/> [Accessed
1066 March 7 2017].

1067 Zhao, X., A. K. Heidinger, and A. Walther (2016), Climatology analysis of aerosol effect on
1068 marine water cloud from long-term satellite climate data records, *Remote Sens.*, 8(4), 300,
1069 doi:10.3390/rs8040300.

Table 1. Statistics of validation between AVHRR and AERONET AOD measurements for SOAR over-water retrievals; n denotes the number of points, R Pearson’s correlation coefficient, f the fraction matching within the EE, and RMSE the root mean square error. The bias is the median AVHRR-AERONET bias. Statistics are given separately for 550 nm and AVHRR bands 1 and 2 (columns labelled 550, 630, 830 respectively).

Satellite	n	R			Bias			f			RMSE		
		550	630	830	550	630	830	550	630	830	550	630	830
NOAA14	1,227	0.92	0.94	0.94	0.022	0.027	0.030	0.64	0.64	0.61	0.071	0.065	0.064
NOAA18	13,412	0.86	0.88	0.90	0.0002	0.009	0.014	0.73	0.74	0.72	0.088	0.076	0.061

Table 2. As Table 1, except for the comparison between SOAR AVHRR retrievals and ship-based AOD measurements.

Satellite	n	R			Bias			f			RMSE		
		550	630	830	550	630	830	550	630	830	550	630	830
NOAA11	80	0.70	0.67	0.61	0.011	0.012	0.016	0.76	0.76	0.79	0.081	0.080	0.077
NOAA14	20	0.98	0.99	0.99	-0.017	-0.004	0.003	1.0	1.0	1.0	0.029	0.021	0.019
NOAA18	252	0.92	0.93	0.92	-0.019	-0.014	-0.007	0.79	0.82	0.83	0.071	0.056	0.047

Table 3. Statistics of validation between NOAA18 AVHRR and AERONET AOD measurements for DB over-land retrievals, globally and by region (as indicated in Figure 6). Statistics are defined as in Table 1, given separately for 550 nm and AVHRR band 1 (columns labelled 550 and 630 respectively).

Region	n	R		Bias		f		RMSE	
		550	630	550	630	550	630	550	630
Global	89,104	0.80	0.81	-0.014	-0.010	0.69	0.74	0.15	0.13
BOR	7,155	0.86	0.85	-0.016	-0.010	0.81	0.87	0.073	0.062
ENA	11,582	0.65	0.64	-0.010	-0.006	0.79	0.84	0.087	0.070
WNA	11,080	0.48	0.47	0.008	0.009	0.66	0.70	0.11	0.094
CSA	5,745	0.90	0.89	0.015	0.014	0.66	0.70	0.12	0.10
EUR	26,319	0.63	0.63	-0.018	-0.013	0.74	0.78	0.099	0.082
NAME	10,451	0.73	0.74	-0.052	-0.045	0.47	0.50	0.28	0.27
SA	2,277	0.70	0.67	-0.021	-0.015	0.72	0.77	0.12	0.098
IND	3,346	0.79	0.79	-0.058	-0.050	0.68	0.71	0.19	0.17
NEA	6,483	0.86	0.86	-0.039	-0.031	0.65	0.68	0.21	0.18
SEA	2,402	0.70	0.69	-0.046	-0.035	0.62	0.65	0.22	0.18
OCE	2,264	0.37	0.36	-0.002	-0.001	0.74	0.78	0.089	0.074

Table 4. As Table 3, except for NOAA14 AVHRR over land, and the OCE row is omitted due to a lack of sites in this region during the 1995-1999 time period.

Region	n	R		Bias		f		RMSE	
		550	630	550	630	550	630	550	630
Global	6,668	0.84	0.82	-0.010	-0.009	0.71	0.74	0.17	0.16
BOR	284	0.66	0.64	-0.004	-0.002	0.92	0.94	0.051	0.042
ENA	2,153	0.81	0.77	0.0003	-0.001	0.78	0.82	0.12	0.11
WNA	1,132	0.77	0.73	-0.002	-0.002	0.77	0.80	0.094	0.082
CSA	628	0.91	0.89	-0.041	-0.031	0.62	0.66	0.24	0.23
EUR	583	0.78	0.77	-0.033	-0.031	0.72	0.76	0.11	0.095
NAME	945	0.73	0.74	-0.067	-0.062	0.42	0.43	0.30	0.28
SA	622	0.86	0.83	-0.029	-0.026	0.78	0.82	0.13	0.11
IND	99	0.76	0.76	0.017	0.006	0.72	0.76	0.12	0.098
NEA	209	0.54	0.55	0.017	0.018	0.62	0.64	0.24	0.21
SEA	13	0.51	0.45	0.013	0.009	0.69	0.85	0.15	0.15

Table 5. Statistics of multi-sensor time series comparison for locations shown in Figures 13 and 14. Columns show the correlation coefficient R and median (AVHRR-other) bias at each location, separately for NOAA14 and NOAA18 AVHRR, between monthly mean 550 nm AOD.

Note the NOAA14 and MODIS Aqua time series do not overlap.

Statistic Comparison NOAA platform	R								Bias			
	AERONET		SeaWiFS		MODIS		AERONET		SeaWiFS		MODIS	
	14	18	14	18	14	18	14	18	14	18	14	18
Ocean sites												
Capo Verde	0.85	0.91	0.72	0.84	-	0.94	0.032	0.026	0.070	0.036	-	-0.036
Wallops	0.95	0.84	0.89	0.76	-	0.75	0.040	0.020	0.039	0.012	-	-0.006
Land sites												
Alta Floresta	0.91	0.99	0.98	0.95	-	0.96	-0.063	-0.029	-0.057	0.005	-	-0.016
GSFC	0.81	0.77	0.91	0.75	-	0.74	0.014	-0.009	0.033	-0.004	-	-0.028
Mongu	0.92	0.93	0.88	0.87	-	0.90	-0.006	-0.024	0.046	0.030	-	-0.032