



# Assessment of the Level-3 MODIS daily aerosol optical depth in the context of surface solar radiation and numerical weather modeling

J. A. Ruiz-Arias<sup>1,2,3</sup>, J. Dudhia<sup>3</sup>, C. A. Gueymard<sup>4</sup>, and D. Pozo-Vázquez<sup>1,2</sup>

<sup>1</sup>Solar Radiation and Atmosphere Modeling Group, Physics Department, University of Jaén, Jaén, Spain

<sup>2</sup>Center of Advanced Studies in Energy and Environment, University of Jaén, Jaén, Spain

<sup>3</sup>Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, Colorado, USA

<sup>4</sup>Solar Consulting Services, Colebrook, New Hampshire, USA

Correspondence to: J. A. Ruiz-Arias (jararias@ujaen.es)

Received: 15 June 2012 – Published in Atmos. Chem. Phys. Discuss.: 7 September 2012

Revised: 30 November 2012 – Accepted: 21 December 2012 – Published: 18 January 2013

**Abstract.** The daily Level-3 MODIS aerosol optical depth (AOD) product is a global daily spatial aggregation of the Level-2 MODIS AOD (10-km spatial resolution) into a regular grid with a resolution of  $1^\circ \times 1^\circ$ . It offers interesting characteristics for surface solar radiation and numerical weather modeling applications. However, most of the validation efforts so far have focused on Level-2 products and only rarely on Level 3. In this contribution, we compare the Level-3 Collection 5.1 MODIS AOD dataset from the Terra satellite available since 2000 against observed daily AOD values at 550 nm from more than 500 AERONET ground stations around the globe. Overall, the mean error of the dataset is 0.03 (17 %, relative to the mean ground-observed AOD), with a root mean square error of 0.14 (73 %, relative to the same), but these errors are also found highly dependent on geographical region. We propose new functions for the expected error of the Level-3 AOD, as well as for both its mean error and its standard deviation. Additionally, we investigate the role of pixel count vis-à-vis the reliability of the AOD estimates, and also explore to what extent the spatial aggregation from Level 2 to Level 3 influences the total uncertainty in the Level-3 AOD. Finally, we use a radiative transfer model to investigate how the Level-3 AOD uncertainty propagates into the calculated direct normal and global horizontal irradiances.

## 1 Introduction

Aerosols play a major role in the earth's energy budget. Predicting the climate system invariably requires an evaluation of the aerosol's radiative forcing, and this, in turn, requires knowledge of the spatial distribution, temporal evolution and optical properties of atmospheric aerosols (Kiehl and Ramanathan, 2006). Currently, the scientific understanding of the aerosols effect on climate is limited (IPCC, 2007), and the climatic perturbations they produce are a source of major uncertainty.

Aerosols are mainly generated at sources near the ground but also in the atmosphere (by clouds and gas-to-particle conversion). A detailed description of the different emitting sources of aerosols can be found in Jaenicke (1993). Near the ground, aerosols are advected into the lower layers of the atmosphere from many different sources such as biogenic emissions, oceans, deserts or human activities. The largest concentrations generally occur in urban and desert areas. These concentrations rapidly decrease with height in the troposphere (Liou, 2002) and can experience abrupt changes in their spatial distribution. Due to their short tropospheric residence times, aerosols generated near the ground mainly act at regional or local scales, albeit their influence can be also extended to broader scales when they are advected into upper layers of the atmosphere (e.g. Chen et al., 2010).

The mesoscale variations of tropospheric aerosols have been studied by Anderson et al. (2003). Based on aerosol optical data, they showed that mesoscale variability (specifically, for horizontal scales of 40–400 km and temporal

scales of 2–48 h) is a common and perhaps universal feature of lower-tropospheric aerosol light extinction. Very recently, based on empirical measurements gathered from the AERONET network (AEosol RObotic NETwork; Holben et al., 1998), Gueymard (2012) showed that, on average, monthly aerosol optical depth data captures only about half of the day-to-day variability. The use of monthly data, which is a common practice, is becoming a limiting factor for many current solar applications, such as solar energy utilization in power plants, which require an accurate determination of the available solar resource for optimal sizing of the production facilities, annual production estimates, and plant operation. Moreover, most current research in solar energy focuses on concentrating technologies (both concentrating photovoltaics and concentrating solar thermal). These rely completely on direct normal irradiance (DNI), which is highly sensitive to aerosol optical depth (AOD). In particular, some studies have concluded that both the demonstrated large inter-annual variability in DNI (Lohmann et al., 2006; Pozo-Vázquez et al., 2011) and the large differences between the current DNI databases (Šúri et al., 2009) can be partially explained by a misrepresentation of the aerosol optical properties (Cebecauer et al., 2011; Gueymard, 2011), particularly when they are represented by monthly climatologies. In fact, Gueymard (2012) also showed that to correctly reproduce the daily variations in DNI over areas strongly impacted by desert dust or smoke, where the induced effect on the inter-annual variability in DNI is estimated to reach 10–30 %, irradiance predictions with radiative models should rely on daily AOD rather than on the more usual monthly-averaged AOD data.

Kaufman et al. (2000) used 7 yr (1993–1999) of AOD data from 50–70 AERONET ground stations to check to what extent measurements acquired at specific times of the day (such as current measurements from satellites Terra and Aqua) could represent the daily average aerosol forcing of climate. For all these stations, they calculated the ratio of the averaged AERONET AOD for the virtual overpass of the satellites (10:00–11:30 LT for Terra's overpass, and 12:00–13:30 LT for Aqua's overpass) to the daily averaged AERONET AOD. They concluded that measurements at these specific times of the day could represent the annual average value within 2 %, independently of the particle size and range of AOD.

In the last few years, the number and time span of the available worldwide aerosol databases have considerably increased, and rapid progress has been made in aerosol science. These developments have been boosted mostly by global ground measurement networks, such as AERONET, and by satellite sensors, such as the MODerate-resolution Imaging Spectroradiometer (MODIS) or the Multi-angle Imaging SpectroRadiometer (MISR), aboard the Terra and Aqua satellites. The length of the MODIS database and the number of AERONET stations worldwide make it possible to address a similar question to that of Kaufman et al. (2000) by a direct validation of the daily MODIS AOD retrievals using ground measurements.

In this contribution, we present a global validation of the combined land and ocean Level-3 (L3) MODIS AOD product (Collection 5.1) at 550 nm from the Terra satellite. Small differences are expected with respect to the Aqua satellite (Levy et al., 2010). The motivations of this study will become clearer in Sect. 1.2 after a brief introduction to the MODIS dataset in Sect. 1.1. The dataset evaluated here spans all the available daily AOD data from the Terra launch (February 2000) to December 2011. For ground truth, we use the complete dataset of daily AOD observations from all possible stations of the AERONET network (Level 2.0; see Sect. 2.2) that have reported data during the 12-yr period of the satellite observations. The dataset thus encompasses a total of 525 stations, but the number of available stations operating any given day is variable.

### 1.1 MODIS aerosol optical depth

The L2 MODIS dataset is a collection of aerosol optical properties, particularly including the spectral AOD at 550 nm, as well as particle size information. The aerosol is evaluated at the time of the satellite overpass with an un-gridded 10-km nominal resolution. It is derived from MODIS radiances using either one of three different algorithms: one over ocean (Remer et al., 2005) and two over land, the Dark-Target (Levy et al., 2007) and the Deep-Blue (Hsu et al., 2004) algorithms – the latter for brighter surfaces. MODIS data are divided into 5-min granules with a swath of  $\approx 2330$  km. The L2 MODIS product is provided with a quality assurance system that flags those pixels that reach exceptional values or reach some error threshold in the intermediate steps of the retrieval algorithm. Lastly, these flags are summarized into a quality assurance confidence (QA) flag, with values from 0 (no-confidence) to 3 (best quality). Prior to the Terra and Aqua satellites launch, some sensitivity studies (Kaufman et al., 1997; Tanré et al., 1997) estimated the expected error of MODIS AOD retrievals as  $\pm(0.03 + 0.05\tau)$  over ocean and  $\pm(0.05 + 0.20\tau)$  over land, where  $\tau$  is the AOD. (By  $\tau$  and AOD we mean the AOD at 550 nm, unless otherwise stated). Subsequently, Chu et al. (2002) suggested that the expected error over land could be reduced to  $\pm(0.05 + 0.15\tau)$ . This has been globally confirmed by Levy et al. (2010) for the Collection 5 and the Dark-Target algorithm.

The L3 MODIS product (Hubanks et al., 2008) is a global daily aggregation of L2 MODIS data into an equal-angle latitude-longitude grid with a spatial resolution of  $1^\circ \times 1^\circ$ . Each L3 MODIS grid cell value is calculated from the L2 MODIS pixels that fall into its cell boundaries, the number of which is known as pixel count. (Note that several  $1^\circ \times 1^\circ$  cells are covered by a single L2 MODIS swath.) The actual area of each  $1^\circ \times 1^\circ$  grid cell shrinks when moving from the equator ( $\approx 12\,321$  km<sup>2</sup>) to the poles ( $\approx 107$  km<sup>2</sup>). Besides, there is a change in the size of the L2 MODIS pixels due to the viewing distortion at high sensor scan angles. As a result,

there are fewer pixels to include in the computation of each L3 MODIS cell over regions swept with high sensor scan angles and/or near the poles. In addition, mostly due to overlapping orbits toward the poles, but also to viewing distortion at high sensor scan angles, the L3 MODIS values poleward of  $23^\circ$  actually become a time average. Between approximately  $23^\circ$  S and  $23^\circ$  N, the reported values are rather a spatial average of all L2 MODIS values obtained during a single satellite overpass every day. Finally, we want to note that there is no minimum pixel count to derive the L3 MODIS product statistics. Typically, for the aerosol products considered here, the pixel count varies from 0 to 121, although even more pixels may be eventually used.

## 1.2 Motivations and objectives

The relatively high spatio-temporal resolution, and the large (near-global) coverage of the L3 MODIS dataset make it especially suitable for use with Numerical Weather Prediction (NWP) models (Lara-Fanego et al., 2011; Ruiz-Arias et al., 2010, 2011, 2012). We will focus here only on AOD because it is the most important aerosol optical property driving solar extinction, and thus the incident surface shortwave irradiance – most importantly the direct irradiance. In particular, at the end of this paper we analyze how the uncertainty in the L3 MODIS AOD estimates propagates into the uncertainty in modeled values of GHI and DNI. To this end, we conduct an initial global evaluation of the L3 MODIS AOD uncertainty based on ground measurements. The Angström exponent, an important parameter in determining the AOD spectral dependence, is also provided within the L3 MODIS dataset. However, it is not considered here, since its use has been discouraged by Levy et al. (2010).

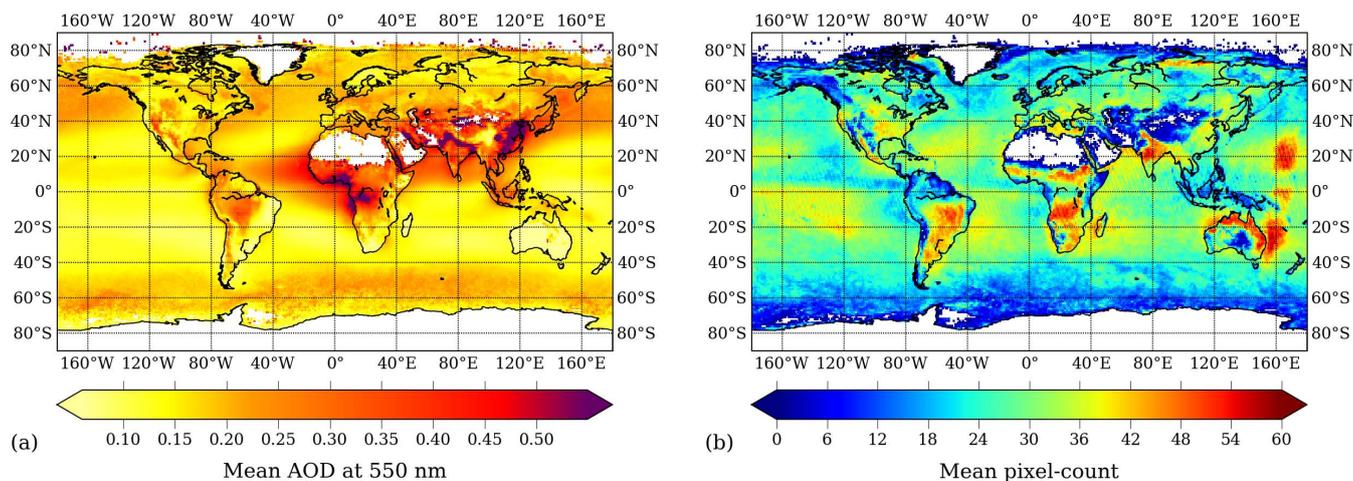
So far, most of the MODIS validation exercises only involved the L2 MODIS dataset (Mishchenko et al., 2008; Remer et al., 2008; Hoelzemann et al., 2009; Jethva et al.; Levy et al., 2010; Li et al., 2010; Mei et al., 2012, to mention just a few). But the L3 MODIS dataset has a number of peculiarities that differentiate it from the L2 MODIS dataset. Usually, L2 MODIS AOD (hereafter, also referred to as L2 AOD) validation studies employ the Ichoku approach (Ichoku et al., 2002) to verify collocation, i.e. when a ground measurement matches a satellite observation. The Ichoku approach compares the spatial average of a grid of  $5 \times 5$  L2 MODIS pixels (in the middle of which the ground station stands) to the temporal average of the data collected by the ground station within  $\pm 30$  min of the satellite overpass. At least 5 L2 MODIS pixels and 2 ground observations are required to accept the collocated data pair. This approach intrinsically assumes a notable time-space correlation between hourly temporal averages and the spatial averages in a region of roughly  $50 \times 50$  km. Kaufman et al. (2000) showed that a single measurement of AOD every day can be used to approximate the annual average daily AOD value. Since the L3 MODIS product is updated daily, it makes sense to validate the L3 MODIS

AOD dataset (hereafter, also referred to as L3 AOD) using daily AOD averages acquired at ground stations. In such a case, the validation time window widens to the daylight hours (and cloudless conditions) relative to the ground station, and the spatial average spreads over a region of  $1^\circ \times 1^\circ$ , i.e., roughly 4 times coarser than for the validation of the L2 AOD product following the Ichoku approach. It is important to know to what extent this broadened time-space correlation assumption affects the performance of the L3 AOD data, and whether the expected error (EE) of the L2 AOD also holds for the L3 AOD data.

In this respect, it is also important to note that the use of point-wise observations may not be the ideal validation reference for the  $1^\circ \times 1^\circ$  L3 AOD values. Strictly speaking, the results of this validation should not be categorized as “errors” because they also encompass differences regarding the coarser spatial aggregation of the L3 AOD, thus neglecting the natural spatial variability of AOD at smaller scales than  $1^\circ \times 1^\circ$ . However, from a large-scale application standpoint, the validation of the L3 AOD product with point-wise observations is sufficiently informative and convenient, and is the only practical global validation that can be done anyway. Therefore, although for the sake of language clarity we refer to the result of the validation in the present framework as “error” (including the term “expected error”), we explicitly acknowledge that it is only a first guess or proxy to the actual L3 AOD error. Section 5 presents an extended discussion on this topic.

We also explore the influence of the L2 AOD pixel counts on the L3 AOD reliability and address the relative importance of the spatial uncertainty caused by the  $1^\circ \times 1^\circ$  AOD sub-pixel variability, also known as spatial representativeness uncertainty. More specifically, the question here is how the spatial representativeness uncertainty compares to the total L3 AOD error. The outcome of this analysis is important because it might suggest the use of an increased spatial resolution in future releases of the L3 MODIS AOD product. Based on the results obtained in the validation exercise, we finally explore how the L3 AOD uncertainty propagates into the irradiance prediction (of GHI and DNI). This result is important to determine the usability and possible benefits of the L3 AOD dataset for surface solar resource assessments, in particular.

In Sect. 2, we describe the L3 MODIS AOD and observational datasets used to conduct this study, as well as the collocation approach to match pairs of data. We also present a brief exploratory analysis of the datasets. Section 3 shows the results of the validation, both globally and constrained by geographical region. In Sect. 3.3 we provide a simple model to estimate the mean error and uncertainty of the L3 AOD measurements. Section 4 investigates the role of pixel count in the reliability of the L3 AOD, whereas Sect. 5 focuses on the contribution of the AOD variability at sub-grid scale to the L3 AOD uncertainty. Finally, Sect. 6 presents a study on the



**Fig. 1.** (a) Mean AOD over the entire validation period; (b) mean pixel count. White areas are the regions without data.

propagation of the L3 AOD uncertainty into the surface solar irradiance. The main conclusions are summarized in Sect. 7.

## 2 Data

Here, we describe both the L3 MODIS and the AERONET AOD datasets used in the study. Additionally, we offer a brief description of the collocation approach and some comments on the spatial distribution of the experimental stations used from the AERONET network.

### 2.1 Level-3 MODIS AOD dataset

This validation covers the period since the deployment of the MODIS instrument aboard Terra satellite (February 2000) to the end of Dec. 2011. The dataset belongs to the Collection 5.1 combined land and ocean L3 MODIS AOD at 550 nm from the Terra satellite (variable name “Optical\_Depth\_Land\_And\_Ocean\_Mean”, available from [ftp://ladsweb.nascom.nasa.gov/allData/51/MOD08\\_D3/](ftp://ladsweb.nascom.nasa.gov/allData/51/MOD08_D3/)). The documentation states that the dataset was derived only from the best quality data (flagged as QA = 3 in the MODIS quality assurance system) for both ocean and land data. However, we were noticed by one reviewer during the revision of the manuscript that there is in fact an error in the documentation. This product was apparently derived including also data with QA = 1 and 2 over land and QA = 0, 1 and 2 over ocean. Nonetheless, a noticeable difference in the results with respect to the use of only QA = 3 data is unlikely.

Overall, the dataset spans 12 yr, totaling data over 4209 days. On average, about 33 % of the earth’s surface is tiled every day with AOD measurements. But if the surface is restricted to only latitudes between 60° S and 60° N (higher latitudes are typically covered with snow and thus prone to missing values), the dataset covers about 46 % of the sur-

face. Missing tiles are mostly caused by clouds and high reflectance surfaces, such as deserts or snow.

Figure 1 shows the mean AOD value and the mean pixel count at each grid cell over the entire validation period. The mean AOD over the whole grid is 0.18 and the standard deviation is 0.09 (i.e., 50 % relative to the mean value). There exists a large difference between the minimum and maximum mean AOD,  $-0.05$  and  $2.78$ , respectively (negative AOD values are non physical for solar radiation applications, and therefore are not considered further). Interestingly, 95 % of all the AOD values are smaller than 0.35, due to the generally low aerosol burden over oceans. The remaining 5 % of all cases, where AOD can be considerably larger than 0.5, mostly corresponds to land areas over Asia, north or central Africa, and eastern South America. In Asia, AOD is particularly high in eastern China (presumably due to industrial aerosols in large part), the Hymalayas and the Indus river basin. In Africa, the mean AOD is specially high in the central rain-forests and the Gulf of Guinea, due to the seasonal burning of biomass, and the Sahel area, due to the dust drag from the desert. Figure 1a shows a deep intrusion of African aerosol into the central Atlantic Ocean that extends up to the Caribbean Sea. Seasonally high values of mean AOD are also found in the Amazonian region, due to biomass burning. Finally, a hot spot appears over the southwestern USA. The high AOD values there completely disagree with the low values of ground truth, as will be discussed in more details in Sect. 3. The incorrect MODIS retrievals in this region are likely due to a misrepresentation of the land surface and/or assumed aerosol properties in the Dark-Target algorithm (Levy et al., 2010).

Figure 1b shows the mean L2 pixel count used in the computation of the L3 AOD values. The mean pixel count over the entire grid is 24.9. On average, 20 % of the L3 AOD values were calculated using no more than 15 pixel counts, and 90 % of them were derived using less than 35 pixel counts. At

very high latitudes (poleward of about  $\pm 60^\circ$ ) the pixel-count number is generally less than 10. A very reduced number of L2 pixels is also found over highly reflective surfaces such as deserts and mountains (for instance, central Australia, central Asia, Middle-East or Sahara) or regions with high cloudiness (northern South-America, Alaska or British Isles). The regions with the highest pixel count are mainly distributed along the equatorial belt area and the Southern Hemisphere.

## 2.2 Observational dataset: AERONET network

The ground truth dataset is part of the AERONET federated network (<http://aeronet.gsfc.nasa.gov>). AERONET provides optical and microphysical properties of aerosols at high temporal resolution and three data quality levels: Level 1.0 (raw, unscreened), Level 1.5 (cloud-screened), and Level 2.0 (cloud screened and quality-assured). In this work we have used all available Version 2.0, Level 2.0 daily AOD data covering the same period as the L3 AOD dataset. Daily averages are calculated only when at least three (cloudless) measurements are available during a day. Direct-sun AOD retrievals are done routinely for (normally) eight different spectral bands, but none of them is centered at 550 nm. Therefore, the AOD at 550 nm is calculated here from the AOD at 675 nm and the Angström exponent derived from the 440 and 870 nm channels. The AOD uncertainty is about 0.01–0.02 (Eck et al., 1999). The collocation between one ground measurement and satellite is assumed to occur only when the ground measurement is within the boundaries of a satellite grid cell that has a non-missing value. In the rare cases where two or more AERONET sites exist inside the same MODIS grid cell, all the ground observations are considered independently.

The total number of stations with at least one collocation during the analyzed period is 525. The availability of ground-truth data changes every day for several reasons, such as cloud-screening, temporary campaigns, experimental problems, or regular off-site calibration. As a result, the average number of collocations per day is 65.8; 90 % of the days have at least 34 collocations, and 44.6 % of the 525 stations have at least one year of collocations.

Many validation studies have found regional differences in the performance of the L2 AOD (e.g., Levy et al., 2010). Consequently, the results of any validation will be invariably bound to the spatial distribution of the experimental stations. In this dataset, up to 290 stations are located in North America and Europe, whereas South America, Africa and Asia have only a total of 164 stations. The remainder (71 stations) is located in Australia and small islands. These regional differences mean that our global results will be more representative of the North American and European regions. However, we have also conducted a regional validation in order to achieve meaningful results for each region (Sect. 3.2). Overall, the use of the entire dataset of available AERONET sta-

tions during the complete period of L3 AOD data from Terra guarantees the maximum possible significance of the results.

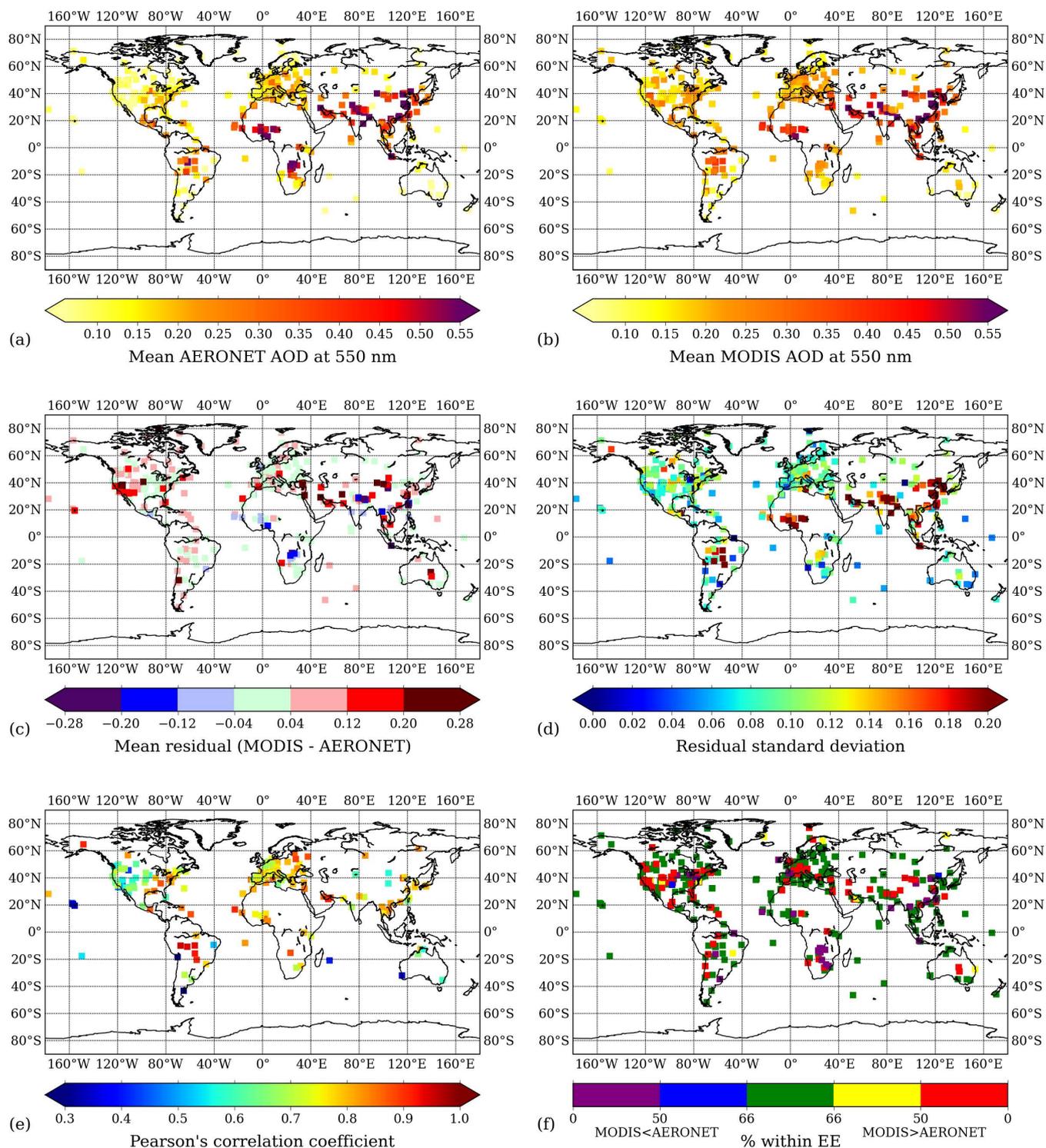
## 3 Comparison against AERONET observations

Figure 2 shows some statistical scores calculated from all the AERONET and L3 AOD collocations during the entire study period. Figure 2a and b show that the mean AERONET and L3 AOD values at 550 nm, respectively, have very similar spatial patterns in general. The most striking difference occurs over the southwestern US, where L3 AOD mean values are significantly larger than AERONET's, as mentioned earlier. Figure 2c and d, respectively, show the residual (L3 AOD minus AERONET) and the residual standard deviation (STDe). Overall, most of the European, African and eastern US locations have a bias below 0.12 ( $\approx 20$ –30 %). MODIS overestimates over the southwestern US, Middle East and some central and coastal Asian locations. The overestimation is high in relative terms ( $\approx 100$  %, not shown here) over the southwestern US, since the measured AOD in this region is very small (0.10–0.15). As suggested by Levy et al. (2010), this could result from a misrepresentation of the surface reflectance. The overestimation over the Middle East and Asia is more likely related to deficiencies in the aerosol model, as also reported by Levy et al. (2010). Some occasional overestimation is found over other regions as well. For instance, over Hawaii and some other coastal areas, some representativeness problems of the local topography at the (relatively coarse) L3 AOD spatial resolution might exist. In Asia and southern Africa, there are areas with large underestimations, up to  $\approx 50$  %.

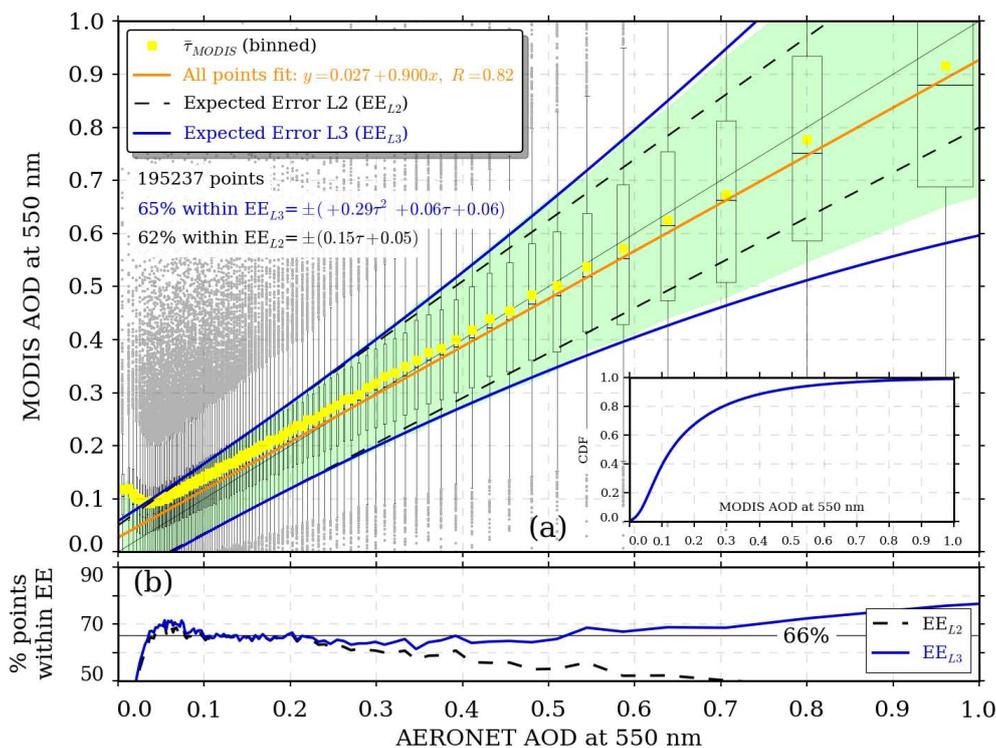
The regions with higher AOD (Asia, central Africa and South America) are also those with higher STDe (above 0.16). This can be explained by the known fact that MODIS estimates do not cover all the AOD dynamic range for high values. However, when STDe is normalized by the mean AERONET AOD value (not shown here), the result is about 40–50 %, which is not the highest relative STDe. The highest values are actually reached over the southwestern US and Australia, where they are close to 100 % because of the very small measured AOD. Figure 2e shows the Pearson's correlation coefficient for all those locations with more than 365 collocations. Overall, the correlation between AERONET sites and the L3 AOD is greater than 0.7, except over some oceanic and coastal locations, central Asia and the western US, where it is less. Interestingly, sites with high AOD and STDe (Fig. 2b and d) have also a high correlation coefficient, which means that changes in AOD are concomitant in sign but not in magnitude.

### 3.1 Expected error in the L3 MODIS AOD dataset

In the literature so far, L2 MODIS AOD estimates have been typically evaluated in relationship with the expected error



**Fig. 2.** Statistical scores calculated at experimental sites from the AERONET and L3 MODIS AOD collocations over the entire study period: (a) mean AERONET AOD at 550 nm; (b) mean L3 AOD at 550 nm; (c) mean residual; (d) standard deviation of the residual; (e) Pearson's correlation coefficient; and (f) frequency of data points within the L2 AOD expected error (EE) interval.



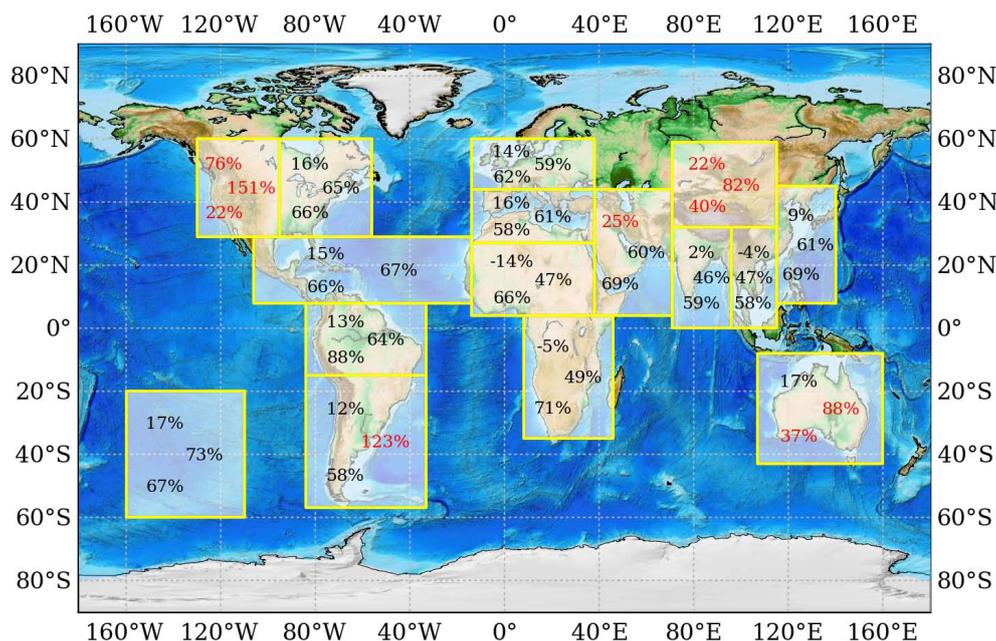
**Fig. 3.** (a) L3 AOD at 550 nm against AERONET AOD at 550 nm. Data are binned by AERONET AOD value in 100 intervals with approximately the same number of samples. The vertical limits of the boxes extend from the lower to the upper quartiles of every bin. The median is indicated by the horizontal line in the middle of the box. The yellow square is the mean L3 AOD value. The whiskers extend to the most extreme point within the (75–25 %) data range. The points beyond the whiskers limits are plotted in gray. The horizontal position of the boxes and their width are the AERONET mean and standard deviation values, respectively. The orange line is the all-point fit. The green shaded area delimits the region containing 66 % of the L3 AOD measurements, and the black dashed line is the L2 AOD EE. Blue lines are the second-order polynomial fit to the green shaded area boundaries. The inset figure provides the cumulative distribution function of the L3 AOD dataset. (b) Rate of points within the L2 (dashed black line) and L3 (solid blue line) AOD expected error intervals.

(EE) defined as  $\pm(0.05 + 0.15\tau)$  where  $\tau$  is the true AOD value. If at least 66 % of the MODIS AOD values are concentrated within the interval defined above by EE, the MODIS retrieval algorithm is considered to behave well. The question is whether this EE still holds for the L3 AOD dataset. Figure 2f shows the rate of points within the L2 AOD EE interval at each experimental site, with a distinction between overestimation and underestimation. Overall, there are many stations over the entire dataset where the L2 AOD EE contains even less than 50 % of the data, usually where MODIS overestimates. This result justifies the necessity of a revision of EE for the L3 AOD dataset.

Many results are summarized in Fig. 3a. In particular, it highlights the interval (green shaded area) that contains 66 % of the L3 AOD estimates as a function of the AERONET AOD for the entire study period and all experimental stations. The L3 AOD values were binned into 100 AERONET AOD intervals with an approximately equal number of samples (about 1950 collocations per bin). For each interval, Fig. 3a shows a box representing some exploratory scores of the data: the vertical limits of the box extend from the lower

to the upper quartiles, the median is indicated by the horizontal line in the middle of the box, and the yellow square is the mean L3 AOD value in the interval. The whiskers extend to the most extreme point within the (75–25 %) data range. That is, the whiskers are as long as the box for each bin. The horizontal position of the box and its width are the AERONET mean and standard deviation, respectively. By comparing the position of the mean satellite values (yellow squares) against the 1:1 line, it is concluded that, on average, the L3 AOD retrievals tend to overestimate the AERONET measurements for a ground-truth AOD below 0.4 and underestimate above 0.5. This effect is also apparent in the all-point fit (orange line). Additionally, since the (theoretical) dynamic range of AOD only sweeps positive values, the spread of the data increases with the mean value for both the MODIS and AERONET measurements. This is shown by the increase of the length of the whiskers and the width of the boxes.

In Fig. 3a, the black dashed line is the EE for the L2 AOD dataset. If the EE limits fitted the boundaries of the green shaded (L3) region, the L2 AOD EE would also hold for the



**Fig. 4.** MBE (top), RMSE (middle) and squared Pearson correlation coefficient  $R^2$  (bottom) for each individual region over the entire study period. The entire dataset scores are shown in the square domain over the South Pacific Ocean.

L3 AOD dataset. From Fig. 3a, it is clear that the L2 AOD EE indeed fits the L3 AOD EE interval only for AOD values below about 0.3–0.4. For AODs greater than 0.3 the interval containing 66 % of the L3 AOD measurements becomes larger than what the L2 AOD EE predicts, meaning that the L3 AOD product is more uncertain than the L2 AOD product. This is also apparent in Fig. 3b, which plots the amount of points (as a percentage) within the EE limits. The reason for the uncertainty increase at large AODs may be found in the fact that the averaging process involved to obtain the L3 data from the L2 dataset actually removes some of the L2 spatio-temporal variability. This is particularly important for large AODs because the departures from the mean value are greater, and so are the errors caused by the averaging process. Nevertheless, according to the inset plot in Fig. 3a, up to 80 % of the data points are below an AOD of 0.3, for which the L2 AOD EE is still valid.

The observed increased uncertainty in the L3 dataset for large AOD values cannot be accommodated with linear EE limits over the whole AOD range. Instead, we propose a quadratic model that fits the L3 data better overall (in the least squares sense):  $\pm(0.29\tau^2+0.06\tau+0.06)$ , which is plotted as the blue line in Fig. 3a. Note that the L2 AOD EE and the L3 AOD EE match well for small AODs. In contrast, the L3 AOD EE follows the actual EE much better for large AODs. According to Fig. 3b, the new quadratic EE described above gives a reasonably good estimate of EE for AODs below  $\approx 0.7$ . This includes more than a 95 % of the dataset (see inset of Fig. 3a).

### 3.2 Regional analysis

It is instructive to study the behavior of the L3 AOD dataset over different regions, so as to provide a more detailed and organized general description than what is summarized in Fig. 2. In this respect, the experimental dataset has been divided into 15 different geographical regions. These are based on previous analyses and global studies, such as those of Remer et al. (2008) and Levy et al. (2010), and are depicted in Fig. 4. For each region, Fig. 4 provides (from top to bottom) the mean bias error (MBE, obtained as L3 AOD minus AERONET), the root mean square error (RMSE), both normalized to the averaged measured value, and the squared Pearson correlation coefficient ( $R^2$ ) calculated using all the stations in each region over the entire study period. Poor results (MBE values above 20 %, RMSE values above 80 %, and  $R^2$  below 50 %) are highlighted in red. Western North America, Central Asia and Australia show such poor results. Particularly, western North America has the poorest performance, with a MBE of 76 %, a RMSE of 151 %, and a  $R^2$  of only 22 %. For all the other regions, the MBE is below 17 %, except in the Middle East, where it reaches 25 %; the RMSE remains in the range 46–67 %, except in South America (123 %), whereas  $R^2$  is always around or above 60 %.

Table 1 provides some additional statistics: number of experimental points ( $N$ ), MBE and RMSE (AOD unit), as well as the rate of samples within the EE intervals of the L2 and L3 AOD datasets. The last column displays the quadratic expression obtained for the L3 AOD EE limits of each region. Overall, the MBE is always below 0.1 (in absolute terms),

**Table 1.** Summary statistics for the entire dataset and each individual region: number of samples ( $N$ ), mean AERONET AOD, MBE and RMSE in AOD unit, fraction of data points within the L2 and L3 AOD expected error intervals (L2 EE and L3 EE, respectively), and L3 AOD expected error limits function.

Region	$N$	Mean		RMSE	L2 EE	L3 EE	L3 EE limits
		AERONET	MBE				
Globe	195 238	0.20	0.03	0.14	62 %	65 %	$\pm(+0.29\tau^2 + 0.06\tau + 0.06)$
West North America	29 451	0.10	0.07	0.14	51 %	66 %	$\pm(+0.98\tau^2 - 0.23\tau + 0.11)$
East North America	24 767	0.17	0.03	0.11	70 %	66 %	$\pm(-0.05\tau^2 + 0.24\tau + 0.03)$
Tropical North Atlantic	12 268	0.21	0.03	0.14	55 %	67 %	$\pm(+0.11\tau^2 + 0.14\tau + 0.07)$
Amazonas	3966	0.34	0.04	0.21	66 %	67 %	$\pm(+0.08\tau^2 + 0.13\tau + 0.05)$
South America	7014	0.15	0.02	0.18	63 %	68 %	$\pm(+0.22\tau^2 + 0.15\tau + 0.05)$
North Europe	25 886	0.20	0.03	0.11	68 %	66 %	$\pm(-0.00\tau^2 + 0.21\tau + 0.04)$
Mediterranean	35 477	0.17	0.03	0.10	70 %	66 %	$\pm(+0.01\tau^2 + 0.20\tau + 0.04)$
Saharan Africa	4304	0.53	-0.07	0.24	47 %	65 %	$\pm(-0.02\tau^2 + 0.35\tau + 0.02)$
South Africa	6358	0.20	-0.01	0.10	72 %	66 %	$\pm(-0.04\tau^2 + 0.25\tau + 0.03)$
Middle East	4057	0.36	0.09	0.22	64 %	66 %	$\pm(+0.42\tau^2 - 0.07\tau + 0.07)$
Central Asia	2960	0.19	0.04	0.16	49 %	66 %	$\pm(+0.43\tau^2 - 0.09\tau + 0.11)$
East Asia	11 706	0.41	0.04	0.25	53 %	65 %	$\pm(+0.01\tau^2 + 0.31\tau + 0.04)$
South Asia	4728	0.42	-0.02	0.20	49 %	66 %	$\pm(+0.17\tau^2 + 0.03\tau + 0.11)$
West Asia	5162	0.46	0.01	0.21	56 %	66 %	$\pm(+0.27\tau^2 - 0.04\tau + 0.10)$
Australia	4643	0.10	0.02	0.09	72 %	66 %	$\pm(+0.24\tau^2 + 0.23\tau + 0.03)$

whereas RMSE exceeds 0.20 only in those regions with a mean AERONET AOD above 0.30, such as over the Amazon, Sahara, Middle East, or some Asian regions. Note that, compared to the mean AERONET AOD, the RMSE is high over western North America, South America, Australia and Central Asia. More particularly, over western North America and Australia, the very small mean measured AOD drastically increases the relative RMSE shown in Fig. 4. The L2 AOD EE,  $\pm(0.15\tau + 0.05)$ , does not represent the interval that contains 66 % of the L3 samples for any region but the Amazon. Conversely, the newly proposed L3 AOD EE always defines an interval encompassing between 65 % and 68 % of the data, for all regions. Overall, and for all regions, the L3 AOD EE indicates a higher EE of the L3 AOD dataset for large AODs. However, for small AODs in some regions, the EE is less than for the L2 AOD dataset, such as over eastern North America, Europe, Africa or Australia. Globally, the EE for small AODs is about 0.06, in similarity with the L2 AOD EE (0.05).

### 3.3 Uncertainty as a function of the L3 AOD value

For some applications the uncertainty associated with the AOD data is needed. However, a description based on EE (Table 1) is not always useful because it is based on the “true” AOD value (in practice, in ground sunphotometric measurements), which is rarely known. It is possible to follow a similar approach to that in the previous section, this time expressing the uncertainty as a function of the retrieved L3 AODs. To that effect, the L3 AOD residuals (L3 AOD minus AERONET AOD) have been grouped in 100 equal-sized

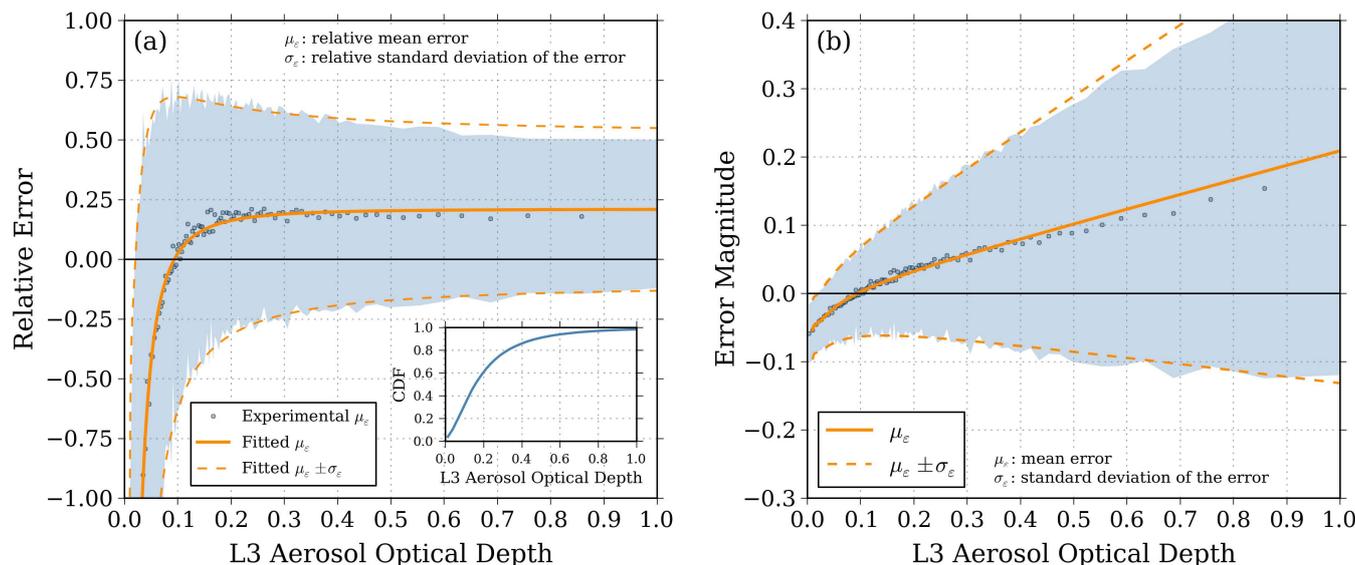
intervals of L3 AOD. In each bin, their mean and standard deviation have been calculated and normalized by the mean L3 AOD for that bin. Figure 5a shows the relative mean error (blue points) and the uncertainty interval defined as the relative mean error plus or minus one relative standard deviation of the error (shaded region), for the entire dataset, as a function of the L3 AOD value. A justification for characterizing the uncertainty with the standard deviation of the error rather than the EE as in Sect. 3.2 is offered in Sect. 5.

Both the relative mean error and relative standard deviation of the error can be modeled using a fifth-order polynomial, such as:

$$\Delta_{L3} = a_0 + a_1\eta + a_2\eta^2 + a_3\eta^3 + a_4\eta^4 + a_5\eta^5, \quad (1)$$

where  $\Delta_{L3}$  is either the relative mean error or the relative standard deviation of the error, and  $\eta = 1/\tau$ , with  $\tau$  being the L3 AOD value. The coefficients  $a_i$  are listed in Tables A1 and A2 separately for the relative mean error and the relative standard deviation of the error, and for each geographical region. All the fits have correlation coefficients ( $R^2$ ) above 92 %. For the case of Fig. 5, i.e., the entire dataset, the fitted curves (orange lines) are very close to those experimental points whose L3 AOD is below 0.5, thus representing  $\approx 90$  % of the dataset (inset of Fig. 5a).

Figure 5a shows a rapid increase of the relative underestimation when the L3 AODs are below 0.1. Above that threshold, a fairly constant relative overestimation is typical in the satellite retrievals. This is also true for the relative standard deviation of the error. In Fig. 5b, this translates into a linear increase of both the mean error and the standard deviation



**Fig. 5.** (a) Relative mean error (L3 AOD minus AERONET AOD) for the whole dataset (blue dots), and uncertainty range as relative mean error plus or minus one relative standard deviation of the error (dashed area). (b) Same as (a) but now the mean error and uncertainty are given in AOD unit. The orange lines are the fitted curves using Eq. (1). Coefficients for the “Globe” region are given in Tables A1 and A2.

for L3 AODs larger than 0.1. For smaller values, the satellite retrievals tend to be too low compared to ground truth. This underestimation increases when AOD decreases. This behavior is general, even for the western US. In this particular case (not shown here), the overestimation becomes noticeable above  $\approx 0.05$ , whereas the asymptotic relative error exceeds 50 % beyond 0.2.

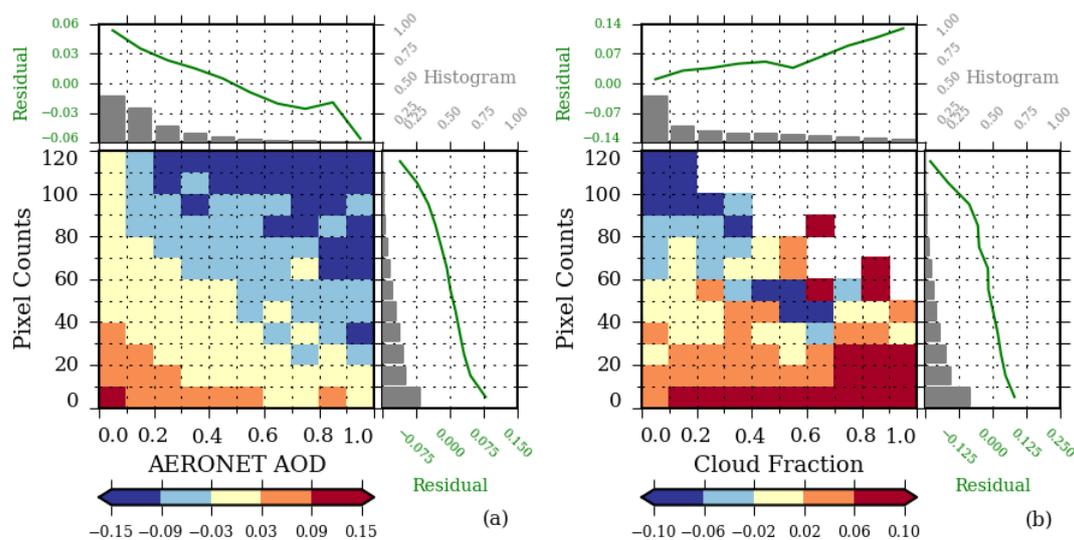
#### 4 The role of pixel count

Section 1.1 describes how the L3 AOD dataset is derived from the L2 AOD data. This procedure intrinsically defines the number of L2 AOD points – or pixel count – used to calculate each L3 AOD value. Levy et al. (2009) conducted a detailed study on the impact that different strategies of averaging could have on spatio-temporal aggregations of L2 AOD values such as for L3 AOD. They concluded that the differences between different averaging approaches may be greater – or, at least, of the same order – than differences between different measurement sensors. One of the cases they analyzed was a weighting scheme based on pixel-count values. The use of pixel count as a weight is reasonable by assuming a higher reliability for those L3 AOD values derived with a greater pixel-count number. However, the direct relationship between pixel count and L3 AOD reliability has apparently not been analyzed in the literature yet. Nonetheless, this relationship can be important for subsequent aggregations of L3 AOD values. Therefore, an important question is whether a higher pixel count necessarily means a more reliable L3 AOD value. If this were true, pixel counts could be used in the data quality assurance procedure, or in a weight-

ing scheme to combine L3 AOD values, for instance. A pixel count is provided for every grid cell in the L3 AOD dataset, and this information is used in what follows.

Figure 6 shows the joint histogram of the L3 AOD residuals (L3 AOD minus AERONET AOD) binned by either a combination of pixel count and AERONET AOD (Fig. 6a), or of pixel counts and cloud fraction (Fig. 6b), for all available collocations during the studied period. The mean residual for all points forming each column bin (green line), and the fraction of points in that bin (grey bars), both appear in the plot above each joint histogram. Likewise, the plot to the right of each joint histogram contains the mean residual for all the points forming each pixel-count bin (green line) and the fraction of points in that bin (grey bars). The joint histogram in Fig. 6a, similarly to Fig. 3, shows that the L3 MODIS retrievals are too high for small AODs and too low for high AODs. This was a known issue in the L2 AOD dataset, which therefore propagates to the L3 product. Overall, since  $\approx 80\%$  of the data have a L3 AOD below 0.3, MODIS retrievals are too high in general. According to Fig. 4, MODIS only underestimates in Africa and some parts of Asia. Possible causes of these biases are discussed by Levy et al. (2010).

Interestingly, a trend in the residual is also observed for the pixel counts (right plot in Fig. 6a). L3 AODs calculated with less than 50 pixel counts (about 90 % of the dataset) tend to overestimate, whereas they underestimate in the case of larger pixel counts. According to Fig. 6b, there is no L3 grid point with both a high cloud fraction and a large pixel count. Thus, the underestimation for points calculated with a large pixel count is dominated by cloudless conditions,



**Fig. 6.** Joint histograms of the L3 AOD residuals (L3 AOD minus AERONET AOD). **(a)** AERONET AOD in the x-axis and pixel counts in the y-axis. Colors of the joint-histogram bins indicate the mean residual of the collocations within that bin. **(b)** Cloud fraction in the x-axis and pixel counts in the y-axis. Colors of the joint-histogram bins indicate the mean residual of the collocations within that bin. In both panels, the top plot contains the mean residual for all points forming each column bin (green line) and the fraction of points in that column bin (grey bars). The plot to the right of each histogram contains the mean residual for all the points forming each row bin (green line) and the fraction of points in that row bin (grey bars).

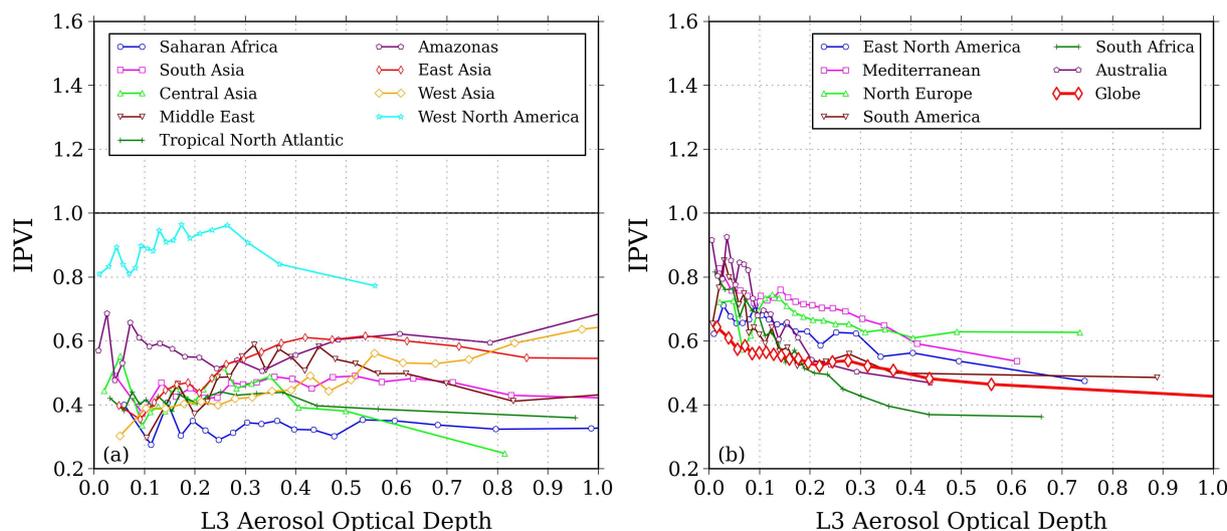
whereas the low pixel count in the other L3 grid points is partly caused by the cloud screening algorithms. It is known that clouds cause an overestimation of AOD in the MODIS algorithm retrievals. This can partially explain the overestimation of L3 AODs when derived with a low pixel count. In contrast, the underestimation observed in the case of large pixel counts could be related to image distortion at high sensor scan angles, and to unresolved surface reflectance variability. This trend relating the residual to pixel count can be observed over almost all regions, but with more or less magnitude depending on regional climatic conditions. A simple third-order polynomial could be used to remove the trend. However, the regional variability of the trend prevents the use of a unique function. Instead, we suggest to analyze the dataset at hand to derive a correction function for that specific dataset or region. For instance, using the entire dataset in this study, the L3 AOD EE after the pixel-count trend correction becomes  $\pm(0.21\tau^2 + 0.12\tau + 0.05)$ , which slightly reduces the EE error for AODs below 0.25 and above 0.5. The correction could be better if applied, as mentioned, at a regional scale. When aggregating L3 AOD values, another possibility is to discard all AODs corresponding to pixel counts larger or smaller than certain limits.

The consequence of the pixel-count trend just described is that great care must be taken when the pixel count is used as a weighting parameter to combine L3 AOD values. In particular, it is interesting to examine the use of the pixel-count weighting average as method to aggregate L3 AOD over time. As discussed earlier, the L3 AOD dataset tends to underestimate for large pixel counts. The use of pixel count

as weight would therefore assign more relative importance to AOD values that are systematically smaller than ground-truth. Furthermore, since the MODIS AOD retrievals generally tend to overestimate ground-truth (Fig. 4), the overall overestimation happens to be corrected just by cancellation of errors. However, this general MODIS overestimation does not occur over all regions. In particular, over any specific region where the satellite retrievals tend to underestimate, the use of the pixel-count weighting scheme would result in less accurate AODs than those from an unweighted average.

## 5 Spatial representativeness uncertainty

Section 3.3 provided an estimate of the total mean bias and uncertainty of L3 AODs by comparison with daily AERONET measurements. This uncertainty is mainly caused by uncertainties in both measurements and retrieval algorithm, but also by the unconsidered AOD variability that occurs at finer spatial and temporal scales than the spatial and temporal resolutions of the L3 MODIS dataset. These errors are also known as spatial and temporal representativeness uncertainties, respectively. One particular aspect of this issue is that L3 AOD values representative of a  $1^\circ \times 1^\circ$  grid cell are compared to spot ground-truth measurements. It can be expected that the importance of this issue increases with topographic complexity and average cell elevation. It is important to know how much these representativeness errors contribute to the overall uncertainty, and whether the spatio-temporal



**Fig. 7.** Intra-pixel variability index (IPVI), defined as the ratio of the standard deviation of the intra-pixel L2 AOD values,  $\sigma_s$ , provided as part of the L3 AOD product, to the standard deviation of the L3 AOD error as compared to ground AERONET measurements,  $\sigma_\varepsilon$  (Sect. 3.3); (a) regions with relatively stable IPVI and, (b) regions with decreasing IPVI for increasing L3 AOD.

resolution of the L3 AOD product is a limiting factor in achieving a smaller uncertainty.

Kaufman et al. (2000) showed that, for any single location, the instantaneous retrievals at overpass times could be used to represent the daily AOD average at that location within reasonable accuracy. Here, we compare the (daily) L3 AOD product at a  $1^\circ \times 1^\circ$  spatial resolution against daily averages at specific AERONET sites (i.e., spot measurements). Can such a large pixel size represent the spatial variability of AOD, and be appropriate for the comparison with spot ground measurements?

Along with the AOD at 550 nm and pixel count, the L3 AOD product provides the standard deviation of the L2 AOD values used in the aggregation leading to the L3 product. This standard deviation can be used as an estimate of the intra-pixel variability of AOD. By doing this, we are intrinsically assuming that the L2 AOD retrievals capture the intra-pixel variability of AOD, even though this does not necessarily mean a negligible overall mean error. Figure 7 shows the “intra-pixel variability index” (IPVI), i.e., the ratio of the standard deviation of the intra-pixel L2 AOD values,  $\sigma_s$ , provided as part of the L3 AOD product, to the standard deviation of the L3 AOD error as compared to ground AERONET measurements,  $\sigma_\varepsilon$ , (Sect. 3.3). Note that this ratio should never be larger than unity, because the spatial representativeness uncertainty is part of the total uncertainty.

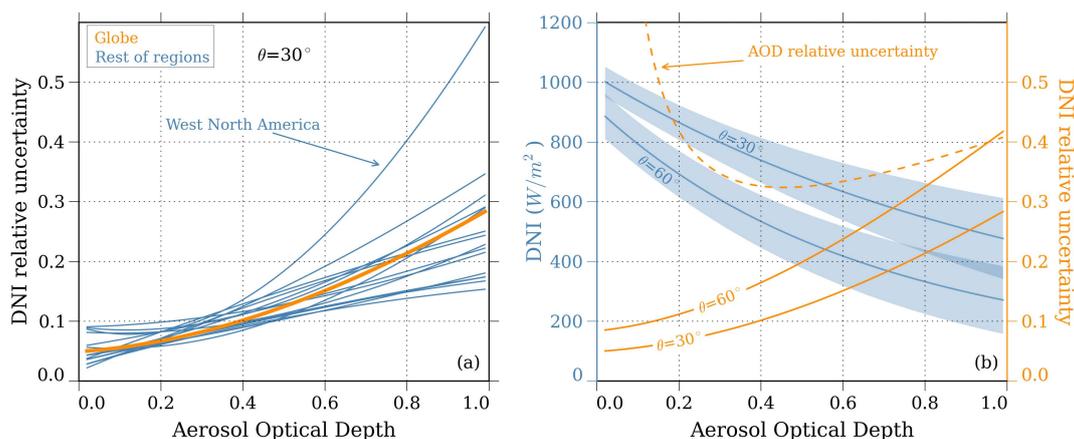
IPVI appears to obey two possible (and very different) patterns, depending on the region. On the one hand, regions with typically high aerosol load, usually dominated by the coarse mode, have a relatively constant IPVI, in the range 30–60% (Fig. 7a). One exception is western North America, where IPVI amounts to more than 80% of the spatial representativeness

uncertainty. This might be a consequence of the problems of the Dark-Target MODIS retrieval algorithm over that region, discussed earlier. On the other hand, in those regions with lower aerosol load and high relative importance of the fine aerosol mode (Fig. 7b), such as eastern North America, Europe, South America and southern Africa, IPVI decreases from  $\approx 80\%$  down to  $\approx 50\%$  when the L3 AOD increases. The entire dataset, as a region (Globe), is also in this group. Australia, however, seems to be a notable exception to this pattern.

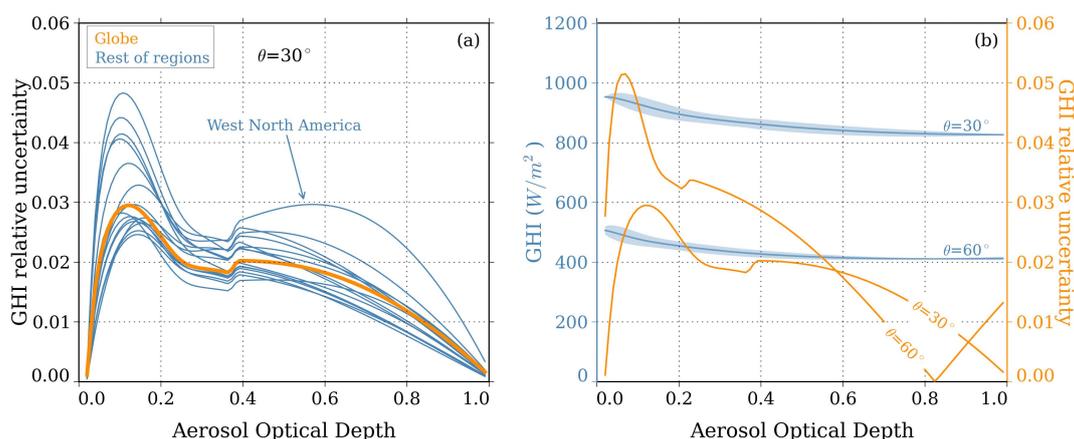
In rough terms overall, the spatial representativeness uncertainty contributes  $\approx 50\%$  to the total uncertainty of the L3 AOD. The remaining  $\approx 50\%$  is attributable to other errors. The only exception to this rule is for AODs below 0.1 over fine-mode-dominated regions, where the importance of the spatial representativeness error increases.

## 6 Induced uncertainty in surface solar irradiance

AOD is the main variable driving DNI under cloudless skies. Its impact is about 3 to 4 times larger on DNI than on GHI (Gueymard, 2012). Therefore, AOD is a key element in DNI modeling. Numerical Weather Prediction and radiative transfer models are applied to vast areas, so that an accurate prediction of DNI (and, to a lesser extent, GHI) requires a gridded AOD dataset with sufficient spatio-temporal resolution and geographical coverage. Gueymard (2012) indicated the convenience of daily AOD to capture the daily DNI variations. Moreover, Sect. 5 showed that the spatial representativeness uncertainty associated with the L3 AOD is about half of the total uncertainty. Therefore, the  $1^\circ \times 1^\circ$  spatial resolution of the L3 MODIS product appears adequate. Besides, it



**Fig. 8.** (a) Relative uncertainty in DNI for each region as a function of AOD, for the entire dataset and for  $\theta = 30^\circ$ . (b) Magnitude (blue) and relative DNI uncertainty (orange) for the entire dataset for  $\theta = 30^\circ$  and  $\theta = 60^\circ$ . The blue-shaded area is the DNI uncertainty in  $W m^{-2}$ . The blue line is for the DNI value. Also shown is the AOD's relative uncertainty (dashed line).



**Fig. 9.** Same as Fig. 8, but for global horizontal irradiance (GHI).

exists with a near global coverage. The remaining question is whether it is worthy to use L3 AODs for DNI modeling, considering how the uncertainty in AOD propagates into the DNI predictions.

A sensitivity analysis can be conducted with the help of the REST2 broadband solar radiation model (Gueymard, 2008), using uncertainty propagation theory:

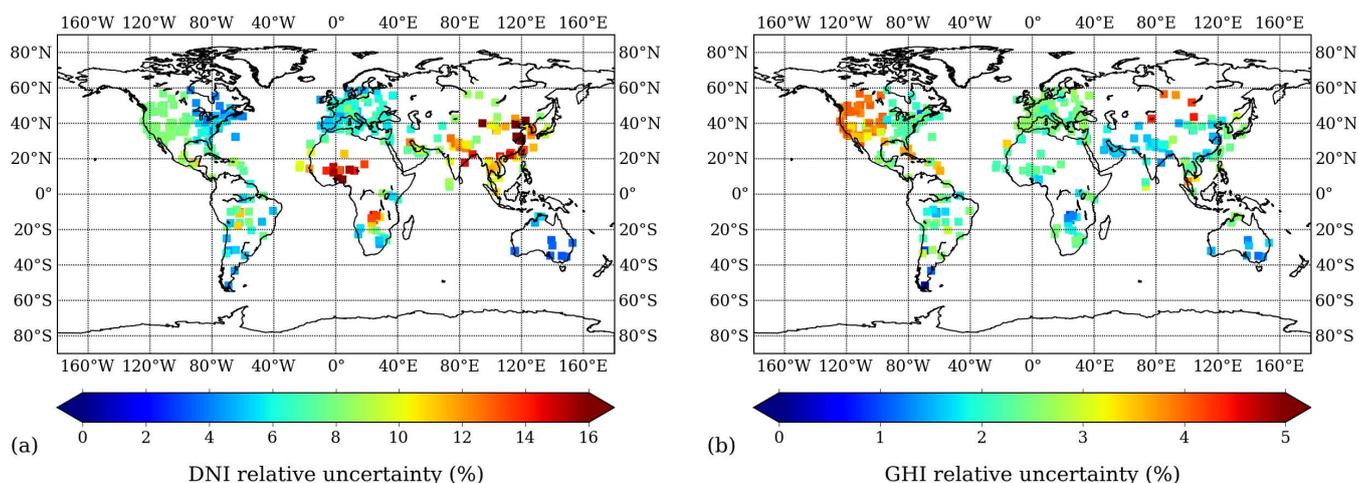
$$\Delta E_b = \left| \frac{\partial E_b}{\partial \tau} \right| \Delta \tau, \quad (2)$$

where  $E_b$  is the DNI (GHI),  $\Delta \tau$  is the AOD uncertainty and  $\Delta E_b$  is the uncertainty in DNI (GHI) due to  $\Delta \tau$  alone.

If we assume a mean AOD value of 0.2 – referred to as “mean global AOD value” in Table 1 – and request a maximum uncertainty of, say, 5% or 10% in DNI, due only to uncertainty in AOD, the maximum acceptable uncertainty in AOD would be 6% and 12%, respectively, i.e., 0.012 or 0.024 in AOD unit. For GHI, the maximum tolerable uncertainty in AOD would increase to 17% or 34%, respectively.

These values were obtained with the REST2 model for a solar zenith angle of  $30^\circ$  and fixed atmospheric standard values: atmospheric pressure = 1013.25 hPa, precipitable water = 1.5 cm, ozone amount = 0.35 atm cm (350 DU), nitrogen dioxide = 0.0002 atm cm, and Angström exponent = 1.3.

Figure 8a shows the relative uncertainty in DNI ( $\Delta E_b/E_b$ ) as a function of AOD for each geographical region. It is obtained from Eq. (2), using the AOD uncertainty expressed as the EE provided in the last column of Table 1. The sensitivity term  $\left| \frac{\partial E_b}{\partial \tau} \right|$  was also calculated with the REST2 model for solar zenith angles  $\theta = 30^\circ$  and  $60^\circ$  and the same reference atmospheric inputs as above, but keeping AOD as the only variable. For  $\theta = 30^\circ$ , a similar relative uncertainty in DNI is found over all regions, i.e., below 15% for AOD smaller than 0.5. For larger AODs, differences between regions increase. In the case of western North America, a large increase of the relative uncertainty occurs for AODs larger than 0.4.



**Fig. 10.** Relative uncertainty in DNI and GHI at AERONET sites. The relative uncertainties are calculated from the mean measured AOD values at each site (Fig. 2a).

However, AOD in this region hardly exceeds 0.5, thus the relative uncertainty is usually below 20 %.

Figure 8b shows the magnitude and relative value of the DNI uncertainty for the entire dataset and the two solar zenith angles ( $30^\circ$  and  $60^\circ$ ). The relative uncertainty increases by 30–40 % when the solar zenith angle increases from  $30^\circ$  to  $60^\circ$ . However, the absolute uncertainty (blue-shaded region) does not simultaneously increase.

Whereas both DNI and GHI are essential in energy applications, and particularly solar resource assessments, GHI is the normal (if not only) radiative output of NWP, reanalysis or climate models. Figure 9 shows the same sensitivity analysis as in Fig. 8 but for GHI. The relative uncertainty induced in GHI by the L3 AOD uncertainty is always smaller than 5 % for any region and a solar zenith angle of  $30^\circ$  (Fig. 9a). But contrarily to DNI, the uncertainty now decreases as AOD increases. This is due to the opposite effect of AOD on DNI and diffuse irradiances: an increase of AOD reduces DNI and simultaneously enhances diffuse irradiance. Therefore, uncertainties in DNI and diffuse irradiance due to AOD tend to cancel each other. The smaller relative impact of AOD on GHI, comparatively to DNI, appears in Fig. 9b: when AOD increases from 0 to 1, GHI decreases by only  $\approx 13\%$ , whereas DNI decreases by  $\approx 50\%$ .

Based on this analysis, Fig. 10 further shows the relative uncertainties in DNI and GHI caused by mean L3 AOD errors but, this time, at AERONET sites. These results are calculated from the mean measured AOD values at each site (Fig. 2a) and with the same approach as in Figs. 8 and 9. The highest mean relative uncertainties in DNI ( $\approx 12\text{--}15\%$ ) occur in Asia, northern and southern Africa (Fig. 10a), coinciding with the highest mean measured AODs, as could be expected from the results in Fig. 8a. Note that, according also to Fig. 8a, the DNI relative uncertainty over western North America is larger (per unit AOD) than in the rest of regions,

at high AODs. However, when using mean measured AODs (Fig. 10a), the relative uncertainty in DNI is only around 8 %, far from the regions with the highest mean uncertainty. An explanation is that the measured AOD over western North America is typically low (with a mean of 0.10), and thus the mean uncertainty in DNI also remains low. In parallel, the sudden change in DNI uncertainty from the western to the eastern North America is a direct consequence of the much poorer L3 AODs in the former region.

Figure 10b shows the relative uncertainty in GHI, also based on mean ground-measured AODs. Now, and contrarily to the DNI case, western North America is the region with the highest mean relative uncertainty (about 5 %). Again, this is a consequence of the very low mean measured AODs. By comparison with Fig. 9a, it can be seen that the relative uncertainty in GHI is maximum for an AOD value of about 0.10. The large difference between western and eastern North America is still present in this case. Interestingly, the large DNI relative uncertainties observed over Asia and Africa correspond to low relative uncertainties in GHI, because of the large AODs over there. As can be understood from Fig. 9a, the relative uncertainty in GHI decreases as AOD increases, and consequently the GHI uncertainty is low in those regions.

## 7 Conclusions

We have presented a global validation of the combined land and ocean Level-3 MODIS AOD product (Collection 5.1) at 550 nm from the Terra satellite, and investigated its suitability for surface solar radiation calculations in numerical weather or radiative transfer modeling.

The study has been conducted at both global and regional scales, based on the Terra satellite dataset (2000–2011) and the collocated daily measurements from the AERONET

radiometric network at 525 stations, used as ground truth. Globally, the L3 AOD dataset overestimates by 17 %, with a RMSE of 73 % and a squared Pearson correlation coefficient of 67 %. On a regional basis (using 15 distinct geographical regions), the MBE varies between 2 % (eastern Asia) and 76 % (western North America). In parts of northern and southern Africa and Asia, the L3 AOD is too low, whereas too large values are found over central Asia, Middle East and (most particularly) western North America. The regional RMSE is always larger than 50 %, with specially high values for western North America (151 %), South America (123 %), central Asia (82 %) and Australia (88 %).

The Level-3 MODIS AOD consists of a spatial aggregation of the Level-2 MODIS AOD, so a key question addressed in this study is about the L3 AOD reliability compared to that of the original L2 AOD. This comparison uses the expected error, EE, which is defined as the interval around the true AOD that contains 66 % of the MODIS AOD estimates. Overall, the aggregation from L2 AOD to L3 AOD increases EE for AODs greater than about 0.3, which approximately correspond to 20 % of the highest values. For the remaining 80 %, the L2 AOD EE and the L3 AOD EE are very similar. A quadratic function is proposed to describe the L3 AOD EE, in lieu of the conventional linear function used for the L2 MODIS EE.

EE provides an estimate of the MODIS AOD uncertainty in terms of the AOD determined from ground truth. Since, in general, this “true” AOD is unknown, we developed a model for the L3 AOD uncertainty that provides both the mean error and standard deviation of the error as a function of the L3 AOD. This model shows that, for the entire dataset and AODs greater than 0.1, the L3 AOD overestimates. Furthermore, both the mean error and the standard deviation of the error increase linearly with L3 AOD. For L3 AODs below 0.1, the satellite retrievals underestimate, and both this underestimation and the standard deviation of the error increase as the L3 AODs decrease. We provide error functions for the entire dataset and for each region separately.

We have also investigated the role of the pixel count in the reliability of the L3 AOD. Overall, the L3 AODs derived with small pixel counts tend to overestimate whereas the L3 AODs derived with large pixel counts tend to underestimate. This occurs in both the global dataset and for each individual region. However, each region has its peculiarities. Consequently, the assumption of a higher reliability in the L3 AODs for larger pixel counts must be adopted only with caution and, ideally, should be checked for each spatio-temporal dataset.

We have explored the relative contribution of the spatial representativeness uncertainty to the overall uncertainty of the L3 AOD. We used the standard deviation of the L2 AOD values used to calculate the L3 AOD as an estimate of the spatial representativeness uncertainty in the L3 AOD. By comparing this against the standard deviation of the error obtained in Sect. 3.3 we found that, overall, the spatial repre-

sentativeness uncertainty is roughly similar to the other errors associated with the L3 AOD. Therefore, the L3 AOD’ spatial representativeness uncertainty is not a limiting factor of its reliability. However, this result suggests that, when the quality of the AOD retrievals eventually improves (as can be anticipated), the spatial resolution of the L3 AOD should be enhanced.

Finally, one key issue addressed in this contribution is the quantification of the uncertainty induced in the predicted DNI when the L3 AOD product is used as input to radiative transfer models. It is found that, overall, the induced uncertainty in DNI is smaller than  $\approx 15$  % for AOD values below 0.5 (90 % of the global dataset) and a solar zenith angle of  $30^\circ$ . Under this same condition, the relative uncertainty induced in GHI is always below 5 %. At regional scale, the L3 AOD is therefore of sufficient overall quality to produce good-enough GHI estimates, but, conversely, this is not generally the case for DNI. In spite of this significant AOD-induced uncertainty in DNI, we consider the daily L3 AOD dataset very valuable because it has the advantage to generate a correct daily variability in DNI, which is important in solar energy applications. Of course, any monthly AOD dataset derived from the daily L3 AOD will also be affected by the uncertainties discussed here.

*Acknowledgements.* The authors would like to thank the AERONET Science Team for establishing and maintaining the sites used in this investigation, and the MODIS Science Team for the processing and public availability of the Collection 5.1 Level-3 aerosol products. J. A. Ruiz-Arias and D. Pozo-Vázquez were funded by a Marie Curie Action under the project PEOF-GA-2010-273648 within the 7th European Community Framework Programme, the Spanish Ministry of Science and Innovation under the project CGL2011-30377-C02-01 and the University of Jaén and Caja Rural de Jaén under the project R1/13/2010/01. This work was carried out while the first author was hosted by NCAR.

Edited by: B. Mayer

## Appendix A

## Coefficients for Eq. (1)

**Table A1.** Best fit (least-square error) coefficients of Eq. (1) for the relative mean error in each geographical region.

Region	$a_0$	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$
Globe	+2.1016e-01	+1.2266e-03	-2.3511e-03	+3.9770e-05	-2.6765e-07	+4.8374e-10
West North America	+7.1735e-01	-4.8607e-02	+1.0321e-04	-1.5191e-06	+6.5441e-09	-7.1457e-12
East North America	+2.5915e-01	-1.3513e-02	-9.8362e-04	+9.7010e-06	-3.8668e-08	+3.4472e-11
Tropical North Atlantic	+5.6089e-02	+7.8105e-02	-1.3864e-02	+6.4374e-04	-1.3055e-05	+7.0825e-08
Amazonas	+2.4388e-01	-4.0352e-02	-4.9475e-04	+1.4915e-06	+3.6795e-09	-5.8018e-12
South America	+3.4466e-01	-3.2560e-02	-6.0864e-04	+4.8308e-06	-1.1458e-08	+3.0188e-12
North Europe	+3.1493e-01	-3.7988e-02	-7.8874e-04	+2.2351e-05	-2.9474e-07	+9.9069e-10
Mediterranean	+3.0271e-01	-3.0121e-02	+1.8079e-05	-2.4451e-05	+3.7300e-07	-1.2706e-09
Saharan Africa	+2.1433e-01	-2.1029e-01	+1.4844e-02	-1.0180e-03	+2.4567e-05	-1.9690e-07
South Africa	+1.1466e-01	-1.7303e-02	-1.3977e-03	+1.7575e-05	-8.3911e-08	+1.0311e-10
Middle East	+3.5175e-01	-8.7605e-02	+5.7959e-03	-1.8348e-04	+0.0000e+00	+0.0000e+00
Central Asia	+3.9875e-01	-2.9181e-02	-3.8358e-03	+7.7667e-05	-5.6089e-07	+1.1000e-09
West Asia	+2.3061e-01	-9.6103e-02	-2.7327e-03	+3.4063e-05	+0.0000e+00	+0.0000e+00
East Asia	+2.1887e-01	-5.0683e-02	-1.1754e-03	-2.0864e-04	+8.6074e-06	-8.6414e-08
South Asia	+1.0831e-01	-2.8966e-02	-1.0586e-02	+4.3893e-04	-7.5255e-06	+4.1292e-08
Australia	+4.6244e-01	-3.4511e-02	-1.5014e-04	+5.5557e-08	+5.6603e-10	-1.3655e-13

**Table A2.** Best fit (least-square error) coefficients of Eq. (1) for the relative standard deviation of the error in each geographical region.

Region	$a_0$	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$
Globe	+3.0673e-01	+3.3666e-02	+1.4334e-04	-2.8216e-06	+2.5807e-08	-5.2532e-11
West North America	+2.4489e-01	+1.9495e-02	+3.3796e-04	-3.1312e-06	+9.8862e-09	-9.2615e-12
East North America	+2.8104e-01	+1.7071e-02	+4.1317e-04	-5.8807e-06	+3.1788e-08	-3.0771e-11
Tropical North Atlantic	+2.6667e-01	+2.7871e-02	+6.1932e-03	-4.4659e-04	+1.0521e-05	-5.9687e-08
Amazonas	+1.1934e-01	+6.0612e-02	-1.6307e-03	+2.3573e-05	-9.8376e-08	+8.7116e-11
South America	+3.7705e-01	+2.7186e-02	+2.0452e-04	-1.5190e-06	+3.7121e-09	-9.8225e-13
North Europe	+2.5034e-01	+1.8410e-02	+1.2646e-03	-4.3194e-05	+5.2326e-07	-1.6751e-09
Mediterranean	+2.0632e-01	+3.4149e-02	-9.3781e-04	+4.3191e-05	-5.6045e-07	+1.8361e-09
Saharan Africa	+2.7400e-01	+4.8876e-02	+7.8326e-03	-9.0040e-04	+3.9994e-05	-5.2303e-07
South Africa	+3.2772e-01	+1.0769e-02	+1.2042e-03	-1.9588e-05	+1.0685e-07	-1.3748e-10
Middle East	+2.2785e-01	-4.6514e-02	+1.4682e-02	-5.9941e-04	+0.0000e+00	+0.0000e+00
Central Asia	+3.3506e-01	+2.5225e-02	+1.9813e-03	-5.8834e-05	+5.8103e-07	-1.2941e-09
West Asia	+1.4205e-01	+8.6013e-02	+6.0449e-03	-1.0170e-04	+0.0000e+00	+0.0000e+00
East Asia	+2.4084e-01	+7.8624e-02	+7.9394e-04	+2.4427e-04	-1.2244e-05	+1.3889e-07
South Asia	+2.2528e-01	+9.9628e-02	-5.5069e-03	+3.4382e-04	-8.1394e-06	+5.2737e-08
Australia	+1.8743e-01	+3.8519e-02	-3.3033e-04	+2.6922e-06	-5.0324e-09	+1.0540e-12

## References

- Anderson, T., Charlson, R., Winker, D., Ogren, J., and Holmén, K.: Mesoscale Variations of Tropospheric Aerosols, *J. Atmos. Sci.*, 60, 119–136, doi:10.1175/1520-0469(2003)060<0119:MVOTA>2.0.CO;2, 2003.
- Cebecauer, T., Šúri, M., and Gueymard, C.: Uncertainty sources in satellite-derived direct normal irradiance: how can prediction accuracy be improved globally?, in: Proc. SolarPACES Conf., Granada, Spain, 2011.
- Chen, S., Wang, S., and Waylonis, M.: Modification of Saharan air layer and environmental shear over the eastern Atlantic Ocean by dust-radiation effects, *J. Geophys. Res.*, 115, D21202, doi:10.1029/2010JD014158, 2010.
- Chu, D., Kaufman, Y., Ichoku, C., Remer, L., Tanré, D., and Holben, B.: Validation of MODIS aerosol optical depth retrieval over land, *Geophys. Res. Lett.*, 29, 8007, doi:10.1175/JAS3385.1, 2002.
- Eck, T., Holben, B., Reid, J., Dubovik, O., Smirnov, A., O’neill, N., Slutsker, I., and Kinne, S.: Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols, *J. Geophys. Res.*, 104, 31333–31349, doi:10.1029/1999JD900923, 1999.
- Gueymard, C.: Uncertainties in modeled direct irradiance around the sahara as affected by aerosols: Are current datasets of bankable quality?, *J. Sol. Energy Eng.*, 133, 031024, doi:10.1115/1.4004386, 2011.
- Gueymard, C.: Temporal variability in direct and global irradiance at various time scales as affected by aerosols, *Sol. Energy*, 86, 3544–3553, doi:10.1016/j.solener.2012.01.013, 2012.
- Gueymard, C. A.: REST2: High-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation – Validation with a benchmark dataset, *Sol. Energy*, 82, 272–285, doi:10.1016/j.solener.2007.04.008, 2008.
- Hoelzemann, J., Longo, K., Fonseca, R., Do Rosario, N., Elbern, H., Freitas, S., and Pires, C.: Regional representativity of AERONET observation sites during the biomass burning season in South America determined by correlation studies with MODIS Aerosol Optical Depth, *J. Geophys. Res.*, 114, D13301, doi:10.1029/2008JD010369, 2009.
- Holben, B., Eck, T., Slutsker, I., Tanre, D., Buis, J., Setzer, A., Vermote, E., Reagan, J., Kaufman, Y., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET – A federated instrument network and data archive for aerosol characterization, *Remote Sens. Environ.*, 66, 1–16, 1998.
- Hsu, N., Tsay, S., King, M., and Herman, J.: Aerosol properties over bright-reflecting source regions, *Geosci. Remote Sens.*, *IEEE Trans.*, 42, 557–569, doi:10.1109/TGRS.2004.824067, 2004.
- Hubanks, P., King, M., Platnick, S., and Pincus, R.: MODIS atmosphere L3 gridded product algorithm theoretical basis document Collection 005 Version 1.1, Tech. Rep. ATBD-MOD-30, NASA, 2008.
- Ichoku, C., Chu, D., Mattoo, S., Kaufman, Y., Remer, L., Tanré, D., Slutsker, I., and Holben, B.: A spatio-temporal approach for global validation and analysis of MODIS aerosol products, *Geophys. Res. Lett.*, 29, 8006, doi:10.1029/2001GL013206, 2002.
- IPCC: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, in: *Climate Change 2007: The Physical Science Basis*, edited by: Solomon, S., Qin, D., Manning, M., Marquis, M., Averyt, K. M. B., Tignor, M., Miller, H. L., and Chen, Z., Cambridge University Press, Cambridge, UK, 2007.
- Jaenicke, R.: Aerosol-cloud-climate interactions, chap. Tropospheric aerosols, *International Geophysics Series*, Academic Press, San Diego, California, 1993.
- Jethva, H., Satheesh, S., Srinivasan, J., and Levy, R.: Improved retrieval of aerosol size-resolved properties from moderate resolution imaging spectroradiometer over India: Role of aerosol model and surface reflectance, *J. Geophys. Res.*, 115, D18213, doi:10.1029/2009JD013218, 2010.
- Kaufman, Y., Tanré, D., Remer, L., Vermote, E., Chu, A., and Holben, B.: Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer, *J. Geophys. Res.*, 102, 17051–17067, doi:10.1029/96JD03988, 1997.
- Kaufman, Y., Holben, B., Tanré, D., Slutsker, I., Smirnov, A., and Eck, T.: Will aerosol measurements from Terra and Aqua polar orbiting satellites represent the daily aerosol abundance and properties?, *Geophys. Res. Lett.*, 27, 3861–3864, doi:10.1029/2000GL011968, 2000.
- Kiehl, J. T. and Ramanathan, V.: *Frontiers of climate modeling*, Cambridge University Press, New York, USA, 2006.
- Lara-Fanego, V., Ruiz-Arias, J., Pozo-Vázquez, D., Santos-Alamillos, F., and Tovar-Pescador, J.: Evaluation of the WRF model solar irradiance forecasts in Andalusia (southern Spain), *Sol. Energy*, 86, 2200–2217, doi:10.1016/j.solener.2011.02.014, 2011.
- Levy, R., Remer, L., Mattoo, S., Vermote, E., and Kaufman, Y.: Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance, *J. Geophys. Res.*, 112, D13, doi:10.1029/2006JD007811, 2007.
- Levy, R., Leptoukh, G., Kahn, R., Zubko, V., Gopalan, A., and Remer, L.: A critical look at deriving monthly aerosol optical depth from satellite data, *IEEE T. Geosci. Remote Sens.*, 47, 2942–2956, doi:10.1109/TGRS.2009.2013842, 2009.
- Levy, R. C., Remer, L. A., Kleidman, R. G., Mattoo, S., Ichoku, C., Kahn, R., and Eck, T. F.: Global evaluation of the Collection 5 MODIS dark-target aerosol products over land, *Atmos. Chem. Phys.*, 10, 10399–10420, doi:10.5194/acp-10-10399-2010, 2010.
- Li, B., Yuan, H., Feng, N., and Tao, S.: Spatial and temporal variations of aerosol optical depth in China during the period from 2003 to 2006, *Int. J. Remote Sens.*, 31, 1801–1817, doi:10.1080/01431160902926665, 2010.
- Liou, K.: An introduction to atmospheric radiation, vol. 84 of *International Geophysics Series*, Academic Press, 2nd Edn., 2002.
- Lohmann, S., Schillings, C., Mayer, B., and Meyer, R.: Long-term variability of solar direct and global radiation derived from IS-CCP data and comparison with reanalysis data, *Sol. Energy*, 80, 1390–1401, doi:10.1016/j.solener.2006.03.004, 2006.
- Mei, L., Xue, Y., Xu, H., Guang, J., Li, Y., Wang, Y., Ai, J., Jiang, S., and He, X.: Validation and analysis of aerosol optical thickness retrieval over land, *Int. J. Remote Sens.*, 33, 781–803, doi:10.1080/01431161.2011.577831, 2012.
- Mishchenko, M., Liu, L., Geogdzhayev, I., Travis, L., Cairns, B., and Lacis, A.: Toward unified satellite climatology of aerosol properties. 3. MODIS versus MISR versus AERONET, *J. Quant. Spectrosc. Ra.*, 111, 540–552, doi:10.1016/j.jqsrt.2009.11.003, 2008.

- Pozo-Vázquez, D., Wilbert, S., Gueymard, C., Alados-Arboledas, L., Santos-Alamillos, F., and Granados-Muñoz, M.: Interannual variability of long time series of DNI and GHI at PSA, Spain, in: Proc. SolarPACES Conf., Granada, Spain, 2011.
- Remer, L., Kaufman, Y., Tanré, D., Mattoo, S., Chu, D., Martins, J., Li, R., Ichoku, C., Levy, R., Kleidman, R., Eck, T., Vermote, E., and Holben, B.: The MODIS aerosol algorithm, products, and validation, *J. Atmos. Sci.*, 62, 947–973, 2005.
- Remer, L., Kleidman, R., Levy, R., Kaufman, Y., Tanré, D., Mattoo, S., Martins, J., Ichoku, C., Koren, I., Yu, H., and Holben, B.: Global aerosol climatology from the MODIS satellite sensors, *J. Geophys. Res.*, 113, S07, doi:10.1029/2007JD009661, 2008.
- Ruiz-Arias, J., Cebecauer, T., Tovar-Pescador, J., and Šúri, M.: Spatial disaggregation of satellite-derived irradiance using a high-resolution digital elevation model, *Sol. Energy*, 84, 1644–1657, doi:10.1016/j.solener.2010.06.002, 2010.
- Ruiz-Arias, J., Pozo-Vázquez, D., Lara-Fanego, V., Santos-Alamillos, F., and Tovar-Pescador, J.: A high-resolution topographic correction method for clear-sky solar irradiance derived with a numerical weather prediction model, *J. Appl. Meteorol. Clim.*, 50, 2460–2472, doi:10.1175/2011JAMC2571.1, 2011.
- Ruiz-Arias, J., Gueymard, C., Dudhia, J., and Pozo-Vázquez, D.: Improvement of the Weather Research and Forecasting (WRF) model for solar resource assessments and forecasts under clear skies, in: Proc. Amer. Sol. Energy Soc., Denver, USA, 2012.
- Tanré, D., Kaufman, Y., Herman, M., and Mattoo, S.: Remote sensing of aerosol properties over oceans using the MODIS/EOS spectral radiances, *J. Geophys. Res.*, 102, 16, doi:10.1029/96JD03437, 1997.
- Šúri, M., Remund, J., Cebecauer, T., Hoyer-Klick, C., Dumoriter, P., Huld, T., Stackhouse, P., and Ineichen, P.: Comparison of direct normal irradiation maps for Europe, in: Proc. SolarPACES Conf., Berlin, Germany, 2009.