NREL
NATIONAL RENEWABLE ENERGY LABORATORY

Improved Gridded Aerosol Data for India

C. Gueymard
Solar Consulting Services

M. Sengupta
National Renewable Energy Laboratory

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Technical Report
NREL/TP-5D00-58762
November 2013

Contract No. DE-AC36-08GO28308

NATIONAL RENEWABLE ENERGY LABORATORY

Improved Gridded Aerosol Data for India

C. Gueymard
Solar Consulting Services

M. Sengupta
National Renewable Energy Laboratory

Prepared under Task No. WFK6.1001

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

Technical Report
NREL/TP-5D00-58762
November 2013

Contract No. DE-AC36-08GO28308

NOTICE

This report was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or any agency thereof.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Available electronically at http://www.osti.gov/bridge

Available for a processing fee to U.S. Department of Energy and its contractors, in paper, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
phone: 865.576.8401
fax: 865.576.5728
email: mailto:reports@adonis.osti.gov

Available for sale to the public, in paper, from:

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
phone: 800.553.6847
fax: 703.605.6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/help/ordermethods.aspx

Cover Photos: (left to right) photo by Pat Corkery, NREL 16416, photo from SunEdison, NREL 17423, photo by Pat Corkery, NREL 16560, photo by Dennis Schroeder, NREL 17613, photo by Dean Armstrong, NREL 17436, photo by Pat Corkery, NREL 17721.

Printed on paper containing at least 50% wastepaper, including 10% post consumer waste.

List of Acronyms

AE	Ångström exponent
AERONET	Aerosol Robotic Network
AOD	aerosol optical depth
AOD550	AOD at 550
CM-SAF	The Satellite Application Facility on Climate Monitoring
GHI	global horizontal irradiance
DNI	direct normal irradiance
MATCH	Model of Atmospheric Transport and Chemistry
MATMOD	MATCH Optical Depth
MISR	Multi-Angle Imaging Spectroadiometer
MODIS	Moderate Resolution Imaging Spectroradiometer

Executive Summary

Using point data from ground sites in and around India equipped with multi-wavelength sunphotometers, as well as gridded data from space measurements or from existing aerosol climatologies, an improved gridded database providing the monthly aerosol optical depth at 550 nm (AOD550) and Ångström exponent (AE) over India is produced. Data from 83 sunphotometer sites are used here as ground truth to calibrate, optimally combine, and validate monthly gridded data from 2000 to 2012.

To compare the satellite and climatology data (at original resolutions of 1° x 1° or 0.5° x 0.5°) to the pointlike ground-truth data, scale height corrections are applied to the AOD550 data to compensate for the effect of local elevation.

Different optimal combinations of the satellite data are obtained for the three main seasons of the Indian subcontinent. Most generally, the Moderate Resolution Imaging Spectroradiometer data are found to be closest from ground truth, albeit with some bias, which is globally corrected for the whole subcontinent, separately for each season. One main difficulty resides in the numerous data breaks that are found in the satellite-based data bases, particularly over Rajasthan and during the monsoon season. An optimal combination of alternate satellite data and climatologies is applied in such cases.

Additionally, a statistical analysis is undertaken to evaluate how to optimally correct the monthly-mean AOD550 so that more accurate direct normal irradiance results can be obtained on a mean monthly basis.

The resulting AOD550 and AE databases are used to derive the necessary input to the State University of New York radiation model, namely a data set of AOD at 700 nm, with a spatial resolution of 0.1° x 0.1°.

Table of Contents

List of Figures

List of Tables

1 Introduction

The goal of this study is to develop a complete data set of aerosol optical depth (AOD) for India, to be used as input to the State University of New York satellite model in the process of deriving high-resolution solar irradiance estimates for India. This constitutes a revision of a previous study, which led to the National Renewable Energy Laboratory's release of a series of solar resource maps in 2010. The period of interest was then 2002 to 2008; it has been extended to 2002 to 2012 for this study, which will lead to a new release of solar resource maps and irradiance time series by the National Renewable Energy Laboratory.

The State University of New York model works at a spatial resolution of 0.1° x 0.1° (about 10 km x 10 km), and provides solar irradiance estimates at hourly intervals, using satellite imagery of cloud cover. Ideally, the aerosol input to the model should be at the same resolution, both spatially and temporarily. Unfortunately, such stringent requirements cannot be matched directly with any currently available source of aerosol data. Other issues are the limited accuracy of the existing sources of AOD data and their frequent missing data gaps. A large part of the present study is devoted to finding acceptable solutions to these issues. Considering the various constraints of this study, the target spatial resolution of the primary data is 0.5° x 0.5° (about 50 km x 50 km), from which the end resolution necessary for the model (0.1° x 0.1°) is derived by interpolation, as discussed in Section 4. The adopted temporal resolution is one month because satellite observations have too many missing points on a daily basis.

The issue of accuracy is critical in the context of solar resource assessment, particularly for concentrating technologies (concentrating solar power or concentrating photovoltaics), because AOD has a strong impact on direct normal irradiance (DNI). It has a smaller impact on global horizontal irradiance (GHI), as discussed in the literature; e.g., [1], which makes flat-plate technologies (such as photovoltaics) much less sensitive to AOD. The specific aerosol conditions of India must also be discussed, because they are dramatically different from those in most other areas of the world.

Asia is known for its large aerosol burden, due to a combination of various causes: (1) large sources of local dust, which tends to rise easily under hot, dry conditions; (2) long-range transport from dry areas or deserts in Asia or Africa; (3) smoke from biomass burning, as a result of agricultural practices in particular; and (4) anthropogenic pollution. This mixture of large quantities of aerosols frequently results in strong haze. When laden with significant industrial and transportation pollution from large population centers, it creates what is referred to as the Asian brown cloud. The finest aerosol particles may be washed out by rain during the monsoon period. This is known as the scavenging effect. Winds can import or export aerosols, depending on season and meteorological conditions. Finally, an important case is that of the Indo-Gangetic basin, in the north of India. Considerable aerosol sources exist there because of the presence of agriculture, industries, and an extremely high population density. Influx of aerosols from desert areas west of India is frequent. These aerosols cannot escape to the north because the Himalayas create an effective barrier, resulting in southward outflow toward the Bay of Bengal (Figure 1). An almost permanent haze is noticeable over the region. This can impair the solar resource significantly. In contrast, most parts of Rajasthan have a significantly lower aerosol load and thus a much better solar resource. A strong spatial gradient in aerosol load exists between the "clean" parts of Rajasthan and the northwest Indo-Gangetic Basin, which complicates the precise solar resource mapping in that area of India. This spatial gradient is generally evident from satellite imagery, actually showing a clear-cut separation between low- and high-turbidity areas over Rajasthan, in particular (Figure 2).

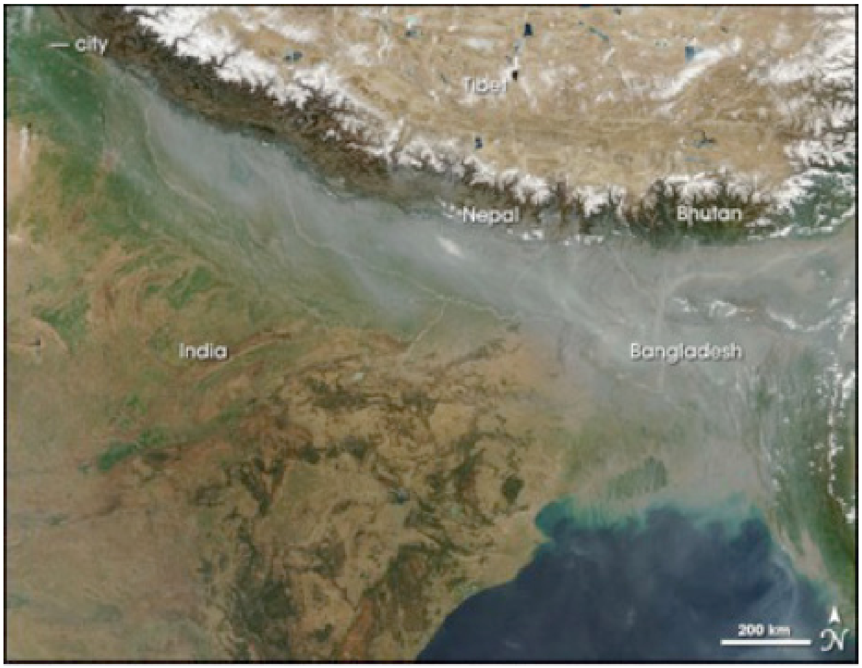


Figure 1. Intense haze over the eastern Indo-Gangetic Basin and the Bay of Bengal, as observed from satellite. ***Photo from NASA***

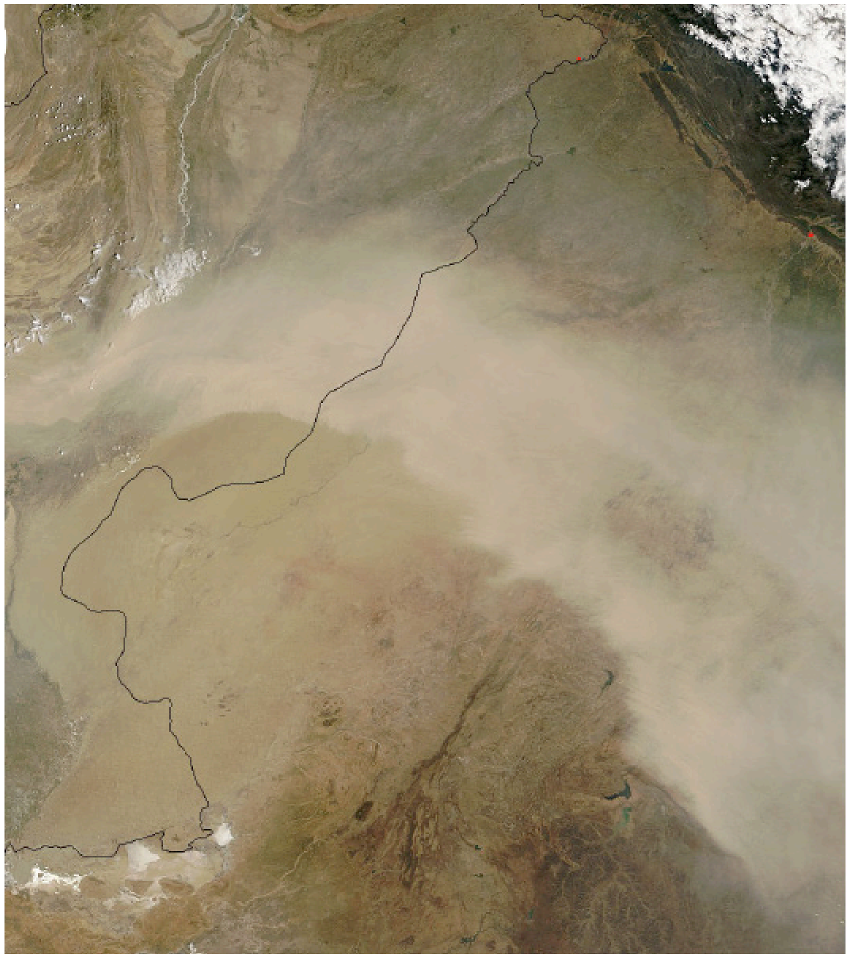


Figure 2. Aerosol brown cloud near or over the Thar Desert in Rajasthan, as observed from satellite. ***Photo from NASA***

Some studies have examined the long-term trends in aerosol emissions and/or AOD, showing significant increases over large parts of India, particularly the Indo-Gangetic Basin [2-4]. The Delhi area appears most affected, with upward trends in AOD exceeding 20% per decade. This can certainly explain, at least in part, the concomitant decrease in measured GHI that has been noted during the last decades [3, 4]. The dimming in GHI over India started around 1960. During the four decades that followed, calculations indicate that the increased AOD caused an average GHI dimming of 3 W m^{-2} to 4 W m^{-2} per decade, which corroborates the negative trend observed in the long-term GHI measurements from the Global Energy Balance Archive network [3]. As shown in Figure 3, similarly pronounced negative trends, of up to 7% per decade, have been reported, based on measured GHI data from the India Meteorological Department [5].

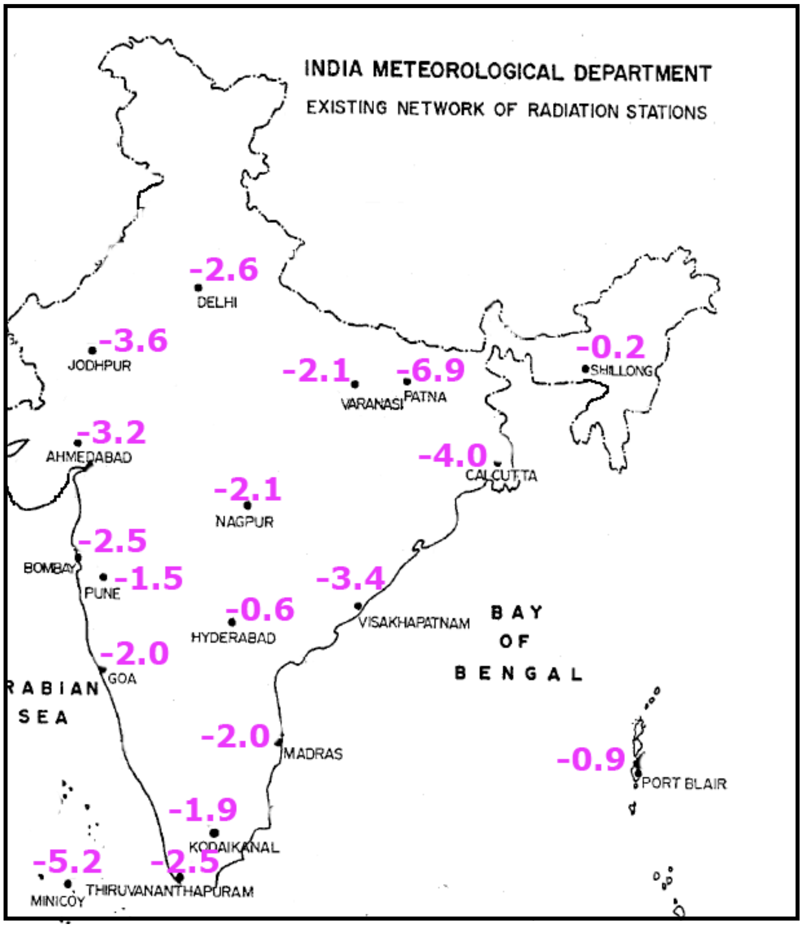


Figure 3. Trends in GHI (in percentage per decade) for some measurement stations. ***Image from Attri, 2008***

2 Aerosol Optical Properties

Two common aerosol optical properties are used here: the AOD at 550 nm (AOD550 or τ_{a550}) and the Ångström exponent (AE or α). The AOD at any other wavelength λ, $\tau_{a\lambda}$, can be derived through

$$\tau_{a\lambda} = \tau_{a550}\,(0.55/\lambda)^{\alpha} \tag{1}$$

where λ is expressed in µm. Equation (1) can be used, for instance, to obtain the AOD500 or the Ångström's turbidity coefficient, β, which is the AOD at 1 µm. Both of them have a strong presence in the literature, and have been around long before 550 nm became a de facto standard wavelength for satellite instruments. Ground-based sunphotometers usually report AOD at various wavelengths between 330 nm and 1,050 nm. (For instance, most Cimel instruments used by the Aerosol Robotic Network (AERONET) report AOD at 340 nm, 380 nm, 440 nm, 500 nm, 670 nm, 870 nm, and 1,020 nm, simultaneously.) AOD550 is most generally *not* measured with these instruments, and therefore must be derived. This can be accomplished by using the log-log transform of Equation (1):

$$\ln \tau_{a\lambda} = \ln \tau_{a550} + \alpha \ln(0.55/\lambda). \tag{2}$$

A linear fit of the simultaneous $\tau_{a\lambda}$ measurements to Equation (1) with the usual least-squares method provides both the intercept, $\ln \tau_{a550}$, and the slope, α. This process is illustrated in Figure 4 for three different months with widely different $\tau_{a\lambda}$ and α in Kanpur, India.

The current version of the State University of New York model relies on the AOD at 700 nm, which can be derived from Equation (1) as:

$$\tau_{a700} = \tau_{a550}\,(0.55/0.70)^{\alpha}. \tag{3}$$

In what follows, a methodology is developed to obtain gridded data of τ_{a550} and α without any missing data break, and with the appropriate statistical representativeness so as to derive hourly DNI and GHI with as little bias as possible. The desired τ_{a700} data set is then simply derived by applying Equation (3).

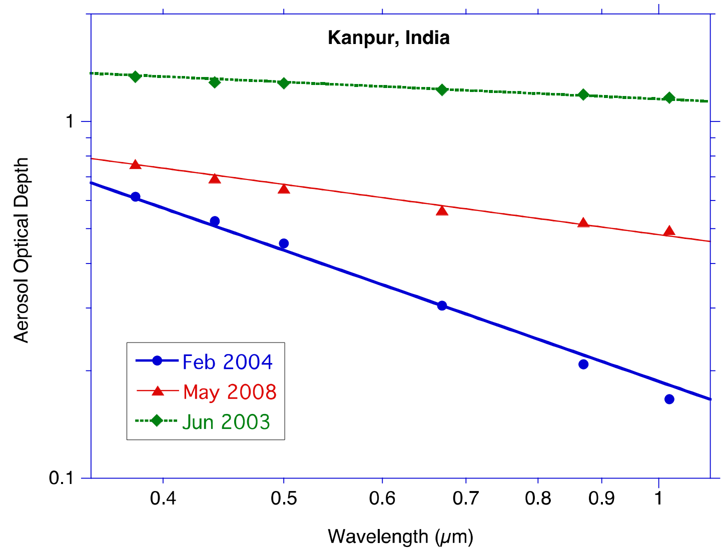


Figure 4. Spectral variation of mean monthly AOD for three different months in Kanpur

3 Sources of Data

Three widely different kinds of data are used here: (1) point data from ground sites equipped with sunphotometers (highest accuracy, but periods of missing data are frequent); (2) gridded data from space observations (low accuracy, with frequent missing data); and (3) gridded data from monthly climatologies (low to intermediate accuracy, no missing data). Data of the first kind are conventionally used as ground truth to validate and optimally combine data from the other sources.

Ground-based multi-wavelength sunphotometric measurements are the most accurate source of AOD data. Individual data points have an estimated uncertainty of approximately ±0.01 AOD unit [6]. Most validation studies of gridded aerosol products have used the National Aeronautics and Space Administration's AERONET data as ground truth. In the present case, this would not have provided enough sites or monthly data points to guarantee sufficient spatial coverage and statistical significance. To increase the number of ground-truth sites, the area of study has been increased to encompass not only India, but a much larger area, defined by latitudes 4° N to 39° N and longitudes 52° E to 103° E. Furthermore, the existence of sunphotometer data from sources other than AERONET has been queried. The literature shows that a sizeable number of sunphotometric stations belonging to various institutions (particularly the Indian Space Research Organization) exist in India. After considerable effort, it has been possible to obtain monthly-average AOD data for 18 stations, thanks to the intervention of the Ministry of New and Renewable Energy. The quality of this data set appears reasonable, but not as consistent or accurate as the AERONET data. Only a few data points had to be rejected as obvious outliers.

To mitigate this problem of insufficient ground-truth data, the literature has been searched thoroughly for any published data of AOD and AE. Of approximately 200 articles that have been scrutinized, about 10% contained valuable data (usually displayed in graphics, which had to be digitized), which ultimately resulted in 15 additional data series. Note that there were some instances in which different authors reported conflicting data for the same site and same period. Attempts to resolve such inconsistencies through personal communications with the concerned authors were unsuccessful (i.e., no response). Thus, dubious or conflicting data points had to be eliminated.

Through a combination of the AERONET and literature data, the total number of potential ground-truth sites then reached 95. From this total, 12 sites were not usable because either they were at too high altitude in the Himalayas (where the spatial variability in AOD is high and may seriously affect the validation); they were decommissioned before the first satellite data became available, in 2000; or their data were found to be of too low quality. As a result, only 83 sites were kept to establish the final ground-truth data set. The location of these sites is shown in Figure 5. It is obvious that the geographical distribution of the existing ground-truth stations is uneven. In particular, the lack of such stations where AOD is low and the solar resource is high (such as in large parts of Rajasthan) creates an important sampling problem. Two of the sites, Kanpur, India, and Anantapur, India, are indicated in Figure 5 by the orange and white circles, respectively. Kanpur is representative of the very hazy conditions of the Indo-Gangetic Basin, and has been the longest-operating AERONET station in India (since 2001). Its monthly-average AOD observations, reduced to a wavelength of 550 nm for comparative purposes, are shown in Figure 6. The seasonal variation of the 12-year long-term average is also shown. In contrast, the monthly AOD is lower and has a different seasonal pattern at Anantapur, in southern India (Figure 7). The data (available during the period from 2001 to 2009 only) for the latter sunphotometric station was extracted from the Ministry of New and Renewable Energy data set.

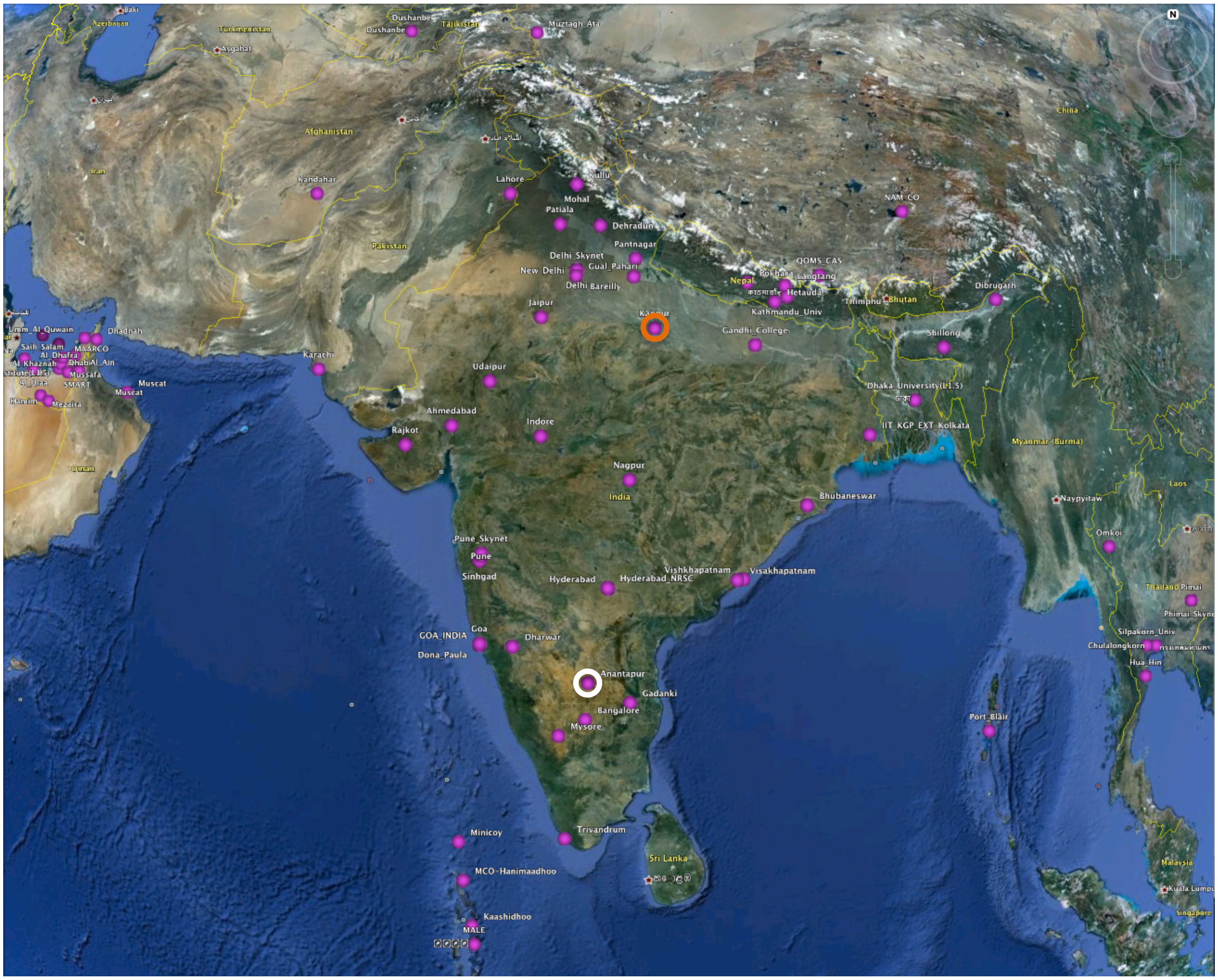


Figure 5. Location of ground-truth sites used in this study. Kanpur and Anantapur are indicated by the orange and white circles, respectively. ***Image © Google Earth***

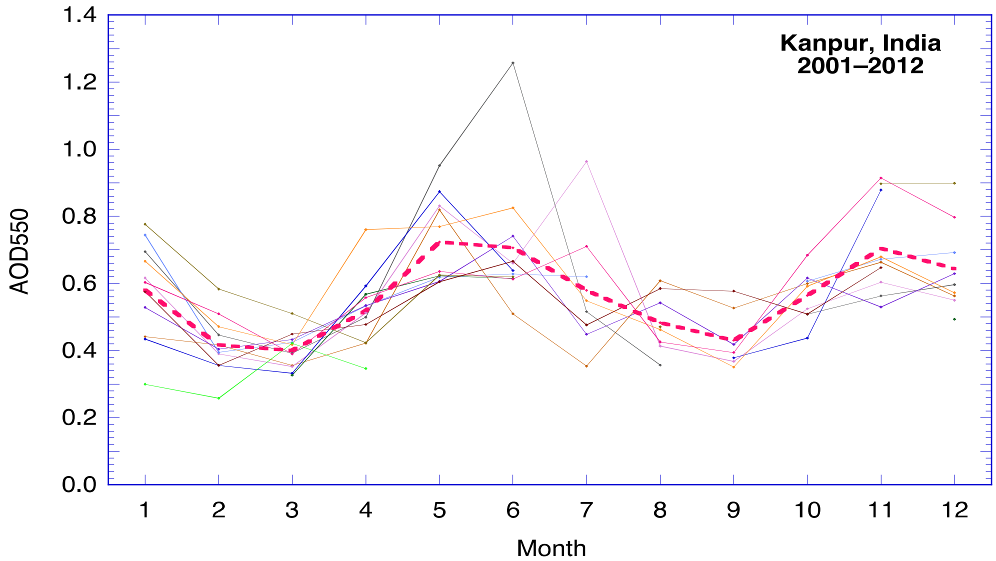


Figure 6. Observed monthly AOD550 at Kanpur for different years and the (thick, broken line) long-term average

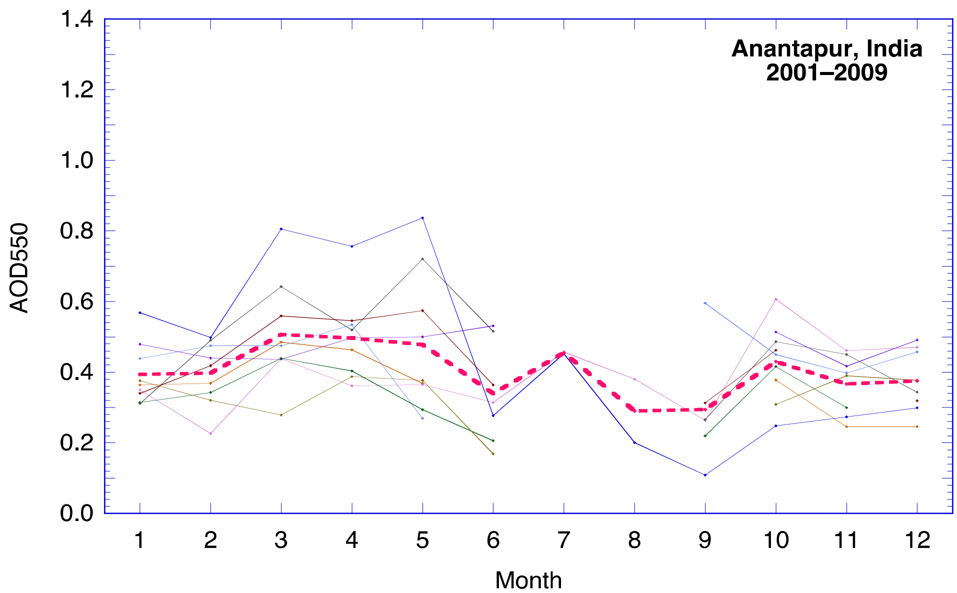


Figure 7. Observed monthly AOD550 at Anantapur for different years and the (thick, broken line) long-term average

In parallel, the sources of gridded data from satellite observations are shown in Table 1. A preliminary comparison of the four satellite AOD data sets described in Table 1 reveals two important things: (1) the magnitude of their distributions of AOD over India differ largely; and (2) there are a lot of missing data pixels, as a result of high ground albedo, high cloudiness, or other factors. These issues are illustrated in Figure 8 for a single common month (June 2008) when AOD is usually very high. Other patterns of absolute differences and missing data are observed for essentially all months, so that a comparison of gridded versus ground-truth AOD is needed to obtain performance statistics and select the best data source. This process is described in Section 5.

Monthly AOD climatologies have been derived from the data sets found in Table 1 by simple averaging of each grid-cell data point over the whole record period (Table 2). These climatologies still have missing data, however. To resolve this issue, continuous climatologies (with no missing data) from other sources have also been considered, as described in Table 2. A comparison of the seven AOD climatologies is shown in Figure 9 for the month of May. Again, it is obvious that significant differences exist between these data sets.

Although the target period for the end-product data set is from 2002 to 2012, all data points available for the period from 2000 to 2012 have been used for development.

Table 1. Sources of Gridded Satellite Data for AOD and AE*

Variable	Data Source	Period	Original Resolution
AOD550	MISR v4	2000–2012	0.5° x 0.5°
	MODIS-Terra v5.1	2000–2012	1° x 1°
	MODIS-Aqua v5.1	2002–2012	1° x 1°
	MODIS-Aqua “Deep Blue” v5.1	2002–2012	1° x 1°
AE	MISR v4	2000–2012	0.5° x 0.5°
	MODIS-Terra Land v5	2000–2010	1° x 1°
	MODIS-Terra Ocean v5	2000–2007	1° x 1°
	MODIS-Aqua Land v5.1	2003–2009	1° x 1°
	MODIS-Aqua Ocean v5.1	2002–2010	1° x 1°

*From the National Aeronautics and Space Administration

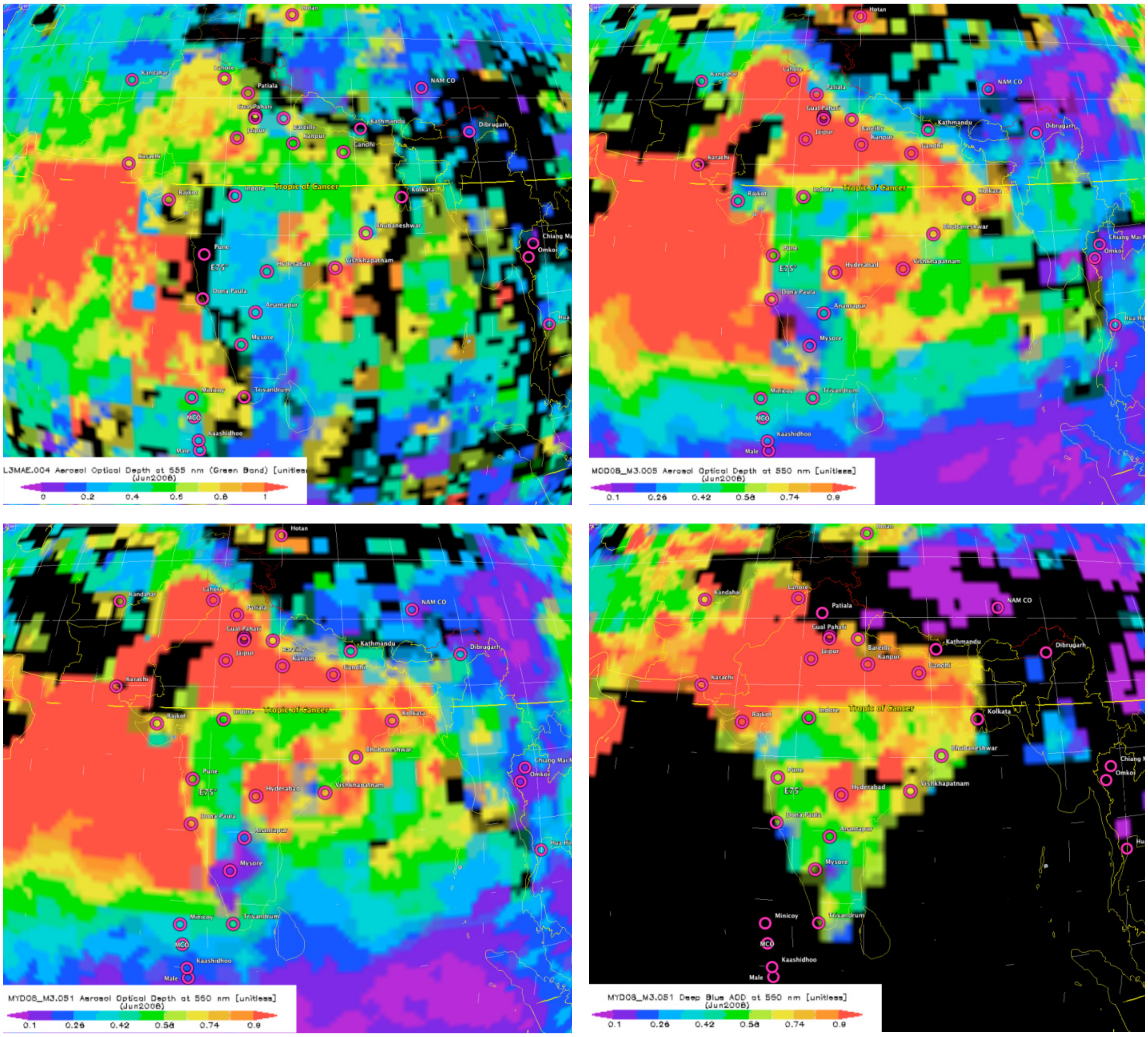
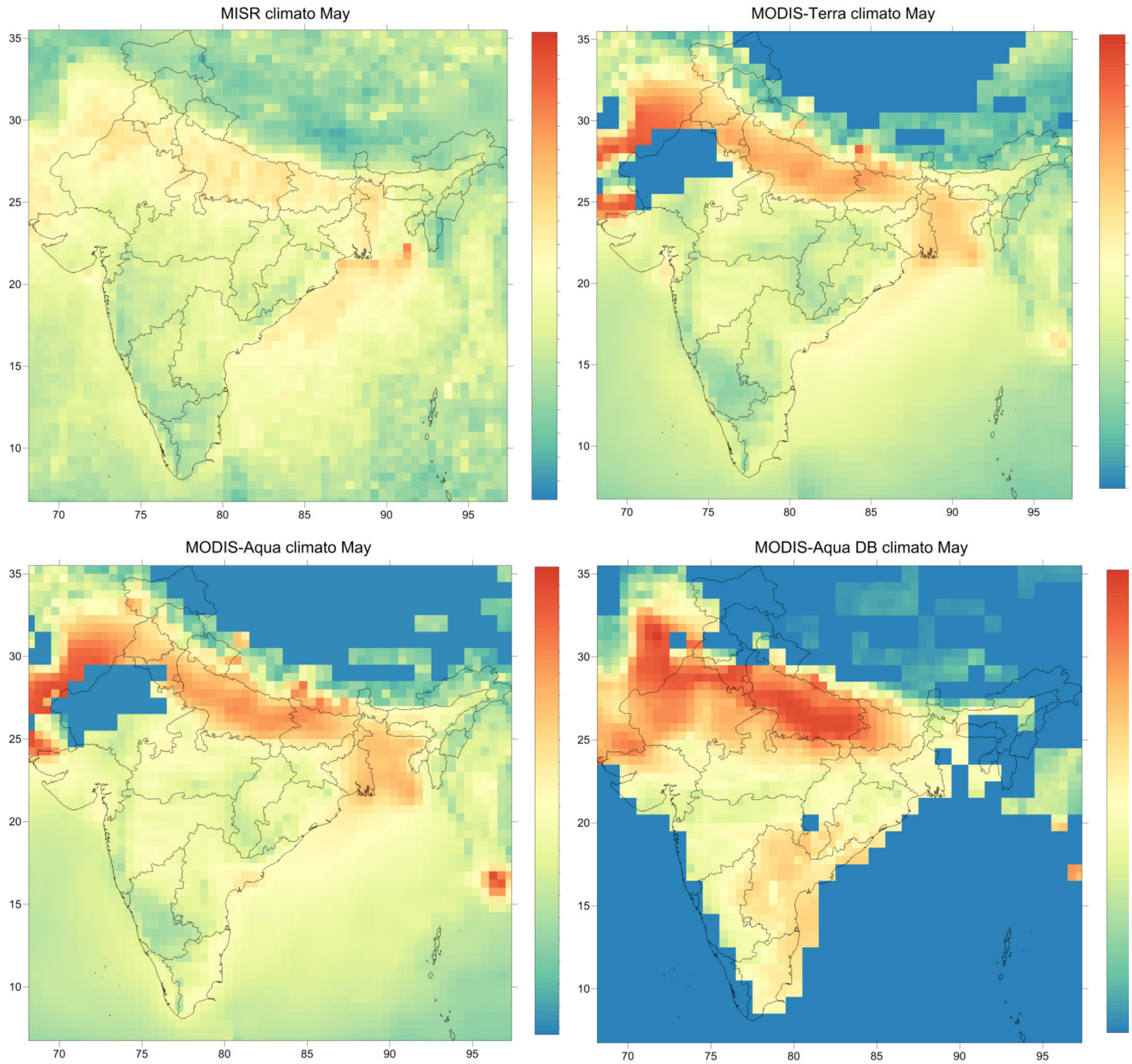


Figure 8. Mean AOD550 over India for June 2008, according to (top left) MISR, (top right) MODIS-Terra, (bottom left) MODIS-Aqua, and (bottom right) MODIS-Aqua Deep Blue

Table 2. Sources of Gridded Climatological Data for AOD and AE

Variable	Data Source	Resolution	Source	Continuous
AOD550	MISR v4	0.5° x 0.5°	Table 1	
	MODIS-Terra v5.1	1° x 1°	Table 1	
	MODIS-Aqua v5.1	1° x 1°	Table 1	
	MODIS-Aqua Deep Blue v5.1	1° x 1°	Table 1	
	CM-SAF	1° x 1°	www.cmsaf.eu/	•
	MATMOD	1° x 1°	[7]	•
	Meteonorm	1° x 1°	[8]	•
AE	MISR v4	0.5° x 0.5°	Table 1	
	MODIS-Terra Land v5	1° x 1°	Table 1	
	MODIS-Terra Ocean v5	1° x 1°	Table 1	
	MODIS-Aqua Land v5.1	1° x 1°	Table 1	
	MODIS-Aqua Ocean v5.1	1° x 1°	Table 1	
	CM-SAF	1° x 1°	www.cmsaf.eu/	•
	MATMOD	1° x 1°	[7]	•


MISR climato May
MODIS-Terra climato May
MODIS-Aqua climato May
MODIS-Aqua DB climato May
35
30
25
20
15
10
70
75
80
85
90
95

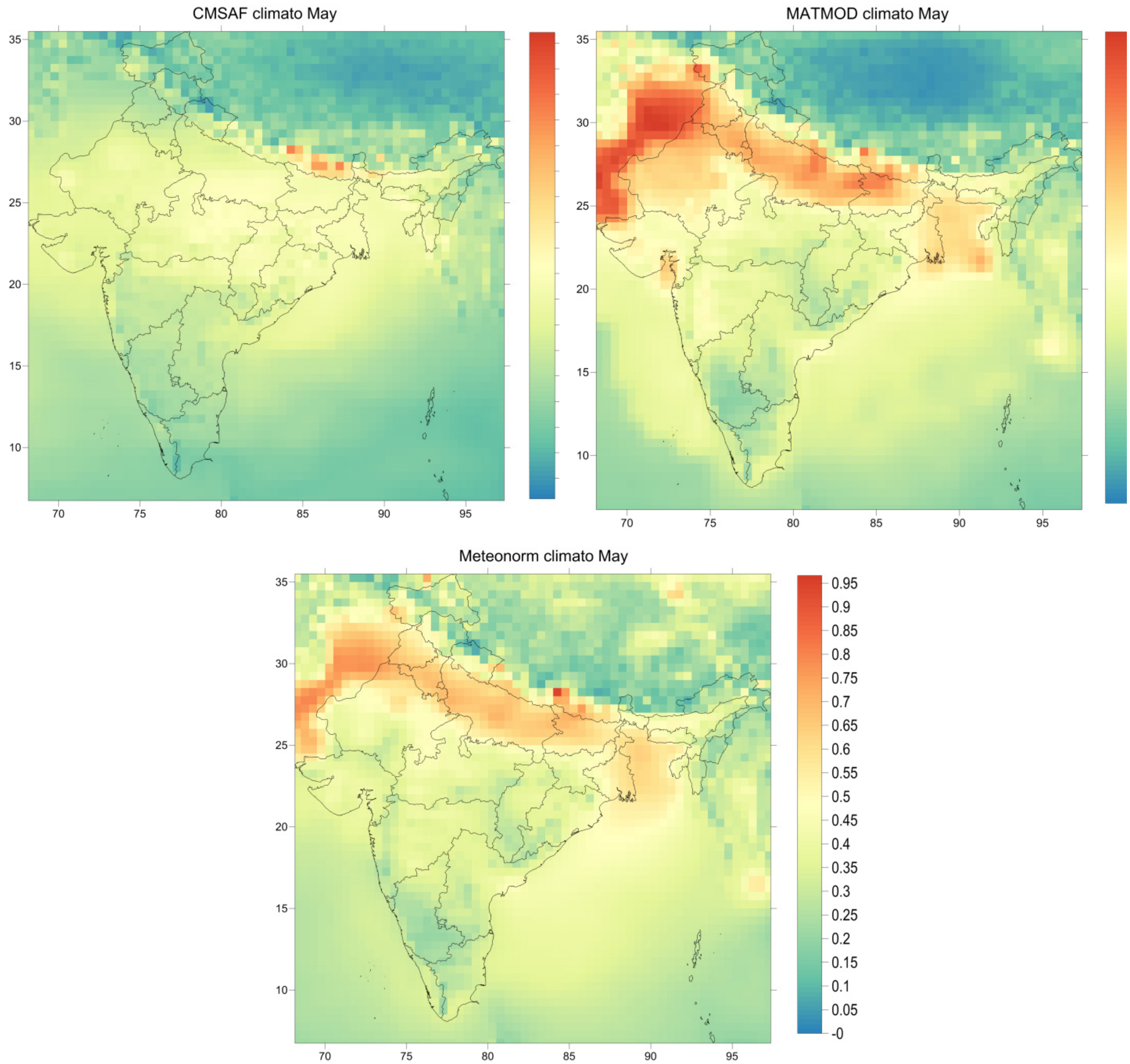


Figure 9. AOD550 climatology over India for May, according to all climatology data sets listed in Table 2, after regridding to 0.5° x 0.5° resolution

Compared to AOD550, AE appears to have a much more limited interannual variation, and a more distinct seasonal variation. This is illustrated for Kanpur in Figure 10. AE's steady decrease during the dry season corresponds to the buildup of dust particles of large size, for which α is typically low (because α is inversely related to the mean aerosol particle size). These large particles are then scavenged (by rain) during the monsoon season, so that α increases steadily until the next dry season.

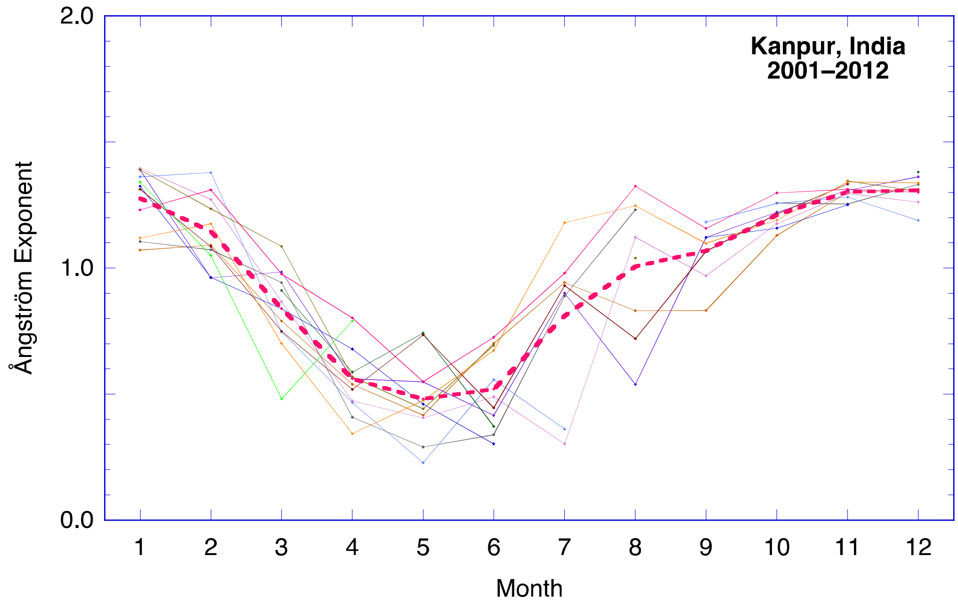


Figure 10. Observed monthly AE at Kanpur for different years and the (thick, broken line) long-term average

The geographical feature that is important to consider in the case of AE is the proximity to the coast. Maritime aerosols have a large size and are thus characterized by a low α, normally less than 0.5. Coastal areas (and even more so, islands) are therefore characterized by a lower α than inland areas in any season. This becomes obvious when comparing the situation at the inland site of Kanpur, India (Figure 10) to that of the coastal site of Karachi, Pakistan (Figure 11). The general U-shape trend of the seasonal variation of α still exists, because maritime areas are still strongly affected by dust and other types of aerosols transported from continental sources. In parallel, Table 1 and Table 2 indicate that the MODIS retrieval algorithms for AE vary depending on the expected presence of maritime aerosols, thus their outputs were separated into "Land" and "Ocean" data sets. Consequently, in contrast to the comparative analysis between satellite data and ground truth that can be carried out only on a seasonal basis for AOD550, that of AE additionally requires a separate analysis for the coastal and inland areas.

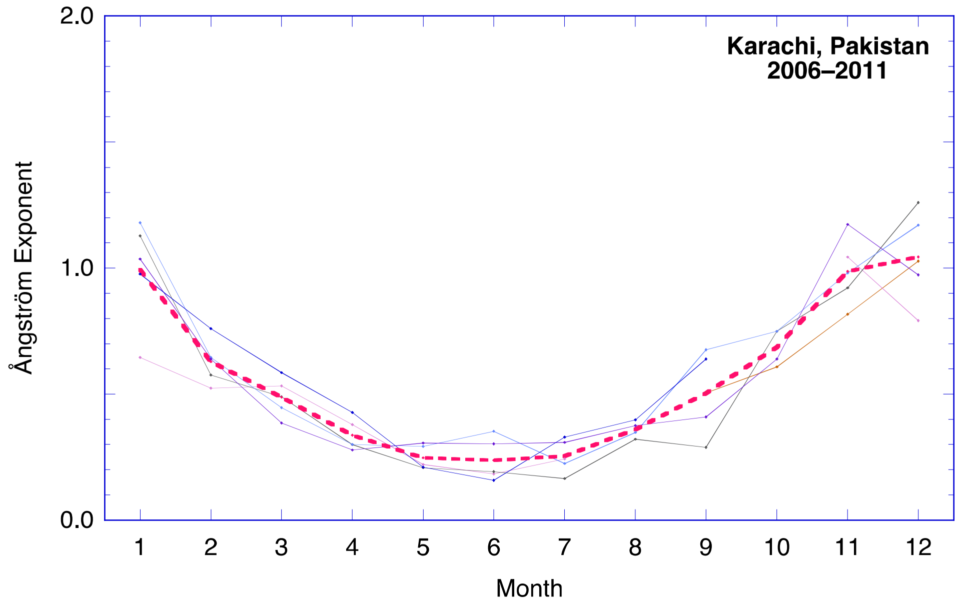


Figure 11. Observed monthly AE at Karachi for different years and the (thick, broken line) long-term average

4 AOD Data Interpolation

Table 1 and Table 2 reveal that the MISR data has better spatial resolution than the other data sources. All the 1° x 1° AOD data sets were thus regridded to 0.5° x 0.5° so that they could be compared to the MISR data. This also improves comparisons with ground observations because the point-source ground-truth data is then compared to gridded data that is spatially averaged over a much smaller cell area, which is therefore more representative of the ground-truth site conditions.

Because AOD is a strong function of elevation, this regridding must take the local elevation into consideration. This can be done via the convenient scale-height method, which assumes that the AOD at elevation h is lower than that at sea level (h=0), in such a way that

$$\mathrm{AOD}(h) = \mathrm{AOD}(0)\exp(-h/H_a) \quad (4)$$

where H_a is the aerosol scale height. The method is convenient, but approximate because H_a varies, depending on region, season, etc. Based on a previously published study [7], a typical value of H_a is approximately 2,500 m.

The overall interpolation is executed in three successive steps: (1) the 1° x 1° data points for each grid cell are first reduced to sea level through the reverse form of Equation (4); (2) a double linear interpolation is used to obtain 0.5° x 0.5° data at sea level; and (3) finally, Equation (4) is used to obtain the gridded values at their final resolution. The first and last step involve a digital elevation model of the study area at 1° x 1° and 0.5° x 0.5°, respectively.

Equation (4) has also been applied to the ground-truth data, so that their corrected values are scaled to the average elevation over the 0.5° x 0.5° cell where the site is located. If the ground site's elevation is h_g and the 0.5° x 0.5° cell's mean elevation is h_i, the corrected ground-truth AOD is obtained from

$$\mathrm{AOD}(h_i) = \mathrm{AOD}(h_g)\exp(-h_i/H_a) / \exp(-h_g/H_a). \quad (5)$$

The effect of elevation on AE is most likely much smaller than that on AOD, but is not known precisely. Therefore, it has been considered that the AE values for 1° x 1° cells are conserved when regridding to 0.5° x 0.5°.

5 Gridded Data Derivation

5.1 Seasonal and Geographical Considerations

The scarcity of ground-truth data, particularly during the cloudier monsoon months, has a negative impact on the number of valuable data points and on the overall statistical significance that would be achieved by the desirable monthly analysis. A seasonal analysis is therefore preferable. The literature on aerosols in India clearly separates the aerosol data into three distinct seasons: pre-monsoon (dry season), monsoon (wet season) and post-monsoon. The exact period and strength of the monsoon season actually varies with latitude and also has a marked interannual variability. For simplicity, the monsoon season is defined here as the four-month period from June to September, after which the three-month post-monsoon season starts, October to December. The dry season is therefore defined as the five-month period from January to May.

Considering the notable differences in AOD between all available satellite-gridded data sets (Figure 8) and between all available climatological data sets (Figure 9), it is important to compare each of them against ground truth. The number of data points for which coincident ground truth and gridded data points are available during the period from 2000 to 2012 is shown in Table 3. It is clear that, for any season, the MODIS-Aqua Dark Blue data set is at a disadvantage as a result of its large fraction of missing data. If a low fraction of data breaks were the only selection criterion, MISR and MODIS-Terra would be the sources of choice.

Table 3. Number of Monthly Data Points Used for the Validation of Each Satellite Data Set From 2000 to 2012 and Coincidence With Ground Truth

Season	Ground Truth	MISR	MODIS-Terra	MODIS-Aqua	MODIS-Aqua Dark Blue
Dry	828	757	729	643	457
Monsoon	479	374	382	355	271
Post-Monsoon	422	371	363	333	248
ALL	**1729**	**1502**	**1474**	**1331**	**976**
Coincident Availability	100.0	86.9	85.3	77.0	56.4

For each of the three identified seasons, Figure 12 provides a visual comparison between the monthly satellite-derived AOD data points (from MISR and MODIS-Terra) and ground truth. The scatter around the ideal 1:1 line is large in all cases. However, the correlation between the satellite-derived and measured AOD data points is much better in the case of MOSIS-Terra than with MISR, even though the latter has a finer nominal resolution. The scatter is more pronounced during the monsoon season, which can be explained by the predominantly cloudy weather, and therefore a higher probability of spatial heterogeneity in cloud cover, yielding a lower probability of spatial and temporal coincidence between the satellite and ground observations. Overall, the two satellite-derived data sets tend to underestimate AOD compared to ground truth. The only potential exception is for MODIS-Terra during the monsoon season, but the light overall overestimation might then be caused by a few outliers, themselves likely resulting from cloud contamination in the satellite data.

A similar analysis shows that, among the seven AOD climatologies mentioned in Table 2, the better ones are MODIS-Terra, MODIS-Aqua, MATMOD, and Meteonorm. It is clear from Figure 9 that the MISR and CM-SAF climatologies tend to severely underestimate the high AOD conditions over the Indo-Gangetic Basin.

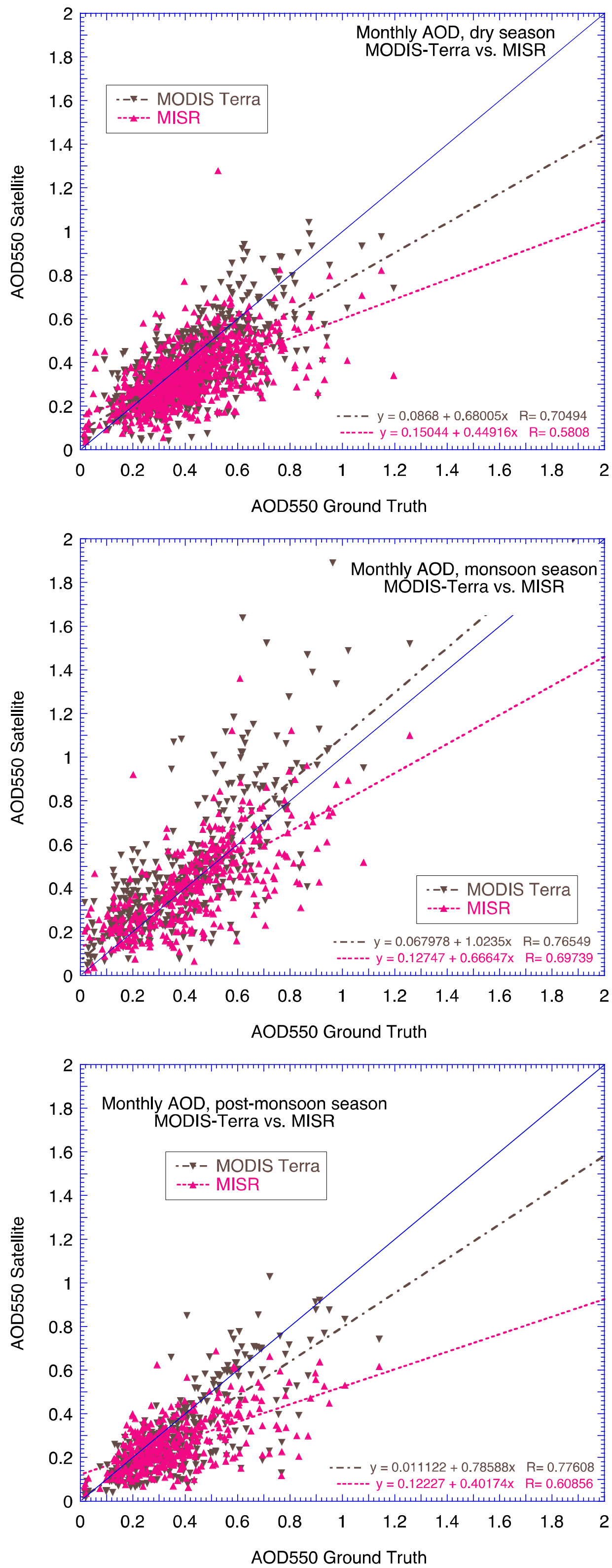


Figure 12. Seasonal validation of monthly AOD from MISR and MODIS-Terra for all ground-truth stations. Top plot: dry season; middle plot: monsoon season; bottom plot: post-monsoon season

A comparison of the satellite-derived α for both coastal and inland sites during the monsoon season is shown in Figure 13, again using data from the MODIS-Terra and MISR instruments. The observed large scatter in these plots confirms previous findings [9] that the quality of AE products from MODIS was particularly low. Unfortunately, none of the satellite-derived data sets for any of the seasons or geographical areas is of good quality. Although some of the bias can be reduced globally, as discussed in the next subsection, the low quality of the satellite data is obviously an area of concern, and a non-negligible source of error in the final product (AOD at 700 nm).

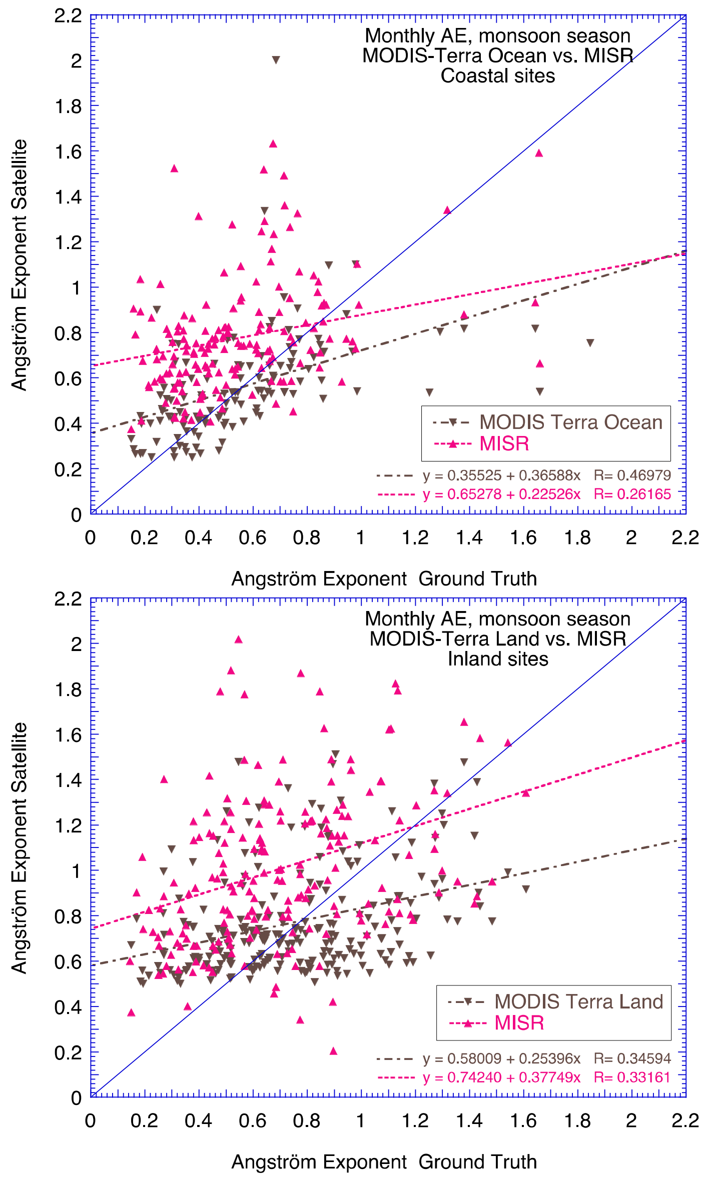


Figure 13. Validation of monthly AE from MISR and MODIS-Terra for all ground-truth stations during the monsoon season. Top plot: coastal sites; bottom plot: inland sites

5.2 Calibration of Satellite Data and Missing Data Fill-In

Figure 12 and Figure 13 indicate that, for both AOD550 and AE, the satellite data manifest both scatter and bias, which may depend on season and geographical location. The scatter is essentially a result of six sources of difference between the data sets being compared, sorted in decreasing order of assumed importance:

1. Errors in the satellite data (miscalibration of the retrieval algorithm, cloud contamination, ground reflectance artifacts, etc.)
2. Spatial mismatch between the satellite-based data (grid-cell average) and the ground-based data (point source)
3. Temporal mismatch between the satellite and ground data (The days used to derive the monthly average may not be the same.)
4. Errors introduced by the elevation correction
5. Errors introduced by the interpolation technique
6. Errors in the ground-truth data

The relative contribution of each source of noise is unknown, but it is likely that a significant part of the total noise observed in Figure 12 and Figure 13 is caused by the sources identified as numbers 2 through 6, and not only by errors in the retrieved aerosol data themselves (number 1). Whereas the noise-causing random errors are essentially unavoidable and impossible to correct a posteriori, it is desirable to correct for the overall bias. This is actually possible under the condition that it is relatively constant over large geographical regions and during a specific season. The correction method relies on scatter plots such as those shown in Figure 12 and Figure 13.

For AOD, the MODIS-Terra and MODIS-Aqua data sets appear to have less bias than the MISR data (Figure 12). To remove such bias as much as possible, a linear correction method is applied, based on the linear fits obtained from scatter plots such as shown in Figure 12 and Figure 13. The major problem with the MODIS data, however, is that some regions (such as around Rajasthan) or periods (monsoon months) have systematic data breaks (Figure 8 and Figure 9). The first attempted solution is to use MODIS data wherever/whenever available and MISR or climatology data elsewhere. However, this results in both temporal and spatial discontinuities or inconsistent results, particularly over Rajasthan. Following unsuccessful attempts of patching MODIS data breaks with various possible arrangements, the best solution is obtained with the Meteonorm data set, after scaling its values with the normalized MISR monthly AOD values, i.e., the values for individual months divided by the corresponding long-term averages. Finally, a weighted average of MODIS-Terra, MODIS-Aqua, and this scaled climatology has been developed for the whole subcontinent, yielding virtually indiscernible discontinuities over Rajasthan or during monsoon months.

For AE, various combinations of the monthly data from MISR and MODIS-Terra, with fill-in provided by combinations of the CM-SAF and MATMOD climatologies, appropriately scaled with normalized monthly values from MODIS or MISR whenever available, were used. The weighting of these combinations depends on location (coastal/inland) and season.

5.3 Production of Monthly-Mean Data Sets

A Fortran code has been written to prepare the production of monthly values of AOD550 at a resolution of 0.5° x 0.5° over India. A similar code has been written to produce the monthly AE values. These codes read all the necessary AOD or AE data sets, detect missing values, select the appropriate source database(s) for each pixel and each month, apply bias corrections and substitutions whenever needed, and print the results in a single text file. Each line corresponds to a latitude/longitude combination, and each column corresponds to a specific month. Sample results for the long-term monthly averages of AOD and AE are shown in Figure 14 and Figure 15, respectively.

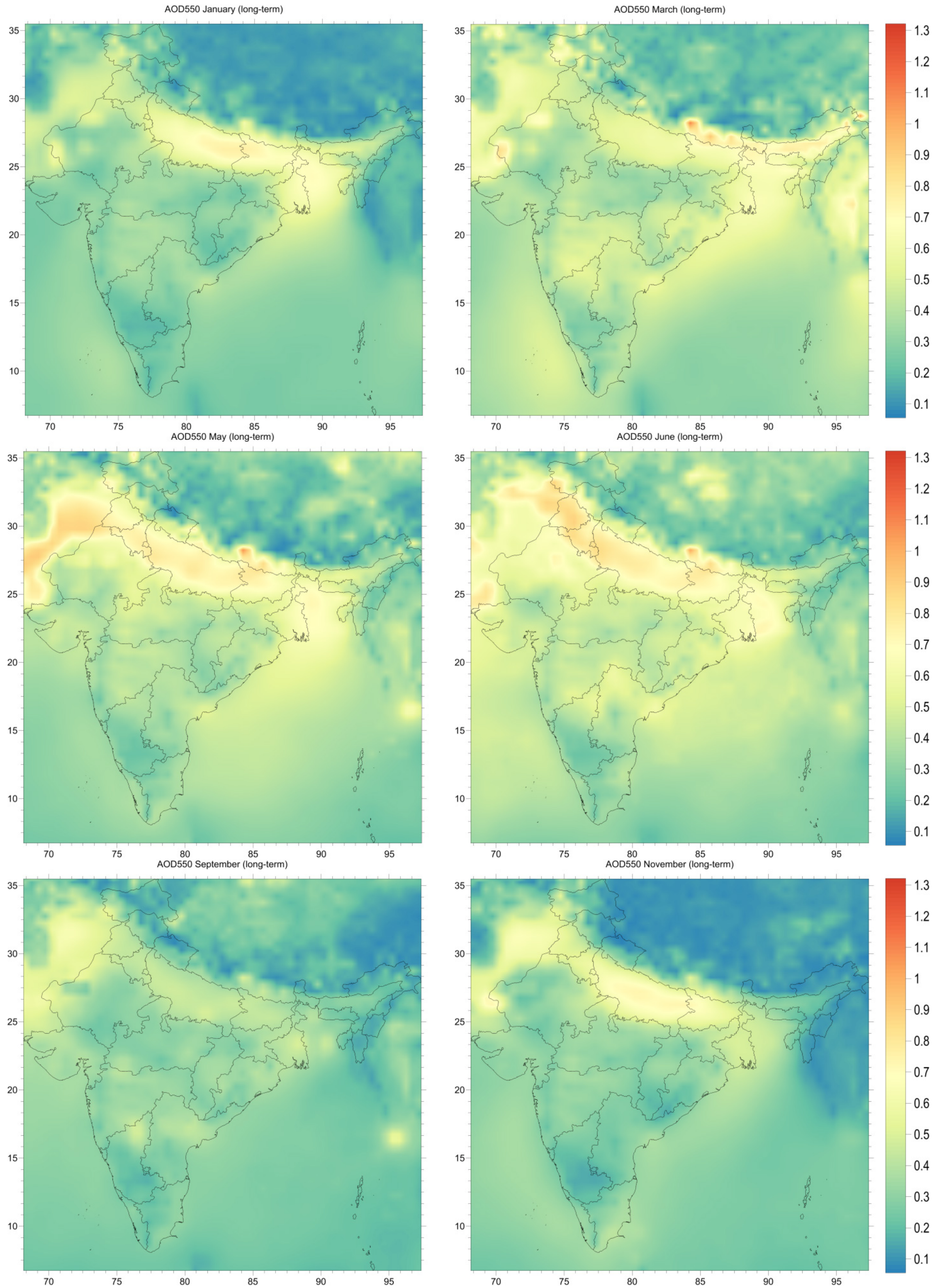


Figure 14. Long-term average AOD550 over India for January, March, May, June, September, and November

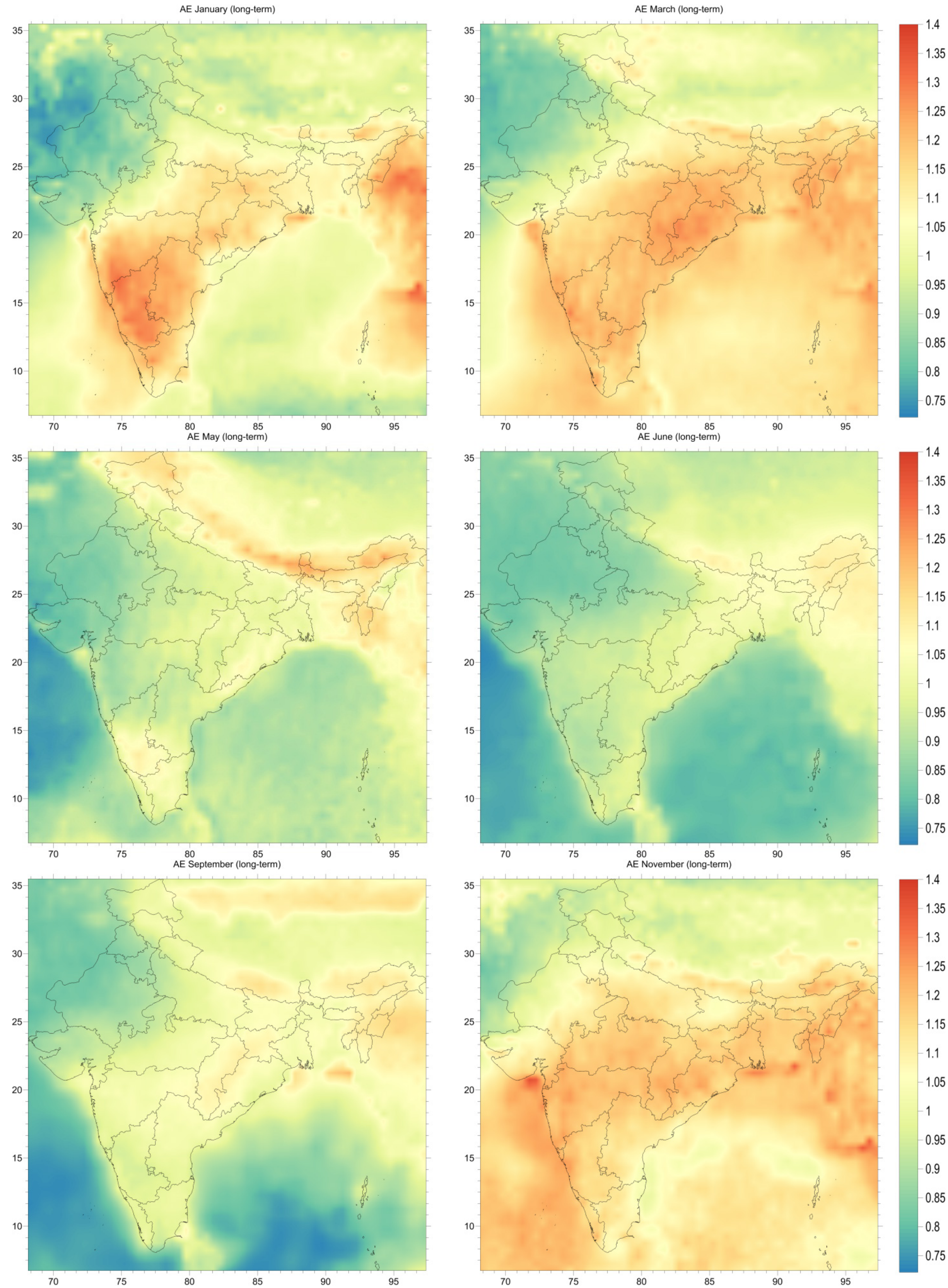


Figure 15. Long-term average AE over India for January, March, May, June, September, and November

6 Aerosol Data Set Validation

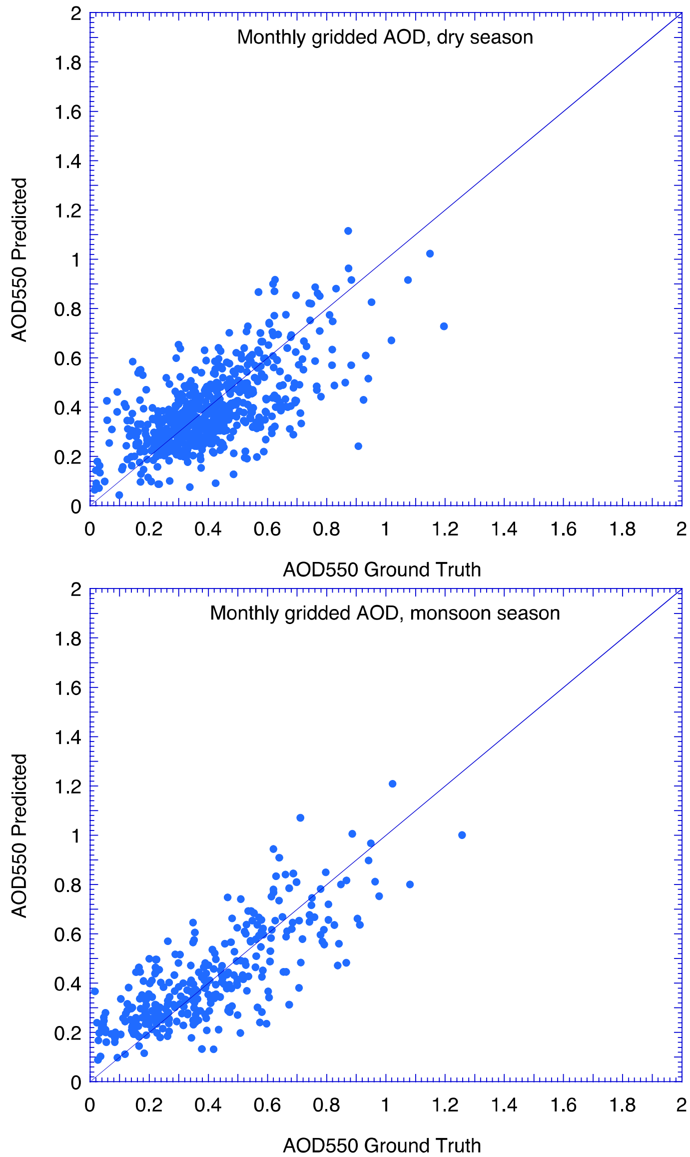
Because of the lack of ground-truth data, the performance of this new data set cannot be assessed against independent data. However, the validation undertaken here uses the *production* data set, which has a smaller size (latitudes 3° to 39° N, longitudes 65° to 100° E) and less control stations (60) than the *development* data set (latitudes 4° to 39° N, longitudes 52° to 103° E, 83 stations), but is also more centered on India.

6.1 Validation of the Monthly-Mean AOD550 Data Set

A comparison between the predicted and measured AOD550 is shown in Figure 16, which is similar to Figure 12, except that the Y-axis now indicates the production values, after correction and/or filling of the available satellite data. Although the original bias and scatter has been reduced, particularly for the monsoon season, it is obviously still far from perfect. A summary of the quality statistics (Mean Bias Deviation (MBD) and Root Mean Square Deviation (RMSD)) is provided in Table 4 for each of the three seasons and for the whole data set. Specific results for Kanpur and Anantapur are also provided separately, showing that they are comparable to the overall results. This could be expected because the ground-truth data was *not* merged with the gridded satellite data (because this would have created undesirable "Swiss-cheese" local artifacts), but used only to calibrate it on a regional basis.

Compared to the AERONET ground-truth data of AOD550, for which the uncertainty is approximately ±0.01 [6], it is estimated that the overall uncertainty in the production data set is 5 to 10 times larger. Availability of more ground-truth sites and of longer data series (so that a finer monthly analysis becomes possible) would help this situation. Still, improved satellite retrievals, particularly complemented by assimilation or optimal combinations with AOD predictions from chemical-transport models, would be the essential key to reduce the current uncertainty.

Another validation approach is to compare the proposed data to time series of ground-truth measurements at specific sites. This is advisable here at only a few ground-truth sites, because most sites have too short periods of record or too many missing months. This is done here at six sites, for which the coordinates and mean measured AOD are shown in Table 5. The time-series comparisons for these six sites are shown in Figure 17 through Figure 22. In general, the predicted AOD correctly follows the actual (large) seasonal and interannual variations that are measured locally. This is particularly the case at Karachi (Figure 19), possibly because it is a coastal location, and AOD satellite retrievals over water are more accurate than over land. Exceptions to this good agreement do occur, however, for some months at most sites. In addition to the relatively high uncertainty in the satellite retrievals already discussed, these exceptions can alternatively be explained by cloud artifacts in the satellite observations (or even in the ground sunphotometer data), by a too-low (or non-matching) number of data days in a month (particularly during the monsoon period), and in some cases by an incorrect sunphotometer calibration. The latter cause could explain the sudden downward step in measured AOD at Anantapur in mid-2006 (Figure 21). Before that time, the experimental data might well have been affected by a high bias of approximately ±0.20 in AOD550. (It is unlikely that natural causes can explain the strange behavior of the measured time series there.) Unfortunately, it is impossible to assess the historical quality of the data recorded at Anantapur or at the other sites in the Ministry of New and Renewable Energy data set because there are too many unknowns about the protocols they follow for the calibration of the instrument, the conditions of operation, the filtering of cloud-contaminated data, or the reduction and a posteriori quality control of the data.



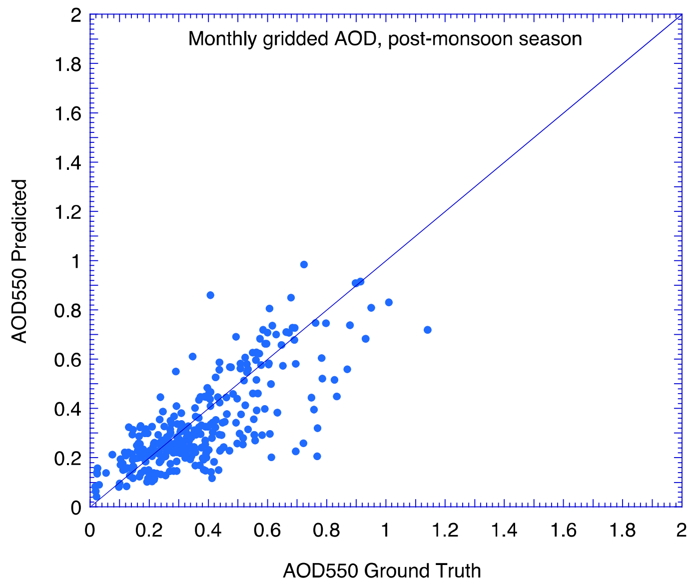


Figure 16. Predicted versus measured monthly-mean AOD550 over India for the three seasons.

Table 4. Performance Statistics of the 2000 to 2012 AOD550 Database for Two Specific Sites and for All Ground-Truth Sites

Statistic	Dry Season		Monsoon Season		Post-Monsoon Season		Year	
	Meas.	Predicted	Meas.	Predicted	Meas.	Predicted	Meas.	Predicted
Kanpur								
N	57	57	35	35	28	28	120	120
Min	0.258	0.348	0.351	0.321	0.437	0.562	0.258	0.321
Max	0.951	0.963	1.257	1.071	0.914	0.915	1.257	1.071
Mean	0.526	0.608	0.563	0.593	0.640	0.699	0.563	0.625
MBD		0.082		0.030		0.059		0.062
RMSD		0.127		0.137		0.092		0.123
Anantapur								
N	45	45	17	17	25	25	87	87
Min	0.193	0.092	0.092	0.098	0.209	0.104	0.092	0.092
Max	0.714	0.445	0.508	0.403	0.517	0.424	0.714	0.445
Mean	0.388	0.300	0.282	0.289	0.335	0.225	0.352	0.277
MBD		-0.087		0.007		-0.109		-0.075
RMSD		0.152		0.105		0.143		0.141
All Sites								
N	641	641	313	313	303	303	1257	1257
Min	0.016	0.044	0.016	0.089	0.017	0.042	0.016	0.042
Max	1.196	1.115	1.257	1.209	1.141	0.984	1.257	1.209
Mean	0.406	0.390	0.399	0.420	0.371	0.335	0.396	0.384
MBD		-0.016		0.021		-0,036		-0.012
RMSD		0.099		0.142		0.144		0.138

Table 5. Stations Used for the Validation of AOD550 Time Series

Station	Latitude	Longitude	Elevation (m)	Source	Period	Mean Annual AOD550
Gandhi College	25.871	84.128	60	AERONET	2006–11	0.539
Kanpur	26.513	80.232	123	AERONET	2001–12	0.563
Karachi	24.87	67.03	49	AERONET	2006–11	0.472
Pune	18.537	73.805	559	AERONET	2008–12	0.356
Anantapur	14.2	77.7	25	India	2001-09	0.413
Dibrugarh	27.3	94.6	111	India	2001-09	0.307

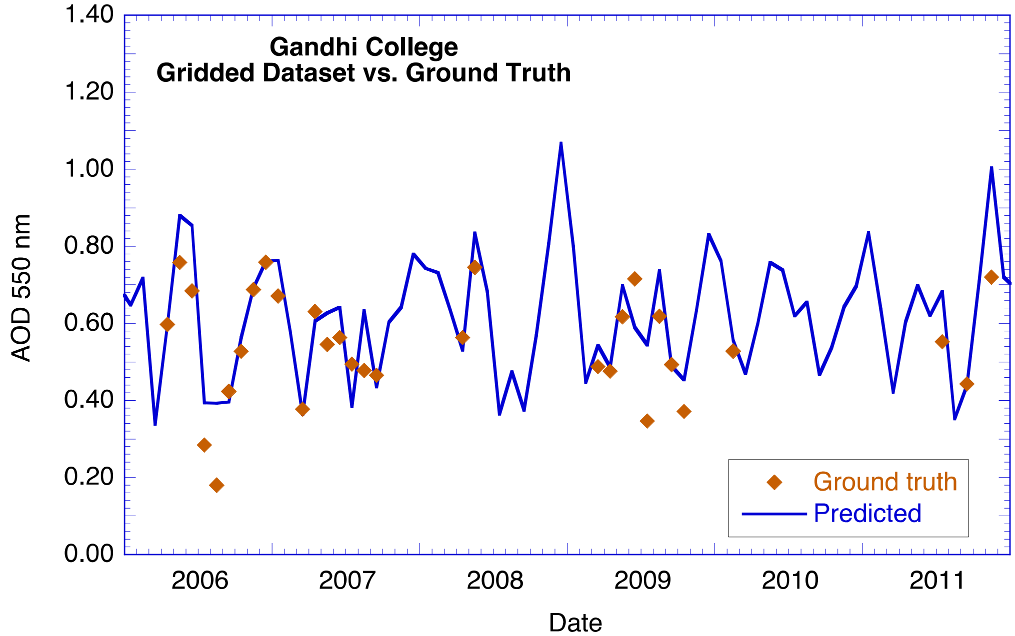


Figure 17. Measured versus predicted time series of mean monthly AOD at Gandhi College from 2006 to 2011

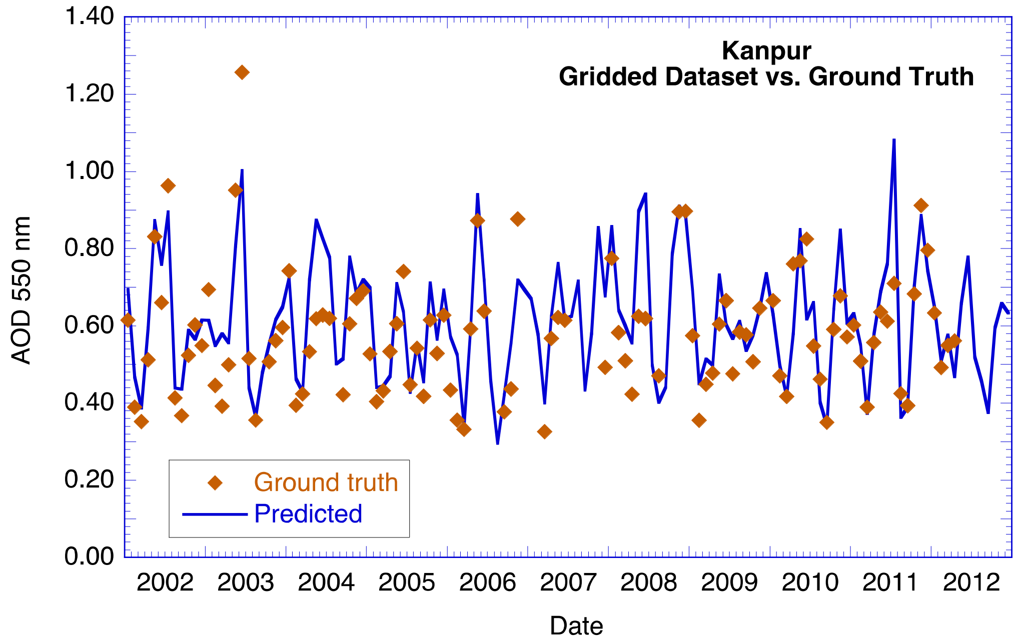


Figure 18. Measured versus predicted time series of mean monthly AOD at Kanpur from 2002 to 2012

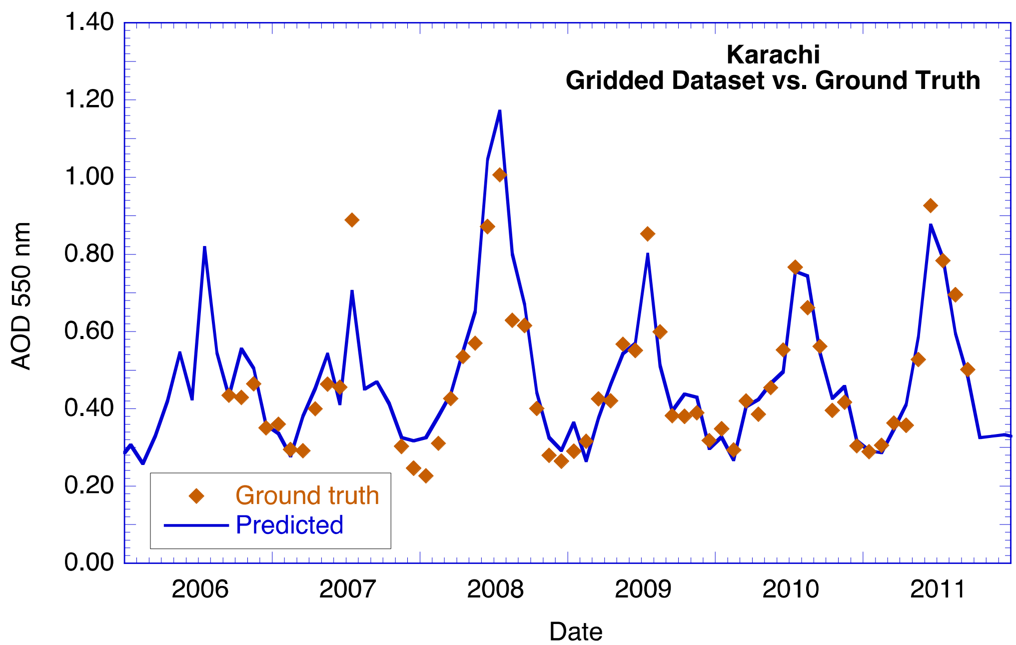


Figure 19. Measured versus predicted time series of mean monthly AOD at Karachi from 2006 to 2011

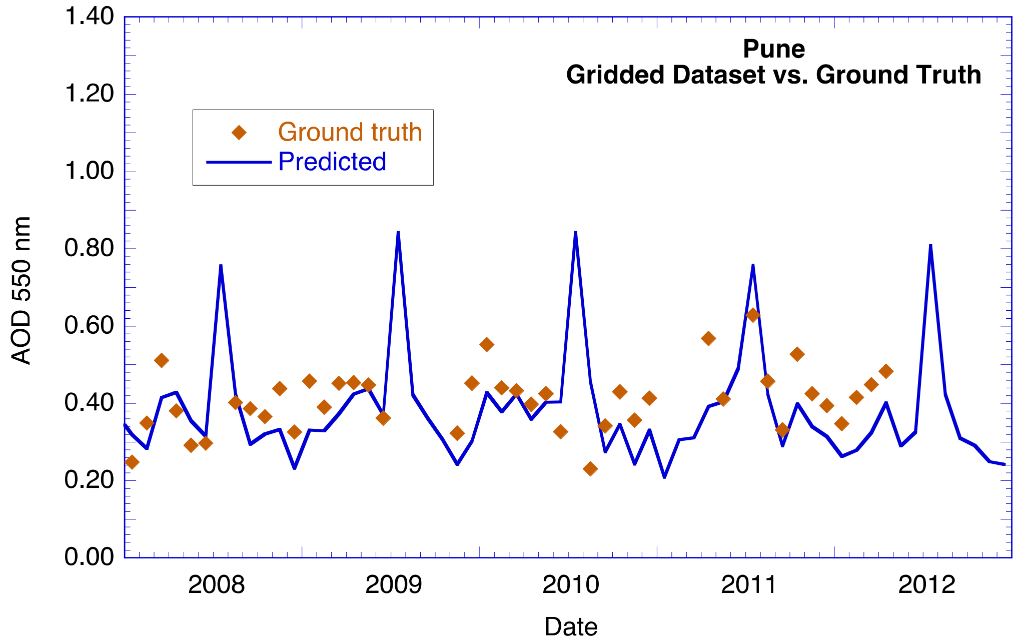


Figure 20. Measured versus predicted time series of mean monthly AOD at Pune, India, from 2008 to 2012

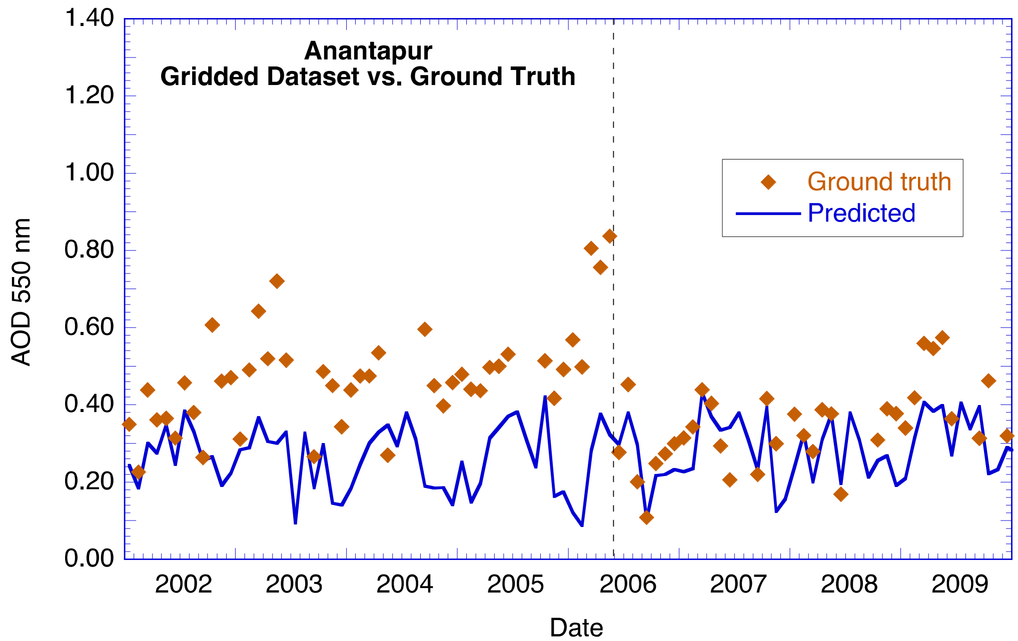


Figure 21. Measured versus predicted time series of mean monthly AOD at Anantapur from 2002 to 2009. The vertical dashed line separates the periods of high and low measured AOD.

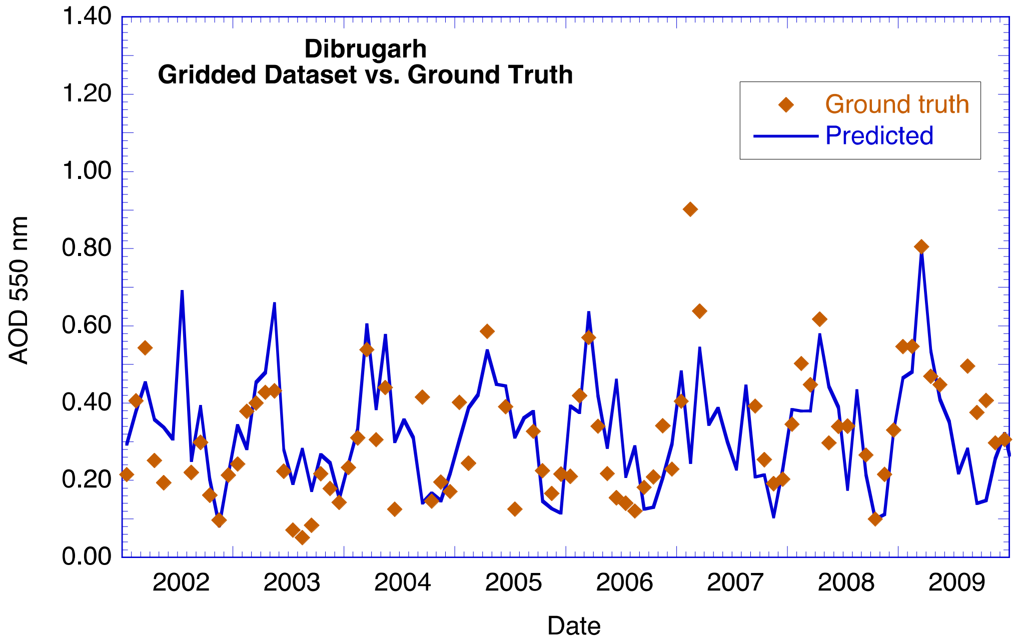


Figure 22. Measured versus predicted time series of mean monthly AOD at Dibrugarh, India, from 2002 to 2009

6.2 Validation of the Monthly-Mean AE Data Set

The monthly AE data produced here are compared to the ground-truth data shown in Figure 23, separately for the coastal and inland locations. It is obvious that the problems of the original data (noted earlier in Section 5.1) are still present here. Although the corrections applied to the original data have reduced its bias, a substantial compression of the monthly dynamic range is still present, which translates into substantial noise and a lack of values below 0.5 or above 1.3. It should be noted, however, that *in relative terms* the experimental uncertainty in AE is larger than that in AOD. This is a consequence of the way that α is calculated in practice (see Equation (2) and Figure 4): a relatively small error in the AOD from one radiometer channel may yield a noticeable change in the slope of the best-fit line, hence in α. Still, the observed large scatter in the satellite-based data can be attributed to the various inadequacies of the current retrieval algorithms, and to the diminishing AOD signal observed by any sensor at long wavelengths (itself conducive of larger errors). The overall performance statistics in Table 6 show a low bias on an annual basis.

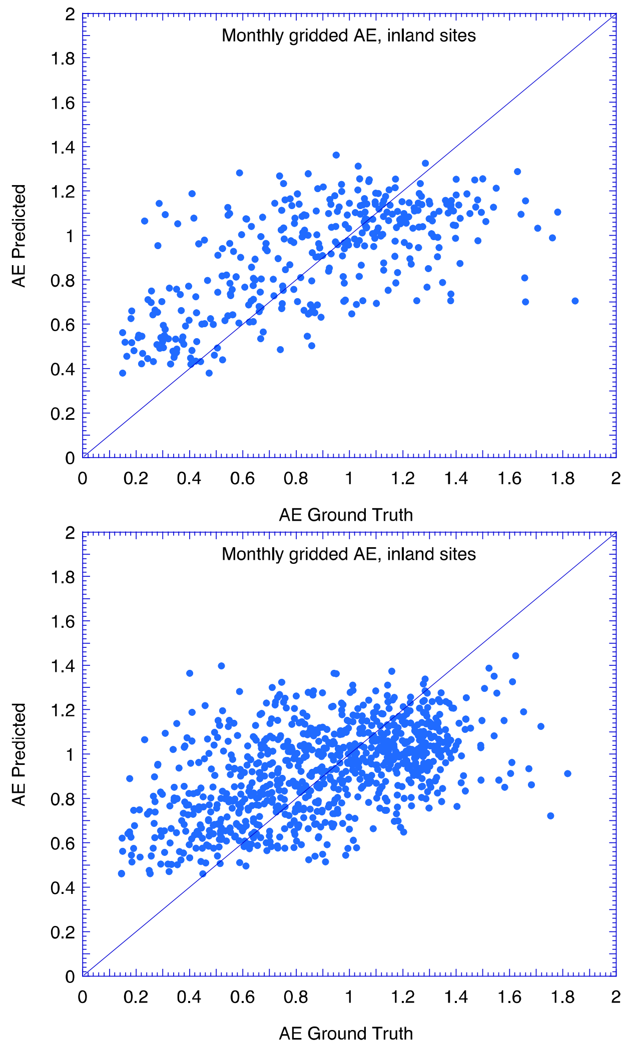


Figure 23. Predicted versus measured monthly-mean AE over India for (top) coastal areas and (bottom) continental areas

Table 6. Performance Statistics of the 2000 to 2012 Monthly AE Database for All Ground-Truth Sites

Statistic	Coastal		Continental		ALL	
	Meas.	Predicted	Meas.	Predicted	Meas.	Predicted
N	356	356	867	867	1194	1194
Min	0.149	0.380	0.145	0.461	0.145	0.380
Max	2.276	1.362	2.252	1.443	2.276	1.443
Mean	0.874	0.903	0.898	0.924	0.896	0.915
MBD		0.029		0.026		0.019
RMSD		0.302		0.292		0.290

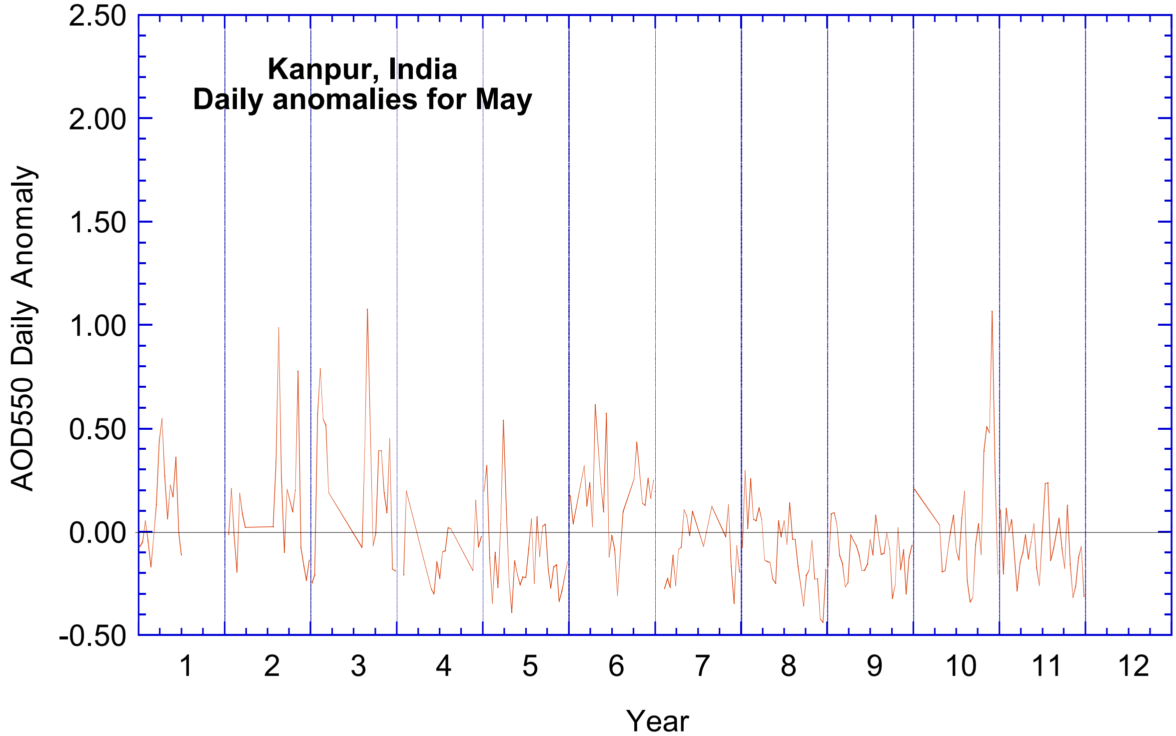
7 Statistical Correction of AOD Data

The analysis described above to produce the AOD data set was based exclusively on *mean* monthly values. This is because the raw satellite data used here are available only as monthly means. However, recent unpublished studies pointed out that using the mean monthly AOD often results in noticeable underestimations of the mean monthly direct irradiance. The National Renewable Energy Laboratory's previous irradiance maps of India were based on the assumption that the monthly *mode* AOD value should be used [10], based on the fact that the monthly modal value of AOD550 is lower than the monthly mean, and its use would thus result in higher direct irradiance. This might have been an overcorrection, however, resulting in too-high direct irradiances over India, according to some observers. A later study [11], which was devoted to the core issue of selecting the most appropriate monthly statistic for AOD, suggested that the desired *monthly effective* AOD was close to the *median* of daily values, and that the *median/mean ratio* was close to 0.8 on average over the world, with however significant scatter. Because that study was not specific to India and because a lot of unanswered questions still subsisted, a new analysis—now specific to the actual aerosol conditions of India—is undertaken here.

In India, the same as over any other geographic area, the aerosol load, and hence AOD550, varies on a daily basis. That daily variability is more or less pronounced depending on geographic area [1]. An example of daily variability of AOD550 during a specific month of the year (May) is shown in Figure 24 for two sites: Kanpur, India, and Banizoumbou, Niger, based on AERONET data obtained during a period of 11 and 12 years, respectively. In these plots, all months of May with available data are artificially concatenated, and the daily AOD550 is shown as an anomaly, i.e., the difference between the actual daily value and the long-term monthly mean. Although the long-term mean AOD550 is comparable at the two locations for that month (0.708 at Kanpur versus 0.676 at Banizoumbou), the daily variations denote a critical difference: the individual daily values show much less excursions around the mean at Kanpur than at Banizoubou. This can be explained by the relatively permanent dust and pollution load over Kanpur, in contrast to the frequent alternation of very clear days and very hazy days (as a result of transported dust or smoke from nearby sources, for example) at Banizoumbou.

The corresponding frequency distributions of AOD550 for these two sites and months are shown in Figure 25. Although both plots display the typical log-normal shape that could be expected [11], it is obvious that this shape is closer to that of a normal distribution at Kanpur than at Banizoumbou—a consequence of the difference in aerosol regime just mentioned. Another visible difference (and a direct consequence of the former observation) is that the spread between the monthly mean, median, and mode is much smaller at Kanpur than at Banizoumbou. For instance, the median/mean ratio for May is 0.931 at Kanpur versus 0.827 at Banizoumbou. Therefore, whereas at the latter site the value of that ratio is close to the world average, it is significantly higher at Kanpur.

Another aspect of the observed differences in variability between Kanpur and Banizoumbou is that the standard deviation of daily AOD values is smaller at Kanpur (0.249 in May) than at Banizoumbou (0.418). Still, the magnitude of the daily standard deviation is high at Kanpur (it even reaches 0.411 in June there). This indirectly provides an estimate of the typical error that is introduced in the predicted DNI when using monthly AOD values rather than (more desirable) daily values. Based on the analysis in [1], this propagation of error from AOD to DNI could amount to approximately ±10% to ±25% in May or approximately ±20% to ±40% in June at Kanpur, which is very significant.



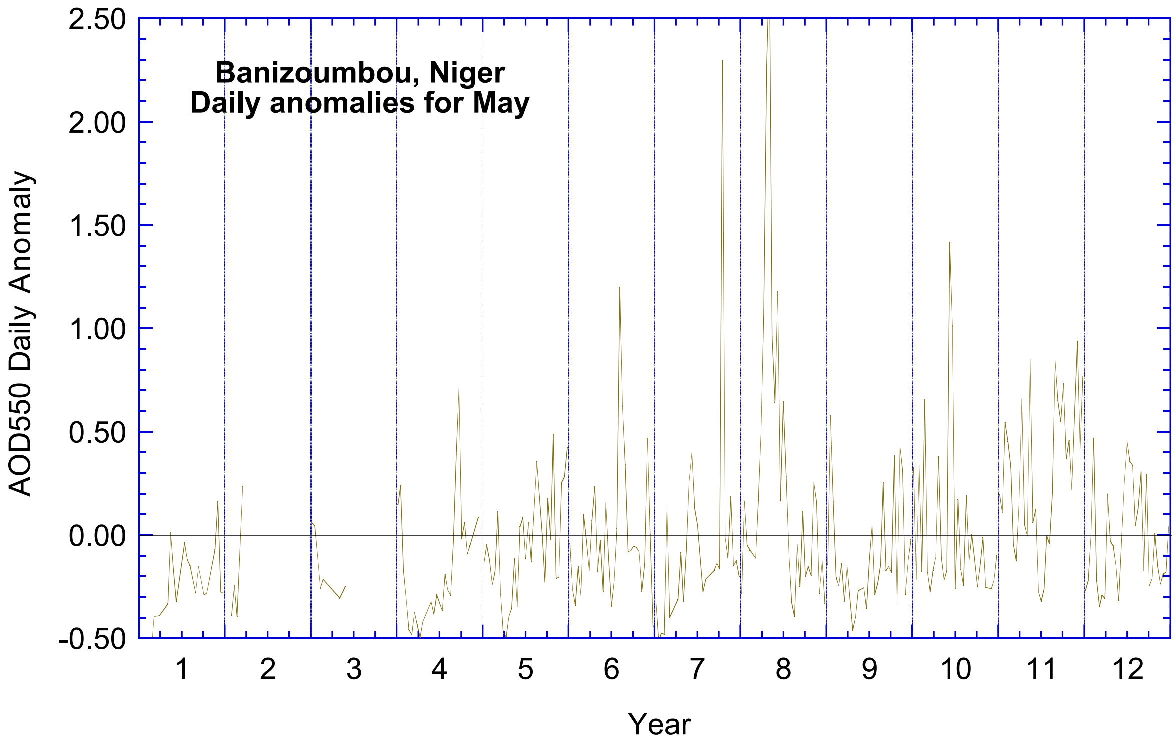


Figure 24. Daily variability of the anomaly in the AERONET AOD550 (relative to the monthly mean) at (top) Kanpur and (bottom) Banizoumbou during all months of May of the record

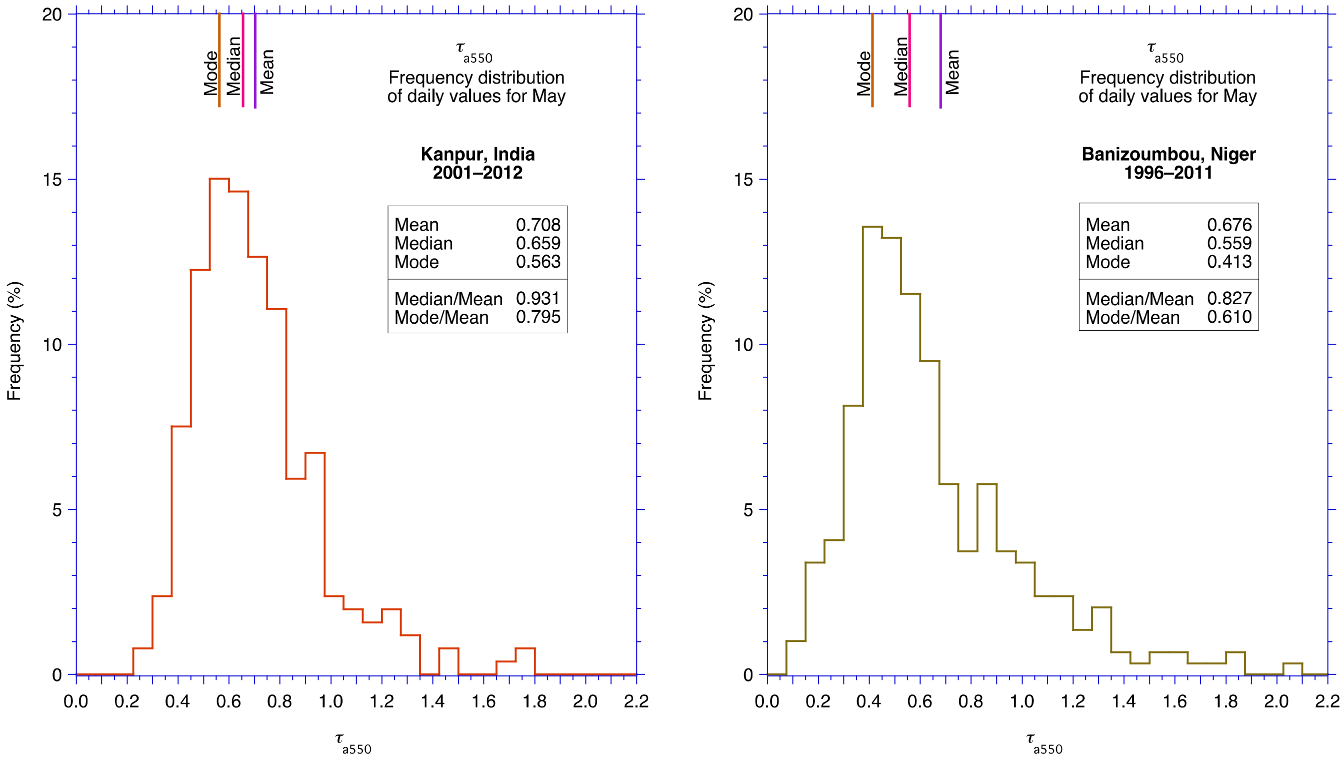


Figure 25. Frequency distributions of the (left) daily measured AOD550 at Kanpur and (right) Banizoumbou during all months of May. Monthly statistics are also indicated.

The study described in [11] has been repeated here, but only with stations in Asia and in or around India. The analysis uses daily AERONET data of AOD and precipitable water, w. The spectral AOD data points are reduced to pairs of (α, β) data using Equation (2) and

$$\beta = \tau_{a550}\, 0.55^{\alpha}. \tag{6}$$

Individual daily values of β, together with monthly-mean values of α and w, are used as inputs to the REST2 radiation model [12] to obtain the daily and monthly-mean DNI. A monthly effective β, β_{eff}, is then obtained as the single (monthly) value of β that yields the same monthly-mean DNI as previously, all other things being conserved. These calculations are repeated for each station and each calendar month with more than 40 daily data points. The present results, shown in Figure 26, confirm those in the previous study [12], to the effect that β_{eff} is very close to the monthly median β. Some scatter exists, which can be attributed to the simplicity of the method, particularly regarding the modulating effects of α and w, which are not taken into account. Interestingly, the ratio between the effective and median β is closer to one over India than at those sites in Asia that are outside India.

A more detailed, seasonal analysis has provided the ratios β_{eff}/β_{mean} for each of the three seasons identified in 5.1. These ratios are found to be 0.950 for the dry season, 0.920 for the monsoon season, and 0.915 for the post-monsoon season.

The monthly "effective" AOD550 values were obtained by multiplying the monthly-mean AOD550 values derived in Section 5.3 by these correction factors. The desired monthly AOD700 data were then obtained through Equation (3). Finally, a double linear interpolation, using the topographic correction scheme of Section 4, was applied to obtain a data set of AOD700 at 0.1° x 0.1° resolution.

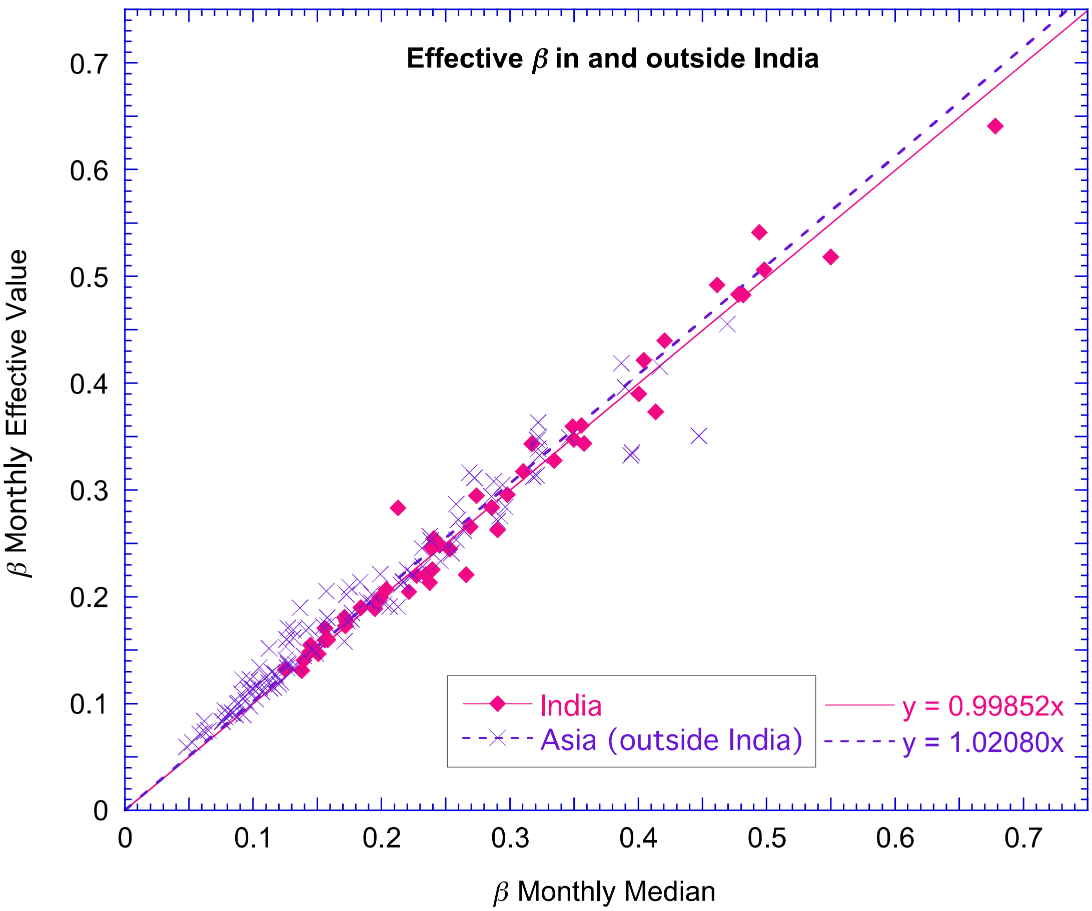


Figure 26. Calculated effective monthly values of ß over Asia, compared to the observed monthly median ß at AERONET sites

The present statistical analysis introduces significant changes compared to the previous AOD data set that was used for the National Renewable Energy Laboratory's 2010 India irradiance maps. The main difference is the much larger value of the monthly effective AOD (close to the median) compared to the monthly mode value that was used previously. That difference is illustrated in Figure 25 for one site and one month, but is a general characteristic. Considering the sensitivity of DNI on AOD [1], it can be expected that the newer modeled DNI data will be lower (typically by about 10 ±5%) than those in the 2010 release. This should actually help improve the accuracy of the newer irradiance maps by removing the high bias in the 2010 maps for DNI and GHI, which was recently found in some validation exercises, particularly in Rajasthan (e.g., [13]).

8 Conclusion and Recommendations

Various existing aerosol data bases have been tested and optimally combined to develop a data set of monthly values of AOD and AE over India. Various existing databases have been tested and optimally combined to develop a data set of monthly values of AOD at 550 nm and of AE for the Indian subcontinent and the period from 2002 to 2012. These data sets are largely based on satellite information (particularly from MODIS Terra), after proper regional scaling to match ground truth as closely as possible, and data filling with climatological values to compensate for missing data. The end product of this investigation is a data set of AOD at 700 nm with a fine spatial resolution of 0.1° x 0.1°, to be used as input of the State University of New York satellite model to derive irradiance time series and solar resource maps.

Current limitations in satellite retrieval techniques and in availability of ground-truth data result in significant scatter in the proposed data set. This issue constitutes a source of error in the eventual irradiance predictions (particularly those of DNI) that use this data set as input. Nevertheless, the present data set represents significant progress compared to previous attempts. In particular, the detailed analysis of the statistical attributes of the AOD frequency distributions that has been conducted here produced a way to correct the low bias that existed in the previous (2010) AOD data set. It is thus anticipated that the forthcoming solar irradiance maps, which will be based on the present data set, will have less bias than the 2010 irradiance maps released by the National Renewable Energy Laboratory. In particular, the DNI data series are expected to be approximately 10% ±5% lower than before, which should be enough to remove their noted high bias.

Considering these current deficiencies in the source data, it is hoped that future versions of the MISR and MODIS algorithms will improve the accuracy of their retrievals. Additionally, a large increase in ground-truth data, using high-quality sunphotometers and experimental procedures, would be necessary to develop better corrections to the satellite data, and a more convincing validation process. Similarly, high-quality DNI measurements would be necessary to assess the performance of the solar resource maps that will be developed from this study. Through an appropriate backward procedure, experimental DNI data could even be used to lower the current uncertainty in the historical time series of AOD data over India.

9 References

1. Gueymard, C.A. "Temporal Variability in Direct and Global Irradiance at Various Time Scales as Affected by Aerosols." *Solar Energy* (86) 2012; pp. 3,544–2,553.

2. Datar, S.V., et al. "Trends in Background Air Pollution Parameters Over India." *Atmospheric Environment* (30) 1996; pp. 3,677–3,682.

3. Porch, W., et al. "Trends in Aerosol Optical Depth for Cities in India." *Atmospheric Environment* (41) 2007; pp. 7,524–7,532.

4. Ramanathan, V., et al. "Atmospheric Brown Clouds: Impacts on South Asian Climate and Hydrological Cycle." *Proceedings of the National Academy of Sciencies* (102) 2005; pp. 5,326–5,333.

5. Attri, S.D. "Carbon Cycle Greenhouse Gas Measurements and Research." *Earth Observations and Earth Sciences for Societal Benefits.* New Delhi: Ministry of Earth Sciences/National Oceanic and Atmospheric Administration, 2008.

6. Holben, B.N., et al. "AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization." *Remote Sensing of Environment* (66) 1998; pp. 1–16.

7. Gueymard, C.A.; Thevenard, D. "Monthly Average Clear-Sky Broadband Irradiance Database for Worldwide Solar Heat Gain and Building Cooling Load Calculations." *Solar Energy* (83) 2009; pp. 1,998–2,018.

8. Remund, J.; Domeisen, D. *Aerosol Optical Depth and Linke Turbidity Climatology*. Final Report of IEA SHC Task 36. Bern, Switzerland: Meteotest, 2010.

9. Levy, R.C., et al. "Global Evaluation of the Collection 5 MODIS Dark-Target Aerosol Products Over Land." *Atmospheric Chemistry and Physics* (10) 2010; pp. 10,399–10,420.

10. Gueymard, C.A.; George, R. "Gridded Aerosol Data for Improved Direct Normal Irradiance Modeling: The Case of India." *Procedings of Solar 2011 Conference.* Raleigh, N.C.: American Solar Energy Society, 2011.

11. Gueymard, C.A.; Sengupta M. "Improving Modeled Solar Irradiance Historical Time Series: What is the Appropriate Monthly Statistic for Aerosol Optical Depth?" *Proceedings of ASES Conference, World Renewable Energy Forum*. Denver, CO: 2012.

12. Gueymard, C.A. "REST2: High-Performance Solar Radiation Model for Cloudless-Sky Irradiance, Illuminance, and Photosynthetically Active Radiation—Validation with a Benchmark Dataset." *Solar Energy* (82) 2008; pp. 272–285.

13. Meyer, R.; Chhatbar, K.; Schwandt, M. *Solar Resource Assessment at MNRE Site in Rajasthan.* Technial Report for the Ministry of New and Renewable Energy of India. Suntrace GmbH, November 2012.