

## RESEARCH ARTICLE

10.1002/2013JD020538

This article is a companion to *Li et al.* [2013] doi:10.1002/jgrd50686.

## Key Points:

- CPCA successfully extracts coupled modes in multiple data sets
- The comparison suggests good agreement among the three data sets
- Differences in the CPCA modes are discussed and explained

## Correspondence to:

J. Li,  
jl2862@columbia.edu

## Citation:

Li, J., B. E. Carlson, and A. A. Lacis (2014), Application of spectral analysis techniques in the intercomparison of aerosol data: Part III. Using combined PCA to compare spatiotemporal variability of MODIS, MISR, and OMI aerosol optical depth, *J. Geophys. Res. Atmos.*, 119, 4017–4042, doi:10.1002/2013JD020538.

Received 9 JUL 2013

Accepted 2 DEC 2013

Accepted article online 9 DEC 2013

Published online 7 APR 2014

# Application of spectral analysis techniques in the intercomparison of aerosol data: Part III. Using combined PCA to compare spatiotemporal variability of MODIS, MISR, and OMI aerosol optical depth

Jing Li<sup>1,2</sup>, Barbara E. Carlson<sup>1</sup>, and Andrew A. Lacis<sup>1</sup>

<sup>1</sup>NASA Goddard Institute for Space Studies, New York, New York, USA, <sup>2</sup>Department of Applied Physics and Applied Math, Columbia University, New York, New York, USA

**Abstract** Satellite measurements of global aerosol properties are very useful in constraining aerosol parameterization in climate models. The reliability of different data sets in representing global and regional aerosol variability becomes an essential question. In this study, we present the results of a comparison using combined principal component analysis (CPCA), applied to monthly mean, mapped (Level 3) aerosol optical depth (AOD) product from Moderate Resolution Imaging Spectroradiometer (MODIS), Multiangle Imaging Spectroradiometer (MISR), and Ozone Monitoring Instrument (OMI). This technique effectively finds the common space-time variability in the multiple data sets by decomposing the combined AOD field. The results suggest that all of the sensors capture the globally important aerosol regimes, including dust, biomass burning, pollution, and mixed aerosol types. Nonetheless, differences are also noted. Specifically, compared with MISR and OMI, MODIS variability is significantly higher over South America, India, and the Sahel. MODIS deep blue AOD has a lower seasonal variability in North Africa, accompanied by a decreasing trend that is not found in either MISR or OMI AOD data. The narrow swath of MISR results in an underestimation of dust variability over the Taklamakan Desert. The MISR AOD data also exhibit overall lower variability in South America and the Sahel. OMI does not capture the Russian wild fire in 2010 nor the phase shift in biomass burning over East South America compared to Central South America, likely due to cloud contamination and the OMI row anomaly. OMI also indicates a much stronger (boreal) winter peak in South Africa compared with MODIS and MISR.

## 1. Introduction

The study of atmospheric aerosols is gaining increased interest from both the scientific and policy community, due to their importance in affecting the energy budget and thus contributing to climate change [Intergovernmental Panel on Climate Change (IPCC), 2007]. In recently decades, many satellite sensors have been developed to monitor the global distribution and temporal variation of aerosols. For example, Moderate Resolution Imaging Spectroradiometer (MODIS) and Multiangle Imaging Spectroradiometer (MISR) are two dedicated instruments that retrieve Aerosol Optical Depth (AOD) and size information at visible to near-IR wavelengths. The Ozone Monitoring Instrument (OMI) makes use of its UV channels to provide measurements of AOD as well as aerosol absorption. The abundance of remote sensing data sets of aerosol properties has greatly boosted the rapid progress in aerosol science. Satellite measured aerosol properties are widely used to validate and constrain aerosol parameterization in General Circulation Models (GCMs), in order to provide more accurate estimations of aerosol forcing. For example, *Kinne et al.* [2003] compared monthly mean aerosol climatology from different models, satellite and ground observations. *Liu et al.* [2006] assessed the GISS GCM aerosol climatology using Level 3 monthly mean products from multiple sensors and Aerosol Robotic Network (AERONET) [Holben et al., 1998]. It is thus essential to understand the reliability of different data sets to represent the spatial and temporal variability of the true aerosol field, especially those associated with source regions, seasonal and interannual variability and trends.

However, aerosol properties are highly variable on both the global and regional scale, due to their different emission sources and emission processes, as well as their interaction with meteorological conditions. As a result, simultaneous comparison of the spatial and temporal variability between different data sets is not an

easy task. Spectral analysis, which aims at decomposing the high-dimensional data matrix and extracting major variability, is an effective technique for this purpose. For example, principal component analysis (PCA) has been extensively used to study climate variability such as El Niño–Southern Oscillation [Kawamura, 1994; Wang and An, 2005; Kao and Yu, 2009], North Atlantic Oscillation [Slonosky et al., 2000; Jung et al., 2003], and the Madden-Julian Oscillation [Kessler, 2001]. With respect to aerosol data, Li et al. [2009] used rotated empirical orthogonal function (EOF) analysis to identify dust and biomass burning source regions. In particular, Li et al. [2013] compared the leading EOF modes of four independent AOD measurements and found good agreement across the data sets in representing major aerosol variability on a global scale.

However, with multiple data sets available, spectrally decomposing the individual data sets and comparing the modes may not be the most efficient means of comparison. On one hand, due to differences in instrument design (which may lead to differences in sampling), differences in cloud screening, and differences in the retrieval algorithm used to process the data from each satellite sensor, the same aerosol variability may appear in different EOF modes, or may even be split into several modes, which makes parallel comparison and evaluation rather difficult. On the other hand, since each mode of variability is represented by the product of a spatial pattern and a time series, it is not possible to make quantitative comparisons on either the spatial distribution or temporal variation as noted by Li et al. [2013]. In this study, we adopt a different approach—the combined principal component analysis (CPCA) to overcome the above two shortcomings associated with the traditional EOF approach. The CPCA is a modification of the EOF by joining different fields into one large data matrix and extracting the common modes of variability from this combined field. In this way, the leading modes maximize the variance explained of the sum of all fields. The amount of agreement is easily found by comparing the spatial patterns of each mode, which is the primary focus when using the data to improve aerosol models. Moreover, for each mode of variability, the spatial patterns for the different data sets are associated with a common time series; therefore, quantitative comparison becomes possible by examining the differences in the spatial patterns. This further offers insights into the capability of each data set in representing the temporal variability.

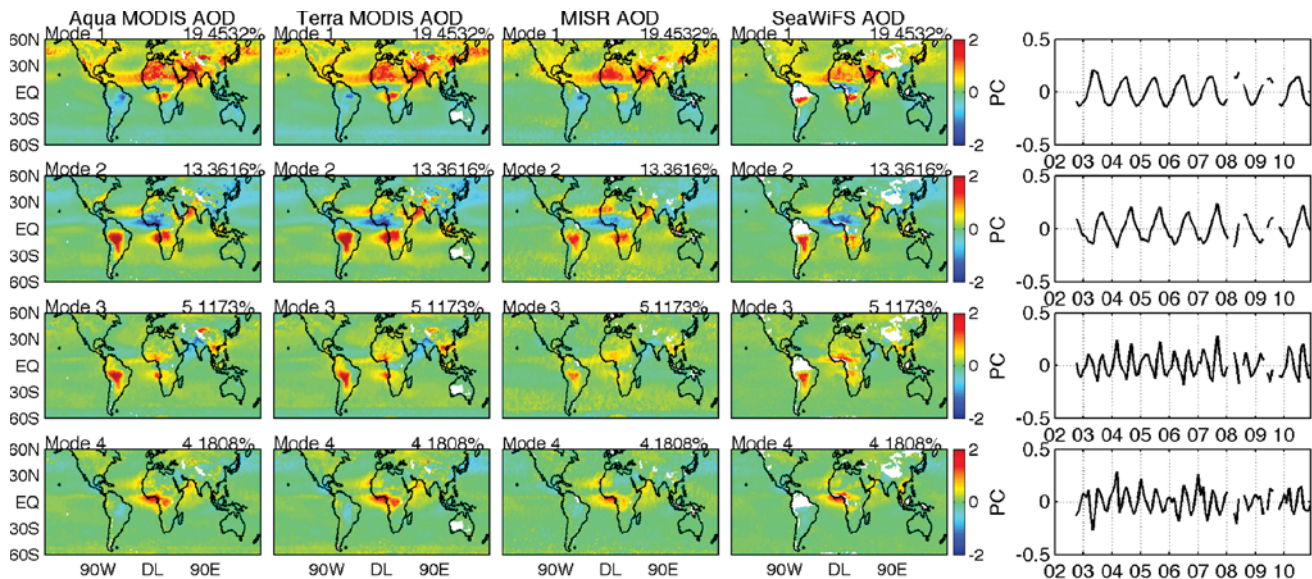
The CPCA method was proposed as early as 1967 by Kutzbach [1967]. Bretherton et al. [1992] compared this technique with several other spectral analysis techniques in finding coupled modes. Wallace et al. [1992] further applied this analysis to investigate the relationship between sea surface temperature and 500 mbar height anomalies. In spite of its early development, this method is not extensively used due to two major limitations: (1) Combining the data implicitly assumes equal weight of each field, but simple normalization may not always satisfy this constraint for fields with very different units and magnitudes of variability; (2) Combining the data also requires the same or close spatial representativeness of the individual data fields. Therefore, it is not suitable for two fields with different observational resolution, such as satellite and surface measurements. However, this method is particularly suitable for the question addressed here—parallel comparison of multiple aerosol optical depth observations—because we are focusing on multiple satellite measurements of the same quantity. Regardless of instrumental and algorithmic differences, the measured quantity— aerosol optical depth—should have the same degree of magnitude and variability that is independent of measuring technique; therefore, there is no question of normalization or representation error. The differences resulting from the comparison should be a reflection of the capability of each sensor in representing the spatial and temporal variability of the real aerosol optical depth field.

In the present paper, we present the application of the CPCA method to three independent AOD data sets: Aqua-MODIS, MISR, and OMI. The analysis is performed on both global data and several representative regions. The study shows that CPCA successfully extracts the dominant modes of variability from all data sets, and associates the comparison with well known aerosol regions/types. The results confirm the agreement of the data sets in representing the variability of many aerosol types and source regions, as well as identifies potential problems in the individual data sets. Section 2 describes the data sets used in this study. Detailed mathematical description of the CPCA method is given in section 3. Section 4 presents global and regional comparisons of leading CPCA modes. Finally, section 5 summarizes the advantage of this technique and our findings from this comparison.

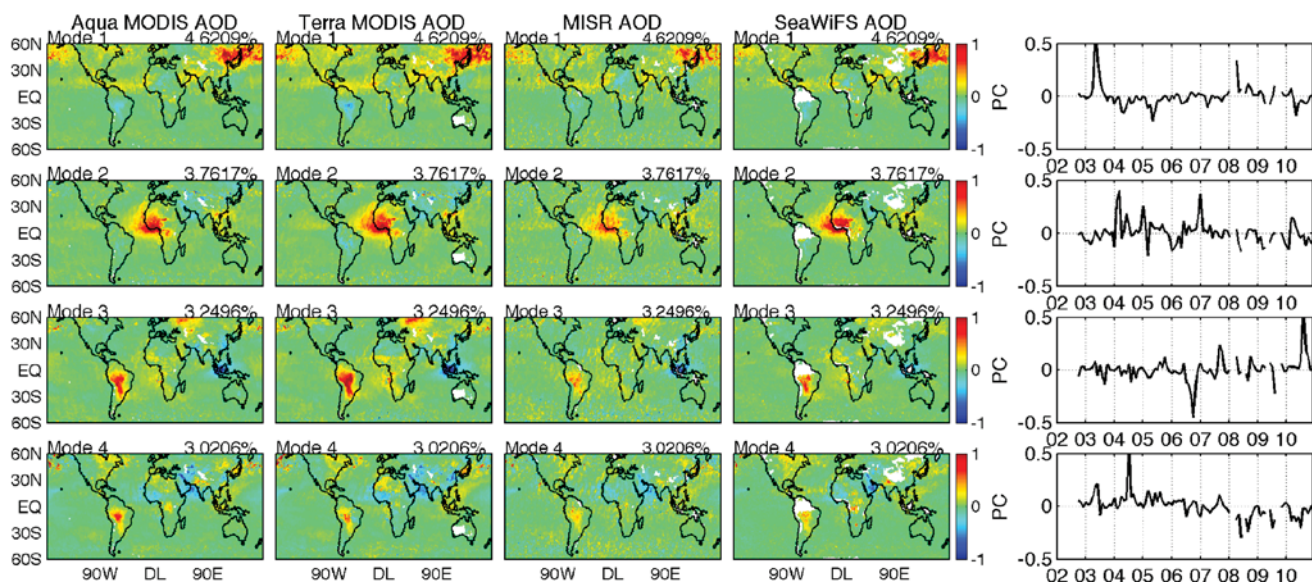
## 2. Data

Different from Part I [Li et al., 2013], in this study, we use Level 3 monthly mean AOD products from Aqua-MODIS, MISR, and OMI UV algorithm. Two data sets used in Part I—Terra-MODIS and Sea-viewing Wide Field-of-view

a



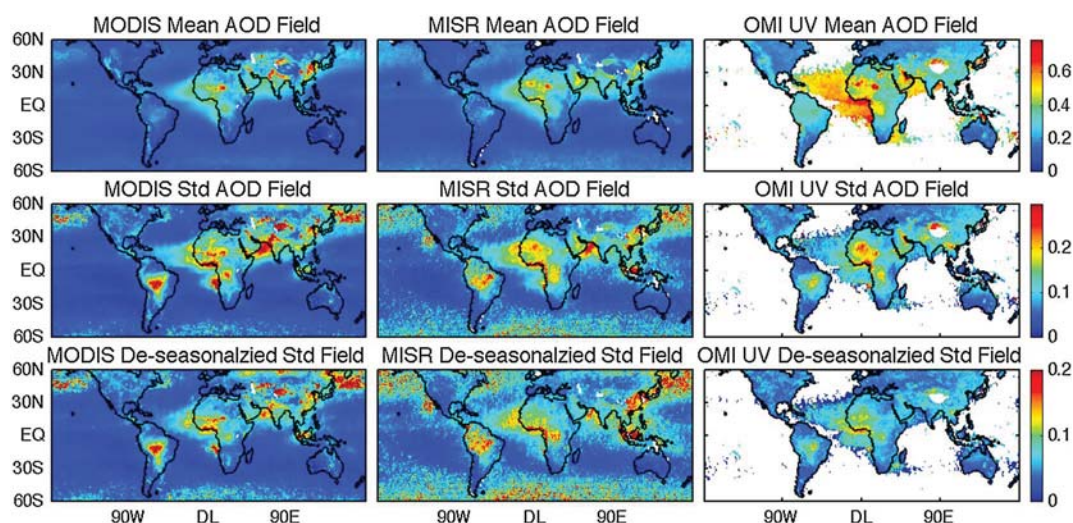
b



**Figure 1.** The first four CPCA modes for the four data sets used in Part I of the series [Li et al., 2013]. (a) CPCA modes of the full data set. (b) CPCA modes of the anomaly data set. The newly released SeaWiFS version 4 data are used. These four modes resemble the four EOF modes of individual data sets shown in Part I. Difference between SeaWiFS V3 and SeaWiFS V4 data are not significant.

Sensor (SeaWiFS)—are dropped in this study due to (1) the concern that the similarity of the Aqua- and Terra-MODIS data may put too much weight on the variance of MODIS AOD in the combined field, and the fact that the Terra-MODIS deep blue record is not as complete as Aqua-MODIS and (2) we noticed that the SeaWiFS AOD has a lot of missing data over the tropical regions, while its influence on the global analysis presented in Part I is minimal, the missing data cause problems in regional analyses by decreasing the weight of the SeaWiFS data set. Moreover, OMI data are added to better illustrate the utility of the CPCA approach in comparing of multiple satellite data sets. Nonetheless, we also present the global CPCA results of the original four data sets in Figure 1 and show that the CPCA modes are quite similar to the EOF modes of each individual data set [Li et al., 2013, Figures 2 and 4]. The study period for this paper is January 2005 to December 2012.





**Figure 2.** The averaged mean field, standard deviation field, and standard deviation field without the seasonal cycle of MODIS, MISR, and OMI over the entire study period. OMI seems to have an overall high bias over North Africa and the Atlantic. It also shows an overall lower variability. MODIS have higher variability over South America, North India, and North Pacific.

The MODIS instrument is a single-view imager with a swath width of 2330 km and near global coverage of 2 days. This high sampling frequency captures most of aerosol variability and microphysics properties. The AOD data used here is from the MODIS on Aqua platform and belongs to Collection 051, available from [ftp://ladsweb.nasacom.nasa.gov/allData/51/MYD08\\_M3](ftp://ladsweb.nasacom.nasa.gov/allData/51/MYD08_M3). To ensure global coverage, we combine ocean, dark target land [Levy *et al.*, 2010; Remer *et al.*, 2008], and deep blue retrievals [Hsu *et al.*, 2004, 2006]. The data resolution is  $1^\circ \times 1^\circ$  and QA weighted averages are selected (variable name with “\*”) [Hubanks *et al.*, 2008]. While the consistency and quality of combining dark target AOD (DT\_AOD) and deep blue AOD (DB\_AOD) retrievals have not been fully evaluated, global coverage is essential in the CPCA analysis and the DB\_AOD product covers many important aerosol source regions. The combination of DT\_AOD and DB\_AOD data follows the procedure described by Levy *et al.* [2013], which determines the selection of DT\_AOD or DB\_AOD according to the MODIS normalized difference vegetation index product and is used to create the new merged data set in MODIS Collection 6 product. Also, to match OMI AOD retrievals, the MODIS data have been interpolated to 500 nm using measurement at 470 nm and 660 nm according to the Angstrom relationship.

The MISR is a multiangle sensor with nine push broom cameras on the EOS Terra platform. The zonal overlap of the common swath of all nine cameras is at least 360 km in order to provide multiangle coverage in 9 days at equator, and 2 days at poles [Diner *et al.*, 1998]. Compared to MODIS, the multiangle view of MISR performs better over bright surfaces, while its lower sampling may not fully resolve small-scale variability. In this study we use version 31 Level 3 gridded monthly products, available from <http://eosweb.larc.nasa.gov>. The original  $0.5^\circ \times 0.5^\circ$  data resolution has been rescaled to  $1^\circ \times 1^\circ$ . The rescaling is performed by assigning equal weights to each subgrid, and the final  $1^\circ \times 1^\circ$  grid is considered valid only when more than half of the subgrids have valid data. The data is also interpolated to 500 nm using measurements at 446 nm, 555 nm, 672 nm, and 865 nm.

The OMI sensor [Levelt *et al.*, 2006] on the EOS Aura satellite has been providing global aerosol measurements since October 2005. It is also a wide-view imager with a swath width of 2600 km and produces daily global coverage. The AOD data used here are derived from the UV algorithm (OMAERUV, OMI near-UV aerosol retrieval algorithm) [Torres *et al.*, 2007]. The AOD is primarily retrieved at 388 nm using the instrument’s two near-UV channels, and the reported 500 nm AOD is converted according to the spectral dependence of the assumed aerosol model [Torres *et al.*, 2007; Ahn *et al.*, 2008]. While the reliability of the 500 nm AOD is affected by aerosol model assumptions, comparison with MODIS and MISR shows reasonable agreements [Ahn *et al.*, 2008]. Here we use Collection 003 data at  $1^\circ \times 1^\circ$ , available from Goddard Earth Sciences Data and Information Services Center (<http://mirador.gsfc.nasa.gov/>). Note that as the current OMAERUV algorithm does not explicitly account for ocean color effects and retrievals over ocean are limited to only high AOD conditions, it

is only used over land and in regional analyses. The wide swath of OMI provides daily global coverage. However, its relatively large footprint ( $13 \times 24 \text{ km}^2$  at nadir) makes cloud contamination a more serious issue in OMI retrievals [Torres *et al.*, 2007].

In addition to the satellite sensors, ground-based measurements at a few stations from the Aerosol Robotic Network (AERONET) [Holben *et al.*, 1998] are also used to facilitate the interpretation of the results. The AERONET measurements are highly accurate and are usually considered as ground truth in satellite data validation. The data used here are Level 2.0 monthly mean AOD and single scattering albedo (SSA) products.

### 3. Methodology

The combined principal component analysis (CPCA) is originally developed to find the relationship between different data fields through the decomposition of the temporal covariance matrix of the joint field. It was introduced by Kutzbach [1967], who used it to analyze sea level pressure, surface temperature, and precipitation fields. Despite its early development, it is not frequently used in data analysis due to the strong underlying assumptions of this method. Specifically, equal weight is assumed for each field in the combining procedure. This assumption is usually addressed by normalization. However, for variables with very different units and variability, effective normalization may be difficult to realize. Moreover, for data sets with different spatial grids, e.g., satellite observations and scattered surface measurements, simply combining the data will introduce errors due to their different spatial representativeness (scale). In this data intercomparison study, however, the CPCA method is very suitable, because the different data sets used in the analysis are measurements of the same variable and with approximately the same spatial representativeness, which satisfies both of the above assumptions. Note that the spatial representativeness referred here is not specific to any sensor, but to the AOD quantity itself, which is closely related to, or determined by, its scale of spatial and temporal variability.

Mathematically, CPCA finds the modes that maximize the variance explained by the sum of the elements in the combined fields [Bretherton *et al.*, 1992]. Assuming the three AOD data fields that we have are  $X$ ,  $Y$ , and  $Z$ , all with size  $N \times M$ , where  $N$  is the number of locations and  $M$  is the number of observations. Before the analysis, each row of  $X$ ,  $Y$ , and  $Z$  is centered by removing the mean. The combined data matrix is constructed as

$$D = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

whose size is  $3N \times M$ . The temporal covariance matrix is

$$C = \frac{1}{M-1} DD^T$$

The spatial patterns or empirical orthogonal functions (EOFs) are found by determining the eigenvectors of  $C$ :

$$C = E \Lambda E^T$$

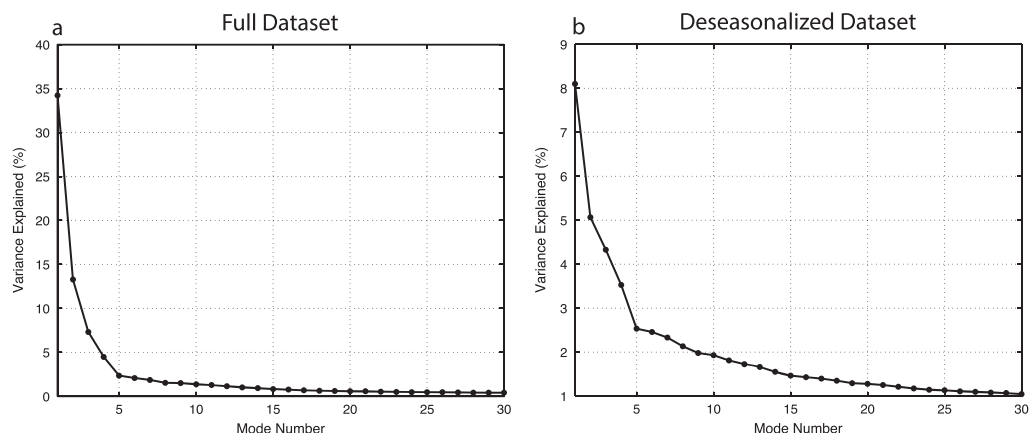
where  $E$  is a  $3N \times 3N$  orthogonal matrix.  $\Lambda$  is a diagonal matrix whose elements are the eigenvalues of  $C$ , sorted in descending order. The first  $N$  elements of each column of  $E$  form the EOFs of  $X$ , the next  $N$  elements form the EOFs of  $Y$  and the last  $N$  elements are the EOFs of  $Z$ . The expansion coefficients of each EOF mode, or principal components (PCs) are determined by projecting the data matrix on to each EOF as

$$\vec{p}_i = D^T \vec{e}_i$$

It can be shown that the PCs are also orthogonal and the elements in the diagonal matrix  $\Lambda$  are their variances. Let  $\lambda_i$  denote the  $i$ th element of  $\Lambda$ , the fraction of variance (FV) explained by the  $i$ th mode is

$$FV = \frac{\lambda_i}{\sum \lambda_i}$$

In this study, we perform the analysis on both the full data set (after removing the mean) to examine seasonal variability and on the AOD anomaly data (after removing the mean and multiyear averaged seasonal cycle) to examine interannual variability. The method to remove mean seasonal cycle is the same as Li *et al.* [2013]. In Figure 2, we briefly describe the mean field, standard deviation field which represents all variability and the anomaly field which represents interannual variability of the three data sets. All fields are averaged over the



**Figure 3.** Variance explained by the first 30 CPCA mode of globe data over land for (a) full data set and (b) anomaly data set. Because after Mode 4, the variances explained by the high-order modes are significantly smaller than the first four modes, in the following analysis only the first four modes are presented for global results.

entire study period. Figure 2 shows reasonable agreement across the three data sets in all three fields. Globally, the dominant aerosol signals come from North Africa, Arabian Peninsula, India, and East Asia. For the mean field, OMI shows an overall high bias over dust-dominant regions of North and West Africa and North Atlantic. MODIS has a slightly high bias over North India, and East Asia. In the seasonal and interannual variability fields, South America, South Africa, and Indonesia also show up with strong signals for MODIS and MISR. However, the OMI variability for these three regions appears lower. Also compared with MISR and OMI, MODIS has an overall higher variability for South America, Central Asia, India, and North Pacific. In the next section, we will compare the variability from different aerosol regimes in greater detail by decomposing the combined field of variability.

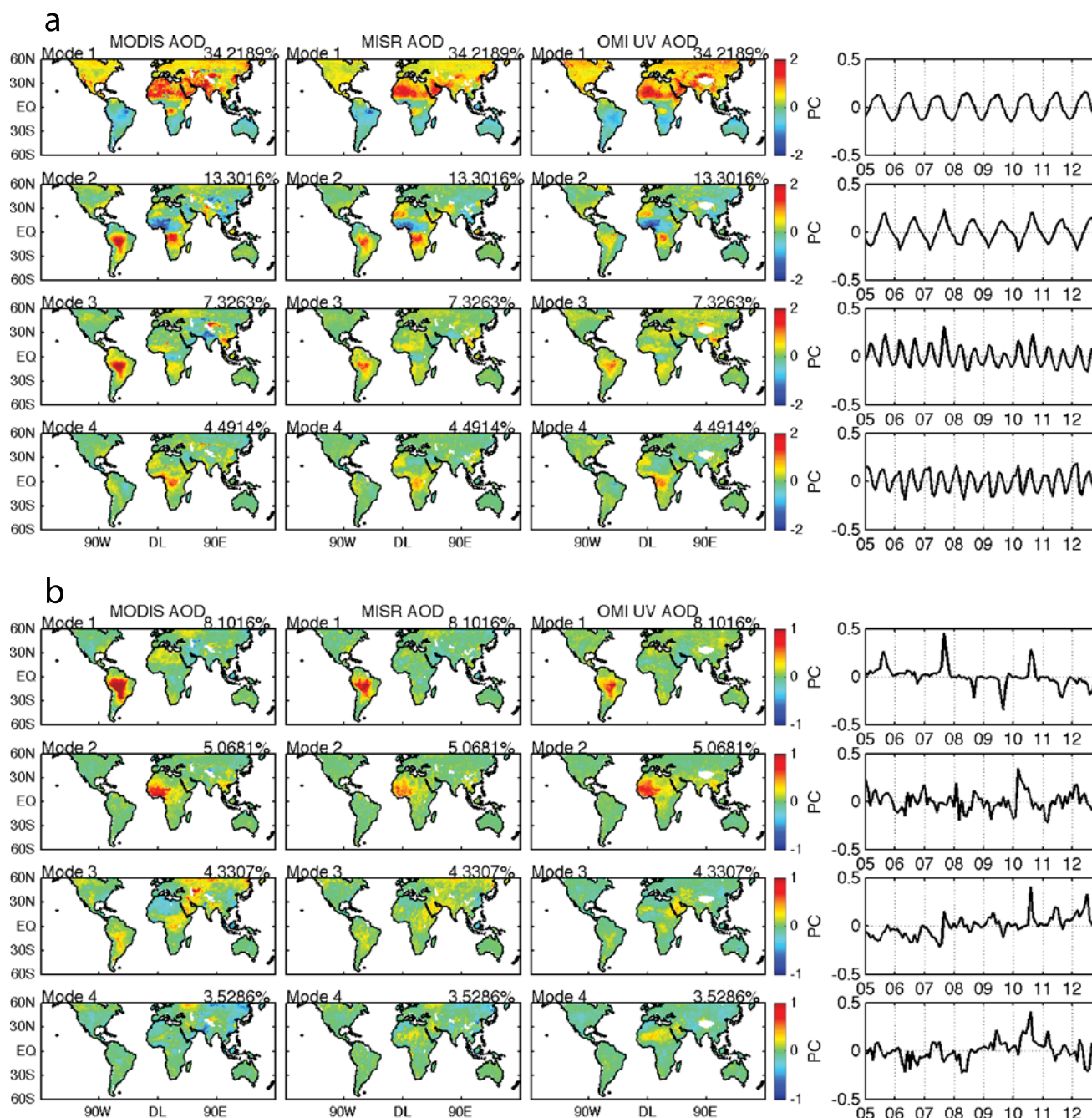
## 4. Results

In this section, we present comparison between major CPCA modes from the three data sets. The analysis is first performed on global scale over land and ocean, and then on three representative regions.

### 4.1. Global Analysis

#### 4.1.1. Land

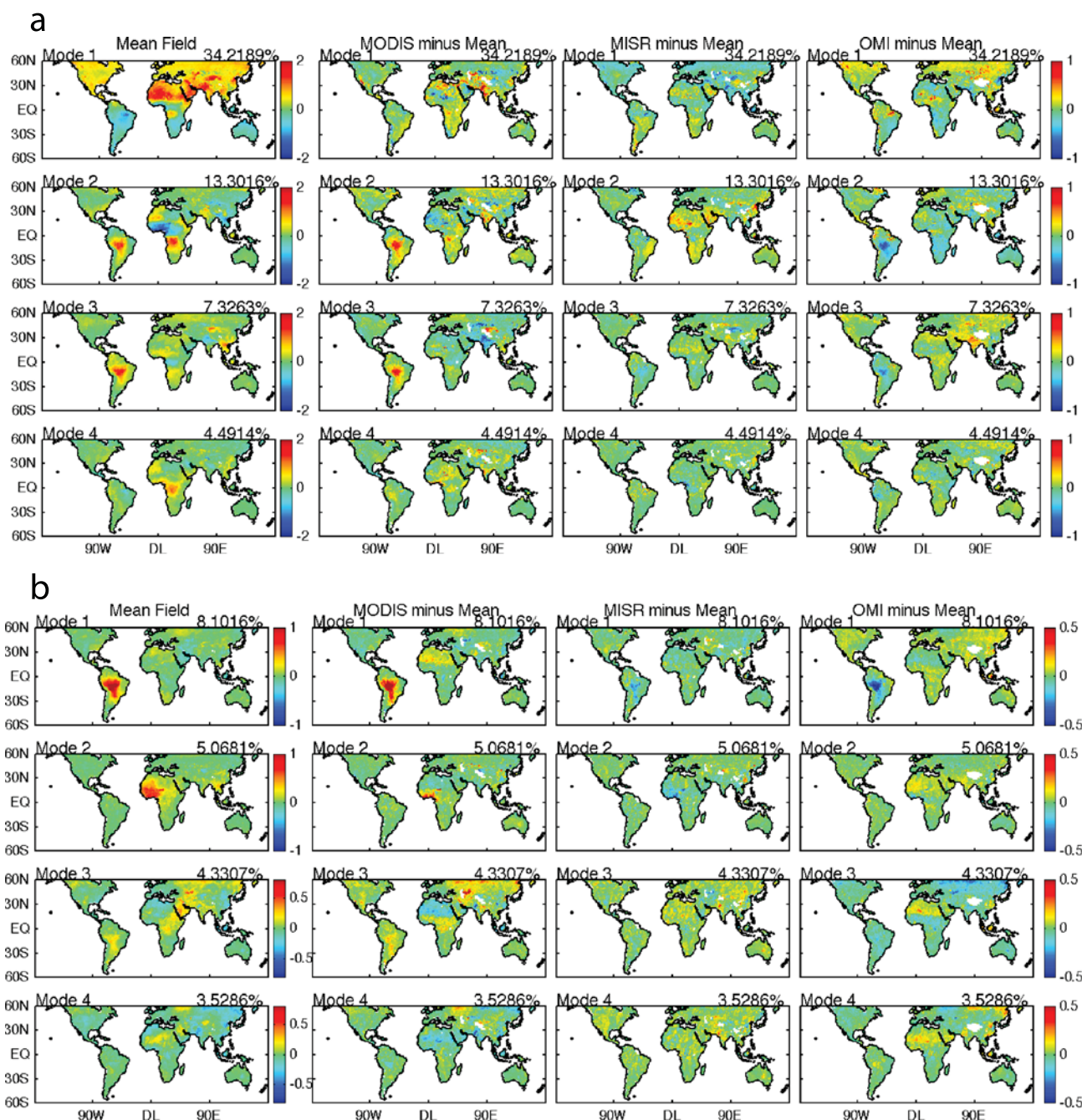
In determining the number of significant modes in spectral analysis, the magnitude and change of variances explained is an important criterion. Figure 3 presents the variance explained by each mode. It is seen that the first four modes for the full data set (Figure 3a) account for  $> 60\%$  of the total variance. Moreover, the variance curve begins to level off starting with Mode 5, each mode higher than 4 only explains a very small portion of the remaining variance. We therefore choose to present only the first four modes here. Figure 4a shows the four leading modes of MODIS, MISR, and OMI over land, which account for  $\sim 50\%$  of the total variance. The spatial patterns and the time series resemble those of EOF analysis of each individual field shown in Part I [Li *et al.*, 2013]. It is clearly seen that all three spatial patterns agree very well across the data sets. The first mode is dominated by dust aerosols over North Africa, Arabian Peninsula, and central Asia. PC 1 displays summer/winter seasonal variability, corresponding to the dust outbreak seasons in the Northern hemisphere, especially the Sahara Desert. The second mode captures biomass burning aerosols over South America, South Africa, and the Sahel. The opposite signals over the Sahel and South Africa are due to their different peak burning seasons. The Sahel burning mainly occurs in the winter-spring season while the burning of South Africa peaks during summer-autumn [Edwards *et al.*, 2006; van der Werf *et al.*, 2006]. A biomass burning signal in South America also appears in Mode 2, but with a stronger signal in Mode 3. The splitting of this signal is because on a global scale, the variability from different regions may influence each other in the analysis, especially when these regions have similar seasonal cycles. The biomass burning season in South America starts in June and peaks in September, and has approximately a one month lag compared to South Africa. As a result, the temporal variability over these two regions is not orthogonal; therefore, Mode 2 captures part of the South American variability. Mode 3 then complements Mode 2 and the two modes



**Figure 4.** The first four modes of CPCA analysis over land, for (a) the full data set and (b) the anomaly data set. The right column shows the PC time series of each spatial mode and the x axis is years. On the spatial maps, the red and blue colors indicate regions vary in-phase and out-of-phase with the PC series, respectively. The number at the upper right corner of each map is the variance explained by that mode.

together recover most of the South American aerosol variability. Similarly, the India region also appears in both Mode 2 and Mode 3 for MODIS and MISR. Mode 2 mainly represents dust variability in this region during the summer season. Mode 3, with semiannual variability, captures dust during premonsoon season and anthropogenic aerosols during the winter. This interaction between different regions in PCA modes is not uncommon. Such problem can be addressed using rotated EOF analysis [e.g., Li *et al.*, 2009], or removing common variability such as the seasonal cycle, or localized analysis focusing on individual regions. The latter two approaches will be discussed later in this section and in section 4.2, respectively. The fourth mode corresponds to the region just south of the Sahel. This region has a local biomass burning season in the



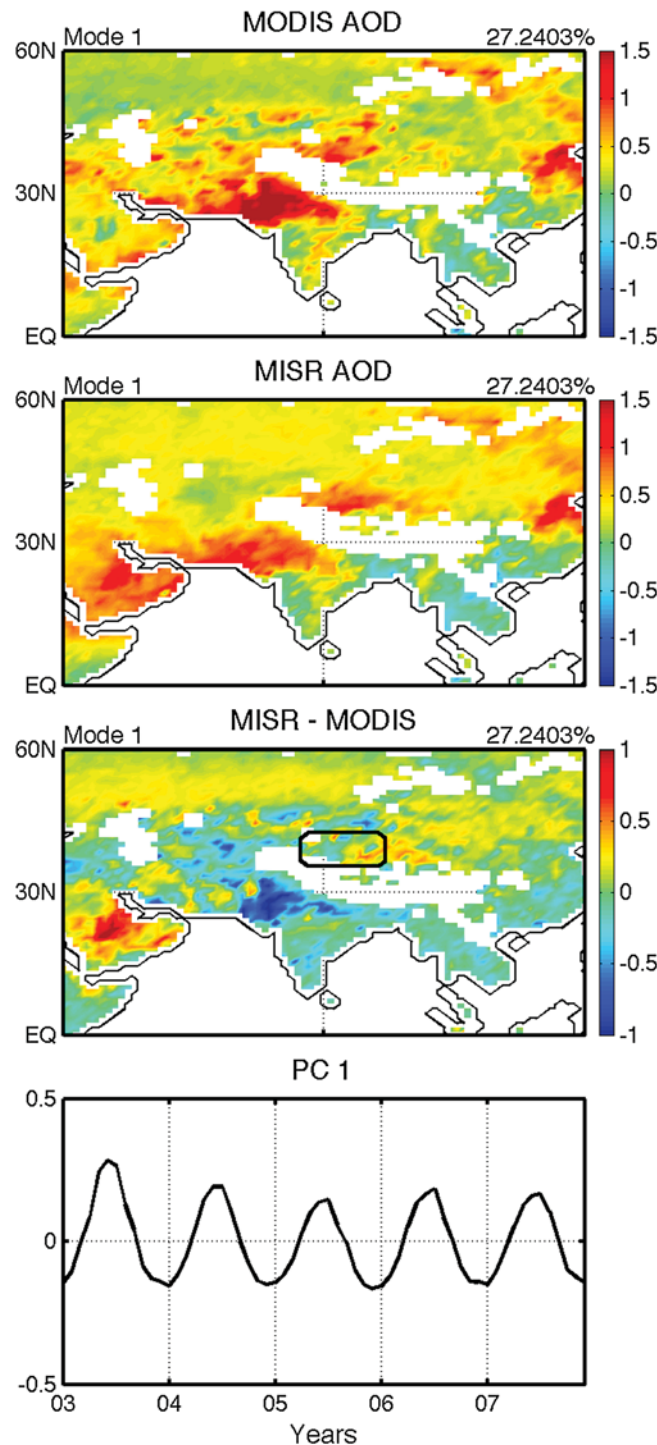


**Figure 5.** The mean field, the difference between the mode of each data set, and the mean for (a) the full data set and (b) the anomaly data set. The differences in Figure 5a mainly include lower variability for MISR over the Taklamakan Desert and higher variability for MODIS over South America and Indian subcontinent; and those in Figure 5b mainly lies in South America and the Sahel where MODIS has higher variability and North Africa where the MODIS variability is lower. Also OMI does not well reflect the Russian fire in Mode 4.

summer but is also influenced by aerosols from the Sahel in the winter; therefore, its time series exhibits semiannual variability with peaks in these two seasons.

The above CPCA analysis suggests good global agreement of the three data sets in presenting major aerosol process and regimes, which is consistent with the EOF analysis of individual fields [Li *et al.*, 2013]. Two advantages of CPCA over PCA approach to find coherent patterns are also clearly demonstrated: (1) EOF analysis may separate the same variability into different modes for different data sets, while in the CPCA results, they all appear in the same mode. (2) In interpreting the results, both the PC and spatial pattern of the





**Figure 6.** The first mode of cosampled Terra-MODIS and MISR AOD and their differences, focusing on Central Asia. The difference between the two data sets over the Taklamakan Desert (box in the third panel) now becomes minimal, suggesting sampling is the reason causing the difference for this region in Figure 6. Nonetheless, other differences still exist, such as North India.

becomes negligible. The remaining differences mainly lie in North India and will be discussed further in regional comparison. Since OMI also has a much wider swath than MISR, it is thus very likely that the MISR underestimation of the Taklamakan AOD variability is associated with sampling. In addition, other major

EOF results must be compared and correlated, while in the CPCA this information is easily obtained through comparison of the resultant the spatial patterns only since they share a common time series.

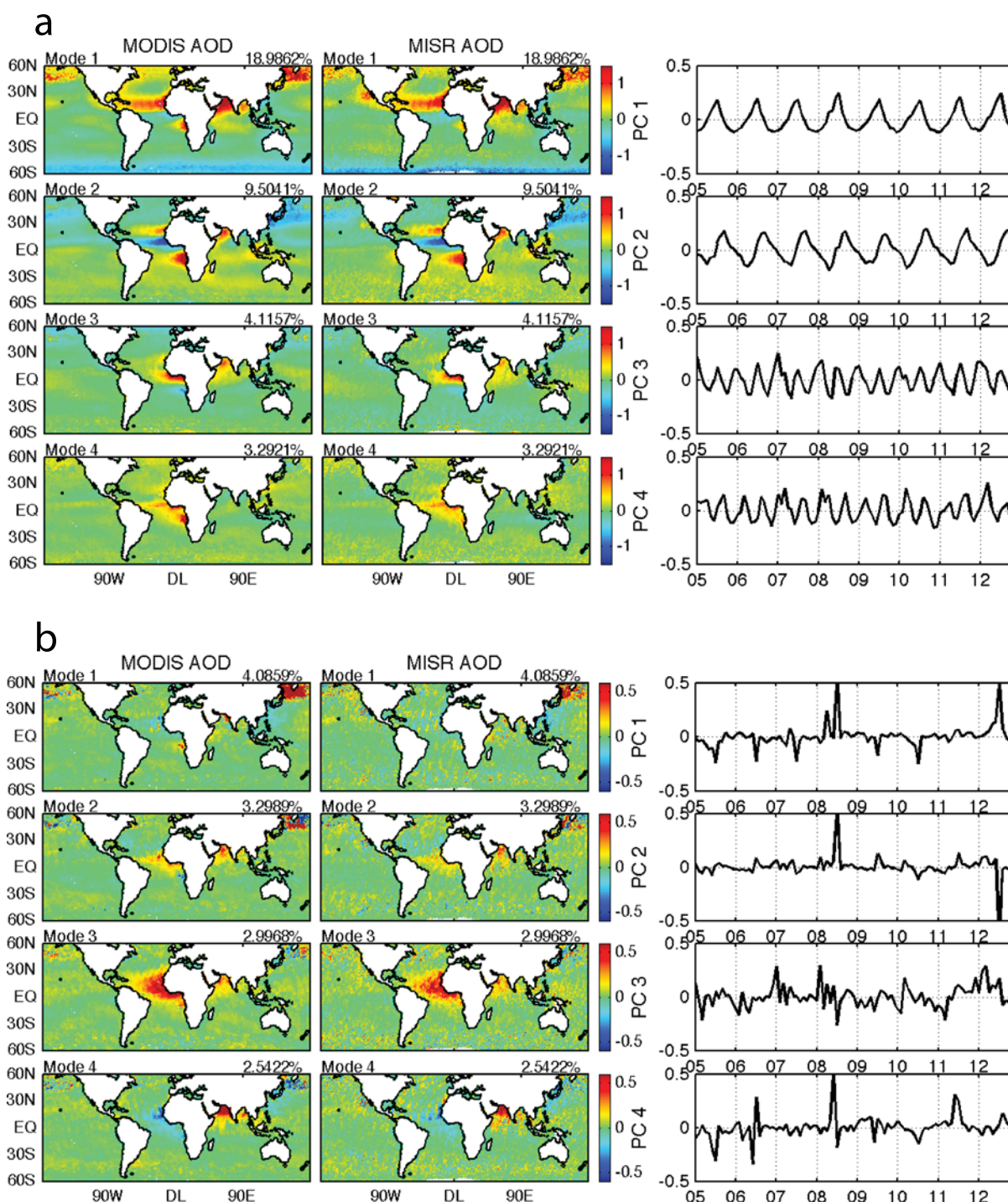
Moreover, since the spatial patterns associated with the different data sets share a common time series, we are able to make more quantitative comparisons by examining the differences between the resultant maps. The mean field of the spatial patterns of each mode, and the differences between MODIS, MISR, and OMI and the mean field, are presented in Figure 5a. Please note that Figure 5a only indicates differences in the magnitude of the variability, rather than the absolute values of the data, which is the nature of spectral analysis. Most differences in Figure 5a are likely noise suggested by the lack of spatially coherent patterns. The major differences include an underestimation by MISR in the variability of the Taklamakan Desert in Mode 1. *Ahn et al.* [2008] pointed out that MISR may miss daily dust storm events or sources of emission because of its narrow swath and may therefore underestimate AOD over desert regions. Here we examine this effect by comparing the cosampled MODIS and MISR AOD. As Aqua-MODIS data cannot be collocated with MISR due to differences in satellite overpass time, the cosampling is performed using Terra-MODIS and MISR data for the 2003–2007 period, following the collocation procedure described by *Liu et al.* [2008]. The collocated data are gridded to  $1^\circ \times 1^\circ$  resolution and CPCA is performed. The first mode of MODIS and MISR, as well as their differences, are presented in Figure 6. We can see that after cosampling with MISR, the dust variability over the Taklamakan Desert as well as North Africa in the MODIS data is also decreased. The difference in the Taklamakan Desert that appeared in Mode 1 of Figure 4a

differences to note include that the fact that MODIS indicates higher aerosol variability over South America in Modes 1 and 3, while OMI has the weakest signal for this region. These differences can be attributable to the conservative cloud screening used in the MISR [Kahn *et al.*, 2009, 2010] and OMI [Ahn *et al.*, 2008] algorithms which underestimates AOD in heavy smoke conditions. Ichoku *et al.* [2003] and Levy *et al.* [2013] also suggested that MODIS tends to overestimate the aerosol loading for the biomass burning regions, as there might be too much absorption in the assumed aerosol model. In the regional analysis we will investigate this further through comparison with AERONET data. MODIS also has larger variability for the Indian subcontinent in Mode 3 compared with both MISR and OMI (this region appears as a negative signal in Mode 3 so the blue color in MODIS minus mean map indicates stronger variability). This region has been associated with greater uncertainty in satellite retrievals due to the high variability in both surface and aerosol types [e.g., Kahn *et al.*, 2010; Levy *et al.*, 2010]. In section 4.2, we will examine some of these regions in greater detail.

The above comparison focuses on seasonal features. Next we examine the data variability and consistency in representing interannual variability using the anomaly data sets. According to the fall-off in the amount of variance explained, we again select the first 4 modes to discuss. Figure 4b shows the four leading modes of the anomaly data. Because the bulk of the variance in the data sets comes from the annual mean and seasonal cycles, the variances explained by each anomaly modes are significantly smaller than those in Figure 4a. However, as mentioned previously, removing the seasonal cycle will reduce the interaction between the variability of different regions and produce more isolated patterns. For example, unlike Figure 4a, South America and West Africa are now separated into different modes, as without the seasonal cycle, aerosol variability of these two regions is affected by different sources, types and meteorological conditions that are largely independent. Globally, South America appears as the strongest source of aerosol interannual variability. This mode is quite consistent across the three data sets. PC 1 shows strong positive anomalies in 2005, 2007, and 2010 and negative anomalies in 2008, 2009, 2011, and 2012. This behavior is consistent with previously documented variability in biomass burning over this region [e.g., Torres *et al.*, 2010; Hooghiemstra *et al.*, 2012; Ma *et al.*, 2012]. Mode 2 represents the dust aerosol source region of West Africa, and the time series generally exhibits small interannual fluctuations. This mode is also very consistent across the three data sets. The strong peak in spring 2010 corresponds to the intense African dust outbreak documented by Jung *et al.* [2013]. More disagreement appears in Modes 3 and 4. For the third mode, the PC has an increasing trend, and its projection on the spatial maps highlights several regions. The Arabian Peninsula shows up in all three data sets, indicating an AOD increase over this region from 2005 to 2012, which is consistent with the results of previous studies using SeaWiFS measurements [Hsu *et al.*, 2012] and AERONET data [Yoon *et al.*, 2012]. MODIS and MISR also indicate increasing trends over North and Northeast Asia. In Mode 4, both MODIS and MISR capture the extensive wildfires in western Russia during the summer of 2010 [Portin *et al.*, 2012; Mielonen *et al.*, 2012], while this feature is absent in the OMI data. The problem in OMI may be associated with cloud contamination and the row anomaly that results in a significant reduction in spatial sampling (<http://www.knmi.nl/omi/research/product/rowanomaly-background.php>).

Again, we examine the differences between the first four modes of the anomaly data in Figure 5b. In Mode 1, the lower aerosol variability of MISR and OMI data over South America identified in Figure 4a becomes more evident in Figure 5b. In addition, MODIS has stronger signals over the Sahel region in Mode 2, but weaker signals over North Africa in Modes 3 and 4. In section 4.2, we will further investigate the difference for Africa and show that they are related to both the seasonal variability and trend in the MODIS DB\_AOD data. As mentioned above, the OMI data does not capture the wildfires in Russia seen in Mode 4 of MODIS and MISR. This is likely due to the contamination of subpixel clouds in the larger OMI footprints.

Finally, note that most of the spatial patterns of the full data sets and the anomalies, as well as the differences, are also reflected in the standard deviation fields presented in Figure 2. This consistency validates the CPCA method in representing the variability in the data. More importantly, the spectral analysis reveals information on the temporal evolution of the stable patterns in the PCs, which is a major advantage over examining only the standard deviation maps shown in Figure 2. Another major advantage of using the CPCA is dimension reduction. If we were to look for both the spatial and temporal variability in these multidimensional data sets without spectral decomposition, we would have to examine every time snapshot of the global map, or the time series of every grid box. However, spectral methods allow us to achieve this goal by examining only the first few modes.



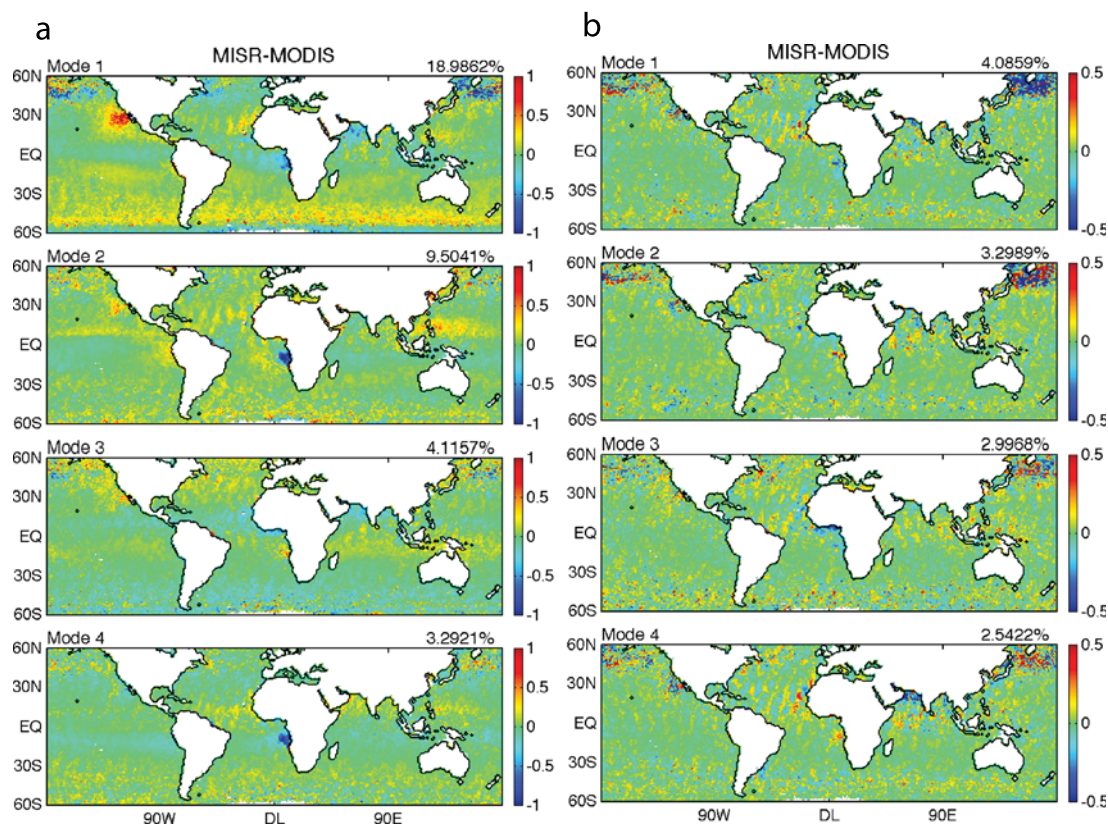
**Figure 7.** The first four modes of CPCA analysis of ocean data. (a) Modes for the full data set. (b) Modes for the anomaly data set. OMI is not included due to its limited coverage over ocean.

#### 4.1.2. Ocean

In this section, we discuss the comparison between the CPCA modes over ocean. Since the ocean coverage for OMI data is limited, this part of the analysis focuses on MODIS and MISR data only.

The ocean CPCA modes also resembles the EOF modes of individual data sets in Part I [Li *et al.*, 2013], and the agreement between the MODIS and MISR data sets is even better than that for the land results (Figure 7). This is not surprising since the retrieval of aerosol optical depth over the ocean is generally an easier task where



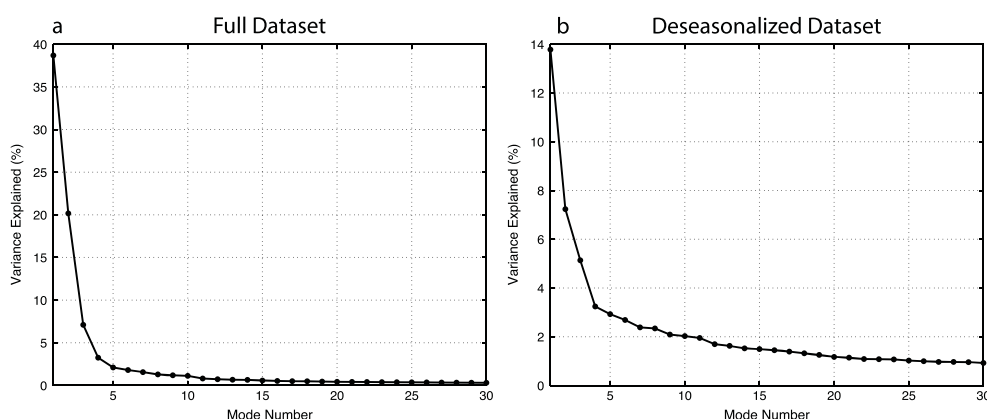


**Figure 8.** The difference between the four modes shown in Figures 7a and 7b. Compared to MODIS, MISR has slightly larger variability off North Mexican coast South African coast in Figure 8a. These areas have high stratocumulus cloud fraction and the difference may come from cloud contamination.

the surface is sufficiently dark. Large AOD loadings in the ocean modes are mostly found in coastal regions where aerosol are transported from land sources. Similar to Figure 4, the first two CPCA modes in Figure 7a both exhibit strong seasonal cycles. The first mode captures dust transport to the North Atlantic and North Indian Oceans and aerosol discharge from East Asia to the North Pacific. The second mode is associated with biomass burning aerosol transported from South Africa, and Southeast Asia. The Sahel region, which has a mixture of dust and biomass burning aerosols during the boreal winter, appears with a negative signal, as its seasonal variability is out-of-phase with PC 2. Both Mode 3 and Mode 4 have semiannual variability and represent aerosol transported from the Sahel and South Africa, respectively. Some regions, such as the Arabian Sea and North Atlantic, appear in multiple modes due to the similar seasonal cycles and the interaction between different aerosol types and sources.

The differences between the CPCA spatial patterns are minimal (Figure 8a). As only two data sets are included in ocean analysis, Figures 8a and 8b are produced by differencing the spatial patterns of MISR and MODIS. According to Figure 8a, MISR tends to have higher variability over the North Mexico coast in Mode 1, but lower variability over the South Africa coast in Modes 2 and 4. These two regions correspond to stratocumulus cloud decks in the summer; therefore, this difference may be due to cloud contamination of the aerosol retrievals. Some of this cloud contamination in MODIS ocean retrieval is likely addressed in the updated Collection 6 product which uses a revised cloud mask for thin-cirrus detection [Levy *et al.*, 2013].

The CPCA modes of the AOD anomaly data over ocean are also highly consistent between MODIS and MISR (Figure 7b). However, both the spatial pattern and time series of the ocean modes appear noisier than those for land (Figure 4b), with the leading modes only each explaining ~3% of the total variance. The difference fields are also mostly noise, only MODIS data display a slightly higher variability over the North Pacific (Figure 8b). Also note that all fields in Figure 8b show satellite orbital tracks globally, especially in the tropical region. Further comparison with Figure 7b, especially Mode 4, suggests that these strips come from MISR



**Figure 9.** Variance explained by the first 30 CPCA mode for North Africa. (a) Full data set. (b) Anomaly data set. According to the shape of the variance curve, the first three modes of both the full data set and anomaly data are considered significant.

data. This feature is a reflection of the different spatial sampling between MODIS and MISR. As MISR has a narrower viewing swath, it tends to leave gaps between orbits, especially for the low latitudes.

As a summary of section 4.1.1 and section 4.1.2, the CPCA analysis effectively extracts major aerosol variability from the three high-dimensional data sets. The decomposed modes reveal and confirm many of the known aerosol sources, variability with season, events and trends, such as extreme biomass burning seasons over South America, strong dust outbreak over West Africa in April 2010 and the increasing AOD trend over the Arabian Peninsula. MODIS, MISR, and OMI agree qualitatively in representing these variabilities in the data. However, the differences between the spatial patterns of different data sets reveal many disagreements that cannot be neglected. Specifically, MODIS has stronger overall variability over South America than MISR and OMI, this is likely associated with both the underestimation of MISR and OMI due to conservative cloud screening, and the overestimation of MODIS AOD due to errors in the absorbing aerosol model used in the retrieval. The MODIS variability is also comparatively high over the Indian subcontinent. However, it indicates lower variability over Northwest Africa. The MISR variability over the Taklamkan Desert and West Africa–Sahel regions is biased low, which is attributed to the limitations in its spatial sampling. The MISR sampling issue also appears in the ocean results. OMI does capture the wildfires in Russia, which is likely the result of cloud contamination and the row anomaly. The coastal regions of North Mexico and South Africa seem to suffer from cloud contamination for MODIS and MISR, respectively. In the next section, we will further examine the consistency as well as disagreements for several representative regions, as some features may not be fully isolated in global analysis.

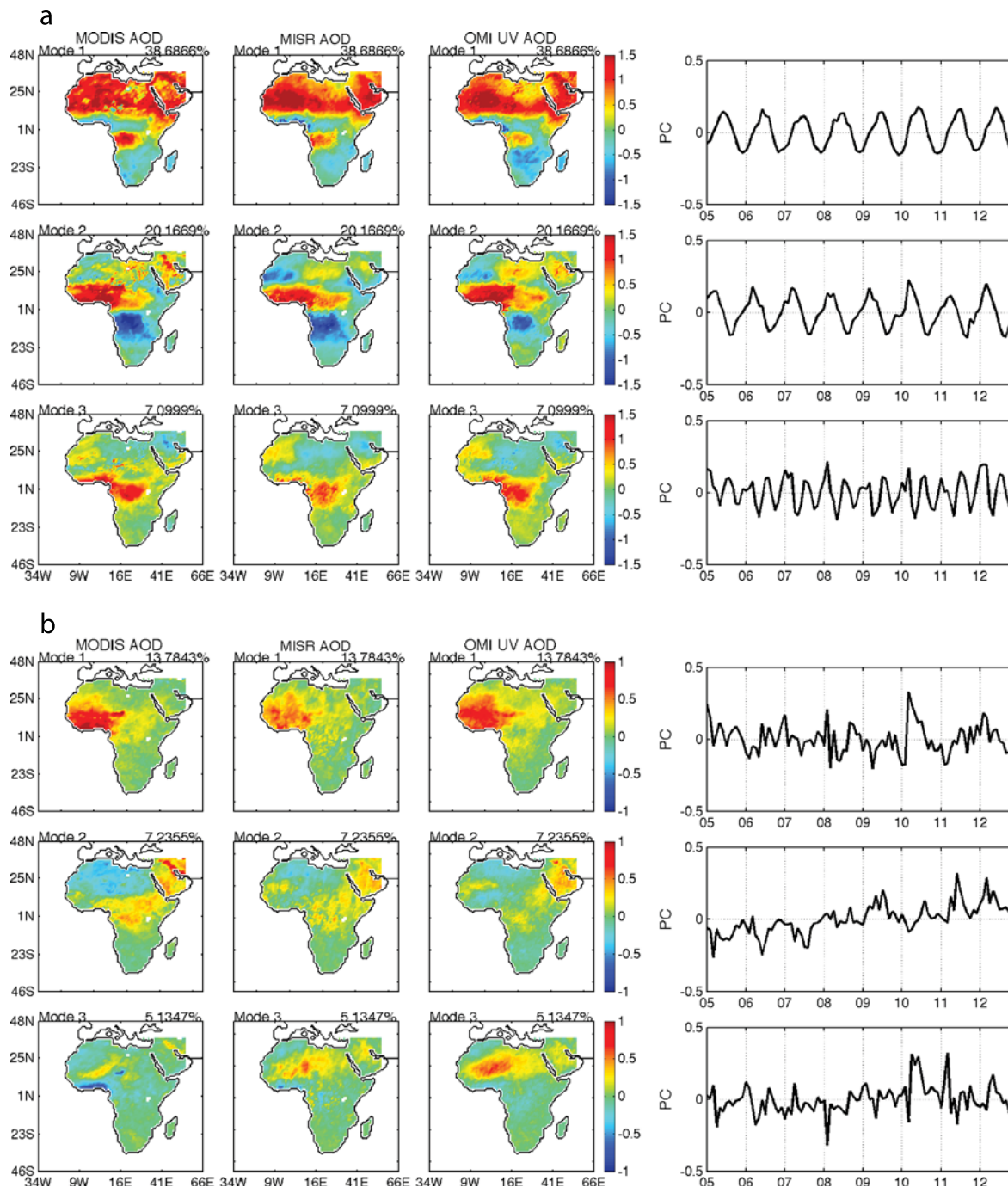
## 4.2. Regional Analysis

In this section, we present and compare the aerosol variability over three regions representing typical aerosol types and sources, in order to better examine more localized aerosol variability and phenomena.

A major difference between the global and regional results is that the global aerosol variability comprises many aerosol types, sources, and meteorological interactions; therefore, the bulk of the variance in the data are usually distributed in more CPCA modes (e.g., 4 modes in this study). Regionally, however, aerosol variability is usually dominated by one or two types, and the first one or two modes capture most of the variance. As a result, in the following discussion, we will focus on the most relevant modes for each region.

### 4.2.1. Africa

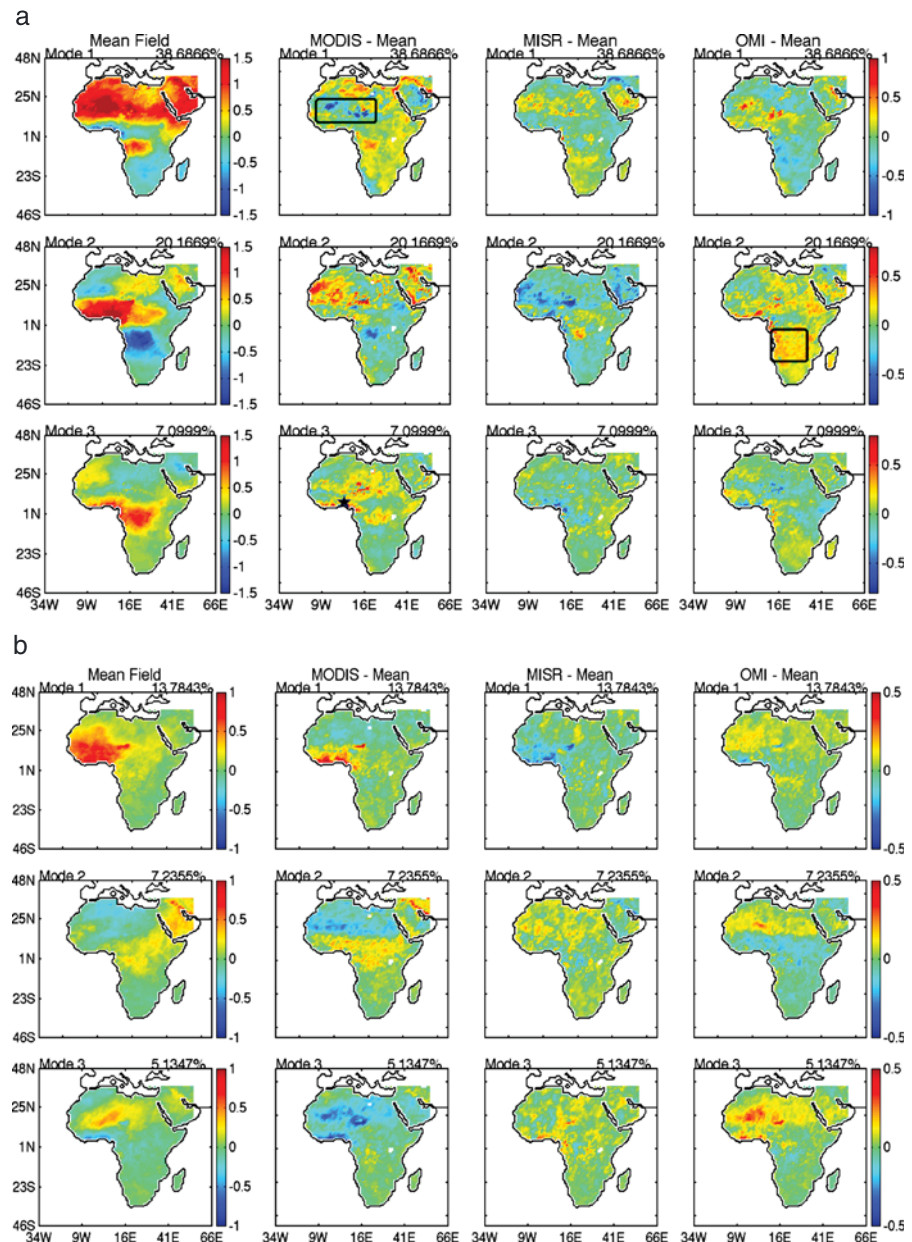
Africa is a very important aerosol source region globally. The Sahara Desert, along with the Arabian Peninsula, in the north is a dominant dust source, and biomass burning is extensive from just north of the equator to southern Africa. By examining the behavior of the variance explained from decomposition results (Figure 9), we determined that the first 3 modes of the full data set and the anomaly data set should be significant to consider. Mode 1 to 3 of the full data set are shown in Figure 10a, which together account for ~70% of the total variance. Similar to the global results, MODIS, MISR, and OMI agree well for all three modes in Figure 10a. Due to its dominant role in global aerosol variability, the spatial patterns of the African modes also resemble those of global modes. Mode 1 represents seasonal variability of Sahara dust and Modes 2 and 3 show the



**Figure 10.** CPCA analysis over Africa for (a) the full data set and (b) anomaly data set.

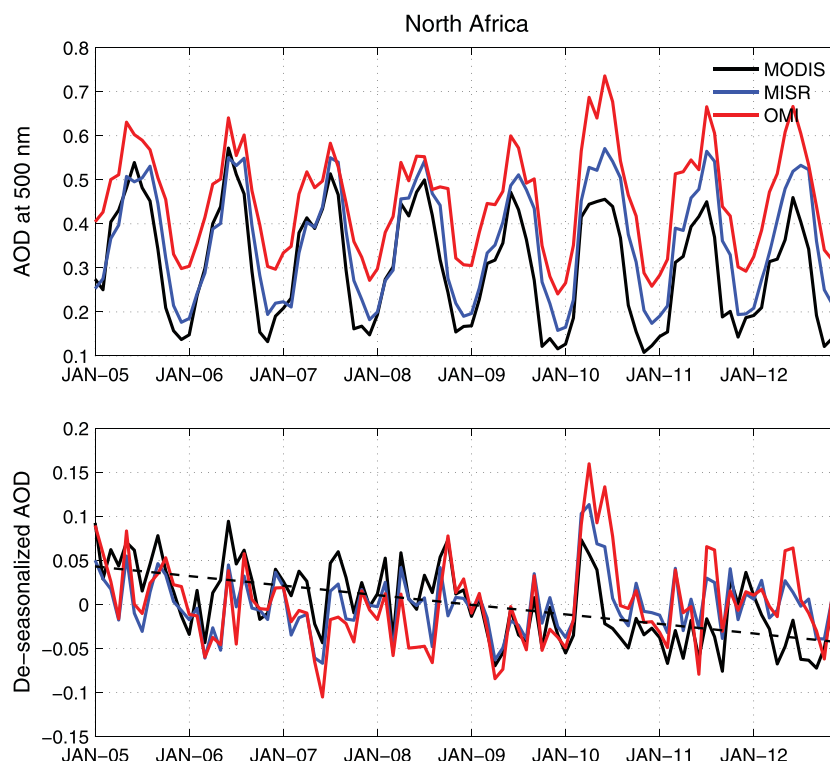
seasonal migration of biomass burning regions from the Sahel to South Africa. The burning starts in December over the Sahel region, gradually shifts to the south with the dry season, and reaches maximum over South Africa in August [Edwards et al., 2006]. Therefore, the Sahel region appears as a negative signal in Mode 2 as the AOD varies out-of-phase with the PC that has summer maximum. The equatorial biomass burning is not completely separated from that in South Africa mainly due to the phase lag in their seasonal cycles. Despite the qualitative agreement, a closer look at the differences (Figure 11a) reveals nonnegligible discrepancies. Specifically, in Mode 1, MODIS DB\_AOD has a low bias in two dust source regions: West Africa and the Bodélé depression while OMI has a slightly higher bias. This points to the different seasonal variability





**Figure 11.** The difference between the modes shown in Figures 10a and 10b. The black boxes on Mode 1 of MODIS-mean and Mode 2 of OMI-mean in (a) mark the places with differences that need further examination. The star on Mode 3 of MODIS-mean in Figure 11a shows the location of Ilorin station where the data will be compared with AERONET.

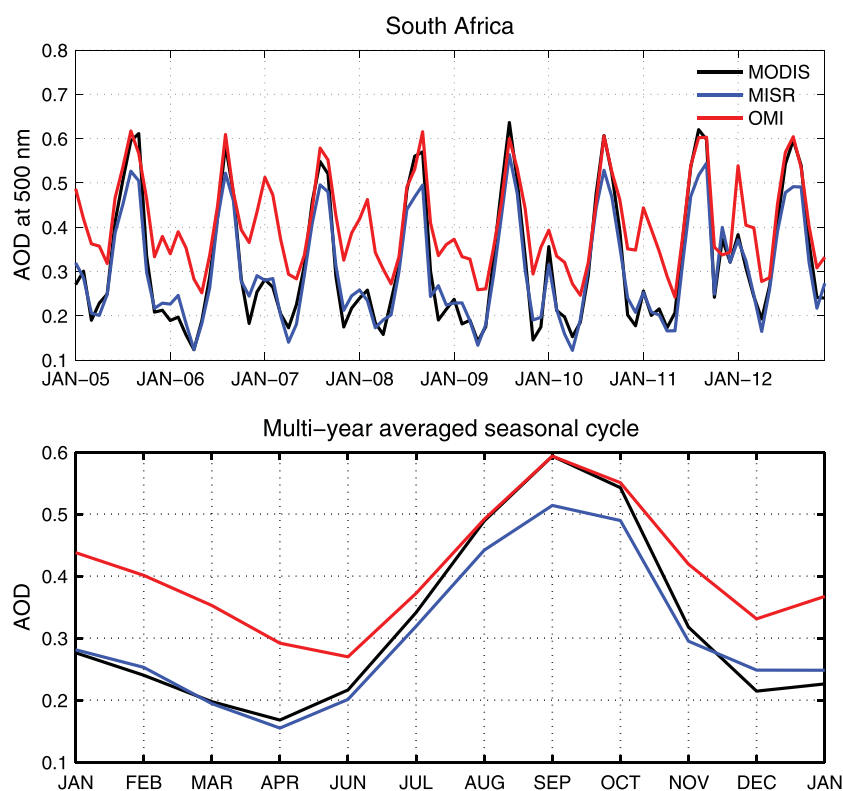
in the three data sets and is further examined by comparing their time series for these regions. Figure 12 shows the averaged AOD time series averaged over the area marked in the black box on the Mode 1 of MODIS-Mean map of Figure 11a. We can see that MODIS DB\_AOD clearly has both a low overall bias and a lower seasonal variability compared to OMI and MISR, especially from 2009 to 2012. Sayer *et al.* [2013] suggested the MODIS underestimation over North Africa might be due to insufficient absorption, surface reflectance overestimation, or overzealous cloud flagging. OMI is biased high, but the magnitude of variability is similar to that in the MISR data. In addition, note that distinct from MISR and OMI, MODIS also has a decreasing trend, which is more clearly seen in the anomaly data shown in Figure 11b. This MODIS trend results in the weaker projection of PC 1 on North Africa, and part of the North Africa variability in MODIS DB\_AOD is mixed with other trend-related modes (Mode 6, figure not shown) and will be better illustrated by the analysis of the anomaly data. However, due to the lack of surface measurements in the Sahara Desert,



**Figure 12.** AOD time series and trend in the anomalies averaged over North Africa marked by the black box in MODIS-mean Mode 1 of Figure 11a. MODIS DB\_AOD clearly has a low bias in both the magnitude and seasonal variability compared to MISR and OMI, especially from 2010 to 2012. In addition, MODIS DB\_AOD appears to have a decreasing trend (dashed black line in lower panel) that is not observed in MISR and OMI.

it is difficult to conclude whether the trend in MODIS DB\_AOD is real or a retrieval artifact. For Mode 2, the lack of negative signal in MODIS mode for North Africa can also be explained by its lower variability shown in Figure 11a. In addition, OMI has an overall low bias in South Africa region. In Figure 13, we further compare the AOD time series averaged over the region marked by the black box in the OMI-mean Mode 2 of Figure 11a. It is seen that MODIS and MISR agree well in both magnitude and variability. OMI also agrees with the other two in the (boreal) spring peak. However, it significantly overestimates for the weaker (boreal) fall peak. This is seen more clearly in the multiyear seasonal cycle. This difference suggests that OMI AOD data over the South Africa region have stronger semiannual variability than MODIS and MISR. Therefore, the projection of PC 2 (spring–fall variability) on OMI is weaker. This also explains the stronger signal in Mode 1 for OMI as this mode captures summer–winter variability. The difference in Mode 3 mainly lies in the weaker variability of MISR for the Sahel region compared to MODIS and OMI. Fortunately several ground measurements from AERONET are available for this region. We select the Ilorin station to compare with the satellite data as it is located right at the point of disagreement and has a long data record. Figure 14 shows the time series of AERONET AOD at Ilorin, and MODIS, MISR, and OMI AOD for the grid box containing the Ilorin station, as well as the multiyear averaged seasonal cycle. We can see that during the peak seasons of boreal winter, MODIS and OMI AOD are comparable to the AERONET data while MISR is usually biased low. Additionally, MISR AOD is slightly higher during a few low AOD periods of August to October. Both factors lead to an overall low seasonal variability of AOD in the MISR data set, which is more evident in the lower panel of Figure 14. Previous comparisons between MISR and AERONET [Kahn *et al.*, 2010] and between MISR and MODIS [Kahn *et al.*, 2009] also identified this region as problematic, and suggested the need to add nonspherical dust and spherical absorbing particles to the MISR aerosol model [Kahn *et al.*, 2010].

Next, we go on to analyze the anomaly data to better isolate different regions. Three major dust source regions Northwest Africa, Arabian Peninsula, and central African around the Bodélé depression are separated into Modes 1–3 (Figure 10b). The three spatial patterns in Mode 1 are reasonably consistent. The difference maps shown in Figure 11b indicate only a slightly lower signal in MISR for the Sahel region, which could be



**Figure 13.** AOD time series and multiyear averaged annual cycle averaged over South Africa marked by the black box in OMI-mean Mode 2 of Figure 11a. While the three data sets agree at the peak of August to October, OMI significantly overestimates during the boreal winter months, resulting in a stronger semiannual variability. Therefore, the projection of PC 2 onto OMI is weaker (Mode 2) and part of OMI the signal for South Africa is distributed to Mode 1 which is characterized by a summer–winter variability.

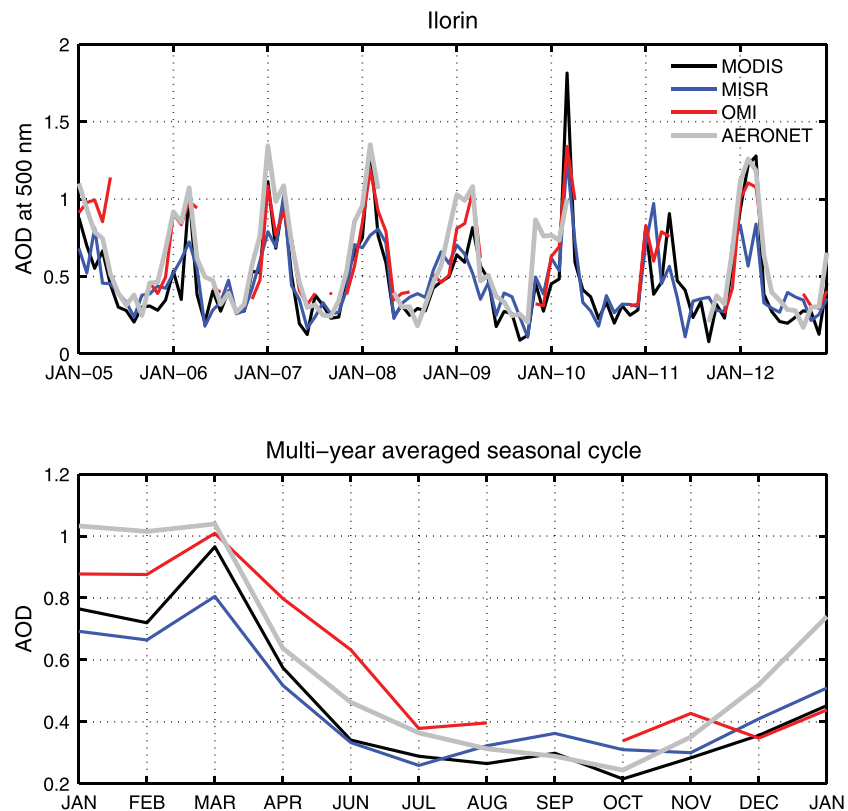
attributed to MISR's low variability for this region as previously inferred from Figure 14. More differences appear in Modes 2 and 3 (Figure 11b). Both modes exhibit an increasing trend in the PC time series. In Mode 2, while the Arabian Peninsula appears in both three data sets, MODIS also has a negative projection in Northwest Africa. This could be the result of the decreasing trend in the MODIS DB\_AOD product for this region as presented in Figure 12. The MODIS trend induces an even larger disagreement in Mode 3, in which both MISR and OMI indicate central North Africa as the major source of variability with the slightly increasing trend, while the MODIS signal over this region is very weak or even opposite in sign.

#### 4.2.2. South America

South America is a dominant biomass burning source region globally. It has multiple burning areas including the southern edge of the Amazonian rain forest, the cerrado grasslands, and the rain forests along the east Brazil coast [Fishman *et al.*, 1996; Tansey *et al.*, 2004]. The burning season starts in June and reaches its maximum in September, when fires are identified extensively across central South America.

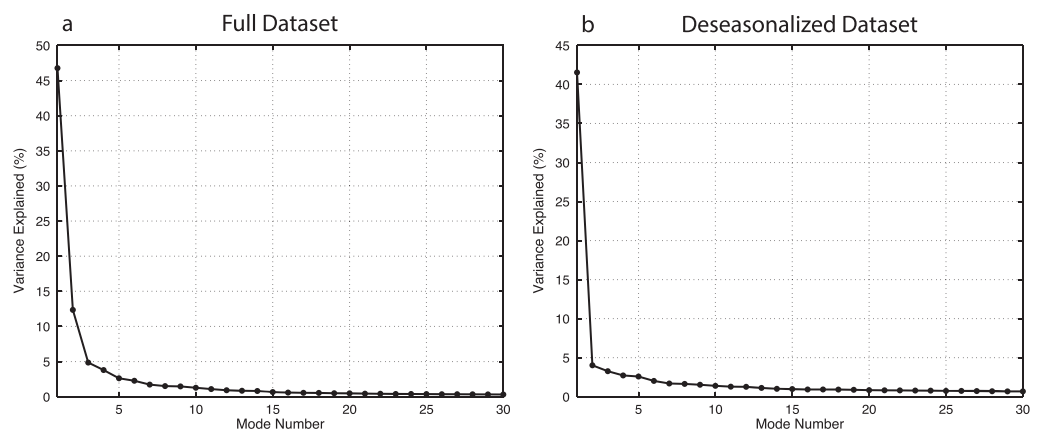
Due to the prevailing dominance of biomass burning aerosols, the first two CPCA modes of the full data set and the first mode of the anomaly data set already account for > 60% of the variance (Figure 15). Not surprisingly, Mode 1 displays the extensive burning region with the expected September peaks. Consistent with global results as shown in Figure 6, the signal in the MODIS data is significantly higher than that in the MISR and OMI data. Here we compare the satellite time series with AERONET measurement at the Alta\_Floresta station (location marked in the first panel of Figure 16) to further examine the bias (Figure 17). According to Figure 17, MISR and OMI consistently underestimate AOD during several peak events in 2003, 2004, 2005, and 2008. They also tend to overestimate during nonpeak season from January to March. Both factors result in their low seasonality in Mode 1. MODIS in general agrees better with AERONET. However, it overestimates AOD during the strong peaks in 2004 and 2005. A further check of AERONET SSA during these two seasons gives an averaged value of 0.89, which is higher than the MODIS assumed value of 0.85 for this region during the



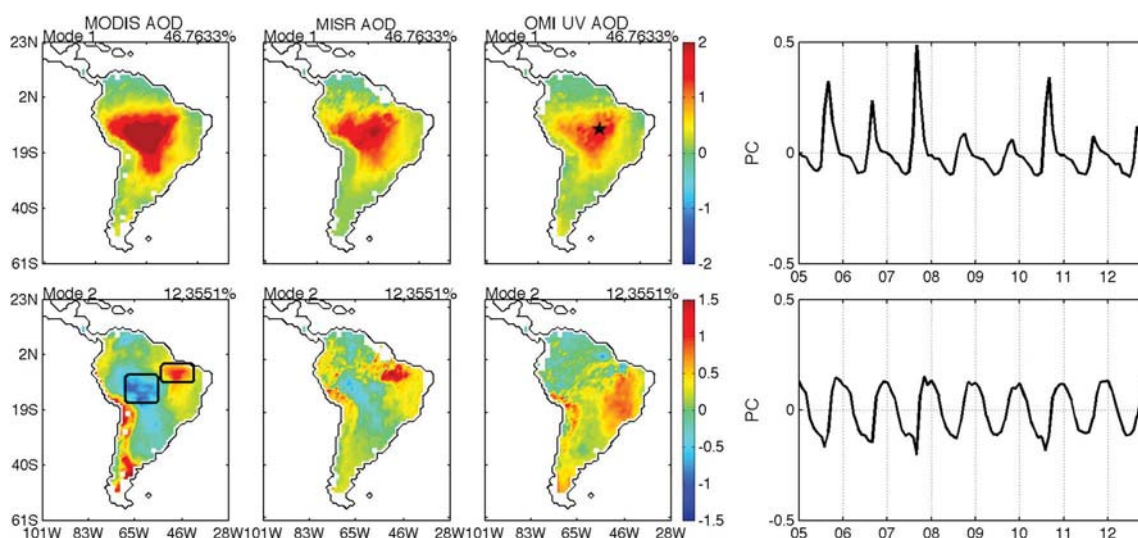


**Figure 14.** AOD time series and multiyear averaged seasonal cycle at the Ilorin station for MODIS, MISR, OMI, and AERONET. MISR has an overall underestimation during the peak season of January to March, while it slightly overestimates during the AOD minimum season of August to October. This contributes to a weaker seasonal cycle in MISR that is responsible for its low variability over the Sahel region found in Mode 2.

biomass burning season [Levy *et al.*, 2007]. This low bias in the assumed SSA likely leads to the overestimation of AOD for MODIS. In Mode 2, interestingly, both MODIS and MISR exhibit a dipole pattern, with a positive center in Northeast Brazil and a negative center in West Brazil and Bolivia. The associated time series has obvious seasonal variability, which is an indication that this dipole is a regular seasonal occurrence. However, Mode 2 of OMI only shows a positive center in West Brazil. In order to trace back the source of this dipole, we show in Figure 18 the time series averaged over the two centers highlighted with black boxes in MODIS Mode 2 in Figure 16. We find that the peak season over East South America (East SA) has a one to two month lag



**Figure 15.** Variances explained by the first 30 CPCA modes for South America. The first two modes of (a) full data set and the first mode of the (b) anomaly data are dominant modes.

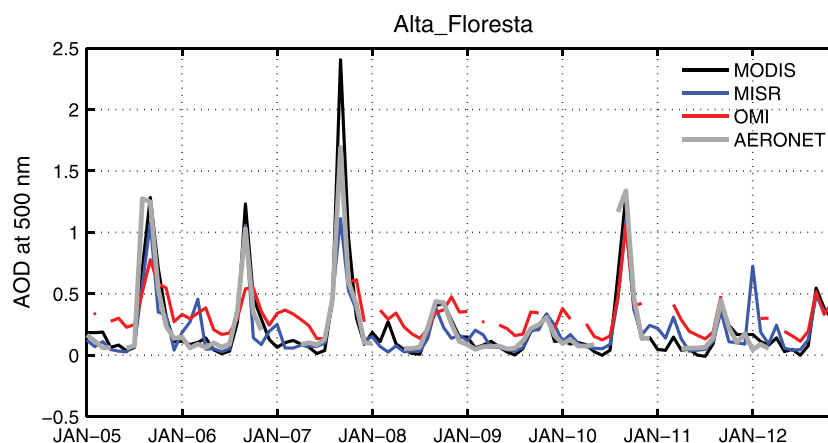


**Figure 16.** The first two modes for South America using full data set. The star on OMI Mode 1 marks the location of the Alta\_Floresta station whose data will be compared to AERONET. Mode 2 of both MODIS and MISR display a east–west dipole feature, and the boxes indicates the regions over which the AOD time series will be averaged and compared.

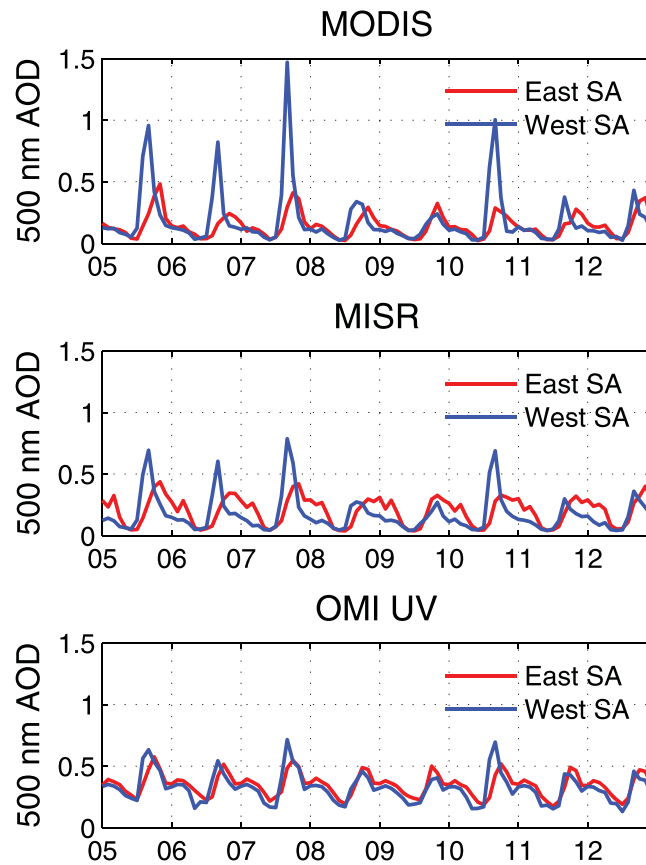
with respect to West SA. This phase shift is most prominent in the MISR data, moderate in MODIS, but almost undetectable in the OMI data except for the last 3 years. The change in the OMI record could be primarily attributed to sampling issues associated with cloud contamination as well as the row anomaly that developed recently. *Videla et al.* [2012] showed that the maximum burning season over the East Brazil area is from October to November using the MODIS fire product, while that for the central South America is from August to October. Moreover, for both MODIS and MISR, the AOD magnitude in the peak month for West SA is much higher than that for East SA, while the OMI AOD has a comparable magnitude for these two regions. This seasonal shift, together with the differences in AOD magnitude, are responsible for the dipole pattern in MODIS and MISR. The stronger negative pole in MODIS should be attributed to the stronger seasonal variability of this data set as previously identified. Analysis of the AOD anomaly data and difference maps does not add any additional information and it is not included here.

#### 4.2.3. India

Aerosol variability over the Indian subcontinent is high influenced by meteorological conditions, and the different aerosol types during the premonsoon/monsoon season and postmonsoon/winter season. During the premonsoon (March–May) and monsoon season (June–August), this region is primarily influenced by



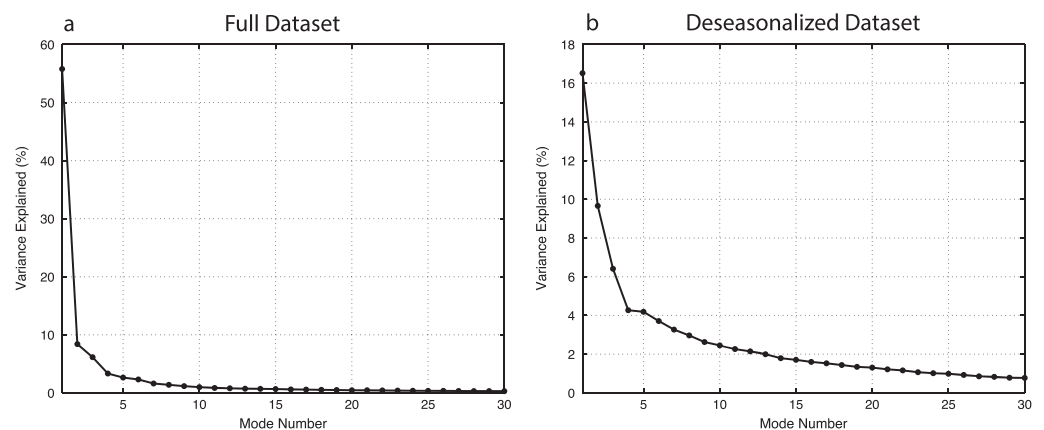
**Figure 17.** AOD time series for MODIS, MISR, OMI, and AERONET at Alta\_Floresta. Compared to AERONET, OMI is biased low during several peak seasons of 2005–2007. MISR is also biased low during the 2007 peak. While MODIS overestimates during the 2006 and 2007 peak seasons. This difference is responsible for the different magnitude of variability in Mode 1 of Figure 16. The reduction of OMI sampling from 2008 should be related to the row anomaly.



**Figure 18.** Averaged time series from the two centers shown in Mode 2 of MODIS and MISR patterns. It is clear that both MODIS and MISR indicate a 1 to 2 months lag between the AOD peaks of East SA and West SA. The AOD during peak season of West SA is also significantly higher than the East SA for MODIS and MISR. These are the factors responsible for dipole pattern in Mode 2. However, the phase shift and magnitude difference is not as clearly observed in OMI data as MODIS and MISR.

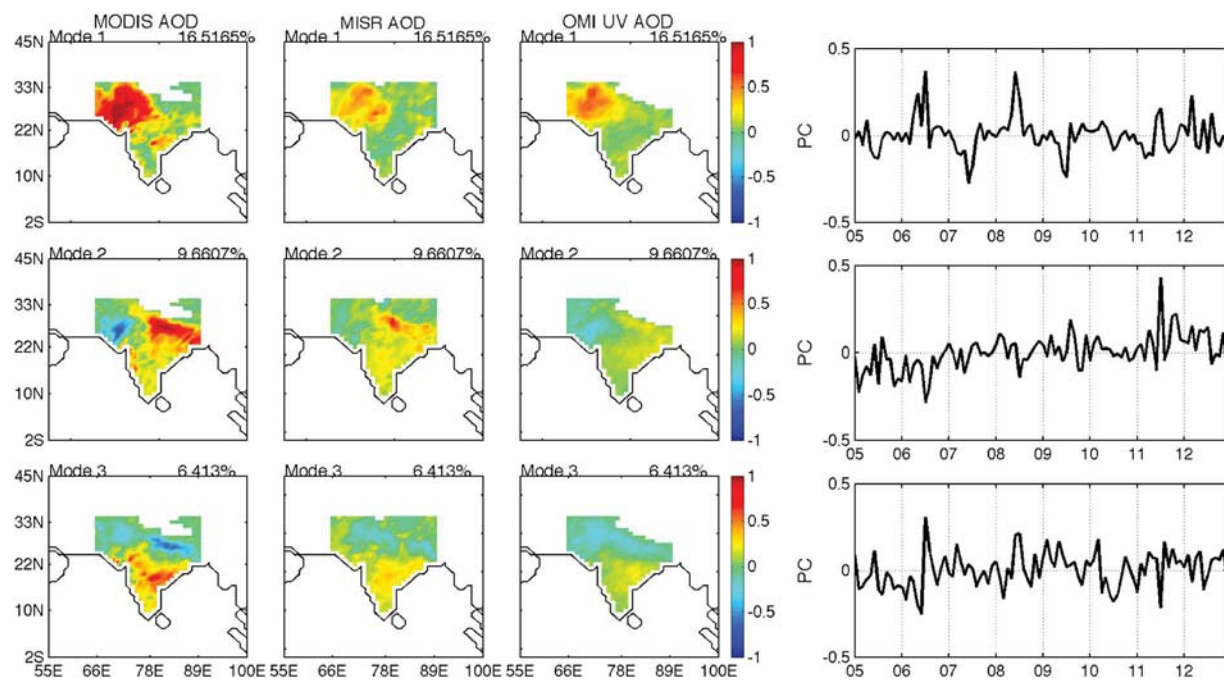
dust aerosols, while during the post-monsoon (September–November) and winter seasons (December and January), anthropogenic aerosols compose a larger fraction of total aerosol loading [Singh *et al.*, 2004; Dey and Di Girolamo, 2010].

From the shape of the variance curve shown in Figure 19, we select the first three modes. These three modes well separate the dust and pollution aerosol from different parts of the India (Figure 20). The decomposition results for the full data set contains the same information but with some interaction between North and



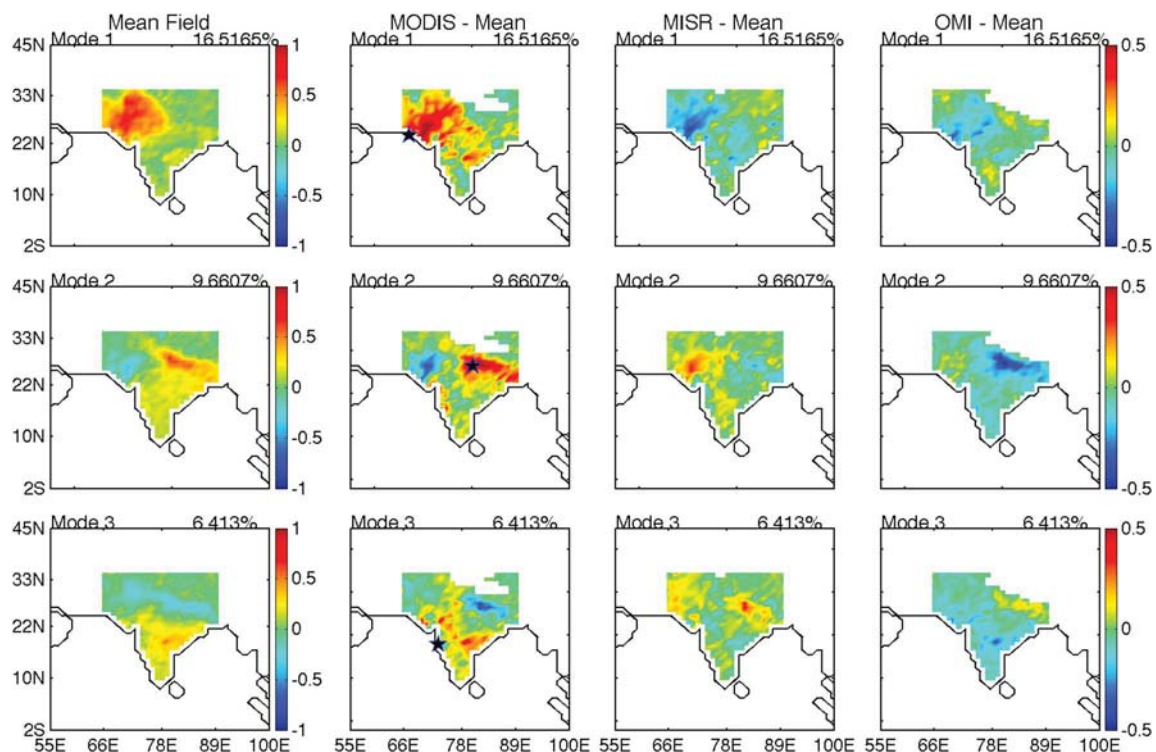
**Figure 19.** Variances explained by the first 30 CPCA modes for India. (a) Full data set. (b) Anomaly data.



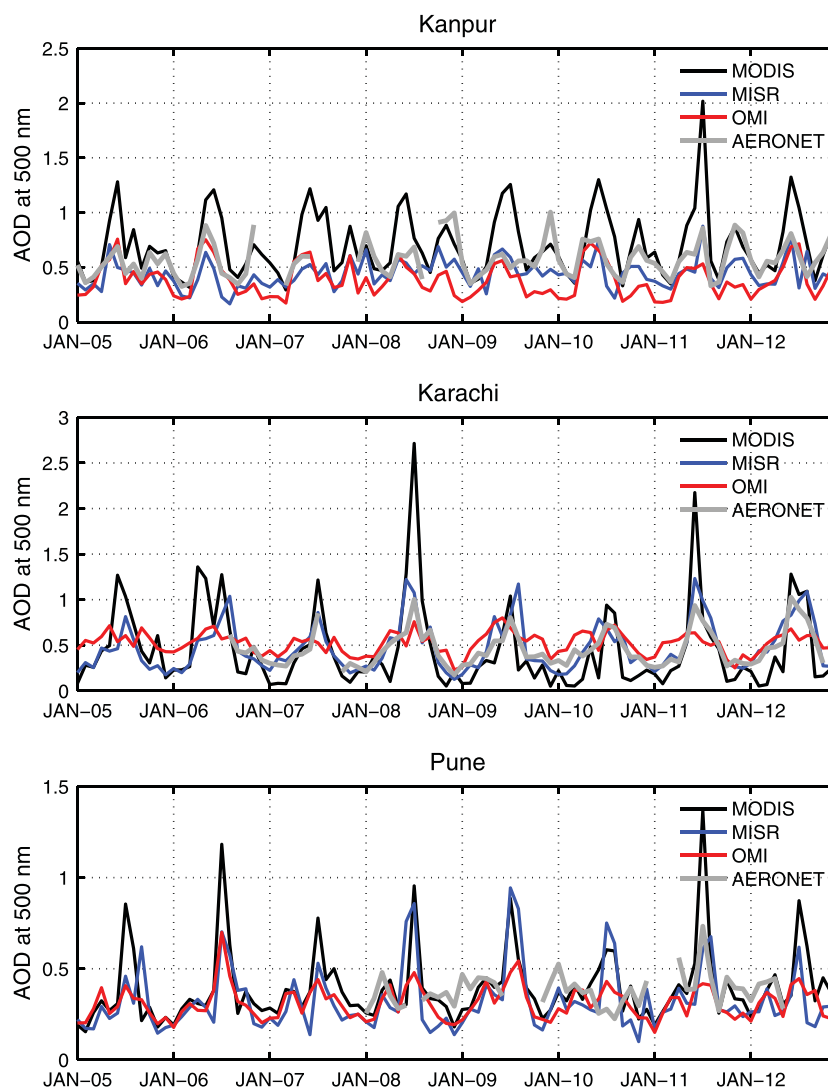


**Figure 20.** The first three modes for India using anomaly data. The full data sets show the same information but not as clearly separated, so they are not shown here.

South India due to similarity in seasonal cycle (Figure not shown). Mode 1 highlights the Thar Desert. PC 1 has two strong positive anomalies in 2006 and 2008 and two negative anomalies in 2007 and 2009. The contrasting negative and positive aerosol anomaly in 2007 and 2008, respectively, is also reported by *Gautam et al.* [2009] and has been attributed to the excess rainfall during the winter preceding the 2007 premonsoon season.



**Figure 21.** The difference between the three modes shown in Figure 20. The stars on MODIS-mean modes indicate locations of the three selected AERONET stations where addition comparison will be made to examine the differences.



**Figure 22.** AOD Time series at Kanpur (Ganga Basin), Karachi (edge of Thar Desert), and Pune (South India). In general, MISR agrees best with AERONET while MODIS seriously overestimates during the summer peaks. For Pune, although the AERONET record is not complete, an underestimation of OMI compared to AERONET in 2011 is still observed. OMI underestimates during most peak seasons but slightly overestimates during the nonpeak season at Karachi. The different behaviors of the three sensors result in strongest signal for MODIS and weakest signal for OMI shown in Figure 21.

Also note that PCs 1 to 3 all have an increasing trend, which is in line with the warming trend induced by increased greenhouse gases and black carbon aerosols [Ramanathan *et al.*, 2007; Ramanathan and Carmichael, 2008].

The three data sets agree qualitatively in the spatial pattern of the three modes. However, the differences, as shown in Figure 21, are still significant. MODIS tends to have higher variability over most of India, especially the Thar Desert (Mode 1) and the Indo-Gangetic Basin (Mode 3), while OMI appears to have the weakest signal for all three regions. Previous studies have identified North India as a major place of discrepancy between MODIS and MISR retrievals of AOD [e.g., Kahn *et al.*, 2009, 2010; Shi *et al.*, 2011]. Tripathi *et al.* [2005] suggested problems in MODIS over India during dust loading seasons, especially the northern part through the comparison with AERONET surface measurements. Jethva *et al.* [2007] also indicated that MODIS tends to overestimate in the presence of dust with respect to AERONET measurement at the Kanpur station in northern India. In addition, Jethva *et al.* [2009] found large discrepancies between MODIS assumed and aircraft measured relationships between surface reflectance at visible channels and the 2100 nm channel, and suggested this could lead to errors in MODIS AOD retrieval. These studies suggest that the disagreements between MODIS and MISR/OMI are likely attributable to problems with the surface parameterization and dust aerosol models. In the

present study, we also use AERONET data to further examine the reason for the differences in the CPCA modes. Three stations, Kanpur, Karachi, and Pune, are selected to represent aerosol property for the Ganga Basin, Thar Desert, and South Indian, respectively, whose locations are marked on the MODIS-mean fields of Figure 21. The time series of AERONET AOD, as well as satellite AOD at these three stations are shown in Figure 22. Note that even for the Thar Desert region, a significant amount of the MODIS AOD still comes from the DT\_AOD product. It is clearly seen that MODIS overestimates AOD for all three stations during the peak season from June to August. It also underestimates AOD at Karachi during the nonpeak season. Overall, MISR has the best agreement with AERONET, while OMI has a low bias for Kanpur and Karachi. Since these stations are primarily influenced by dust aerosols in the summer, the results suggest problems in the dust model used in MODIS DT\_AOD retrieval, which is consistent with the findings of previous studies. The underestimation of OMI over the Indian subcontinent, especially in July, November, and December has also noted by *Ahn et al.* [2008], who attributed it to the minimum interference of subpixel cloud contamination.

## 5. Conclusion

In this paper, we presented the CPCA technique as an effective approach to the comparison of spatiotemporal variability across multiple data sets by using it to examine the consistency as well as discrepancies among MODIS, MISR, and OMI AOD products. This method is especially suitable in parallel comparison between different measurements of the same variable with the same spatial representativeness. This study well demonstrates the advantages of the CPCA approach, including (1) Like the traditional PCA method, the CPCA greatly reduces data dimension and limits the comparison to the few modes that account for most of the data variance in the combined field; (2) The leading modes physically relate to specific aerosol types, sources, or events; (3) Compared with PCA, CPCA extracts common modes of variability and provides robust verification or identification of certain phenomena in the data when the same pattern appears in multiple independent data sets. The agreements found in the spatial patterns are the most believable and reliable aerosol variability that should be primarily used to constrain the aerosol climatologies used in GCMs; (4) Because all spatial modes from the combined field share a common time series, we are able to make quantitative comparison by examining their differences and are able to identify problems in the individual data sets. This is also a major advantage over traditional EOF analysis of individual fields.

The comparison across the CPCA modes of the three data sets reveals primary agreement in representing seasonal and interannual variability of major aerosol regimes both globally and regionally, including dust over North Africa and Central Asia, biomass burning over South America and South Africa, mixed aerosol variability for the Sahel and India and Asian pollution. These regions and the associated temporal variability are the most confident information from current multisensor aerosol measurements, and should be the primary focus of model representations. In addition, the analysis also uncovers interesting phenomena that are not easily observed by non-spectral methods, such as the east–west dipole patterns for South America.

CPCA also enables the examination of the differences. Note the differences referred to here are associated with a specific region and its temporal variability, rather than the data accuracy. A summary of the problems identified in each data set includes

1. Compared to AERONET, MODIS has a higher variability over South America during the biomass burning season, while MISR and OMI are biased low during AOD maximums, in both global and regional analysis. This is likely due to too much absorption in the MODIS absorbing aerosol model, as well as conservative cloud screening for the other two data sets;
2. MODIS also has a slight high bias during the AOD maximum over the Sahel, compared with AERONET, MISR, and OMI at the Ilorin station. The low bias in the MODIS prescribed SSA compared to AERONET is found to be primarily responsible for its higher AOD. This problem, as well as that for South America, are likely improved in the MODIS Collection 6 data which uses a better tuned aerosol type assignments as a function of location and season [*Levy et al.*, 2013];
3. The MODIS DB\_AOD product shows an overall lower seasonal variability over most of North and Northwest Africa than MISR and OMI. In addition, a decreasing trend from 2005–2012 is found in MODIS DB\_AOD for North Africa, while no significant trend exists in MISR and OMI. Due to the lack of AERONET stations in this region, it is difficult to determine whether the MODIS trend is real or a retrieval artifact.



4. The variability of MODIS AOD for the Indian subcontinent is significantly higher than MISR, OMI, and AERONET. A further examination of the time series suggests that the major problem is associated with dust aerosols during the summer months, when MODIS significantly overestimates AOD. Adjustments in the dust model are required to improve MODIS retrieval over this region. Surface parameterization could also be a potential factor.
5. In the global results, MISR has a lower variability over the Taklamakan desert. This is due to the narrow swath and limited spatial sampling the MISR instrument which may miss dust storms or emission events, as inferred from the comparison with cosampled Terra-MODIS results.
6. The aerosol variability over the Sahel is also biased low for MISR, through comparison with MODIS, OMI, and AERONET. Kahn *et al.* [2009, 2010] also identified this region to be problematic in MISR retrievals and recommended the inclusion of a mixed aerosol type to the algorithm;
7. OMI fails to capture the intense fires in Russia in 2010 as found in the global Mode 4 of MODIS and MISR. This may be due to sampling issues associated with cloud contamination and row anomaly;
8. Both MODIS and MISR exhibit phase lags and differences in aerosol loading during the peak burning seasons of East Brazil and Central South America, which produce a dipole like pattern in the CPCA Mode 2 of South America region. However, this phenomenon is absent in OMI data;
9. The OMI data shows a significant overestimation during the boreal fall season over South Africa compared to MODIS and MISR, which results in a weaker annual variability but a stronger semiannual variability of OMI for this region;
10. An overall low bias is found in OMI data over most of the Indian subcontinent, especially during the summer peak seasons. The row anomaly could play a role, as there is an obvious reduction of sampling after 2008.

It should be kept in mind that the main focus of this paper is to present the usefulness of CPCA technique in finding the common patterns in multiple data sets, confirming real variability and signals, and providing clues and insights of the capability of each data set. Nonetheless, while many problems are identified through this comparison, the exact sources of uncertainty can be difficult to identify and needs to be investigated in greater detail in a future study, or through a reexamination of the retrieval algorithm assumptions. Finally, the CPCA method can easily be extended to include more measurements or applied to other variables. It can also be used for model-data or model-model comparison. With the continuous development of aerosol remote sensing, many current records will be extended and new data sets are likely to appear. Therefore, this method will become a useful tool in future comparison studies.

## Acknowledgments

The authors would like to thank the MODIS, MISR, and OMI team for providing the AOD data used in this study. We thank the AERONET for providing the reference data. Thanks are also given to the anonymous reviewers for providing helpful comments and suggestions. The study is funded by NASA grant 509496.02.08.04.24. Jing Li is funded by the NASA Postdoctoral Program (NPP).

## References

- Ahn, C., O. Torres, and P. K. Bhartia (2008), Comparison of ozone monitoring instrument UV aerosol products with Aqua/Moderate Resolution Imaging Spectroradiometer and Multiangle Imaging Spectroradiometer observations in 2006, *J. Geophys. Res.*, **113**, D16S27, doi:10.1029/2007JD008832.
- Bretherton, C. S., C. Smith, and J. M. Wallace (1992), An intercomparison of methods for finding coupled patterns in climate data, *J. Clim.*, **5**, 541–560.
- Dey, S., and L. Di Girolamo (2010), A climatology of aerosol optical and microphysical properties over the Indian subcontinent from 9 years (2000–2008) of Multiangle Imaging Spectroradiometer (MISR) data, *J. Geophys. Res.*, **115**, D15204, doi:10.1029/2009JD013395.
- Diner, D. J., et al. (1998), Multiangle Imaging Spectroradiometer (MISR) description and experiment overview, *IEEE Trans. Geosci. Remote Sens.*, **36**, 1072–1087.
- Edwards, D. P., et al. (2006), Satellite-observed pollution from Southern Hemisphere biomass burning, *J. Geophys. Res.*, **111**, D14312, doi:10.1029/2005JD006655.
- Fishman, J., J. M. Hoell Jr., R. D. Bendura, R. J. McNeal, and V. W. J. H. Kirchhoff (1996), NASA GTE TRACE A experiment (September–October 1992): Overview, *J. Geophys. Res.*, **101**(D19), 23,865–23,879, doi:10.1029/96JD00123.
- Gautam, R., Z. Liu, R. P. Singh, and N. C. Hsu (2009), Two contrasting dust-dominant periods over India observed from MODIS and CALIPSO data, *Geophys. Res. Lett.*, **36**, L06813, doi:10.1029/2008GL036967.
- Holben, B. N., et al. (1998), AERONET—A federated instrument network and data archive for aerosol characterization, *Remote Sens. Environ.*, **66**(1), 1–16.
- Hooghiemstra, P. B., M. C. Krol, T. T. vanLeeuwen, G. R. van derWerf, P. C. Novelli, M. N. Deeter, I. Aben, and T. Röckmann (2012), Interannual variability of carbon monoxide emission estimates over South America from 2006 to 2010, *J. Geophys. Res.*, **117**, D15308, doi:10.1029/2012JD017758.
- Hsu, N. C., S. C. Tsay, M. D. King, and J. R. Herman (2004), Aerosol properties over bright reflecting source regions, *IEEE Trans. Geosci. Remote Sens.*, **42**, 557–569.
- Hsu, N. C., S. C. Tsay, M. D. King, and J. R. Herman (2006), Deep blue retrievals of Asian aerosol properties during ACE-Asia, *IEEE Trans. Geosci. Remote Sens.*, **44**, 3180–3195.
- Hsu, N. C., R. Gautam, A. M. Sayer, C. Bettenhausen, C. Li, M. J. Jeong, S.-C. Tsay, and B. N. Holben (2012), Global and regional trends of aerosol optical depth over land and ocean using SeaWiFS measurements from 1997 to 2010, *Atmos. Chem. Phys.*, **12**, 8037–8053.
- Hubanks, P., M. King, S. Platnick, and R. Pincus (2008), MODIS atmosphere L3 gridded product algorithm theoretical basis document, Collection 005 Version 1.1, *Tech. Rep. ATBD-MOD-30*, NASA.

- Ichoku, C., L. Remer, Y. Kaufman, R. Levy, D. Chu, D. Tanre, and B. Holben (2003), MODIS observation of aerosols and estimation of aerosol radiative forcing over southern Africa during SAFARI 2000, *J. Geophys. Res.*, *108*(D13), 8499, doi:10.1029/2002JD002366.
- Intergovernmental Panel on Climate Change (IPCC) (2007), 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 S. Solomon et al., Cambridge Univ. Press, Cambridge, U. K., and New York.
- Jethva, H., S. K. Satheesh, and J. Srinivasan (2007), Assessment of second-generation MODIS aerosol retrieval (Collection 005) at Kanpur, India, *Geophys. Res. Lett.*, *34*, L19802, doi:10.1029/2007GL029647.
- Jethva, H., S. K. Satheesh, J. Srinivasan, and K. K. Moorthy (2009), How good is the assumption about visible surface reflectance in MODIS aerosol retrieval over land? A comparison with aircraft measurements over an urban site in India, *IEEE Trans. Geosci. Remote Sens.*, *47*(7), 1990–1998.
- Jung, E., B. Albrecht, J. M. Prospero, H. H. Jonsson, and S. M. Kreidenweis (2013), Vertical structure of aerosols, temperature, and moisture associated with an intense African dust event observed over the eastern Caribbean, *J. Geophys. Res. Atmos.*, *118*, 4623–4643, doi:10.1002/jgrd.50352.
- Jung, T., M. Hilme, E. Ruprecht, S. Kleppek, S. K. Gulev, and O. Zolina (2003), Characteristics of the recent eastward shift of interannual NAO variability, *J. Clim.*, *16*, 3371–3382.
- Kahn, R. A., D. L. Nelson, M. Garay, R. C. Levy, M. A. Bull, J. V. Martonchik, D. J. Diner, S. R. Paradise, E. G. Hansen, and L. A. Remer (2009), MISR aerosol product attributes, and statistical comparison with MODIS, *IEEE Trans. Geosci. Remote Sens.*, *47*(12), 4095–4114.
- Kahn, R. A., B. J. Gaitley, M. J. Garay, D. J. Diner, T. F. Eck, A. Smirnov, and B. N. Holben (2010), Multiangle Imaging Spectroradiometer global aerosol product assessment by comparison with the Aerosol Robotic Network, *J. Geophys. Res.*, *115*, D23209, doi:10.1029/2010JD014601.
- Kao, H. Y., and J. Y. Yu (2009), Contrasting eastern-Pacific and central-Pacific types of ENSO, *J. Clim.*, *22*, 615–632.
- Kawamura, R. (1994), A rotated EOF analysis of global Sea surface temperature variability with interannual and interdecadal scales, *J. Phys. Oceanogr.*, *24*, 707–715.
- Kessler, W. S. (2001), EOF representations of the Madden-Julian Oscillation and its connection with ENSO, *J. Clim.*, *14*, 3055–3061.
- Kinne, S., et al. (2003), Monthly averages of aerosol properties: A global comparison among models, satellite data, and AERONET ground data, *J. Geophys. Res.*, *108*(D20), 4634, doi:10.1029/2001JD001253.
- Kutzbach, J. (1967), Empirical eigenvectors in sea-level pressure, surface temperature, and precipitation complexes over North America, *J. Appl. Meteorol.*, *6*, 791–802.
- Levelt, P. F., G. H. van den Oord, M. R. Dobber, A. Malkki, H. Visser, J. de Vries, P. Stammes, J. O. V. Lundell, and H. Saari (2006), The ozone monitoring instrument, *IEEE Trans. Geosci. Remote Sens.*, *44*, 1093–1101.
- Levy, R. C., L. A. Remer, S. Mattoo, E. F. Vermote, and Y. J. Kaufman (2007), Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance, *J. Geophys. Res.*, *112*, D13211, doi:10.1029/2006JD007811.
- Levy, R. C., L. A. Remer, R. G. Kleidman, S. Mattoo, C. Ichoku, R. Kahn, and T. F. Eck (2010), Global evaluation of the Collection 5 MODIS dark-target aerosol products over land, *Atmos. Chem. Phys.*, *10*, 10,399–10,420.
- Levy, R. C., S. Mattoo, L. A. Munchak, L. A. Remer, A. M. Sayer, and N. C. Hsu (2013), The Collection 6 MODIS aerosol products over land and ocean, *Atmos. Meas. Tech. Discuss.*, *6*, 159–259, doi:10.5194/amtd-6-159-2013.
- Li, J., B. E. Carlson, and A. A. Lacis (2009), A study on the temporal and spatial variability of absorbing aerosols using Total Ozone Mapping Spectrometer and Ozone Monitoring Instrument Aerosol Index data, *J. Geophys. Res.*, *114*, D09213, doi:10.1029/2008JD011278.
- Li, J., B. E. Carlson, and A. A. Lacis (2013), Application of spectral analysis techniques in the intercomparison of aerosol data: 1. An EOF approach to analyze the spatial-temporal variability of aerosol optical depth using multiple remote sensing data sets, *J. Geophys. Res. Atmos.*, *118*, 8640–8648, doi:10.1002/jgrd.50686.
- Liu, L., and M. I. Mishchenko (2008), Toward unified satellite climatology of aerosol properties: Direct comparisons of advanced level 2 aerosol products, *J. Quant. Spectrosc. Radiat. Transfer*, *109*(14), 2376–2385.
- Liu, L., A. A. Lacis, B. E. Carlson, M. I. Mishchenko, and B. Cairns (2006), Assessing Goddard Institute for Space Studies ModelE aerosol climatology using satellite and ground-based measurements: A comparison study, *J. Geophys. Res.*, *111*, D20212, doi:10.1029/2006JD007334.
- Ma, X., K. Bartlett, K. Harmon, and F. Yu (2012), Comparison of AOD between CALIPSO and MODIS: Significant differences over major dust and biomass burning regions, *Atmos. Meas. Tech. Discuss.*, *5*, 8343–8367, doi:10.5194/amtd-5-8343-2012.
- Mielonen, T., H. Portin, M. Komppula, A. Leskinen, J. Tamminen, I. Jalongo, J. Hakkarainen, K. E. J. Lehtinen, and A. Arola (2012), Biomass burning aerosols observed in Eastern Finland during the Russian wildfires in summer 2010—Part 2: Remote sensing, *Atmos. Environ.*, *47*, 279–287.
- Portin, H., T. Mielonen, A. Leskinen, A. Arola, E. Pärjälä, S. Romakkaniemi, A. Laaksonen, K. E. J. Lehtinen, and M. Komppula (2012), Biomass burning aerosols observed in Eastern Finland during the Russian wildfires in summer 2010—Part 1: In-situ aerosol characterization, *Atmos. Environ.*, *47*, 269–278.
- Ramanathan, V., and G. Carmichael (2008), Global and regional climate changes due to black carbon, *Nat. Geosci.*, *1*, 221–227.
- Ramanathan, V., M. V. Ramana, G. Roberts, D. Kim, C. Corrigan, C. Chung, and D. Winker (2007), Warming trends in Asia amplified by brown cloud solar absorption, *Nature*, *448*, 575–578.
- Remer, L. A., et al. (2008), Global aerosol climatology from the MODIS satellite sensors, *J. Geophys. Res.*, *113*, D14S07, doi:10.1029/2007JD009661.
- Sayer, A. M., N. C. Hsu, C. Bettenhausen, and M.-J. Jeong (2013), Validation and uncertainty estimates for MODIS Collection 6 “Deep Blue” aerosol data, *J. Geophys. Res. Atmos.*, *118*, 7864–7872, doi:10.1002/jgrd.50600.
- Shi, Y., J. Zhang, J. S. Reid, E. J. Hyer, T. F. Eck, B. N. Holben, and R. A. Kahn (2011), A critical examination of spatial biases between MODIS and MISR aerosol products—Application for potential AERONET deployment, *Atmos. Meas. Tech.*, *4*, 2823–2836.
- Singh, R. P., S. Dey, S. N. Tripathi, V. Tare, and B. Holben (2004), Variability of aerosol parameters over Kanpur, northern India, *J. Geophys. Res.*, *109*, D23206, doi:10.1029/2004JD004966.
- Slonosky, V. C., P. D. Jones, and T. D. Davies (2000), Variability of the surface atmospheric circulation over Europe, 1774–1995, *Int. J. Climatol.*, *20*, 1875–1897.
- Tansey, K., et al. (2004), Vegetation burning in the year 2000: Global burned area estimates from SPOT VEGETATION data, *J. Geophys. Res.*, *109*, D14S03, doi:10.1029/2003JD003598.
- Torres, O., A. Tanskanen, B. Veihelmann, C. Ahn, R. Braak, P. K. Bhartia, P. Veefkind, and P. Levelt (2007), Aerosols and surface UV products from Ozone Monitoring Instrument observations: An overview, *J. Geophys. Res.*, *112*, D24S47, doi:10.1029/2007JD008809.
- Torres, O., Z. Chen, H. Jethva, C. Ahn, S. R. Freitas, and P. K. Bhartia (2010), OMI and MODIS observations of the anomalous 2008–2009 Southern Hemisphere biomass burning seasons, *Atmos. Chem. Phys.*, *10*, 3505–3513, doi:10.5194/acp-10-3505-2010.
- Tripathi, S. N., S. Dey, A. Chandel, S. Srivastva, R. P. Singh, and B. Holben (2005), Comparison of MODIS and AERONET derived Aerosol Optical Depth over the Ganga basin, India, *Ann. Geophys.*, *23*, 1093–1101.

- van der Werf, G. R., J. T. Randerson, L. Giglio, G. J. Collatz, P. S. Kasibhatla, and A. F. Arellano Jr. (2006), Interannual variability in global biomass burning emissions from 1997 to 2004, *Atmos. Chem. Phys.*, *6*, 3423–3441, doi:10.5194/acp-6-3423-2006.
- Videla, F. C., F. Barnaba, F. Angelini, P. Cremades, and G. P. Gobbi (2012), The relative role of Amazonian and non-Amazonian fires in building up the aerosol optical depth in South America: A five year study (2005–2009), *Atmos. Res.*, *122*, 298–309.
- Wallace, J., C. Smith, and C. Bretherton (1992), Singular value decomposition of wintertime sea surface temperature and 500-mb height anomalies, *J. Clim.*, *5*, 561–576.
- Wang, B., and S.-I. An (2005), A method for detecting season-dependent modes of climate variability: S-EOF analysis, *Geophys. Res. Lett.*, *32*, L15710, doi:10.1029/2005GL022709.
- Yoon, J., W. von Hoyningen-Huene, A. A. Kokhanovsky, M. Vountas, and J. P. Burrows (2012), Trend analysis of aerosol optical thickness and Ångström exponent derived from the global AERONET spectral observations, *Atmos. Meas. Tech.*, *5*, 1271–1299, doi:10.5194/amt-5-1271-2012.