# Multispectral Analysis of Aerosols Over Oceans Using Principal Components

Thomas A. Jones and Sundar A. Christopher

***Abstract*—Applying principal component analysis (PCA) to one month of Moderate Resolution Imaging Spectroradiometer (MODIS) narrow-band short-wave radiance data and comparing with the Goddard Global Ozone Chemistry Aerosol Radiation Transport (GOCART) model simulations, we show that aerosol size and speciation information can be inferred from multispectral radiance information without having to use other parameters, such as a fine mode fraction (FMF), that are difficult to validate. PCA was applied to seven highly correlated MODIS solar channels (0.47, 0.55, 0.66, 0.86, 1.24, 1.64, and 2.12 $\mu$m) to extract noncorrelated pseudochannels, each with a unique interpretation. The first pseudochannel (PC1) can be interpreted as the mean radiance across the seven channels, which is directly proportional to the aerosol concentration. The second pseudochannel (PC2) is sensitive to the aerosol size since different aerosol types scatter and absorb differently across the seven MODIS short-wave channels. PC3 is inversely related to the aerosol optical thickness (AOT) and the FMF and appears most sensitive to changes in sulfate and maritime sea-salt concentrations. Results indicate that high values of PC1 are indicative of high dust aerosol concentrations comprising more than 40% of the total AOT, whereas high values of PC2 indicate anthropogenic aerosol concentrations (deduced from GOCART) in excess of 60%. Compared to simple 0.55-$\mu$m FMF thresholds, the PC channels are much more sensitive to dust aerosol concentrations and certain aspects of anthropogenic aerosols, with very low FMF values alone ($< 0.2$) being the best indicator of predominately sea-salt aerosol concentrations. Our results indicate that PCA could be used as an alternate method for inferring aerosol speciation information in future research over ocean and more complex land surfaces.**

***Index Terms*—Aerosols, statistics.**

## I. Introduction

With the advent of reliable satellite-derived data sets of global aerosol properties, the climate effect of specific aerosol types, particularly those anthropogenic in origin, has become an important focus for research during the past 5 years [1]. Moderate Resolution Imaging Spectroradiometer (MODIS) provides radiance measurements in 36 channels from the ultraviolet to the thermal part of the electromagnetic spectrum. The MODIS aerosol algorithm over ocean uses the solar radiance at seven channels (0.47, 0.55, 0.66, 0.86, 1.24, 1.64, and 2.12 $\mu$m) [2] and then retrieves the aerosol optical thickness (AOT) using a lookup table approach [2]. The multiwavelength capability provides other retrievals such as the size, which is useful for separating natural from anthropogenic aerosols [1]. The spectral values of the AOT ($\tau$) contain information about the aerosol size, described by the Angstrom power law, i.e.,

$$\tau = \beta\lambda^{-\alpha}. \tag{1}$$

Here, $\alpha$ is the wavelength exponent, $\beta$ is the turbidity parameter, and $\lambda$ is the wavelength in micrometers. The value of $\tau$ is dependent on the ratio of large to small aerosols, with large $\tau$ indicating the presence of high concentrations of small aerosols (e.g., smoke aerosols from biomass burning), and small $\tau$ indicating the presence of large aerosols, e.g., dust. Observational studies have made use of this size information such as small $(0.1 < r_e < 0.25\ \mu\text{m})$ and coarse modes $(1.0 < r_e < 2.5\ \mu\text{m})$ [sac1]AOT at 0.55 $\mu$m to discriminate between various aerosol types (e.g., [1]). The ratio of small to total AOT is known as the fine mode fraction (FMF) and has proven quite successful in separating predominately naturally occurring (coarse mode) from anthropogenic (small mode) aerosols [1], although it is a difficult parameter to validate [3]. In this method, MODIS AOT and FMF at 0.55 $\mu$m are used to separate the total AOT into various aerosol components using several thresholds [1]. AOT components are separated by first assuming that the maritime AOT is solely a linear function of near-surface wind speed over the ocean. The dust and anthropogenic components of the AOT are then solved using a mathematical relationship derived between the total AOT, the maritime AOT, the observed FMF, and the FMF thresholds. However, there are many assumptions required when separating the total AOT into various components such as the FMF thresholds themselves, and assuming that all small-mode aerosols are anthropogenic also leads to uncertainties in aerosol component radiative effects [3], [4].

The aerosol classification used by Kaufman *et al.* [1] and others generally uses aerosol information from only a single channel (usually at 0.55 $\mu$m), overlooking differences in the multispectral response for different aerosol types. Since each channel is highly correlated, we apply the principal component analysis (PCA) to extract the independent signals that are present in the seven-channel radiance data. Tanre *et al.* [5] used a similar PCA approach using theoretical radiance values for MODIS aerosol channels and observed that the AOT for small-mode aerosols decreases as a function of the wavelength much more quickly than for coarse-mode aerosols. Zubko *et al.* [6] extended this technique to other parameters, such as the single

Manuscript received November 27, 2007; revised February 23, 2008. This work was supported in part by the National Aeronautics and Space Administration (NASA) Radiation Sciences Program, in part by the NASA Interdisciplinary Sciences Program, in part by an Earth Observing System Grant, and in part by the Atmospheric Chemistry Modeling and Analysis Program.

The authors are with the Department of Atmospheric Science, The University of Alabama in Huntsville, Huntsville, AL 35805 USA (e-mail: tjones@nsstc.uah.edu; sundar@nsstc.uah.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2008.920019

scattering albedo, the viewing angles, and the surface wind speed. Results from both efforts observed that two or three independent pieces of information can be retrieved from MODIS multispectral radiance values through the PCA, and that these pieces of information account for more than 90% of the total variance present within the raw data. However, very little research has occurred that utilizes the multispectral characteristics of various aerosol species in aerosol classification using the PCA. In this study, the PCA is applied to the Terra MODIS radiance data at each channel, thereby producing a new set of independent channels, with each new channel sensitive to different aerosol characteristics. These data are then compared with the Goddard Global Ozone Chemistry Aerosol Radiation Transport (GOCART) modeled aerosol speciation during the same month and the Aerosol Robotic Network (AERONET) data from several sites that are known to have a single predominant aerosol species. The goal is to determine if multispectral PCA will provide information on aerosol species that cannot be observed using the 0.55-$\mu$m AOT and/or FMF alone.

## II. Data and Methods

Terra MODIS radiance values at 0.47, 0.55, 0.66, 0.86, 1.24, 1.64, and 2.12 $\mu$m are obtained over the global oceans only from the MOD04, Collection 5 aerosol product [7], with a spatial resolution of 10 km$^2$ at nadir. We use the raw radiance values as inputs to the PCA rather than the retrieved AOT values at each wavelength since AOT retrievals are already a function of all seven radiance channels. For increasing aerosol concentrations, for scattering types of aerosols, larger amounts of short-wave radiation are scattered back to the satellite, resulting in a higher observed top of atmosphere radiance values. The AOT values over the ocean are retrieved from MODIS radiance values using a combination of four small-mode and five coarse-mode aerosol models with the relative amounts of small-mode to coarse-mode aerosols adjusted until a solution is found, which minimizes the error across the seven radiance values [2]. As a result, the reported AOT values at a certain wavelength are dependent on the radiance values from other channels. Using the retrieved AOT would induce uncertainties and biases that are present in the MODIS aerosol models into the research presented here. Thus, we chose to use the measured MODIS cloud-free radiance values as inputs to the PCA rather than any derived products. MODIS data are collected for August 2004 due to the variety of differing aerosol types around the globe during this month. MODIS AOT and FMF from the MOD04 level 2 product from the Terra satellite along with the AERONET AOT values are used to help interpret the PCA components. The AERONET is a network of well-calibrated ground-based sun photometers that provide spectral AOT and other aerosol properties from hundreds of locations around the world [8]. We use the data for August 22, August 16, and August 10, 2004, from AERONET sites Capo Verde, Gosan (Korea), and Ascension Island, which are compared and colocated with the MODIS AOT data for that day. Data from these days were selected based on times when both good quality AERONET and MODIS data exist. For a particular day, MODIS AOT data within $\pm 1^\circ$ of the AERONET site were averaged and compared to the AERONET observations. Each site location represents where the primary aerosol source (other than maritime sea salt) is known *a priori* from GOCART, including dust, fossil fuel combustion, and biomass burning, respectively.

The satellite-retrieved AOT and FMF are compared with the aerosol species simulations produced by the GOCART model [9]. GOCART simulates the transport of aerosols and their type over a global domain on a monthly averaged timescale. GOCART uses global emissions of aerosols and assimilated meteorological fields to separately calculate the mass loading of each aerosol species, which are then converted to the AOT using mass extinction coefficients. Aerosols are categorized as black carbon (BC), organic carbon (OC), sulfate (SU), dust (DU), and sea salt (SS). Organic carbon and sulfate are also split into anthropogenic and naturally occurring components, the latter generally from dimethyl sulfide (DMS). Monthly averaged data for August 2004 were acquired for an ocean only, global domain and were interpolated to a $2^\circ \times 2^\circ$ grid.

The statistical technique of the PCA enables the reduction of multiple and highly correlated data attributes into a fewer number of independent variables, each with a unique physical interpretation [10], [11]. The initial step in the PCA is the calculation of a correlation matrix between each of the input channels, which are the radiance values at all seven MODIS channels. Once the correlation matrix is computed, eigenvalues and vectors are calculated from the correlation matrix to determine the weighting coefficients, which are then applied to the radiance data to produce the new independent pseudochannels (or PC channels). Each PC channel is ordered in such a way that the first channel (PC1) accounts for the greatest variance in the raw data, with the next (PC2) accounting for the next highest amount of variance, and so on. For this paper, the total number of pseudochannels created is equal to the number of input channels, which is seven. For illustration purposes, PC values are normalized to a 0–1.0 scale and averaged to a $2^\circ \times 2^\circ$ resolution grid to correspond with the monthly averaged GOCART data resolution. To visualize the sensitivity of the PC channels to aerosol properties, we compare GOCART and MODIS data as a function of the PC value. PC values are binned at 0.05 intervals, and, for each bin, the ratio of the GOCART component AOT to the total AOT is calculated for grid cells within a particular PC bin. To compare the effectiveness of the PC classification with the 0.55-$\mu$m FMF classification, a similar procedure is used but using a 0.05 FMF interval bin.

## III. Results

### A. PC Weightings

As expected, solar radiance from each of the seven MODIS aerosol channels is highly correlated, although the correlation does decrease to $< 0.5$ when comparing 0.47- to 2.13-$\mu$m radiance values. Analysis of the eigenvalues associated with each PC channel indicates that 78.4% of the variance is explained by PC1, 19.7% by PC2, and 1.4% by PC3, with PC1–PC3 accounting for $\sim$99% of the total variance, similar to that expected in [5] and [6]. The amount explained by PC4–PC7 combined is less than 1%, with no consistent patterns observed

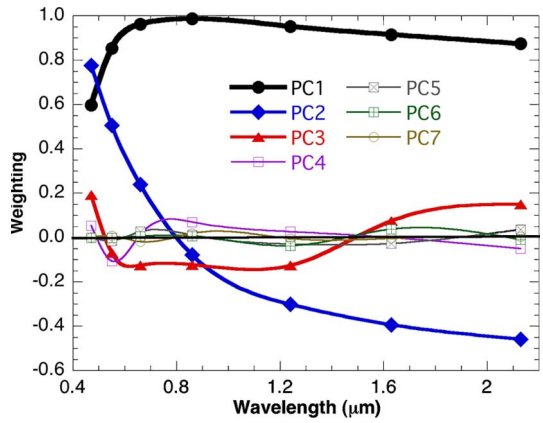


Fig. 1. Weighting coefficients derived from the PCA for PC channels 1–7. Magnitudes indicate the relative importance of each original channel to the resulting PC channel. Note that weightings for PC4–PC7 are very small compared to those for PC1–PC3.

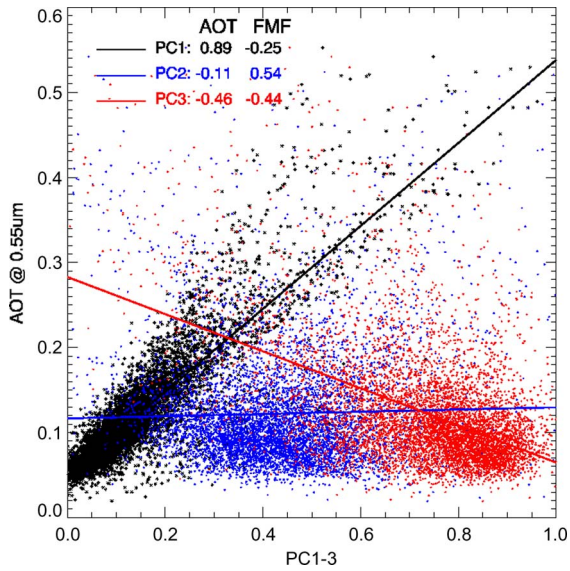


Fig. 2. Scatterplot of monthly averaged MODIS AOT at 0.55 $\mu$m as a function of normalized PC1, PC2, and PC3 values. Best-fit lines along with linear correlation coefficients between PC values and MODIS AOT and FMF are also given.

in the weighting coefficients. Thus, we are confident that nearly all of the physical signals present in the seven-channel radiance data are taken into account by the first three pseudochannels (PC1–PC3). These values are only representative of the one-month data set examined here; however, they remain relatively constant compared to the remaining 11 months of 2004: $74\% < \text{PC1} < 84\%$; $15\% < \text{PC2} < 24\%$; $1\% < \text{PC3} < 2\%$. The physical interpretation of PC1–PC3 also remains the same despite the presence of differing aerosol concentrations and types from month to month.

Raw weighting coefficients as a function of the MODIS wavelength are plotted in Fig. 1. The coefficients indicate the relative contribution of each original channel to the new pseudochannel. PC1 has, by far, the highest weightings compared to PC2 and PC3. Physically, PC1 represents a weighted mean or "multispectral" radiance value that is directly proportional to the aerosol concentration, with low values of PC1 corresponding to low values of the AOT (for all original wavelengths), whereas high PC1 values correspond to high values of the AOT at 0.55 $\mu$m (Fig. 2). The correlation between PC1 and the AOT at 0.55 $\mu$m is high ($r = 0.89$) and is generally linear for AOT $< 1.0$. This linear relationship is similar to that observed between the AOT and the broadband short-wave flux, which itself is highly correlated with the MODIS narrow-band short-wave observations [12].

A more interesting relationship exists in PC2. Weighting coefficients for PC2 decrease as the MODIS wavelength increases, and range from 0.78 at 0.47 $\mu$m (blue) to $-0.46$ at 2.12 $\mu$m (near infrared) with a weighting of near zero at 0.86 $\mu$m (Fig. 1). Due to its small weighting coefficient, the radiance at 0.86 $\mu$m contributes very little to PC2 values. These weightings indicate that PC2 is the most sensitive to the difference between blue and red radiance values, which would be greatest for aerosols that scatter more radiation in the blue portion of the spectrum and for aerosols that scatter proportionally less radiation in the near-infrared portion of the spectrum. Compared to PC1, PC2 is only weakly correlated with the AOT at 0.55 $\mu$m, with a correlation coefficient of only $-0.11$ (Fig. 2). However, the correlation between PC2 and the FMF at 0.55 $\mu$m is much higher, i.e., 0.54, indicating the dependence on the aerosol size. A similar result was observed by Tanre *et al.* [5], who also noted that this relationship was the most important for small-sized aerosols ($r_e < 0.3$ $\mu$m). Previous research has observed that anthropogenic aerosols, which are primarily small mode in nature, often have higher blue versus near-infrared radiance values (e.g., [13]). Thus, PC2 could be useful for separating naturally occurring from anthropogenic aerosol regions since it is correlated with aerosol size properties.

The interpretation of PC3 is not as clear as those of PC1 and PC2. Weighting coefficient values are quite small ($< \pm 0.2$) and are positive for MODIS 0.47-, 1.24-, and 2.12-$\mu$m channels, but negative for the remaining four channels. PC3 shows a modest inverse correlation with the AOT and the FMF at 0.55 $\mu$m ($r = -0.46, -0.44$; Fig. 2). Interestingly, the shape of the weighting coefficient curve mirrors that observed for PC1 to some extent, which is consistent with the reversed sign of the AOT correlation. Although PC3 only explains a small proportion of the total variance, the significance of the AOT and FMF correlations suggests that it is indeed an actual signal and not random noise. The correlation between PC4–PC7 and the AOT or the FMF at 0.55 $\mu$m never exceeds 0.15, making the physical interpretation of these channels difficult if such an interpretation exists at all.

Additional evidence for the physical interpretation of PC channels can be found when comparing multispectral MODIS and AERONET AOT values for locations where the predominant aerosol source is known. The locations selected are Capo Verde, Gosan (Korea), and Ascension Island, whose aerosol concentrations are *largely* from dust, fossil fuel combustion, and biomass burning, respectively. Fig. 3 shows MODIS and AERONET AOT values as a function of the wavelength. For all three locations, MODIS AOT decreases as a function of the wavelength. However, the decrease is much more significant (compared to the average AOT across all channels) for Gosan

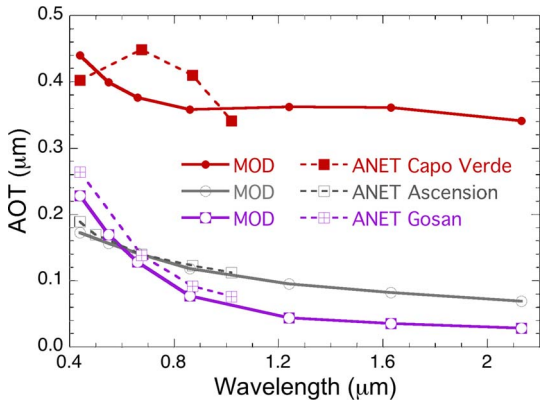


Fig. 3. MODIS (MOD) and AERONET (ANET) AOT between 0.4 and 2.2 $\mu$m for three sites, Capo Verde, Gosan (Korea), and Ascension Island, each representing significant dust, industrial pollution, and biomass burning aerosol concentrations, respectively. Note the similarities between Gosan and Ascension Island AOT and PC2 weighting coefficients.

(Korea) and Ascension Island locations, where small-mode aerosol concentrations are significant [14]. AERONET AOT for these two sites closely follows that of MODIS AOT, although some differences were observed at Capo Verde. Levy *et al.* [15] noted the differences in the multispectral response to dust aerosols between MODIS and AERONET measurements. Their differences are different than those observed here; however, we use Collection 5 data, which have improved coarse-mode aerosol models not used by Levy *et al.* [15]. However, the spectral response for small-mode aerosols for both MODIS and AERONET AOT remains clearly significant. Also, the decreasing AOT as a function of increasing wavelength for these aerosols corresponds quite well with the weighting coefficients for PC2, indicating that it is indeed sensitive to the aerosol size (Figs. 1 and 3).

The usefulness of this multispectral approach toward aerosol classification is evident when creating a three-band overlay using normalized PC1, PC2, and PC3 values (Fig. 4). Here, high values of PC1, PC2, and PC3 are denoted by red, green, and blue colors, respectively. Brightness increases as the values for PC1–PC3 become larger. It is immediately clear that aerosol speciation information is evident when utilizing the first three PC channels. Regions normally associated with dust (eastern North Atlantic and western Arabian Sea) in August appear as pink and red, which indicates that high values of PC1 correspond to high concentrations of dust aerosols compared to other aerosol types. Yellow-green regions are present in the South Atlantic Ocean off west of Africa and, to a lesser degree, in Indonesia and east of China. Anthropogenic aerosols, in the form of sulfates and organic carbon, are the predominant aerosol types in these regions. The signal is most evident in the South Atlantic, where large amounts of aerosols from biomass burning are transported over the ocean. Since the AOT is high and the FMF is low, PC1 and PC2 values are high, resulting in the yellow color observed. Light blues and greens are also evident off the east coasts of the U.S. and northern Europe. Compared to the dust and biomass burning regions, the total AOT is low (low red values), but the anthropogenic (particularly sulfate) proportion of aerosols is high resulting in the greenish appearance. Regions over the open oceans appear blue, corresponding to high values of PC3. Here, aerosol concentrations are generally low (low PC1) and consist of maritime sea salt and DMS (low PC2). In the Southern Hemisphere south of 35 °S, somewhat higher values of PC1 were observed, possibly resulting from larger sea-salt concentrations due to high wind speeds or cloud contamination [16].

### B. PC Values Compared to GOCART Aerosol Components

To determine the sensitivity of each PC channel to specific aerosol properties, PC1–PC3 values are compared to monthly averaged GOCART model-generated aerosol speciation over the global oceans for August 2004. Fig. 5(a)–(c) shows average MODIS AOT and FMF at 0.55 $\mu$m as a function of normalized PC1–PC3, binned at 0.05 intervals. Also plotted is the percentage of the total GOCART AOT at 0.55 $\mu$m accounted for by one of the six aerosol species, including sea salt (SS), dust (DU), black carbon (BC), anthropogenic organic carbon (OC), anthropogenic sulfate (SU), and naturally occurring sulfate (NSU). MODIS AOT increases from 0.08 to 0.60 as PC1 increases from 0 to 1.0, as should be the case given the high correlation between PC1 and short-wave radiance at most channels [Fig. 5(a)]. The FMF remains roughly constant at $\sim 0.5$ when $0.2 < \mathrm{PC1} < 0.8$, although it does vary at the low and high extremes of PC1. Comparing PC1 values to GOCART aerosol speciation shows several interesting features. The naturally occurring portion of organic carbon was always small, never being greater than 3%, and, thus, was not plotted. DMS accounts for up to 15% when $\mathrm{PC1} < 0.2$, but less than 10% thereafter. The anthropogenic sulfate concentration remains between 20% and 25% for all PC1 bins. However, the dust aerosol concentration does substantially change as a function of PC1, ranging from 15% to 50% at the expense of sea salt and natural and organic sulfate concentrations. Since low aerosol concentrations often occur over the open oceans, the high sea-salt concentration at low PC1 values should exist, although, even for low PC1 values, the total sea-salt portion of the AOT does not exceed 30%. The low values of PC1 $(\mathrm{PC} < 0.2)$ correspond to relatively pristine ocean conditions, where sea salt and DMS are the primary aerosol species, whereas the high values of PC1 correspond to high dust aerosol concentrations. PC1 values show little sensitivity to black and organic carbon concentrations.

Aerosol properties also substantially change as a function of PC2 values [Fig. 5(b)]. Unlike PC1, PC2 is sensitive to changes in the aerosol size, evidenced by the increasing FMF values as a function of PC2. Moreover, unlike PC1, PC2 is not very sensitive to the total aerosol concentration with a very little change in MODIS AOT when $0.2 < \mathrm{PC2} < 0.8$, whereas the FMF increases from 0.3 to 0.7 over that same range. Corresponding to this increase in the FMF, anthropogenic organic carbon and sulfate, which are both generally considered small-mode aerosols, substantially increase accounting for over 50% of the total AOT when $\mathrm{PC2} > 0.8$. The portion of the AOT due to black carbon also slightly increases, but remains below 10%. The dust aerosol concentration generally represents less than 20% of the total AOT for all PC2 bins, with no apparent trends. The sea-salt concentration rapidly decreases with values

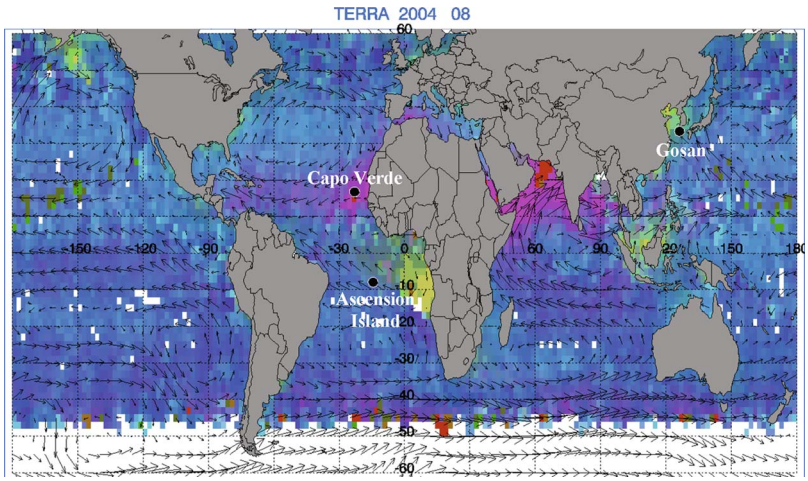


Fig. 4. Three-band overlay of normalized (red) PC1, (green) PC2, and (blue) PC3 values for August 2004. Locations of AERONET sites (Capo Verde, Gosan, and Ascension Island) are represented by black circles.

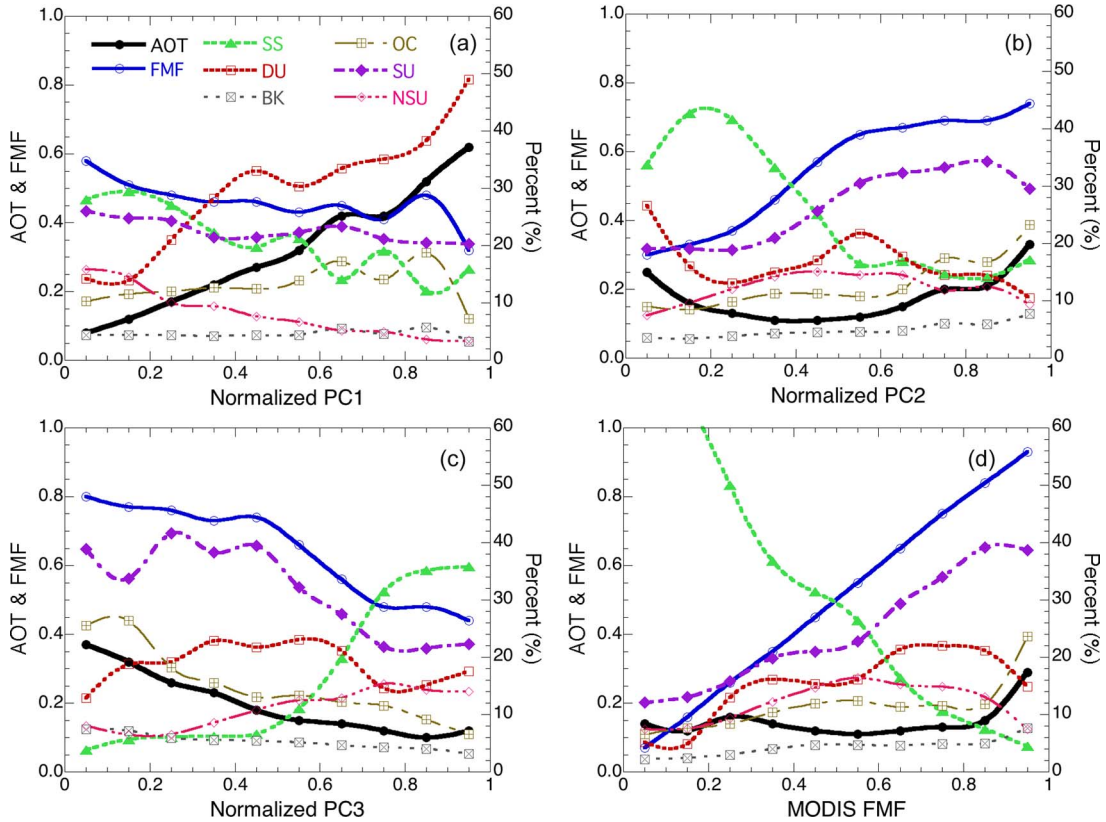


Fig. 5. MODIS AOT, FMF, and GOCART-modeled contribution to the total AOT for six major aerosol species (sea salt, dust, black carbon, anthropogenic organic carbon, anthropogenic sulfate, and naturally occurring sulfate) as a function of a normalized PC value for (a) PC1, (b) PC2, (c) PC3, and (d) FMF at 0.55 $\mu$m.

exceeding 40% for PC2 $< 0.3$, but decreasing to less than 20% thereafter. These results indicate that PC2 are the sensitive aerosols produced from anthropogenic sources, particularly biomass burning occurring off the African coast (Fig. 4). For PC2 greater than 0.8, the total anthropogenic component of the AOT $(\mathrm{BC} + \mathrm{OC} + \mathrm{SU})$ represents over 60% of the total AOT modeled by GOCART, whereas the combined dust and sea-salt contribution represents less than 30% of the total AOT.

The third PC channel, i.e., PC3, is an interesting case. Although it only accounts for 1.4% of the total variance, the weighting coefficients show a distinct pattern, somewhat mirroring the shape of PC1 weightings (Fig. 1). Fig. 2 suggests that PC3 is equally sensitive to the aerosol concentration as well as the aerosol size parameters. Comparing PC3 values against GOCART-modeled aerosol species also reveals that this PC is also useful in the separation of aerosol species. Unlike PC1 and PC2, sea-salt and DMS contributions increase for larger PC3 values that when combined, become $> 50\%$ for PC3 $>$ 0.8 [Fig. 5(c)]. A corresponding decrease in anthropogenic organic carbon and sulfate is also observed, with little overall

sensitivity to the dust aerosol concentration. The decrease in anthropogenic aerosol concentrations corresponds with similar decreases in MODIS FMF and AOT values [Figs. 2 and 5(c)]. Thus, PC3 can be said to be inversely proportional to the aerosol size and the total aerosol concentration, PC1 is primarily sensitive to the total aerosol concentration, and PC2 is primarily sensitive to the aerosol size but not to the concentration, consistent with Fig. 2. PC3 is statistically independent to PC2; thus, the multispectral radiance values are providing two independent sources of aerosol size and concentration information.

The exact physical mechanism behind PC3 remains unclear; however, spatial plots of PC3 alone (not shown) clearly show high values in maritime regions, whereas the plots show very low values in regions associated with aerosols transported from biomass burning sources. One possible hypothesis is the sensitivity of the 0.47-$\mu$m channels toward ocean reflectance, which also happens to have a positive weighting coefficient in PC3 (Fig. 1). Thus, it is possible that the high values of PC3 correspond to low AOT concentrations when the ocean reflectance signal is the strongest. Similarly, PC3 is minimized when the AOT is high, masking the ocean reflectance signal from the satellite. However, this hypothesis still does not explain the positive coefficients present at 1.63- and 2.12-$\mu$m wavelengths. This remains a question for future research to answer. For higher order PC channels (PC4–PC7), there appears to be very little sensitivity to aerosol properties. With the exception of weak sensitivity to the natural sulfate concentration for PC5, GOCART-modeled aerosol properties showed no consistent trends as a function of PC values (not shown). For the purposes of aerosol classification, these channels provide little in the way of useful information.

### C. MODIS FMF Values Compared to GOCART Aerosol Components

To determine if principal components are indeed useful, we perform the same analysis as above, but for MODIS FMF values [Fig. 5(d)]. It is quite evident that FMF values are an excellent way to distinguish among sea-salt and anthropogenic sulfate and organic carbon aerosols. When FMF $< 0.2$, the proportion of the AOT from sea salt exceeds 60%, whereas the anthropogenic aerosol contribution approaches 60% for FMF $> 0.8$, with most of the anthropogenic contribution being from sulfate. However, the FMF alone is not very sensitive to the dust aerosol concentration. There is an increase in the dust proportion from $\sim$10% at a low FMF to $\sim$20% at a high FMF, which actually contradicts the expectations. Generally, dust aerosols have larger particle sizes than the anthropogenic components: thus, one would expect the dust concentrations to be greatest somewhere around FMF $\approx 0.5$ [1]. At least for the time period used here, dust aerosols are sufficiently mixed with other aerosol types for their signal to be not apparent in MODIS FMF observations alone. However, performing the PCA on the raw MODIS radiance values produces a pseudochannel (PC1) that is sensitive to the dust aerosol concentration to a much higher degree. Since PC1 is highly correlated with the total AOT, one might expect that the total AOT is also sensitive to the dust aerosol concentration; however, this, again, was not observed here (not shown). Rather, we believe that it is the multispectral response of dust aerosols compared to anthropogenic types that produces this signal. These differences are apparent in Fig. 1, with the AERONET AOT dust spectral dependence being of different shape than the other two aerosol types. This shape is quite similar to the weighting value curve plotted in Fig. 1; thus, high concentrations of dust are most likely to result in high values of PC1.

### D. Application of Weighting to Independent Data

To access the applicability of the weighting coefficients to independent data, we apply the coefficients derived from the Terra data above to the MODIS data from the Aqua satellite acquired for the same month (August 2004). Aside from the Aqua overpass being 3–4 h after Terra, everything else related to the retrievals of aerosol radiance should be the same. If we assume that aerosol properties do not significantly change in this short time, applying the Terra weighting coefficients to Aqua data should produce good results. The Terra weighting coefficients are applied to the Aqua MODIS radiance values producing one set of PC data, which are compared with another set of PC data produced from the Aqua radiance values using the weighing coefficients derived from the Aqua data itself. The resulting correlation between these data is greater than 0.99 for PC1–PC3. For higher order PC channels, the correlation substantially decreases, as random noise becomes a more significant factor.

## IV. Conclusion

We have shown that multispectral AOT data from the MODIS satellite in conjunction with GOCART can be used to extract important aerosol type information without the use of small-mode AOT values. The PCA condenses seven highly correlated radiance channels into a smaller number of independent PC channels, each with its own physical interpretation. Using the monthly averaged GOCART data as a guide, it becomes apparent that different PC channels are sensitive to various aerosol species and mixtures thereof. We show that PC1 and PC3 can distinguish between pristine (sea salt and DMS) and heavy dust aerosol concentrations, whereas PC2 is very sensitive to anthropogenic aerosol concentrations, with high PC2 values corresponding to a 60% or greater anthropogenic component to the total AOT. PC data, particularly PC1, are much more sensitive to the dust aerosol concentration compared to the use of simple FMF thresholds. Since dust aerosol concentrations are often the most difficult type to extract on an objective basis, perhaps future works could take advantage of the approach examined here. It is likely that the use of higher spatial and temporal resolution aerosol speciation modeling allows for even better multispectral aerosol characteristics to be drawn. With higher resolution GOCART data, higher order PC channels may also be able to provide additional information into the observed aerosol properties. Given the problems noted with satellite-derived aerosol parameters over land, the use of multispectral information provides another, possibly more effective, avenue for aerosol classification over land. Future research will analyze

this possibility using larger data sets encompassing multiple years of observations.

## ACKNOWLEDGMENT

The Clouds and the Earth's Radiant Energy System (CERES) Single Scanner Footprint data, which contain the merged MODIS and CERES and the Measurements of Pollution in the Troposphere data were obtained through the NASA Langley Distributed Active Archive Systems. The MODIS daily products were obtained from the Goddard Distributed Active Archive Systems. The authors would like to thank Dr. M. Chin for providing the GOCART results.

## REFERENCES

[1] Y. J. Kaufman, O. Boucher, D. Tanre, M. Chin, L. A. Remer, and T. Takemura, "Aerosol anthropogenic component estimated from satellite data," *Geophys. Res. Lett.*, vol. 32, no. 17, L17 804, 2005. DOI:10.1029/2005GL023125.

[2] L. A. Remer, Y. J. Kaufman, D. Tanré, S. Mattoo, D. A. Chu, J. V. Martins, R.-R. Li, C. Ichoku, R. C. Levy, R. G. Kleidman, T. F. Eck, E. Vermote, and B. N. Holben, "The MODIS aerosol algorithm, products, and validation," *J. Atmos. Sci.*, vol. 62, no. 4, pp. 947–973, Apr. 2005.

[3] T. L. Anderson *et al.*, "An 'A-Train' strategy for quantifying direct climate forcing by anthropogenic aerosols," *Bull. Amer. Meteorol. Soc.*, vol. 86, no. 12, pp. 1795–1809, Dec. 2005.

[4] T. A. Jones and S. A. Christopher, "Statistical variability of top of atmosphere cloud-free shortwave aerosol radiative effect," *Atmos. Chem. Phys. Discuss.*, vol. 7, pp. 2937–2948, 2007.

[5] D. Tanre, M. Herman, and Y. J. Kaufman, "Information on aerosol size distribution contained in solar reflected spectral radiances," *J. Geophys. Res.*, vol. 101, no. D14, pp. 19 043–19 060, 1996.

[6] V. Zubko, Y. J. Kaufman, R. I. Burg, and J. V. Martins, "Principal component analysis of remote sensing of aerosols over oceans," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 730–745, Mar. 2007.

[7] L. A. Remer, D. Tanre, and Y. Kaufman, *Algorithm for Remote Sensing of Tropospheric Aerosol From MODIS: Collection 5*, 2006.

[8] B. N. Holben *et al.*, "AERONET—A federated instrument network and data archive for aerosol characterization," *Remote Sens. Environ.*, vol. 66, no. 1, pp. 1–16, Oct. 1998.

[9] M. Chin *et al.*, "Aerosol distributions and radiative properties simulated in the GOCART model and comparisons with observations," *J. Atmos. Sci.*, vol. 59, pp. 461–483, 2002.

[10] T. A. Jones, K. M. McGrath, and J. T. Snow, "Association between NSSL mesocyclone detection algorithm-detected vortices and tornadoes," *Weather Forecast.*, vol. 19, no. 5, pp. 872–890, Oct. 2004.

[11] D. S. Wilks, *Statistical Methods in the Atmospheric Sciences*. San Diego, CA: Academic, 2006.

[12] N. G. Loeb and N. Manalo-Smith, "Top-of-atmosphere direct radiative effect of aerosols over global oceans from merged CERES and MODIS observations," *J. Climate*, vol. 18, no. 17, pp. 3506–3526, Sep. 2005.

[13] T. F. Eck *et al.*, "Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols," *J. Geophys. Res.*, vol. 104, no. D24, pp. 31 333–31 349, 1999.

[14] A. Smirnov, B. N. Holben, Y. J. Kaufman, O. Dubovik, T. F. Eck, I. Slutsker, C. Pietras, and R. N. Halthore, "Optical properties of atmospheric aerosol in maritime environments," *J. Atmos. Sci.*, vol. 59, no. 3, pp. 501–523, Feb. 2002.

[15] R. C. Levy, L. A. Remer, S. Mattoo, E. F. Vermote, and Y. J. Kaufman, "Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of moderate resolution imaging spectroradiometer spectral reflectance," *J. Geophys. Res.*, vol. 112, no. D13, D13 211, Jul. 2007. DOI: 10.1029/2006JD007811.

[16] J. Zhang, J. S. Reid, and B. N. Holben, "An analysis of potential cloud artifacts in MODIS over ocean aerosol optical thickness products," *Geophys. Res. Lett.*, vol. 32, no. 15, L15 803, Aug. 2005. DOI: 10.1029/2005GL023254.

**Thomas A. Jones** received the B.S. and M.S. degrees in meteorology from the University of Oklahoma, Norman, in 2000 and 2002, respectively, while focusing on the development of regional thunderstorm climatologies, and the Ph.D. degree from The University of Alabama in Huntsville, Huntsville, in 2006 while specializing in improving hurricane intensity change forecasts using satellite-derived passive microwave imagery.

Since May 2006, he has been a Research Scientist with The University of Alabama in Huntsville. He has a wide range of research interests within the atmospheric sciences community. He is currently researching the effect of aerosols, and their subspecies, on short-wave and long-wave radiative effects using MODIS and CERES data. He is extending this research to include the effects of meteorological conditions on aerosols and the aerosol's effect to clouds.

**Sundar A. Christopher** received the M.S. degree in meteorology from the South Dakota School of Mines and Technology, Rapid City, in 1989, the M.S. degree in industrial/organizational psychology from The University of Alabama in Huntsville (UAH), Huntsville, in 2002, and the Ph.D. degree in atmospheric sciences from the Colorado State University, Fort Collins, in 1995.

He is currently a Professor with the Department of Atmospheric Sciences and the Associate Director of the Earth System Science Center, both at UAH. His research interests include satellite remote sensing of clouds and aerosols and their impact on air quality and global and regional climate.

# Multispectral Analysis of Aerosols Over Oceans Using Principal Components

Thomas A. Jones and Sundar A. Christopher

***Abstract*—Applying principal component analysis (PCA) to one month of Moderate Resolution Imaging Spectroradiometer (MODIS) narrow-band short-wave radiance data and comparing with the Goddard Global Ozone Chemistry Aerosol Radiation Transport (GOCART) model simulations, we show that aerosol size and speciation information can be inferred from multispectral radiance information without having to use other parameters, such as a fine mode fraction (FMF), that are difficult to validate. PCA was applied to seven highly correlated MODIS solar channels (0.47, 0.55, 0.66, 0.86, 1.24, 1.64, and 2.12 $\mu$m) to extract noncorrelated pseudochannels, each with a unique interpretation. The first pseudochannel (PC1) can be interpreted as the mean radiance across the seven channels, which is directly proportional to the aerosol concentration. The second pseudochannel (PC2) is sensitive to the aerosol size since different aerosol types scatter and absorb differently across the seven MODIS short-wave channels. PC3 is inversely related to the aerosol optical thickness (AOT) and the FMF and appears most sensitive to changes in sulfate and maritime sea-salt concentrations. Results indicate that high values of PC1 are indicative of high dust aerosol concentrations comprising more than 40% of the total AOT, whereas high values of PC2 indicate anthropogenic aerosol concentrations (deduced from GOCART) in excess of 60%. Compared to simple 0.55-$\mu$m FMF thresholds, the PC channels are much more sensitive to dust aerosol concentrations and certain aspects of anthropogenic aerosols, with very low FMF values alone ($< 0.2$) being the best indicator of predominately sea-salt aerosol concentrations. Our results indicate that PCA could be used as an alternate method for inferring aerosol speciation information in future research over ocean and more complex land surfaces.**

***Index Terms*—Aerosols, statistics.**

## I. Introduction

With the advent of reliable satellite-derived data sets of global aerosol properties, the climate effect of specific aerosol types, particularly those anthropogenic in origin, has become an important focus for research during the past 5 years [1]. Moderate Resolution Imaging Spectroradiometer (MODIS) provides radiance measurements in 36 channels from the ultraviolet to the thermal part of the electromagnetic spectrum. The MODIS aerosol algorithm over ocean uses the solar radiance at seven channels (0.47, 0.55, 0.66, 0.86, 1.24, 1.64, and 2.12 $\mu$m) [2] and then retrieves the aerosol optical thickness (AOT) using a lookup table approach [2]. The multiwavelength capability provides other retrievals such as the size, which is useful for separating natural from anthropogenic aerosols [1]. The spectral values of the AOT ($\tau$) contain information about the aerosol size, described by the Angstrom power law, i.e.,

$$\tau = \beta\lambda^{-\alpha}. \tag{1}$$

Here, $\alpha$ is the wavelength exponent, $\beta$ is the turbidity parameter, and $\lambda$ is the wavelength in micrometers. The value of $\tau$ is dependent on the ratio of large to small aerosols, with large $\tau$ indicating the presence of high concentrations of small aerosols (e.g., smoke aerosols from biomass burning), and small $\tau$ indicating the presence of large aerosols, e.g., dust. Observational studies have made use of this size information such as small $(0.1 < r_e < 0.25\ \mu\text{m})$ and coarse modes $(1.0 < r_e < 2.5\ \mu\text{m})$ [sac1]AOT at 0.55 $\mu$m to discriminate between various aerosol types (e.g., [1]). The ratio of small to total AOT is known as the fine mode fraction (FMF) and has proven quite successful in separating predominately naturally occurring (coarse mode) from anthropogenic (small mode) aerosols [1], although it is a difficult parameter to validate [3]. In this method, MODIS AOT and FMF at 0.55 $\mu$m are used to separate the total AOT into various aerosol components using several thresholds [1]. AOT components are separated by first assuming that the maritime AOT is solely a linear function of near-surface wind speed over the ocean. The dust and anthropogenic components of the AOT are then solved using a mathematical relationship derived between the total AOT, the maritime AOT, the observed FMF, and the FMF thresholds. However, there are many assumptions required when separating the total AOT into various components such as the FMF thresholds themselves, and assuming that all small-mode aerosols are anthropogenic also leads to uncertainties in aerosol component radiative effects [3], [4].

The aerosol classification used by Kaufman *et al.* [1] and others generally uses aerosol information from only a single channel (usually at 0.55 $\mu$m), overlooking differences in the multispectral response for different aerosol types. Since each channel is highly correlated, we apply the principal component analysis (PCA) to extract the independent signals that are present in the seven-channel radiance data. Tanre *et al.* [5] used a similar PCA approach using theoretical radiance values for MODIS aerosol channels and observed that the AOT for small-mode aerosols decreases as a function of the wavelength much more quickly than for coarse-mode aerosols. Zubko *et al.* [6] extended this technique to other parameters, such as the single

Manuscript received November 27, 2007; revised February 23, 2008. This work was supported in part by the National Aeronautics and Space Administration (NASA) Radiation Sciences Program, in part by the NASA Interdisciplinary Sciences Program, in part by an Earth Observing System Grant, and in part by the Atmospheric Chemistry Modeling and Analysis Program.

The authors are with the Department of Atmospheric Science, The University of Alabama in Huntsville, Huntsville, AL 35805 USA (e-mail: tjones@nsstc.uah.edu; sundar@nsstc.uah.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2008.920019

scattering albedo, the viewing angles, and the surface wind speed. Results from both efforts observed that two or three independent pieces of information can be retrieved from MODIS multispectral radiance values through the PCA, and that these pieces of information account for more than 90% of the total variance present within the raw data. However, very little research has occurred that utilizes the multispectral characteristics of various aerosol species in aerosol classification using the PCA. In this study, the PCA is applied to the Terra MODIS radiance data at each channel, thereby producing a new set of independent channels, with each new channel sensitive to different aerosol characteristics. These data are then compared with the Goddard Global Ozone Chemistry Aerosol Radiation Transport (GOCART) modeled aerosol speciation during the same month and the Aerosol Robotic Network (AERONET) data from several sites that are known to have a single predominant aerosol species. The goal is to determine if multispectral PCA will provide information on aerosol species that cannot be observed using the 0.55-$\mu$m AOT and/or FMF alone.

## II. Data and Methods

Terra MODIS radiance values at 0.47, 0.55, 0.66, 0.86, 1.24, 1.64, and 2.12 $\mu$m are obtained over the global oceans only from the MOD04, Collection 5 aerosol product [7], with a spatial resolution of 10 km$^2$ at nadir. We use the raw radiance values as inputs to the PCA rather than the retrieved AOT values at each wavelength since AOT retrievals are already a function of all seven radiance channels. For increasing aerosol concentrations, for scattering types of aerosols, larger amounts of short-wave radiation are scattered back to the satellite, resulting in a higher observed top of atmosphere radiance values. The AOT values over the ocean are retrieved from MODIS radiance values using a combination of four small-mode and five coarse-mode aerosol models with the relative amounts of small-mode to coarse-mode aerosols adjusted until a solution is found, which minimizes the error across the seven radiance values [2]. As a result, the reported AOT values at a certain wavelength are dependent on the radiance values from other channels. Using the retrieved AOT would induce uncertainties and biases that are present in the MODIS aerosol models into the research presented here. Thus, we chose to use the measured MODIS cloud-free radiance values as inputs to the PCA rather than any derived products. MODIS data are collected for August 2004 due to the variety of differing aerosol types around the globe during this month. MODIS AOT and FMF from the MOD04 level 2 product from the Terra satellite along with the AERONET AOT values are used to help interpret the PCA components. The AERONET is a network of well-calibrated ground-based sun photometers that provide spectral AOT and other aerosol properties from hundreds of locations around the world [8]. We use the data for August 22, August 16, and August 10, 2004, from AERONET sites Capo Verde, Gosan (Korea), and Ascension Island, which are compared and colocated with the MODIS AOT data for that day. Data from these days were selected based on times when both good quality AERONET and MODIS data exist. For a particular day, MODIS AOT data within $\pm 1^\circ$ of the AERONET site were averaged and compared to the AERONET observations. Each site location represents where the primary aerosol source (other than maritime sea salt) is known *a priori* from GOCART, including dust, fossil fuel combustion, and biomass burning, respectively.

The satellite-retrieved AOT and FMF are compared with the aerosol species simulations produced by the GOCART model [9]. GOCART simulates the transport of aerosols and their type over a global domain on a monthly averaged timescale. GOCART uses global emissions of aerosols and assimilated meteorological fields to separately calculate the mass loading of each aerosol species, which are then converted to the AOT using mass extinction coefficients. Aerosols are categorized as black carbon (BC), organic carbon (OC), sulfate (SU), dust (DU), and sea salt (SS). Organic carbon and sulfate are also split into anthropogenic and naturally occurring components, the latter generally from dimethyl sulfide (DMS). Monthly averaged data for August 2004 were acquired for an ocean only, global domain and were interpolated to a $2^\circ \times 2^\circ$ grid.

The statistical technique of the PCA enables the reduction of multiple and highly correlated data attributes into a fewer number of independent variables, each with a unique physical interpretation [10], [11]. The initial step in the PCA is the calculation of a correlation matrix between each of the input channels, which are the radiance values at all seven MODIS channels. Once the correlation matrix is computed, eigenvalues and vectors are calculated from the correlation matrix to determine the weighting coefficients, which are then applied to the radiance data to produce the new independent pseudochannels (or PC channels). Each PC channel is ordered in such a way that the first channel (PC1) accounts for the greatest variance in the raw data, with the next (PC2) accounting for the next highest amount of variance, and so on. For this paper, the total number of pseudochannels created is equal to the number of input channels, which is seven. For illustration purposes, PC values are normalized to a 0–1.0 scale and averaged to a $2^\circ \times 2^\circ$ resolution grid to correspond with the monthly averaged GOCART data resolution. To visualize the sensitivity of the PC channels to aerosol properties, we compare GOCART and MODIS data as a function of the PC value. PC values are binned at 0.05 intervals, and, for each bin, the ratio of the GOCART component AOT to the total AOT is calculated for grid cells within a particular PC bin. To compare the effectiveness of the PC classification with the 0.55-$\mu$m FMF classification, a similar procedure is used but using a 0.05 FMF interval bin.

## III. Results

### A. PC Weightings

As expected, solar radiance from each of the seven MODIS aerosol channels is highly correlated, although the correlation does decrease to $< 0.5$ when comparing 0.47- to 2.13-$\mu$m radiance values. Analysis of the eigenvalues associated with each PC channel indicates that 78.4% of the variance is explained by PC1, 19.7% by PC2, and 1.4% by PC3, with PC1–PC3 accounting for $\sim$99% of the total variance, similar to that expected in [5] and [6]. The amount explained by PC4–PC7 combined is less than 1%, with no consistent patterns observed

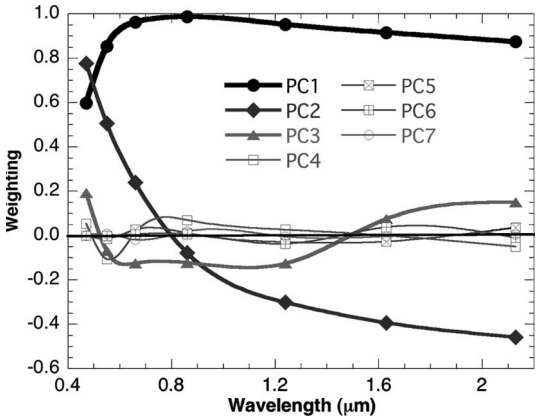


Fig. 1. Weighting coefficients derived from the PCA for PC channels 1–7. Magnitudes indicate the relative importance of each original channel to the resulting PC channel. Note that weightings for PC4–PC7 are very small compared to those for PC1–PC3.

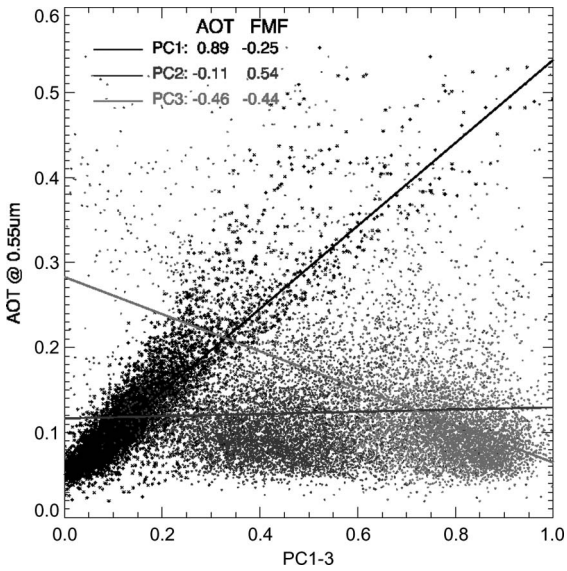


Fig. 2. Scatterplot of monthly averaged MODIS AOT at 0.55 $\mu$m as a function of normalized PC1, PC2, and PC3 values. Best-fit lines along with linear correlation coefficients between PC values and MODIS AOT and FMF are also given.

in the weighting coefficients. Thus, we are confident that nearly all of the physical signals present in the seven-channel radiance data are taken into account by the first three pseudochannels (PC1–PC3). These values are only representative of the one-month data set examined here; however, they remain relatively constant compared to the remaining 11 months of 2004: $74\% < \text{PC1} < 84\%$; $15\% < \text{PC2} < 24\%$; $1\% < \text{PC3} < 2\%$. The physical interpretation of PC1–PC3 also remains the same despite the presence of differing aerosol concentrations and types from month to month.

Raw weighting coefficients as a function of the MODIS wavelength are plotted in Fig. 1. The coefficients indicate the relative contribution of each original channel to the new pseudochannel. PC1 has, by far, the highest weightings compared to PC2 and PC3. Physically, PC1 represents a weighted mean or "multispectral" radiance value that is directly proportional to the aerosol concentration, with low values of PC1 corresponding to low values of the AOT (for all original wavelengths), whereas high PC1 values correspond to high values of the AOT at 0.55 $\mu$m (Fig. 2). The correlation between PC1 and the AOT at 0.55 $\mu$m is high ($r = 0.89$) and is generally linear for AOT $< 1.0$. This linear relationship is similar to that observed between the AOT and the broadband short-wave flux, which itself is highly correlated with the MODIS narrow-band short-wave observations [12].

A more interesting relationship exists in PC2. Weighting coefficients for PC2 decrease as the MODIS wavelength increases, and range from 0.78 at 0.47 $\mu$m (blue) to $-0.46$ at 2.12 $\mu$m (near infrared) with a weighting of near zero at 0.86 $\mu$m (Fig. 1). Due to its small weighting coefficient, the radiance at 0.86 $\mu$m contributes very little to PC2 values. These weightings indicate that PC2 is the most sensitive to the difference between blue and red radiance values, which would be greatest for aerosols that scatter more radiation in the blue portion of the spectrum and for aerosols that scatter proportionally less radiation in the near-infrared portion of the spectrum. Compared to PC1, PC2 is only weakly correlated with the AOT at 0.55 $\mu$m, with a correlation coefficient of only $-0.11$ (Fig. 2). However, the correlation between PC2 and the FMF at 0.55 $\mu$m is much higher, i.e., 0.54, indicating the dependence on the aerosol size. A similar result was observed by Tanre *et al.* [5], who also noted that this relationship was the most important for small-sized aerosols ($r_e < 0.3$ $\mu$m). Previous research has observed that anthropogenic aerosols, which are primarily small mode in nature, often have higher blue versus near-infrared radiance values (e.g., [13]). Thus, PC2 could be useful for separating naturally occurring from anthropogenic aerosol regions since it is correlated with aerosol size properties.

The interpretation of PC3 is not as clear as those of PC1 and PC2. Weighting coefficient values are quite small ($< \pm 0.2$) and are positive for MODIS 0.47-, 1.24-, and 2.12-$\mu$m channels, but negative for the remaining four channels. PC3 shows a modest inverse correlation with the AOT and the FMF at 0.55 $\mu$m ($r = -0.46, -0.44$; Fig. 2). Interestingly, the shape of the weighting coefficient curve mirrors that observed for PC1 to some extent, which is consistent with the reversed sign of the AOT correlation. Although PC3 only explains a small proportion of the total variance, the significance of the AOT and FMF correlations suggests that it is indeed an actual signal and not random noise. The correlation between PC4–PC7 and the AOT or the FMF at 0.55 $\mu$m never exceeds 0.15, making the physical interpretation of these channels difficult if such an interpretation exists at all.

Additional evidence for the physical interpretation of PC channels can be found when comparing multispectral MODIS and AERONET AOT values for locations where the predominant aerosol source is known. The locations selected are Capo Verde, Gosan (Korea), and Ascension Island, whose aerosol concentrations are *largely* from dust, fossil fuel combustion, and biomass burning, respectively. Fig. 3 shows MODIS and AERONET AOT values as a function of the wavelength. For all three locations, MODIS AOT decreases as a function of the wavelength. However, the decrease is much more significant (compared to the average AOT across all channels) for Gosan

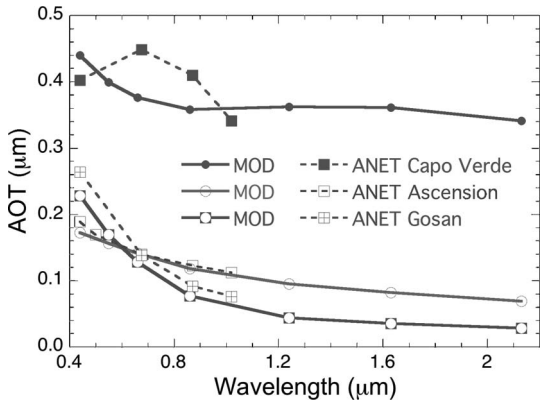


Fig. 3. MODIS (MOD) and AERONET (ANET) AOT between 0.4 and 2.2 $\mu$m for three sites, Capo Verde, Gosan (Korea), and Ascension Island, each representing significant dust, industrial pollution, and biomass burning aerosol concentrations, respectively. Note the similarities between Gosan and Ascension Island AOT and PC2 weighting coefficients.

(Korea) and Ascension Island locations, where small-mode aerosol concentrations are significant [14]. AERONET AOT for these two sites closely follows that of MODIS AOT, although some differences were observed at Capo Verde. Levy *et al.* [15] noted the differences in the multispectral response to dust aerosols between MODIS and AERONET measurements. Their differences are different than those observed here; however, we use Collection 5 data, which have improved coarse-mode aerosol models not used by Levy *et al.* [15]. However, the spectral response for small-mode aerosols for both MODIS and AERONET AOT remains clearly significant. Also, the decreasing AOT as a function of increasing wavelength for these aerosols corresponds quite well with the weighting coefficients for PC2, indicating that it is indeed sensitive to the aerosol size (Figs. 1 and 3).

The usefulness of this multispectral approach toward aerosol classification is evident when creating a three-band overlay using normalized PC1, PC2, and PC3 values (Fig. 4). Here, high values of PC1, PC2, and PC3 are denoted by red, green, and blue colors, respectively. Brightness increases as the values for PC1–PC3 become larger. It is immediately clear that aerosol speciation information is evident when utilizing the first three PC channels. Regions normally associated with dust (eastern North Atlantic and western Arabian Sea) in August appear as pink and red, which indicates that high values of PC1 correspond to high concentrations of dust aerosols compared to other aerosol types. Yellow-green regions are present in the South Atlantic Ocean off west of Africa and, to a lesser degree, in Indonesia and east of China. Anthropogenic aerosols, in the form of sulfates and organic carbon, are the predominant aerosol types in these regions. The signal is most evident in the South Atlantic, where large amounts of aerosols from biomass burning are transported over the ocean. Since the AOT is high and the FMF is low, PC1 and PC2 values are high, resulting in the yellow color observed. Light blues and greens are also evident off the east coasts of the U.S. and northern Europe. Compared to the dust and biomass burning regions, the total AOT is low (low red values), but the anthropogenic (particularly sulfate) proportion of aerosols is high resulting in the greenish appearance. Regions over the open oceans appear blue, corresponding to high values of PC3. Here, aerosol concentrations are generally low (low PC1) and consist of maritime sea salt and DMS (low PC2). In the Southern Hemisphere south of 35 °S, somewhat higher values of PC1 were observed, possibly resulting from larger sea-salt concentrations due to high wind speeds or cloud contamination [16].

### B. PC Values Compared to GOCART Aerosol Components

To determine the sensitivity of each PC channel to specific aerosol properties, PC1–PC3 values are compared to monthly averaged GOCART model-generated aerosol speciation over the global oceans for August 2004. Fig. 5(a)–(c) shows average MODIS AOT and FMF at 0.55 $\mu$m as a function of normalized PC1–PC3, binned at 0.05 intervals. Also plotted is the percentage of the total GOCART AOT at 0.55 $\mu$m accounted for by one of the six aerosol species, including sea salt (SS), dust (DU), black carbon (BC), anthropogenic organic carbon (OC), anthropogenic sulfate (SU), and naturally occurring sulfate (NSU). MODIS AOT increases from 0.08 to 0.60 as PC1 increases from 0 to 1.0, as should be the case given the high correlation between PC1 and short-wave radiance at most channels [Fig. 5(a)]. The FMF remains roughly constant at $\sim 0.5$ when $0.2 < \mathrm{PC1} < 0.8$, although it does vary at the low and high extremes of PC1. Comparing PC1 values to GOCART aerosol speciation shows several interesting features. The naturally occurring portion of organic carbon was always small, never being greater than 3%, and, thus, was not plotted. DMS accounts for up to 15% when $\mathrm{PC1} < 0.2$, but less than 10% thereafter. The anthropogenic sulfate concentration remains between 20% and 25% for all PC1 bins. However, the dust aerosol concentration does substantially change as a function of PC1, ranging from 15% to 50% at the expense of sea salt and natural and organic sulfate concentrations. Since low aerosol concentrations often occur over the open oceans, the high sea-salt concentration at low PC1 values should exist, although, even for low PC1 values, the total sea-salt portion of the AOT does not exceed 30%. The low values of PC1 $(\mathrm{PC} < 0.2)$ correspond to relatively pristine ocean conditions, where sea salt and DMS are the primary aerosol species, whereas the high values of PC1 correspond to high dust aerosol concentrations. PC1 values show little sensitivity to black and organic carbon concentrations.

Aerosol properties also substantially change as a function of PC2 values [Fig. 5(b)]. Unlike PC1, PC2 is sensitive to changes in the aerosol size, evidenced by the increasing FMF values as a function of PC2. Moreover, unlike PC1, PC2 is not very sensitive to the total aerosol concentration with a very little change in MODIS AOT when $0.2 < \mathrm{PC2} < 0.8$, whereas the FMF increases from 0.3 to 0.7 over that same range. Corresponding to this increase in the FMF, anthropogenic organic carbon and sulfate, which are both generally considered small-mode aerosols, substantially increase accounting for over 50% of the total AOT when $\mathrm{PC2} > 0.8$. The portion of the AOT due to black carbon also slightly increases, but remains below 10%. The dust aerosol concentration generally represents less than 20% of the total AOT for all PC2 bins, with no apparent trends. The sea-salt concentration rapidly decreases with values

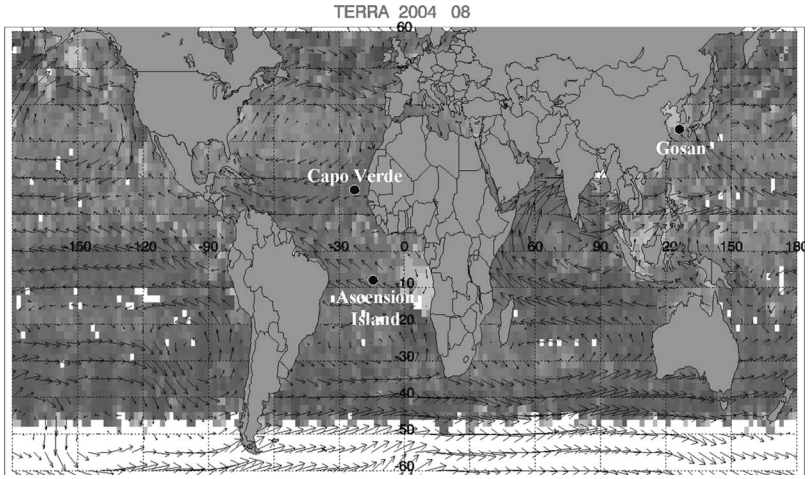


Fig. 4. Three-band overlay of normalized (red) PC1, (green) PC2, and (blue) PC3 values for August 2004. Locations of AERONET sites (Capo Verde, Gosan, and Ascension Island) are represented by black circles.

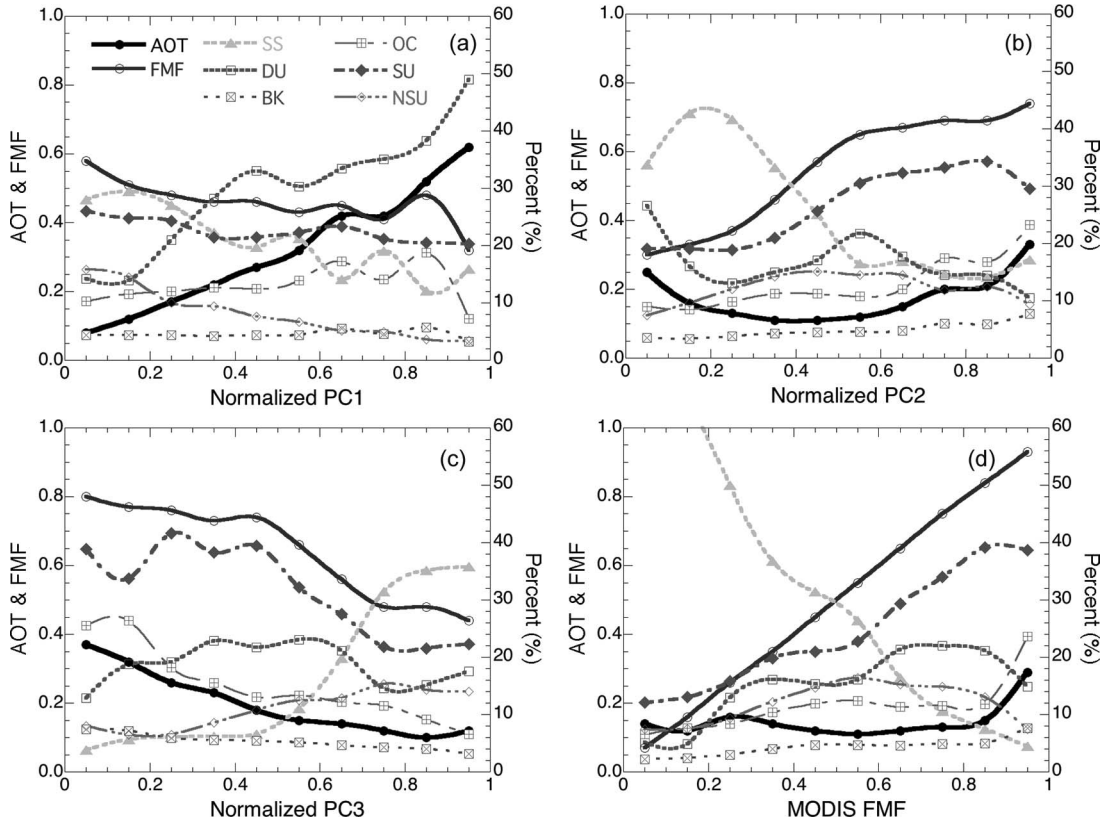


Fig. 5. MODIS AOT, FMF, and GOCART-modeled contribution to the total AOT for six major aerosol species (sea salt, dust, black carbon, anthropogenic organic carbon, anthropogenic sulfate, and naturally occurring sulfate) as a function of a normalized PC value for (a) PC1, (b) PC2, (c) PC3, and (d) FMF at 0.55 $\mu$m.

exceeding 40% for PC2 $< 0.3$, but decreasing to less than 20% thereafter. These results indicate that PC2 are the sensitive aerosols produced from anthropogenic sources, particularly biomass burning occurring off the African coast (Fig. 4). For PC2 greater than 0.8, the total anthropogenic component of the AOT $(\mathrm{BC} + \mathrm{OC} + \mathrm{SU})$ represents over 60% of the total AOT modeled by GOCART, whereas the combined dust and sea-salt contribution represents less than 30% of the total AOT.

The third PC channel, i.e., PC3, is an interesting case. Although it only accounts for 1.4% of the total variance, the weighting coefficients show a distinct pattern, somewhat mirroring the shape of PC1 weightings (Fig. 1). Fig. 2 suggests that PC3 is equally sensitive to the aerosol concentration as well as the aerosol size parameters. Comparing PC3 values against GOCART-modeled aerosol species also reveals that this PC is also useful in the separation of aerosol species. Unlike PC1 and PC2, sea-salt and DMS contributions increase for larger PC3 values that when combined, become $> 50\%$ for PC3 $>$ 0.8 [Fig. 5(c)]. A corresponding decrease in anthropogenic organic carbon and sulfate is also observed, with little overall

sensitivity to the dust aerosol concentration. The decrease in anthropogenic aerosol concentrations corresponds with similar decreases in MODIS FMF and AOT values [Figs. 2 and 5(c)]. Thus, PC3 can be said to be inversely proportional to the aerosol size and the total aerosol concentration, PC1 is primarily sensitive to the total aerosol concentration, and PC2 is primarily sensitive to the aerosol size but not to the concentration, consistent with Fig. 2. PC3 is statistically independent to PC2; thus, the multispectral radiance values are providing two independent sources of aerosol size and concentration information.

The exact physical mechanism behind PC3 remains unclear; however, spatial plots of PC3 alone (not shown) clearly show high values in maritime regions, whereas the plots show very low values in regions associated with aerosols transported from biomass burning sources. One possible hypothesis is the sensitivity of the 0.47-$\mu$m channels toward ocean reflectance, which also happens to have a positive weighting coefficient in PC3 (Fig. 1). Thus, it is possible that the high values of PC3 correspond to low AOT concentrations when the ocean reflectance signal is the strongest. Similarly, PC3 is minimized when the AOT is high, masking the ocean reflectance signal from the satellite. However, this hypothesis still does not explain the positive coefficients present at 1.63- and 2.12-$\mu$m wavelengths. This remains a question for future research to answer. For higher order PC channels (PC4–PC7), there appears to be very little sensitivity to aerosol properties. With the exception of weak sensitivity to the natural sulfate concentration for PC5, GOCART-modeled aerosol properties showed no consistent trends as a function of PC values (not shown). For the purposes of aerosol classification, these channels provide little in the way of useful information.

### C. MODIS FMF Values Compared to GOCART Aerosol Components

To determine if principal components are indeed useful, we perform the same analysis as above, but for MODIS FMF values [Fig. 5(d)]. It is quite evident that FMF values are an excellent way to distinguish among sea-salt and anthropogenic sulfate and organic carbon aerosols. When FMF $< 0.2$, the proportion of the AOT from sea salt exceeds 60%, whereas the anthropogenic aerosol contribution approaches 60% for FMF $> 0.8$, with most of the anthropogenic contribution being from sulfate. However, the FMF alone is not very sensitive to the dust aerosol concentration. There is an increase in the dust proportion from $\sim$10% at a low FMF to $\sim$20% at a high FMF, which actually contradicts the expectations. Generally, dust aerosols have larger particle sizes than the anthropogenic components: thus, one would expect the dust concentrations to be greatest somewhere around FMF $\approx 0.5$ [1]. At least for the time period used here, dust aerosols are sufficiently mixed with other aerosol types for their signal to be not apparent in MODIS FMF observations alone. However, performing the PCA on the raw MODIS radiance values produces a pseudochannel (PC1) that is sensitive to the dust aerosol concentration to a much higher degree. Since PC1 is highly correlated with the total AOT, one might expect that the total AOT is also sensitive to the dust aerosol concentration; however, this, again, was not observed here (not shown). Rather, we believe that it is the multispectral response of dust aerosols compared to anthropogenic types that produces this signal. These differences are apparent in Fig. 1, with the AERONET AOT dust spectral dependence being of different shape than the other two aerosol types. This shape is quite similar to the weighting value curve plotted in Fig. 1; thus, high concentrations of dust are most likely to result in high values of PC1.

### D. Application of Weighting to Independent Data

To access the applicability of the weighting coefficients to independent data, we apply the coefficients derived from the Terra data above to the MODIS data from the Aqua satellite acquired for the same month (August 2004). Aside from the Aqua overpass being 3–4 h after Terra, everything else related to the retrievals of aerosol radiance should be the same. If we assume that aerosol properties do not significantly change in this short time, applying the Terra weighting coefficients to Aqua data should produce good results. The Terra weighting coefficients are applied to the Aqua MODIS radiance values producing one set of PC data, which are compared with another set of PC data produced from the Aqua radiance values using the weighing coefficients derived from the Aqua data itself. The resulting correlation between these data is greater than 0.99 for PC1–PC3. For higher order PC channels, the correlation substantially decreases, as random noise becomes a more significant factor.

## IV. Conclusion

We have shown that multispectral AOT data from the MODIS satellite in conjunction with GOCART can be used to extract important aerosol type information without the use of small-mode AOT values. The PCA condenses seven highly correlated radiance channels into a smaller number of independent PC channels, each with its own physical interpretation. Using the monthly averaged GOCART data as a guide, it becomes apparent that different PC channels are sensitive to various aerosol species and mixtures thereof. We show that PC1 and PC3 can distinguish between pristine (sea salt and DMS) and heavy dust aerosol concentrations, whereas PC2 is very sensitive to anthropogenic aerosol concentrations, with high PC2 values corresponding to a 60% or greater anthropogenic component to the total AOT. PC data, particularly PC1, are much more sensitive to the dust aerosol concentration compared to the use of simple FMF thresholds. Since dust aerosol concentrations are often the most difficult type to extract on an objective basis, perhaps future works could take advantage of the approach examined here. It is likely that the use of higher spatial and temporal resolution aerosol speciation modeling allows for even better multispectral aerosol characteristics to be drawn. With higher resolution GOCART data, higher order PC channels may also be able to provide additional information into the observed aerosol properties. Given the problems noted with satellite-derived aerosol parameters over land, the use of multispectral information provides another, possibly more effective, avenue for aerosol classification over land. Future research will analyze

this possibility using larger data sets encompassing multiple years of observations.

## ACKNOWLEDGMENT

The Clouds and the Earth's Radiant Energy System (CERES) Single Scanner Footprint data, which contain the merged MODIS and CERES and the Measurements of Pollution in the Troposphere data were obtained through the NASA Langley Distributed Active Archive Systems. The MODIS daily products were obtained from the Goddard Distributed Active Archive Systems. The authors would like to thank Dr. M. Chin for providing the GOCART results.

## REFERENCES

[1] Y. J. Kaufman, O. Boucher, D. Tanre, M. Chin, L. A. Remer, and T. Takemura, "Aerosol anthropogenic component estimated from satellite data," *Geophys. Res. Lett.*, vol. 32, no. 17, L17 804, 2005. DOI:10.1029/2005GL023125.

[2] L. A. Remer, Y. J. Kaufman, D. Tanré, S. Mattoo, D. A. Chu, J. V. Martins, R.-R. Li, C. Ichoku, R. C. Levy, R. G. Kleidman, T. F. Eck, E. Vermote, and B. N. Holben, "The MODIS aerosol algorithm, products, and validation," *J. Atmos. Sci.*, vol. 62, no. 4, pp. 947–973, Apr. 2005.

[3] T. L. Anderson *et al.*, "An 'A-Train' strategy for quantifying direct climate forcing by anthropogenic aerosols," *Bull. Amer. Meteorol. Soc.*, vol. 86, no. 12, pp. 1795–1809, Dec. 2005.

[4] T. A. Jones and S. A. Christopher, "Statistical variability of top of atmosphere cloud-free shortwave aerosol radiative effect," *Atmos. Chem. Phys. Discuss.*, vol. 7, pp. 2937–2948, 2007.

[5] D. Tanre, M. Herman, and Y. J. Kaufman, "Information on aerosol size distribution contained in solar reflected spectral radiances," *J. Geophys. Res.*, vol. 101, no. D14, pp. 19 043–19 060, 1996.

[6] V. Zubko, Y. J. Kaufman, R. I. Burg, and J. V. Martins, "Principal component analysis of remote sensing of aerosols over oceans," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 730–745, Mar. 2007.

[7] L. A. Remer, D. Tanre, and Y. Kaufman, *Algorithm for Remote Sensing of Tropospheric Aerosol From MODIS: Collection 5*, 2006.

[8] B. N. Holben *et al.*, "AERONET—A federated instrument network and data archive for aerosol characterization," *Remote Sens. Environ.*, vol. 66, no. 1, pp. 1–16, Oct. 1998.

[9] M. Chin *et al.*, "Aerosol distributions and radiative properties simulated in the GOCART model and comparisons with observations," *J. Atmos. Sci.*, vol. 59, pp. 461–483, 2002.

[10] T. A. Jones, K. M. McGrath, and J. T. Snow, "Association between NSSL mesocyclone detection algorithm-detected vortices and tornadoes," *Weather Forecast.*, vol. 19, no. 5, pp. 872–890, Oct. 2004.

[11] D. S. Wilks, *Statistical Methods in the Atmospheric Sciences*. San Diego, CA: Academic, 2006.

[12] N. G. Loeb and N. Manalo-Smith, "Top-of-atmosphere direct radiative effect of aerosols over global oceans from merged CERES and MODIS observations," *J. Climate*, vol. 18, no. 17, pp. 3506–3526, Sep. 2005.

[13] T. F. Eck *et al.*, "Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols," *J. Geophys. Res.*, vol. 104, no. D24, pp. 31 333–31 349, 1999.

[14] A. Smirnov, B. N. Holben, Y. J. Kaufman, O. Dubovik, T. F. Eck, I. Slutsker, C. Pietras, and R. N. Halthore, "Optical properties of atmospheric aerosol in maritime environments," *J. Atmos. Sci.*, vol. 59, no. 3, pp. 501–523, Feb. 2002.

[15] R. C. Levy, L. A. Remer, S. Mattoo, E. F. Vermote, and Y. J. Kaufman, "Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of moderate resolution imaging spectroradiometer spectral reflectance," *J. Geophys. Res.*, vol. 112, no. D13, D13 211, Jul. 2007. DOI: 10.1029/2006JD007811.

[16] J. Zhang, J. S. Reid, and B. N. Holben, "An analysis of potential cloud artifacts in MODIS over ocean aerosol optical thickness products," *Geophys. Res. Lett.*, vol. 32, no. 15, L15 803, Aug. 2005. DOI: 10.1029/2005GL023254.

**Thomas A. Jones** received the B.S. and M.S. degrees in meteorology from the University of Oklahoma, Norman, in 2000 and 2002, respectively, while focusing on the development of regional thunderstorm climatologies, and the Ph.D. degree from The University of Alabama in Huntsville, Huntsville, in 2006 while specializing in improving hurricane intensity change forecasts using satellite-derived passive microwave imagery.

Since May 2006, he has been a Research Scientist with The University of Alabama in Huntsville. He has a wide range of research interests within the atmospheric sciences community. He is currently researching the effect of aerosols, and their subspecies, on short-wave and long-wave radiative effects using MODIS and CERES data. He is extending this research to include the effects of meteorological conditions on aerosols and the aerosol's effect to clouds.

**Sundar A. Christopher** received the M.S. degree in meteorology from the South Dakota School of Mines and Technology, Rapid City, in 1989, the M.S. degree in industrial/organizational psychology from The University of Alabama in Huntsville (UAH), Huntsville, in 2002, and the Ph.D. degree in atmospheric sciences from the Colorado State University, Fort Collins, in 1995.

He is currently a Professor with the Department of Atmospheric Sciences and the Associate Director of the Earth System Science Center, both at UAH. His research interests include satellite remote sensing of clouds and aerosols and their impact on air quality and global and regional climate.