The invigoration of deep convective clouds over the Atlantic: aerosol effect, meteorology or retrieval artifact?

I. Koren\textsuperscript{1}, G. Feingold\textsuperscript{2}, and L. A. Remer\textsuperscript{3}

\textsuperscript{1}Department of Environmental Science and Energy Research, Weizmann Institute of Science, Rehovot 76100, Israel
\textsuperscript{2}NOAA Earth System Research Laboratory, Boulder, CO, USA
\textsuperscript{3}Laboratory for Atmospheres, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Received: 11 January 2010 – Accepted: 25 January 2010 – Published: 10 February 2010

Correspondence to: I. Koren (ilan.koren@weizmann.ac.il)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

The effects of the aerosol on convective clouds in the tropical Atlantic Ocean are explored using satellite remote sensing, a chemical transport model, and a reanalysis of meteorological fields. Two of the most challenging problems are addressed: the potential for elements of the cloud field to be erroneously ascribed to aerosol optical depth; and the potential for correlations between aerosol and cloud parameters to be erroneously considered to be causal. Results show that there is a robust positive correlation between cloud fraction or cloud top height and the aerosol optical depth, regardless of whether a stringent filtering of aerosol measurements in the vicinity of clouds is applied, or not. These same positive correlations emerge when replacing the observed aerosol field with that derived from a chemical transport model. A correlation exercise between the full suite of meteorological fields derived from model reanalysis and satellite-derived cloud fields shows that observed cloud top height and cloud fraction correlate best with pressure updraft velocity and relative humidity. Observed aerosol optical depth does correlate with meteorological parameters but usually different parameters from those that correlate with observed cloud fields. The result is a near-orthogonal influence of aerosol and meteorological fields on cloud top height and cloud fraction. The results strengthen the case that the aerosol does play a role in invigorating convective clouds.

1 Introduction

Aerosol effects on clouds are recognized as contributing substantially to anthropogenic effects on climate and the water cycle. Understanding the different cloud feedbacks initiated by changes in aerosol properties poses one of the greatest challenges in climate, cloud and precipitation physics and radiative transfer (Ramanathan et al., 2001; Kaufman, 2002). The strong sensitivity of the climate system to clouds, and the steadily
increasing pressure on water resources, makes this a problem of major importance (IPCC, 2007).

But why are aerosol-cloud interactions so difficult to quantify? Some of the important aspects of this problem are enumerated below:

Complexity: The inherent complexity of clouds is such that the system is not amenable to analytical solution, nor to observation or model simulation at the full range of temporal and spatial scales. The sensitivity to initial and boundary conditions in the form of thermodynamic, radiative, and aerosol properties is inherently non-linear, so that small changes in the initial conditions can propagate to large ones in the size, shape, microphysical properties and evolution of the cloud. It is not always clear which of the non-linear feedbacks will be ignited by changes in the initial or boundary conditions of the system. The aerosol can modify cloud radiative properties (Twomey, 1977) but also the ability of the cloud to precipitate (Albrecht, 1989; Rosenfeld et al., 1999; Andreae et al., 2004) and the pathways via which precipitation develops (e.g., Khain et al., 2001; Lee et al., 2008). Modification to precipitation influences the dynamics of the environment by changing the vertical distribution of latent heat. It has been suggested that clouds growing in polluted environments are characterized by more vigorous convection owing to a combination of suppression of early rainout of the cloud and its attendant stabilization, together with suppression of freezing, and the eventual release of latent heat at higher altitudes (Koren et al., 2005; Khain et al., 2005). Aerosol perturbations to deep convective systems may influence secondary convection (Seifert et al., 2006; van den Heever and Cotton, 2007; Lee et al., 2008) and other dynamical responses that magnify the initial microphysical perturbation. Conversely, the feedbacks sometimes exhibit multiple microphysical, dynamical and radiative effects that counter one another (Kaufman and Koren, 2006; Jiang and Feingold, 2006; Koren et al., 2008a) and yield a relatively small overall effect. It has been suggested recently that aerosol-cloud interactions occur within a buffered system so that the response of the system to the aerosol is much smaller than might have been expected had internal interactions not been accounted for (Stevens and Feingold, 2009). The challenge is to
identify geographical regions or distinct regimes where the aerosol effect is likely to be largest.

**Measurement Uncertainties:** Clouds are extremely inhomogeneous and form complex, three-dimensional (3-D) structures, making them difficult to characterize from in situ measurements and causing remote sensing retrievals to be a true challenge (Platnick et al., 2003). The aerosol may be more homogeneous than clouds, but its measurable signal is weak relative to the background and instrument noise (Tanré et al., 1996, 1997; Kahn et al., 2005). The challenge grows even larger when measuring aerosol properties in the vicinity of clouds. When attempting to study cloud-aerosol interactions from observations, we often ask for the impossible: on the one hand we strive to measure the aerosol as close as possible to clouds in order to reflect the relevant aerosol properties that interact with the cloud. On the other hand, we ask for very accurate measurements of aerosol loading and properties, which is very difficult to achieve in the vicinity of clouds, especially by satellite remote sensing.

The inter-cloud region, sometimes referred to as the “cloud twilight zone”, is usually considered cloud-free, but has been shown to comprise sheared cloud fragments, tiny growing or decaying clouds, and hydrated aerosol. Thus the aerosol measured in this region is significantly different from the aerosol far-removed from cloud fields (Koren et al., 2007; Charlson et al., 2007). This difference introduces a series of complexities in interpreting remotely sensed aerosol properties near clouds. The likelihood of the remotely sensed aerosol properties being contaminated by a contribution from undetectable clouds (the abovementioned small clouds or clouds with weak optical signature) increases as a function of the distance from detectable clouds (Koren et al., 2008b, 2009). Moreover, aerosol particles may change their (true and apparent) optical properties near clouds as the humidity increases (Charlson et al., 2007; Twohy et al., 2009) and due to enhancement in the mean photon flux as a result of the clouds serving as secondary photon sources illuminating the cloud field from their edges (Marshak et al., 2006; Wen et al., 2007). All of the above effects: undetectable clouds, aerosol humidification and the 3-D cloud effects, yield an apparent larger Aerosol Optical Depth
Aerosol optical depth (AOD) measured from space, and interfere with properly characterizing the aerosol measured from ground-based remote sensing or even in situ measurements.

**Causality:** Even assuming that clouds and aerosol can be measured correctly, the last and the ultimate problem is the strong coupling of clouds and aerosols to meteorology (environmental properties). Are the observed aerosol-cloud relationships a result of the aerosol effect on clouds or does meteorology drive the changes in both aerosol and clouds properties? Meteorological conditions control most of the cloud properties. Variables such as temperature, humidity, surface fluxes and winds largely determine the depth of convection and the size of clouds. The major challenge is to find the cloud response to perturbations in the aerosol properties buried beneath the significant natural variability due to meteorology. Numerical models that simulate the same scenario and change only the aerosol properties are often used to establish causality but it is difficult to generalize the results for different regimes. It is not always clear if the observed relationships are applicable to a wide variety of meteorological regimes, requiring great effort to decouple basic meteorological properties (humidity and temperature) from aerosol effects (Teller and Levin, 2008; Altaratz et al., 2008).

An extreme test: Detecting the interaction between aerosol and convective clouds is an extreme test of all of the above challenges. Convective clouds exhibit the highest inhomogeneity and are extremely sensitive to changes in the environmental conditions (e.g. level of instability). Due to their high variability, in a given convective cloud field one can expect to find clouds at various stages of their lifecycle, which causes measuring/retrieving cloud properties difficult and introduces artifacts to aerosol measured/retrieved in the vicinity of detectable clouds. Lastly, convective clouds serve as a secondary photon source, illuminating the area between them thus making the 3-D cloud effects a significant consideration to both cloud and aerosol measurements and retrievals. Nevertheless, convective clouds are a major source for precipitation and their radiative effect can vary from cooling for low cumulus to warming in the deep convective cells and their related anvils. Therefore any change in their properties has the potential for a strong climate impact.
1.1 Aerosol effects on deep convective clouds

Deep convective clouds are manifestations of strong forcing in a potentially very unstable atmosphere. When changes in the aerosol properties affect microphysical processes they may modify the vertical distribution of condensate and energy, and thus modify cloud development.

Several observational and modeling studies support a hypothesis that suggests that higher aerosol loading leads to invigoration of convective clouds (Andreae et al., 2004; Khain et al., 2008). More aerosol particles provide more cloud condensation nuclei (CCN). Larger concentrations of CCN result in more, and smaller cloud droplets with smaller size variance (Twomey, 1977) that are less likely to collide and coalesce, and more likely to be lofted higher in the cloud because their fall velocities are smaller. Conversely, clean aerosol conditions promote the formation of larger droplets that can overcome updrafts, collecting enough droplets to form precipitation and its attendant surface cooling, which acts to stabilize the atmosphere. Therefore, polluted clouds are less likely to form early (warm) precipitation (Gunn and Phillips, 1957; Warner, 1968) and the updrafts are more likely to carry the smaller drops upwards and higher into the atmosphere, resulting in taller clouds that last longer (Koren et al., 2005; Fan et al., 2009). Once the cloud passes the freezing level (supercooled water), a chain of events is ignited in the mixed- and cold-phases. Heterogeneous ice nucleation is less efficient for smaller droplets. Therefore, heterogeneous freezing (and in colder environments, homogeneous freezing) is likely to take place higher in the atmosphere in a colder environment (Rosenfeld and Woodley, 2000). Less effective ice nucleation will further delay precipitation initiated by cold processes (Andreae et al., 2004). Thus it is thought that cold rain suppression may lead to longer cloud lifetime and larger cold cloud fraction (Koren et al., 2005; Lindsey and Fromm, 2008). Another positive feedback is expected to be initiated when the larger number of droplets (and larger net water mass) eventually freezes at colder temperatures releasing the latent heat at higher altitudes, enhancing buoyancy, and therefore, again promoting stronger updrafts that loft
the ice phase higher (Rosenfeld and Woodley, 2000; Jenkins and Pratt, 2008). Finally, the stronger evaporative cooling associated with aerosol-perturbed clouds results in stronger cold-pool development and more intense secondary convection (Seifert et al., 2006; van den Heever and Cotton, 2007; Lee et al., 2008). The net result of this hypothesis is that higher concentrations of aerosol generate larger and taller convective clouds. The fact that microphysical and dynamical processes in convective clouds are so complex, and that systems cycle through multiple stages of convection begs the question: is there an imprint of the aerosol on the macroscale features of clouds such as cloud top height and cloud fraction, and is it measureable with current tools?

Because of their global coverage and ability to observe literally millions of clouds in a variety of aerosol regimes, satellites provide a substantial source of data for investigation of the concept of invigoration. Analyses of these data provide evidence in support of the hypothesis of aerosol invigoration of convective clouds (Devasthale et al., 2005; Koren et al., 2005, 2008a; Meskhidze et al., 2009; Quaas et al., 2009). These studies identify correlative relationships between satellite-retrieved cloud and aerosol products that show cloud fraction and cloud top height systematically increasing as aerosol loading increases. The studies interpret these associations between variables as indicators of possible cause and effect, and conclude that aerosol has a detectable influence on the strength of the convection.

This conclusion is often challenged because of two of the most difficult problems associated with satellite-based studies of aerosol-cloud interaction: contamination of the retrieved aerosol product (Yang et al., 2008; Tian et al., 2008; Loeb and Manalo-Smith 2005; Zhang et al., 2005; Kaufman et al., 2005b), and the case for causality (Matsui et al., 2006; Loeb and Schuster, 2008; Mauger and Norris, 2007; Stevens and Brenguier, 2009). We define “contamination” as any influence on the retrieved AOD that would interfere with its ability to provide a stable proxy for CCN. “Proving causality” implies separating the cloud response due to aerosol from the response due to natural meteorological factors. The cloud type of interest will be deep convective clouds, and specifically the ITCZ region of the tropical Atlantic. While the difficulties we address
here have universal application to a wide variety of measurement-based studies of aerosol-cloud interaction, we focus on satellite-based studies and specifically the use of MODIS and MODIS-like imagers on polar-orbiting satellites that provide snapshot views of the Earth in multiple channels from the visible through the near-infrared and out to the thermal infrared.

The intent of this study is to attempt to discern a clear and significant net aerosol effect embedded within the large convective cloud variance that cannot be explained by contamination or meteorology. The study does not attempt to quantify the net effect of the aerosol, which depends strongly on the location, season, and type of clouds.

2 Methods

2.1 Analysis of aerosol-cloud interactions from satellite observations

There is a solid body of scientific literature addressing aerosol-cloud interaction using satellites as the primary tool (Coakley Jr. et al., 1987; Kaufman and Nakajima, 1993; Kaufman and Fraser, 1997; Wetzel and Stowe, 1999; Nakajima et al., 2001; Bréon et al., 2002; Feingold et al., 2001; Koren et al., 2004, 2005; Kaufman et al., 2005a; Berg et al., 2008; L’Ecuyer et al., 2009). Before the year 2000 these studies were based mostly on AVHRR reflectances, but beginning in 2000, after the launch of the NASA Terra satellite, aerosol-cloud studies mostly switched to using the Earth Observing System (EOS) satellite sensors including the MODerate resolution Imaging Spectroradiometer (MODIS) that flies on NASA’s Terra and Aqua satellites. Other EOS sensors have been used in these studies, as well, including the active sensors (lidar and radar) CALIOP and CloudSat. While the challenges we address here have universal application to a wide variety of measurement-based studies of aerosol-cloud interaction, the current investigation and analysis focuses on the use of MODIS and MODIS-like imagers such as AVHRR that view the Earth in multiple channels from the
visible through the near-infrared and out to the thermal infrared. These imagers are onboard polar orbiting satellites that provide instantaneous views of the aerosol cloud field at relatively coarse temporal resolution. MODIS views the Earth in 36 channels from 0.41 µm to 14 µm at a variety of spatial resolutions (250 m, 500 m and 1000 m). Among the hundreds of products derived from MODIS-measured radiances is a suite of aerosol products (Remer et al., 2005; Levy et al., 2007) and another set of cloud products (King et al., 2003; Platnick et al., 2003), including aerosol optical depth (AOD), cloud top pressure and cloud fraction.

Often the AOD is used as a proxy for the cloud condensation nucleus (CCN) concentration (Andreae, 2009). The reliability of this proxy depends on the uniformity of the aerosol size, composition, vertical distribution, but may in many cases be used as a first approximation. AOD is provided as a 10 km product, and the cloud products are provided at either 5 km or 1 km resolution.

To study the interactions between aerosol and clouds from satellite observations, information on both is needed, for the same time and location. However, aerosol cannot be retrieved beneath clouds under cloudy conditions by most satellite sensors, and cannot be retrieved above clouds by MODIS. To solve such a problem the characteristic spatial scale of the aerosol field is assumed to be larger than the scale of a cloud so that it is sufficient to measure aerosols in the vicinity of clouds. Often the retrieved resolution is decreased so the grid square size covers a much larger area compared with the retrieval resolution, increasing the likelihood of information on clouds from the cloudy part and on aerosol from the cloud-free part within the larger grid square.

Grid squares with larger coverage are also useful for variance reduction. Due to the high instability and the high thermodynamic energy involved, convective clouds exhibit large variance in their physical properties. This adds great difficulty to the detection and estimation of aerosol impacts on such clouds. To reduce the complexity, datasets that average millions of clouds are needed. To simplify the data handling and analysis and to reduce the variance, pixels with large spatial area (1 degree) are used (King et al., 2003). Although lower resolution increases the likelihood of having both significant
cloud and aerosol information, it may suffer from unavoidable mixing between cloud types, and increase the possibility of contamination.

2.2 Cloud contamination

The MODIS aerosol algorithm was designed to minimize cloud contamination by using a sensitive cloud detection algorithm (Martins et al., 2002; Gao et al., 2002). Furthermore, the aerosol algorithm removes inadvertent cloudy pixels and cloud shadows that escape the cloud mask by eliminating the darkest and brightest 25% of every cluster of 400 pixels after all detectable clouds have been removed (Remer et al., 2005). Nevertheless it has been shown that the MODIS AOD retrievals may contain residual artifacts introduced by clouds (Kaufman et al., 2005b; Zhang et al., 2005), by 3-D effects near clouds (Marshak et al., 2006; Wen et al., 2007) and by physical changes to the “aerosol” in the vicinity of clouds (Tackett and Di Girolamo, 2009).

The consequences of potential contamination in the MODIS aerosol product are investigated in two independent ways. First by rejecting aerosol measurements with a higher likelihood to be in the vicinity of clouds, and second by using aerosol information taken from the GOCART transport model (Chin et al., 2000a,b) instead of MODIS data.

2.2.1 Filtering AOD pixels with high cloud content

To estimate possible effects of cloud contamination on the aerosol retrieval we use collocated cloud and aerosol retrievals over the ITCZ (tropical Atlantic) for the months of July and August 2007 (0–10° N and 20–30° W). Two aerosol data bases are created. The first is parallel to the operational MODIS product and is constructed from all MODIS Level_2, 10-km AOD retrievals, no matter how cloudy the original 10 km product. The second data base starts from the same initial 10 km aerosol retrievals, but includes only those retrievals that report less than 20% cloud fraction within the 10 km box. Previous studies have suggested that the likelihood for cloud contamination or other biases
in the retrieved AOD is proportional to the distance from detectable clouds (Koren et al., 2007). We note that the 20% cloud fraction cut-off corresponds to an average distance of 5 km between a pixel used by the AOD retrieval and an identified cloudy pixel. There is no guarantee that this procedure eliminates all possible cloud contamination by imposing this arbitrary cloudiness cutoff, however most of the pixels likely to be cloud-contaminated are filtered out (Fig. 1). Other thresholds that range from 80% to 20% were tested and all showed similar results. Here the results pertaining to the most stringent (20%) filter are shown. The 20% cutoff is therefore a conservative measure that should eliminate most, if not all, artificial relationships between retrieved AOD and retrieved cloud parameters. However, such filtering while reducing cloud effects on the aerosol retrieval can result in significant loss of useful data. Because the tropical sky is very cloudy such filtering results in a rejection of more than 40% of the 10-km aerosol retrievals and often a whole aerosol outbreak that is not necessarily cloud contaminated is rejected (Fig. 1).

The MODIS aerosol algorithm retrieves directly two aerosol parameters namely the AOD and the fine-fraction (fraction of the optical thickness at 550 nm contributed by fine particles). Often, the fine-fraction information allows classification of aerosol types (Kaufman et al., 2005c). Cloud contamination in the aerosol retrievals is expected to decrease the fine-fraction, by introducing larger particles in the form of cloud droplets and decreasing the Angstrom exponent. Figure 2 shows a scatter plot of the mean fine-fraction from each 1-degree×1-degree grid of the aerosol products subjected to the 20% cloud fraction cut-off plotted against the standard retrieval. The regression slope is 1.043 with an intersect of 0.0075, and a correlation of $R=0.91$. The average AOD and fine-fraction are 0.246 and 0.446, respectively, for the standard retrieval, and 0.227 and 0.473 for the stringently less cloudy data set. The less cloudy data set does have smaller average AOD and larger fine-fraction than the standard product, consistent with cloud contamination in the standard product. However, even the aerosol retrievals with the 20% cloud fraction threshold exhibit a physical, negative correlation between AOD and fine-fraction and therefore, part of the ~5% reduction in fine-fraction
can be attributed to true rejection of cloud contaminated pixels but part is likely due to rejection of non-contaminated pixels with higher AOD. We note the strong correlation between the data sets, the nearly 1:1 slope and small offset, despite the less cloudy data set losing 40% of the original retrievals, suggesting that the properties of the aerosol retrieval are not strongly affected by the close proximity of clouds. The real question is not the relationship between the aerosol retrievals in the two data sets, but how the aerosol-cloud correlations differ when using either the standard or filtered data sets.

2.2.2 Selection of deep convective clouds

High resolution analysis of the cloud top pressure (CTP) height distribution over the study area (Fig. 3) suggests the existence of 3 cloud classes: Low boundary layer clouds with tops less than 850 hPa; clouds of medium height – free tropospheric clouds with cloud-top height between 850 and 600 hPa; and high clouds with heights above 600 hPa. The data are analyzed first for all clouds and then only for high clouds.

The data are divided into four subsets:

1. “All data”: data is gathered into 1-degree grid squares (similar to the operational MODIS level 3 product).

2. “All data filtered”: only those retrievals that were retained after the imposition of the 20% cloud fraction cut-off (at the 10 km pixel resolution) are aggregated into 1-degree grid squares.

3. “High only mode” – only those retrievals of clouds with CTP <600 hPa (high clouds) were averaged into the 1-degree resolution grid squares.

4. “High, filtered mode”: only those retrievals of clouds with CTP <600 mb and cloud fraction <20% were gathered into the 1-degree resolution grid.

Each of the above datasets is further divided into 3 AOD levels of 0<AOD<0.15, 0.15<AOD<0.35 and 0.35<AOD<1. The levels were selected to represent clouds in...
relatively clean, moderately polluted and heavy aerosol conditions, with bins defined such that there are comparable numbers of samples in each one.

3 Results

3.1 Correlations between MODIS-derived cloud top pressure, cloud fraction and aerosol optical depth

The histograms as well as the mean of the CTP and cloud fraction (CFR) are calculated for each AOD subset and presented in Fig. 4 for all clouds, and Fig. 5 for the high-only subset. Note that the CFR presented here is a standard MODIS cloud product, produced independently from the aerosol algorithm using a separate set of MODIS channels (Platnick et al., 2003). This means that the filtering of potentially cloud-contaminated aerosol products using a diagnostic of the aerosol algorithm (Reimer et al., 2005), does not affect the cloud fraction calculation.

The all-data have higher cloud tops (lower CTP) and higher CFR than the data set subjected to the 20% cloud fraction filter. However, the relationships between cloud parameters and AOD is robust. Higher aerosol loading tends to push the distributions and means toward higher clouds and larger cloud fractions, regardless of whether the AOD is calculated from all-data, or the stringently less cloudy data sets.

When focusing on the high-cloud data sets, the data are noisier due to the significant reduction in the sample size. Nevertheless, there is a clear, increase in cloud top height associated with increased aerosol loadings, regardless of whether the data set includes all-data or only the data after the 20% cloud fraction filter is applied. There are subtle differences between the data sets; the less-cloudy filtered data show a stronger relationship between CTP for high clouds and increasing AOD than do the all-data, but a weaker relationship between cloud fraction and AOD.
3.2 Correlations between GOCART AOD and MODIS cloud properties

The satellite aerosol loading measured over the same area of the clouds is now replaced with numerical simulations of aerosol loading calculated by the Goddard Chemistry Aerosol Radiation Transport (GOCART) model (Chin et al., 2000a,b). Similar attempts to investigate aerosol effects on clouds using a chemical transport model to represent the aerosol and satellite measurements of clouds have been performed by Chameides et al. (2002) and Schwartz et al. (2002). The GOCART model simulates emission, chemistry, turbulent mixing, advection, moist convection, dry and wet deposition for major tropospheric aerosol types, including sulfate, dust, organic carbon, black carbon, and sea salt aerosols (Chin et al., 2002). The model has a horizontal resolution of 2° latitude by 2.5° longitude and 20–30 vertical sigma layers, and uses the assimilated meteorological fields generated from the Goddard Earth Observing System Data Assimilation System (GEOS DAS, Schubert et al., 1993).

The use of simulated instead of measured AOD depends heavily on the accuracy of the assimilated dynamical data and the aerosol processes. The model and measurements have shown good agreement on seasonal and higher variations on daily scales (Yu et al., 2004). However the big advantage of model output, in our case, is that it is completely free of any cloud contamination.

In order to increase the sampling statistics both GOCART and MODIS data were included and the study area was increased from the Atlantic ITCZ only, to a box encompassing (0° N to 14° N; 18° W to 45° W) for August 2007. The GOCART AOD model output was interpolated from 2° by 2.5° to 1° spatial resolution to match the MODIS level-3 data, which corresponds to the “all-data” data set of the previous analysis.

While significant differences in the magnitudes of the parameters are expected, the trends between cloud properties and AOD are similar regardless of whether the AOD is retrieved from MODIS measurements (Figs. 4 and 5) or GOCART simulations (Fig. 6). The above analyses give strong indication that cloud contamination of the aerosol product is not artificially producing correlations between aerosol and cloud parameters in
the MODIS data sets, for the region and season of interest. The observed relationships of increasing cloud top height and cloud fraction with increasing AOD holds whether or not all data or only data selected for low cloud fraction are analyzed, and this is equally true for all clouds in the region or applied to only the highest clouds. Moreover the trends are similar when using modeled AOD suggesting a clear increase in cloud cover and cloud vertical development with increasing AOD. The above analysis is not a definitive result for all regions and seasons, nor is it a statement about cloud contamination in the MODIS AOD product. The analysis only shows that using the MODIS AOD product to find associations between aerosols and clouds in this particular situation will not lead to false conclusions due to cloud contamination in the AOD retrieval.

3.3 Causality

When performing analysis from observations one can detect correlations but an understanding of the causal relationships is much more elusive. Often numerical cloud models are used to isolate aerosol effects, while controlling all other variables, and to resolve the physical explanation for the observed relationships. Another way to try to separate variables that may control cloud properties is to narrow the variability in individual components of the system. This process is analogous to running a model for the same conditions of one variable set (say meteorology) and comparing the results to those obtained when changing the other one (aerosol loading). Here we have used the NOAA-NCEP Global Data Assimilation System (GDAS) at 1° resolution to define the meteorology of the region of interest. The GDAS data set is assimilated 4 times a day, for 13 pressure levels from the surface to 20 hPa using surface and satellite observations with model data (Kanamitsu, 1989; Parrish and Derber, 1992). The GDAS is the final run in the series of NCEP operational model runs. We have used the 12 UTC set which, on average, covers a period of time less than 2 h from the satellite observation time.
Reanalysis data have previously been used to define the meteorology in aerosol-cloud studies from satellite and to estimate the meteorological contribution to cloud properties (e.g., Kaufman et al., 2005a; Brioude et al., 2009; Savtchenko, 2009). Here we use this approach to find the variables that best describe the ITCZ convective cloud properties (vertical development and cloud fraction) as measured by MODIS. After finding the meteorological variables that have the most statistically significant correlations with the measured cloud properties, the meteorological variance will be restricted by limiting the allowed range of the selected variables. Since convective systems are expected to have a well-defined set of meteorological properties (Jakob and Tselioudis, 2003), such a direct approach is preferred to a multi-regression analysis (Kaufman et al., 2005a) where many meteorological variables are mixed together to completely span the meteorological variance space. The multi-regression approach may yield better overall correlations with the free (convective cloud) variables but when mixing many different variables with different units one may lose the physical reasoning and the likelihood of statistically insignificant correlations appearing is higher.

Since we seek meteorological variables that represent cloud properties at several scales, ranging from the highest resolution that is limited by the assimilation resolution (1°) to a synoptic scale, the same study area as for the GOCART analysis is used (Atlantic ITCZ; 0° N to 14° N; 18° W to 45° W; August 2007), along with the MODIS level-3 data which is also given at 1° spatial resolution. By using this coarser resolution data we lose the ability to separate the three cloud modes seen in the Fig. 3 histogram that was constructed from much finer resolution retrievals (5 km). Thus, the specific focus on the highest cloud top category previously done in the analysis concerning cloud contamination cannot be repeated in this study of causality. Still the Atlantic ITCZ region offers a situation with deep convection comprising a prevalent component of the “all cloud” situation. Identifying robust relationships between aerosol and cloud properties here, will provide indirect support that aerosol is affecting the deep convective elements in the system.
The GDAS output has 286 meteorological variables. The key ones such as geopotential heights, temperature, winds (horizontal and vertical), relative humidity, vorticity and water mixing ratio are all given at 13 pressure levels. Other variables are specified at specific pressure levels (such as variables that pertain to the boundary layer) or as an average of a few levels (the complete list of GDAS variables is given in Appendix A).

First we perform a correlation test over our study area between MODIS-derived CTP, CFR and AOD against all of the GDAS output variables. In order to check sensitivity to spatial and temporal scales this is done for the finest assimilation resolution (1°) and also for coarser resolutions of (2°, 4° and 8°), and for data measured at the same time as the GDAS output and at time lags of 1 day and 5 days. The significance of each correlation test was tested using the p-value test (Schervish, 1996). The p-value test is the probability of getting a correlation as large as the observed values by random chance, when the true correlation is zero (the null hypothesis). Lower p-values correspond to more significant results. Often, p-value thresholds of <0.05 are used to define correlations as truly significant (Wilks, 2005). In our study a p-value threshold of 0.01 is used to be even more conservative.

Figure 7 shows the correlations between MODIS CTP and all of the GDAS output variables at 4 different spatial resolutions (1°, 2°, 4°, 8°). The plot area is divided by background color to signify the key variable regimes. The left part of each regime refers to the lower part of the atmosphere, gradually increasing in height (decreasing in pressure). Note how two regimes present remarkably high correlations, namely the pressure vertical velocity (ω) variables (53–73) and relative humidity (RH) variables (74–94). Two of the other variables that show high correlations are related (var. 228 precipitable water and var. 245–248 relative humidity at different pressure level ranges). Decreasing the spatial resolution from 1° to 4° to 8° sometimes has an effect on the correlations, but not enough to change the basic finding: MODIS-observed cloud top pressure is most robustly linked to broad scale vertical velocity and humidity variables, especially towards the upper troposphere.
Similarly to Fig. 7, Fig. 8 shows the correlations between MODIS-derived CTP, CFR and AOD and all of the GDAS output variables at 1° spatial resolution. Indeed AOD does correlate well with some of the meteorological parameters but the parameters are different than those that correlate with the clouds. Moreover the correlation between $\omega$ and RH are of opposite sign, suggesting that the same meteorology that favors higher aerosol loading (say midlevel continental transport) will oppose development of deep convective clouds. For example, the analysis suggests that the correlation between AOD and RH is negative in this region because aerosol transport to the tropical Atlantic is often accompanied by dry continental air masses. This appears to dominate the otherwise expected positive correlation between AOD and RH due to water vapor uptake by the aerosol.

Because $\omega$ (550 hPa) and RH (350 hPa) yielded the highest correlations with the cloud properties, these parameters will be used to constrain the meteorological conditions. The CTP and CFR data were divided into 3 $\omega$ ranges: $\omega<-0.1$ Pa/s for the strong convective regime, $-0.1<\omega<0$ Pa/s for the weak convective regime and $0<\omega$ for the subsidence regime. Similarly the data were divided into 3 RH bins: $0<\text{RH}<25\%$, $25\%<\text{RH}<55\%$ and $55\%<\text{RH}$. The correlations with AOD are presented for CTP and CFR restricted by $\omega$ and RH in Fig. 9. Note how indeed stronger uplifting and RH correlates with taller clouds that have larger cloud fraction suggesting that these variables do reflect correctly their expected effect on the convective cloud properties. Note also how the general relationships between increasing AOD and cloud properties are not dramatically different between the different meteorological subsets and the original full data set (dotted blue lines). These curves suggest that the aerosol and meteorology exert a near-orthogonal forcing on the convective cloud properties with the y-intercept indicating the meteorological effect that is represented by these GDAS variables and the slopes representing the aerosol effect.
4 Discussion and conclusions

In this paper two of the most pernicious problems that emerge when studying aerosol-cloud interactions from satellite observations have been tackled. Whenever correlations between aerosol and cloud properties are observed in satellite data the first question that arises is: Are these correlations physically-based, or are they an artifact of the retrievals? The second question is: Are these correlations due to aerosol effects or is the aerosol acting as a tracer for specific meteorological conditions that are the true modifier of the cloud properties? The objective has been to present two methodologies that can help confront these important questions. The results strengthen the case that the observed increase in cloud top height and cloud fraction associated with higher aerosol loadings in convective clouds over the Atlantic ITCZ in the data set studied is due to aerosol processes and not to retrieval artifacts or meteorology. The observed response to aerosol perturbations is referred to as “invigoration” and comprises complementary parts: a suppression of the stabilizing effects of surface rain generated by warm processes; a delay in the freezing process that results in latent heat release at higher levels in the cloud (destabilization); and stronger secondary convection associated with more vigorous cold pool outflow.

The first problem addressed is that of the potential bias in aerosol measurements as a result of cloud contamination. Analyses have been performed with the standard MODIS retrieval products as well as eliminating all 10 km aerosol retrievals with internal CFR>20%, representing an inordinate effort to remove contaminated aerosol retrievals. When removing these retrievals more than 40% of the AOD retrievals are removed and information on complete aerosol regimes that are not necessarily affected by clouds is lost (see example in Fig. 1).

For the “all-data set”, i.e., the standard MODIS products, the cloud top pressure for the highest AOD category (AOD=0.5) is 110±30 hPa lower (~800 m higher) than for the lowest AOD category (AOD=0.1). After removing the aerosol data set with CFR>20%, height differences still exist, but are reduced to 60+/−30 hPa (~480 m). Both all-data
and the minimally-cloudy filtered data sets show similar magnitudes of increasing CFR with increasing AOD; CFR is $0.35 \pm 0.05$ higher for the highest AOD category as compared with the lowest AOD category for the all-data set and $0.30 \pm 0.08$ for the <20% cloud fraction filtered subset. In the high-only cloud-top case the results suggest enhancement of the high AOD/low AOD differences for the filtered subset, as compared with the all-data case. Here the average cloud tops rise $55 \pm 15$ hPa higher ($\sim 1000$ m) when the AOD changes from 0.10 to 0.24 in the minimally-cloudy filtered subset and only $25 \pm 10$ hPa ($\sim 500$ m) in the all-data set. The saturation of the microphysical effect after AOD $> \sim 0.25$ is also visible in the high-only data sets. The high-only CFR relationships with AOD are reduced from $0.35 \pm 0.05$ to $0.23 \pm 0.08$, from the “filtered” to “all-data” analyses. Indeed the data for all clouds and high-only clouds show clear and significant positive correlations between cloud top height and CFR and aerosol loading, whether all-data or minimally-cloudy filtered aerosol data sets are used.

Use of the GOCART-modeled AOD output has the advantage of removing any concerns of cloud contamination in the reported AOD. Nevertheless correlations of CFR and CTP with AOD show similar trends with larger magnitude. When the AOD changes from 0.05 to 0.28 the cloud tops are higher by $330 \pm 30$ hPa and the cloud fraction is significantly larger ($0.4 \pm 0.1$). Again this part of the analysis suggests that the correlations between MODIS-derived aerosol and cloud products in this region and season of interest are not due to artifacts of the MODIS aerosol retrievals.

The second problem addressed is that of causality. To this end GDAS assimilation data are used to provide the context for the observed aerosol-cloud correlations. The GDAS assimilation data are used as an approximation for the meteorological conditions at the same time and location of the satellite observations. The assimilation data include surface, radiosondes and space measurements in conjunction with an atmospheric general circulation model to represent a wide variety of meteorological data on a three-dimensional grid. The GDAS data may contain some information on aerosol radiative effects (if these affected the measurements used by the assimilation) but it should be minor and will not contain direct information on the aerosol microphysical
effects that are the focus of this paper. Therefore it is argued to be a suitable tool to represent the dynamical component.

The differences in the correlation patterns between the cloud properties and AOD (Fig. 8) already suggest that, at least over this area, the positive correlations between AOD and deep convective vertical and horizontal development occur in spite of the meteorology. The aerosol does correlate with some of the assimilation variables but mostly with ones different than those that correlate with convective clouds, and in many cases if there is a joint correlation it is with opposite sign, suggesting that AOD will be high in conditions that are not favorable for deep convective cloud development.

The potential for new insights into understanding deep convection based on correlations between the assimilation variables and satellite-derived cloud parameters is enormous. Each significant correlation should be explained and its scale-dependence understood. Here we have used only a minor portion of this information to restrict the variance of convective cloud meteorology. The variables that give the highest and the most consistent correlations with CTP and CFR are the large-scale pressure vertical velocities (\(\omega\)) and the relative humidity (RH) of the upper troposphere. Intuitively these variables should have good agreement with convective clouds and so the fact that their correlations are the highest serves as a sanity check.

Restriction of the data to \(\omega\) and RH regimes showed that, as expected, stronger broad-scale rising motion (low level convergence) and RH correlate with taller clouds and higher cloud fraction. Relationships between the satellite-derived cloud properties and AOD have been shown here to be almost invariant to the meteorological binning. For any of the meteorological scenarios spanned by the \(\omega\) and RH ranges, the ITCZ convective clouds are taller and larger when the aerosol loading increases. Moreover, the slopes (as a function of the AOD) are similar for all subsets (CTP moreso than CFR) suggesting that the aerosol and meteorology exert a near-orthogonal forcing on convective invigoration. In such variable space, the aerosol effect defines a family of lines with a given slope and their intercept depends on the meteorology.
The analysis does include a number of imperfections. First, the AOD, the basic metric used for aerosol loading cannot always be monotonically linked to aerosol concentration, and suffers from ambiguity associated with water vapor uptake. In spite of the positive correlation between AOD and surface CCN measurements, errors can be substantial (Kapustin et al., 2006; Andreae, 2009). The quality of the AOD as a measure of CCN concentration depends on the aerosol type, uniformity and vertical distribution of the aerosol, and the merits of this approach may vary greatly. Second, it is not possible to completely separate meteorological influences from aerosol influences on clouds. This work can therefore only provide further evidence of the aerosol playing a significant role in increasing cloud top height and cloud fraction, but cannot quantify the relative contributions with confidence.

Often the suggested approach to the complex phenomena of convective cloud-aerosol interaction is to use numerical models to answer all questions. The problem is, of course, that in reality it is very hard (likely impossible) to build a model with sufficient complexity to properly simulate all physical pathways, at all scales. Models need constraints from observations. Some constraints can and should be imposed based on laboratory or field observations, but others that involve regional or seasonal scales can and should be imposed from robust satellite observations. We have shown here that aerosol-cloud relationships derived from satellite data can be held to strict scrutiny and still deliver robust, physically significant results. These robust relationships can be coupled with scientific insight to provide hypotheses from which models can start to explore the detailed physics behind the observations, as well as to constrain the model results. Thus, to answer the difficult questions concerning aerosol-cloud interaction, and how particles can alter climate and the hydrological cycle through their modification of clouds, we need a comprehensive approach that includes modeling, in situ measurements and satellite-based observational studies.
Appendix A

GDAS model parameters associated with the abscissae of Figs. 7 and 8

The model includes 26 pressure levels (P)

(i=1, 26): P(1000, 975, 950, 925, 900, 850, 800, 750, 700, 650, 600, 550, 500, 450, 400, 350, 300, 250, 200, 150, 100, 70, 50, 30, 20, 10) (hPa)

(i=27, 52): Temperature (at P levels i=1,26)

(i=53, 73): Pressure vertical velocity (at P levels i=1,20)

(i=74, 94): Relative humidity (at P levels i=1, 20)

(i=95, 120): Absolute vorticity (at P levels i=1, 26)

(i=121, 126): Ozone mixing ratio (at P levels i=21, 26)

(i=127, 147): Cloud water mixing ratio (at P levels i=1, 20)

(i=148): 5-wave geopotential height at i=13 (P=500 mb)

(i=149, 200): (U; V) horizontal velocity pairs (at P levels i=1, 26)

(i=201): Temperature at 30 hPa above ground

(i=202): Relative humidity at 30 hPa above ground

(i=203): Specific humidity at 30 hPa above ground

(i=204): U-wind at 30 hPa above ground

(i=205): V-wind at 30 hPa above ground

(i=206): Temperature at 1829 m above MSL

(i=207): Temperature at 2743 m above MSL
(i=208): Temperature at 3658 m above MSL
(i=209): U-wind at 1829 m above MSL
(i=210): V-wind at 1829 m above MSL
(i=211): U-wind at 2743 m above MSL
(i=212): V-wind at 2743 m above MSL
(i=213): U-wind at 3658 m above MSL
(i=214): V-wind at 3658 m above MSL
(i=215, 216): Geopotential height, unknown height
(i=217, 218): Temperature, unknown
(i=219, 220): Pressure, unknown
(i=221, 222): Vertical speed shear, unknown
(i=223, 226): (U; V) wind pairs, unknown
(i=227): Surface pressure
(i=228): Precipitable water
(i=229): Relative humidity, unknown
(i=230): Geopotential height, tropopause
(i=231): Temperature, tropopause
(i=232): Pressure, tropopause
(i=233): Vertical speed shear, tropopause
(i=234): Surface Lifted Index, surface
(i=235): Convective Available Potential Energy (CAPE), surface
\( i=236 \): Convective Inhibition (CIN), surface
\( i=237 \): Best (4-layer) lifted index, surface
\( i=238 \): Convective Available Potential Energy (CAPE), 180 hPa above ground
\( i=239 \): Convective Inhibition (CIN), 180 hPa above ground
\( i=240 \): Geopotential height at max wind level
\( i=241 \): Temperature at max wind level
\( i=242 \): Pressure at max wind level
\( i=243 \): Geopotential height at surface
\( i=244 \): Pressure reduced to MSL
\( i=245 \): Relative humidity at sigma 0.44–1.00
\( i=246 \): Relative humidity at sigma 0.72–0.94
\( i=247 \): Relative humidity at sigma 0.44–0.72
\( i=248 \): Relative humidity at sigma 0.33–1.00
\( i=249 \): Potential temperature at sigma=0.995
\( i=250 \): Temperature at sigma=0.995
\( i=251 \): Pressure vertical velocity at sigma=0.995
\( i=252 \): Relative humidity at sigma=0.995
\( i=253 \): Total ozone
\( i=254 \): Cloud water mixing ratio (unknown)
\( i=255 \): Geopotential height at 0°C isotherm
(i=256): Relative humidity at 0°C isotherm
(i=257): Geopotential height (unknown)
(i=258): Relative humidity (unknown)
(i=259): U-wind (tropopause)
(i=260): V-wind (tropopause)
(i=261): U-wind maximum
(i=262): V-wind maximum
(i=263): U-wind at sigma=0.995
(i=264): V-wind at sigma=0.995
(i=265): Temperature at surface
(i=266): Volumetric soil moisture (0–10 cm down)
(i=267): Volumetric soil moisture (10–40 cm down)
(i=268): Volumetric soil moisture (40–100 cm down)
(i=269): Volumetric soil moisture (100–200 cm down)
(i=270): Temperature (0–10 cm down)
(i=271): Temperature (10–40 cm down)
(i=272): Temperature (40–100 cm down)
(i=273): Temperature (100–200 cm down)
(i=274): Accumulated snow at surface
(i=275): Land cover fraction (land=1; sea=0)
(i=276): Surface ice fraction (ice=1; no ice=0)
The invigoration of deep convective clouds over the Atlantic
I. Koren et al.

Acknowledgements. IK acknowledges a visiting fellowship from the Cooperative Institute for research in the Environmental Sciences (CIRES) for supporting this work. GF and IK acknowledge support from NOAA’s Climate Goal. LR and IK acknowledge support from NASA’s Interdisciplinary Science program.

References

(i=277): Temperature at 2 m above ground
(i=278): Specific humidity at 2 m above ground
(i=279): Planetary boundary layer height
(i=280): Total cloud cover
(i=281): Relative humidity at 2 m above ground
(i=282): U-wind at 10 m above ground
(i=283): V-wind at 10 m above ground
(i=284): Geopotential height anomaly at 1000 hPa
(i=285): Geopotential height anomaly at 500 hPa
(i=286): 5-wave Geopotential height anomaly at 500 hPa


Fig. 1. Example of the consequences of removing MODIS aerosol retrievals collocated with more than 20% cloud fraction. Left – part of the 12 August 2007 MODIS visible image over the Atlantic ITCZ. Center: the MODIS operational AOD retrieval for the scene. Right – the same retrieval after eliminating all AOD retrievals collocated with higher than 20% cloud fraction. Note how AOD retrievals in the vicinity of any cloud are removed leaving only AOD in cloud-free islands.
Fig. 2. Correlations between aerosol fine fraction ratio: the standard product (x-axis) and after applying a stringent 20% cloud fraction threshold to retain the least cloudy scenes (y-axis).
Fig. 3. Distribution of cloud top heights over the study area constructed from 5-km retrievals.
Fig. 4. Cloud properties vs. AOD for 3 AOD ranges: blue 0<AOD<0.15, red 0.15<AOD<0.35 and black 0.35<AOD<1. Normalized by area histograms of (a) cloud top pressure for “all data”, (b) cloud fraction for “all data”, (c) as in (a) but for the data subjected to the 20% cloud fraction filter (“all data filtered”) and (d) as in (b) but for the data subjected to the 20% cloud fraction filter only (“all data filtered”). (e) the average cloud top pressure vs. the average AOD for all-data (cyan) and “all data filtered” (magenta). (f) average cloud fraction vs. the average AOD for all-data (cyan) and “all data filtered” (magenta). The vertical extent of the lines in (e) and (f) represents ± one standard error.
Fig. 5. As in Fig. 4 but for high clouds.
Fig. 6. MODIS cloud data vs. GOCART AOD analysis for 3 AOD ranges: blue 0<AOD<0.1, red 0.1<AOD<0.2 and black 0.2<AOD<1. Normalized-by-area histograms of (a) cloud top pressure, (b) cloud fraction, (c) the average MODIS cloud top pressure vs. the average GOCART AOD) (d) average MODIS cloud fraction vs. the average GOCART AOD. The vertical extent of the lines in (c) and (d) represents +/- one standard error.
Fig. 7. Correlations between MODIS cloud top pressure and the GDAS assimilation outputs for 4 model spatial resolutions: blue 1°, red 2°, black 4° and magenta 8°. The zero correlation data points are the variables that did not pass the p-value<0.01 test.
Fig. 8. Correlations between MODIS cloud top pressure (upper), cloud fraction (middle) and aerosol optical depth (lower) and the GDAS assimilation outputs for 1° spatial resolution. The zero correlation data points are those variables that did not pass the p-value<0.01 test.
Fig. 9. MODIS cloud properties plotted as a function of AOD restricted by $\omega$ (550 hPa) and RH (350 hPa). The $\omega$ subsets are all in [Pa/s] and divided as following: blue – all data, red $\omega<-0.1$, black $-0.1<\omega<0$ and magenta $0<\omega$. The RH subsets are: blue – all data, red $0<\text{RH}<25\%$, black $25<\text{RH}<55\%$ and magenta $55<\text{RH}$. (a) CTP vs. AOD for $\omega$ subsets, (b) CFR vs. AOD for $\omega$ subsets, (c) CTP vs. AOD for RH subsets and (d) CFR vs. AOD for RH subsets. Note how close to parallel the relationships are (more so for the CTP than for the CFR) suggesting that the meteorology can be separated from aerosol effects. The offset of the lines (along the y-axis) represents the meteorology effect while the slope is attributed to aerosols.