ABSTRACT

Title of Dissertation:	MULTI-SENSOR CLOUD AND AEROSOL RETRIEVAL SIMULATOR AND ITS APPLICATIONS
	Galina Wind, Doctor of Philosophy, 2016
Dissertation directed by:	Professor Ross J. Salawitch, Department of Atmospheric and Oceanic Science Dr. Steven Platnick, NASA Goddard Space Flight Center Code 610 Dr. Arlindo M. da Silva, NASA Goddard Space Flight Center, Global Modeling and Assimilation Office

Executing a cloud or aerosol physical properties retrieval algorithm from controlled synthetic data is an important step in retrieval algorithm development. Synthetic data can help answer questions about the sensitivity and performance of the algorithm or aid in determining how an existing retrieval algorithm may perform with a planned sensor. Synthetic data can also help in solving issues that may have surfaced in the retrieval results. Synthetic data become very important when other validation methods, such as field campaigns,are of limited scope. These tend to be of relatively short duration and often are costly. Ground stations have limited spatial coverage whilesynthetic data can cover large spatial and temporal scales and a wide variety of conditions at a low cost.

In this work I develop an advanced cloud and aerosol retrieval simulator for the MODIS instrument, also known as Multi-sensor Cloud and Aerosol Retrieval Simulator (MCARS). In a close collaboration with the modeling community I have seamlessly combined the GEOS-5 global climate model with the DISORT radiative transfer code, widely used by the remote sensing community, with the observations from the MODIS instrument to create the simulator.

With the MCARS simulator it was then possible to solve the long standing issue with the MODIS aerosol optical depth retrievals that had a low bias for smoke aerosols. MODIS aerosol retrieval did not account for effects of humidity on smoke aerosols. The MCARS simulator also revealed an issue that has not been recognized previously, namely the value of fine mode fraction could create a linear dependence between retrieved aerosol optical depth and land surface reflectance. MCARS provided the ability to examine aerosol retrievals against "ground truth" for hundreds of thousands of simultaneous samples for an area covered by only three AERONET ground stations.

Findings from MCARS are already being used to improve the performance of operational MODIS aerosol properties retrieval algorithms. The modeling community will use the MCARS data to create new parameterizations for aerosol properties as a function of properties of the atmospheric column and gain the ability to correct any assimilated retrieval data that may display similar dependencies in comparisons with ground measurements.

MODIS CLOUD AND AEROSOL RETRIEVAL SIMULATOR AND ITS APPLICATIONS

by

Galina Wind

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2016

Advisory Committee:
Professor Ross J. Salawitch, Chair
Dr. Steven Platnick
Dr. Arlindo da Silva
Professor Rachel T. Pinker
Research Assistant Professor Timothy P. Canty
Dean's Representative: Professor Ronald A. Yaros

Acknowledgements

I would like to thank everyone who collaborated and supported me on this project:

Dr. Steve Platnick for funding this effort and helping out with all sorts of great advice over the years.

Dr. Bryan Baum at the University of Wisconsin – Madison for sharing the original DISORT-5 simulator code with our research group.

Dr. Arlindo da Silva and Dr. Peter Norris for making the integration with GEOS-5 possible.

Brad Wind for laying down the groundwork for simplifying the modifications to the Collection 5 operational MODIS cloud optical and microphysical retrieval code that made creation of CHIMAERA retrieval system possible and for floating the original idea for what eventually became MCARS way back in 2004.

Shana Mattoo for running countless MODIS aerosol retrievals with MCARS output.

Leigh Munchak for providing MODIS – AERONET comparisons.

Dr. Peter Colarco for providing the OPAC information.

Mara Jade Skywalker, without whom I would've never had the guts to even think of getting here.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AERONET	AErosol RObotic NETwork
AGCM	Atmospheric General Circulation Model
AOD	Aerosol Optical Depth
BRDF	Bi-directional Reflectance Distribution Function
CHIMAERA	Cross-platform HIgh resolution Multi-instrument AtmosphEric Retrieval Algorithms
CWP	Condensed Water Path
DISORT	DIScrete Ordinate Radiative Transfer
DMS	DiMethyl Sulfide
ECMWF	European Center for Medium-range Weather Forecasts
EOS	Earth Observing System
ESMF	Earth System Modeling Framework
FRP	Fire Radiative Power
GAAS	Goddard Aerosol Assimilation System
GOES	Geostationary Operational Environmental Satellite
GEOS-5	Goddard Earth Observing System model version 5
GMAO	Global Modeling and Assimilation Office
GOCART	Goddard Chemistry, Aerosol, Radiation and Transport
GSFC	Goddard Space Flight Center
GSI	Grid-point Statistical Interpolation
HDF	Hierarchical Data Format
HG	Henyey-Greenstein phase function model
ICA	Independent Column Approximation
IGBP	Inernational Geosphere-Biosphere Programme
IPCC	Intergovernmental Panel on Climate Change
L1B	MODIS calibrated radiance and reflectance product designation
LUT	Look-Up Table
MISR	Multi-angle Imaging SpectroRadiometer
MOD03	MODIS geolocation product
MOD06	MODIS cloud optical and microphysical properties product
MOD04	MODIS aerosol properties product
MOD35	MODIS cloud mask product
MOD08	MODIS 1-degree resolution gridded product
MODAPS	MODIS Adaptive Processing System
MODIS	MODerate resolution Imaging Spectroradiometer
MODTRAN	MODerate resolution atmospheric TRANsmittance
NCCS	NASA Center for Climate Simulations
NCEP	National Center for Environmental Prediction
NOAA	National Oceanic and Atmospheric Administration
NPP	NPOESS Preparatory Mission
NRT	Near Real Time
OPAC	Optical Properties of Aerosols and Clouds
PW	Precipitable Water
QFED	Quick Fire Emission Dataset

RT	Radiative Transfer
SDS	Scientific Data Set
SWIR	Short-Wave InfraRed
TOA	Top Of Atmosphere
TPW	Total column Precipitable Water
VNSWIR	Visible, Near- or ShortWave InfraRed
WGNE	Working Group on Numerical Experimentation

Chapter 1: Introduction

1.1 Background

Throughout my career I have been asked the same question on multiple occasions: what more can we learn about the atmosphere? I have historically found the question quite difficult to answer. After all, we know quite a lot as it is and unlike faraway galaxies that I used to study fifteen years ago, Earth's atmosphere is something that you are actually in touch with and Earth's weather is something you can experience by taking a step outside your front door. You don't have to go very far. Or do you? Can you really touch it? Well, when you really think about it the topic actually gets quite complicated. You cannot experience that weather everywhere. Even if you could stand and take direct measurements of various atmospheric properties from every single point on land at exact same time, you still have that two-thirds of the planet left covered by water that we can't very well stand on.

A great deal of the information that we get about our atmosphere comes from wide-angle imaging instruments on earth observing spacecraft like EOS Terra, Aqua (Barnes et al., 1998) and Aura (Schoeberl et al., 2006), Suomi NPP (Hillger et al., 2013), Meteosat (Schmetz et al., 2002), GOES (Schmit et al., 2001) and many others. Imaging instruments on board those space-crafts take a variety of spectral measurements to obtain timely information about the state of the atmosphere over a wide area. However, there are major issues with information we get from space. Most of them are related to the fact that radiative transfer equations have a unique solution only if solved in order to obtain a radiance that would result from a known

set of conditions, but are ill-conditioned for retrieving that same set of conditions when presented with the radiance.

For a given set of atmospheric conditions the radiative transfer equation can be solved to obtain one unique value of top of the atmosphere radiance or reflectance that corresponds to those specific conditions and cloud and aerosol properties in the atmospheric column for a particular wavelength. However, many different sets of atmospheric conditions are capable of producing the same exact values of spectral radiance (or something very close to the same values). This means that when presented with sensor radiance measurements, the non-uniqueness of the inversion problem prevents the easy determination of the exact conditions in the atmospheric column that resulted in those radiances. The inversion problem can be very complex for both clouds (Nakajima and King, 1990) and aerosols (Kaufman et al, 1997). Clouds can consist of ice crystals and water droplets of various sizes and shapes. They come in multiple overlapping layers at various altitudes in the atmospheric column. Those layers can dramatically vary in thickness (Platnick et al., 2003). Aerosols come from many different sources. (Levy et al., 2007); different types of aerosols have very different scattering properties. Some aerosol types are primarily reflective; some have significant absorption. Some aerosols chemically interact with the atmosphere while others do not (Hess et al., 1998). Aerosols can be embedded in clouds or located in layers above or below clouds.

Knowledge of properties of aerosols and clouds is very important for both numerical weather prediction (Wu et al., 2002; Kleist et al., 2009) and climate studies (IPCC, 2013). Radiative forcing due to clouds and aerosols has the largest uncertainties as far

as determination of future Earth's climate is concerned (IPCC, 2013). Different types of clouds result in different climate effects.

Marine boundary-layer stratocumulus clouds exhibit an overall negative radiative effect because they are highly reflective, but in a warming Earth scenario many models indicate a net positive feedback because their fraction is predicted to decrease due to increased temperature of the marine boundary layer (Chang and Coakley, 2006). On the other hand, cirrus clouds exhibit an overall positive radiative forcing because they trap outgoing infrared radiation more than they reflect solar radiation (Stephens et al., 1990). However, most climate models appear to decrease the high cloud amount in low and mid-latitudes in a warming climate (Trenberth and Fasullo, 2009). This effect leads to a negative feedback from ice clouds, but combined with positive feedback in visible and near infrared, the contribution of ice clouds is predicted to be mostly neutral (Meehl et al., 2007).

Aerosols from anthropogenic pollution largely have a negative radiative forcing on climate though in some regions the overall affect can be positive, especially from human-caused biomass burning. Anthropogenic aerosols are also able to interact with clouds, acting as seeds for cloud particle nucleation as can be seenfrom aircraft contrails in the sky for the most visible manifestation of this effect. since smoke from aircraft engines causes a cloud to form behind it. Soluble aerosols produced in the exhaust of large cargo and military ships (in particular from diesel engines) under right conditions can form bright tracks in fields of marine boundary layer stratocumulus clouds that commonly form off the coast of California, Peru and Namibia (Radke et al, 1989). However it is very hard to quantify the effects of these

aerosols (Haywood and Boucher, 2000) and their uncertainties (Forster et al., 2007), (IPCC, 2013). There are many research articles about the effects of aerosols and the latest IPCC Workgroup I report (2013) (chapter 7 in particular) can serve as a great source of information.

To help reduce uncertainties in cloud and aerosol radiative forcing and feedbacks, we would like to have as much information about physical properties of clouds and aerosols as possible. In order to obtain that information we have to make assumptions about some of the properties of clouds and aerosols. These assumptions allow us to obtain a unique set of cloud optical and microphysical properties given a set of sensor spectral radiances (Platnick et al., 2003). Another set of assumptions can be made about some of the properties of aerosols in a given area (Levy et al., 2007) so that unique value of aerosol optical depth can be retrieved when presented with a specific spectral profile of measured sensor radiances.

Of course one must question whether the assumptions are sufficient and accurate. There are many methods used for validation of retrieved cloud and aerosol properties. Field campaigns are used to obtain direct in-situ measurements of cloud and aerosol properties with coordinated satellite sensor under-flights by aircraft carrying sensors similar to the ones on the satellite (King et al., 2010) (Chiriaco et al., 2007). However field campaigns are quite expensive and satellite under-flight opportunities during a single field campaign are limited. Ground sites such as AERONET (Holben et al., 1998) provide an ongoing stream of in-situ measurements of aerosol properties, however locations for such ground sites are limited and consistent measurements over ocean are problematic.

A direct, large-scale simulation of sensor radiances from known sets of clouds, aerosols and atmospheric properties can help in reducing the uncertainties. This can be accomplished due to the sheer volume of simulated data, "ground truth" and the ability to execute controlled experiments with single parameter variations that can be used to test the assumptions made in order to perform retrievals of cloud and aerosol properties from measured sensor radiances.

This thesis is a detailed description of such a simulator, the Multi-sensor Cloud and Aerosol Retrieval Simulator (MCARS) that was created in close collaboration with the modeling and remote sensing communities (Wind et al., 2013, 2016).

1.2 History of MCARS

1.2.1 The beginning of MCARS

In this work I present to the reader the development and evolution of the Multi-sensor Cloud and Aerosol Retrieval Simulator (MCARS).

The particular implementation of MCARS discussed throughout this work is based on simulating the MODerate resolution Imaging Spectroradiometer (MODIS) instrument currently flown on board EOS AM-1 (Terra) and EOS PM-1 (Aqua) satellites. MODIS is a wide-angle instrument with view zenith angle of 50 degrees and total swath width of 2330 km. The MODIS instrument has 36 spectral channels between 0.41μm and 14.2μm. The instrument takes data at three different resolutions of 250 m for channels 1 and 2, 500 m for channels 3-7 and the rest are acquired at 1 km native resolution. All channels are geolocated and aggregated to 1 km resolution

to create a MODIS Level-1B (L1B) radiance file MOD021KM for Terra MODIS and MYD021KM for Aqua MODIS. These product designation codes can be used to search and order the data directly from the Atmosphere Archive and Distribution System (LAADS), Level 1. These 1 km radiances are the basis for cloud and aerosol properties retrievals. The MODIS instrument is described in more detail in Barnes et al. (1998).

The code that eventually became the MCARS simulator was first described in Wind et al. (2010). This paper is presented in this work as Chapter 2. A very brief summary of this work is outlined below.

The code that became known as MCARS evolved from a need to develop a solid theoretical basis for the operational MODIS multilayer cloud detection algorithm (available as Cloud_Multi_Layer_Flag in the M(O/Y)D06 MODIS cloud optical and microphysical properties product). This algorithm allowed to create maps where thin clouds made of ice crystals overlapped clouds made of liquid water drops. The algorithm, was originally developed empirically. However in addition to demonstrating that the algorithm works one must demonstrate how well it performs and under which conditions it may encounter issues. Specifically, one would like to know when the algorithm would be likely to incorrectly identify single layer clouds as being multilayer, also known as giving a false-positive result. Alternatively the opposite could happen and the algorithm would miss a problematic overlap situation, thus giving a false negative result.

In order to answer these questions, a set of simulations has been set up where a user could feed various information about the desired test case to the DISORT-5

(DIScrete Ordinate Radiative Transfer version 5) (Stamnes et al., 1998) radiative transfer code and receive radiances and reflectances that would correspond to a given surface albedo, atmospheric column and cloud information. The individual cases originally had to be compiled by a manual process. That process has been simplified and partially automated to carry out testing of the MODIS multilayer cloud detection algorithm. It had been automated to the point where it was possible to build a database of multilayered cloud radiances that spanned almost 200,000 unique cases (Wind et al., 2010). This database is available to any interested users upon request.

There were a number of organizational issues with the resulting database. It was not maintainable. It could not be easily accessed to retrieve data for specific conditions. It was not possible to just share it with another researcher because it would have required that researcher to do a lot of upfront work in order to be able to use any of the data.

The simulator code produced output in plain text format that then had to be converted to HDF (Hierarchical Data Format). The CHIMAERA (Cross-platform HIgh resolution Multi-instrument AtmosphEric Retrieval Algorithms) system could handle cloud retrievals from these radiances, but no other code could do so without major algorithm changes. Full operational cloud retrieval could not run on that database because major modifications needed to be made to an extremely large amount of upstream code. A modified version of the cloud retrieval code, particularly for cloud mask and cloud top properties, had to be executed.

All these issues needed to be remedied because I could see how useful a pixel-level instrument simulator could be for remote sensing applications, algorithm

validation in particular. It would have been very convenient to be able to submit the data to the MODIS Adaptive Processing System (MODAPS) as if it were actual instrument-acquired data without requiring any changes to the system. Research data is also only useful as far as it can be shared with others. A data format identical to instrument-acquired radiances would allow for easy sharing because any existing MODIS data reader code would be able to read the data transparently.

1.2.2 Partnership with GMAO

Working towards the goal of creating a general-purpose synthetic radiance and reflectance product, I partnered with the Goddard Modeling and Assimilation Office (GMAO) to combine the core of the radiative transfer simulator with the GEOS-5 (Goddard Earth Observing System) global climate model. This combination code is now known as MCARS. The core of MCARS is described in Wind et al. (2010). This paper in its entirety is presented in Chapter 3. A brief summary of that work is presented in what follows.

GEOS-5 model is capable of providing virtually unlimited amounts of data about the atmospheric column, cloud and aerosol vertical profiles. I conducted a general experiment in which the satellite instrument was assumed to be flying over the model fields instead of the actual planet Earth. To that purpose I took the actual MODIS geometry files (MOD03) and used the information about time, geographic location, solar and sensor geometry to sample the global model fields and give the simulator core the required angle information. I then used the Independent Column Approximation (ICA) method to distribute the contents of individual 28km GEOS-5 grid boxes into 1km subcolumns of MODIS instrument resolution while conserving

the total grid box amounts of various quantities such as liquid water content and so on (Norris et al., 2008, 2016) (Wind et al., 2013,2016).

The resulting radiances were stored under appropriate spectral channels in the corresponding MODIS radiance file. The MODIS instrument stores geometry information separate from radiances in what is called a M{O/Y}D03 file that can also be ordered from the LAADS system. All the metadata required by automated data processing system had been preserved and the simulation output files became completely transparent also to other algorithms capable of reading data from the MODIS instrument (a research code for a new or updated retrieval method or one used in satellite data operations to produce an automated, continuous stream of physical data products for use by the public).

The first implementation of MCARS as described in Wind et al. (2013) included contributions from clouds and atmosphere. It did not include any information about aerosols and the simulator then had a name of Multi-sensor Cloud Retrieval Simulator (MCRS). MCRS code was then extended to include information about aerosols and is currently known as MCARS. In addition to clouds and atmosphere, the latest MCARS code directly includes contributions from a mixture of 15 species of aerosol provided by GEOS-5 model in 26 simulated layers. This extension is now published as a paper by Wind et al. (2016). This manuscript is presented entirely in Chapter 4 together with some of the new preliminary results that the MCARS simulator continues to provide.

I then applied the fully developed MCARS simulator to a number of cases that were considered to be good potential test cases for the simulator's performance. I

presented resulting simulated radiance files to the operational MODIS aerosol properties retrieval. Then the retrieved aerosol optical depth at 550nm was compared to "ground truth" aerosol optical depth provided by GEOS-5. The relationship between retrieval and "ground truth" for the smoke aerosol cases in Brazil looked surprisingly similar to comparisons of MODIS retrievals with AERONET ground truth. There is a known low bias in MODIS aerosol retrievals when compared with AERONET for smoke aerosol.

MCARS allows the user to apply a very fine level of control over the simulation experiment. Using that control a series of experiments was performed testing some of the individual assumptions I mentioned in section 1.1.1 that make the retrievals possible in the face of a problem of having too many unknowns and not enough information to determine the unique set of conditions that led to a particular radiance value being acquired. The shape of the smoke aerosol phase function had little effect on retrieval value. The land surface albedo inhomogeneity had some impact on the scatterplot spread, but not on shape. The cloud masking used by the MODIS aerosol optical properties code was sufficient so the product did not appear to have cloud contamination issues. The single scattering albedo value was the reason for the underestimate of retrieved aerosol optical depth at 550nm as compared to AERONET at high optical depth. The MODIS aerosol properties retrieval assumed a constant 80% column relative humidity over the entire globe at all times. That assumption may be good for aerosols that do not interact with water vapor, but smoke does chemically interact with water vapor and its single scattering albedo changes dramatically with humidity in shortwave infrared channels that the MODIS aerosol

properties retrieval uses. When the single scattering albedo was corrected, the satellite-derived retrievals aligned with in-situ data. So the assumption of constant 80% relative humidity was indeed incorrect and aerosol optical depth retrieval quality would benefit greatly if the MODIS aerosol team takes relative humidity into account when performing the retrieval (Wind et al, 2016).

The simulation experiment that ran initially on a Macintosh G4 personal computer, became the MCARS code that currently runs on the NCCS (NASA Center for Climate Simulations) Discover supercomputer on 144 processors where it is utilized in its full implementation. We are currently working on extending the MCARS code to produce synthetic data for geostationary instruments. We hope that data produced by MCARS will become an integral part of operational retrieval algorithm testing and development. The MCARS project stands as a bridge between modeling and remote sensing and it is hoped that the modeling community can also draw benefit from utilizing MCARS data.

Chapter 2: Multilayer Cloud Detection with the MODIS Near-Infrared Water Vapor Absorption Band

2.1 Introduction

Plane-parallel single-layered cloud radiative transfer (RT) models are used by global passive imager algorithms like MODIS (MODerate resolution Imaging Spectroradiometer) (Barnes et al. 1998) for cloud thermodynamic phase, cloud-top pressure/temperature, and optical and microphysical properties retrievals (King et al. 2003; Platnick et al. 2003). The use of such a RT model works reasonably well as confirmed by many field campaigns and theoretical calculations (King et al. 2004; Mace et al. 2005; Chiriaco et al. 2007; Bedka et al. 2007; Otkin et al. 2008). The model can work for some retrievals if there are multilayered clouds in a vertical column (e.g., an ice cloud overlapping a liquid water cloud) and the uppermost layer is optically thick. In particular, use of the RT model can result in biases with cloud effective radius retrievals when liquid water clouds are overlaid by relatively thin cirrus clouds (Davis et al. 2009). The retrieved effective radius of what is thought to be single layer ice clouds decreases significantly in areas overlying the water clouds. When the cirrus is too optically thin to dominate the upwelling radiance and the cloud is identified as being liquid water phase, the retrieval tends towards abnormally large water droplets. There is not a large detrimental effect on cloud optical thickness to the extent that the combined optical thickness of all layers is retrieved with little dependence on the assumed phase.

It is important to flag areas where there are problematic effective radius retrievals due to multilayer clouds of differing thermodynamic phases since those retrievals can affect cloud statistics in a significant fashion and should be excluded

from further analysis.

There have been other algorithms designed to identify multilayer clouds with passive imagers. The algorithm developed by Pavolonis and Heidinger (2004) is a pixel-level algorithm that uses ratios and differences of reflectances and brightness temperatures in various bands. This approach can be applied to historical and current multispectral imager data such as the AVHRR (Advanced Very High Resolution Radiometer) on the NOAA spacecraft and MODIS. Such an approach may also be continued with future measurements from the VIIRS (Visible/Infrared Imager Radiometer Suite) that will be flown on the NPOESS (National Polar-orbiting Operational Environmental Satellite System) platforms. This multispectral approach uses ratios and differences of reflectances and brightness temperatures. We will show a comparison of this AVHRR-VIIRS algorithm with the MODIS algorithm in section 5.

The algorithm developed by Baum and Nasiri (Baum et al. 2000; Nasiri and Baum 2004) is a statistically based algorithm that is executed in shifting steps over a box area of user-defined size, typically 200 × 200 pixels, with a restriction that some clear sky is available in the area; the algorithm retrieves a probability that the cloud is multilayered. This algorithm was developed for the MODIS instrument, but has not been used extensively outside of case studies. The need to use a large area to work on and a requirement for presence of clear sky pixels within each work area reduces the effective algorithm resolution and usefulness as many multilayered cloud retrievals occur within synoptic systems that span a wide area with extensive cloud cover. We will show a comparison of our algorithm with the Nasiri-Baum algorithm in section 5.

Another approach for multilayer cloud detection has been presented by Chang and Li (Chang and Li 2005a,b). The method of Chang and Li uses an estimation of cirrus cloud emissivity based on the difference of cloud top temperature retrieved by using the CO_2 slicing result (not assuming an opaque cloud) and the 11-μm band (i.e.,

assuming an opaque cloud). The algorithm relies on being able to identify single-layer liquid water clouds and clear sky pixels in an area of 250 × 250 km centered on the point of interest. The cloud effective emissivity is then computed, from which the infrared (IR) cloud optical thickness is derived. If that cloud optical thickness is significantly different from the cloud optical thickness retrieved using a visible or shortwave infrared (SWIR) band, the cirrus cloud likely has a liquid water cloud underneath it.

The MODIS operational multilayer cloud detection algorithm relies on a difference in above-cloud precipitable water retrievals obtained from using the 0.94 μm band versus precipitable water computed from the CO_2 slicing-derived cloud top altitude. The 0.94 μm band is relatively insensitive to optically thin cirrus and so the column moisture is integrated from the top of the atmosphere (TOA) to the lower level cloud, if such is present. The CO_2 slicing retrieval of cloud top height, and subsequent calculation of the above-cloud precipitable water from a forecast model profile, occurs from the TOA to the level of the higher cloud. From that difference, and several other tests such as the difference between retrieved IR and SWIR cloud thermodynamic phases and reflectance ratios to screen for single layer clouds over bright surfaces, a determination is made as to whether or not the cloud is multilayered in a way that affects the applicability of the plane-parallel single layer cloud models used in retrievals of cloud effective radius.

In the following discussion we present the MODIS operational multilayer cloud detection algorithm, describe how the multilayer cloud information is stored in the MOD06/MYD06 Level-2 HDF (Hierarchical Data Format) files, present results of executing the algorithm on data produced by forward simulations of multilayered clouds, and compare the algorithm to other methods.

A useful cloud simulation data set was developed using a set of moist atmospheres from the ECMWF (European Centre for Medium-Range Weather

Forecasts) 40-year reanalysis data set. Selected profiles were chosen at grid points that contained sufficient amounts of cloud to create a more realistic setting into which well-separated cloud layers were inserted. The profiles created in this fashion are available at http://modis-atmos.gsfc.nasa.gov/MOD06_L2/validation.html.

To summarize briefly the discussion that will follow, section 2.2 describes the algorithm and the data format in which the results are stored. Section 2.3 presents the details of the RT simulations and describes in detail the method used to create the simulation data set. Section 2.4 provides results from applying the MODIS operational multilayer cloud detection algorithm over the simulated scenes as well as selected MODIS data granules, and also provides an example of global statistical aggregation of multilayer cloud data. Section 2.5 discusses a direct comparison of our results with other passive remote sensor methods for detecting multilayer clouds. Conclusions, ongoing work and future directions are discussed in section 2.6.

2.2 Algorithm Description

The operational MODIS multilayer cloud retrieval uses a number of bands in addition to individual retrievals of physical quantities such as above-cloud precipitable water and cloud optical thickness, to arrive at a decision. The main component of the retrieval is a test for the difference of above-cloud precipitable water retrievals obtained by two different methods.

The first method is based on the cloud top pressure retrieval obtained from CO_2 slicing using ratios of MODIS bands 33, 34, 35 and 36 that are centered between 13.3 and 14.2 μm (Menzel et al. 2008). The retrieved cloud top pressure is then used to obtain above-cloud water vapor amount (PW_{CO2}) by adding up the layer averaged water vapor amounts from the NCEP (National Centers for Environmental Prediction)

global 6-hour atmospheric profile product, produced at 1° resolution. Due to the nature of CO_2 absorption, the algorithm is sensitive to high clouds of optical thickness (τ_c) greater than 0.5 (Menzel et al. 2008) when multilayer clouds are present and will return a low value of above-cloud precipitable water.

The second method uses water vapor absorption in the MODIS 0.94-μm band. Above-cloud precipitable water is retrieved using an iterative approach. That is possible because cloud reflectance is flat in the spectral range between 0.86 and 0.94 μm and the difference in measured cloud reflectance is due to the water vapor amount between the cloud and the sensor. If the visible optical thickness of thin cirrus layer is less than 6, the 0.94 μm band is sensitive to the low clouds when multilayer clouds are present and will return a higher value of above-cloud precipitable water than the CO_2 slicing method would. The discrepancy in retrieved amounts of above-cloud precipitable water can be attributed to the presence of multilayered clouds.

The MODIS operational multilayer algorithm first assumes that a single layered cloud exists, with a cloud top temperature based on the 11-μm brightness temperature. Cloud-top pressure is then inferred by mapping the temperature into the NCEP pressure profile. The mapping is done from the top downward so as to avoid the high likelihood of temperature inversions nearer the surface.

This cloud top pressure together with the view geometry is used to index a MODIS atmospheric transmittance table for 0.86 μm and 0.94 μm, which is generated by using the ECMWF ERA-40 atmospheric profile database as input to MODTRAN (MODerate resolution atmospheric TRANsmission) version 4.2r1 (Berk et al. 1998). This lookup results in a vector of two-way atmospheric transmittance as a function of

above-cloud precipitable water for each band. These transmittance vectors are then applied to the measured reflectances. The dominant contributor to absorption in the 0.94 μm band is water vapor. If there were no water vapor between observer and cloud, measured reflectances can be assumed to be identical. Using that assumption we look for a point where the two vectors intersect. The closest table index value of precipitable water at the intersection point is our retrieval of above-cloud precipitable water ($PW_{0.94}$). We choose to neglect a very small amount of ozone absorption in the 0.86 μm band (<0.001 additional absorption amount) as it has no discernible impact on location of the intersection point due to lookup table resolution.

We then use the retrieved water vapor amount to perform a crude atmospheric emission correction on the 11 μm radiance. Measured 11μm radiance consists of three components: emission from ground, emission from cloud and emission from atmosphere above cloud. We assume that cloud emissivity is unity, therefore we do not deal with emission from ground. This is the exact assumption made by MODIS CO_2 slicing-based cloud top properties retrieval method. Now we must subtract the atmospheric emission from measurement and also correct the result for water vapor absorption in the 11 μm channel. So the final corrected radiance takes on the following form:

$$I_{corr} = \frac{I_{meas} - B(T_{mean\_above\_cloud}) * (1 - trans)}{trans},$$

where I_{meas} is the measured radiance, $T_{mean\_above\_cloud}$ is the integrated layer mean temperature from given atmospheric profile and *trans* is the 1-way atmospheric transmittance at 11 μm.

The entire process is repeated using the corrected 11 μm radiance as a source of

cloud top temperature. We have found that one additional iteration is enough for the retrieval to converge to within 0.25K, which we consider to be a sufficient degree of accuracy for our purpose. This same type of retrieval, but without iteration, is also performed one additional time with the assumption that the cloud in question is located at 900 hPa ($PW_{0.94@900}$). If a high, cold cloud ($T_c < 265$ K) with little water vapor above it is moved vertically in the atmosphere, its retrieved temperature and pressure stay nearly constant because atmospheric transmittance for amounts of water vapor less than 0.5 cm shows very little dependence on pressure. Moving such cloud from 200 mb down to 900 mb changes 11μm transmittance by only 0.8%, 0.94μm transmittance by 1.05% and 0.86 μm by 0.01%. However, this is not so for a warm, low cloud with a significant amount (>1 cm) of water vapor above it, which is fairly typical for boundary layer clouds. For such cloud 11μm and 0.86 μm transmittances change by about the same amount as for a high cloud, but the 0.94 μm transmittance changes by 8% if such cloud with 1 cm of precipitable water above it is moved between 600 and 900 mb. The error in retrieved precipitable water amount for the lower level cloud will increase as the optical thickness of the overlaying ice cloud increases. The result is similar regardless of where the lower-level cloud lies between 800 and 1000 hPa. A low-level cloud pressure of 900 hPa is chosen as the default value. We mitigate the effect of ground elevation due to the fact that the NCEP profiles extrapolate every profile down to 1000 mb level, regardless of terrain. A precipitable water retrieval based on this assumption acts to mitigate the 'cooling' effect of an upper ice cloud and results in the inference of a more realistic high precipitable water amount above the lower level cloud. This process does not affect

the results for single-layered ice clouds or multilayered clouds where the upper ice cloud layer is optically thick, and permits the tracking of more multilayered clouds.

As both 0.94 μm and 0.86 μm channels are much more sensitive to the presence of lower-level clouds in multilayer situations, the retrieved precipitable water value is quite different from the same retrieval performed based on the inference of high clouds from CO_2 slicing. That difference, weighted by the total column precipitable water (TPW), is a key determinant of whether or not there may be multilayered clouds present. A value of

$$\frac{|PW_{0.94} - PW_{CO2}|}{TPW} > 8\%$$

is used as the threshold for marking the pixel as potentially containing multilayered clouds based on case studies and estimates regarding the occurrence of effective radius biases (see following example). Forward radiative transfer simulations, discussed in sections 3 and 4, confirm that this is an appropriate choice.

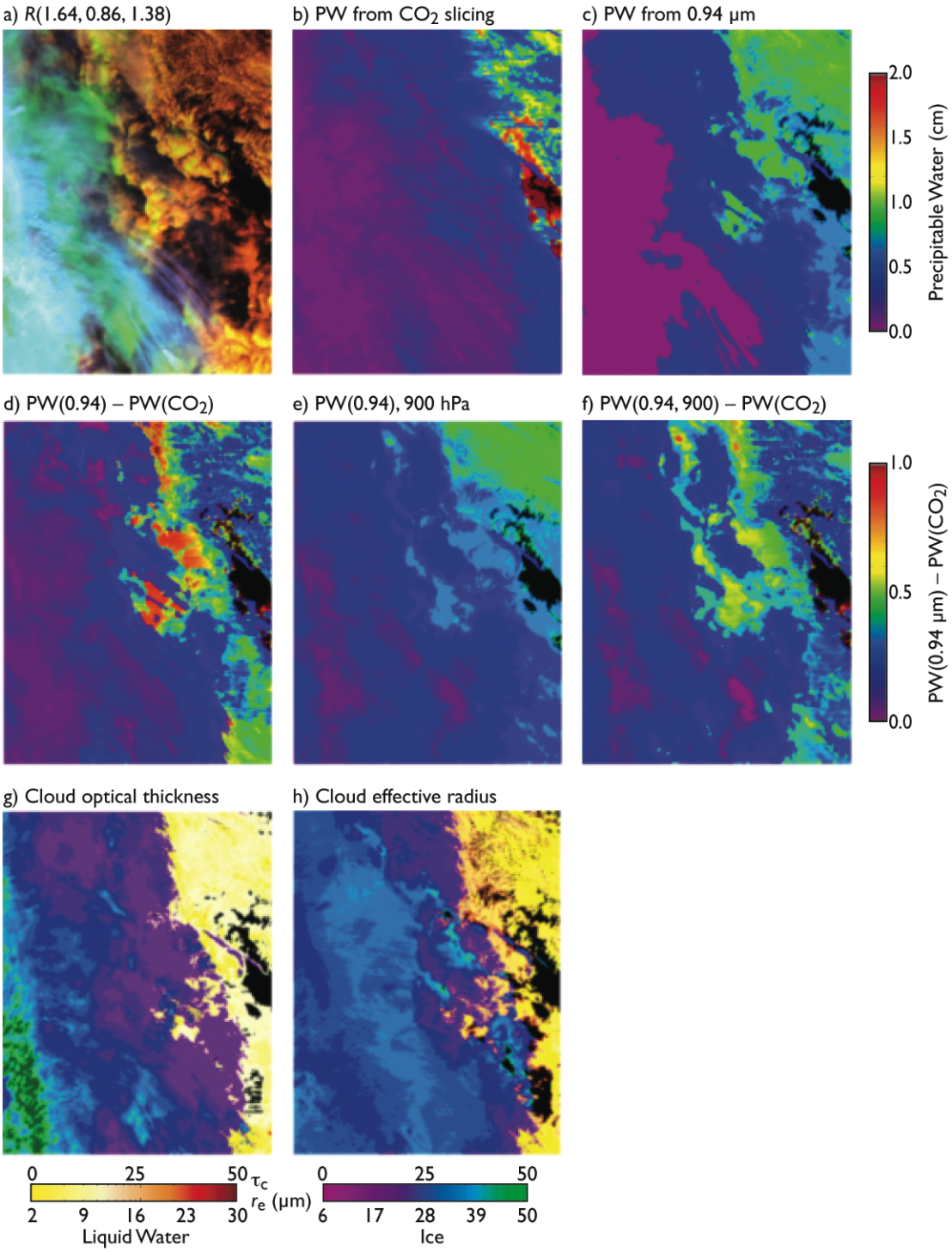

Figure 2-1. Cloud and water vapor properties over the western Pacific Ocean off Japan as acquired by *Terra* MODIS on 25 October 2008 at 00:15 UTC. The false-color composite in panel a) indicates where ice clouds (blue) overlap liquid water clouds (yellow and orange). The effect of that overlap can be seen in the other panels as abnormally small ice cloud effective radii.

Figure 2-1 illustrates the effect of this retrieval on a portion of a MODIS data granule. These data were collected from *Terra* MODIS on 25 October 2008 at 00:15 UTC in the western Pacific Ocean just east of Japan. The panels show the process of obtaining a multilayer result using the precipitable water method. Figure 2-1a shows a false-color image of MODIS bands 6, 2, 26 (1.64, 0.86, and 1.38 μm, respectively). Thin cirrus is advecting over a field of cumuliform clouds. In this false-color composite, liquid water clouds appear gold, ice clouds appear blue and white, and the ocean surface appears black. A number of areas where thin cirrus overlaps the liquid water clouds are visible in the image and take on a greenish hue. Figure 2-1b is an image of above-cloud precipitable water from the MODIS cloud top properties algorithm that uses CO_2 slicing (PW_{CO_2}). The figure shows a strong preference for high cloud properties in the overlap region, and thus a low water vapor amount is derived above clouds. There is barely a trace of the low-level clouds in the image. Very low values of precipitable water are seen for the high clouds as expected. Figure 2-1c is an image of the standard 0.86-0.94 μm retrieval of precipitable water ($PW_{0.94}$), which is more sensitive to low clouds and so gives higher precipitable water values that are more typical for those clouds. Figure 2-1d is the difference image between the precipitable water from the CO_2 slicing and the 0.94 μm algorithm. Outlines of low-level clouds are becoming clearly visible in the difference image. Small differences in precipitable water correspond to either thicker cirrus, which is not sensitive to multilayer clouds, or breaks in the low-level cumulus clouds. But more cloud could be flagged as the cirrus becomes thicker to the west and is affecting the

vertical placement of the cumulus. Figure 2-1e shows the precipitable water retrieval in which the low-level clouds are assumed to be at the 900 hPa level. It is not that different from the main 0.86-0.94 μm result with the exception that it captures some of the cloud features covered by somewhat thicker cirrus to the west. Even though the clouds are thicker, they still contain some contribution from the underlying low-level cloud. Figure 2-1f shows the difference image resulting from the 900 hPa retrieval versus that from the CO_2 slicing.

The final two images are the retrieved cloud optical thickness and effective radius for the scene. The warm colors indicate liquid water clouds with cold colors for ice cloud retrievals. The optical thickness image indicates that the cirrus is quite thin and fairly uniform over the overlap area. There is no significant impact of multilayered clouds on optical thickness as the overlying cirrus is thin and its contribution to the combined visible optical thickness is very small. In contrast, the impact on the cloud effective radius retrieval is much greater. The outlines of low-level clouds are clearly seen in the effective radius image as areas of small ice effective radii. The breaks of open water in the cumulus cloud fields return effective radius values of around 25 μm, so it is unlikely that the actual cloud microphysics is changing in the overlap area.

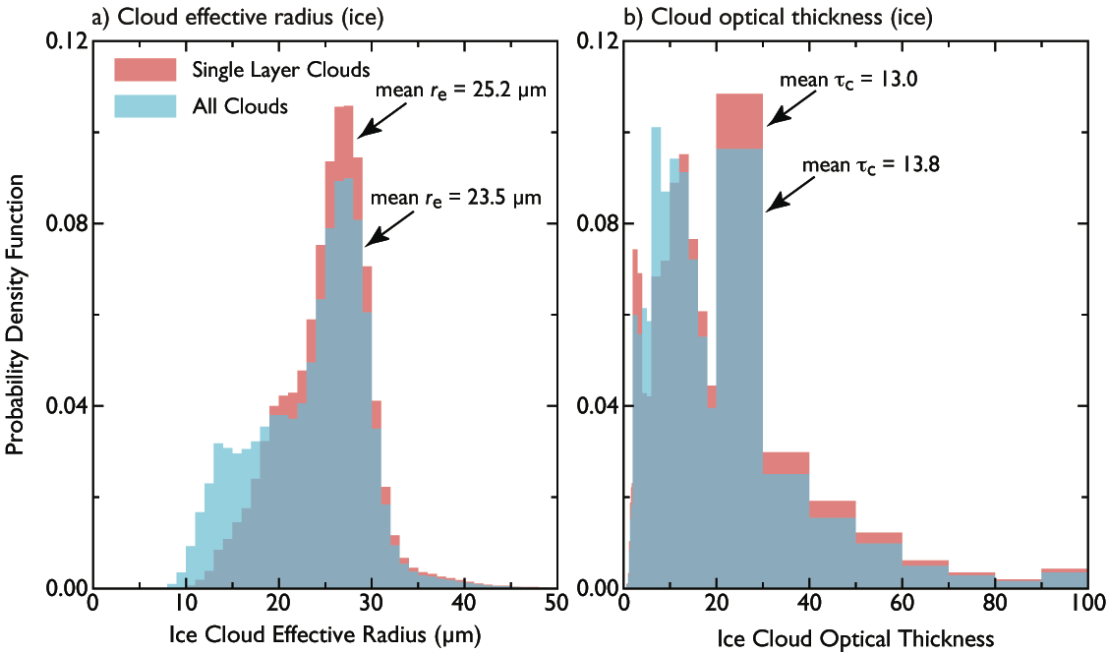

Figure 2-2. Histograms of optical thickness and effective radius for ice clouds within the scene presented in Figure 2-1. Notice the second mode of the cloud effective radius histogram at around 12µm or so. That mode is caused by increased reflectance due to bright liquid water clouds underneath.

Figure 2-2 shows the net statistical effect of multilayer clouds on cloud optical thickness and cloud effective radius. While there is not a large effect on cloud optical thickness, there is a significant shift in effective radius distribution towards smaller radii when multilayered clouds are not removed from the scene. In this particular case 19.2% of ice cloud in the scene was multilayer and ~54,000 pixels were removed from the distribution.

The MODIS CO_2 slicing algorithm is applied with the most confidence for clouds at pressures lower than about 700 hPa (Menzel et al. 2008). In a typical MODIS scene, however, the CO_2 slicing algorithm is rarely applied for clouds at pressures larger than 600 hPa. If the CO_2 slicing algorithm is unable to converge on a solution, the 11-µm band is used under the assumption that there is a low-level opaque cloud present. The choice was made to ignore CO_2 slicing results at pressures

larger than 550 hPa to minimize the potential for false positive retrievals. In light of improvement in vertical resolution to 101 levels used in MODIS CO_2 slicing algorithm beginning with Collection 6, this 550 hPa restriction may be eased in the future, although uncertainties due to resolution of the NCEP profiles will remain.

Due to uncertainties in inferring cloud emissivity from passive sensors, it is possible to obtain a false positive multilayer retrieval for the case when an optically thin cirrus cloud is present with $\tau_c < 4$. If the cloud is very optically thin, upwelling radiance from surface will cause that cloud to be placed at pressure much higher than truth. That means the 0.94 μm cloud top properties method will retrieve much higher precipitable water amount than CO_2 slicing would because of surface contamination and not because of multilayer situation. We assume that if the total column optical thickness is < 4, the likelihood is that there is not a lower cloud underneath it. The liquid water cloud layer underneath would most likely push the total optical thickness above 4. If a liquid water cloud is so thin that threshold of 4 is not reached, then we would have difficulty with retrieving effective radius due to shape of forward library space (Platnick, et.al. 2003), any multilayer situation aside. False negatives do arise from use of this threshold, but with overall effective radius retrieval uncertainty being well above 20% for thin clouds, the weight of such retrievals should be greatly reduced in any statistical studies anyhow.

We also must consider cases of single-layer clouds over bright surfaces. It is possible for the algorithm to mistake a thin cirrus cloud over a bright surface for a cloud that is multilayer. The 0.65 μm and the 1.24 μm reflectances are used to check for vegetation and snow/ice, respectively. Cloud reflectance is reasonably flat in that

spectral region, while surface albedo changes significantly. So for a true multilayer cloud situation, the reflectance ratio would be close to 1.0, but not so for a single layer of thin cirrus over a bright surface. It is useful to use ratios of 0.86 μm reflectance to 0.65 μm and 1.24 μm reflectance to check for bright surfaces, with thresholds set as follows:

$$\frac{R(0.86\ \mu m)}{R(0.65\ \mu m)} < 1.25$$
$$\frac{R(0.86\ \mu m)}{R(1.24\ \mu m)} < 1.3$$

These thresholds were empirically derived on the basis of case studies; however our forward simulations indicate that a parameterization based on ecosystem type may be more appropriate in the future. We will investigate such parameterization in MODIS data for collection 6.

In addition to the precipitable water difference, another test is based on retrievals of cloud thermodynamic phase from two different methods. The first method is the MODIS SWIR thermodynamic phase (SP) algorithm (Platnick et al. 2003) that uses a number of cloud mask tests and reflectance ratios in visible, NIR and SWIR bands to arrive at cloud thermodynamic phase. The second method is the IR bi-spectral cloud phase (IP) algorithm based on brightness temperature differences between 8.5 and 11 μm bands, which is a modification of the Baum IR tri-spectral algorithm (Baum et al. 2000). When these two methods infer different thermodynamic phases, that can be an indication of a multilayered cloud situation. This particular test tends to be sensitive to cirrus over liquid water clouds in which thin cirrus is too thin to result in an ice phase retrieval, but still biases the liquid water cloud retrievals as the cloud effective

radius retrieval is larger than expected.

The main uncertainty associated with using the thermodynamic phase test tends to arise in polar regions. At latitudes above 60°, the IR method results in quite a few undetermined phase answers due to inherent difficulties of an IR method over very cold surfaces, so we assign a lower degree of confidence to multilayered clouds that are flagged only by the cloud phase test and no other test.

The 0.94 μm precipitable water retrieval performed at both pressure at cloud top and at 900 mb, together with a test on retrieved cloud thermodynamic phase combine to create a final integer answer that tells the user whether the multilayer detection algorithm arrived at a positive result and what method(s) were positive as shown in Table 2-1.

Result Value	Description
0	Clear sky
1	Single layer cloud or cloud too thin ($\tau_c < 4$)
2	Multi-layer. Cloud phase test positive
3	Multi-layer. Precipitable water with retrieved pressure test positive
4	Multi-layer. Precipitable water with pressure fixed at 900 hPa test positive
5	Multi-layer. Both 3 and 4
6	Multi-layer. Both 2 and 3
7	Multi-layer. Both 2 and 4
8	Multi-layer. All three tests positive

Table 2-1. Listing of discrete values in SDS Cloud_Multi_Layer_Flag and definitions.

We store the final value in the MOD06/MYD06 Level-2 HDF file in two places. The values from table 2-1 are stored in a Scientific Data Set (SDS) named Cloud_Multi_Layer_Flag. The multilayer cloud information is also stored in the 5th byte of the Quality_Assurance_1km SDS as information about the thermodynamic phase of the cloud and its multilayer status. The full description of the Quality_Assurance_1km SDS is given in Hubanks (2006) and a brief listing of relevant values is given in table 2-2.

Result Value	Description
0	Cloud mask undetermined
1	Not processed (typically clear)
2	Single layer liquid water cloud
3	Multilayer liquid water cloud
4	Single layer ice cloud
5	Multilayer ice cloud
6	Single layer undetermined phase cloud
7	Multilayer undetermined phase cloud

Table 2-2. Listing of discrete values in the 5th byte of SDS Quality_Assurance_1km and definitions.

The discussion in this section is summarized in Figure 2-3. The algorithm flow chart shows the overall logical flow of the algorithm.

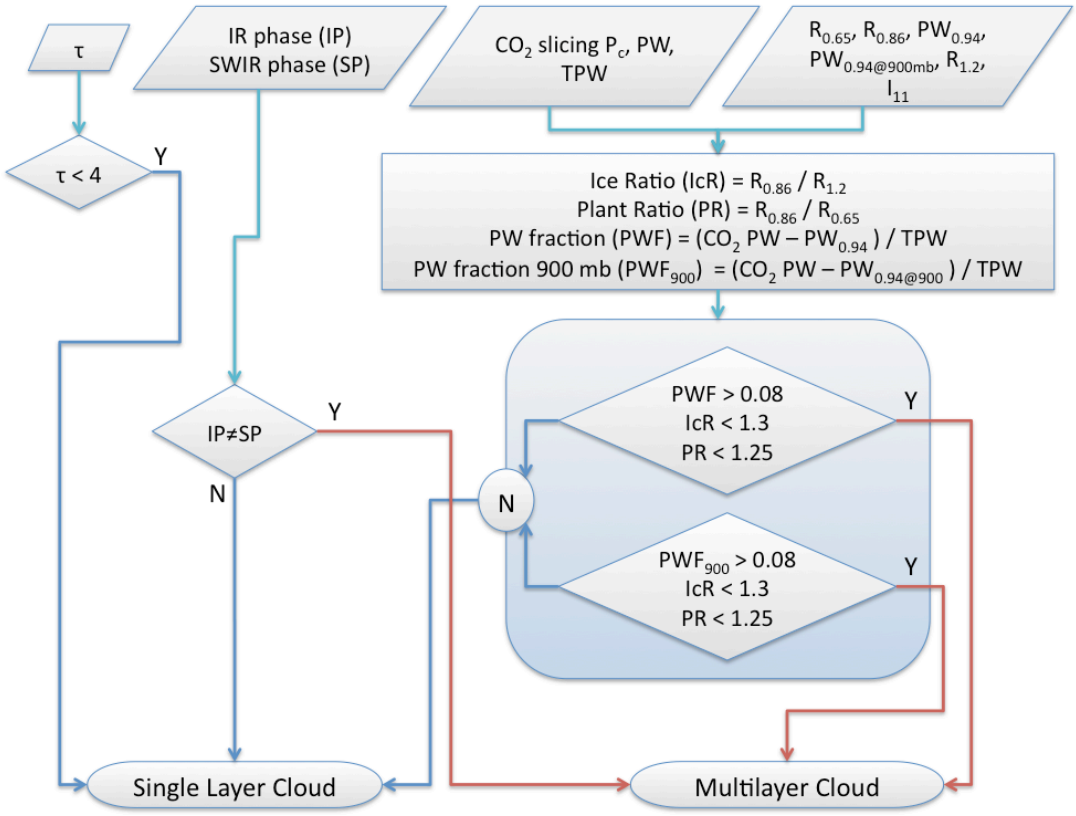

Figure 2-3. Flowchart for determining the presence of multilayer clouds using MODIS (collection 5). Red lines indicate the path to multilayer cloud detection.

2.3 Radiative Transfer Models

We have conducted an extensive set of forward RT modeling studies of multilayer clouds under varying atmospheric conditions, layer separations, surface types and layer thicknesses to thoroughly test the sensitivities and skill of the MODIS multilayer cloud detection algorithm.

Zonal and temporal average profiles are calculated from the ECMWF sampled

60-level global atmospheric profile database aggregated from ERA-40 data over 48 days for two years using 1st and 15th of each month between January 1992 and December of 1993 (Chevallier, 2001). The database profiles were separated to represent a typical midlatitude summer (MLS), midlatitude winter (MLW), tropical atmosphere (TRP), and polar oceanic (POL) profile. Profiles over polar landmasses, dominated by profiles from the Antarctic continent, were not included as they would contain strong inversions and would be likely used disproportionately for pressures lower than 700 hPa. The polar oceanic profile consists of daytime profiles only. Nighttime profiles are not used since for our purposes, cloud optical and microphysical property retrievals are performed in daytime only. We define the tropical region as 30^{o}S < latitude < 30°N, midlatitudes as 30° < |latitude| < 60° and the polar regions as above 60^{o} latitude. For midlatitudes, winter profiles occur between 1 November and 30 April; summer profiles are the remainder of the year. Within each latitude belt, profiles are chosen from regions that had cloud fraction (CF) > 0.85 to match the conditions of interest. Profiles were separated further by land and ocean using the ECMWF land fraction flag with threshold set at 0.5.

Given these averaged profiles, chosen levels were saturated with cloud having an appropriate cloud thermodynamic phase by setting relative humidity at levels that were chosen to contain cloud to 100%. The profiles were interpolated from the native 60-level resolution to 36 levels spaced at 1 km vertically between 0 and 25 km with sparser resolution in the upper atmosphere.

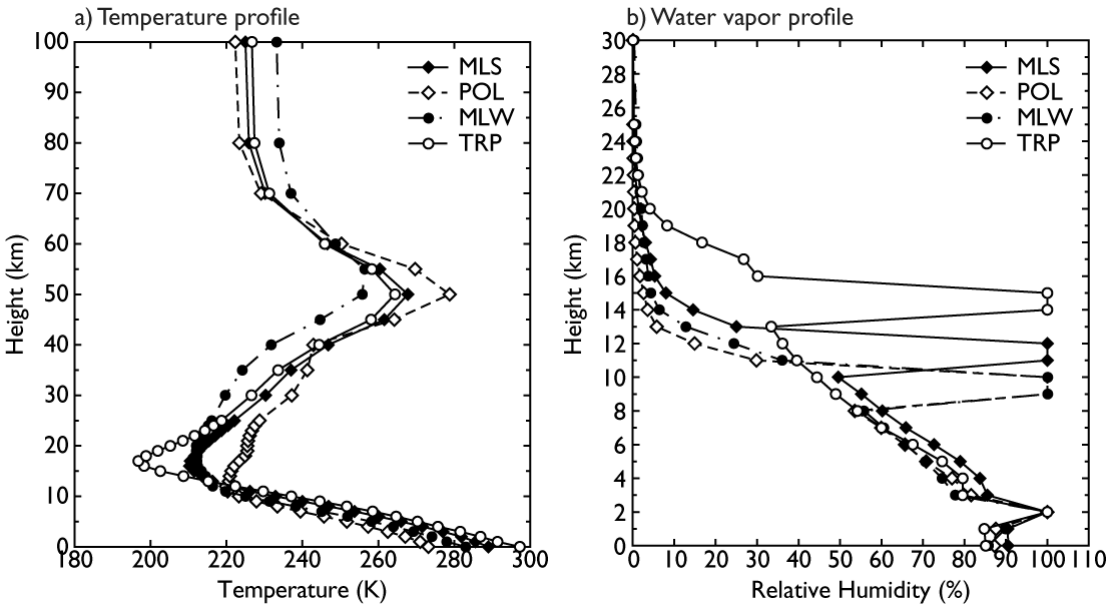

Figure 2-4. Vertical profiles of (a) temperature and (b) moisture used as a basis for the forward models. The standard profiles had been saturated with respect to appropriate phase to account for the cloud layers present.

Figure 2-4 shows a combined plot of the temperature and moisture profiles used in the simulations. These particular plots show the liquid water cloud layer at 2 km. Simulations were run for a variety of solar and view zenith angles with the solar zenith angles appropriate for the time of year in question. We sampled the solar zenith angle from the MODIS Level-3 global monthly product (Hubanks et al. 2008). The cosine of the view zenith angle corresponded to $\mu = 1.0, 0.8$ and 0.6. For detailed examination, simulations were run for ice cloud effective radii of 10, 30 and 50 μm and water radii of 6, 10 and 20 μm. An ice cloud layer of 2 km physical thickness was fixed at the base of the tropopause as indicated by temperature in each of the different profiles shown in Figure 2-4: 8 km (MLW and POL), 12 km (MLS) and 14 km (TRP) . The ice cloud optical thickness varied between 0 and 20, with increments as follows: 0.0, 0.1, 0.25, 0.5, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 10.0, 20.0. These increments were chosen specifically to examine the thin cirrus region and also to

appropriately capture the point where the ice cloud becomes too thick to be affected by the underlying liquid water cloud. Water cloud layers were assumed to be 1 km thick and were placed at two different altitudes: 2 km and 4 km. For liquid water clouds, optical thicknesses ranged as follows: 0.0, 2.0, 5.0, 10.0, 20.0.

Radiances were simulated for 16 MODIS bands, which was necessary to perform the relevant cloud mask tests (Ackerman et al. 2006; Frey et al. 2008), in particular the 3.7-11 μm brightness temperature test, the CO_2 slicing cloud top properties retrieval, and the full MODIS cloud optical and microphysical property retrievals. The set included MODIS bands 1, 2, 5, 6, 7, 19, 20, 22, 26, 29, 31, 32, 33, 34, 35 and 36 (Ackerman et al. 2006).

Each simulation was repeated over a wide variety of surfaces. The oceanic profiles only had one option (dark ocean with surface albedo of 0.05) with the exception of polar ocean that also included a sea ice surface. The land surface profiles presented options of vegetated, desert, or snow cover. Midlatitude land included mixed forest and desert with or without snow, appropriately, while tropical land included desert and evergreen broadleaf forest. All classifications were based on definitions of the International Geosphere-Biosphere Programme (IGBP) ecosystem map and the surface albedo values taken from MOD43-based 1-km resolution surface albedo product (Moody et al. 2005, 2007, 2008).

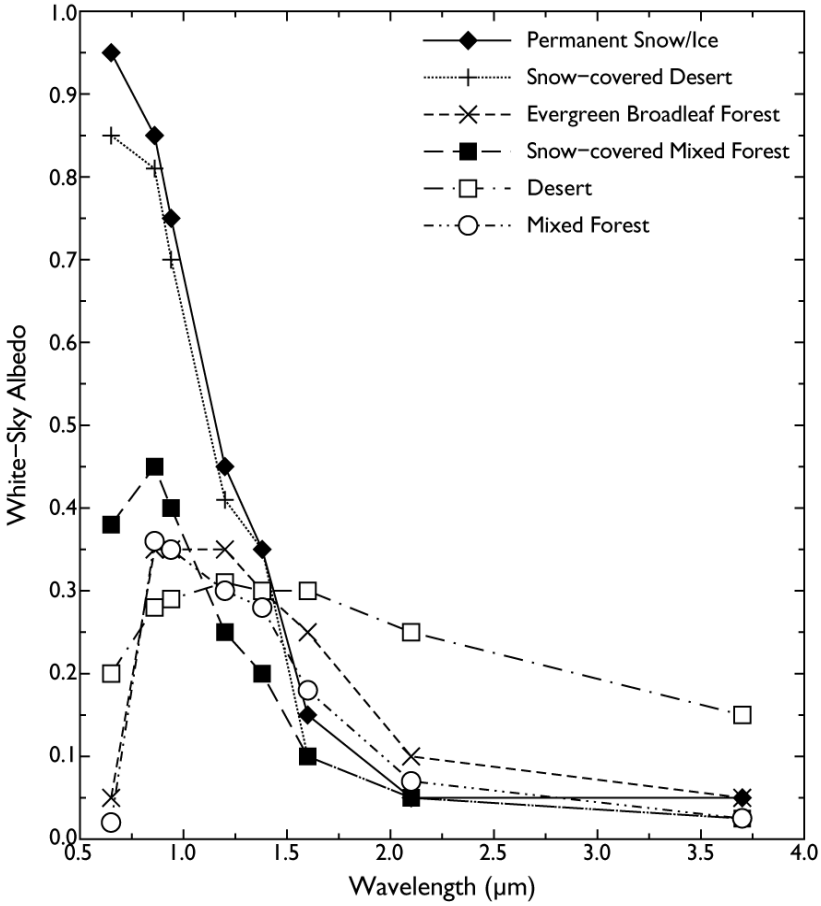

Figure 2-5. White-sky albedo as a function of wavelength for selected IGBP ecosystem classifications used in the forward calculations.

Figure 2-5 shows a plot of the white-sky (diffuse) surface albedo as a function of wavelength for the various surfaces considered in this investigation. MODIS bands that contain no solar component were given a zero surface albedo.

The RT simulations were performed using DISORT (DIScrete Ordinate Radiative Transfer) code (Stamnes et al. 1988) using liquid water cloud phase function results from Mie calculations based on the water droplet size distributions using a gamma distribution with an effective variance of 0.1 (Platnick et al. 2003) and bulk ice cloud phase functions developed by Baum et al. (2005a,b). The same phase functions for both ice and liquid water are used in the LUTs employed in the MODIS

cloud optical and microphysical properties algorithm for collection 5. The correlated-*k* method (Kratz, 1995) was used to account for water vapor and other gaseous absorbers. The DISORT code, in conjunction with the correlated-*k* method, then produced the simulated MODIS band radiances. We used 32 streams in our radiative transfer calculations, which, together with truncation of strong forward peaks and use of delta-fit method by Hu et al. (2000), can be considered sufficient computational accuracy as described by Ding et al. (2009).

With the parameter ranges described above, the forward RT calculations resulted in 26 files corresponding to combinations of atmospheric profiles and surface types. Each file contained 7560 individual data points for each geometry, optical thickness and effective radius tested. Results are provided in the following section for application of the MODIS multilayer cloud detection algorithm to a cross-section of this database of simulated MODIS radiances. In section 2.5 similar results are provided for the Pavolonis-Heidinger and Nasiri-Baum algorithms to this same dataset with comparison to the results from the MODIS operational algorithm.

2.4 Results

2.4.1 MODIS multilayer cloud retrieval

In this section we show results of applying the MODIS multilayer cloud retrieval simulated MODIS data. We show the results from a cross-section of our forward RT simulations.

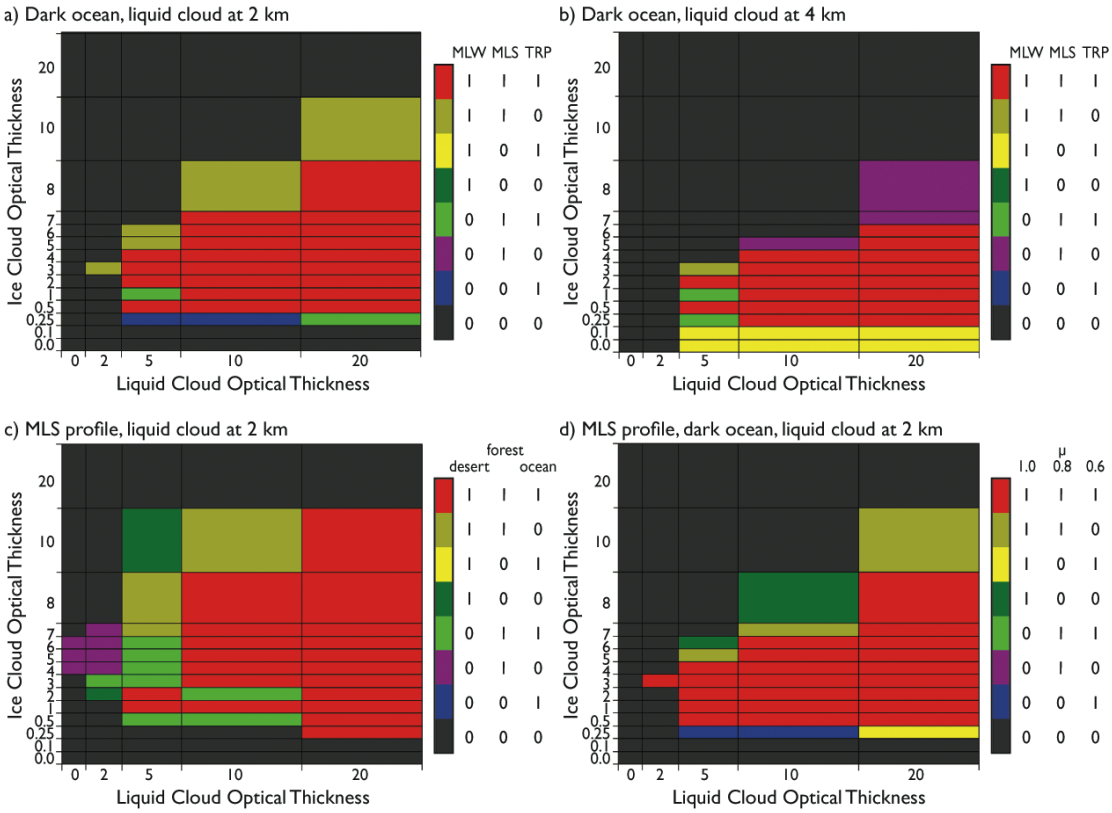

Figure 2-6. MODIS multilayer cloud detection over various surfaces, water vapor content, and view zenith angle for a cross-section of DISORT simulations.

Figure 2-6 shows a set of combined results from the DISORT forward simulations. To facilitate the interpretation of results, we group individual runs having all but one identical parameters to illustrate the effect of the differing parameter on the multilayer cloud retrieval result. Figure 2-6a combines the results of simulations conducted with a nadir view, solar zenith at 32°, dark ocean surface and liquid water cloud located at an altitude of 2 km. The atmospheric profile is varied in terms of the overall column moisture content. The plot in Figure 2-6a effectively shows multilayer cloud detection as a function of the total column water vapor. The 'bits' in the effective binary numbers that result from this data combination indicate whether or not a multilayer cloud was detected. The bit significance was arranged as

a function of the column moisture with the least significant bit for the most moisture. For example, a value of 011, which is light green in the plot, means that a multilayer cloud was detected under the conditions specified above using TRP and MLS profiles, but no multilayer cloud was detected for the MLW profile. The algorithm is more likely to detect a multilayer situation when the ice cloud is optically thin if the atmospheric moisture content is higher.

Figure 2-6b shows the same basic situation as Figure 2-6a with the exception that the altitude of the lower-layer liquid water cloud was placed at 4 km and thus decreases the cloud layer separation. When the cloud layer separation is smaller, the amount of atmospheric water vapor between the two cloud layers is also lower and so the absorption in the 0.94 μm channel is decreased over the previous case. The sensitivity of the algorithm decreases as the ice optical thickness increases compared to the case where the liquid water cloud is at 2 km altitude. Some false positives occur in which multilayer cloud is detected for thicker liquid water clouds where there is no ice cloud above. These false positives come from growing uncertainties in retrieving IR cloud phase and CO_2 cloud top properties as the cloud gets colder. The detection results can be inspected further by looking at individual tests, some of which have lower confidence than others as mentioned in section 2.2. The detection status is reported as a binary answer and may result in a false positive result.

Figure 2-6c illustrates the multilayer detection result as a function of underlying surface type, assuming a single MLS profile and a liquid water cloud placed at 2 km altitude. The surface types are arranged such that the least significant bit corresponds to the lowest overall surface albedo with no snow on the ground. The

plot shows that multilayered clouds are not detected for a desert ecosystem with thin liquid water clouds below, since the liquid cloud emissivity is likely somewhat less than 1.0, thereby indicating that we may need a separate detection threshold for deserts since the surface albedo of deserts is significantly different in spectral shape from vegetation and snow/ice surfaces. The desert spectral albedo tends to be somewhat flatter than vegetation, as Figure 2-5 shows, and so may require a somewhat different approach. The effect of this on our global statistics is not very significant as the actual cloud fraction over deserts is rather low (cf. Figure 2-9).

Figure 2-6d shows multilayered cloud detection as a function of cosine of the viewing zenith angle (μ) for a MLS profile with a dark ocean surface and a lower-layer liquid water cloud placed at 2 km. The points are ordered in μ-space such that a more oblique angle, i.e., lower μ, is the least significant bit in the binary number displayed. The relative azimuth angle for this comparison was set to 0°. The figure indicates that the algorithm is more likely to detect a thinner ice cloud over a liquid water cloud at more oblique angles. On the other hand, it is possible to flag cases with higher ice cloud optical thicknesses at more nadir view angles.

2.4.2 The Pavolonis-Heidinger algorithm

The Pavolonis-Heidinger method, originally developed for the AVHRR and adapted for the upcoming VIIRS instrument, uses a series of reflectance and brightness temperature difference thresholds described in detail in (Pavolonis et al. 2004). For the algorithm comparison purposes we have been provided with their most recent development of the method, with improvements and modifications made since the publication of their paper. Similar to the MODIS operational algorithm, it is a

single-pixel method that works on samples individually without using any spatial aggregation. Because of this similarity we were able to execute the Pavolonis-Heidinger algorithm on the results of our DISORT simulations of multilayer clouds.

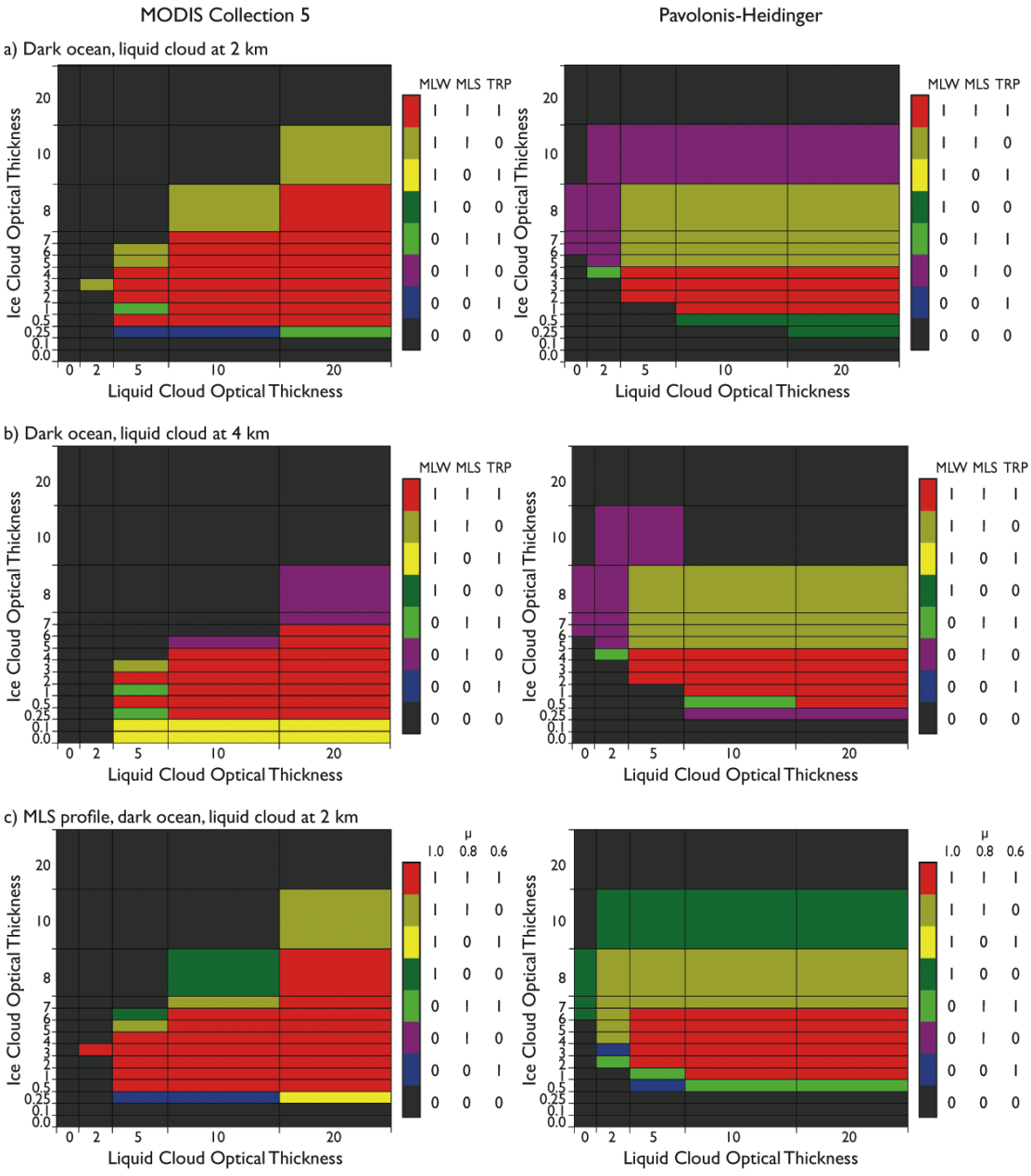

Figure 2-7. Results of MODIS (left) and Pavolonis-Heidinger (right) multilayer cloud detection for a cross-section of DISORT simulations. The Pavolonis-Heidinger algorithm appears to consistently identify clouds that are either single layer or too thick to show the multilayer cloud radiative effect as multilayer.

Figure 2-7 shows the comparison of these results. The figures on the left are the MODIS results from Figure 2-6a, 2-6b and 2-6d, and on the right are corresponding Pavolonis-Heidinger results. We compared the algorithms for three out of four database cross-sections shown in Figure 2-6. It was not possible to perform the exact comparison for the surface-type section, since the Pavolonis-Heidinger algorithm uses a 0.41 μm band over desert regions that we did not include in the original DISORT band set. The Pavolonis-Heidinger algorithm uses a lookup table (LUT) derived from simulations of multilayered clouds over a various surfaces. The LUT includes the difference in brightness temperatures (BTD) between the 11 and 12 μm bands. A threshold function is defined since the multilayered clouds (i.e., ice over water cloud) display a BTD as a function of visible reflectance that is quite different from single-layered liquid water and ice clouds. In addition to that threshold, a number of constraints are placed on reflectances at 0.65 and 1.38 μm to help with the identification of single layer clouds over a variety of surfaces. The 1.65 μm band is used by the algorithm to aid in identifying the thermodynamic phase of clouds since ice clouds have greater absorption than liquid water clouds at 1.65 μm.

There are similarities in the results as well as some differences, but overall the comparison is favorable. The MODIS algorithm has a somewhat wider section where multilayer clouds are detected for the entire range of the varied conditions, be it atmospheric moisture content or view angle. However the detection rate generally drops off as the ice cloud thickens with only the thickest simulated liquid water cloud showing at ice cloud optical thickness of 10. Both algorithms show that once the ice cloud optical thickness reaches 20, no detection of multilayer clouds is possible. Both

algorithms also show that detection is a function of layer separation with detection rate being lower when the liquid water cloud is placed at 4 km as opposed to 2 km cloud top altitude. The Pavolonis-Heidinger algorithm shows more detection when both cloud layers thicken, but not as much when the cloud layers are thin.

We could not run the Nasiri-Baum algorithm on DISORT simulations because it is a statistical aggregate algorithm that depends on natural variability of the data within a certain box area. More specifically, the algorithm uses a brightness temperature difference between 8.5 and 11 μm to confidently determine the clouds that are liquid water phase. Liquid water clouds tend to have a large, negative brightness temperature difference. The data are analyzed on 200 × 200 pixel tiles that must meet a number of conditions to attempt the retrieval. The tile must contain clouds and must contain at least 10 pixels each of ice cloud, liquid water cloud, and clear sky. Additionally, a scatter plot of 11-μm brightness temperature versus the 2.1 μm reflectance is created, wherein pixels that belong to single layer ice clouds and single layer liquid water clouds create two distinct lines. The clear sky points lay on the intersection of those lines. The angle between the lines must be greater than 20°. The grid of tiles undergoes successive systematic shifts to increase the number of times a particular pixel is processed; the more times a pixel is flagged as multilayered, the higher the confidence of the final answer. As DISORT results are single points, there is no way to create an appropriate analysis box that would satisfy the data requirements of the Nasiri-Baum algorithm.

Our overall conclusion from examining all these results is that the MODIS multilayer cloud detection algorithm is robust and performs as intended under a wide variety of conditions.

2.5 Analysis and comparison with other methods

In this section we show an example case study from MODIS and comparisons of our method against two other multilayer cloud detection algorithms, which we mentioned in section 2.1.

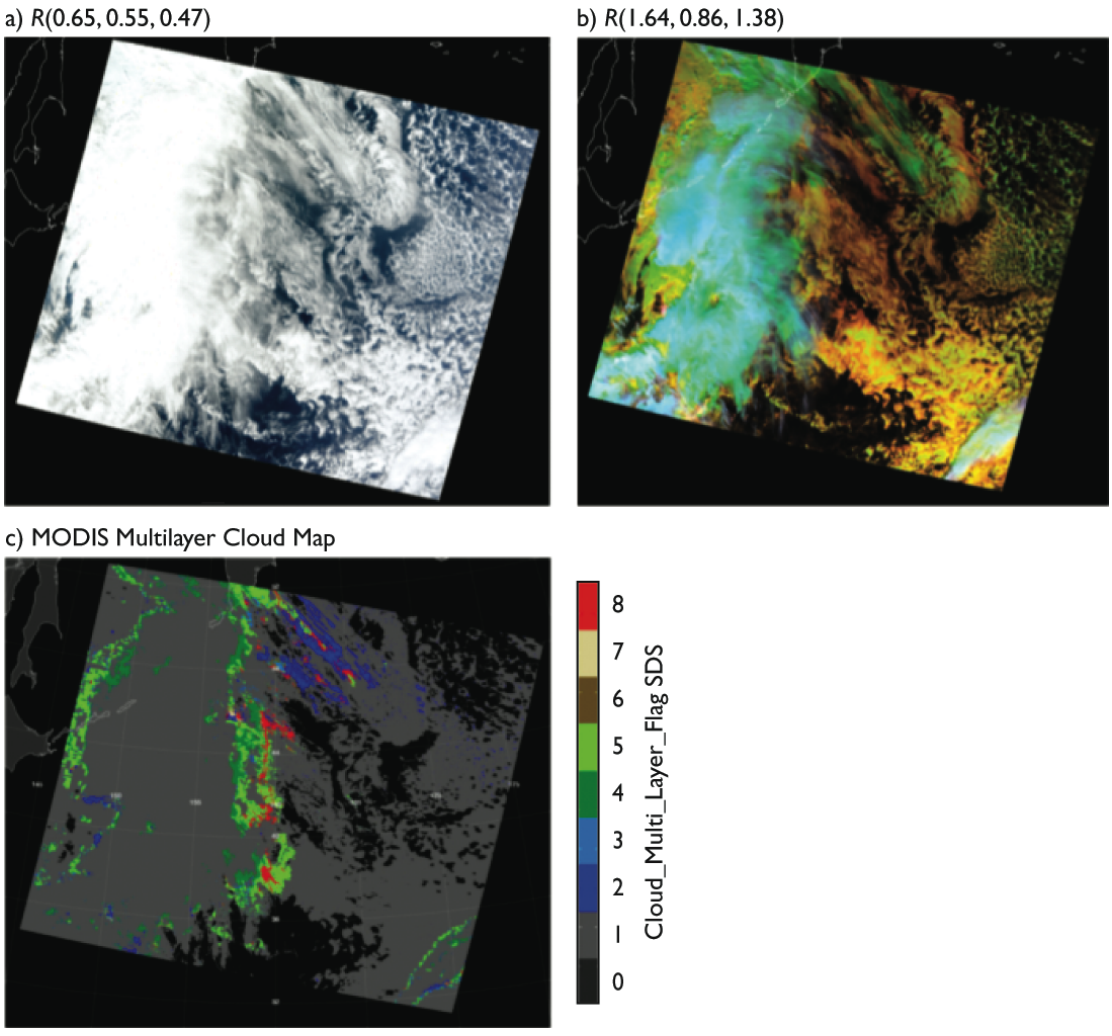

Figure 2-8. Multilayer cloud over the western Pacific Ocean off Japan on 25 October 2008 at 00:15 UTC from Terra MODIS. Same granule example has

been used in figure 2-1.

Figure 2-8 shows an example of multilayer cloud detection for a *Terra* MODIS granule acquired on 25 October 2008 off the coast of Japan at 0015 UTC. This is a full granule, a portion of which was shown in Figure 2-1. Figure 2-8a shows an atmospherically corrected true color image formed as a composite of MODIS bands 1, 4, and 3 (0.65, 0.55, and 0.47 μm, respectively). The false color image indicates that the scene has a high cloud content, but tells us very little about the actual complexity of the scene. Figure 2-8b shows a false color image formed as a composite of bands 6, 2, and 26 (1.64, 0.86, and 1.38 μm, respectively), which more readily separates clouds of different thermodynamic phase by color. There is a significant amount of multilayer cloud in this scene, indicated by areas where the yellow liquid water clouds show through the more blue and white ice clouds. Figure 2-8c shows the multilayer cloud retrieval. Different values on the color scale correspond to different tests flagging the cloud as clear sky (0), single layer cloud (1), and multilayer (2 through 8) cloud, as described in Table 2-1. These results are not an absolute measure of multilayer cloud amount, but rather a map of areas where presence of multilayer clouds negatively affects cloud effective radius retrievals.

MODIS multilayer cloud retrievals are aggregated to the global level-3 daily, eight-day, and monthly products as an average of data down-sampled to 5 km and aggregated into a 1° grid. The multilayer cloud fraction is stored, combined and separated by thermodynamic phase, and also includes mean values of cloud optical and microphysical properties retrievals with and without multilayer clouds. (Hubanks et al. 2008)

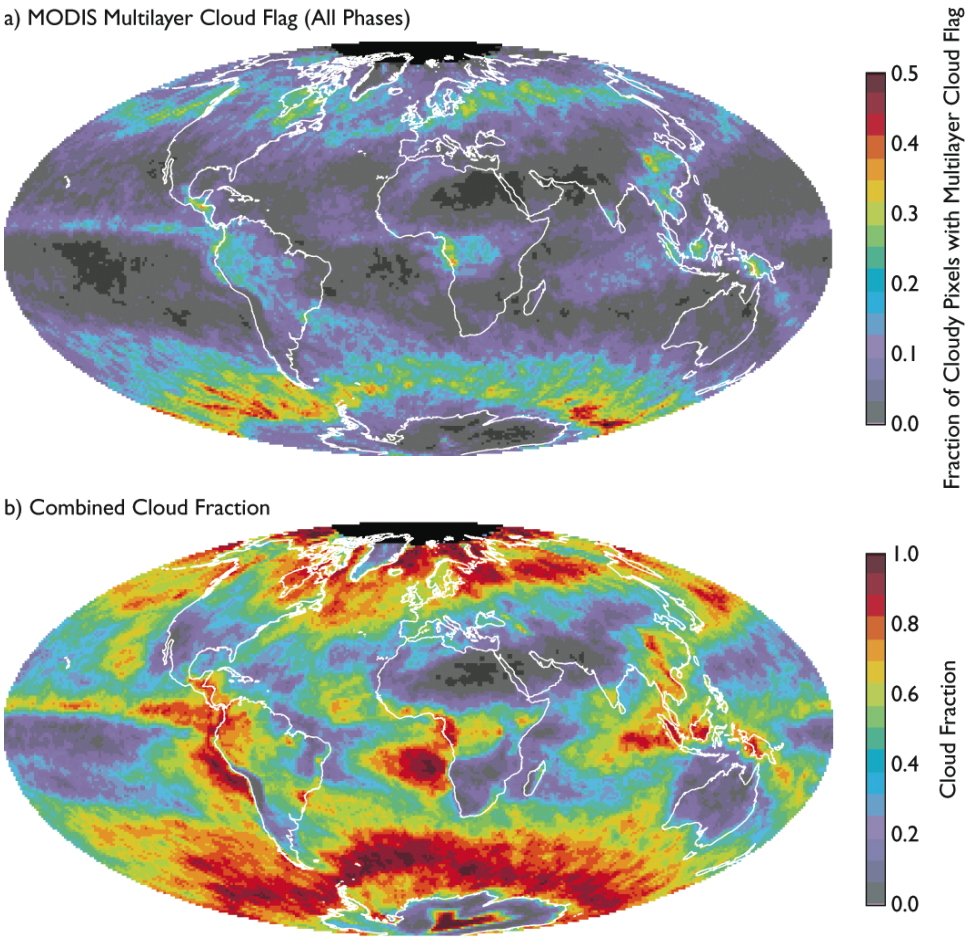

Figure 2-9. *Terra* MODIS monthly level-3 global products for October 2008. Panel a) shows the multilayer cloud fraction and panel b) is the overall cloud fraction. The multilayer cloud fraction increases in the vicinity of areas of deep convection where anvil cirrus would likely be transported over underlying low-level clouds.

Figure 2-9 shows an example of such an aggregation for the month of October 2008 derived from *Terra* MODIS data. Figure 2-9a shows the fraction of all cloudy pixels that have the multilayer flag set, and Figure 2-9b shows the mean monthly cloud fraction. The small black area on the very top of the images corresponds to polar darkness or low sun where no retrievals are attempted (cosine of the solar zenith angle $\mu 0 < 0.15$).

A monthly global map like this is useful for providing the spatial distribution of multilayered clouds. Based on observational evidence, one might expect a higher frequency of multilayered clouds to occur in the vicinity of low-pressure systems and their frontal boundaries. Higher frequencies of multilayer clouds tend to occur in the Southern Ocean and in the northern midlatitudes. The anvil cirrus from the ITCZ (InterTropical Convergence Zone) is also a likely candidate to create multilayer cloud situations. A good portion of the ITCZ is flagged as multilayer in the eastern Pacific Ocean. Strong convective zones over rainforest areas also tend to generate anvil cirrus, resulting in high frequencies of multilayered clouds in the Congo basin, Borneo, and New Guinea. One can also see the effect of advection of anvil cirrus over the marine stratocumulus zones in the Southern Hemisphere off the coasts of Peru and Ecuador, and in the Gulf of Guinea.

The three multilayer cloud detection algorithms available to us are now applied to the MODIS granule shown previously in Fig. 2-8, with the results shown in Figure 2-10.

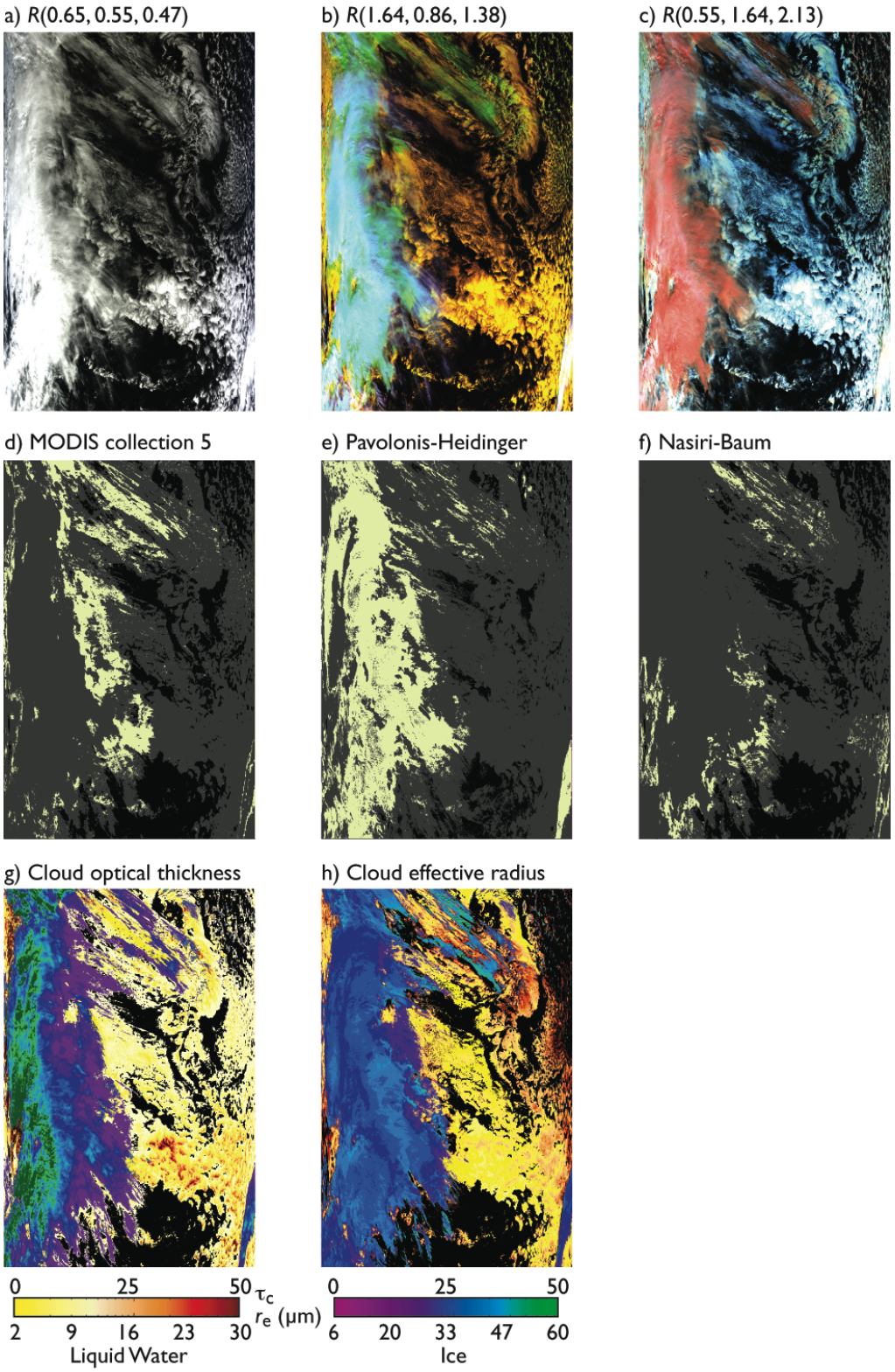

Figure 2-10. Multilayer cloud analysis and cloud optical properties over the western Pacific Ocean off Japan as acquired by Terra MODIS on 25 October 2008 at 0015 UTC. Panels a), b) and c) show a variety of color composites that

bring out different aspects of the cloud field under consideration. Panel b) shows the cloud overlap better, while panel c) enhances the presence of ice phase clouds (red). Panels d), e) and f) show the multilayer cloud detection maps for three different multilayer cloud detection algorithms. Panels g) and h) show the operational Collection 5 retrievals of cloud optical thickness and cloud effective radius using the 0.86-2.1μm band combination.

Figure 2-10a shows the true color composite constructed from bands at 0.65, 0.55, and 0.47 μm, Figure 10b the false color composite constructed from bands at 1.64, 0.86, and 1.38 μm, and Figure 2-10c the false color composite constructed from bands at 0.55, 1.64, and 2.13 μm. Figures 2-10(d-f) show the results of applying the multilayer cloud detection using the d) MODIS operational algorithm, e) Pavolonis-Heidinger algorithm, and f) Nasiri-Baum algorithm.

As there is a wide range of options, described in the code documentation, that the Nasiri-Baum algorithm can be executed under, for the purposes of this comparison we took the suggested default values. The Nasiri-Baum algorithm can only be executed under conditions that some clear sky, liquid water cloud and ice cloud exists within the box being currently analyzed, so the algorithm does not attempt retrievals over a portion of this granule. The Nasiri-Baum algorithm also outputs its result as a probability of overlap. For clarity we display non-zero overlap probability as a positive answer. We performed a similar procedure with the results from the MODIS operational multilayer cloud algorithm, combining the multilayer values 2 through 8 into a single positive identification value. The Pavolonis-Heidinger algorithm returns its result as a single value so no additional data conversion was necessary to visualize the results.

Overall many of the same areas flagged as multilayer, even though the results may not look exactly the same, as the different multilayer algorithms had been developed with different purposes in mind. The Nasiri-Baum algorithm gives the fewest multilayer occurrences, but that can be attributed to a limited area over which the algorithm attempted retrievals. The main disagreement between Pavolonis-Heidinger and our algorithm arises in the flagging of thicker high clouds as being part of multilayer scenes (left side portion of the granule). One can make a valid argument that most, if not all, ice phase clouds in that part of the granule are multilayered clouds because of the apparent wide presence of low clouds in that region as well as there being some indication in the 1.38 μm false color composite. The result given by the Pavolonis-Heidinger algorithm is consistent with detection achieved for simulated DISORT data, where clouds with combined extinction optical thickness as large as 30, with the upper layer thickness of 10, can be flagged as multilayer, as shown in Figure 2-7. The decision whether to flag a cloud as multilayer depends on the issue being addressed. In our case, we are looking for multilayer clouds that challenge the applicability of our single-layer plane-parallel cloud models used in cloud optical and microphysical properties retrievals. Our goal is to create a map of areas where the model application is problematic. From our RT simulations we have found that the effect of ice cloud overlapping a liquid water cloud on cloud effective radius retrieval diminishes quite rapidly with ice cloud optical thickness and is barely detectable when ice cloud optical thickness becomes greater than about 6. So, whereas it may in fact be true that the thicker upper level clouds in this granule are also multilayer,

having those clouds flagged as such does not address our primary objective regarding microphysical biases.

2.6 Conclusions and future directions

In this chapter we presented the MODIS operational multilayer cloud detection algorithm used in the MODIS collection 5 cloud optical and microphysical properties product. The multilayer cloud detection method was developed to address a need to indicate areas of cloud where an assumption of single-layer plane-parallel cloud models was challenged due to the presence of two distinct cloud layers with differing thermodynamic phases and the upper cloud layer having low extinction optical depth. Such situations manifest themselves as areas of abnormal cloud effective radius retrievals. Our method used the difference between retrieved above-cloud precipitable water amounts from the 0.94 μm band and from CO_2 slicing together with a number of other tests. The physical basis of the multilayered cloud detection algorithm was provided, with examples of results from forward simulations as well as case studies involving MODIS data and global aggregations of results. Results from this approach were compared to two other methods of multilayer cloud detection. We also presented a set of standard cloudy atmospheres that we developed in order to perform our studies. Wherever possible we performed all comparisons utilizing a single source dataset, so the differences in retrieved results would be solely due to differences in methodology.

Our results and analysis indicate that the multilayer cloud detection algorithm presents a reliable means of identifying situations that would create difficulties for

retrievals of cloud effective radius. The forward simulations indicate that there are very few false-positive results and that they arise under conditions, which would give high retrieval uncertainty due to one of the cloud layers being extremely thin. Forward simulations, performed under a wide variety of surface and atmospheric conditions, indicate that the algorithm gives a reliable result regardless of observation considered. These forward simulations laid a solid foundation for the MCARS simulator described in detail in the two chapters following.

Chapter 3: Equivalent Sensor Radiance Generation and Remote Sensing from Model Parameters.

3.1 Introduction

Accurate knowledge of cloud cover and cloud properties is important in model studies that involve Earth's radiative budget, climate prediction and numerical weather prediction. High clouds are observed to have a net warming effect on the atmosphere because of their low albedo and low temperature. Low clouds have a net cooling effect due to their high albedo and relatively small temperature contrast with the surface. Clouds and their interactions with aerosols are significant sources of uncertainty in climate prediction studies (IPCC, 2007). In addition, clouds continue to be the main source of climate feedback uncertainty and hence climate sensitivity (e.g., Bony et al., 2006).

The Goddard Earth Observing System Version 5 (GEOS-5) earth system model is maintained by the Global Modeling and Assimilation Office (GMAO) at NASA Goddard Space Flight Center (GSFC). GEOS-5 contains components for atmospheric circulation and composition (including atmospheric data assimilation), ocean circulation and biogeochemistry, and land surface processes. Components and individual parameterizations within components are coupled under the Earth System Modeling Framework (ESMF, Hill et al. 2004). In addition to traditional meteorological parameters (winds, temperatures, etc., Rienecker et al. 2008), GEOS-5 includes modules representing the atmospheric composition, most notably aerosols (Colarco et al. 2010) and tropospheric/stratospheric chemical constituents (Pawson et al. 2008), and the impact of these constituents on the radiative processes of the

atmosphere. GEOS-5 has a mature atmospheric data assimilation system that builds upon the Grid-point Statistical Interpolation (GSI) algorithm jointly developed with NCEP (Wu et al. 2002, Derber et al. 2003, Rienecker et al. 2008). The GSI solver was originally developed at NCEP as a unified 3D-Var analysis system for supporting global and regional models. GSI includes all the in-situ and remotely sensed data used for operational weather prediction at NCEP. GEOS-5 also includes assimilation of Aerosol Optical Depth (AOD) observations from the MODerate resolution Imaging Spectroradiometer (MODIS) imager on the NASA Earth Observing System (EOS) *Terra* and *Aqua* spacecraft; an algorithm for assimilating cloud property information from measurements in the visible and infrared portions of the spectrum is currently under development (Norris and da Silva 2013). While the GEOS-5 meteorological assimilation includes a wide variety of space borne sensor data, traditionally samples containing clouds are carefully screened out. The near real-time GEOS-5 data assimilation and forecasting system runs at a nominal horizontal resolution of 25 km with 72 vertical layers (Rienecker et al. 2008, Molod et al. 2012).

The MODIS instrument (Barnes et al. 1998) is a passive imager, producing a wide variety of remotely sensed data products for land, ocean and atmosphere disciplines from 36 spectral channels. Data Collection 5.1 processing includes algorithms for retrieving cloud cover amount (Ackerman, et al. 2006, Frey et al. 2008), cloud top properties such as cloud top pressure and temperature (Menzel et al. 2008) and cloud optical and microphysical properties such as cloud optical thickness, cloud effective radius and cloud water path (Platnick et al. 2003; Wind et al. 2010; Zhang and Platnick 2011; King et al. 2013).

In this chapter we present a technique that brings together remote sensing methods and model-generated fields. We use MODIS geolocation data to sample GEOS-5 fields as if the MODIS instrument were flying over the model fields instead of the Earth surface. Once the sampling is complete, we generate equivalent sensor radiance data for the MODIS footprint. We then replace the contents of the 1 km, 500 m and 250 m resolution MODIS Level-1B (Xiong, et al. 2006) radiance files with these simulated radiances and insert the resulting alternate data stream into the start of the MODIS Adaptive Processing System (MODAPS) operational algorithm processing chain for the atmosphere discipline cloud products mentioned above (product designation MOD06 and MYD06 for Terra and Aqua MODIS, respectively). The data stream is fully transparent to the system so that pixel-level (Level-2) retrievals can be aggregated through the same gridded (Level-3) 1°x1° code (Hubanks et al. 2006; King et al. 2003, 2013) used in MODAPS production (MOD08 and MYD08 for Terra and Aqua, respectively). We have performed tests of processing MOD06 data produced from MCRS radiances through to the Level 3 code. The Level 3 code transparently ingested the data as the retrieval files were produced by the official operational code with all appropriate metadata produced as if the data granule came from an actual sensor. There are many potential uses for the resulting Level-2 and Level-3 data. Level-3 model aggregations can be compared to archived MODIS Level-3 and GEOS-5 source data fields directly as a means of model validation and study of model biases that could exist. Level-2 data can be used to study some aspects of retrieval algorithm behavior and sensitivities since all retrievals are performed with known (prescribed) *truth*.

The equivalent sensor data framework had been developed with instrument flexibility in mind, so that by simple substitution of spectral response functions and data reader, the MODIS instrument can be replaced by other space borne or airborne sensors, currently in operation or part of a future concept, and a different sensor data stream can be produced. Thus, identical products from different sensors or different retrieval algorithms for the same sensor can be compared and analyzed in a controlled environment, which can provide insight and lead to improvements in remote sensing algorithms.

This flexibility extends to model data as well. Any climate or weather prediction model fields can be used as long as a means of ingesting the necessary parameters is provided. Thus synthetic retrievals based on multiple models can be compared and analyzed using the same sensor interface in a controlled environment, leading to a consistent diagnostic toolset. Furthermore, this detailed simulation capability can function as a test bed for very fast simulators such as the Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package (COSP) (Bodas-Salcedo, et al. 2011) or the hyperspectral simulator of Feldman et al. (2011).

In section 2 we describe the model–sensor interface using the GEOS-5 model and MODIS imager. Section 3 shows an example of simulation and retrieval of cloud properties on a sample Level-2 data granule. In section 4 we elaborate on future directions for the software suite.

3.2 Radiance simulations at scales smaller than the model's grid spacing

We start the process by selecting an area and time period of study. It can be as small as a few-pixel subsection of a single MODIS granule or as large as an entire year of MODIS data. The study size is only limited by availability of computing resources. As far as model itself is concerned, there is no need for actual MODIS data to be present, but we specifically want the actual MODIS radiances to be available, so that retrievals from simulated and actual radiances can be compared directly. Similarities and differences in those retrievals can be analyzed and results applied on a variety of levels in order to improve both the model and the sensor retrieval algorithm.

For simplicity's sake in all subsequent references and illustrations the study area will be taken to be a standard 5 minute MODIS data granule (approximate 2000 km in the along track direction by 2300 km). Once the granule is chosen, we proceed to choose model output files that bound the granule time. For example, for the granule at 0200 UTC, we would select model output at 0000 and 0300 UTC. We use a MODIS standard geolocation file (MOD03 product) to define the spatial locations to sample the model fields. Solar and view angle information contained in the same MODIS geolocation file is also used in the simulation. For the examples shown in this chapter we used the GEOS model v.5.7.2 output. A listing of specific GEOS-5 fields and products used in this suite is given in Table 3-1.

Field Code	Description
U10M	U-component of wind speed at 10m altitude
V10M	V-component of wind speed at 10m altitude
FRSEAICE	Sea ice fraction
FRSNO	Snow fraction
PS	Surface pressure
T2M	Temperature at 2m altitude
SLP	Mean sea-level pressure
QV2M	Specific humidity at 2m altitude
O3	Ozone concentration profile
T	Temperature profile
DELP	Level pressure differential profile
RH	Relative humidity profile
CLOUD	Radiative cloud fraction profile
QLLS	Large scale cloud liquid water mixing ratio
QLAN	Anvil cloud liquid water mixing ratio
QILS	Large scale cloud ice mixing ratio
QIAN	Anvil cloud ice mixing ratio

Table 3-1. GEOS v.5.7.2 fields and products used in simulations. Additional information about the field codes can be obtained from Rienecker et al (2008).

3.2.1 Surface albedo determination

In order to save on computational time we pre-determine surface albedo for the area of study. The surface albedo data comes from a variety of sources. Over ice-free ocean, MODIS geometry and model wind speed are used in a Cox-Munk ocean surface BRDF model (Cox and Munk, 1954) to produce cloud-free ocean surface reflectance. This model reflectance is calculated for four cardinal wind directions and then averaged into a lookup table (LUT), that is a function of wavelength, wind speed, solar and sensor zenith angles and relative azimuth angle. We do the calculation at three wind speeds of 3, 7 and 15 m/s. The LUT has 33 solar zenith values, 28 view zenith values and 37 relative azimuth values. We linearly interpolate as needed to obtain surface spectral albedo in the selected fifteen MODIS channels

that have a shortwave reflective component. The MODIS channels used in the simulation and their central wavelengths are listed in Table 3-2.

Channel number	Central wavelength (μm)
1	0.65
2	0.86
3	0.47
4	0.55
5	1.24
6	1.63
7	2.13
8	0.41
9	0.44
17	0.91
18	0.94
19	0.94
20	3.7
22	3.9
26	1.38
27	6.2
28	7.3
29	8.5
31	11.0
32	12.0
33	13.2
34	13.4
35	13.8
36	14.2

Table 3-2. MODIS channels used in simulations.

The ocean reflectance LUT contains data for channels 1-22 and 26 from the table. For the IR (infrared) channels that have no reflective component (27-36), a constant value of 0.015 is used. This value is based on the ocean surface emissivity value suggested by the MODIS cloud top properties algorithm.

Over land several methods are utilized to model the radiances for all MODIS channels. We use the MODIS land surface spectral albedo gap-filled dataset (Moody

et al. 2005, 2008) that has been updated for MODIS data Collection 6 and is derived from both Aqua and Terra data. In addition to providing 16-day time period averages every 8 days, the gap-filled albedo files are generated for each year separately (instead of aggregating all years together as was done previously). Further, spatial resolution has been improved to 1 km. These files are derived from the Collection 5 MCD42B product (C. Schaaf et al., personal communication). This MODIS land surface albedo product is used directly for channels 1-7 and interpolated linearly to cover other MODIS spectral channels that fall within the 0.47 – 2.14μm wavelength range. For wavelengths that are longer than 2.14μm we use the surface emissivity dataset used in MODIS clear-sky profile retrievals (Seemann et al. 2008). For wavelengths shorter than 0.47μm we use a seasonally averaged surface albedo database utilized by the MODIS deep-blue algorithm (Hsu et al. 2004) to obtain the albedo for the 0.41μm channel and then interpolate for the 0.44 μm channel.

Over snow and sea ice we use the MODIS zonal snow/ice albedo dataset (Moody et al. 2007). The table lookup is determined by the MODIS pixel latitude, the International Geosphere–Biosphere Programme (IGBP) ecosystem type and snow/ice fractions from GEOS-5 FRSNO and FRSEAICE model fields.

The resulting surface albedo values are written out to file for each study area so that they can be referenced later for different simulation scenarios involving same area and time. A good example of such varying scenarios would be repeated experiments with or without the presence of aerosols.

3.2.2 Water vapor and other gaseous absorbers

After the surface albedo is determined, we proceed to ingest the profiles of temperature, relative humidity, ozone concentration and atmospheric pressure. These profiles are downsampled to 27 atmospheric levels, as listed in Table 3-3, from the GEOS-5 native 72 vertical levels and sent to an atmospheric transmittance module that uses the correlated-*k* method (Kratz, 1995) to calculate weights and optical thicknesses for each atmospheric layer due to water vapor and other gaseous absorbers.

Level number	Level altitude (km)
1	80
2	60
3	50
4	45
5	40
6	35
7	30
8	25
9	20
10	18
11	16
12	15
13	14
14	13
15	12
16	11
17	10
18	9
19	8
20	7
21	6
22	5
23	4
24	3
25	2
26	1
27	0

Table 3-3. Vertical levels used in simulations.

In cases where the surface is encountered at an altitude higher than 0 km, the profile is trimmed accordingly and the surface level is inserted as the last level to be used. We perform the vertical downsampling in order to save on the computational cost of the equivalent sensor radiance simulation step, with the bulk of downsampling occurring above the tropopause. We preserve finer vertical resolution in the troposphere. We find this to be permissible as radiance data stored in a MODIS L1B file has an accuracy to only the 5th decimal place and the uncertainty due to varying the number of vertical levels in the upper atmosphere is less than this data storage accuracy.

3.2.3 Water vapor and other gaseous absorbers

Sampling of model cloud-related fields to the MODIS pixel scale is not straightforward because cloud properties typically vary on scales not adequately resolved by the operational 0.25° GEOS-5 resolution. To sample cloud fields, 1 km MODIS pixels for each GEOS-5 grid column are collected and the same number of pixel-like sub-columns is generated using a statistical model of sub-grid column moisture variability. The general approach of Norris et al. (2008) is followed, namely using a parameterized probability density function (PDF) of total water content for each model layer and a Gaussian copula to correlate these PDFs in the vertical.

In this application, we use the skewed triangle PDF, which allows a simple inclusion of moisture variability skewness, a ubiquitous feature of atmospheric

boundary layers. This PDF has a simple scalene form characterized by three parameters: a lower and upper bound and a mode. Under some circumstances, these three parameters can be directly diagnosed from the *layer mean* total water and condensate contents, q_t and q_c, and cloud fraction f, but in many cases some adjustments are necessary to f, and possibly q_c, to achieve consistency. The details of this calculation are described fully in Norris and da Silva (2013). Approximations must also be made in the case of clear or overcast layers, when the triangle is under-determined.

For the Gaussian copula we use a correlation matrix with a fixed vertical decorrelation scale of 100 hPa, further modified by a multiplicative Riishojgaard (1998) flow-dependent correlation in total water that permits sharper decorrelation across inversion features. Further details are given in Norris and da Silva (2013). Once the correlation matrix is specified, the Gaussian copula correlated ranks of each of the grid column's layers are easily generated (Norris *et al.,* 2008) and then inverted with the cumulative distribution function (CDF) of each layer's skewed triangle distribution. The net result is an ensemble of subcolumns of total moisture content that sample the specified layer PDFs and have the specified vertical correlations and accompanying cloud and condensate overlap properties. The transformation of total moisture content to vapor, liquid water and ice contents assumes the vapor is capped at the GEOS-5 saturation vapor content and that the excess moisture is condensate, split between the phases using an ice fraction linear in temperature over the -35 to 0°C range. It is these subcolumn condensates, combined with GEOS-5 diagnostic effective radii, that are used to evaluate subcolumn (or "pixel") liquid water and ice

optical thicknesses for each layer. These are input to the MODIS radiance simulator code.

Note that the subcolumns generated in this way are horizontally independent (the independent column approximation or ICA), but are subsequently "clumped," or rearranged, to give horizontal spatial coherence, by using a horizontal Gaussian copula applied to condensed water path. This clumping acts to give the generated clouds a reasonable horizontal structure, such that the cloudy pixels in a grid column are actually grouped into reasonable looking clouds, rather than being randomly distributed. This is important because the MODIS cloud optical and microphysical properties retrieval algorithm has some spatial variance tests for potentially partially-cloudy pixels, removing cloud edges by the so-called "clear-sky restoral" (Zhang and Platnick 2011; Pincus et al. 2012). If clumping is not used, then individual points generated by ICA stand an exceptionally high chance of being eliminated by the clear sky restoral unless a model grid box has a nearly 100% cloud fraction.

In practice, the clumping algorithm works as follows: a correlation matrix C is a generated between all pixels in a grid column based on the distance between the actual pixels in the MODIS granule and assuming a nominal 5 km decorrelation length. If there are N pixels, C is an $N \times N$ matrix. This matrix is used by a Gaussian copula to generate N correlated ranks, which are subsequently used to sample (effectively, to re-order) a list of the N *simulated* pixels that has previously been sorted by column condensed water path (CWP). Because horizontally nearby pixels are more correlated by C, they will have a higher chance of having similar ranks, and therefore similar values of CWP. In this way the pixels are grouped together

horizontally into coherent clouds. [Note that this clumping acts on subcolumns as a whole, and independent of the preexisting vertical correlations in the ICA subcolumns, so the clumping will work better for single cloud layers. For multilayer clouds, the layer that dominates the CWP will dominate the clumping.]

The sub-grid column cloud generator described above is, of course, only one of many possible generators. A less complicated example, very much akin to the internal GEOS-5 treatment of cloud overlap, would be the following "homogeneous cloud, maximum-random overlap" generator: divide the atmosphere into pressure bands, e.g., low, middle and high bands, with interfaces at 700 and 400 hPa. Say we again wish to generate N subcolumns, $n = 1,...,N$, for the grid column. Then for each pressure band, generate a set of N uniform random numbers $\{r_n\}$ on [0,1], and for each model layer k falling within the band, assign cloudiness to layer k of subcolumn n if $r_n < f_k$, where f_k is that layer's cloud fraction. The fact that the *same* set $\{r_n\}$ is used for each layer k in the band enforces maximum cloud overlap *within* the band. But choosing independent sets of $\{r_n\}$ for each pressure band enforces random overlap *between* the bands. Finally, every subcolumn which is cloudy at layer k, shares the same homogeneous in-cloud condensate contents $\frac{q_{(i,l)k}}{f_k}$, where $q_{(i,l)k}$ are the layer mean condensate contents.

Note that this simple generator, as with the earlier more sophisticated generator, produces subcolumns of *condensate*. The specification of optical thicknesses from condensate contents proceeds in the same way in both generators, as presented earlier. We emphasize this strategy because the reader should be very aware of the potential traps associated with using diagnostic layer cloud optical thicknesses *directly* from

GCM (e.g., GEOS-5) output files. When using diagnostic layer cloud optical thicknesses directly, one must know whether they are in-cloud or "layer mean", and if in-cloud, for what cloud fraction. For example, for GEOS-5, the layer cloud optical thicknesses TAUCLW and TAUCLI are "in-cloud" values consistent with the maximum cloud fraction f_{max} of the layer's pressure band, *not* with the layer's actual cloud fraction f_k. This is because the GEOS-5 diagnostics internally dilutes each layer's in-cloud cloud optical thickness (in an approximately radiatively-consistent manner) by stretching each layer's cloud fraction to its band's f_{max}, in order that it may simply add the diluted layer "in-cloud cloud optical thicknesses" within a band to produce a "band in-cloud cloud optical depth", such as GEOS-5's "TAULOW". Because of this, the GEOS-5 diagnostic TAULOW, for example, can be regarded as the low-band in-cloud cloud optical thickness consistent with the model's low band cloud fraction CLDLOW. Similarly for TAUMID and TAUHGH. But note that one cannot simply add TAULOW, TAUMID and TAUHGH to get a column in-cloud optical thickness, because the actual column value depends on the overlap of these bands. Currently this overlap is random, and so we could express the column in-cloud optical thickness in terms of the sum over the 2^3 combinations of cloud/clear in the three bands, each with their respective fractions (e.g., a fraction CLDLOW*CLDMID*CLDHGH of the grid column would have a column TAU of TAULOW+TAUMID+TAUHGH, a fraction CLDLOW*CLDMID*(1-CLDHGH) would have a column TAU of TAULOW+TAUMID, etc.), and finally all normalized by the column cloud fraction CLDTOT = 1 – (1-CLDLOW)*(1-CLDMID)*(1-CLDHGH).

Not surprisingly this gives

$$TAUTOT = \frac{CLDLOW * TAULOW + CLDMID * TAUMID + CLDHGH * TAUHGH}{CLDTOT}$$

for the column in-cloud optical thickness, because we are assuming that we can average optical thickness (including zeros) in the horizontal, to get a layer mean optical thickness. That assumption in itself is rather questionable, since cloud radiative properties are non-linear in optical thickness. It is therefore much more accurate to deal with radiative averages over subcolumn ensemble optical thicknesses, generated as we have described, than with "layer mean" or "band in-cloud" optical thicknesses directly.

The point we are ultimately making is that one cannot simply interpret the column consequences of model layer cloud diagnostics without a knowledge of the model's cloud overlap. This is why GEOS-5 now includes the COSP simulator suite to produce satellite observables with an inbuilt treatment of the model overlap.

3.2.4 Radiative transfer calculation

Now that we've collected all the necessary information about atmosphere and cloud layers, we begin the simulation process. The radiative transfer calculations were performed using the Discrete Ordinate Radiative Transfer (DISORT) code (Stamnes et al. 1988) with liquid water cloud phase function results from Mie calculations based on gamma distribution water droplet size distributions with an effective variance of 0.1 and bulk ice cloud phase models developed by Baum et al. (2005), both consistent with MOD06. We have experimented with a different number of computational streams in order to balance speed and desired accuracy. We found

that only 16 streams were required to achieve the needed precision. Generally a large number of streams is required to accurately model the forward peak of the phase function and multiple scattering components. Initial calculations were done with 32 streams, however the execution time was rather prohibitive. We settled on 16 streams as a balance between execution time and precision as the difference in resulting equivalent sensor radiance between 32 and 16 stream simulations was less than 0.5%.. Also the forward peak is further truncated and use of the delta-fit method of Hu et al. (2000) can be considered sufficiently accurate, as described by Ding et al (2009), for calculations where there is no stored accuracy limit such as the multilayer cloud simulations in Wind et al (2010). As we pre-calculate the surface spectral albedos, we can save further time by calling DISORT in Lambertian mode with predetermined values. When we encounter cloud subcolumns over the ocean, however, we must adjust the computed Cox-Munk surface albedo to compensate for the diffuse illumination that the presence of the cloud creates. A good value for the diffuse illumination albedo of a water surface is 0.05 (Platnick et al. 2003). We then linearly fit surface albedo as a function of cloud optical thickness, with full diffuse illumination at a total column cloud optical thickness of 3 and full Cox-Munk surface albedo at total column cloud optical thickness of 0.

3.3 Example retrievals

In this section we discuss a couple of example results of radiance simulations and subsequent cloud property retrievals. We performed the simulation on the NASA Center for Climate Simulations (NCCS) Discover system using 12 Intel Westmere nodes with 12 cores each. The memory footprint of the software suite is very small,

around 80 Mb peak usage, but the process is quite CPU-heavy. A full-resolution 1 km simulation using a full MODIS granule as a study area took about three and a half hours wall clock time to complete.

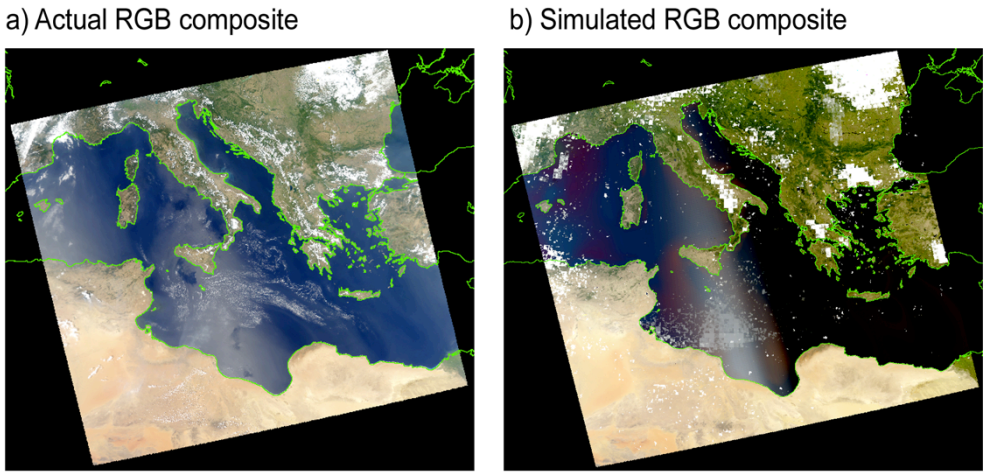

Figure 3-1. Equivalent sensor radiance simulation together with an actual MODIS granule that was used as study area. Panel a) shows the true-color composite of an actual MODIS granule and panel b) shows the simulation that used GEOS-5 model fields that were concurrent with the real granule.

Figure 3-1 shows resulting true-color RGB image of sample MODIS granule 2012 day 228 (August 15th) at 12:00 UTC together with the true-color image of the actual MODIS granule before its channel data was replaced. Panel a) shows the actual data acquired by Aqua MODIS and panel b) shows the simulation result. GEOS-5 does not assimilate cloudy radiances and so there should be little expectation of a granule-level feature match. However, in this case the model does remarkably well with cloud placement. Bands of cloud over southern France are present and located properly, as are clouds over the northern Balkans and southern Asia Minor. The orographic clouds over Italy and Greece are also present, as are scattered clouds over the Sahara desert. There are some rather important differences between the model and

the actual data, however, when it comes to cloud properties.

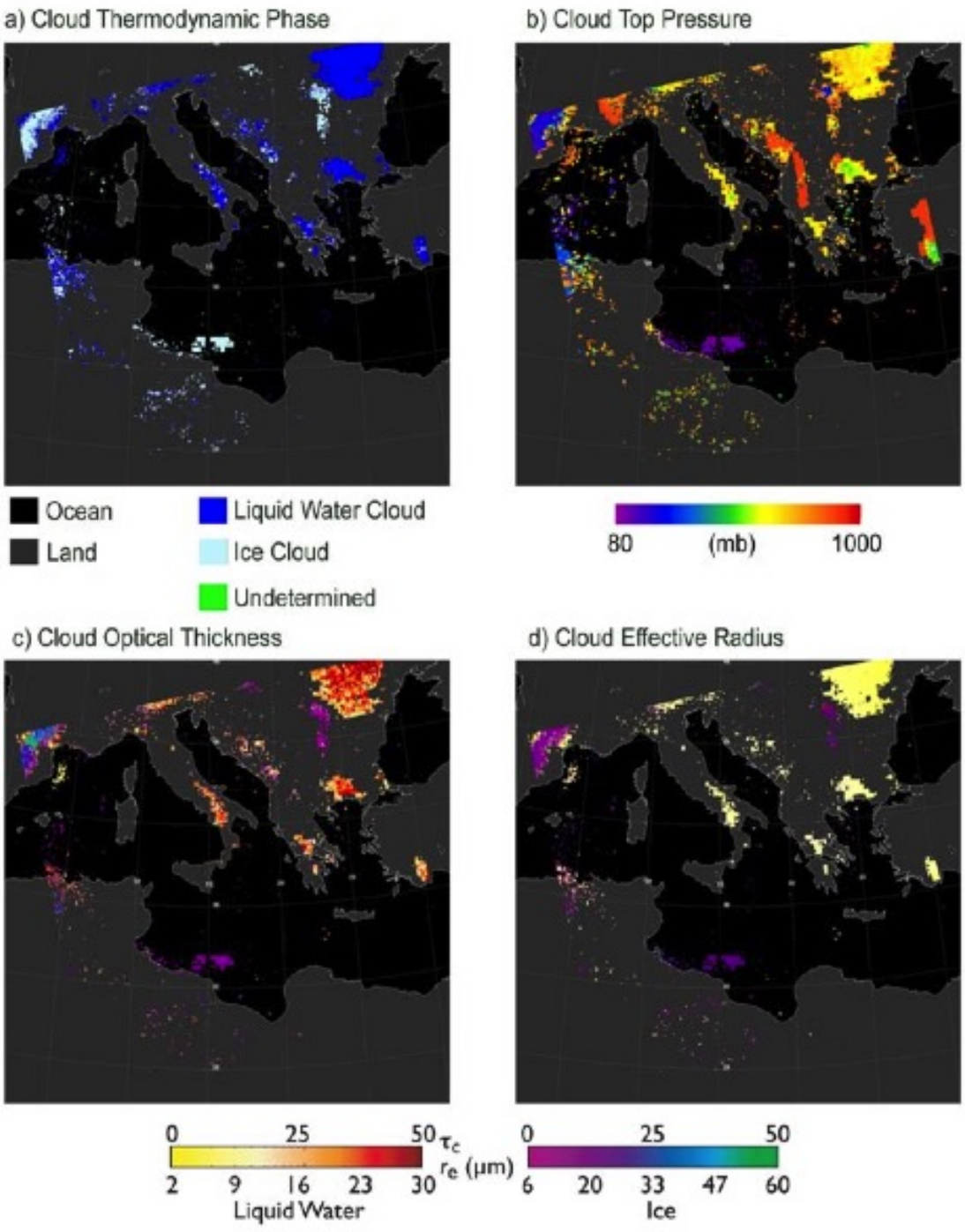

Figure 3-2. Example operational Data Collection 5 cloud top, optical and microphysical properties retrieval for synthetic granule covered by Aqua MODIS 2012 day 288 12:00 UTC.

Figure 3-2 shows the results of running the Data Collection 5.1 operational retrieval chain on the resulting L1B file from model fields. Panel a) shows the cloud thermodynamic phase, panel b) the cloud top pressure, panel c) the cloud optical thickness and panel d) the cloud effective radius retrieved with the VNSWIR (Visible, Near- or ShortWave InfraRed) and 2.1μm channel combination. Figure 3-3 shows the actual Aqua MODIS retrieval for that same granule using identical panel arrangement. The cloud field over the central Mediterranean is given improper vertical location by the model. The actual cloud field is retrieved as liquid water and is low cloud, with cloud top pressures of 800-900mb. The model generates a thin cirrus cloud in that location with cloud top pressure of about 100mb and of course ice thermodynamic phase. This has serious implications for outgoing radiation The cloud field over Romania has a consistent phase, but the model indicates the cloud to be positioned somewhat higher in altitude than the observation and also significantly optically thicker than what is observed. The same is true for the cloud field over NW Turkey.

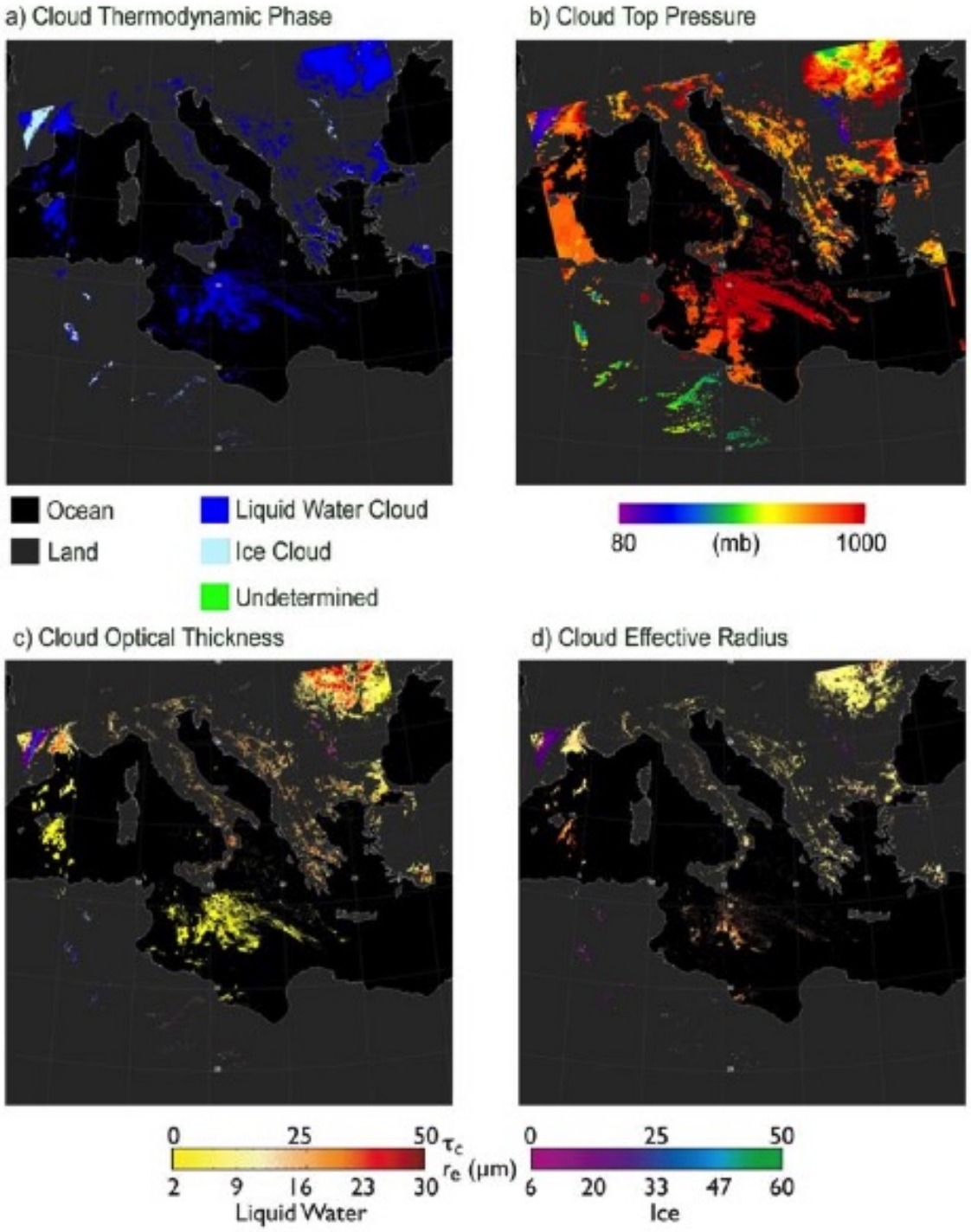

Figure 3-3. Actual operational Data Collection 5 cloud top, optical and microphysical properties retrieval for sensor-acquired granule Aqua MODIS 2012 day 228 at 12:00 UTC .

Figure 3-4 shows a cloud top pressure / cloud optical thickness joint histogram for the actual granule in panel a) and simulated one in panel b). Whereas in this comparison we are not necessarily looking for quantitative evaluation of model parameters, some things do tend to jump out.

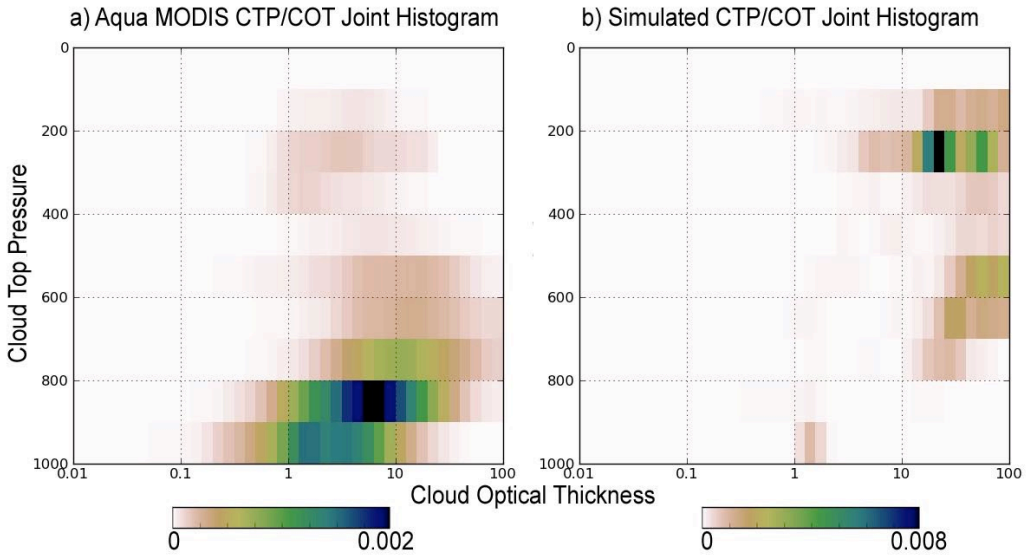

Figure 3-4. Joint histograms of cloud optical thickness vs. cloud top pressure for actual (a) and model-based (b) cloud fields covered by Aqua MODIS 2012 day 228 at 12:00 UTC. In this case the model is placing clouds too high in the atmosphere compared to the observed cloud field.

The actual MODIS granule has mostly low clouds that are moderately thick. The simulated granule on the other hand lacks low clouds almost entirely and instead produces thicker clouds at high altitude. The RGB images look very similar in this case, so the model is performing well on geographical cloud placement, but fails rather badly when it comes to proper cloud placement in altitude. This kind of disconnect can have some significant implications for Earth radiative budget calculations. It is one of our future goals to determine just how frequently such

disconnects occur on the global scale.

Figures 3-5 through 3-8 show another simulation example, this time from Terra MODIS 2013 day 151 (05.31.2013) at 11:15UTC.

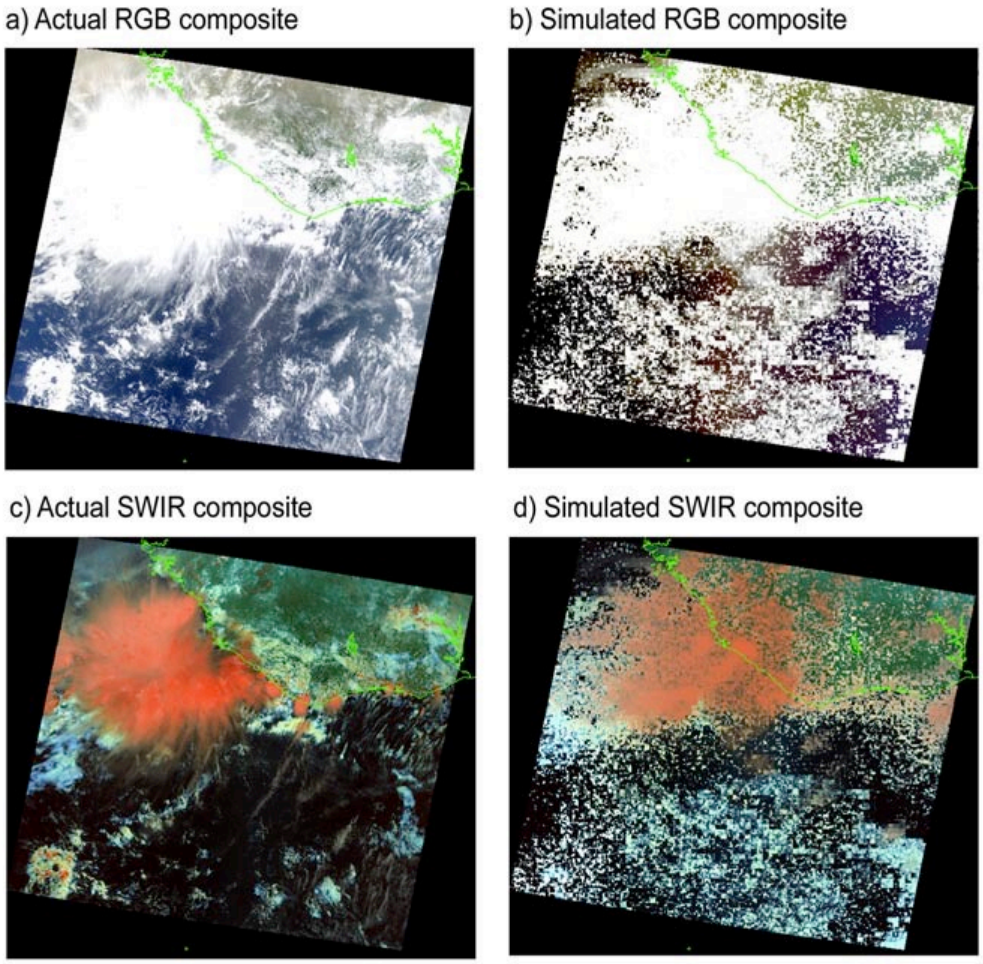

Figure 3-5. Equivalent sensor radiance simulation together with an actual MODIS granule that was used as study area. Terra MODIS granule 2013 day 151 at 11:15 UTC. Panels a) and c) show the true-color and a false-color that enhances differences in cloud thermodynamic phase composites for an actual observed granule. Panels b) and d) show same color composites for a synthetic granule produced by the MCRS code.

Figure 3-5 shows a true-color RGB image for simulated and actual MODIS granule together with the SWIR composite that allows user to visually estimate cloud

thermodynamic phase. Ice clouds appear red in such image. We can not show one for Aqua MODIS because of detector issues with Aqua MODIS 1.6μm channel that is needed to create the image.

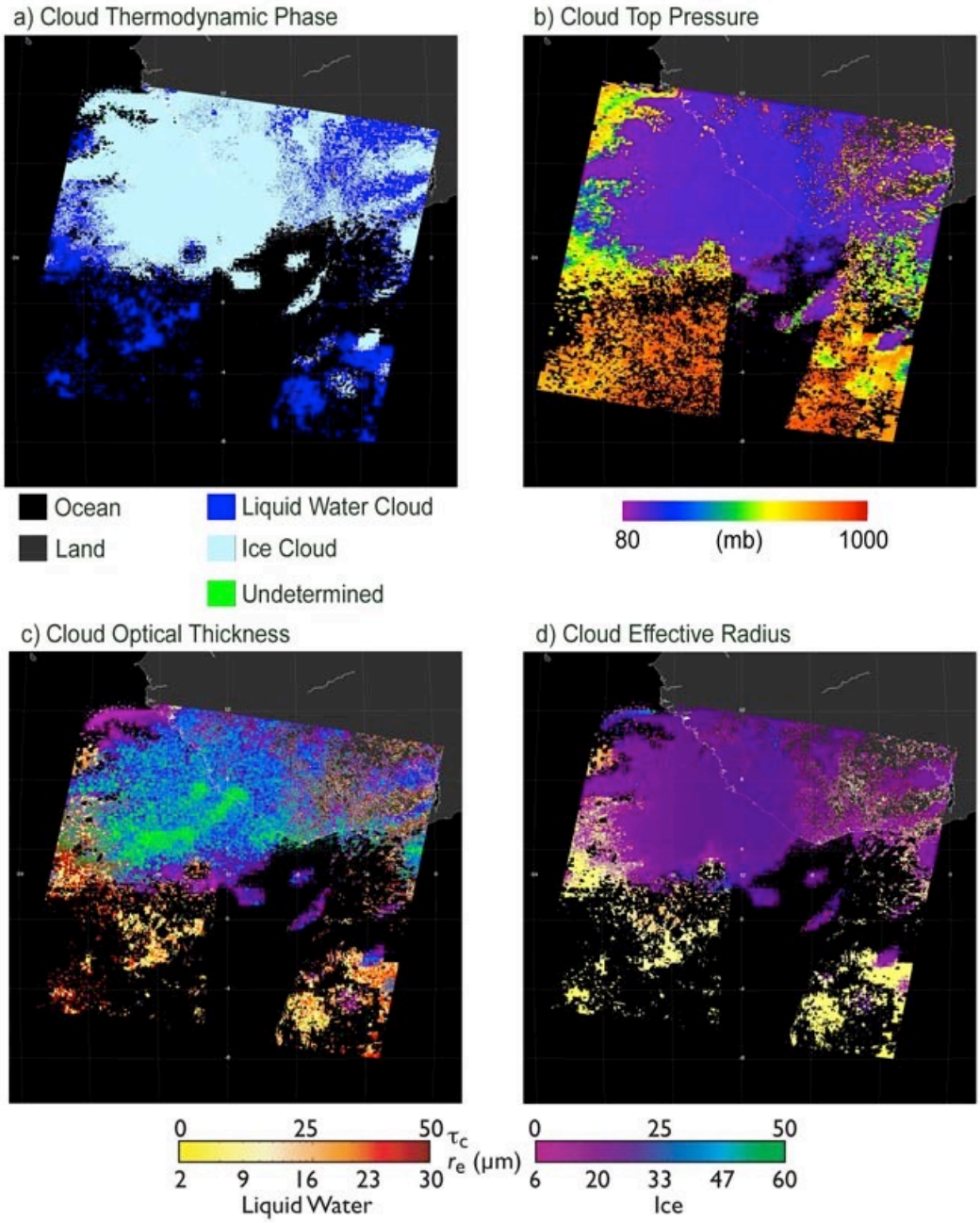

Figure 3-6. Example operational Data Collection 5 cloud top, optical and microphysical properties retrieval for synthetic granule covered by Terra MODIS 2013 day 151 11:15 UTC.

Figure 3-6 shows retrieval results for the simulated granule, figure 3-7 shows the actual Terra MODIS granule retrievals.

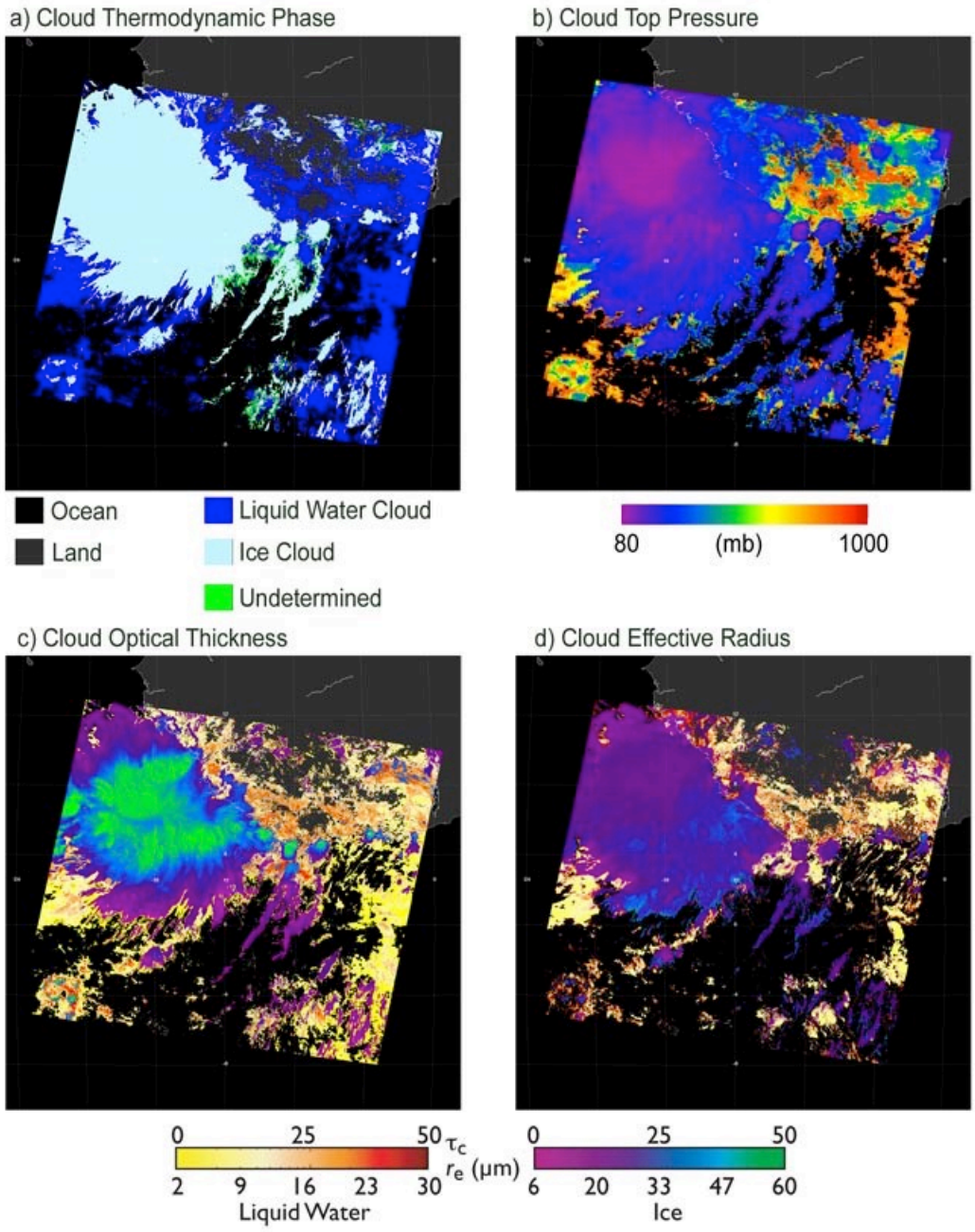

Figure 3-7. Actual operational Data Collection 5 cloud top, optical and microphysical properties retrieval for sensor-acquired granule Terra MODIS

2013 day 151 at 11:15 UTC .

Figure 3-8 shows the joint histograms of cloud top pressure and cloud optical thickness. In this case there is actually reasonable agreement between sensor measurement and model cloud field representation both geographically and vertically. The model could have benefitted from producing somewhat more mid-level clouds, but overall the large convective system that dominates the scene is represented reasonably well as are the broken clouds around it.

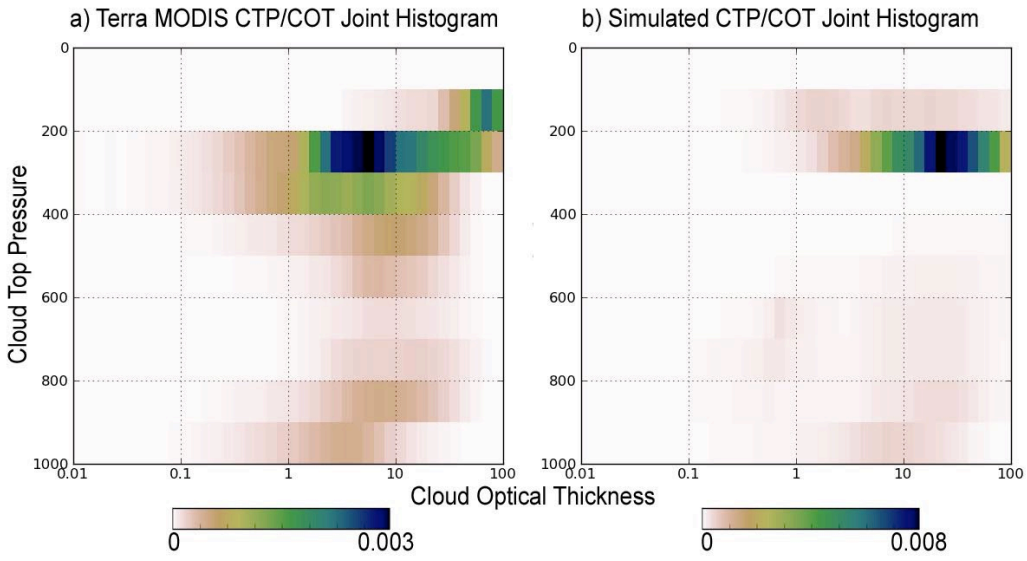

Figure 3-8. Joint histograms of cloud optical thickness vs. cloud top pressure for actual (a) and model-based (b) cloud fields covered by Terra MODIS 2013 day 151 at 11:15 UTC. In this case the model appears to place clouds at the altitude consistent with direct observation.

In this case the MODIS operational cloud mask had some trouble detecting clouds in the sun glint region, but sun glint quite often can be a challenging area for retrieval algorithms.

The MODIS cloud top pressure retrieval is quite sensitive to ancillary

atmospheric profile information (Menzel et al 2008) and some of differences found in retrievals could be a result of different representations of the atmospheric profile by GEOS-5 and the NCEP Mesoscale Meteorological Model (MM5)-derived model profiles used during the MODIS retrieval.

Situations in which cloud optical thickness retrievals show significant differences tend to be more indicative of significant differences in cloud structure. Unlike cloud top pressure, cloud optical thickness retrievals have very little dependence on atmospheric profile information as there is very little atmospheric absorption in the 0.65 and 0.86μm channels used to retrieve this quantity.

Cloud effective radius retrievals from the 2.1μm channel depend somewhat on the atmospheric profile, but differences in that retrieval are also mainly due to differences in cloud microphysics present in the model and in the actual atmosphere. Retrieved cloud effective radius appears to be somewhat smaller overall for GEOS-5 data than for MODIS. Even though the clumped-ICA cloud formation method allows us to model some of the scene inhomogeneity normally encountered in actual MODIS data, in the present implementation of GEOS-5 it would be highly unlikely to encounter effective radius artifacts such as ones appearing in MODIS (Zhang and Platnick, 2011). GEOS-5 uses a prescribed value of cloud effective radius in order to achieve a particular optical thickness and ice / water path that would be consistent with grid box atmospheric profile. In our preliminary results, there appears to be generally somewhat less variability in retrieved cloud effective radius from GEOS-5 than from real MODIS data.

3.4 Conclusions and future directions

We have evolved the simple simulator of chapter 2 and developed a flexible software suite that allows us to interface model fields to operational satellite remote sensing retrieval algorithms. We have presented an example of its operation using the GEOS-5 model and MODIS instrument. In the next chapter we will show the aerosol properties simulation extension and applications of this software. We would like to examine aerosol-cloud interactions by performing simulations with and without model aerosol fields. Once operational cloud and aerosol retrieval algorithms are applied to such data, we may be able to quantify some aerosol effects on clouds and maybe even find some ways to retrieve aerosols above clouds.

We are confident that there are many various applications for this software that will be found in the future besides ones outlined above and that it will become a valuable tool for both the remote sensing and modeling communities.

Chapter 4: Multi-sensor cloud and aerosol retrieval simulator

4.1 Introduction

Aerosols in the atmospheric column are a significant source of uncertainty for passive remote-sensing (e.g. from a satellite) retrievals of cloud optical and microphysical properties. Thick aerosol layers can be wrongly identified as clouds, and aerosols above clouds will lead to biases in cloud retrievals (Meyer et al. 2013). Biases in cloud detection and retrievals of cloud microphysics will lead to uncertainties in properties important for quantifying Earth's radiative budget. On the other hand, clouds wrongly identified and retrieved as aerosol may have similar impacts on estimates of aerosol radiative forcing and effects on climate and clouds. The Moderate-resolution Imaging Spectroradiometer (MODIS; Barnes et al. 1998) has been flying on the polar orbiting (at 705 km altitude) satellites known as Terra (since 2000) and Aqua (since 2002). Viewing a 2300 km swath, split into 5-minute granules, MODIS measures radiance (or reflectance) in 36 spectral channels, of which 19 are in reflective solar bands, with the other 17 being terrestrial infrared emission. All bands are in at least 1 km spatial resolution. Based on MODIS observations, separate teams have created high-quality retrievals of both cloud (e.g. the M{O/Y}D06_L2 (MxD06); Platnick et al, 2003) and aerosol (M{O/Y}D04_L2 (MxD04; Levy et al., 2013) properties. Current operational cloud retrieval includes methods for clearing the aerosols mis-identified as clouds from retrieval attempts.

(Zhang and Platnick 2011; Pincus et al. 2012). Similarly for aerosol retrievals, much effort is made to reclassify as "not cloudy" scenes that are in fact, heavy dust or smoke. Therefore, for both teams, uncertainty whether a particular sample is cloud-covered or contains primarily aerosols, and how to propagate this uncertainty into retrieval products, remains a topic of great interest. A major problem is that there is no absolute ground-truth to confirm or deny these decisions in all cases. Ground based instrumentation such as sun photometers (Holben et al 1998) may not be able to accurately distinguish between aerosol and thin clouds due to limited spectral range, generally reaching only up to a wavelength of 1.024μm. Newer sun photometers do provide information up to 1.64μm, but they are not present at every ground site. The ground sites in Brazil that fall within the area we studied in this paper carry the older instrumentation. The best wavelengths for detecting cirrus clouds are located around 1.38 and 1.8 μm. There are also efforts to retrieve aerosol optical depth above cloud layers (Meyer and Platnick 2015, Meyer et al. 2013). Validation for such algorithms is often done using lidar and radar data (Ackerman, et al. 2008, Notarnicola, et al. 2011). However as current space borne lidar and radar instruments have fixed nadir view, the amount of such data acquired in tandem with an instrument like MODIS is rather limited.

While a global meteorological model cannot be directly used to validate observations and retrievals due to the many assumptions and simplifications commonly made in the dynamic core and physics parameterizations (Rienecker et al. 2008), one can use such a model to learn about sensitivities of retrieval algorithms. As global models such as the Goddard Earth Observing System Model, Version 5

(GEOS-5; Rienecker et al. 2008, Molod et al. 2012), become increasingly realistic when simulating aerosols and clouds over complex surface terrain, we can apply detailed radiative transfer (RT) to simulate how these scenes would appear to a satellite such as MODIS, and how operational algorithms would in turn retrieve the specified conditions. Since the specified model aerosol and cloud properties of the scene are known, one can then characterize the ability (and uncertainties) of standard (e.g. MxD04 or MxD06) retrievals in these scenes. Thus, one can evaluate the current (and possibly historical) performance of cloud and aerosol properties retrievals. Application and evaluation of these simulation capabilities for known instruments is also an important step in development of Observing System Simulation Experiments for future observing missions.

The Multi-Sensor Cloud and Aerosol Retrieval Simulator (MCARS; Wind et al., 2013) is a modular, flexible tool, in which model output is coupled with a radiative transfer code in order to simulate Top Of Atmosphere (TOA) radiances that may be measured by a remote sensing instrument if it were passing over the model fields. In principle, MCARS can be applied to any model / visible-IR radiometer combination. The simulation complexity is only limited by computer power. However, in this paper, the MCARS continues to use the combination of GEOS-5 model and Discrete Ordinate Radiative Transfer (DISORT) code (Stamnes et al. 1988) to simulate MODIS radiances. In Wind et al. (2013), the MCARS simulated only clouds; here we add microphysical properties of aerosols present in scenes we examine.

The approach we take is to populate the operational MODIS Level 1B calibrated radiance files with TOA radiances simulated from GEOS-5 model output and

DISORT. For a given time and location, MODIS provides a particular geometry of observation. Since GEOS-5 simulates clouds and aerosols interactively, we can replace the MODIS-observed reflectance data with the simulated radiance product derived from the model. Then we run the standard aerosol (MxD04_L2) and cloud (MxD06_L2) retrieval codes and compare retrieval result to the known GEOS-5 source data. The discrepancies diagnosed by this device can then be contrasted to discrepancies obtained by comparing the real operational retrievals to independent, trusted observations (e.g., AOD from AErosol RObotic NETwork (AERONET)). To the extent that simulated and real statistical comparisons match, we can use capabilities of the MCARS code to examine the causes for such discrepancies, and hopefully identify opportunities for algorithm improvement. Since the aerosol retrieval is under-determined (Levy et al. 2013) and a number of assumptions must be made, the MCARS simulation approach is highly valuable as individual assumptions can be tested in isolation. The MCARS code has sufficient flexibility to test impacts of settings of single operational retrieval code parameters without interference from other components.

Section 4.2 describes the GEOS-5 aerosol properties and their addition into MCARS. Section 4.3 describes the MODIS aerosol product. Section 4.4 discusses case selection for the current analysis. It shows the selected scenes simulated by MCARS and describes other special simulation settings available that provide additional analysis capabilities. This section also presents analysis of retrieved aerosol properties as compared to the specified "ground" truth that served as input to

the simulations. Finally, section 4.5 discusses next steps in the continuing MCARS development.

4.2 GEOS-5 aerosol model and data assimilation systems

4.2.1 System description

Global aerosol, cloud, surface and atmospheric column fields from the GEOS-5 model and data assimilation system serve as the starting point for radiance simulations. The GEOS-5 system contains components for atmospheric circulation and composition (including aerosol and meteorological data assimilation), ocean circulation and biogeochemistry, and land surface processes. Components and individual parameterizations within components are coupled under the Earth System Modeling Framework (ESMF, Hill et al. 2004). This study is based on the near real-time (NRT) configuration of GEOS-5 where sea surface temperature and sea ice are specified from observations (Molod et al. 2012). The Goddard Chemistry Aerosol Radiation and Transport (GOCART, Colarco et al. 2010, Chin et al. 2002) bulk aerosol scheme is used in the GEOS-5 NRT aerosol forecasting system in this paper. A version of GOCART is run online and affects atmospheric radiative heating and budget in GEOS-5. GOCART treats the sources and sinks of dust, sulfate, sea salt and black

and organic carbon aerosols. Total mass of sulfate, and hydrophobic and hydrophilic modes of carbonaceous aerosols are tracked. Dust and sea salt have an explicit particle size distribution with five non-interacting size bins for each constituent.

Emission functions of both dust and sea salt depend on wind speed. Sulfate and carbonaceous species have contributions primarily from fossil fuel combustion, biomass burning, and biofuel consumption, with additional biogenic sources of organic carbon. Sulfate has additional chemical production from oxidation of SO_2 and dimethyl sulfide (DMS). We additionally include a database of volcanic SO_2 emissions and injection heights (Diehl et al. 2012). For all aerosol species, optical properties are obtained primarily from the commonly used Optical Properties of Aerosols and Clouds (OPAC) data set (Hess et al. 1998). We have recently updated our dust optical properties data set to incorporate non-spherical dust properties based on the work of Meng et al. (2010), Colarco et al. (2013) and Buchard et al. (2014). The aerosol transport is consistent with the underlying atmospheric dynamics and physical parameterizations (e.g., moist convection and turbulent mixing) of the model.

The GEOS-5 meteorological data assimilation is based on the Grid-point Statistical Interpolation (GSI) analysis scheme, jointly developed with National Oceanic and Atmospheric Administration National Center for Environmental Prediction (NOAA/NCEP) (Wu et al. 2002; Kleist et al. 2009). While the current GEOS-5 operational algorithm is based on a hybrid ensemble-variational scheme, the results reported here are based on the original 3D-Var implementation (Rienecker et al 2008). The aerosol re-analysis is produced at three-hour intervals, with assimilation of bias-corrected aerosol optical depth from MODIS, and has been evaluated against ground-base sun photometer measurements (Holben et al., 1998)

and against the Multi-angle Imaging Spectroradiometer (MISR) satellite instrument (Kahn et al, 2007).

4.2.2 Fire emissions

The fire emissions used in our simulations come from the Quick Fire Emission Dataset (QFED) Version 2.4 (Darmenov and da Silva, 2015). The QFED emissions are based on a top-down approach relating satellite retrieved Fire Radiative Power (FRP) at the top of the atmosphere to the amount of gases and particulate matter being emitted at the burning surface. The QFED emission factors are tuned so as to promote agreement among modeled and observed AOD. Another unique feature of QFED is how it handles areas obstructed by clouds when estimating grid-box mean emission rates. A sequential, minimum-variance algorithm keeps track of the fractional obscured area of given grid box. Emissions under the obscured area are then obtained by means of damped persistency model. Details can be found in Darmenov and da Silva (2015).

4.2.3 Case study selection

The WMO's Working Group on Numerical Experimentation (WGNE) has organized an exercise to evaluate the impact of aerosols on Numerical Weather Prediction (NWP) (Freitas et al. 2015.) This exercise involves testing of regional and global models currently used for weather forecasting by operational centers worldwide. The authors of this exercise selected 3 strong or persistent events of aerosol pollution worldwide that could be fairly represented by current NWP models. These cases were specifically selected to facilitate evaluation of the aerosol impact on

weather prediction. We chose one of the specified WGNE events: an extreme case of biomass burning smoke in Brazil, as the focus of this paper.

4.3 MODIS aerosol product

The MODIS "dark-target" (DT) aerosol product is described in detail in Levy, et al. (2013) and references therein. In this section we will give a brief overview of the DT algorithm as applied to MODIS observations.

The standard MODIS aerosol properties retrieval algorithm is a 10 km resolution product calculated from a detailed analysis of 10x10 boxes of 1km MODIS pixels. A necessary constraint for the algorithm is that the underlying surface is dark in visible and shortwave-IR wavelengths. There are two separate algorithm paths for ocean and land.

Pixels that are suspected to be cloudy or too bright or too noisy are discarded using conditions described in (Levy et al, 2007). Once the data sample is prepared, a spectral profile of average TOA spectral reflectance is created and compared against a lookup table. If a match is found, values for aerosol optical depth (AOD) and fine-mode aerosol weighting (FMW) are then returned.

In this paper we will focus on the land algorithm. Full description of the ocean algorithm can be found in Levy, et al (2013). Over land, even though there is greater variability of underlying surface than over ocean and thus greater uncertainty in retrieved aerosol properties, aerosol retrieval is still achievable. Over vegetated and dark-soiled surfaces, Kaufman et al. (1997) found that surface reflectance values for red (e.g. 0.65 μm) and blue (0.47 μm) wavelengths are correlated with the surface reflectance in a short-wave infrared (SWIR) band (e.g. 2.13 μm). The land algorithm

uses 0.47, 0.65 and 2.13μm channels for the main retrieval and 0.55, 0.86 and 1.24μm channels to give additional surface constraints.

The aerosol LUT is calculated for black surfaces and sea-level pressure. There are three fine particle model types and one coarse particle model type of aerosols used for dust based on climatology of AERONET inversion data (Dubovik et al, 2002). Each model type is multi-lognormal and is represented by size distribution, particle shape and complex refractive indices. The three fine-dominated models are differentiated primarily by single scattering albedo (SSA) in mid-visible wavelengths: urban/industrial type (SSA~0.95), near-source biomass burning (SSA~0.85) and a moderately absorbing type (SSA~0.90) to cover all other cases. For each aerosol type, the LUT includes TOA reflectance for a variety of angles and AOD referenced to 0.55μm.

Even with the constraints on surface reflectance, the aerosol retrieval does not have enough information to select between different aerosol types. Therefore, the relative proportion of fine-mode and coarse-mode aerosols must be prescribed so that, coupled with surface constraints, a best match can be found in the LUT for TOA spectral reflectance in the blue, red and SWIR wavelengths. The difference between TOA and nearest LUT reflectance is the fitting error.

With Levy et al., (2013) and previous studies, the primary validation of the MODIS product is by detailed co-location with ground-based sun photometer data, especially the Aerosol Robotic Network (AERONET; Holben et al., 1998). In this way, Levy et al., (2013) have defined the expected error (EE) envelope for the 0.55 μm AOD as ±(0.05 + 15%;). While spectral surface reflectance is also retrieved, it

does not tend to compare well with values obtained from the sun photometers. Note that the EE is defined upon mutually retrieved data. This means that satellite and sun photometer both observe enough clear-sky to retrieve AOD.

Also, while AERONET is well distributed about the globe, there are many situations for which MODIS retrieves aerosol, but there are no AERONET data available to compare with. Thus, there is no way to determine whether the MODIS aerosol retrieval has made reasonable choices, either for pixel selection, for cloud screening, or for aerosol model type and surface reflectance assumptions.

This motivates our use of the MCARS. Having full knowledge of underlying atmospheric, cloud, aerosol and surface parameters MCARS allows us to see deeper than AERONET would and over a much wider spatial area.

4.4 MCARS simulations

4.4.1 The MCARS software

We produced the simulation input data in accordance with the methods outlined in Wind et al. (2013). The GEOS-5 model output is split into 1-km subcolumns using the ICA method as described in detail in Wind et al. (2013). Here we give a brief summary of the model data preparation methodology.

Sampling of model cloud-related fields to the MODIS pixel scale is not straightforward because cloud properties typically vary on scales not adequately resolved by the operational 0.25° GEOS-5 resolution. To sample cloud fields, 1 km MODIS pixels for each GEOS-5 grid column are collected and the same number of

pixel-like sub-columns are generated using a statistical model of sub-grid column moisture variability. The general approach of Norris et al. (2008) is followed, namely using a parameterized probability density function (PDF) of total water content for each model layer and a Gaussian copula to correlate these PDFs in the vertical. Full details of the calculation of this PDF are described fully in Norris and da Silva (2016).

The subcolumns generated in this way are horizontally independent, but are subsequently "clumped," or rearranged, to give horizontal spatial coherence, by using a horizontal Gaussian copula applied to condensed water path. This clumping acts to give the generated clouds a reasonable horizontal structure, such that the cloudy pixels in a grid column are actually grouped into reasonable looking clouds, rather than being randomly distributed. This is important because the MODIS cloud optical and microphysical properties retrieval algorithm has some spatial variance tests for potentially partially-cloudy pixels, removing cloud edges by the so-called "clear-sky restoral" (Zhang and Platnick 2011; Pincus et al. 2012). If clumping is not used, then individual points generated by ICA stand an exceptionally high chance of being eliminated by the clear sky restoral unless a model grid box has a nearly 100% cloud fraction.

The layer aerosol properties are obtained using the independent column approximation with the same PDF of total water content as used for clouds. The MCARS code uses a species file, produced from the GEOS-5 model output, which for each simulated MODIS pixel gives individual aerosol optical depths by aerosol type. The OPAC database (Hess et al, 1998) is then queried in order to obtain the

aerosol phase function for each of the 15 aerosol species and the properties such as single-scattering albedo are then augmented by profile of subcolumn relative humidity. The result of this query is a set of Legendre coefficients and a single-scattering albedo that correspond to the combined effect of all 15 aerosol species.

Model parameters such as profiles of temperature, pressure, ozone and water vapor together with layer information about clouds (and now aerosols) are combined with solar and view geometry of the MODIS instrument. Surface information is also a combination of GEOS-5 information of surface temperature, snow and sea ice cover and MODIS-derived spectral surface albedo (Moody et al. 2007, 2008). All these parameters are transferred to the DISORT-5 radiative transfer code and reflectances and radiances in 24 MODIS channels are produced. They are output into a standard MODIS L1B file that corresponds to the source MODIS geolocation file we used to sample the model output with. All metadata is preserved in this process and so the MCARS output is indistinguishable from a real MODIS granule except in how it may appear to the user's eye. These synthetic reflectances and radiances are completely transparent to any operational or research-level retrieval algorithm code and can be used for any purpose that real sensor data can.

In order to produce these simulations we use the NASA Center for Climate Simulations (NCCS) supercomputer Discover. It takes 5.5 hours of wall clock time on 144 processors to produce one complete simulation. The performance can be improved if the user limits the simulation scope to fit a particular investigation they are working on. For example, an aerosol researcher would not likely need to simulate the MODIS channels that they would not use and thus reduce execution time by at

least half. Because these simulations are simultaneously used for both cloud and aerosol work, we simulate all the channels that would be used by both cloud and aerosol disciplines.

4.4.2 Granule selection

In order to perform tests of the MCARS aerosol module we have selected Aqua MODIS granules from time period corresponding to WGNE selection for biomass burning in Brazil. In this chapter we specifically present results from simulations based on two granules of smoke in Brazil 2012 day 252 17:30 UTC and day 254 17:20 UTC subsequently referred as “Brazil 1” and “Brazil 2”.

4.5 Analysis

For each granule, we ran the simulations in several modes with varied run-time option settings. For example, the cloud-only mode corresponds to a clean atmosphere with no aerosols; this mode was the only one considered in Wind et al. (2013). In the current paper we consider additional options afforded by the implementation of the aerosol effect. The cloud-free option runs atmosphere and aerosols without any clouds. When clouds are turned off, we do not alter the humidity profiles to dry the atmosphere out; because of the high relative humidity conditions where clouds are present, aerosol hygroscopic effects are pronounced there as well. The full simulation option includes atmosphere (temperature, humidity and ozone profiles), all clouds and all aerosols. There is also an additional option where the user can remove both clouds and aerosols and be left with just the atmosphere itself. Rayleigh scattering is always included by default but user also has control over whether or not to turn it off. While

this no-cloud/no-aerosol mode could be useful for studies of atmospheric correction methods, we do not exercise it here, as our primary goal here is to investigate the performance of the MODIS aerosol algorithms.

The cloud-free mode of operation is convenient when complex cloud and aerosol scenes are being investigated and one wishes to quantify or remove possible impacts of cloud contamination on the retrieval.

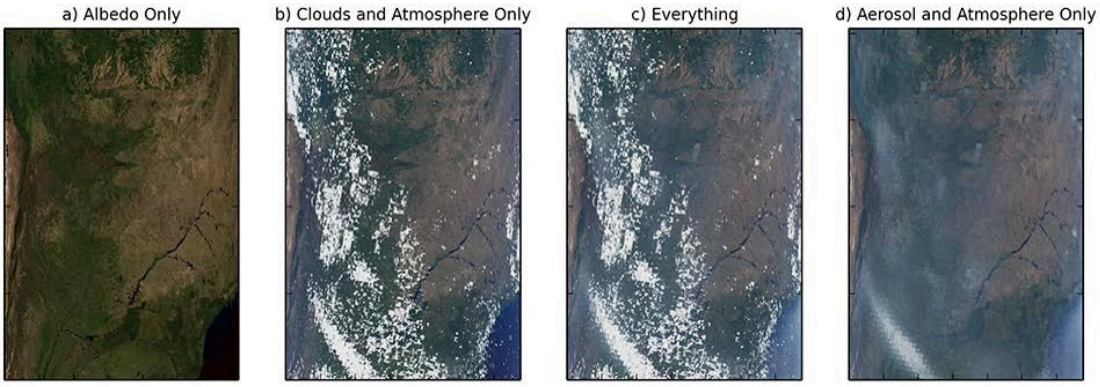

Figure 4-1. Example of various execution modes of the MCARS code using the "Brazil 1" case 2012 day 252 17:30UTC. Panel a) shows the albedo-only run with no atmospheric constituents of any kind. Panel b) shows the run with clouds and atmosphere included, but no aerosols. Panel c) shows the run with all available atmospheric constituents present and panel d) has radiative contributions from the atmosphere and aerosols only.

Figure 4-1 shows RGB images constructed from simulated MODIS L1B for the different modes of execution for the "Brazil 1" case. MODIS aerosol retrievals were produced for radiance simulations including atmosphere, cloud and aerosols (Figure 4-1c) and for radiance simulations excluding clouds (Figure 4-1d). Rayleigh scattering is included in these simulations.

These different execution modes can also be highly useful for looking at aerosol direct radiative effects when "clean" run without any aerosols is compared to data for the same scene produced with aerosols present.

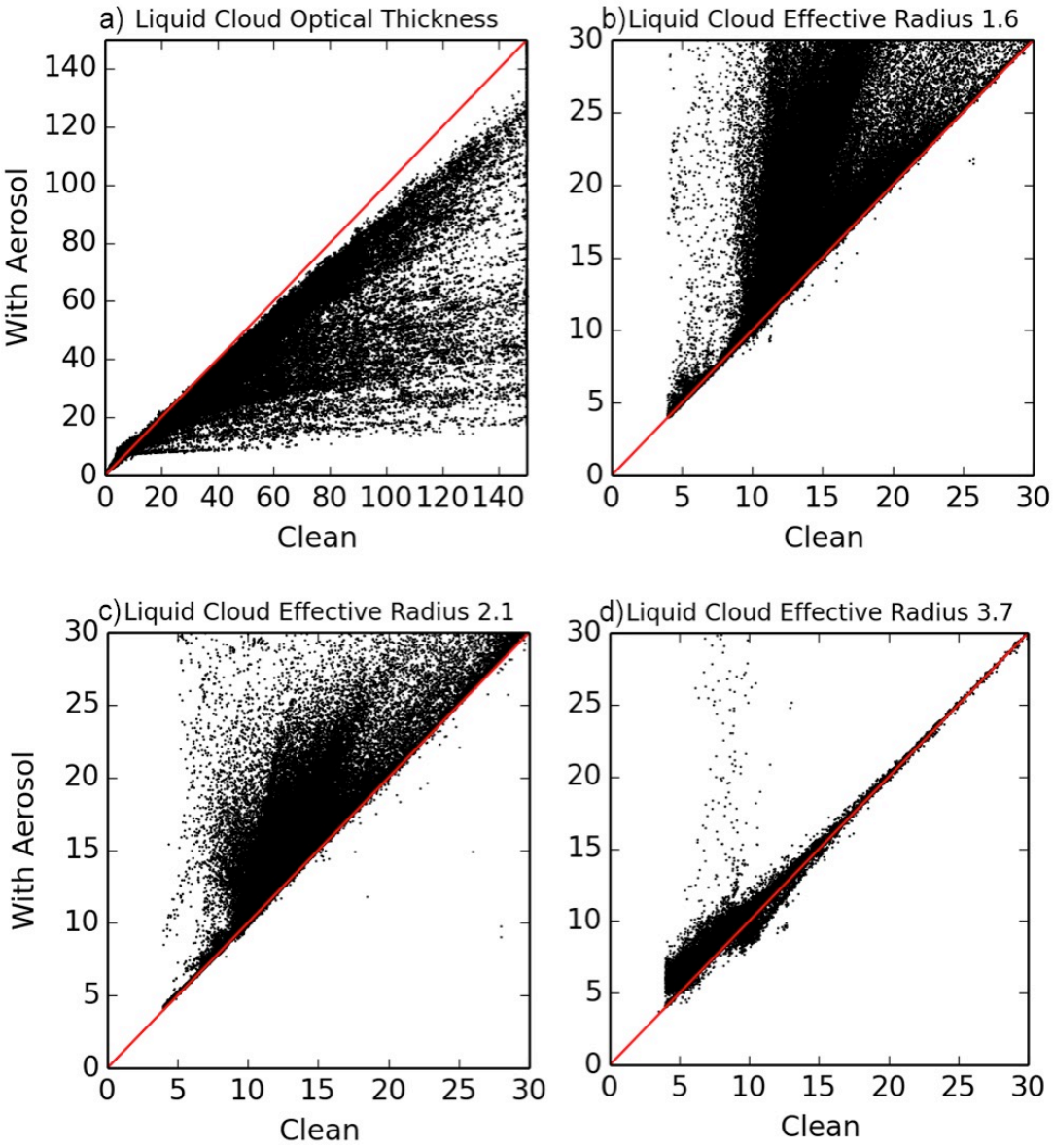

Figure 4-2. A preliminary example of using the different MCARS run modes of "Brazil 1" case in Figure 4-1 to illustrate effects of above-cloud aerosols on cloud optical and microphysical properties retrievals. Panel a) shows the impact of above-cloud aerosols on retrieved cloud optical thickness. Panels b), c) and d) show the impacts on retrievals of cloud effective radius using combination of 0.65µm channel and the different SWIR channels of 1.6, 2.1 and 3.7µm respectively.

Figure 4-2 shows a scatterplot of retrieved cloud optical thickness and cloud effective radius in different SWIR channels. It is generally thought, but not directly

quantified that aerosol impact is greater at shorter wavelengths. The panels of figure 4.2 show that it indeed is the case. The deviation from the 1/1 line is greatest for the 1.6μm retrieval of cloud effective radius in the presence of above-cloud aerosols and almost completely disappears by the time 3.7μm is reached. We have just started this particular study and we expect to have many more cases relating to the upcoming field campaign in the Southern Atlantic ocean. These comparisons of retrievals of cloud optical and microphysical properties indicate that in the presence of above-cloud absorbing aerosols, retrieved cloud optical thickness would be biased low and cloud effective radius would be biased high. This kind of comparison, with more cases included, could be used statistically to create a parameterization that could allow removal of cloud retrieval biases due to absorbing aerosols above clouds.

These Brazil cases came from source MODIS Aqua granules and had been processed using the MODIS Aqua aerosol properties retrieval algorithm. Therefore in this section we will use MYD04 designation for the MODIS aerosol properties retrieval result. There are some slight differences between the MODIS Terra (MOD04) and MODIS Aqua (MYD04) algorithms due to calibration differences between the two instruments (Levy et al, 2013).

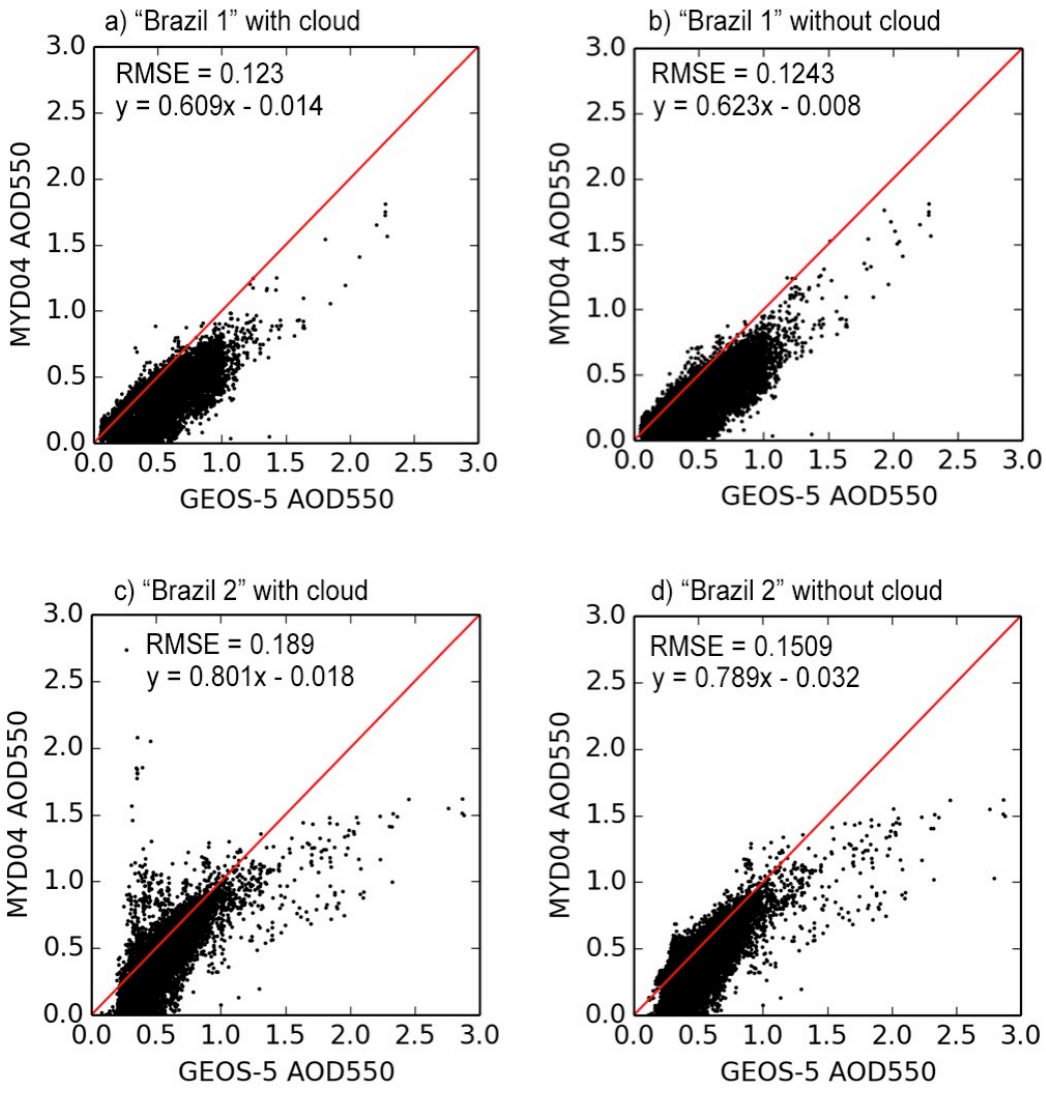

Figure 4-3. MYD04 retrieval vs. ground "truth" of GEOS-5 aerosol optical depth. Presence of clouds does not significantly impact the aerosol optical depth retrieval. In these cases the operational MYD04 cloud mask appears to function optimally in removing cloud contamination from radiances selected for retrieval.

The scatter diagrams in Figure 4-3 compare AOD retrieved using the MYD04 algorithm to the specified GEOS-5 AOD, which is considered the ground truth in this case. MODIS aerosol retrievals are commonly compared to co-located AERONET AOD measurements (Correia and Pires 2006, Levy, et al. 2007, Remer et al. 2005) for validation. Unlike comparisons of actual MODIS data with AERONET, the match

ups in Figure 4-3 did not require any temporal averaging or aggregation because for every MYD04 retrieval there is a directly corresponding input data point with all aerosol, cloud and atmospheric properties readily available. The overall shape of resulting scatter plots turned out to be quite similar to existing MYD04 – AERONET comparisons for this region such as those that appear in Correia and Pires (2006) and Figure 4-4.

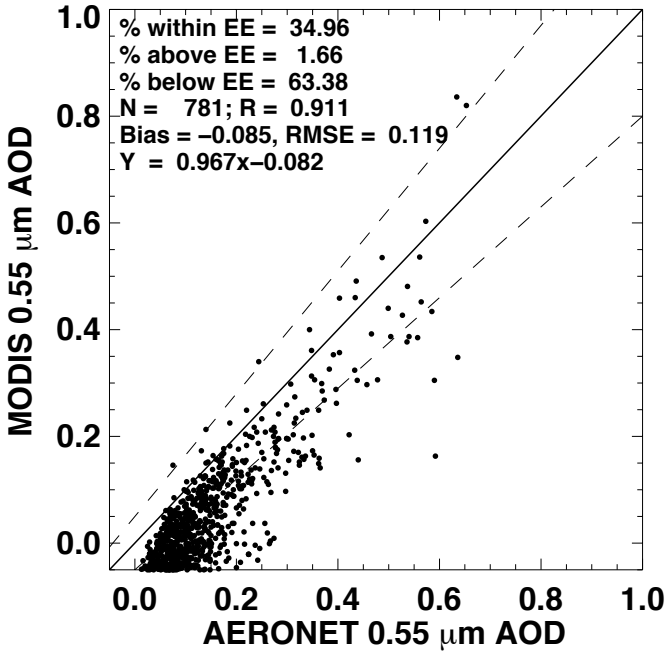

Figure 4-4. Comparison of actual AERONET measurements and operational Aqua MODIS Collection 6 aerosol product for Brazil sites Campo_Grande_SONDA, Sao_Paulo and CUIABA-MIRANDA in the general area of MCARS granules.

Figure 4-4 shows an actual comparison for AERONET observations for months of July and August and all available Aqua MODIS collocated observations from year 2002 through 2015. The chosen AERONET sites: Campo_Grande_SONDA, Sao_Paulo and CUIABA-MIRANDA fall in the general area of the two Brazil cases selected for study. They of course represent a tiny sample of the geographical area

covered by the MCARS data, just three points out of 2.7 million collocated samples that MCARS provides, but they display a similar shape of the relationship between ground truth and MYD04 retrieval.

MCARS is a fully configurable system where source input for all synthetic radiances can be controlled at all times, so that any resulting retrieval can be examined in great detail insofar as the particular setup of model input and radiative transfer core allows. For these smoke cases we used these capabilities to investigate further the specific reasons why the MYD04 retrievals tend to underestimate AOD for smoke aerosol.

The first test we made was to examine the performance of MYD04 cloud mask, which is an aerosol specific product (Remer et al, 2005), different from the operational MODIS cloud mask product (Ackerman et al, 2006). The main purpose of this analysis was to ascertain whether cloud contamination could account for some of the discrepancies. Individual panels in Figure 4-3 show the results of retrievals run with and without the cloud layers. Panels a) and b) show result for “Brazil 1” and panels c) and d) are for “Brazil 2”. “Brazil 1” case does not show any significant cloud contamination. The MYD04 cloud mask does a very good job of avoiding cloud. “Brazil 2” does show some very minor cloud contamination as evident by a small cluster of high MYD04 AOD and low GEOS-5 AOD that disappears when clouds are removed from simulation. However the overall shape of the scatter plot when clouds are removed remains unchanged.

The aerosol models used in the MYD04 retrievals make assumptions about the smoke aerosol optical properties, which may not match the aerosol optical

assumptions in GEOS-5 (Levy et al, 2007). In cases of complex aerosol mixtures or if the model selected by the MYD04 algorithm does not correspond to the aerosols provided by GEOS-5, large retrieval errors should result.

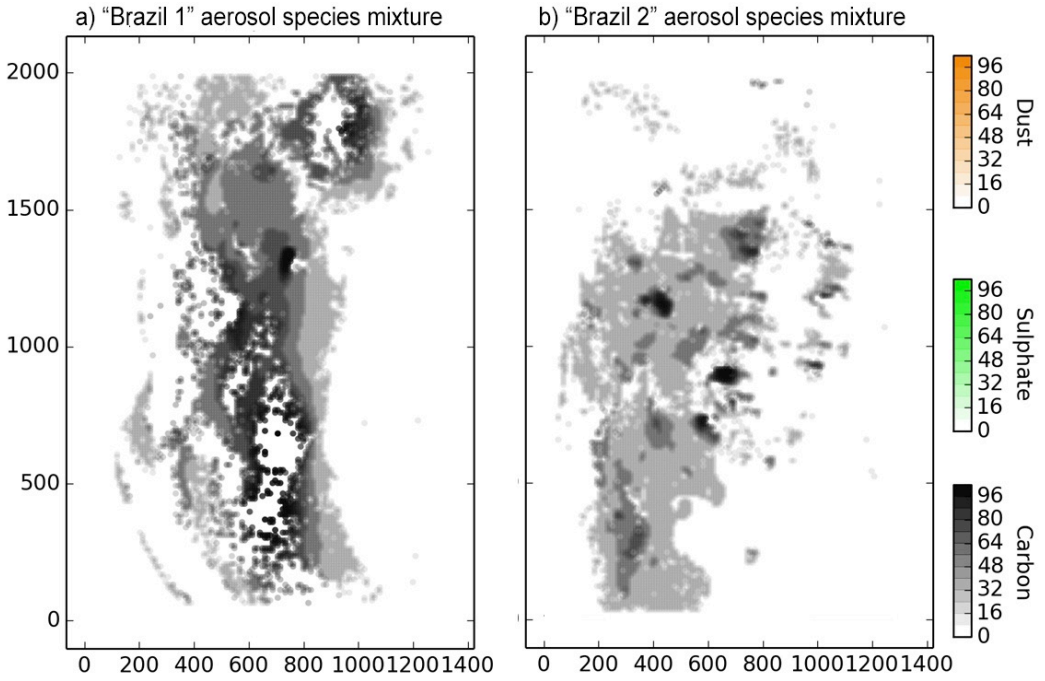

Figure 4-5. GEOS-5 aerosol species mixture for attempted MYD04 retrievals in figure 4-3. The area is fully dominated by absorbing carbon from biomass burning. The other constituents are also present, but in such negligible quantities that they are not visible on this image.

Figure 4-5 shows the species mixture for "Brazil 1" (a) and "Brazil 2" (b) cases. They are both dominated by carbon, organic carbon from smoke in particular, with very little, if any contribution from other species. Therefore these particular cases can be treated as having a single aerosol type present without significant error. MYD04 retrieval output indicates that either moderately or strongly absorbing smoke had been selected, which is very appropriate for the selected granules. Thus any discrepancy in selection of aerosol model does not explain the scatter plot shape.

Another candidate source of retrieval error is any difference between the phase functions assumed by MYD04 and GEOS-5. We ran the initial simulations simply using the Henyey-Greenstein (HG) phase function approximation and then repeated the same simulation using the phase functions provided by the OPAC database described in section 4-2.

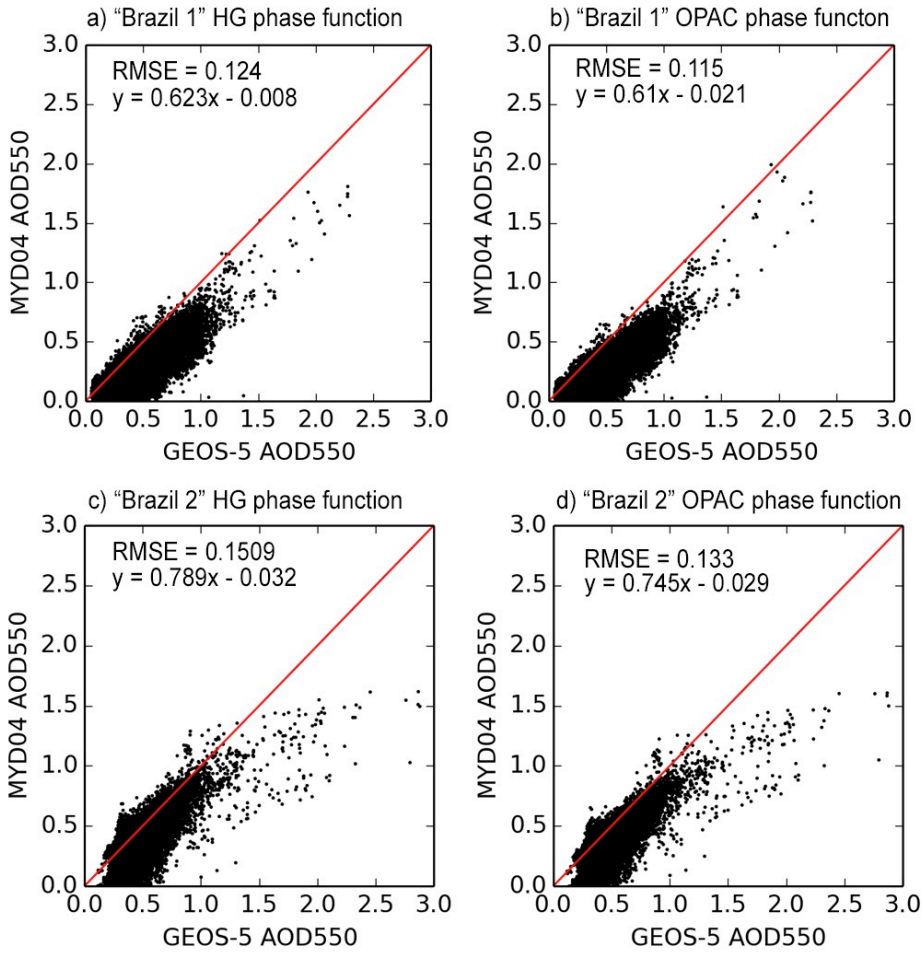

Figure 4-6. Effect of aerosol phase function shape on Brazil smoke cases. Simulations in panels a) and c) have been executed with the simple Henyey-Greenstein aerosol phase function. Panels b) and d) are using the actual OPAC phase function for smoke. As expected there is no significant difference between the two sets of simulations for either case.

Figure 4-6 shows the result for "Brazil 1" and "Brazil 2" cases using the cloud-free run with HG phase function versus OPAC phase function. For the smoke aerosol cases studied, the specific phase function shape does not appear to have a significant impact on the differences seen between MYD04 and GEOS-5.

An additional potential source of error for aerosol retrievals over land is the surface albedo and its variation over a 10x10 km area. We performed a simulation where we selected a single surface albedo profile from a successful MYD04 retrieval and fixed the surface albedo to that particular surface albedo profile for the entire granule. The test albedo profile used is listed in Table 4-1.

MODIS channel	Central Wavelength (μm)	Surface Albedo
1	0.65	0.027
2	0.86	0.288
3	0.47	0.017
4	0.55	0.037
5	1.24	0.252
6	1.63	0.146
7	2.13	0.054
8	0.41	0.014
9	0.44	0.022
17	0.91	0.283
18	0.94	0.280
19	0.94	0.280
20	3.7	0.038

22	3.9	0.038
26	1.38	0.216

Table 4-1. Constant surface albedo setting used in smoke AOD retrieval investigation

The profile corresponds to a very dark vegetated surface, the ideal conditions for the MYD04 land algorithm.

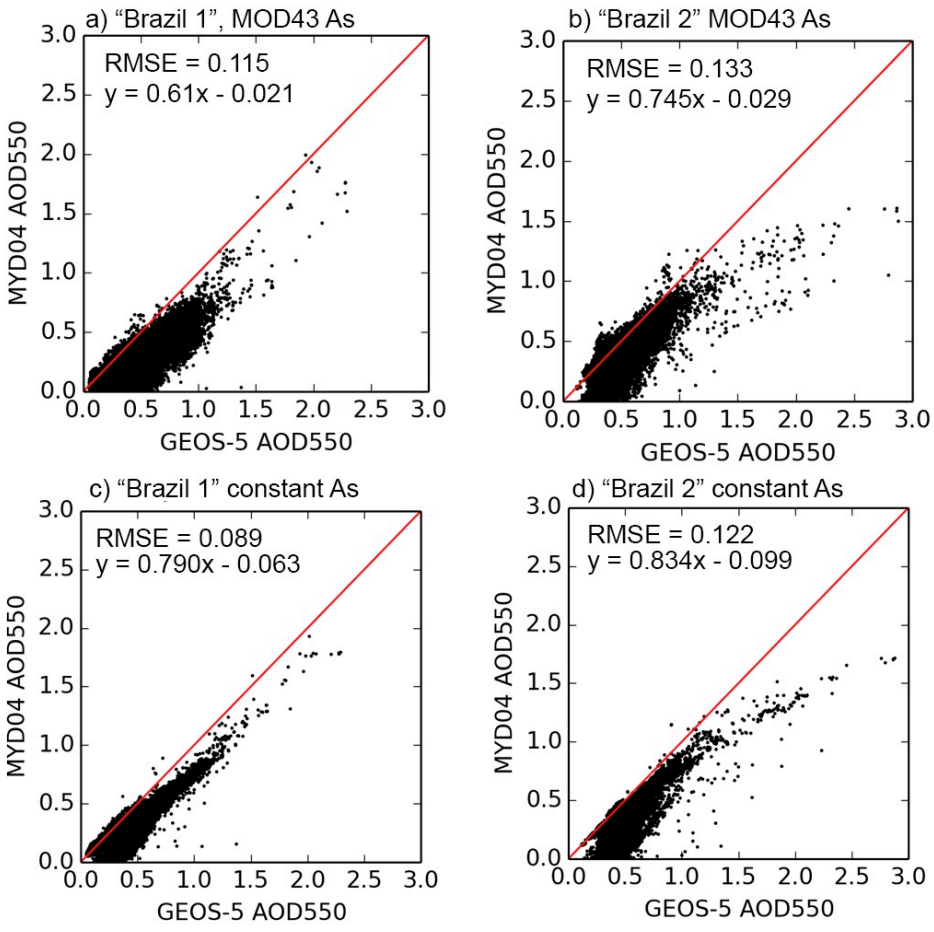

Figure 4-7. Surface albedo effect on Brazil smoke cases. Panels a) and b) show

the simulations performed using the 1-km resolution MOD43 surface albedo product from Boston University. Panels c) and d) have been run with a single, constant spectral albedo profile applied to every pixel.

Figure 4-7 shows the effect of using a constant surface albedo for "Brazil 1" and "Brazil 2" cases. Whereas use of constant surface albedo reduces the scatterplot spread and so allows us to potentially quantify the effect of surface inhomogeneity on MYD04 land retrievals, it does not alter the overall bias characteristics of scatter plots.

With all the factors of model selection, surface parameters and cloud contamination taken into account, we now turn our attention to the aerosol scattering properties, the spectral single scattering albedo (SSA) in particular. Figures 4-8 and 4-9 show the spectral profile of aerosol SSA for "Brazil 1" and "Brazil 2" cases respectively for the first seven MODIS channels. This aerosol SSA is a bulk quantity, integrated over all layers and combines all 15 available aerosol species. However the cases under consideration are heavily dominated by carbon with negligible amounts of dust and sulfate. In this particular case the additional uncertainties that would arise from a mixture of aerosols with different scattering properties do not present an issue. The single scattering albedo remains quite high until we reach the 1.2μm channel, MODIS band 5, and beyond.

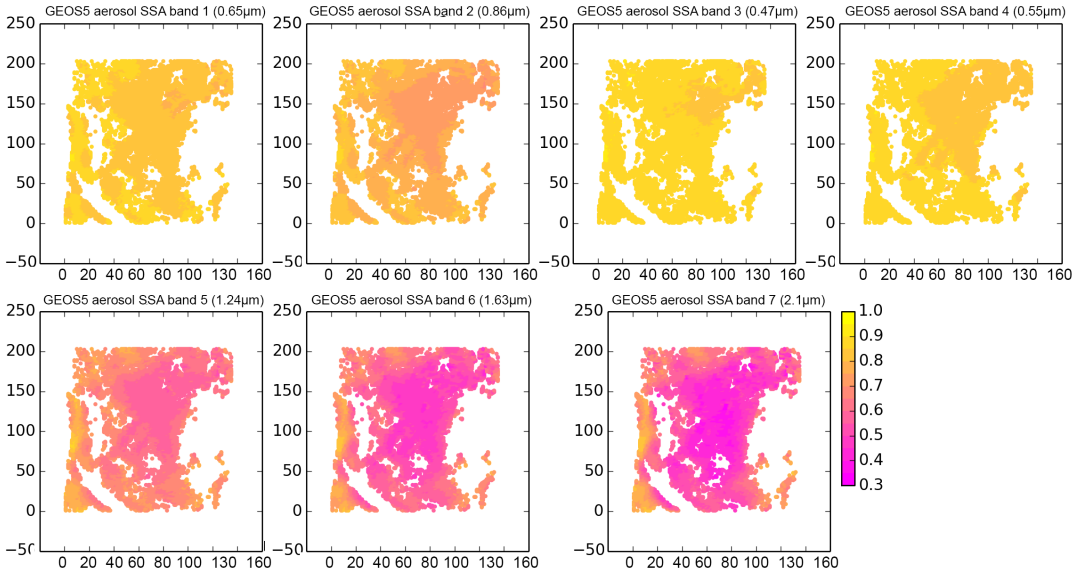

Figure 4-8. Aerosol single scattering albedo for "Brazil 1" case for MODIS channels 1-7. The single scattering albedo remains fairly constant across the visible and NIR channels, but drops precipitously in the SWIR bands.

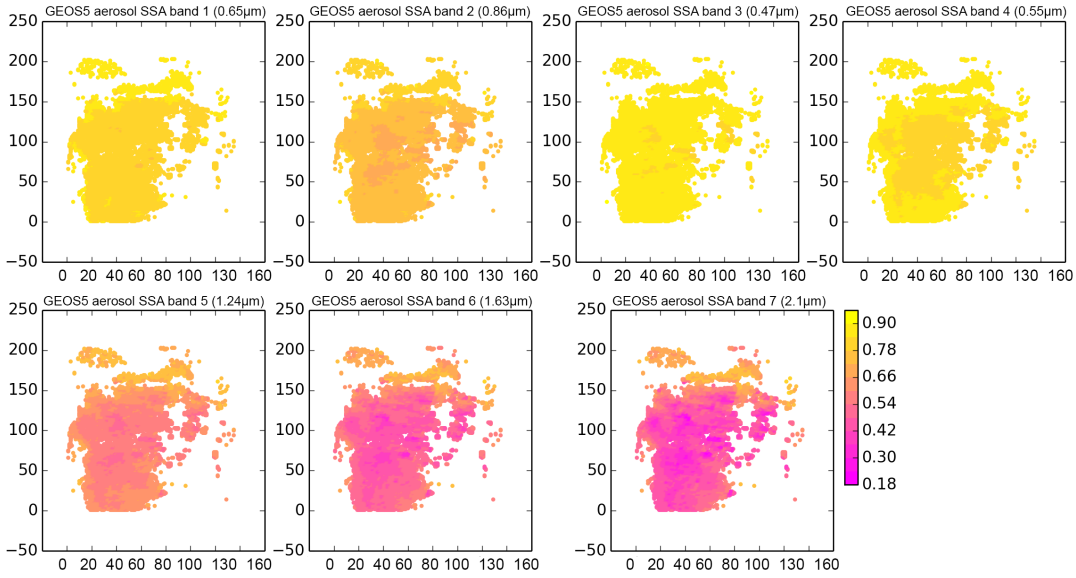

Figure 4-9. Aerosol single scattering albedo for "Brazil 2" case for MODIS channels 1-7. The single scattering albedo remains fairly constant across the visible and NIR channels, but drops precipitously in the SWIR bands.

Then it drops precipitously. AERONET is only able to provide direct inversion retrievals of single scattering albedo for four wavelengths out to a maximum

wavelength of 1.024 μm (Dubovik and King, 2000; Dubovik et al., 2002).. The rapid change in single scattering albedo for smoke aerosol modeled in GEOS-5 is related to aerosol humidification effects, both dilution effects and hygroscopic growth (Colarco et al. 2010, 2013). The net effect is that when humidity decreases, so does the single scattering albedo.

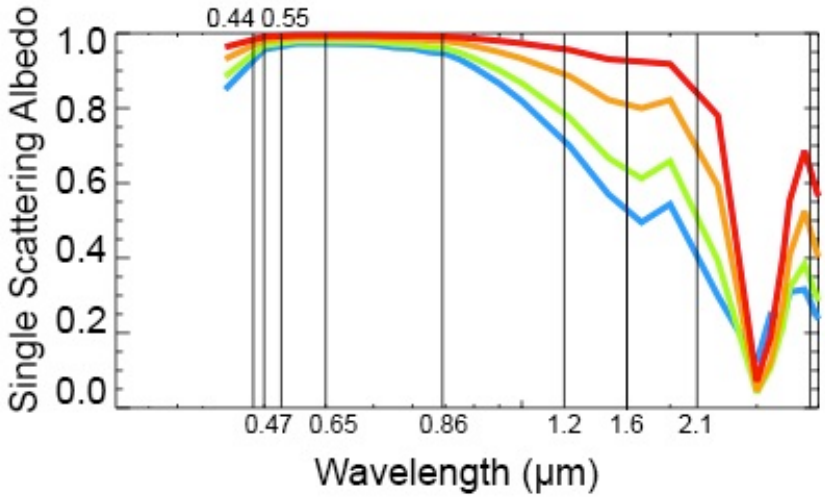

Figure 4-10. OPAC single scattering albedo as a function of humidity (color) and wavelength. The various relative humidity levels are in order (red, orange, green and blue) for 95, 80, 30 and 0% column relative humidity.

Figure 4-10 shows a plot of OPAC single scattering albedo for a variety of column relative humidity values as a function of wavelength. (Colarco, et al 2013) The operational MODIS aerosol code assumes a constant 80% relative humidity when the lookup tables are generated (Levy et al, 2007). It is a reasonable assumption as long as one does not attempt to use channels with wavelengths that are longer than 0.8μm. The MYD04 algorithm however does use the 2.1μm MODIS channel in retrieval, a channel that is sensitive to humidity. MCARS is particularly well suited to

test for humidity impact on the retrieval accuracy. We made another experiment with fixed surface albedo, OPAC aerosol phase function shape but we used the constant single scattering albedo values from the MODIS aerosol algorithm in the reflectance calculation that serves as input to the retrieval algorithm. The result is shown in figure 4-10.

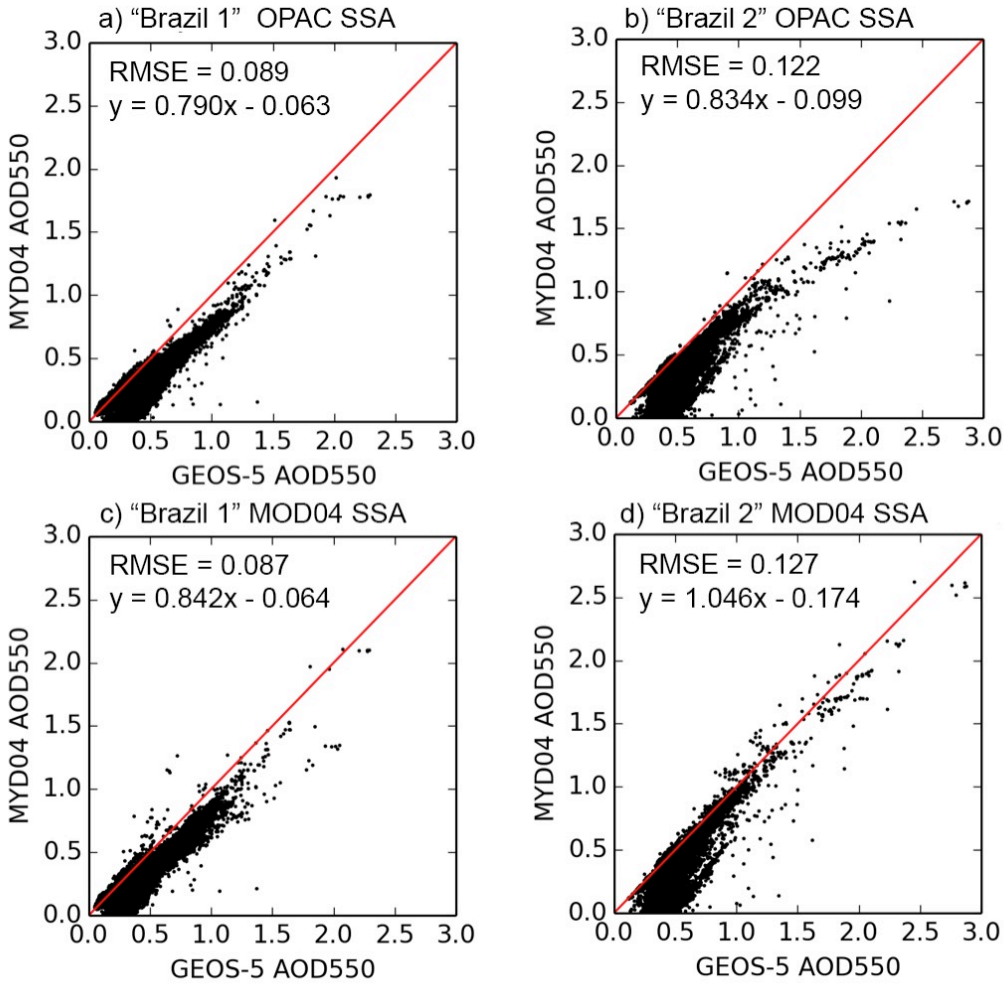

Figure 4-11. Impact of humidity on MOD04 retrieval illustrated via single scattering albedo selection. When humidity's effect on smoke aerosol single scattering albedo is properly considered, the low bias at higher AOD virtually disappears.

When humidification effects are not taken in consideration in the SSA

calculation, MYD04 retrieval results closely line up with synthetic GEOS-5 source data. The underestimate of aerosol optical depth at higher end of the scale disappears, with “Brazil 2” showing the most dramatic improvement. It appears that if MYD04 were to take into account humidification effects and implement a correction for single scattering albedo value as a function of column relative humidity, the result of comparison between MODIS and AERONET could be significantly improved for biomass burning cases in Brazil and other locations with similar synoptic conditions.

The improvement is limited however to AOD higher than about 0.5. Relative humidity does not appear to have an effect on retrieved low AOD values. MYD04 product does not provide pixel-level retrieval uncertainty estimates. It is possible that the inherent uncertainty in performing retrieval using such small signal is so high that it drowns out other effects. More studies may be conducted as to attempt to create a pixel-level estimate of retrieval uncertainty for aerosol optical properties retrievals.

The MODIS aerosol product performs a simultaneous retrieval of land surface reflectance and aerosol optical depth. After looking at the behavior of aerosol optical depth and making a recommendation for a possible improvement in the retrieval algorithm, we examined the retrieval of land surface reflectance. The MODIS aerosol product provides a simultaneous retrieval of aerosol optical depth at 0.55μm and land surface reflectance in the 0.47, 0.65 and 2.1μm channels. We looked at the land surface reflectance for the simulation of figure 4-11 panels c) and d) that now matched the source aerosol optical depth reasonably well. The simulation was run under constant surface albedo conditions and we would have expected to see a result, with some degree of uncertainty of course, that would match the given constant

surface albedo. However the retrieved land surface reflectance appeared to be a near-linear function of aerosol optical depth. The humidity effect does not seem to have an impact on this apparent linear relationship. The relationship persisted through different adjustments of single scattering albedo. One possible explanation for this behavior may involve the assumed fraction of coarse-mode aerosol in the aerosol model mixture. To examine this hypothesis we performed a MYD04 retrieval using an aerosol model setting so that MYD04 retrieval only used fine mode particles.

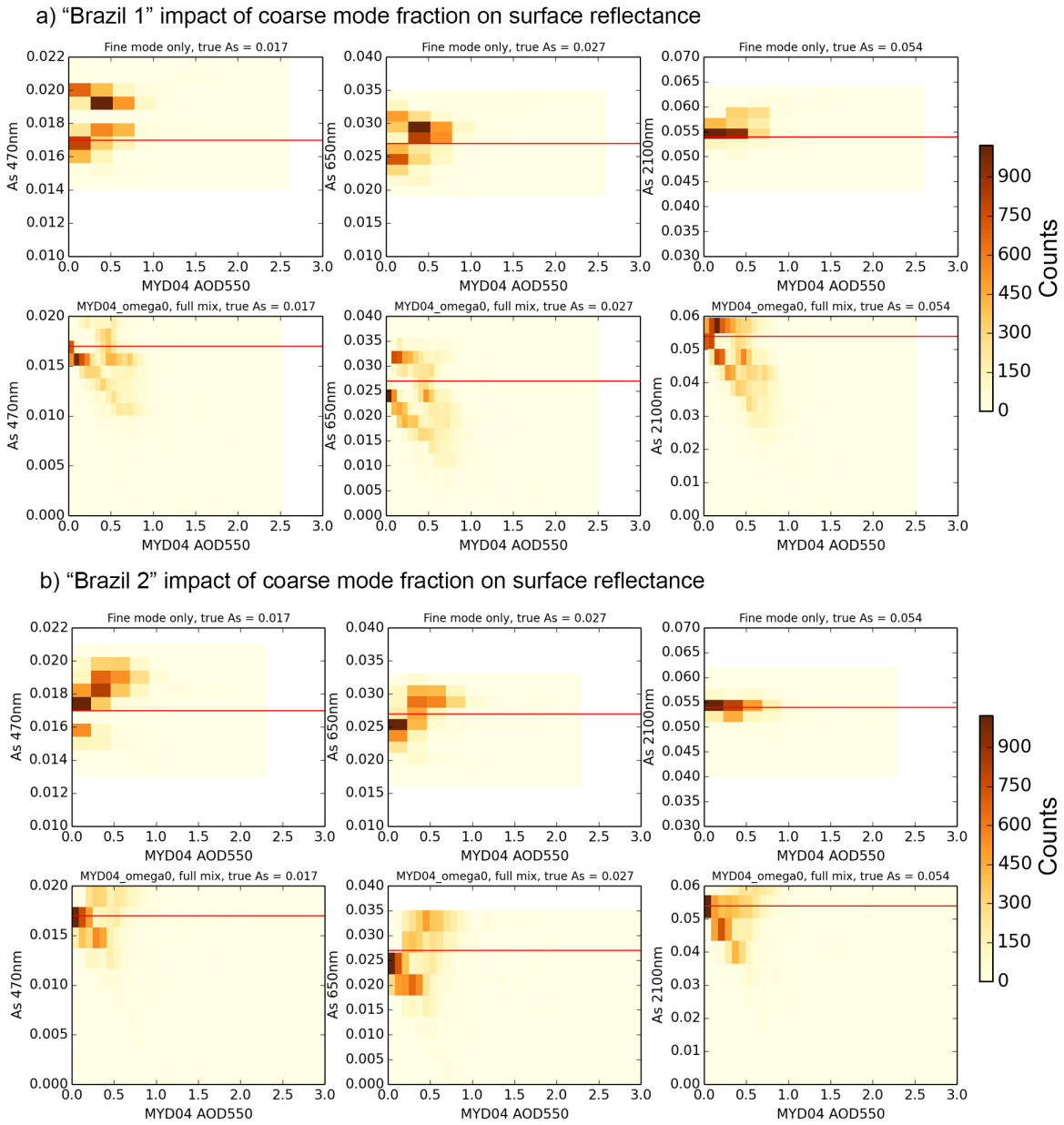

Figure 4-12. Impact of the value of fine mode fraction on MOD04-retrieved land surface reflectance. When the fine mode fraction is set to unity the apparent linear relationship between aerosol optical depth and land surface reflectance disappears and retrievals cluster properly around the prescribed constant spectral surface albedo values.

The retrieval results, depicted in figure 4-12 confirm that the near co-linearity of surface reflectance and AOD was indeed directly related to fraction of coarse mode particles, such as dust, in the assumed aerosol mixture. Of course there is no way to

know exactly what fraction of coarse mode particles may be present in the mixture as the MODIS DT algorithm does not have enough information content to constraint the fine/coarse mode fraction over land (Levy et al, 2007). However, it can be noted that if such co-linearity is seen during a specific local aerosol study maybe during a field campaign, it may be suggested that the coarse mode fraction assumed operationally for that particular region may be too high. An analysis of MODIS operational retrievals to identify locations and times where this co-linearity exists may be useful to identify regions where the assumed coarse/fine mode fraction might need to be adjusted. Figure 4-12 illustrates the impact of coarse-mode fraction selection on land surface reflectance retrievals for "Brazil 1" and "Brazil 2" cases. The fine-to-coarse mode ratio does not appear to have an impact on the low bias of MYD04 AOD retrieval vs. "ground truth" comparisons presented in the earlier figures.

4.6 Conclusions and future directions

This chapter is a continuation of work started in chapter 3. The multi-sensor cloud retrieval simulator code (MCRS) had been extended to add aerosol effects to radiance simulations. The current implementation of the MCARS code generates synthetic radiances by sending GEOS-5 model fields and MODIS sensor geometry and location information to the DISORT-5 radiative transfer core. The radiance and reflectance data are output in a standard MODIS Level 1B format that can be transparently ingested by any retrieval or analysis code that reads data from the MODIS instrument.

After the aerosol properties module had been added to the MCARS code we used the simulator to perform detailed analysis of performance of the operational MODIS

dark target aerosol properties retrieval product for the Aqua MODIS instrument (MYD04). We found the cause of known low bias in MYD04 retrieved AOD for smoke when compared to in-situ measurements. We suggest that the MYD04 retrieval might consider using column relative humidity from ancillary data when performing retrievals in regions that are defined to be dominated by smoke aerosols. The mismatch between the aerosol single scattering albedo assumed by MYD04 and the given synthetic single scattering albedo is the cause of the low bias at higher AODs. The impact of surface inhomogeneity is also quantifiable. Whereas it may not be possible to make an operationally actionable item from retrieval behavior when surface is made homogeneous, it may be possible to deduce an estimate of retrieval uncertainty due to land surface effects.

This study is a good example of capabilities of the MCARS code. We are planning many more studies of retrieval algorithm performance.

The MCARS results give a relationship between aerosol single scattering albedo, bias in retrieved aerosol optical depth and column relative humidity. One of our future directions is to examine further this relationship and possibly establish a solid parameterization that could be used by the modeling community to reduce biases in assimilated observations that might display a similar low bias when compared to in-situ measurements.

The MCARS simulator is currently being extended to calculate synthetic radiances for the Meteosat Second Generation Spinning Enhanced Visible Infrared Radiometer Imager.

Chapter 5: What's next?

In the preceding chapters I showed how a simple MODIS radiance simulation code with a very basic manual interface evolved into the powerful Multi-sensor Cloud and Aerosol Retrieval Simulator (MCARS) that interfaces directly with the GEOS-5 climate model. The MCARS project stands as a solid bridge between modeling and remote sensing.

The very first run-through of MCARS shed light on the causes of low bias in MODIS retrieved 550nm AOD when compared to in-situ measurements for smoke aerosols. It has been known for a considerable amount of time that in comparisons of AOD retrievals from Terra and Aqua MODIS and in-situ AERONET, MODIS-provided AOD is biased low for high AOD when smoke is present (Correia and Pires, 2006). When I compared result of executing the operational MODIS aerosol property algorithm to GEOS-5 "ground truth" aerosol properties used to create the MCARS radiances, I found a relationship that matched reality. M(O/Y)D04 retrieval from MCARS data underestimated the aerosol optical depth. Unlike planet Earth, MCARS allows us full control over all constituents of the atmospheric column and surface properties. I ran a number of experiments in which I examined effects of cloud masking, land surface inhomogeneity, phase function selection and finally aerosol single scattering albedo. The operational MODIS AOD retrieval assumes a constant 80% column relative humidity over the entire globe (Levy et al. 2007) (Levy et al 2013). Whereas this assumption is very reasonable for scattering at wavelengths less than 1μm, as figure 4-10 indicates, it is not correct when SWIR channels are used.

Operational Data Collection 6 MODIS aerosol properties algorithm uses the 2.1μm MODIS band 7. When the aerosol single scattering albedo was adjusted to fit the assumptions of MODIS AOD algorithm, the bias at high AOD disappeared. This strongly suggests that the MODIS aerosol team should consider adding a dimension of humidity to their look-up tables or use some other means such as a correction curve, to adjust their spectral reflectance profiles for effects of column humidity before retrieval is attempted.

So what is next for MCARS?

The MCARS code is publicly available under the NASA Open Source Agreement. Technically any user with sufficient computing power can immediately download and install the GEOS-5 model from the GMAO distribution server and request a copy of the MCARS code from Galina Wind (gala.wind@nasa.gov). MCARS code will eventually be available from the MODIS-Atmospheres website resource section.

If one does decide to become a registered user and continue extension and experimentation with this simulator, there is a long list of studies that can be done, some more technical, others more scientific. Here are some examples of what an MCARS user might attempt.

One may work on quantifying the uncertainty in MODIS aerosol optical depth retrievals over land due to land surface inhomogeneity in the 10x10 km boxes that the standard algorithm uses and then repeat the analysis for the 3 km resolution product.

One also may apply the MCARS code to other areas that are dominated by different aerosol types such as dust, sulfate or pollution aerosols and examine possible reasons for any retrieval and in-situ differences for those aerosols.

Another potentially interesting project would involve replacing the DISORT radiative transfer core with a vector core of a user's choice that is capable of full polarization calculations such as VLIDORT (Spurr, 2006) and performing studies of the assumptions made by the MODIS Deep-Blue aerosol algorithm (Hsu et al, 2004). It is necessary to use a vector core when working with wavelengths shorter than 500nm because polarization becomes an important factor in aerosol scattering calculations (Levy et al, 2004).

It could be also interesting to study aerosol radiative effects directly in studies that would be localized spatially, but extended temporally where a large time period is simulated over a small geographical area with a known aerosol transport such as marine boundary layer off Namibia. One could study the changes in cloud and aerosol retrievals and possibly come up with a temporal parameterization of cloud microphysics as related to monthly and even weekly aerosol variations. Parameterization like that could be valuable for climate studies because it could reduce some of the uncertainties in cloud and aerosol feedbacks and forcing.

Another approach altogether would be to create a hybrid sensor using properties of an existing one. A vector radiative transfer core would likely be required for that kind of project. The MISR instrument is very powerful for dark-target aerosol studies because it does not suffer from the sun glint issue. (Kahn and Gaitley, 2015). If MISR instrument encounters sun glint it simply switches to a different camera

angle that does not have the sun's reflection. However MISR is limited by the number of channels it carries (Diner et al., 2005). It would be possible however to add some of the channels from the MODIS instrument (Barnes et al. 1998) and create a hybrid multi-angle sensor that would include the SWIR channels. This kind of hybrid sensor could provide valuable information about 3D structure of clouds and aerosols over an area far wider than a single lidar footprint. This project would be rather complex, but I think it could be achieved by a small, sufficiently motivated team.

My next planned aerosol study is to examine the aerosol properties over China.

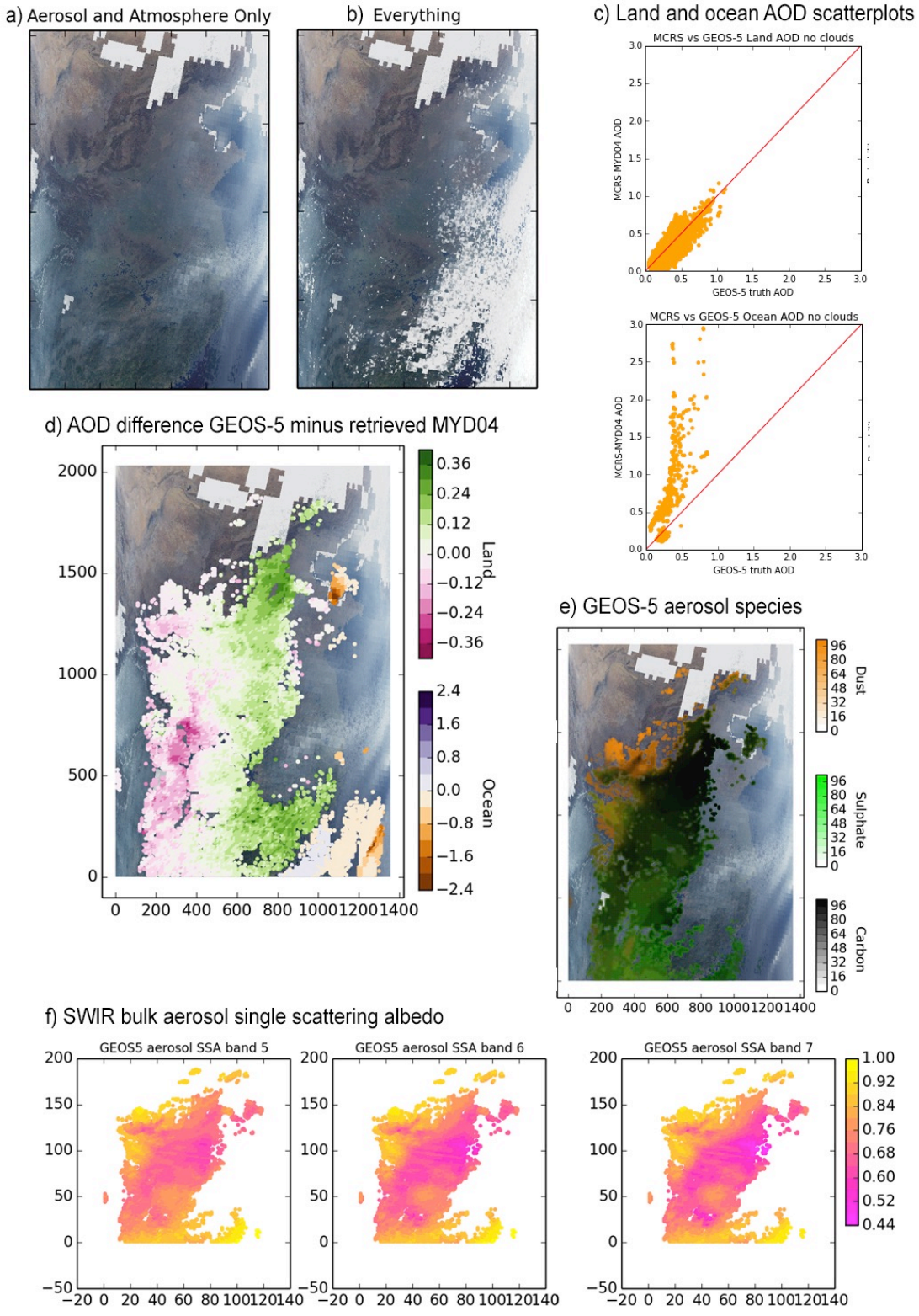

Figure 5-1. Preliminary analysis of WGNE case over China equivalent to Aqua MODIS 2013 day 012, 04:50 UTC. Panels a) and b) show the cloud-free and combined MCARS run modes. Panel c) shows the scatterplots of MYD04

AOD retrievals over land and ocean. Notice very high bias for ocean AOD retrievals. Panel d) shows the differences between retrieved and specified AOD. Panel e) shows the aerosol species mixture in this case. Panel f) shows the bulk SSA for the aerosols in SWIR bands.

Figure 5-1 shows MCARS data from a WGNE case from China, 2013 day 012, 04:50 UTC. Panels a) and b) illustrate the two relevant MCARS run modes for this case. As this is a preliminary result that had not yet received the amount of attention the Brazil cases did, all analysis shown had been done using the "everything" run mode that had all atmospheric constituents present including clouds, which explains absence of retrievals in the eastern part of the granule, mostly over ocean. Panel c) shows the scatterplots of AOD retrievals separated by land and ocean. This is the key panel of interest. The land scatterplot appears to be aligned on the 1:1 line, but that is because it includes all available retrievals over land without localized area sampling. The ocean scatterplot has a large bias where MYD04 retrieval is far larger than the source "ground truth" data. The cause of this bias is currently unknown. There are a few AERONET stations on some of the small islands in the area: Gageocho_Station (June 2011 – May 2012), Gosan_SMU (Dec 2000-present), Fukue (June 2012 – Feb 2013) and Ieodo_Station (Nov 2013-present). However their data availability appears to be quite limited with the exception of Gosan_SMU and Ieodo_Station. That particular ocean area is also frequently cloudy. This would further reduce availability of sun photometer measurements. As part of the study, all available AERONET data would need to be examined to see if this kind of bias is confirmed. This is where the fun part begins. If we move on to panel d) and stay over land, we will see that the reason the land scatterplot appears to sit so neatly around

the 1:1 line is because the areas where MYD04 underestimates AOD are balanced by areas where AOD is overestimated. If one were to make more localized scatterplots, a logarithmic-looking shape similar to AERONET-MYD04 comparisons over Brazil would emerge, particularly towards the NE. Panel e), which shows the specific aerosol species present in the column, shows how complex this case is compared to Brazil, which was just black carbon particles. In this case, dust dominates towards the west, sulfate to the south and carbonaceous particles from pollution join the sulfate and dust in the north. The area of carbon dominance also matches the areas on panel f) where the SWIR single scattering albedo is greatly reduced compared to what MYD04 would assume and therefore a hypothesis can be made that if MYD04 were accounting for relative humidity, the underestimate of AOD (green in panel d)) would disappear. Would the scatterplot then curve upward and maybe start to look more like the one over ocean? What does that mean about the aerosol and atmospheric state of the area? Is all this being caused by the complex aerosol mixture in the scene? Is this kind of mixture typical for the area? If yes, then if it does not match the one assumed by MYD04, if we adjust the mixture or altogether remove certain species and re-run MCARS, can we find the cause of the bias? These are the exciting questions that MCARS allows us to ask and maybe find answers to. If one does not execute, one will never know. MCARS allows us to completely control the atmosphere presented to operational atmospheric remote sensing retrieval algorithms. By comparing the model atmosphere input to MCARS to the atmospheric state presented to us by results of retrieval algorithms we could start getting a handle on biases that may be present in those retrieval results and consider how those biases could be corrected in

order to arrive at a state of atmosphere that may be close to the actual state present at a location of interest or even possibly globally.

There are other applications to testing of operational and experimental retrieval algorithms such as above-cloud aerosol optical depth and MODIS cloud optical and microphysical properties product (MOD06). I would like to perform further investigations of the aerosol effects on cloud retrievals like the ones in figure 4-2. I would like to quantify the effects and see if a parameterization can be made that together with the above-cloud AOD retrieval could be used to correct cloud properties retrievals in areas like marine boundary layer stratocumulus fields off the coast of Namibia that are frequently affected by biomass burning aerosol layers above the clouds. This study is also in its initial stages, less far along than the aerosols over China.

I have successfully developed simulated observations for passive imagers such as MODIS. I would expect that MCARS geolocated radiance observations could be developed for other types of sensor measurements such as active sensors or polarimeters. However using a scalar, plane-parallel radiative transfer core such as DISORT (Stamnes et al., 1988) would not be sufficient to successfully model a measurement that would be taken from something like a polarimeter or a lidar. I am aware that quite a few instruments have developed stand-alone measurement simulators along the same lines as our original MODIS simulator that was used to produce test data in chapter 2. I would be thrilled to work with other instrument teams if they would be interested in interfacing their stand-alone instrument simulators with the GEOS-5, or any other, model fields. MCARS is structured in such a way that a

different type of a pixel-level simulator could be inserted with minimal amount of effort.

There are modeling applications as well. Just using the existing MCARS data from Brazil, an interested user may build a parameterization that would allow their model to correct space borne aerosol optical depth retrievals that show a similar low bias when compared to AERONET as the MOD04 product.

The ultimate goal of MCARS is to allow the modeling community to develop methodologies for assimilating cloudy radiances and make more direct use of the cloud optical and microphysical properties retrievals provided by the many existing passive and active sensors. Then the diagnostic, or pre-set, cloud microphysical properties commonly used in GEOS-5 (Rienecker et al, 2008) could potentially be considerably improved on. Cloudy radiance assimilation is an open research problem (Bauer et al 2011), particularly assimilation of information about cloud properties and any contributions to solving it would be of importance first to numerical weather prediction and possibly to climate research. It would be reasonable to assume that weather forecasts would be improved if retrievals of cloud microphysics and information about precipitation were assimilated into the model making the forecast. Everyone is interested if it will rain tomorrow. A task of producing an accurate forecast without taking into account actual observed physical states of clouds already present in the area must be a challenge. I would like to see if the close relationship between modeling and remote sensing that MCARS fosters could be used to make the job easier.

If I were to develop model-instrument simulators for other existing sensor types, not just other passive imagers, it could be a very interesting project to then hybridize those simulators. They would all be developed using the same modular system and a blended module could be created. Then one could examine impacts of various hardware choices on resulting radiance "measurements" and possibly get help in designing a future Earth Science mission.

I welcome extensions and changes to MCARS and only would like to see wider use of the software.

I would also like to note, that even though in the first and this final chapter of this work, I say that "I did this" and "I did that", as the rules of the University of Maryland require me to say, the reality is not actually so. No one works in a vacuum completely on their own. Most research of value is a result of close collaboration, ideally of folks from different disciplines, because when we can understand and reconcile our differences and combine our strengths, that is when we can contribute the most to the community's understanding of this beautiful planet.

And finally, as one fictional character wisely said: anything that is possible might as well be tried and one always, always should know their territory.

Good luck, future researcher.

Bibliography

Ackerman, A., K. Strabala, P. Menzel, R. Frey, C. Moeller, L. Gumley, B. Baum, S. W. Seemann, and H. Zhang, 2006: Discriminating clear-sky from cloud with MODIS Algorithm Theoretical Basis Document (MOD35). ATBD Reference Number: ATBD-MOD-06. modis-atmos.gsfc.nasa.gov/reference_atbd.php. LAD:07.23.2013

Ackerman, S. A., R.E. Holz, R. Frey, E. W. Eloranta, B.C. Maddux, M. McGill, 2008: Cloud Detection with MODIS. Part II: Validation. J. Atm. Ocn. Tech, 25, 1073-1086, doi: 10.1175/2007JTECHA1053.1

Barnes, W. L., T. S. Pagano, and V. V. Salomonson, 1998: Prelaunch characteristics of the Moderate Resolution Imaging Spectroradiometer (MODIS) on EOS-AM1. *IEEE Trans. Geosci. Remote Sens.*, **36**, 088–1100.

Bauer, P., G. Ohring, C. Kummerow and T. Auligne, 2011: Assimilating satellite observations of clouds and precipitation into NWP models. BAMS, DOI: 10.1175/2011BAMS3182.1

Baum, B. A., A. J. Heymsfield, P. Yang, and S. T. Bedka, 2005a: Bulk scattering models for the remote sensing of ice clouds. Part 1: Microphysical data and models. *J. Appl. Meteor*., **44**, 1885–1895.

Baum, B. A., P. F. Soulen, K. I. Strabala, M. D. King, S. A. Ackerman, W. P. Menzel, and P. Yang, 2000: Remote sensing of cloud properties using MODIS airborne simulator imagery during SUCCESS. II. Cloud thermodynamic phase. *J. Geophys. Res.*, **105**, 11781–11792.

Baum, B. A., P. Yang, A. J. Heymsfield, S. Platnick, M. D. King, Y. X. Hu, and S. T. Bedka, 2005b: Bulk scattering models for the remote sensing of ice clouds. Part 2: Narrowband models. *J. Appl. Meteor*., **44**, 1896–1911.

Bedka, S. T., W. F. Feltz, A. J. Schreiner, and R. E. Holz, 2007: Satellite-derived cloud top pressure product validation using aircraft-based cloud physics lidar data from the ATReC field campaign. *Int. J. Remote Sens.*, **28**, 2221–2239.

Berk A., L. S. Bernstein, G. P. Anderson, P. K. Acharya, D. C. Robertson, J. H. Chetwynd, and S. M. Adler-Golden 1998: MODTRAN cloud and multiple scattering upgrades with application to AVIRIS. *Remote Sens. Environ*., **65**, 367–375.

Bony, S., R. Colman, V. M. Kattsov, R. P. Allan, C. S. Bretherton, J. L. Dufresne, A. Hall, S. Hallegatte, M. M. Holland, W. Ingram, D. A. Randall, B. J. Soden, G. Tselioudis and M. J. Webb, 2006: How well do we understand and evaluate climate change feedback processes?, J. Climate, 19, 3445-3482.Baum, B. A., and J. D. Spinhirne, 2000: Remote sensing of cloud properties using MODIS Airborne Simulator imagery during SUCCESS. III. Cloud overlap. *J. Geophys. Res.,* **105***,* 11793–11804.

Buchard, V., A. da Silva, P. R. Colarco, A. Darmenov, C. A. Randles, R. Govindaraju, O. Torres, J. Campbell, and R. Spurr, 2015. Using the OMI aerosol index and absorption aerosol optical depth to evaluate the NASA MERRA aerosol reanalysis. *Atmos. Chem. Phys*, **15**, 5743-5760. doi: 10.5194/acp-15-

5743-2015.
Buchard, V., A. M. da Silva, P. Colarco, N. Krotkov, R.R. Dickerson, J. W. Stehr, G. Mount, E. Spinei, H. L. Arkinson, and H. He, 2014: Evaluation of GEOS-5 sulfur dioxide simulations during the Frostburg, MD 2010 field campaign, *Atmos. Chem. Phys*., **14**, 1929–1941, doi:10.5194/acp-14-1929-2014
Chang, F-L., J. A. Coakley Jr. 2006: Relationships between marine stratus cloud optical depth and temperature: inferences from AVHRR observations. *J. Clim*. **20**, 2022-2036.
Chang, F. L., and Z. Li, 2005: A new method for detection of cirrus overlapping water clouds and determination of their optical properties. *J. Atmos. Sci*., **62**, 3993–4009.
Chang, F. L., and Z. Li, 2005: A near-global climatology of single-layer and overlapped clouds and their optical properties retrieved from Terra/MODIS data using a new algorithm. *J. Clim*ate, **18**, 4752–4771.
Chevallier, F. 2001: Sampled databases of 60-level atmospheric profiles from the ECMWF analyses. SAF program, Research report #4.
Chin, M., P. Ginoux, S. Kinne, O. Torres, B. N. Holben, B. N. Duncan, R. V. Martin, J. A. Logan, A. Higurashi, and T. Nakajima, 2002: Tropospheric Aerosol Optical Thickness from the GOCART Model and Comparisons with Satellite and Sun Photometer Measurements. *J. Atmos. Sci*., **59**, 461–483.
Chiriaco, M., H. Chepfer, P. Minnis, M. Haeffelin, S. Platnick, D. Baumgardner, P. Dubuisson, M. McGill, V. Noel, J. Pelon, D. Spangenberg, S. Sun-Mack, and G. Wind, 2007: Comparison of CALIPSO-like, LaRC and MODIS retrievals of ice cloud properties over SIRTA in France and Florida during CRYSTAL-FACE. *J. Appl. Meteor. Climatol*., **46**, 249–272.
Colarco, P., A. da Silva, M. Chin, T. Diehl, 2010: Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth. *J. Geophys. Res*., **115**, D14207, doi:10.1029/2009JD012820
Colarco, P. R., E. P. Nowottnick, C. A. Randles, B. Yi, B., P. Yang, K.-M. Kim, J. A. Smith, and C. G. Bardeen, 2013: Impact of Radiatively Interactive Dust Aerosols in the NASA GEOS-5 Climate Model: Sensitivity to Dust Particle Shape and Refractive Index, *J. Geoph. Res.*, doi:10.1002/2013JD020046
Correia, A., C. Pires 2006: Validation of aerosol optical depth retrievals by remote sensing over Brazil and South America using MODIS. Anais do XIV Congresso Brasileiro de Meteorologia.
Cox, C., W. Munk, 1954: Measurement of the Roughness of the Sea Surface from Photographs of the Sun's Glitter. *J. Opt. Soc. Am*., **44**, 838-850
Darmenov, Anton, and Arlindo da Silva, 2015. The Quick Fire Emissions Dataset (QFED): Documentation of versions 2.1, 2.2 and 2.4. NASA/TM–2015–104606, **38**.
Davis, S. M., L. M. Avallone, B. H. Kahn, K. G. Meyer, and D. Baumgardner, 2009: Comparison of airborne in situ measurements and Moderate Resolution Imaging Spectroradiometer (MODIS) retrievals of cirrus cloud optical and microphysical properties during the Midlatitude Cirrus Experiment (MidCiX), *J. Geophys. Res.,* **114**, D02203, doi:10.1029/2008JD010284.

Derber, J. C., R. J. Purser, W.-S. Wu, R. Treadon, M. Pondeca, D. Parrish, and D. Kleist, 2003: Flow-dependent Jb in a global grid-point 3D-Var. Proc. ECMWF annual seminar on recent developments in data assimilation for atmosphere and ocean. Reading, UK, 8-12 Sept. 2003.

Diner, D. J., B. H. Braswell, R. Davies, N. Gobron, J. Hu, Y. Jin, R. A. Kahn, Y. Knyazikhin, N. Loeb, J.-P. Muller, A. W. Nolin, B. Pinty, C. B. Schaaf, G. Seiz, and J. Stroeve, 2005: The value of multiangle measurements for retrieving structurally and radiatively consistent properties of clouds, aerosols, and surfaces", *Rem. Sens. Env.*, **97**, 495-518.

Ding, S., Y. Xie, P. Yang, F. Weng, Q. Liu, B. Baum, Y. Hu, 2009: Estimates of radiation over clouds and dust aerosols: Optimized number of terms in phase function expansion. *JQSRT*, **110**, 1190-1198

Dubovik, O., B. N. Holben, T. F. Eck, A. Smirnov, Y. J. Kaufman, M. D. King, D. Tanré, I. Slutsker, 2002: Variability of absorption and optical properties of key aerosol types observed in worldwide locations *J. Atmos. Sci.,* **59**, 590-608

Dubovik, O., M.D. King, 2000: A flexible inversion algorithm for retrieval of aerosol optical properties from sun and sky radiance measurements *J. Geophys. Res*., **105**, 20673-20696

Feldman, D. R., C. A. Algieri, J. R. Ong, W. D. Collins, 2011: CLARREO shortwave observing system simulation experiments of the twenty-first century: Simulator design and implementation. *J. Geophys. Res.*, **116**, D10107, doi: 10.1029/2010JD015350

Freitas, S. A. da Silva, A. Benedetti, G. Grell, O. Jorba, M. Mokhtari, 2015: Evaluating Aerosol Impacts on Numerical Weather Prediction: A WGNE Initiative. Symposium on Coupled Chemistry-Meteorology/Climate Modeling, Switzerland 23-25 February 2015

Frey, R. A., S. A. Ackerman, Y. Liu, K. I. Strabala, H. Zhang, J. R. Key, and X. Wang, 2007: Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5. *J. Atmos. Oceanic Tech*., **25**, 1057–1072.

Hess, M., Koepke, P. and Schult, I. 1998: Optical properties of aerosols and clouds: The software package OPAC. *B. Am. Meteorol. Soc*., **79**(5), 831–844.

Hill, C., C. DeLuca, V. Balaji, M. Suarez, A. da Silva, 2004: The architecture of the Earth System Modeling Framework, *Comp. Sci. Engr*., **6**(1), 18-28.

Hillger, D., T. Kopp, T. Lee, D. Lindsey, C. Seaman, S. Miller, J. Solbrig, S. Kidder, S. Bachmeier, T. Jasmin and T. Rink, 2013: First-Light Imagery from Suomi NPP VIIRS. *BAMS*, **94**(7), 1019-1029.

Holben, B.N., T. F. Eck, I. Slutsker, D. Tanre, J.P. Buis, A. Setzer, E.F. Vermote, J. A. Reagan, Y.J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, A. Smirnov, 1998: AERONET – A federated instrument network and data archive for aerosol characterization. *Rem. Sens. Env*., **66**, n1, p1-16.

Hsu, N.C., S. Tsay, M.D. King and J.R. Herman, 2004: Aerosol Properties Over Bright-Reflecting Source Regions. IEEE Trans. Geosci. Rem. Sens., 42, 557-569

Hu, Y.-X., B. Wielicki, B. Lin, G. Gibson, S.-C. Tsay, K. Stamnes, T. Wong, 2000: δ-Fit: A fast and accurate treatment of particle scattering phase functions with weighted singular-value decomposition least-squares fitting. *JQSRT*, **65**, 681-690

Hubanks, P. A. 2006: MODIS Atmosphere QA Plan for Collection 005. Available

from modis-atmos.gsfc.nasa.gov/MOD08_M3/atbd.html.
Hubanks, P. A, M. D. King, S. A. Platnick, and R. A. Pincus, 2008: MODIS Atmosphere L3 Gridded Product Algorithm Theoretical Basis Document. ATBD Reference Number: ATBD-MOD-30. Available from modis-atmos.gsfc.nasa.gov/MOD08_M3/atbd.html.
Intergovernmental Panel on Climate Change (IPCC), 2007: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by S. Solomon et al. Cambridge Univ. Press, Cambridge, UK
Intergovernmental Panel on Climate Change (IPCC), 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by T.F. Stocker et al. Cambridge Univ. Press, Cambridge, UK
Kahn, R. A., B. J. Gaitley 2015: An analysis of global aerosol type as retrieved by MISR. *J. Geophys. Res. Atmos*. **120**(9), 4248-4281.
Kaufman, Y.J., D. Tanre, H. R. Gordon, T. Nakajima, J. Lenoble, R. Frouin, H. Grassl, B. M. Herman, M. D. King and P. M. Teilliet 1997: Passive remote sensing of tropospheric aerosol and atmospheric correction for the aerosol effect. *J. Geophys. Res.* **102**(D14), 16815-16830.
Kaufman, Y. J., A. E. Wald, L. A. Remer, B. C. Gao, R. R. Li, L. Flynn, 1997: The MODIS 2.1μm channel - Correlation with visible reflectance for use in remote sensing of aerosol *IEEE Trans. Geosci. Remote Sens.*, **35**, 1286-1298
King, M. D., W. P. Menzel, Y. J. Kaufman, D. Tanré, B. C. Gao, S. Platnick, S. A. Ackerman, L. A. Remer, R. Pincus, and P. A. Hubanks, 2003: Cloud and aerosol properties, precipitable water, and profiles of temperature and humidity from MODIS. *IEEE Trans. Geosci. Remote Sens.*, **41**, 442–458.
King, M. D., S. Platnick, W. P. Menzel, S. A. Ackerman, and P. A. Hubanks: Spatial and Temporal Distribution of Clouds Observed by MODIS onboard the Terra and Aqua Satellites, 2013: *IEEE Trans. Geosci. Remote Sens.* **51**, pp. 3826-3852
King, M. D. S. Platnick, G. Wind, G. T. Arnold, R. T. Dominguez, 2010:Remote sensing of the radiative and microphysical properties of clouds during TC4: Results from MAS, MASTER, MODIS, and MISR *J. Geophys. Res*., **115**, D00J07, doi:10.1029/2009JD013277
King, M. D., S. Platnick, P. Yang, G. T. Arnold, M. A. Gray, J. C. Riedi, S. A. Ackerman, and K. N. Liou, 2004: Remote sensing of liquid water and ice cloud optical thickness, and effective radius in the arctic: Application of airborne multispectral MAS data. *J. Atmos. Oceanic Technol*., **21**, 857–875.
Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W-S. Wu, S. Lord, 2009: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wthr and Fcst*., 1691-1705, DOI: 10.1175/2009WAF2222201.1
Kratz, D. P., 1995: The correlated k-distribution technique as applied to the AVHRR channels. *J. Quant. Spectrosc. Radiat. Transfer*, **53**, 501–517.
Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu, N. C., 2013: The Collection 6 MODIS aerosol products over land and ocean, *Atmos. Meas. Tech*., **6**, 2989-3034, doi:10.5194/amt-6-2989-2013
Levy, R. C., L. A. Remer, Y. J. Kaufman, 2004: Effects of neglecting polarization on

the MODIS aerosol retrieval over land. *IEEE Trans. Geosci. Remote Sens.* **42**(11), pp. 2576-2583

Levy, R. C., L.A. Remer, S. Mattoo, E. F. Vermote, Y. J. Kaufman, 2007: Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance. *J. Geophys. Res.-Atmos*., **112**, D13211, doi: 10.1029/2006JD007811.

Mace, G. G., Y. Zhang, S. Platnick, M. D. King, P. Minnis, and P. Yang, 2005: Evaluation of cirrus cloud properties derived from MODIS data using cloud properties derived from ground-based observations collected at the ARM SGP site. *J. Appl. Meteor*., **44**, 221–240.

Meehl, G. A., et al., 2007: Global climate projections. In: Climate Change 2007: ThePhysical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor and H. L. Miller (eds.)]Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 747–843.

Meng, Z., P. Yang, G. W. Kattawar, L. Bi, K. N. Liou and I. Laszlo, 2010.: Single-scattering properties of tri-axial ellipsoidal mineral dust aerosols: A database for application to radiative transfer calculations, *J. Aerosol Sci*., **41**, 501–512

Menzel, W. P., R. Frey, H. Zhang, D. Wylie, C. Moeller, R. Holz, B. Maddux, B. A. Baum, K. Strabala, and L. Gumley, 2008: MODIS global cloud-top pressure and amount estimation: Algorithm description and results. *J. Appl. Meteor. Climatol*., **47**, 1175–1198.

Meyer, K. G., S. E. Platnick. 2015: Simultaneously inferring above-cloud absorbing aerosol optical thickness and underlying liquid phase cloud optical and microphysical properties using MODIS. *J. Geophys. Res. Atmos*, **120** (11): 5524–5547

Meyer, K. G., S. E. Platnick, L. Oreopoulos, and D. Lee. 2013: Estimating the direct radiative effect of absorbing aerosols overlying marine boundary layer clouds in the southeast Atlantic using MODIS and CALIOP. *J. Geophys. Res. Atmos*., **118** (10): 4801-4815

Molod, A., L. Takacs, M. Suarez, J. Bacmeister, I.-S. Song, and A. Eichmann, 2012: The GEOS-5 Atmospheric General Circulation Model: Mean Climate and Development from MERRA to Fortuna. *Tech. Rep. S. Gl. Mod. Data Assim*., **28**

Moody, E. G., M. D. King, S. Platnick, C. B. Schaaf, and F. Gao, 2005: Spatially complete global spectral surface albedos: Value-added datasets derived from Terra MODIS land products. *IEEE Trans. Geosci. Remote Sens*., **43**, 144–158.

Moody, E. G., M. D. King, C. B. Schaaf, D. K. Hall, and S. Platnick, 2007: Northern Hemisphere five-year average (2000-2004) spectral albedos of surfaces in the presence of snow: Statistics computed from Terra MODIS land products. *Remote Sens. Environ*., **111**, 337–345.

Moody, E. G., M. D. King, C. B. Schaaf and S. Platnick, 2008: MODIS-derived spatially complete surface albedo products: Spatial and temporal pixel distribution and zonal averages. *J. Appl. Meteor*. *Climatol*., **47**, 2879–2894.

Nakajima, T., M. D. King, 1990: Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I:

Theory. *J. Atmos. Sci.*, **47**, 1878-1893
Nasiri, S. L., and B. A. Baum, 2004: Daytime multilayered cloud detection using multispectral imager data. *J. Atmos. Ocean. Technol*., **21**, 1145–1155.
Norris, P. M., L. Oreopoulos, A. Y. Hou, W.-K. Tao, X. Zeng, 2008: Representation of 3D heterogeneous cloud fields using copulas: Theory for water clouds. *J. Q. R. Meteorol. Soc.* **134**: 1843–1864. doi:10.1002/qj.321.
Norris, P. M. and A. M. da Silva, 2016: Monte Carlo Bayesian inference on a statistical model of sub-grid column moisture variability using high-resolution cloud observations. Part I: Method. Accepted to J. Q. R. Meteorol. Soc.
Notarnicola, C., D. Di Rosa, F. Posa, 2011: Cross-Comparison of MODIS and CloudSat Data as a Tool to Validate Local Cloud Cover Masks. Atmos., 2, 242-255, doi:10.3390/atmos2030242
Otkin, J. A., and T. J. Greenwald, 2008: Comparison of WRF model-simulated and MODIS-derived cloud data. *Mon. Wea. Rev.*, **136**, 1957–1970.
Pavolonis, M. J., and A. K. Heidinger, 2004: Daytime cloud overlap detection from AVHRR and VIIRS. *J Appl. Meteor*., **43**, 762–778.
Pawson, S. R., S. Stolarski, A. R. Douglass, P. A. Newman, J. E. Nielsen, S. M. Frith, M. L. Gupta, 2008: Goddard Earth Observing System chemistry-climate model simulations of stratospheric ozone-temperature coupling between 1950 and 2005, *J. Geophys. Res*. **113**, D12103, doi:10.1029/2007JD009511
Pincus, R., S. Platnick, S. A. Ackerman, R. S. Hemler, and R. J. P. Hofmann, 2012: Reconciling simulated and observed views of clouds: MODIS, ISCCP, and the limits of instrument simulators. *J. Climate*, **25**, 4699-4720. doi:10.1175/JCLI-D-11-00267.1.
Platnick, S., M. D. King, S. A. Ackerman, W. P. Menzel, B. A. Baum, J. C. Riedi, and R. A. Frey, 2003: The MODIS cloud products: Algorithms and examples from Terra. *IEEE Trans. Geosci. Remote Sens*., **41**, 459–473.
Radke, L. F., J. A. Coakley Jr., M. D. King, 1989: Direct and remote sensing observations of the effects of ships on clouds. *Science*, **246**, 1146-1149
Remer, L. A., Y. J. Kaufman, D. Tanre, S. Mattoo, D. A. Chu, J.V. Martins, 2005: The MODIS aerosol algorithm, products, and validation. *J. Atm. Sci*. **62**, 947–973. doi:10.1175/JAS3385.1
Rienecker, M.M., M. J. Suarez, R. Todling, J. Bacmeister, L. Takacs, H.-C. Liu, W. Gu, M. Sienkiewicz, R. D. Koster, R. Gelaro, I. Stajner, and J. E. Nielsen, 2008: The GEOS-5 Data Assimilation System - Documentation of Versions 5.0.1, 5.1.0, and 5.2.0. *Tech. Rep. S. Gl. Mod. Data Assim*., **27**
Riishojgaard, LP. 1998: A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. Tellus 50A: 42–57.
Roberts, Y. L., P. Pilewskie, B. C. Kindel, D. R. Feldman, W. D. Collins 2013: Quantitative comparison of the variability in observed and simulated shortwave reflectance. *Atmos. Chem. Phys*. **13**, 3133-3147
Schmetz, J., P. Pili, S. Tjemkes, D. Just, J. Kerkmann, S. Rota and A. Ratier, 2002: An introduction to Meteosat Second Generation (MSG). *BAMS*, **83**(7), 977-992.
Schmit, T. J., E. M. Prins, A. J. Schreiner, J. J. Gurka, 2001: Introducing the GOES-M imager. *Nat. Weath. Dig.* **25**(3,4), 28-37.
Schoeberl, M. R., A. R. Douglass, E. Hilsenrath, P. K. Bhartia, R. Beer, J. W. Waters,

M. R. Gunson, L. Froidevaux, J. C. Gille, J. J. Barnett, P. F. Levelt, P. DeCola, 2006: Overview of the EOS Aura Mission. *IEEE Trans on Geosc. Rem. Sens*. **44**(5), 1066-1074.

Seemann, S. W., E. E. Borbas, R. O. Knuteson, G. R. Stephenson, H-L. Huang, 2008: Development of a Global Infrared Land Surface Emissivity Database for Application to Clear Sky Sounding Retrievals from Multispectral Satellite Radiance Measurements. *J. Appl. Meteor. Climatol*., **47**, 108-123.

Spurr, R. J. D., 2006: VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media, *JQSRT*, **102**, 316-342

Stamnes, K., S. C. Tsay, W. Wiscombe, and K. Jayaweera, 1988: Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. *Appl. Opt.* **27**, 2502–2509.

Stephens, G. L., S. C. Tsay, P. W. Stackhouse, P. J. Flatau, 1990: The relevance of the microphysical and radiative properties of cirrus clouds to climate and climatic feedback *J. Atmos. Sci*., **47**, 1742-1753

Trenberth, K. E., and J. T. Fasullo, 2009: Global warming due to increasing absorbed solar radiation. *Geophys. Res. Lett.*, **36**, L07706.

Wind, G., A.M. da Silva, P. M. Norris and S. Platnick, 2013: Multi-sensor cloud retrieval simulator and remote sensing from model parameters – Part 1: Synthetic sensor radiance formulation, *Geosci. Model Dev*., **6**, 2049-2062, doi:10.5194/gmd-6-2049-2013.

Wind, G., A. M. da Silva, P. M. Norris, S. Platnick, S. Mattoo and R. C. Levy, 2016: Multi-sensor cloud and aerosol retrieval simulator and remote sensing from model parameters – Part 2: Aerosols, *Geosci. Model Dev*., **9**, 2377-2389, doi:10.5194/gmd-9-2377-2016.

Wind, G., S. Platnick, M.D. King, P.A. Hubanks, M.J. Pavolonis, A.K. Heidinger, P.Yang, B. Baum, 2010: Multilayer Cloud Detection with the MODIS Near-Infrared Water Vapor Absorption Band. *J. App. Met. Clim.* **49**, 2315-2333

Wu, W. S., R. J. Purser, D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev*., **130**, 2905-2916

Xiong, X., A. T. Isaacman, and W. Barnes, 2006: MODIS Level-1B Products. Earth Science Satellite Remote Sensing, *Science and Instruments*, **1**, 33-49

Zhang, Z., and S. Platnick, 2011: An assessment of differences between cloud effective particle radius for marine water clouds from three MODIS spectral bands. *J. Geophys. Res*., **116**, D20215, doi:10.1029/2011JD016216.