ASSESSING THE PERFORMANCE OF SEVERAL MODIS LEVEL-2 CLOUD AND AEROSOL PRODUCTS AGAINST A SURFACE BASED POLARIZATION CLOUD LIDAR AT FAIRBANKS, AK

By

Patrick Cobb

RECOMMENDED: _____

_____

_____
Advisory Committee Chair

_____
Chair, Department of Atmospheric Sciences

APPROVED: _____
Dean, College of Natural Science and Mathematics

_____
Dean of the Graduate School

_____
Date

ASSESSING THE PERFORMANCE OF SEVERAL MODIS LEVEL-2 CLOUD AND AEROSOL PRODUCTS AGAINST A SURFACE BASED POLARIZATION CLOUD LIDAR AT FAIRBANKS, AK

A

THESIS

Presented to the Faculty

of the University of Alaska Fairbanks

in Partial Fulfillment of the Requirements

for the Degree of

MASTER OF SCIENCE

By

Patrick Cobb, B.A.

Fairbanks, Alaska

May 2008

Abstract

Global Circulation Models (GCMs) currently suffer from inadequate cloud and aerosol parameterizations due to a shortage of information regarding how clouds and aerosols are globally distributed in space and time. Although surface and airborne instruments can offer reliable cloud and aerosol measurements, such methods are expensive, and their scope is rather limited.

Satellite instruments, such as the MODerate resolution Imaging Spectroradiometer (MODIS), may provide an alternative. MODIS measures spectral radiance in thirty-six discreet spectral bands. A variety of cloud and aerosol properties are derived at the pixel scale from these measurements using algorithms developed from radiative transfer theory. A comprehensive validation campaign is underway to assess the quality of these data in hopes that it can be assimilated into the GCMs.

This thesis provides one such study comparing the MODIS cloud mask, cloud top properties, cloud thermodynamic phase and aerosol type with overlapping surface data from a polarization cloud lidar at Fairbanks, AK. Most of the sample is comprised of cirrus clouds, which are difficult to detect with satellites. Results are mixed, but are generally consistent with past studies. The MODIS products are fairly dependable, but the algorithms could be improved, especially over the polar regions.

Table of Contents

Page

Page

Page

Page

List of Figures

Page

Page

Page

List of Tables

Page

Acknowledgements

There are a number of people I wish to thank who helped get me through to the finish line. First and foremost, my advisor, Dr. Ken Sassen, provided valuable direction, patience and support, both financial and moral, throughout the course of this work. I would also like to thank my committee members, Dr. Shusun Li and Dr. Richard Collins. Dr. Li's experience with MODIS data was invaluable, and he was always available and happy to answer any question I threw at him. Dr. Collins also provided valuable input and somehow found time in his busy schedule to read through this behemoth the week-end before the defense. Both provided extremely valuable corrections and comments to the pre-rough draft of this manuscript.

I would also like to thank the other members of the Sassen group, Jiang Zhu, Ramaswamy Tiruchirapalli, Dr. James Campbell, Pavan Kankanala and Lillia Canonica for their help and support. I am especially indebted to both Jiang and Ramaswamy, who collectively, put in nearly a 40-hour work week helping me to learn IDL and debugging my programs, and for dousing the flames of all the technical and programming emergencies that would constantly arise. Dr. Campbell also saved my thesis defense at the last minute from certain catastrophe by helping me narrow down sixty-five powerpoint slides to thirty-three.

I would also like to thank the talented faculty of the Department of Atmospheric Science for their help and encouragement during my time at UAF. Finally, I wish to thank my family for their unconditional support.

This research was funded under NASA grant NNG04GF35G and NSF grant ATM-0630506.

Chapter 1 Introduction

1.1 Cloud Radiative Parameters

Clouds play a key role in the Earth's radiation budget. They cool the planet by reflecting back to space incoming shortwave radiation. They also warm the planet by absorbing and re-emitting outgoing longwave radiation back towards the surface. At any given moment, their net effect on the radiation balance (cooling or warming) will depend on the magnitudes of these two competing processes.

The degree to which clouds interact with radiation is a function of their optical depth, geometric thickness, water content, altitude, and phase, as well as the number density, shape and size distribution of the particles within the cloud (Cox, 1971; Stephens et al., 1990; Fu and Liou, 1993; Chylek et al., 2006). Additionally, the composition, concentration and size distribution of cloud condensation nuclei (CCN) can alter the aforementioned microphysical properties (Twomey, 1977; Rosenfeld, 2000; Bréon et al., 2002). Ice clouds generally absorb and thus emit more infrared (IR) radiation than water clouds of a similar water content due to ice having a significantly larger imaginary index of refraction than water over certain IR spectral regions (Menzel et al., 2002). Also, the albedo (reflectivity) of a cloud is inversely proportional to the size of the particles of which it is composed, owing to the concept that a large number of small particles provide a much larger scattering surface than a small number of large particles (Twomey, 1977; Bréon et al., 2002).

Ramanathan et al. (1989) describes the mechanism of cloud radiative forcing on the atmosphere using the following set of equations. At any given moment, the radiative heating of a column of atmosphere can be expressed by

$$H = S(1 - \alpha) - F, \tag{1.1}$$

where H denotes the net heating, the term $S(1 - \alpha)$ represents the warming from absorbed solar radiation given a solar irradiance, S, and albedo, α, and F represents the cooling from longwave emission. The cloud radiative forcing (CRF) is defined as the difference in the instantaneous heating rates between a cloudy and a clear atmosphere,

$$CRF = H_{cld} - H_{clr}, \tag{1.2}$$

where the subscripts 'cld' and 'clr' represent cloud and clear respectively. The total CRF is simply the sum of the shortwave component, CRF_{sw}, plus the longwave component, CRF_{lw}. Furthermore, the CRF_{sw} and the CRF_{lw} can be defined as follows

$$CRF_{sw} = S(\alpha_{clr} - \alpha_{cld}), \tag{1.3}$$

$$CRF_{lw} = F_{clr} - F_{cld}. \tag{1.4}$$

Since the clouds tend to be brighter than the earth's surface ($_{cld}$ > $_{clr}$), CRF_{sw} in Equation 1.3 tends to be negative, which denotes a cooling effect. Conversely, cloud top temperatures tend to be colder than the Earth's surface. Hence when clouds are present, the magnitude of the outgoing longwave radiation, CRF_{lw} in Equation 1.4 is reduced, which leads to a smaller F value in Equation 1.1. To compensate, the temperature of the earth-atmosphere system must rise in order to increase the magnitude of F, a process more commonly known as the *greenhouse effect*.

Overall, the cooling from the cloud albedo effect surpasses the greenhouse warming. Globally averaged annual satellite measurements collected during the 1980's for the Earth Radiation Budget Experiment (ERBE) indicate that clouds generate a net surface cooling of around -20 W m^{-2} (Ramanathan et al., 1989), which combines a -50 W m^{-2} cooling from shortwave reflection with a +30 W m^{-2} warming from longwave emission. However, large regional and seasonal differences in CRF are observed. For example, during April of 1985, deficits as high as -100 W m^{-2} over the northern high latitude oceans as a consequence of persistent stratus cloud cover were measured, while some areas of Africa experienced a net +25 W m^{-2} warming from cirrus cloud. Over the tropics, the net cloud forcing was found to balance out between the +100 W m^{-2} induced warming by cirrus with the -100 W m^{-2} induced cooling by optically thick convective clouds.

Bear in mind that these forcings are one to two orders of magnitude larger than the projected + 4 W m^{-2} warming caused by a doubling of the current CO_2 concentrations. Clearly, any slight deviation in the amount, type and distribution of clouds in response to

rising greenhouse gas concentrations could potentially offset, or even amplify, the predicted warming from carbon dioxide.

1.2 Cloud-Climate Feedbacks

A cloud-climate feedback is defined as a change in the type or distribution of clouds in response to an atmospheric temperature change. Feedbacks can be both positive, amplifying the initial temperature perturbation, or negative, offsetting it. For example, an increase in surface air temperatures would lead to an increase in evaporation. A more humid atmosphere could generate more global cloud cover as well as clouds containing significantly higher water contents, which could increase the planetary albedo and reduce temperature. Such a mechanism illustrates a negative feedback.

Conversely, an increase in temperature may alter the convection and synoptic weather patterns such that the global coverage of cirrus clouds, which tend to have a warming affect, would increase, further augmenting the initial rise in temperature. Moreover, a rise in temperature could potentially reduce the global cloud amount since the amount of water vapor needed to achieve saturation in the atmosphere, as expressed by the Clausius-Clapyron relationship, increases exponentially with temperature (Pruppacher and Klett, 1997). Reducing cloud cover while concurrently increasing the amount of water vapor, an extremely active greenhouse gas, could have major consequences for the Earth's climate. These processes illustrate a positive feedback.

1.3 Clouds and Climate Models

To predict how climate will ultimately respond to rising greenhouse gas concentrations, scientists must rely on Global Circulation Models (GCMs). A GCM is a computer model that utilizes the fundamental equations of energy and motion to simulate the general patterns of temperature, pressure, moisture and motion of the atmosphere. The latest Intergovernmental Panel on Climate Change (IPCC) (2007) assessment, based on model projections, predicts a 1.8^{o}C to 4.0^{o}C increase in the average global temperature over the next century. Uncertainties in the warming predictions can largely be attributed to the difficulties associated with modeling clouds and cloud feedbacks (Kerr, 1989; Groisman et al., 2000). Significant improvements in computer processing efficiency over the last twenty years unfortunately have not led to corresponding improvements in climate forecasts (Zhang et al., 2005a). Different models tend to be sensitive to different cloud feedbacks, and large variations in the predictions can exist even within an individual model depending upon how clouds and other variables are chosen to be parameterized (Cess et al., 1989; Senior and Mitchell, 1993). Zhang et al. (2005b) recently compared the seasonal and latitudinal variations in low, mid and high level clouds from ten separate GCM simulations using global cloud climatologies derived from satellites. They notice large discrepancies in global cloud distributions from one model to the next and between the models and the satellite measured climatologies. It was found that the models tend to underestimate (overestimate) optically thin (thick) clouds, while seasonal variations in cloud cover, especially low clouds, do not match well with observations.

To improve the confidence in the GCM predictions, scientists clearly need to improve the parameterization of clouds in climate models. Satellites can provide frequent measurements of cloud and cloud properties at the global scale, yet inferring three-dimensional cloud structures from two dimensional passive radiation measurements is inherently difficult. Algorithms, generated from satellite data, used to calculate cloud radiative parameters such optical thickness, phase, altitude, water content, and particle size distribution must continually be refined through validation studies using a combination of field observations, remote sensing data and radiative transfer theory. Without a reasonable inventory of global cloud distributions, cloud parameters can only be crudely represented in the models.

1.4 Cirrus Clouds

The majority (> 70%) of cloud cases for this research involve high optically thin cirrus clouds. Therefore their unique impacts on the climate need to be addressed. Cirrus clouds are composed of ice crystals that form in the dry upper troposphere in temperatures well below -30^{o}C, where the air is subsaturated with respect to water but supersaturated with respect to ice. Typical saturation vapor pressures at -40^{o}C with respect to water and ice are around 0.202 hPa and 0.129 hPa respectively, a roughly 64% split between these two values. In contrast, saturation vapor pressures for water at 15^{o}C fall around 20 hPa.

Approximately 20-30% of the earth is covered by cirrus at any given moment (Fu and Liou, 1993). Cirrus are typically associated with mid-latitude synoptic scale

cyclones. Rising parcels of moist air cool below the ice saturation point as they reach the cold upper troposphere. Water vapor either directly condenses into ice crystals, or forms haze particles, which spontaneously freeze. New ice crystals typically form in the top of the cloud and then slowly descend. As they fall, they grow by the diffusion of water molecules in saturated air and can clump together with other crystals. Cirrus can also form by thunderstorm convection, airplane contrails and as a result of orographic lifting of moist air over a mountain barrier (Sassen, 2002a).

In recent years, cirrus have become the focus of much attention because they have been shown to enhance the greenhouse effect (Stephens et al., 1990). Even subvisual cirrus, defined by Sassen et al. (1989) as having a visible optical depth less than 0.03, can strongly absorb and emit infrared radiation. Figure 1.1 depicts the spectral signature of an exceptionally thin cirrus cloud labeled 'Small P.S. ICE-CLD' in the 10-13μm range. This spectral data was collected during a 1996 field campaign over the Great Plains using several airborne high resolution radiometers and a cloud lidar (Smith et al., 1998). Note that the cloud exhibits relatively high transmission (warm brightness temperatures) at 10.5μm (980 cm^{-1}), but high absorption (cold brightness temperatures) at 12.5μm (800 cm^{-1}). In contrast, optically thick clouds like those depicted in Fig. 1.1 ('Low cloud', 'Semi-Tranp.CLOUD' and 'Opaque CLOUD') exhibit a relatively constant emissivity (brightness temperature) over the same 10-13μm (750 – 1000 cm^{-1}) region.

Ray tracing and Mie scattering calculations reveal that the optical thickness and albedo of ice clouds are extremely sensitive to particle shape (Mishchenko et al., 1996). Unlike liquid clouds, which are comprised of spherical droplets, ice clouds can consist of

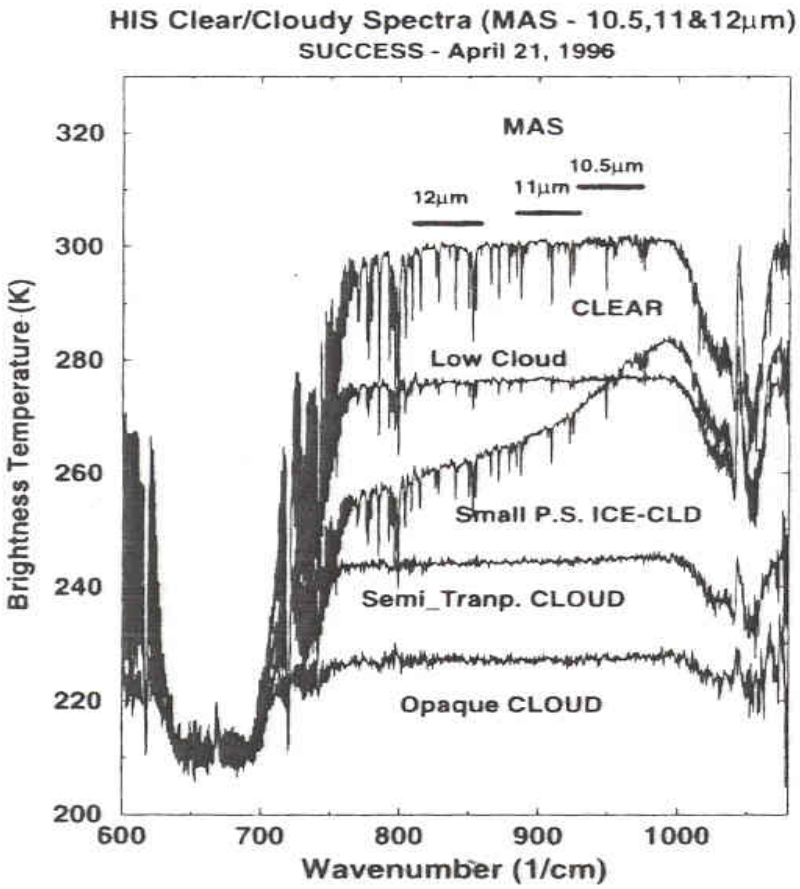


Figure 1.1: IR measurements of a thin ice cloud. High resolution Interferometer Sounder (HIS) measurements of several cloud cover conditions during the Subsonic aircraft Contrail and Cloud Effects Special Study (SUCCESS). The x-axis depicts the wavenumber (cm^{-1}) and the y-axis shows brightness temperature (K). The plot also includes, for convenience, the location of the 10.5, 11 and 12 micron (m) MODIS Airborne Simulator (MAS) channels above the 'CLEAR' curve. The curve labeled 'Small P.S. ICE-CLD' denotes the nearly invisible cirrus cloud (Smith et al, 1998)

a wide variety of regular and irregularly shaped particles including plates, needles, hollow columns, dendrites, bullet rosettes and complex aggregates. Determining the most appropriate shape and size distribution of ice clouds in climate models is a focus of ongoing research (Chou et al., 2002). Several algorithms developed from polarization lidar data can provide rough estimates of crystal shape (Noel et al., 2002, 2004), but the extraction of particle shape using passive techniques is extremely difficult.

Parameterization of particle shape into models can be incorporated using the asymmetry factor, *g*, which provides information on the bulk directionality of the scattered radiation as expressed by the phase function, $P(\cos\Theta)$, integrated over all directions. The phase function allows one to predict the probability by which an incident photon will get scattered into a particular direction. For rigorous mathematical definitions of *g* and P(cos), refer to *An Introduction to Atmospheric Radiation* by K. N. Liou (2002).

Each particle possesses a distinct phase function, which depends on the size, shape and refractive index of the particle as well as the wavelength of the incident light. Since liquid clouds are composed primarily of spherical droplets, the scattering behaviors of such clouds are relatively easy to predict. But different ice particle habits (shapes) can scatter radiation in different ways. Figure 1.2 depicts several common idealized ice crystal shapes and their associated phase functions for 0.63μm light. As with liquid droplets, forward scattering ($\Theta \approx 0^o$) dominates, but side and back scattering varies from one habit to the next. Subtle differences in the side scattering profiles can significantly affect cloud albedo by as much as a factor of three (Mishchenko et al., 1996).

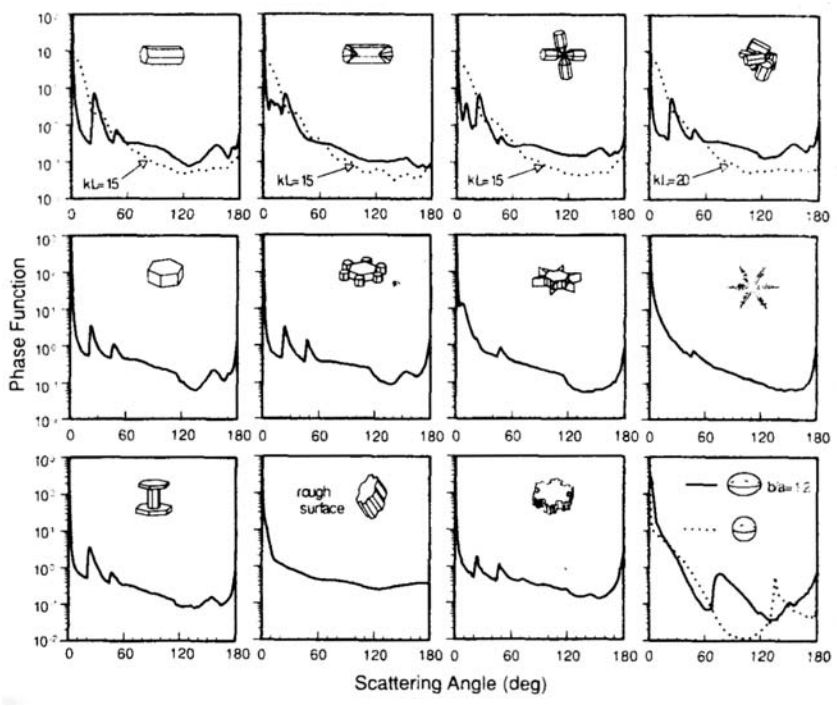


Figure 1.2: Phase functions of various ice crystal habits. Eleven ice crystal shapes and their associated 0.63μm phase functions. Solid lines are solutions from geometric ray tracing and dashed lines are from finite difference calculations. Scattering angles of 0^o and 180^o denote forward and backward scattered radiation respectively. The phase functions for spherical and spheroid droplets are included for comparison (Liou, 2002).

The detection and retrieval of radiative properties of cirrus clouds from satellites continues to be a challenge. There exists an assortment of algorithms, utilizing a broad part of the electro-magnetic spectrum from the visible to the thermal infrared wavelengths, designed to identify cirrus clouds and estimate their microphysical properties. However, these algorithms over the Polar regions are stymied by an assortment of factors such as limited sunlight availability, visible and thermal camouflage by cold ice covered backgrounds and a dry lower atmosphere, which inhibits the use of high cloud detection tests (e.g. the 1.38 m reflectance test) that rely on adequate water

vapor absorption in the lower troposphere (Gao et al., 1998; King et al., 2004; Liu et al., 2004; Berendes et al., 2004).

1.5 Aerosols

The reduction of visibility and the degradation of air quality due to forest fire smoke, Asian dust and Arctic haze are a familiar occurrence in Alaska (Tiruchirapalli, 2006). Smoke, dust and haze reside under the category of aerosols defined as small particles suspended in the atmosphere. They can be either solid or liquid and can originate from various sources, both natural and anthroprogenic. Major sources of aerosols include forest fires, volcanic eruptions, industrial pollution, airplane contrails, windblown dust, and plant pollens. In the oceans, phytoplankton blooms can generate large amounts of dimethyl sulfide (DMS), which subsequently oxidizes into sulfate aerosols (Charlson et al., 1987). Also, the mechanical action of waves at the ocean-atmosphere interface can cause bursting bubbles to expel miniature salt water droplets into the air that eventually evaporate, leaving behind large salt particles. Aerosols can also form by gas-to-particle conversions in the atmosphere (Wallace and Hobbs, 1977).

Aerosols yield considerable influence on the Earth's radiation budget in two ways. First, like clouds, they can scatter and absorb solar radiation, described as the "direct effect." Second, they can alter both the physical and hence radiative properties of clouds, also known as the "indirect effect." The latest IPCC report (2007) estimates the global radiative forcing from the direct aerosol effect to be around -0.5 ± 0.4 W m^{-2} "with a *medium-low* level of scientific understanding," while the indirect forcing ranges

between -0.3 to -1.8 W m^{-2} "with a *low* level of scientific understanding". Although the radiative forcing from aerosols is ordinarily negative, smoke particles have been shown to induce both a direct and an indirect positive forcing (Hsu et al., 2003; Koren et al., 2004).

Aerosols are vital to cloud formation. In a perfectly clean atmosphere, water vapor saturation values required to homogenously nucleate a water droplet directly from the vapor phase are extremely high due to the curvature effect on equilibrium vapor pressure (Rogers and Yau, 1989). In nature, saturation values rarely exceed 1 - 2%, thus requiring the presence of cloud condensation nuclei (CCN) or ice nuclei (IN) to aid in cloud formation, a process known as heterogeneous nucleation. CCN provide a surface upon which water molecules can condense while significantly lowering the equilibrium vapor pressure required to sustain water droplets in the ambient air. The ultimate survival of cloud droplets requires that the equilibrium vapor pressure adjacent to the droplet's surface be greater than or equal to the environmental vapor pressure. This principle can be expressed by the simplified Köhler equation (Rogers and Yau, 1989),

$$\frac{e_s(r)}{e_s(\ )} \quad 1 \quad \frac{a}{r} \quad \frac{b}{r^3}, \tag{1.5}$$

where $e_s(r)$ represents the equilibrium vapor pressure near the surface of the droplet, $e_s(\infty)$ is the environmental saturation vapor pressure, r is the droplet radius and both a and b are constants. The a/r term in Equation 1.5 represents the curvature effect, while the b/r^3

term represents the solution effect. For the case of pure water ($b/r^3 = 0$), as the droplet size decreases ($r \rightarrow 0$), the level of supersaturation [$(e_s(r)/e_s(\infty) - 1)*100$] required to sustain the droplet approaches infinity. When a solute is present ($b > 0$), the solution term on the right significantly reduces the droplet's equilibrium vapor pressure, thereby increasing the likelihood that a small droplet will survive and grow. As the droplet grows, the influence of the solution term diminishes, while the curvature term becomes dominant. Differentiating Equation 1.5 with respect to 'r' and setting it equal to zero allows one to calculate the critical radius, defined as the minimum radius that a droplet must achieve before it can be "activated." The term "activated" refers to the ability of a droplet to grow continuously by condensation without having to increase the ambient vapor pressure.

Aerosols can affect cloud microphysical properties in many ways. Several studies (Twomey, 1977; Coakley et al., 1987; Kuang and Yung, 2000) show that tropospheric sulfate pollution can increase the number of small cloud particles, while reducing the particle size deviation. These modifications can make clouds more reflective, suppress precipitation and prolong cloud duration (Albrecht, 1989; Rosenfeld, 2000). Conversely, satellite observations of smoke plumes over southeast Asia indicate that black carbonaceous aerosols can increase the absorption of solar radiation by cloud droplets (Hsu et al., 2003). Koren et al. (2004) observed that atmospheric heating as a result of solar absorption by smoke over the Amazon can significantly inhibit cloud formation altogether. Also smoke contaminated thunderstorms are known to discharge a much

higher frequency of positive lightning strikes, which are hotter and are more likely to ignite additional forest fires, than negative strikes (Lyons et al., 1998).

For temperatures between 0^{o}C and -40^{o}C, cloud droplets require some sort of freezing nucleus, either in the form of an ice embryo or an aerosol with an ice-like shape, in order to freeze (Wallace and Hobbs, 1977). In the upper troposphere, any kind of wettable aerosol can behave like an ice nucleus due to homogenous freezing in temperatures below -40^{o}C. In temperatures above -40^{o}C, ice nuclei (IN) provide a surface upon which liquid water can freeze, or water vapor can be deposited directly into an ice crystal. In addition, super-cooled liquid droplets can freeze when they come in contact with ice particles or IN which have a shape similar to that of an ice crystal. This is known as contact nucleation.

Volcanoes and dust storms supply significant quantities of IN to the atmosphere. Sulfates injected into the atmosphere by volcanoes can linger for many years in the stratosphere, reflecting shortwave radiation while seeding cirrus clouds (Jensen and Toon, 1992; Lohmann et al., 2003). Researchers using polarization lidars have observed unusually warm ice clouds with tops as warm as -8^{o}C associated with Saharan dust layers (Sassen et al., 2003; DeMott et al., 2003). Sassen (2002b) reported the transformation of a mixed phase alto-cumulus into an all ice cirrus fibratus after coming in contact with an Asian dust layer over Utah.

1.6 Objective

The primary objective of this research is to compare cloud and aerosol data collected by satellite with a reliable set of measurements collected concurrently with a surface instrument, in order to test the reliability of the satellite derived products. This validation study presents just one of many ongoing projects currently underway to test the accuracy of cloud microphysical and radiative properties calculated from satellite measurements.

Cloud and aerosol data obtained with a Polarization Cloud Lidar (PCL) is acquired at the Arctic Facility for Atmospheric Remote Sensing (AFARS) on the roof of the Geophysical Institute (64.86 N, -147.84 W) located on the west ridge of the University of Alaska, Fairbanks. The data is collected over the course of three and a half years between February 2004 and August 2007. Most of the PCL acquisitions are timed to coincide with the MODIS (MODerate resolution Imaging Spectroradiometer) satellite instrument. MODIS is a thirty-six channel scanning passive radiometer onboard both the Aqua and Terra Earth Observing System (EOS) platforms. Detailed descriptions of both the MODIS and PCL instruments and their data products are given in chapter two.

A total of 549 lidar-MODIS overlaps take place, wherein AFARS gets captured in the MODIS field of view during the lidar acquisition. The PCL dataset is considered to be the ground truth with which to compare the MODIS derived cloud products. However the methodologies used to estimate cloud top height and particle phase from the PCL data are rather subjective, and are thus error prone as will be discussed in chapter three. The

sample is collected under a variety of sky conditions, sensor viewing angles, solar illumination and land cover conditions throughout the year.

MODIS generates a plethora of microphysical and radiative cloud parameters, however only the cloud mask, cloud top properties (altitude, pressure, temperature), cloud phase and aerosol type are considered for this study. These datasets are extracted from the MOD04, MOD06 and MOD35 level-2 data files. The most fundamental of these datasets, the MODIS cloud mask, serves as a fundamental input to both the cloud phase and cloud top algorithms. A survey of the MODIS atmosphere products are covered in chapter two.

In addition, numerous cases of aerosol pollution, primarily from forest fire smoke and Asian dust, are captured, including smoke cases from the record breaking 2004 and 2005 forest fire seasons. Even volcanic aerosol is captured during the February 2006 eruption of Mt. Augustine. These lidar aerosol observations are used both to compare with the MODIS derived aerosol types from the MOD04 product, and to evaluate cases where MODIS could be mistakenly identifying aerosols as clouds or vice versa.

The variety of clouds within this sample is limited mostly to thin ice cloud layers. Because the PCL emits a shortwave ruby red (0.694 μm) laser, the beam is not designed to penetrate cloud and aerosol layers that have an integrated optical depth larger than about three ($\tau > 3$). Although several cases are included, most low and mid-level water and mixed phase clouds block the laser from penetrating much beyond a hundred or so meters into the cloud. This bias towards thin clouds will likely increase the amount of error associated with the MODIS derived cloud estimates, whose algorithms tend to be

most accurate for optically thick clouds. However, given the importance of cirrus clouds in the climate system, this sample should prove useful in testing the upper limits of MODIS's cloud detection capabilities.

1.7 Cloud Detection Over Fairbanks: Special Considerations

The MODIS cloud detection algorithm relies on a set of ancillary inputs so that it can select the appropriate cloud detection tests for a given surface type and illumination conditions. Previous and ongoing validation studies try to minimize the contaminating influences of mixed surface types by evaluating the performance of the products over simple homogenous areas, such as over the ocean, snow, or vegetation (Gao et al., 1998; Mace et al., 2005). However for this study, surface mixing unfortunately is unavoidable. The PCL site is located on the roof of a building, which is adjacent to other buildings, roads and urban features, several open fields, a lake and boreal forest. The 1 km, 5 km and 10 km resolution MODIS data will contain some or all of these cover types at different proportions for any given overpass.

Cloud detection over Fairbanks becomes even more complicated in winter, due not only to low sun angles and snow-cover contamination, but also to strong surface inversions. Inversions are defined as an increase in temperature with altitude ($dT/dZ > 0$). Cold surface inversions may effectively camouflage clouds in the thermal infrared by reducing the contrast between the cold cloud tops and the ground. The town of Fairbanks, elevation 136 m above mean sea level (MSL), is located in the bottom of the flat Tanana river valley. The northern and western periphery of town is flanked by large

hills some of which exceed 600 m in elevation. In the valley, surface inversions occur over 90% of the time in winter (Wendler and Nicpon, 1975) and temperature differentials can exceed 20^{o}C over an altitude difference of just 60 m (Jayaweera et al., 1975). However, the PCL site is located slightly above the valley at approximately 200 m above MSL, but the mostly 1 and 5 km spatial resolution of the MODIS pixels will most certainly overlap with the colder adjacent valley. Recently, cloud detection tests specifically designed to counteract the influence from surface inversions (Liu et al., 2004) have been incorporated into the new versions of the MODIS cloud detection algorithm, and have been found to significantly improve cloud detection over several Arctic and Antarctic stations.

By conducting a comprehensive long term study of the MODIS level-2 atmospheric products against a variety of less than ideal conditions, one hopes to accomplish the following:

1. Learn about the various biases and deficiencies associated with the MODIS atmosphere products so that the algorithms can be improved.
2. Provide modelers with a general overview about the reliability of several MODIS level-2 atmosphere products, primarily focusing on cases of thin cirrus cloud over a sub-arctic mixed urban environment.

Chapter 2 Theory and Instrumentation

2.1 Fundamentals of Passive Satellite Remote Sensing

Satellites have become a valuable tool in a wide variety of environmental research fields. Satellite sensors measure the amount of radiation reflected or emitted by the earth over finite spectral bands. By correlating the satellite measurements with the physical characteristics of the interacting media, a plethora of geophysical parameters can be estimated, indirectly, using algorithms derived from radiative transfer theory. For example, researchers are able to approximate plant biomass and leaf area index (LAI), defined as the total leaf surface area per unit ground surface area, from satellite measurements and using physical models that describe how shortwave radiation interacts with vegetation canopies (Myneni et al., 2002).

Remote sensing instruments are generally divided into two types: active and passive sensors. Active sensors, which include radars and lidars, create their own transmission signal using an electronic device, such as a laser or magnetron, to generate and transmit electro-magnetic waves. Passive sensors, such as cameras or radiometers, measure the existing light within the environment. The light arriving at the sensor typically comes in the form of reflected light from the sun, or in the form of emitted infrared radiation from the objects within the sensor's field of view (FOV). The light gets focused onto a detector array and the images are processed into discrete boxes or picture elements (pixels). Each pixel is assigned a grayscale brightness level proportional to the intensity of light or the number of photons falling onto the detector array.

Passive satellite radiometers typically convert the photon counts into a *spectral radiance*. The units of *spectral radiance* are watts per square meter per steradian per micron (W m^{-2} sr^{-1} μm^{-1}). This quantity is defined as the rate at which photons or light energy (Joule sec^{-1} = watt, W) of a given wavelength (per micron, μm^{-1}) originating from a particular direction (per steradian, sr^{-1}) crosses a unit area (per square meter, m^{-2}). A steradian is an element of solid angle equal to the ratio of unit area on the surface of a sphere over the radius squared (dA/r^2). A 3-dimensional steradian of a sphere can be considered analogous to the 2-dimensional radian of a circle. Converting from *spectral radiance* to *radiance* can be accomplished by integrating the former over the minimum and maximum wavelengths of the given spectral band.

In the field of thermal infrared remote sensing, defined for $\lambda > 8$ μm, spectral radiances are traditionally converted to brightness temperatures (BT's) by substituting the radiance measurement into Planck's function and solving for the temperature in degrees Kelvin (K). In the wavelength domain, this relationship can be expressed as (Dozier, 1990)

$$BT \quad \frac{hc}{k \ \ln(1 \quad y)} \tag{2.1}$$

where c is the speed of light, h is Planck's constant, k is Boltzmann's constant and y is a dimensionless constant that is inversely proportional to the sensor measured radiance. Note that BT depends on the wavelength; if the radiating body behaves like a blackbody over a given wavelength interval, then BT equals *kinetic* temperature. Similarly, the

radiant temperature (T_{rad}) of an object equals the BT integrated over all wavelengths. Hence *radiant* temperature is wavelength independent. It is related to an object's *kinetic* temperature by the simple relation (Lillesand and Kieffer, 1994)

$$T_{rad} = \varepsilon^{1/4} T_{kin} \tag{2.2}$$

where ε is the object's emissivity. However, most materials in nature do not radiate as blackbodies, hence both the BT and the *radiant* temperature of the source tend to be slightly cooler than the actual *kinetic* temperature. By definition, a *blackbody* is a perfect absorber and emitter of radiation.

The fundamental equation of passive remote sensing that describes the various sources of radiation arriving at the sensor is known as the *sensor radiance equation*. In order to design a reliable algorithm that can compute a specific scientific parameter, for example cloud thermodynamic phase, the relevant radiation source, in this case the cloud, must be identified and properly separated from the other contaminating terms in the equation, such as the surface and atmospheric terms.

The following figure and derivations for the solar and terrestrial *sensor radiance equation*s are reproduced from Michael J. Collins's fundamentals of remote sensing chapter five lecture notes on the *sensor radiance equation* available in PDF at *http://www.geomatics.ucalgary.ca/~collins/Courses/engo655/engo655_index.html* (Collins, 2007). A similar derivation can be found in chapter four of Schott's *Remote Sensing, The Image Chain Approach* (Schott, 1997).

The photons reaching the sensor can follow multiple paths. The important solar and terrestrial pathways are illustrated below in Figure 2.1.

Traditionally, the shortwave and longwave components are analyzed separately due to the limited overlap between the two domains. The shortwave includes the visible (VIS) and near-infrared (NIR) wavelengths and the terrestrial includes the mid-infrared (MIR) and thermal infrared (TIR) wavelengths. The terrestrial contribution to the shortwave radiation (VIS + NIR) is typically twenty to thirty orders of magnitude smaller than the solar contribution (Schott, 1997). Likewise, the solar contribution of longwave (TIR) ($\lambda > 8$ μm), though proportionately larger, is still two to three orders of magnitude

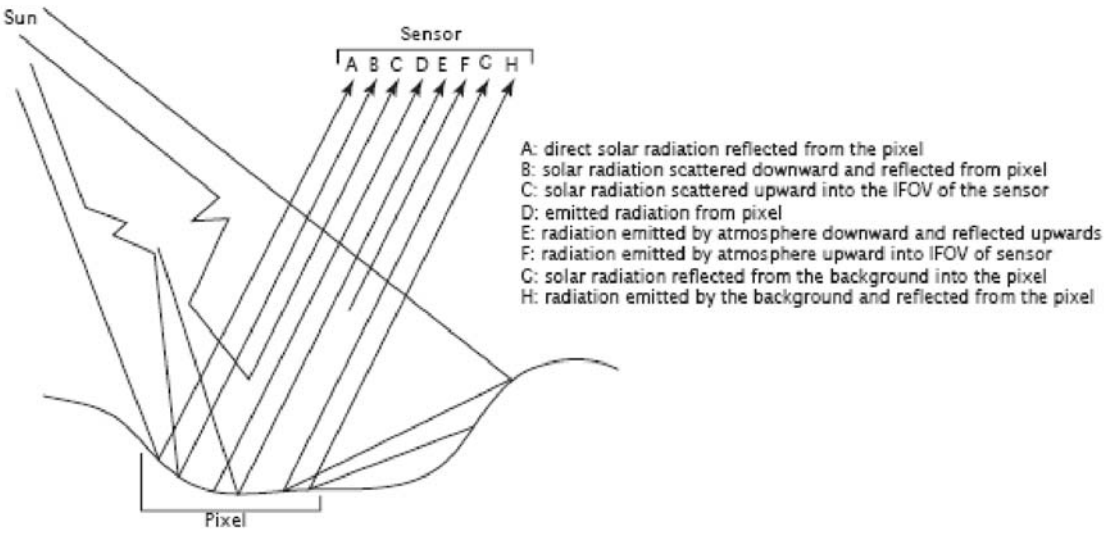


Figure 2.1 Photon pathways en route to the sensor. The shortwave (solar) sources include A, B, C and G; the longwave (terrestrial) sources are D E, F and H (courtesy of Michael J. Collins, 2007).

smaller than the terrestrial component and thus can be ignored (Schott, 1997). However, there exists some overlap between the solar and terrestrial terms when operating in the mid-infrared range (3-5 μm), which should be carefully considered when working in these wavelengths (Schott, 1997).

The reflected shortwave radiance arriving at the sensor is the sum of the *A*, *B*, *C* and *G* terms in Figure 2.1 expressed as (Collins, 2007)

$$L_{sw} \quad ([E \quad \overset{(A^*)}{T} \quad (\ )\cos(\ ) \quad F \overset{(B^*)}{E}_d \quad (1 \quad F) \quad \overset{(C^*)}{L}_b \ ]\frac{r}{—})T(\ ) \quad \overset{(G^*)}{L}_u \tag{2.3}$$

where L_{sw} is the total shortwave at sensor spectral radiance (W m^{-2} sr^{-1} μm^{-1}), and the *A*, *B*, *C* and *G* terms from Figure 2.1 are appropriately labeled above the corresponding term and marked with an asterisk. The direct and diffuse solar radiation terms (*A** and *B**) usually contribute the most, while the path radiance term (*C**) and the background term (*D**) are less substantial. E_λ and $E_{d\lambda}$ represent the top-of-the-atmosphere (TOA) direct and diffuse (skylight) solar spectral irradiances (W m^{-2} μm^{-1}) respectively. $T_\lambda(\theta)$ and T( ) are the atmospheric spectral transmissivity as a function of solar zenith (θ) and sensor zenith angles ( ). $L_{b\lambda}$ represents the element of reflected radiation originating from the adjacent topography or background, and $L_{u\lambda}$ corresponds to the radiation scattered into the sensor's field of view by the atmosphere (solar path radiance). F represents the fraction of unobstructed sky "above the pixel," while (1 – F) represents the fraction of the sky blocked by the background (trees, mountains, et cetera). Finally,

assuming that the surface behaves like a Lambertian reflector (equally scattering in all directions), the fraction of the total incoming radiation reflected by the pixel is derived by multiplying the A^*, B^* and C^* components with the factor r/π, where r is the surface reflectivity.

Moreover, the spectral transmissivity of the atmosphere, T_λ, is expressed as (Collins, 2007)

$$T(,) \quad e^{\sec(,)} \tag{2.4}$$

where the optical depth, τ, in a homogenous atmosphere equals (Liou, 2002)

$$k \quad s. \tag{2.5}$$

The term κ_λ represents the mass extinction (absorption + scattering) cross-section ($m^2\ g^{-1}$), ρ is the mass density of the medium ($g\ m^{-3}$) and s is the path length (m) through the medium. The optical depth (thickness) of the atmosphere is a unitless quantity that indicates the probability that a particular photon will be either scattered or absorbed during its transit through the medium. The total optical depth of the atmosphere combines the attenuation contributions from molecules, aerosols and clouds

$$\tau_{atm} = \tau_{mol} + \tau_{aer} + \tau_{cld.} \tag{2.6}$$

In the terrestrial domain, the primary radiation pathways arriving at the sensor are from the *D*, *E*, *F* and *H* components with the direct terrestrial and atmospheric emissions, *D* and *F*, being the dominant parts. In comparison, the indirect atmospheric and background emissions, *E* and *H*, are minor. The longwave terrestrial *sensor radiance equation* is given by (Collins, 2007)

$$L_{lw} \quad ([F \overset{(E^*)}{E}_d \quad (1 \quad F) \overset{(H^*)}{L}_b \quad] \frac{r}{} \quad \overset{(D^*)}{L}_{bb}(T))T(\ ) \quad \overset{(F^*)}{L}_u \tag{2.7}$$

with the relevant components from Figure 2.1 (*E*, *H*, *D*, and *F*) identified with an asterisk. The additional terms are as follows; $E_{d\varepsilon\lambda}$ represents the longwave radiation falling on a pixel emitted by the atmosphere; $L_{b\varepsilon\lambda}$ corresponds to the emittance onto the pixel from the background terrain; $\varepsilon_\lambda L_{bb}$ signifies the radiation emitted from the surface as a function of temperature, T; $L_{u\varepsilon\lambda}$ is the longwave radiation emitted by the atmosphere directly into the sensor's field of view (thermal path radiance). Although the terrestrial domain involves longwave emitted radiation, the surface reflectivity remains important since the background surface and atmospheric emissions falling on the pixel are partially reflected.

As was mentioned previously, the relevant component of radiance emanating from the source of interest must be isolated and accurately measured in order to serve as a reasonable input for deriving the desired scientific variable. Many of the unknown terms in Equations 2.3 and 2.7 can either be obtained from radiative transfer models or from direct field measurements. For instance, satellite navigational systems automatically record, in real time, the solar and sensor zenith angles (θ,). The TOA solar spectral

irradiance, E_λ, can be easily calculated either from Planck's function or extrapolated using surface sun-photometer measurements and a Langley plot (Liou, 2002). Atmospheric correction models, such as MODTRAN (MODerate spectral resolution TRANsmittance), offer highly accurate emission, scattering and transmittance calculations for a wide variety of atmospheric temperature, moisture and trace gas profiles, aerosol distributions and cloud cover conditions. The input profiles may originate from atmospheric radiosonde measurements or selected from a list of pre-determined conditions selected by the user (refer to *http://www.kirtland.af.mil/library/factsheets/factsheet.asp?id=7915*; MODTRAN 4, 2008). Finally, ground-truthing methods, such as surface radiometer measurements, can provide accurately calibrated values for surface radiances for homogeneous surfaces (Schott, 1997).

The amount of radiance arriving at each pixel represents an average from a fixed geometric area over the earth. The physical size of the geometric area determines the *spatial* resolution of the sensor. As the pixel size decreases, the amount of geometric detail that can be resolved in the scene increases. For example, one might be able to distinguish the roads and buildings of an urban scene in a 10 m resolution image whereas one could not at 1 km resolution.

A special problem associated with poor spatial resolution (large pixel sizes) is the issue of pixel mixing (Campbell, 2002). A mixed pixel possesses more than one radiating feature. If the physical dimensions of the pixel are larger than the dimensions of the feature, the feature cannot be spatially resolved in the image. Mixed pixels are a common occurrence near geographical boundaries, such as coastlines, or over urban

areas. The proportion of mixed pixels in an image will be dictated by the pixel size and complexity of the surface. Images of a poor spatial resolution will experience a greater degree of pixel mixing, which can be especially problematic when trying to extract the individual contributions in the sensor radiance equations (Equations 2.3 and 2.7).

Other design considerations which affect instrument performance are the *spectral* and *radiometric* resolutions of the sensor. The spectral resolution refers to both the spectral coverage and the width of the individual channels. For example, the spectral "fingerprint" of some atmospheric trace gas will reveal narrow and unique absorption features, the strength of which will depend upon the type and concentration of the gas. In order to identify the gas and estimate its concentration, the sensor must possess adequate spectral coverage of both the absorption feature(s) and the adjacent atmospheric window(s), which are regions of high transmissivity. The channels must also be narrow enough so that the strength of the absorption can be resolved from the surrounding signal.

The radiometric resolution refers to the sensitivity of an instrument at perceiving differences in temperature or brightness. The resolution is usually denoted by the number of bits associated with a given channel. For example, an 8-bit channel can provide 2^8 or 256 different shades of gray spanning the brightness range from black to white.

In order to extract the maximum amount of detail from the scene, engineers wish to design sensors that ideally contain a high degree of spatial, spectral and radiometric resolution. However, improving the resolution usually increases the amount of noise contamination from extraneous background emissions relative to the scene radiance, leading to a poor (low) *signal-to-noise* ratio (S/N) (Lillesand and Kiefer, 1994). To

achieve an acceptably high S/N ratio, either the spectral or the spatial resolution of the instrument will need to be sacrificed at the expense of the other.

2.2 MODIS and Its Gridded Atmospheric Products

2.2.1 Sensor Specifications: MODIS

The MODerate resolution Imaging Spectroradiometer (MODIS) gathers high spectral resolution images over the entire globe every other day. At the poles, the coverage occurs more frequently (multiple times per day). MODIS measures radiances in thirty-six discreet bands (channels) spanning the 0.405-14.385 μm region of the spectrum, a range that encompasses both the reflected solar bands and the thermally emitted terrestrial bands. Two channels (bands 1-2) are acquired at 250 m resolution, five (bands 3 – 7) at 500 m, and the remaining twenty-eight (8 – 36) at 1 km, at nadir. The channels provide a 12-bit radiometric resolution (2^{12} = 4096 grayscale levels), which translates to a sensitivity of less than 1^{o}C in the thermal channels. A full list of all thirty-six spectral channels and their primary functions are provided in Table 2.1.

There are two MODIS instruments currently in operation, one on board the Terra platform and the other on board Aqua. Terra was launched in December 18, 1999 and Aqua on May 4, 2002. The satellites' polar orbits are synchronized so that Terra's equatorial crossing time occurs at 10:30 AM in descending mode and Aqua's at 1:30 PM in ascending mode, which allows the instruments to maximize the best possible range in sun angles over the Earth. Both orbits are approximately sun-synchronous with inclination angles of 98^{o}. Both fly at an altitude of 705 km and have an orbital period of

Table 2.1 The 36 MODIS spectral channels. This table and additional sensor specifications can be found at *http://modis.gsfc.nasa.gov/about/specifications.php* (MODIS, 2008a).

Primary Use	Band	Bandwidth[1]	Spectral Radiance[2]	Required SNR[3]
Land/Cloud/Aerosols Boundaries	1	620 - 670	21.8	128
	2	841 - 876	24.7	201
Land/Cloud/Aerosols Properties	3	459 - 479	35.3	243
	4	545 - 565	29	228
	5	1230 - 1250	5.4	74
	6	1628 - 1652	7.3	275
	7	2105 - 2155	1	110
Ocean Color/ Phytoplankton/ Biogeochemistry	8	405 - 420	44.9	880
	9	438 - 448	41.9	838
	10	483 - 493	32.1	802
	11	526 - 536	27.9	754
	12	546 - 556	21	750
	13	662 - 672	9.5	910
	14	673 - 683	8.7	1087
	15	743 - 753	10.2	586
	16	862 - 877	6.2	516
Atmosperic Water Vapor	17	890 - 920	10	167
	18	931 - 941	3.6	57
	19	915 - 965	15	250
Primary Use	Band	Bandwidth[1]	Spectral Radiance[2]	Required NE[delta]T(K)[4]
Surface/Cloud Temperature	20	3.660 - 3.840	0.45(300K)	0.05
	21	3.929 - 3.989	2.38(335K)	2
	22	3.929 - 3.989	0.67(300K)	0.07
	23	4.020 - 4.080	0.79(300K)	0.07
Atmospheric Temperature	24	4.433 - 4.498	0.17(250K)	0.25
	25	4.482 - 4.549	0.59(275K)	0.25
Cirrus Clouds Water Vapor	26	1.360 - 1.390	6	150(SNR)
	27	6.535 - 6.895	1.16(240K)	0.25
	28	7.175 - 7.475	2.18(250K)	0.25
Cloud Properties	29	8.400 - 8.700	9.58(300K)	0.05
Ozone	30	9.580 - 9.880	3.69(250K)	0.25
Surface/Cloud Temperature	31	10.780 - 11.280	9.55(300K)	0.05
	32	11.770 - 12.270	8.94(300K)	0.05
Cloud Top Altitude	33	13.185 - 13.485	4.52(260K)	0.25
	34	13.485 - 13.785	3.76(250K)	0.25
	35	13.785 - 14.085	3.11(240K)	0.25
	36	14.085 - 14.385	2.08(220K)	0.35

1 Bands 1 to 19 are in nm; Bands 20 to 36 are in μm

2 Spectral Radiance values are (W/m2-μm-sr)

3 SNR = Signal-to-noise ratio

4 NE(delta)T = Noise-equivalent temperature difference

Note: Performance goal is 30-40% better than required

approximately 98 minutes.

The MODIS instrument utilizes a scanning mirror to collect and focus the light from the Earth. The mirror rotates at 20.3 rpm, and generates a field of view of ± 55 degrees. This translates to a swath width of about 2330 km. This type of design is known as a "whiskbroom," in which the scanning mirror generates image strips that are perpendicular to the direction of flight. MODIS possesses a 10 km along track field-of-view which translates into ten 1 km, twenty 500 m, or forty 250 m resolution lines for each scan. In addition, MODIS possesses a sophisticated onboard calibration system which monitors the spatial, spectral and radiometric performance of all thirty-six channels during each mirror rotation.

A more detailed description of the MODIS sensor can be found at *http://modis.gsfc.nasa.gov/about/* (MODIS, 2008b). Barnes et al. (1998) also provide a thorough description of the operational hardware design of MODIS.

2.2.2 The MODIS Atmosphere Products

MODIS currently generates over forty different operational data products, six of which are specific to the atmosphere, including the aerosol product (MOD04), the cloud product (MOD06) and the cloud mask (MOD35), which are being evaluated for this research. The products are stored and delivered in Hierarchical Data Format (HDF) files. By convention, files acquired by the Aqua platform are labeled with a 'MYD' prefix, and those acquired by Terra receive a 'MOD' prefix. The various products are identified by numbers such as '04' for the aerosol product, '06' for the cloud product and '35' for the

cloud mask. The HDF filename also includes the acquisition date and start time, collection version, which will be discussed briefly, as well as the date and time that the data was processed. A detailed description of the HDF file naming system can be found at *http://modis-atmos.gsfc.nasa.gov/MOD06_L2/filename.html* (HDF, 2008).

Within each HDF file exists multiple layers of variables called *scientific data sets* (SDS's) which typically constitute, along with the science parameters or channel radiances, geographical information (latitude and longitude), precise overpass times, solar and sensor viewing geometries and quality assurance (QA) data for every pixel. Each SDS consists of a 2-dimensional array of floating point numbers or integers, each array element representing a pixel value. Also each SDS is associated with an attribute list that provides supplemental information such as the name, variable type (float, long, etc.), units, range, and fill values. A full list of all cloud product SDS's and their associated attributes can be found at *http://modis-atmos.gsfc.nasa.gov/_specs/* (MODIS, 2008c).

Some SDS's require a scale factor or an offset to convert the raw pixel value, or digital number (DN), to the correct units. For example, the MODIS cloud product (MOD06) consists of forty-six SDS's, the sixteenth of which stores the cloud top temperature in degrees Kelvin. According to the MOD06 product specification, the range of digital numbers for cloud top temperature spans from 0 to 20,000. To convert these DN values to the appropriate Kelvin temperature, an *add_offset* constant of 15,000 must be added to the DN and then multiplied by a *scale_factor* of 0.01. Hence the 0 – 20,000 range corresponds to a temperature range of 150 K – 350 K.

The spatial dimensions of each MODIS image granule, which represent a five minute data acquisition, encompass an area 2330 km wide by 2030 km long. This translates to 1354 columns by 2030 rows or 270 x 406 for the 1 km or 5 km resolution products respectively. MODIS also generates 250 m, 500 m and 10 km resolution products. Note that the spatial resolution only refers to the pixel size at nadir, the position directly below the sensor along the flight path. Towards the edges, the pixels become increasingly stretched out due to the Earth's curvature and the steeper viewing angles. At $\pm 55^{o}$ sensor viewing angles, the pixel dimensions of the 1 km x 1 km product are stretched to approximately 2 km x 4.8 km. This increase in pixel size is associated with a special type of edge distortion known as the *bowtie effect*. It is mainly the result of pixel overlap which causes the same region of the earth to be sampled more than once. The bowtie effect can be corrected for using a reprojection algorithm. For a discussion of the bowtie effect, refer to: *http://www.sat.dundee.ac.uk/modis-faq.html* (Bowtie effect, 2008).

Data products generated from the Earth Observing System (EOS) satellites come in five standard levels, levels 0 – 4, based on the amount of processing that a product has undergone (refer to *http://ecsinfo.gsfc.nasa.gov/sec3/ProductLevels.html*; MODIS, 2008d). Level-0 is the rawest form of data, which has experienced only minimal processing. Level-0 data is not available for public use as it does not really provide any practical research applications.

Level-1 (A and B) files typically provide the channel radiances as well as the navigational and viewing geometry information during the acquisition. These data have

been radiometrically and geometrically calibrated, but still contain the edge distortions from the bowtie effect.

Level-2 products are geophysical variables produced at the pixel scale for every 5-minute granule of data. They are derived from the level-1 channel radiances and from additional level-2 SDSs (e.g. the cloud mask). Level-2 data include the products labeled MOD04 to MOD44 listed here: *http://modis.gsfc.nasa.gov/data/dataprod/index.php* (last referenced, (MODIS, 2008e). Depending on the SDS, level-2 data can be produced at the 1 km, 5 km or 10 km resolutions.

Level-3 data are essentially level-2 resampled to a courser 1^{o} latitude by longitude grid and averaged over the daily, 8-day and monthly timescales. In the field of climate modeling, level-3 products are generally considered the most useful as they provide statistical information (mean and standard deviation) within each 1^{o} x 1^{o} box globally.

Finally level-4 data, which represent the highest level of processing, are derived from model outputs using lower level inputs. Examples of level-4 data include land Evapotranspiration (MOD 16) and Ocean Primary Productivity (MOD 27). Currently, there are no level-4 atmospheric products. The cloud and aerosol products being evaluated for this research are level-2 type data.

MODIS level-1 and higher level atmospheric products can be ordered free of charge through the level-1 and Atmosphere Archive Distributions System (LAADS) at the following website: *http://ladsweb.nascom.nasa.gov/index.html* (LAADS, 2008a). The search interface allows the user to select the product type and to specify the date and overpass time as well as the geographic coordinates of the region of interest.

The algorithms used to derive the various MODIS products are constantly changing as efforts to improve the existing algorithms and to incorporate new ones are continually evolving. Every few years, the MODIS science team reprocesses all of the archived MODIS data using a revised version of the algorithms. The algorithm version is denoted by the *collection* number in the product name. The most recent version of the *collection* type is version 005. All of the cloud product data for this research falls under the collection 005 category, while the aerosol data is collection 004. Algorithm modifications can found at *http://modis-atmos.gsfc.nasa.gov/products_C005update.html* (MODIS, 2008f).

2.3 The MODIS Algorithms

2.3.1 Cloud Mask

2.3.1.1 Cloud Mask: Description

The most fundamental of all MODIS level-2 products is the cloud mask, a 1-km product denoted as MOD35 or MYD35 for the Terra and Aqua products respectively. Many remote sensing applications rely on the cloud mask to screen out cloud obstructed pixels. The MODIS cloud mask combines earlier cloud detection techniques developed from previous satellite campaigns like the Advanced Very High Resolution Radiometer (AVHRR) Processing scheme Over cLoudy Land and Ocean (APOLLO), the International Satellite Cloud Climatology Project (ISCCP) and the Cloud Advanced Very high resolution Radiometer (CLAVR), with newer tests that exploit MODIS's additional spectral band (Ackerman et al., 2002).

The cloud mask computes a 'Q' score which assigns one of four levels of clear-sky confidence to each pixel as follows: ($Q \geq 0.99$) confident clear, ($0.95 \leq Q < 0.99$) probably clear, ($0.66 \leq Q < 0.95$) probably clouds and ($Q < 0.66$) confident clouds. The algorithm is an amalgamation of numerous cloud detection tests, each using a set of radiance (visible) or brightness temperature (infrared) thresholds to determine the likelihood that a give pixel is cloud covered. The thresholds for an individual test may vary depending on the surface conditions (e.g. snow cover) and land cover types (e.g. vegetation, ocean). Clouds generally tend to be brighter in the visible and dimmer in the infrared than the surrounding surface, so many of the detection tests exploit these common attributes. If the results of the spectral tests are inconclusive, which is common in cases of broken clouds or in the vicinity of cloud edges, spatial and temporal variability tests can be applied as a last resort. According to the latest version (2002) of the cloud mask theoretical basis document (ATBD), only the spatial variability tests are applied, and only over large bodies of water, where the surface tends to have a uniform thermal radiance and where sunglint effects can be accounted for. However, the most current listing of cloud mask bits (R. Frey private communication, 2006) specifies that a temporal variability test is now included along with several new land based spatial variability clear-sky restoral tests.

Results of the cloud mask algorithm are stored in binary format as individual bytes. There mask provides six bytes of information, numbered 0 to 5; each byte contains eight bits, numbered 0 to 7. The first byte, denoted as 'Byte #0' illustrated in Table 2.2, corresponds to the main cloud mask byte. The remaining five bytes store the results of the

Table 2.2 Translation of the main cloud mask byte (Byte #0). This and other byte fields can be found at *http://modis-atmos.gsfc.nasa.gov/MOD35_L2/format.html* (MODIS, 2008g).

Cloud_Mask Bit-Field Interpretation		
Byte #0 (1km Cloud Mask & Processing Path Flags)		
Bits	Field Description	Bit Interpretation Key
0	Cloud Mask Flag	0 = Not Determined 1 = Determined
1-2	Unobstructed FOV Quality Flag	0 = Confident Cloudy 1 = Probably Cloudy 2 = Probably Clear 3 = Confident Clear
3	Day/Night Flag	0 = Night 1 = Day
4	Sunglint Flag	0 = Yes 1 = No
5	Snow/Ice Background Flag	0 = Yes 1 = No
6-7	Land/Water Background Flag	0=Water 1=Coastal 2=Desert 3=Land

individual cloud tests. Nine additional quality assurance (QA) bytes contain supplemental information regarding the confidence of the cloud mask overall, whether or not an individual test was applied, and the ancillary input data sources.

In order to extract the bit information, the pixel value, an integer ranging from 0 to

255, must first be converted to its binary equivalent, which is a string of eight consecutive ones and zeros. By convention, the ordering of the bits begins with the last bit (bit 7) first (bit 0). For example, assume that a pixel value for the first byte of the cloud mask happens to be '249'. Represented as a binary, this integer translates to '11111001'. After reordering the string, the conventional translation becomes '10011111'. Refering to Table 2.2, a value of 249 decodes to the following:

bit 0 = '1', the cloud mask was *Determined*;

bits 1-2 = '00' = *0*, the cloud mask denotes *Confident Cloudy*;

bit 3 = '1', the overpass time occurred during the *Day* defined for solar zenith angles $\leq 85^{o}$

bit 4 = '1', denotes *No Sunglint* (sunglint is only flagged over water at certain sensor viewing geometries);

bit 5 = '1', denotes *No Snow/Ice Background*;

bits 6-7 = '11' = *3*, denotes *Land* as the surface type.

To compute the main cloud mask, stored in bits 1-2, the algorithm must assign a *domain* to a given pixel based on the surface type and the solar illumination conditions. The domain is specified by the input flags in bits 3 – 7 of Byte #0, and determines which spectral tests to include during the processing path. It would not make any sense, for example, to apply the visible reflectance tests on a *night* flagged pixel. Also, some threshold values for a given test might need to be adjusted depending on the surface type

or if snow cover exists. Domain inputs, like snow cover and ecosystem type, may originate from outside data sources. The sources used are specified in the QA bytes. For example, information regarding snow cover may come from the National Snow and Ice Data Center (NSIDC) snow cover map, the National Oceanic and Atmospheric Administration (NOAA) snow cover product, or from MODIS, itself, which can apply a normalized snow difference index (NSDI) calculation using the visible and near-infrared channels. Additionally the algorithm uses an ecosystem map to provide a more detailed surface cover type that helps identify the pixel background (bits 6 and 7). The ecosystem types are also identified in the QA bytes.

Some tests specialize on detecting particular types of clouds located in a particular region of the atmosphere. For example, the 1.38 μm near-infrared reflection test, stored as bit 16 in Byte #2, is used primarily to detect high thin cirrus. Since this wavelength is located near a water vapor absorption band, the 1.38 μm wavelength usually gets absorbed by moisture in the lower atmosphere, and thus the signature from low level clouds is minimal. Water vapor absorption is quite small in the dry upper troposphere where cirrus clouds exist, so cirrus, if present, will reflect a significant portion of this wavelength back to the sensor. This test has proven useful at detecting extremely thin translucent cirrus clouds (Gao et al., 1998).

The cloud detection tests are assembled into five groups, based upon the type of cloud cover conditions for which they are suited to detect. The groups aim to be independent of one another, but some tests may overlap with more than one group. For a

complete list and brief descriptions of each of the various cloud detection tests, refer to the MOD35 algorithm theoretical basis document (ATBD) (Ackerman et al., 2002).

A set of threshold values for each test determines the likelihood that a pixel is obscured by cloud. For example, Figure 2.2 provides the brightness temperature thresholds for the 11 μm brightness temperature (BT_{11}) cloud detection test over the ocean, stored as bit 30 in the most current bit field description. The confidence level is plotted as a function of the brightness temperature. Note that instead of just one temperature threshold, there are three. If the 11 μm brightness temperature of the pixel is colder than the 267K minimum or above the 273K maximum, the test will yield either a 0% confidence of clear or a 100% confidence of clear. However, if the temperature falls in between these values, the confidence level is interpolated using a simple linear relationship, illustrated below.

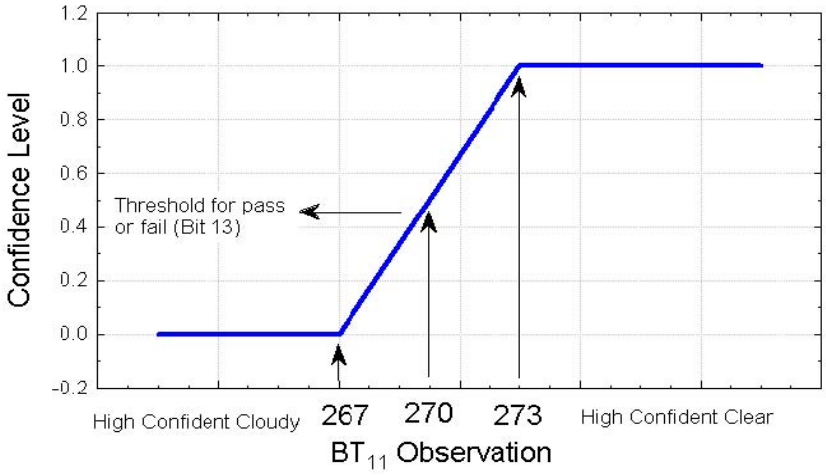


Figure 2.2: Temperature thresholds for the BT_{11} test over the ocean. The x-axis gives temperature (K) and the y-axis gives confidence level (Ackerman et al., 2002)

Once all of the appropriate tests have been performed, the minimum confidence scores from each group are collected,

$$G_{i=1,5} = \min[F_i] \tag{2.8}$$

where G_i represents the minimum confidence score from group i. The final cloud mask can then be computed by taking the N^{th} root of the product of the minimum scores;

$$Q \quad \sqrt[N]{\prod_{i\ 1}^{5} Gi}\ . \tag{2.9}$$

Note that the algorithm is "clear sky conservative" so that if one of the test results happens to score a zero, Q is automatically zero regardless of the other test results. Also, it is important to realize that a final cloud mask indicating *confident clear* or *probably clear* does not necessarily mean that the result of any individual test passed the clear threshold. The individual test results are provided in the bytes of the MOD35 file, but the actual confidence value, which can range from zero to one, is not included. Only a Boolean 'yes' or 'no' as to whether or not the pixel passed the clear sky threshold is provided.

Given the importance of the cloud mask in many fields, the main byte (Byte #0) is included with other level-2 products including the cloud product. However the original 2030 x 1354 1-km pixel array has been resized to a courser 406 x 270 5-km dataset. The resampling method assigns the center pixel of each 5 x 5 1-km box as the representative 5-km pixel (R. Frey private communication, 2006). However, this method could

potentially have some drawbacks in that the cloud cover conditions represented in the 1-km^2 region may not necessarily be representative of the cloud cover conditions over the entire 25-km^2 area. Therefore this study intends to examine the cloud mask using three different techniques. First, the accuracy of the 5-km pixel nearest to the ground site will be examined, using the cloud mask included with the MOD06 product. Second, the accuracy of the 1-km pixel, from the actual MOD35 cloud mask, located above the ground site will be examined. Third, the mode value of a 5x5 1-km pixel array, also derived from the actual cloud mask, above the ground site will be evaluated.

2.3.1.2 Cloud Mask: Possible Sources of Error

Cloud detection is most problematic when the contrast in reflectance or temperature between the cloud and the underlying surface is poor (Ackerman et al., 2002). In the visible, surfaces with a high albedo such as snow and ice covered areas, deserts and salt flats can sometimes be misinterpreted as clouds. In such cases, the ancillary inputs will denote areas of desert or snow covered regions so that the threshold values of the cloud detection tests can be modified accordingly. Sunglint, caused by the specular reflections of sunlight off of water, can also contaminate the field-of-view. MODIS includes a "sunglint" flag to warn the user when the magnitude of the reflected sun angle over the ocean is less than 36^o.

Poor thermal contrast usually results when the surface is covered in snow or ice, or if a surface temperature inversion is present. Both conditions are common over Antarctica, Greenland, and the northern Polar latitudes during winter. Problems with

cloud detection are exacerbated by long nights and low sun angles during winter in the northern regions. A scene that gets classified as *night* automatically reduces the number of threshold tests that can be applied.

The detection of optically thin nearly transparent cirrus has always been difficult for passive instruments. Fortunately, there are several NIR reflectance and brightness temperature tests which are designed specifically to sense thin cirrus. The 1.38 μm near-infrared reflectance test mentioned earlier is one such test that uses a reflectance threshold to look for high clouds. This test is only effective when there is enough humidity in the air to absorb the radiation upwelling from the surface. Over polar and high altitude regions, where the integrated precipitable water is less than 1 cm, a much larger component of the surface reflectance is allowed to reach the sensor, thereby confusing the cloud signal (Ackerman et al., 2002).

The cloud mask also includes a non-cloud obstruction flag, which is reserved for cases suspected of being obstructed by an aerosol, such as smoke, dust or a volcanic plume. It uses several brightness temperature and reflectance tests which help to discriminate aerosol from cloud, however, many cases will arise where it may be impossible to separate the two.

Finally, cases of broken clouds or cloud edges resulting in pixel that is less than 100% cloud-covered, can yield erroneous results. Cloud shadows, which lower the surface albedo, are also flagged for all confident clear sky daytime scenes. However, the cloud shadow flag is considered experimental (Ackerman et al., 2002).

2.3.1.3 Cloud Mask: Validation Studies

Validation studies are currently underway to examine the performance of the MODIS cloud mask. Some field studies employ active and passive ground instruments to collect cloud cover data concurrently at the time of a MODIS overpass. In other field tests, researchers have flown concurrently with onboard instruments. The instruments collect both remotely sensed and in situ cloud data, which can then be compared with the MODIS derived products.

Yet another validation technique will evaluate the MODIS products against the products derived from other satellite sensors. In addition to MODIS, the Aqua and Terra platforms contain a host of instruments suitable for validation studies. The recently launched Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) and CloudSat satellites, part of the A-Train constellation that includes the Aqua platform, contain an active lidar and a millimeter wavelength radar that can provide unambiguous cloud detection data to compare with the MODIS Aqua cloud products.

An exploratory study over southern Africa during the summer of 2000 observed that the MODIS Airborne Simulator (MAS) derived cloud mask was able to differentiate clouds from aerosols over both the land and the ocean (King et al., 2003). Another study (Berendes et al., 2004) examined the accuracy of the MODIS cloud mask against a variety of cloud cover measuring ground instruments at the Atmospheric Radiation Measurement (ARM) site in Barrow, AK from February to September of 2001. The study compared 3650 cloud cover cases between MODIS and a set of ground-based instruments which included a Vaisala Ceilometer (VCEIL), a Micropulse Lidar (MPL), a

Whole-Sky Imager (WSI), and a Normal Incidence Pyrheliometer (NIP). The site encompassed a 700 km^2 area. Researchers found that the MODIS cloud cover agreed to within ± 20% approximately 77% of the time with the VCEIL estimate, 83% of the time with the MPL, 81% of the time with the ARSCL product and 74% of the time with the WSI. They also concluded that MODIS was better than the surface instruments at detecting high thin cirrus, but had problems detecting low level cumulous and fog.

To confront the challenges of cloud detection over snow and surface temperature inversions, Liu et al. (2004) examined the performance of the MODIS cloud mask over three stations in the Arctic and one in the Antarctic during 2001 and 2002. Two new brightness temperature tests were included to supplement the regular cloud detection tests. With the additional tests, the MODIS false positive cases were reduced from 44.2% to 16.3% over the Arctic and from 19.8% to 2.7% over the Antarctic. Most of the false positives occurred during the nighttime overpasses, when inversions were most likely to be present.

One of the first satellite to satellite validation studies to examine the cloud mask was conducted in 2003 using the Geoscience Laser Altimeter System (GLAS) 1064 nm laser on board the Ice Cloud and land Elevation Satellite (ICESat) platform (Mahesh et al., 2004). Although the laser failed a little more than a month after it began collecting data in February of 2003, researchers were able to gather ninety-nine cases of GLAS-Aqua overlap and twenty-four cases of GLAS-Terra overlap. Overall, the MODIS generated cloud mask agreed 77% of the time with the GLAS instrument, however, the agreement with the MODIS-Terra dataset was significantly poorer because most of the

Terra-GLAS overlaps occurred at night. In addition, agreement over high latitude areas suffered because of snow cover.

Another study devised a maximum likelihood (ML) cloud and surface cover classification that used the MODIS cloud mask to define the initial training sites in the scene (Li et al., 2003). A ML classification is a statistical method that attempts to assign a random pixel belonging to an unknown type into a known group (class) by matching its pixel statistics with those to the set of pre-existing groups (Lillesand and Kieffer, 1994). The ML was applied to a variety of daytime scenes over the Midwestern United States. Based on qualitative image interpretation, the maximum likelihood classification using the MODIS cloud mask seemed to improve the initial cloud mask classification, especially over snow and desert covered areas where the cloud mask occasionally misidentified the bright surface as cloud.

2.3.2 Cloud Top Properties

2.3.2.1 Cloud Top Properties: Description

The cloud top temperature and pressure datasets are computed globally every one to two days at 5 km resolution and are included with the MODIS cloud product. Calculation of the cloud top properties entails the use of the MODIS cloud mask, radiance measurements in five of MODIS's infrared channels located near the 15 μm CO_2 absorption band, and surface and atmospheric temperature and moisture profiles derived from National Centers for Environmental Prediction (NCEP) or from NASA's Data Assimilation Office (DAO). The cloud top pressure calculation employs two

algorithms; the carbon dioxide density slicing algorithm, henceforth labeled CO_2, is applied to optically thin and thick mid to high level clouds, and the 11 micron brightness temperature method, henceforth known as BT_{11}, is used with optically thick low clouds.

The CO_2 algorithm has proven to be a robust means of estimating cloud top pressure for both optically thick and thin clouds above 700 hPa in altitude (Frey et al., 1999; Zhang et al., 2002). The underlying theory behind this method is based on the varying degrees of absorption of certain spectral bands located near the edge of the broad 15 m CO_2 absorption line. The algorithm uses MODIS channels 31 and 33 – 36 corresponding to the band centered wavelengths of 11.0, 13.3, 13.6, 13.9 and 14.2 μm respectively. Band absorption increases as the 15 μm CO_2 absorption line is approached. As absorption increases, the altitude where the source of radiation originates also increases.

The cloud top pressure is calculated by taking the ratio of the cloudy minus the clear sky radiances for two adjacent spectral bands, expressed as (Menzel et al., 2002)

$$\frac{R(\nu_1) \quad R_{clr}(\nu_1)}{R(\nu_2) \quad R_{clr}(\nu_2)} \quad \frac{NE_1 \int_{Ps}^{Pc} (\nu_1,p) \frac{dB[\nu_1,T(p)]}{dp} dp}{NE_2 \int_{Ps}^{Pc} (\nu_2,p) \frac{dB[\nu_2,T(p)]}{dp} dp} \qquad 2.10$$

where R is the cloudy radiance measured by the sensor for the bands centered at frequencies ν_1 and ν_2, R_{clr} is the corresponding clear sky radiance for a given surface reflectance, atmospheric temperature and moisture profile, N is the cloud cover fraction

in the field-of-view, E_i is the emissivity for band i, τ is the atmospheric transmissivity for frequency ν at pressure level p, and B is the frequency dependent Planck radiance as a function of temperature and pressure. The term on the right is integrated from the pressure at the surface, P_s, to the pressure at the cloud top, P_c. One can assume that, for bands that are spectrally close together ($\nu_1 \approx \nu_2$), the spectral emissivities are nearly identical ($E_1 \approx E_2$). The R_{clr} values are obtained from the Pressure layer Fast Algorithm for Atmospheric Transmittances (PFAAST) radiative transfer model. The other variables utilize atmospheric temperature and moisture profiles from NCEP reanalysis data. P_c can be computed through forward calculations of the integral. Once P_c is determined, the cloud top temperature can be interpolated from the NCEP profile.

In order to calculate a P_c for a given 5 x 5 1-km pixel field-of-view, a minimum of four of the twenty-five 1-km cloud mask pixels must be flagged as either *probably cloudy* or *confident cloudy*. If this basic requirement is met, then five separate spectral ratios using Equation 2.10 are calculated using the following five band ratios: (B36/B35), (B35/B34), (B35/B33), (B34/B33), and (B33/B31), the latter being reserved for ice clouds. The value of the cloud top pressure, P_c, which provides the best solution for all five ratios is assigned as the optimal solution.

The CO_2 slicing algorithm only works if the cloudy minus clear sky radiance difference falls outside of the instrument noise levels (Menzel et al., 2002). This also applies to low cloud cases possessing tops below 700 hPa, where signal from the cloud is too weak to generate a large enough difference. For these cases, the cloud top is

computed by matching the 11 μm brightness temperature with the temperature in the atmospheric profile, assuming an optically thick cloud that behaves like a blackbody.

2.3.2.2 Cloud Top Properties: Possible Sources of Error

The CO_2 slicing algorithm generally tends to overestimate the cloud top pressure (underestimate the geometric height), because it measures the "radiative center-of-mass" of the cloud (Menzel et al., 2002, 2006). For thin translucent cirrus, this usually falls near the geometric center of the cloud, but approaches the actual cloud top as the optical depth increases. The algorithm also assumes a single cloud layer within the field-of-view. Cases of multiple cloud layers, which occur over 50% of the time (Baum and Wielicki, 1994), will contaminate the results, especially for cases of thin cirrus clouds overlying optically thick water clouds. This influence will also tend to underestimate the cloud top altitude. Additional sources of error include incorrect surface temperature estimates, erroneous temperature profiles, instrument noise and any influences that could confuse the estimates of clear sky radiance (Menzel et al., 2002). However, these sources are thought to be relatively minor compared to the first two.

Menzel et al. (2002) report that most P_c errors tend to fall within the 50 – 100 hPa range of the actual cloud top for mid to high level clouds. The largest estimated error generated from a radiative transfer model was computed to be 220 mb (hPa), which occurred for a thin ice cloud with an effective emissivity, defined as the product N*E (Equation 2.10), of 0.10 overlying an optically thick water cloud located at an altitude of 700 mb.

2.3.2.3 Cloud Top Properties: Validation Studies

The CO_2 algorithm has been around for at least three decades. Smith and Platt (1978) used it to infer cloud top heights using data collected by an infrared radiometer on board the Nimbus 5 satellite. Data from the meteorological Geostationary Operational Environment Satellite (GOES), High resolution Infrared Radiometer Sounder (HIRS), High resolution Interferometer Sounder (HIS), Visible-infrared spin scan radiometer Atmospheric Sounder (VAS), and most recently from the MODIS Airborne Simulator (MAS), an airborne prototype of MODIS, have all used the CO_2 technique to estimate cloud top pressures. Validation methods have varied since cloud top heights can be measured directly using ground, airborne and satellite based lidars and radar, or indirectly using stereo parallax views of clouds and cloud shadows from visible satellite imagery (Naud et al., 2005).

Frey et al. (1999) compared over 4700 simultaneous observations of cloud top heights, 66% of which were multi-layer cases, using the MAS instrument and a dual polarization CLS (Cloud Lidar System) on board the NASA ER-2 aircraft during the Subsonic Aircraft Contrail and Cloud Effects Special Study (SUCCESS) field experiment over the Southern Great Plains (SGP) of the U.S. The MAS CO_2 derived cloud top heights fell to within 500 m and 1500 m for 32% and 64% of the cases respectively. For optically thin clouds those accuracies fell around 30% and 63% respectively. The researchers concluded that inaccurate clear-sky radiance estimates pose a greater risk to P_c estimates than multi-layer clouds.

Using a ground based Millimeter Cloud Radar (MMCR) and an Atmospheric Emitted Radiance Interferometer (AERI) over the SGP Atmospheric Radiation Measurement (ARM) site, Mace et al. (2005) compared cirrus cloud properties derived from the ground based instruments with those derived from MODIS. They concluded that the MODIS cirrus height retrievals tend place the cloud lower in the atmosphere by about 100 – 200 hPa.

Another study compared the MODIS and MISR (Multi-angle Imaging Spectro-Radiometer) derived cloud top heights with those from a ground based MMCR at two sites, one located over the ARM SGP site, which tends to have a large frequency of high clouds, and the other at the Chilbolton Facility for Atmospheric and Radio Research (CFARR) in the U.K., a site with a high frequency of low level marine clouds (Naud et al., 2005). Differences between the MODIS derived CO_2 retrievals and the MMCR retrievals for mid to high level clouds fell around -1.2 ± 1.0 km and 0.6 ± 1.3 km for the SGP and CFARR sites respectively. When the BT_{11} method was used, the correlation between the MODIS and the MMCR were found to be -1.4 ± 2.8 km and -0.5 ± 2.3 km for the SGP and CFARR sites respectively. They also found cloud top height agreements to within 1 km between the MODIS and MISR.

2.3.3 Cloud Thermodynamic Phase

2.3.3.1 Cloud Thermodynamic Phase: Description

One of the more important 5 km resolution datasets included with the MOD06 product is the cloud thermodynamic phase. Particle phase significantly influences how

radiation interacts with clouds, which is important for modeling the atmospheric radiation budget. Additionally, particle phase is a required input to both the cloud particle size (r_{eff}) and the optical thickness (τ) algorithms (King et al., 1997; Platnick et al., 2003; Lee et al., 2006). Calculation of the cloud particle phase primarily employs two algorithms which exploit the differences in the refractive indices of water and ice over discreet spectral regions. The tri-spectral brightness temperature difference (BTD) method uses MODIS bands 29, 31 and 32 centered at 8.55, 11.03 and 12.02 μm respectively, allowing the algorithm to operate during both day and night overpasses. The visible near-infrared bands ratio (BR) method computes the reflectance ratio between one of two near-infrared channels (1.61 or 2.13 μm) and a visible channel (0.66 μm). This method augments the BTD algorithm during the daytime overpasses. In addition, the cloud phase algorithm incorporates the results of several cloud detection tests from the cloud mask algorithm, tests that are designed specifically to identify a particular cloud type (e.g. cirrus) (Platnick et al., 2003). Finally, a "sanity check" gets applied where all *uncertain* pixels with cloud top temperatures below -40^{o}C or above 0^{o}C get grouped under *ice* or *water* respectively (King et al., 2004; Chylek et al., 2006).

As with the cloud top algorithm, the cloud phase implements the MODIS cloud mask to identify the cloud covered pixels in its field-of-view (FOV). The 1-km pixels are then assembled into 5 x 5 boxes. If four of the twenty-five pixels in each box are determined to be either *probably* or *confident cloudy*, then the box receives a 'cloudy' label. The algorithm then computes the mean brightness temperature or reflectance values within that box and uses those to determine the final cloud phase category

(Platnick et al., 2003). Detailed flow charts for both the collection 004 and 005 cloud phase algorithms can be downloaded from *http://modis-atmos.gsfc.nasa.gov/MOD06_L2/atbd.html* (MODIS, 2008h). The basic scientific principles underlying the operation of the BTD and BR methods are discussed below.

The cloud phase yields five cloud groups: *ice*, *water*, *mixed*, *uncertain* or *clear*. The differences between *mixed* and *uncertain* appear to be arbitrary. According to Dr. Baum of the MODIS atmosphere science team, *mixed* and *uncertain* can be used interchangeably (B. Baum private communication, 2006). Therefore to simplify matters, the *uncertain* group has been combined with the *mixed* group.

The scattering and absorption of light by a medium is governed by the complex index of refraction, which is given by the equation (e.g. Menzel et al., 2002)

$$m = n_r - in_i \tag{2.11}$$

where m represents the complex index of refraction, n_r is the real part typically defined as the ratio of the speed of light through the medium over the speed of light in a vacuum, and n_i, the imaginary part, denotes absorption. Furthermore, the wavelength dependent absorption coefficient, κ, can be calculated directly from n_i as follows,

$$\frac{4\ n_i}{} \tag{2.12}$$

where λ is the wavelength of light (e.g. Baum et al., 2000). Knowing κ, the optical depth, , can be inferred.

The bulk measurements of n_i over the 7 – 13 μm range are given in Figure 2.3. Between 8 – 10 μm, absorption by liquid and ice is nearly identical. Beyond 10 μm, the absorption curves diverge and reach a maximum separation at around 11.5 μm. At this wavelength, absorption by ice is nearly five times greater than that of liquid water. This large discrepancy in n_i provides the theoretical foundation underlying the BTD method. First, the measured radiances in the 8.55, 11.03 and 12.02 μm bands for a cloud covered FOV are converted to brightness temperature. When the [8.55 – 11.03 μm] brightness temperature difference ($BTD_{8.5\text{-}11}$) is plotted as a function of ($BTD_{11\text{-}12}$), the *ice*, *water*

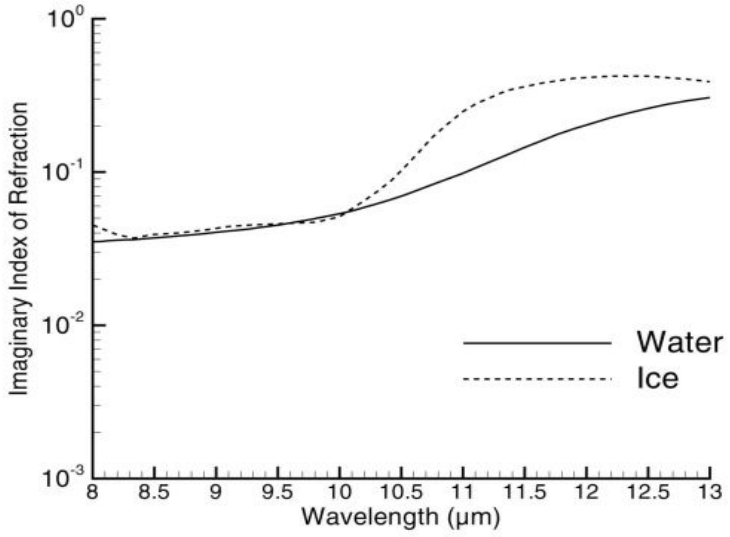


Figure 2.3: Imaginary refractive index of water and ice from 8-13 μm. Imaginary index of refraction curves over a planar surface of water (solid) and ice (dotted) in the thermal infrared**.** The x-axis gives the wavelength (m) and the y-axis provides the imaginary refractive index (n_i from Equation 2.11) on a logarithmic scale (Baum et al., 2000).

and *clear* pixels tend to cluster into distinct groups, with the *ice* clustering above a line of slope one and the *liquid* below. This behavior is illustrated in Figure 2.4, which shows the scatter-plot of the $BTD_{8.5\text{-}11}$ versus $BTD_{11\text{-}12}$ for a thin ice and a thick water cloud. Track A shows the distribution of BTD for a pure ice cloud over the central plains and Track D, the distribution for a pure water cloud over the Gulf of Mexico. This data was collected by the MAS sensor. Clear pixels tend to exhibit a negative $BTD_{8.5\text{-}11}$ because the emissivity of the surface at 11μm is usually larger than at 8.5μm (Baum et al., 2000).

Phase discrimination of mixed phase clouds can be extremely difficult. Figure 2.6 depicts the clustering behaviors of various cloud cover groups, which are represented

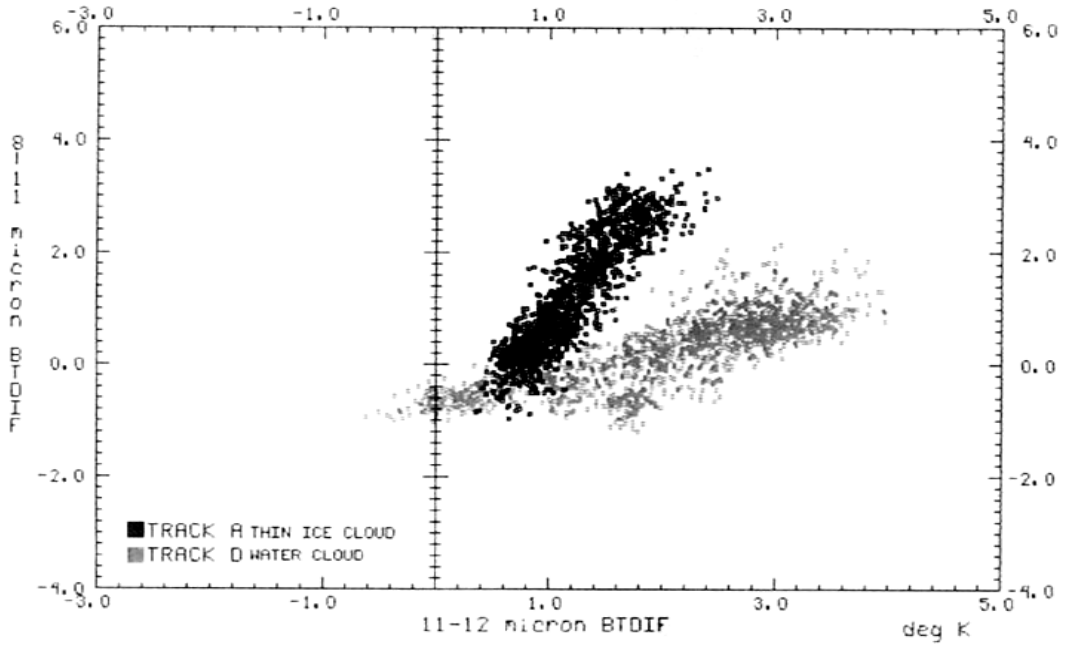


Figure 2.4: $BTD_{8\text{-}11}$ versus $BTD_{11\text{-}12}$ for an ice and water cloud. Scatter-plot of the $BTD_{8\text{-}11}$ versus $BTD_{11\text{-}12}$ pixels for an ice cloud (black dots) and a liquid water cloud (gray dots) collected during two MODIS Airborne Simulator (MAS) flights over the Great Plains (Strabala et al., 1994).

by the boxed regions in Figure 2.5 (Strabala et al., 1994). Under complex sky conditions, such as broken cloud or ice cloud overlying water cloud, groups can overlap with one another making it difficult to unequivocally determine the particle phase.

The BR method uses a similar approach to distinguish between ice and water clouds, however, it exploits the differences in n_i in the shortwave NIR, using MODIS bands six and seven centered at 1.64 and 2.13 μm as shown in Figure 2.7 (Baum et al., 2000).

Similar to the BTD principle, ice absorbs more radiation over these two spectral regions than liquid. A band ratio is calculated using either the 1.24 μm or the 1.64 μm channel with a visible channel. Taking the ratio reduces the dependence of the measured

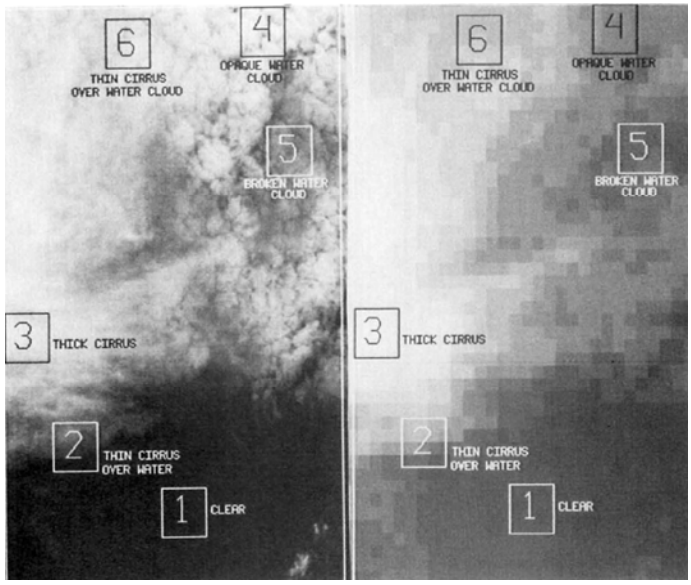


Figure 2.5: Six cloud types captured by the MAS instrument. A (left) visible channel 50 m resolution mixed cloud image derived from the MAS and its corresponding (right) IR image (11 μm) at 500 m resolution (Strabala et al., 1994).

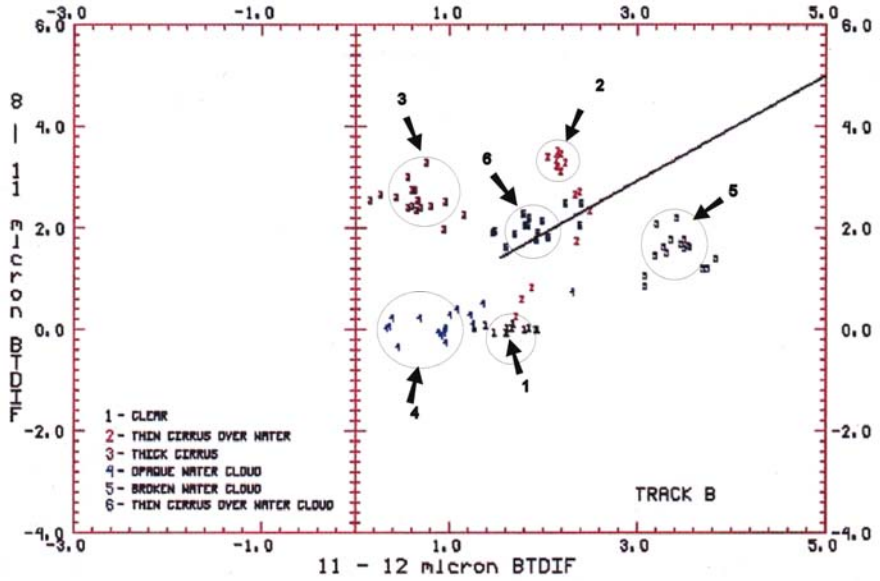


Figure 2.6: $BTD_{8\text{-}11}$ versus $BTD_{11\text{-}12}$ scatter-plot of six cloud types. The BTD scatter-plot corresponding to the six boxed regions represented in Figure 2.5. The main clusters are circled and the numbers reprinted for clarity (Strabala et al., 1994).

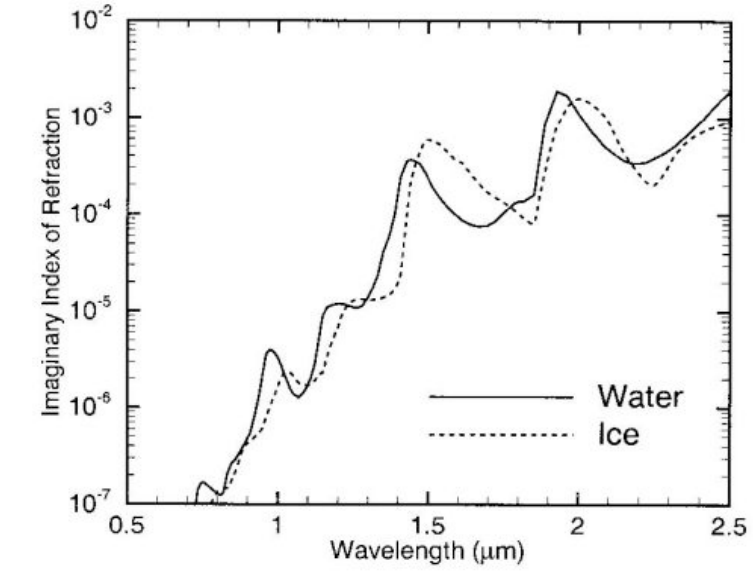


Figure 2.7: Imaginary refractive index for water and ice from 0.5 - 2.5μm. Imaginary refractive index (n_i) over a planar surface of water (solid) and ice (dotted) in the NIR**.** The plot gives the logarithm of n_i as a function of wavelength from 0.50 – 2.5 m (Baum et al., 2000).

radiances on particle shape, size, and number density, and MODIS channel 1 (0.645 μm) is often used because it offers the lowest signal to noise ratio (Chylek et al., 2006).

2.3.3.2 Cloud Thermodynamic Phase: Possible Sources of Error

The bulk scattering properties of clouds depend not only on phase, but also on the particles size distribution and, in the case of ice clouds, particle shape (Baum et al., 2000). Figure 2.8 depicts the $BTD_{8.5\text{-}11}$ and the $BTD_{11\text{-}12}$ for both water and ice clouds as a function of 11 μm brightness temperature derived from radiative transfer calculations (Baum et al., 2000). The water and ice clouds are assigned top temperatures of 284 K and 235 K. Four effective radius (r_{eff}) size distributions for the water and ice cloud models are plotted. Also, the 0.65 μm optical thicknesses are indicated by the numbers ranging from 0.20 to 20 respectively.

First, Figure 2.8 illustrates that there exist multiple solutions of BTD, which will depend on both the optical depth and on the particle size distributions. Note that the $BTD_{8.5\text{-}11}$ curves for the low-level water cloud (Figure 2.8a) all tend to be negative whereas these values all tend to be positive for the ice cloud (Figure 2.8c). This behavior helps to delineate the phase type. On the other hand, $BTD_{11\text{-}12}$ curves for both the water and ice clouds (Figure 2.8b, Figure 2.8d) are similar and are both positive. For mixed phase clouds or for mid-level clouds whose top temperatures fall within the super-cooled range ($-40^{o}C < T < 0^{o}C$), unambiguous phase discrimination becomes nearly impossible (Baum et al., 2000).

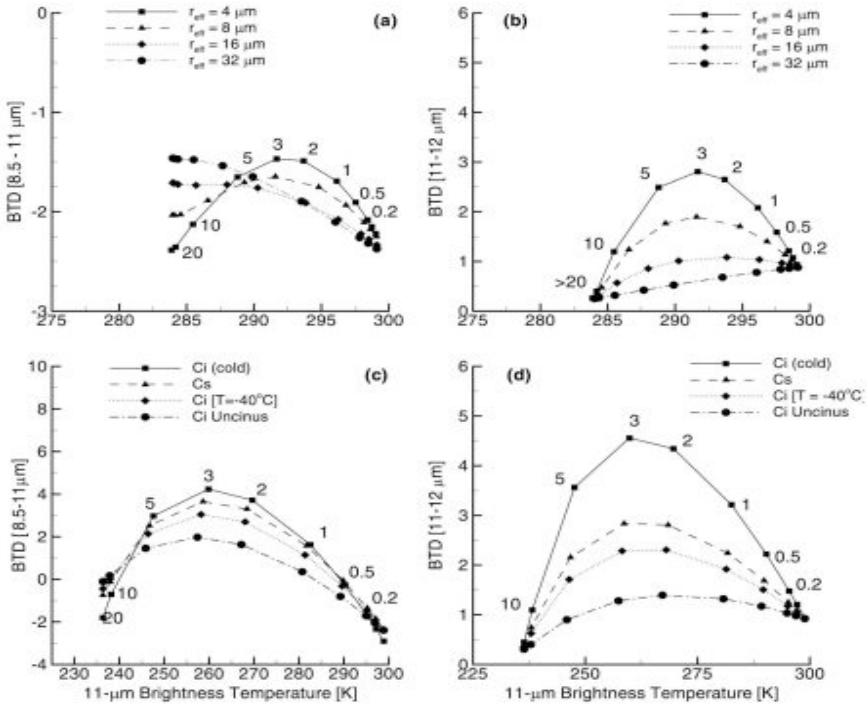


Figure 2.8: Theoretical BTD curves of various water and ice cloud scenarios. Curves of (a) $BTD_{8.5\text{-}11}$ versus BT_{11} for a water cloud, (b) $BTD_{11\text{-}12}$ versus BT_{11} for a water cloud (c) $BTD_{8.5\text{-}11}$ versus BT_{11} for an ice cloud, and (d) $BTD_{11\text{-}12}$ versus BT_{11} for an ice cloud. Each curve represents one of four particle effective radiuses (r_{eff}) or ice cloud types. The numbers next to the symbols signify values of 0.65 m optical depth from 0.20 to 20 (Baum et al., 2000).

Similar problems arise when using the NIR bands (Figure 2.9). In general, for a given optical depth, liquid clouds are more reflective in the 1.64 µm and 2.13 µm regions than ice clouds with reflectance being inversely proportional to particle size (Baum et al., 2000). However, there exists a great deal of overlap between the ice and liquid water solutions, especially in the 1.63 µm plot, which makes it difficult to conclusively differentiate liquid and ice clouds without knowledge their respective particle size distributions and the optical depth. To help minimize the effect of particle size and

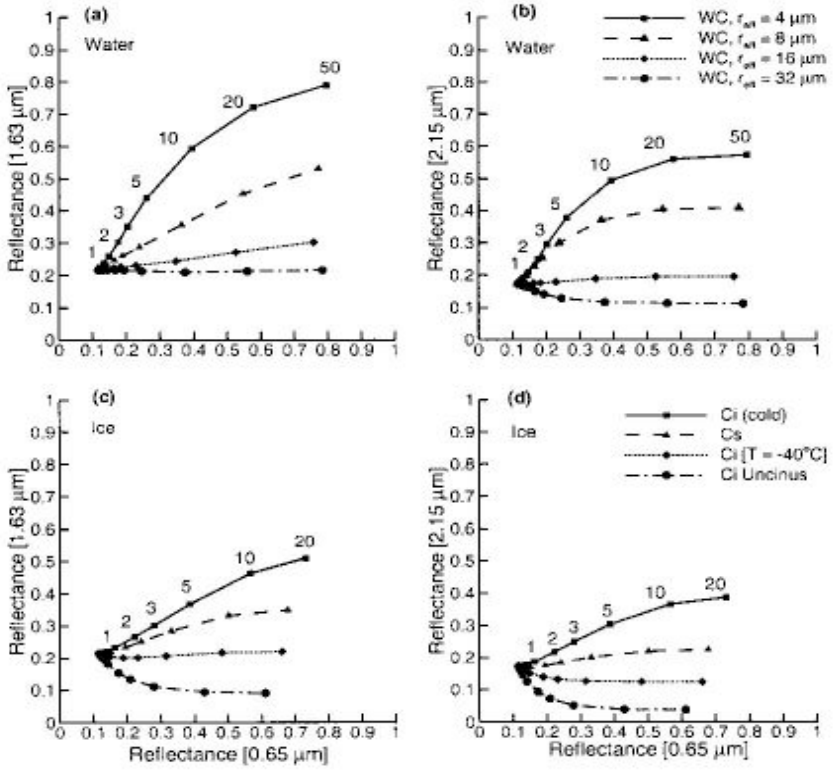


Figure 2.9: Theoretical curves of reflectance for various water and ice cloud scenarios. Model calculations of (a) 1.63 m versus 0.65 m reflectance of a water cloud, (b) 2.15 m versus 0.65 m reflectance for a water cloud, (c) 1.63 m versus 0.65 m reflectance for an ice cloud, and (d) 2.15 m versus 0.65 m reflectance for an ice cloud. As in Figure 2.8, each curve represents one of four droplet effective radiuses or ice cloud types; the numbers signify 0.65 m optical depths from 1 to 50 for water and 1 to 20 for ice (Baum et al., 2000).

shape, both the BTD and the BR methods take a ratio instead of relying solely on absolute radiance measurements (Baum et al., 2000).

Furthermore, phase detection of optically thin cirrus clouds is difficult, not only due to the problems associated with thin cirrus detection, but also due to the fact that a large component of the measured radiance is originating from either the surface or from underlying cloud layers. The problem of multi-layer cloud contamination on CO_2

derived cloud top estimates have not been completely resolved, but such issues are addressed by Baum and Wielicki (1994), Baum et al. (2000), Menzel et al. (2002) and Baum et al. (2003).

In addition, trace gas absorption by water vapor, ozone, methane, nitrous oxide and carbon dioxide can affect the radiance measurements and contaminate the NIR reflectance or the IR brightness temperature estimates. These effects can be mitigated using an atmospheric correction model that employs temperature and moisture profiles derived from reanalysis data or by assuming certain concentration profiles of trace gases in the atmospheric column.

Finally, as with any other cloud property extraction, if the pixel field-of-view is not entirely cloud covered, additional sources of radiation will contaminate the radiance measurements leading to the problem of pixel mixing.

2.3.3.3 Cloud Thermodynamic Phase: Validation Studies

To date, few validation studies have been performed to examine the MODIS cloud phase product. This is due, in part, to the difficulties associated with phase retrievals derived from remotely sensed data. The most accurate way to measure the water or ice content of a cloud is with an in situ device on an aircraft that can directly sample the cloud particles. However, these operations are expensive and thus cannot be conducted very frequently. Depolarization lidar techniques have proven effective at discriminating cloud phase (Sassen, 2005). Depolarization lidars emit light that is plane polarized and then measure both the horizontal and the vertical components of the light

backscattered by molecules, aerosols and cloud particles in the atmosphere. Light that gets scattered by solid ice particles is significantly more depolarized than light scattered by liquid droplets or aerosols. However, the shortwave light emitted by lidars gets completely attenuated when the optical depth is larger than three ($\tau > 3$). This restriction prevents lidars from probing very far into optically thick low and mid level water and mixed phase clouds. Other validation methods may include comparing the phase retrievals to known cloud types identified visually in the corresponding satellite image. Thin cirrus and low-level marine stratus have a distinct appearance, and can be easily identified in satellite imagery (Platnick et al., 2003).

The BTD technique was first evaluated in 1991 using the MODIS Airborne Simulator (MAS) data collected on board a NASA ER-2 aircraft during the First ISCPP Regional Experiment (FIRE) over the Oklahoma (Strabala et al., 1994). Examination of the $BTD_{8.5\text{-}11}$ versus the $BTD_{11\text{-}12}$ first revealed the clustering behaviors of ice cloud, water cloud and clear pixels in the scatter-plot (see Figures 2.4 – 2.6), which led to the development of the current MODIS BTD algorithm.

The inclusion of the BR method as a means to augment the BTD was first carried out in 1996, again using the MAS instrument during the Subsonic Aircraft Contrail and Cloud Effects Special Study (SUCCESS) (Baum et al., 2000). A case study demonstrated that the BR method helped reduce the number of *uncertain* and *mixed* pixels classified by the BTD algorithm, and improved the detection of thin cirrus clouds with an optical depth less than two ($\tau < 2$).

Platnick et al. (2003) present several cloud product datasets for a scene over the west coast of South America from the MODIS Terra instrument on July 18, 2001. The scene contained a mixture of cloud types that were evident in the visible satellite image, including low level marine stratocumulus, high level cirrus and deep convective clouds. In general the MODIS cloud phase algorithm could delineate the low level marine clouds as water and the high level cirrus and areas of deep convection as ice. However, many pixels got classified under *mixed* or *uncertain*, and were subsequently determined to be regions of thin cirrus overlying low-level water clouds.

Chylek et al. (2006) compared the performance of the BTD method with the BR method for five separate cloud scenes. Their findings indicate that the BR algorithm consistently overestimates the amount of ice pixels, even some "warm" pixels possessing 11 μm brightness temperatures above 275 K. They conclude, however, that both algorithms, when used in conjunction with a temperature "sanity check" provide the best results. The "sanity check" ensures that all pixels with cloud top temperatures above 273 K get classified as *water* and all pixels below 233 K get classified as *ice*.

More recent phase determination methods employ additional water vapor absorption channels into the algorithm that help determine the temperature and moisture profiles of the atmosphere above the cloud layers (Spangenberg et al., 2006). The technique, called the Multi-spectral Mixed phase Detection Technique (MMDT) appears to delineate mixed phase clouds with a high degree of accuracy. Observations of cloud phase obtained from ground, radiosonde and airborne instruments at the North Slope of Alaska (NSA) site near Barrow, AK, and also collected during the Mixed-Phase Arctic

Cloud Experiment (MPACE) were used to evaluate the MMDT method. This new method assumes that different cloud types, be it ice, mixed or liquid, exhibit specific and unique temperature and moisture signatures in the atmospheric column directly above the cloud. Overall, the MMDT cloud phase classification matched the ground-based results 90, 86 and 84% of the time for the liquid, mixed and ice cloud phases respectively. Furthermore, 62 of the 69 liquid cases were of super-cooled liquid droplets. Due to its success, the MMDT method will likely be incorporated into future MODIS cloud phase algorithms.

Finally, the aforementioned lidar depolarization technique, is also being implemented with other field campaigns to validate cloud phase. The University of Wisconsin High Spectral Resolution Lidar (HSRL) with polarization capabilities has been deployed to several study sites to validate various cloud properties, including particle phase for optically thin clouds (Holz et al., 2001; Eloranta and Uttal, 2006).

2.3.4. Aerosol Type

The following discussion briefly outlines the operation of the MODIS aerosol detection algorithm. The information presented here is obtained from either the *Algorithm For Remote Sensing of Tropospheric Aerosol From MODIS* (Kaufman and Tanré, 1998) or from the 2006 revision of this ATBD document, which elaborates on the algorithm modifications applied to the collection 005 (C005) aerosol dataset (Remer et al., 2006). For this research, the C004 MOD04 files are used, because it appeared the C005 aerosol type dataset was yielding erroneous results. Also, validation studies

performed specifically to fingerprint different aerosol *types* in the troposphere are practically nonexistent since researchers are more interested in the bulk scattering characteristics. Hence the *Validation Studies* and *Sources of Error* sections included with the earlier algorithm descriptions are omitted from this discussion.

The MODIS atmosphere team generates an operational level-1 10-km aerosol product (MOD04), which provides several key climate related radiation parameters such as aerosol optical depth, mass concentration and particle size distribution. The aerosol algorithm implements two entirely separate procedures, one devoted strictly to detecting aerosols over the ocean, and the other over land. The land algorithm is limited to daytime overpasses over dark, snow and ice-free backgrounds a method known as the "dark target approach".

Initially, a set of look-up tables (LUTs) is generated from radiative transfer calculations utilizing various pre-existing *aerosol models*. These models are given in Table 2.3. Each model assumes a particular aerosol type and a set array of physical and optical parameters such as particle size distribution, refractive index, asymmetry factor, single scattering albedo, and particle sphericity. The tenth SDS in the MOD04 file yields five aerosol types: *mixed*, *dust*, *sulfate*, *smoke*, *heavy_smoke*. These are chosen based on a geographical zone and season using an aerosol climatology derived mostly from a network of automated surface observations using sun-photometers, for example the AErosol RObotic NEtwork (AERONET) . The distribution of aerosol types is provided in Figure 2.10. However, it is not apparent exactly how the five MODIS aerosol types provided with the MOD04 file match the specific aerosol types in Table 2.3

Table 2.3: **MODIS aerosol models.** Outline of the fundamental aerosol models employed by the MODIS aerosol algorithm. These models are derived from extended surface observations. The models are separated into four main types and are subdivided by subtype. The parameters are as follows r_g, mean radius; r_v, volume mean radius; , standard deviation of the lognormal size distribution; V_o, volume of particles per cross-sectional area; $_o$, single scattering albedo (Kaufman and Tanré, 1998).

Continental aerosol	r_g(μm)	r_V(μm)	σ	$V_o(10^6\ cm^3/cm^2)$	ω_o (670nm)
Water Soluble*	0.005	0.176	1.090	3.050	0.96
Dust Like	0.500	17.60	1.090	7.364	0.69
Soot	0.0118	0.050	0.693	0.105	0.16
Biomass burning	r_g(μm)	r_V(μm)	σ	$V_o(10^6\ cm^3/cm^2)$	ω_o(670nm) †
Accumulation	0.061	0.130	0.500	$-2.4+45\tau$	0.90
Stratospheric	0.380	0.510	0.310	0.984	0.98
Coarse	$1.0-1.3\tau$	$6.0-11.3\tau+61\tau^2$	$0.69+0.81\tau$	$2.4-6.3\tau+37\tau^2$	0.84
Industrial/urban aerosol	r_g(μm)	r_V(μm)	σ	$V_o(10^6\ cm^3/cm^2)$	ω_o (670nm)
Accumulation 1	0.036	0.106	0.60	$-2.0+70\tau-196\tau^2+150\tau^3$	0.96
Accumulation 2	0.114	0.210	0.45	$0.34-7.6\tau+80\tau^2-63\tau^3$	0.97
Stratospheric	0.430	0.550	0.29	0.73	0.98
Salt	0.990	1.300	0.30	$-0.16+4.12\tau$	0.92
Coarse	0.670	9.500	0.94	1.92	0.88
Dust aerosol	r_g(μm)	r_V(μm)	σ	$V_o(10^6\ cm^3/cm^2)$	ω_o (670nm)
Dust Background					
1st mode	0.0010	0.0055	0.755	$6.0x10^{-6}$	0.015
2nd mode	0.0218	1.230	1.160	1.0	0.95
3rd mode	6.2400	21.50	0.638	0.6	0.62

Initially, several 0.47, 0.66, 2.1 and 3.8 μm reflectance threshold tests and a preliminary aerosol optical thickness determination are made using the 0.47 and 0.66 μm channels. Based on these results, the algorithm assesses which aerosol *model*, and thus LUT, is appropriate for the given sensor measured radiances. According to the 1998 ATBD, the algorithm "fingerprints" the aerosol type by examining the ratio of the aerosol single scattering path radiance in the red channel, $L_{p\text{-}red}$, with the single scattering path radiance

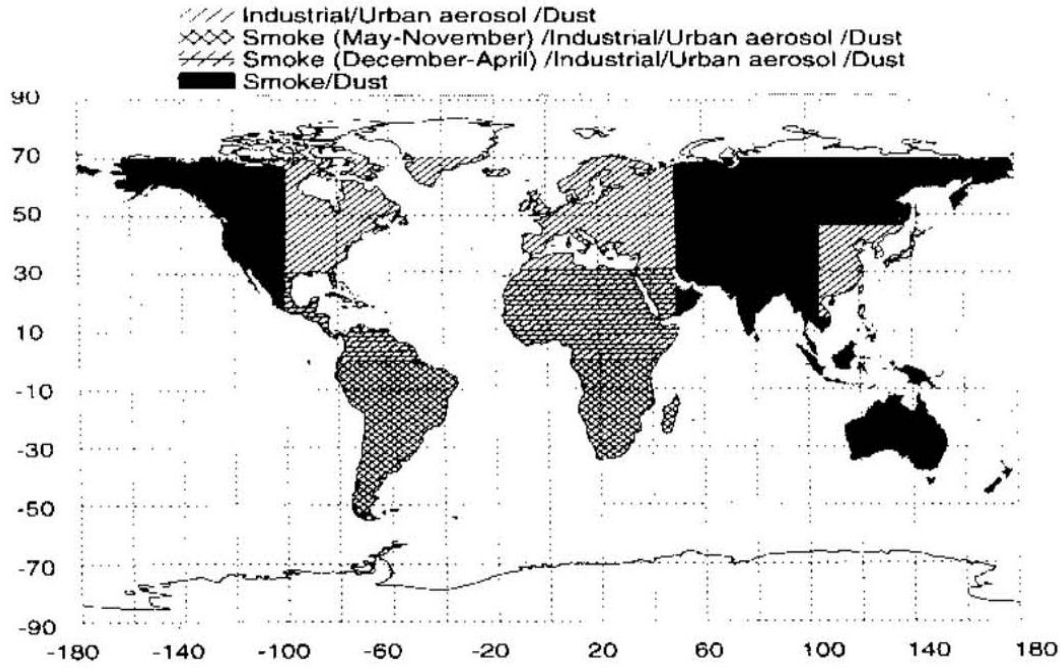


Figure 2.10: **MODIS aerosol types distributed by region and season.** Map representing the geographical distributions of aerosol types employed by the MODIS aerosol detection algorithm for land. This map is developed from in situ aerosol observations (Kaufman and Tanré, 1998).

matched with the radiances in the LUT and the physical and radiative aerosol parameters in the blue, $L_{p\text{-}blue}$. The definition of path radiance is given in section 2.1. After the correct model and LUT are chosen, the initial optical thickness calculations from the satellite measurements are modified to fit the model. Then the measured radiances are (e.g. particle size distribution) used in the LUT calculations are assigned to the pixel (Kaufman and Tanré, 1998).

The aerosol retrieval algorithm tries to account for cloud cover, cirrus contamination, water vapor content, and stratospheric aerosol concentrations. Exactly how the algorithm screens for clouds and handles cloud contamination is somewhat

unclear, and was only briefly mentioned. Apparently, the MODIS cloud mask is not used; an independent spatial variability cloud test gets applied. The radiative effects from cirrus clouds are corrected for using the 1.38 μm channel. However, as discussed in section 2.3.1.1, this method only works when there is a sufficient amount of water vapor absorption in the lower atmosphere. Issues of mistaken identity between aerosol and clouds can be expected, such as the problem of misidentifying cirrus cloud as dust (Roskovensky and Liou, 2005). Water vapor concentrations are estimated using a 2.1 μm absorption band and a solar band. Also, attenuation by stratospheric aerosols is ignored unless a major volcanic eruption has recently occurred.

The primary function of this research with respect to aerosol product is to find out how the MODIS aerosol type dataset (SDS #10) provided with the 10-km aerosol product (MOD04) match the aerosol types documented from the ground and lidar observations. Although this comparison is admittedly rudimentary, aerosol type is nonetheless somewhat important in verifying, first, if the appropriate aerosol model even exists and second, if the correct model is being chosen. If, for example, the MODIS derived aerosol type consistently classifies dust (e.g. Asian dust) as smoke, then it could be important to recognize this error since dust and smoke have different shapes, size distributions and chemical compositions and thus, attenuate radiation differently. At this stage, one can only speculate how the misidentification of aerosol type would trickle down into the cloud and aerosol algorithms.

As of the summer of 2007, the latest version of the aerosol product available for download was the collection 005 (C005) data. According to the latest version of the

aerosol theoretical basis document, the algorithm applied to land is a "complete overhaul" from the previous algorithms, incorporating new channels and threshold tests (Remer et al., 2006). Also, dust remote sensing over land is performed in the C005 data whereas before, dust detection was limited to the ocean surfaces. However, after ordering and extracting the C005 aerosol type determinations over AFARS, it was evident that something was amiss with the newest version; while the C004 data correctly identifies *smoke* or *heavy_smoke* for numerous smoke observations during the summer months, the C005 data yielded significantly fewer aerosol type identifications, and the only aerosol type returned was *sulfate* which presumably is associated with the "urban" aerosol model. Therefore, the MODIS-lidar aerosol type comparison is limited to the 341 cases of C004 data collected from February of 2004 to June of 2006. C004 data is unavailable beyond June of 2006.

2.4. Fundamentals of Active Remote Sensing: Lidars

LIght Detection And Ranging (Lidar) instruments have proven to be a valuable research tool for studying the atmosphere. Lidars emit pulses of shortwave nearly monochromatic laser light and measure the component of the light scattered into the sensor's field-of-view (FOV) by the atmospheric constituents along the path of the beam. Simple elastic-backscatter lidars can discern the location and microphysical characteristics of clouds, precipitation, and aerosol layers. More sophisticated instruments can provide atmospheric profiles of temperature, density and trace gas concentrations. Since the visible and near-infrared wavelengths of lidar laser light cannot

penetrate much beyond an integrated optical depth, τ, larger than three, probing the internal structure of thick clouds simply is not feasible. However, lidars can thoroughly penetrate aerosol and optically thin cloud layers.

The invention of lasers (*l*ight *a*mplification by *s*timulated *e*mission of *r*adiation) in 1960 greatly accelerated the field of lidar research (Wandinger, 2005). Lidars employ laser technology to generate the concentrated nearly monochromatic beam of intense light required for research applications. As the name implies, lasers work by exploiting the principle in quantum mechanics of stimulated emission. A simple instrument design contains a cylindrical cavity which holds a laser rod composed of glass which has been "doped" with some kind of metallic element. Surrounding the rod is a flash lamp which is able to generate extremely powerful and bright pulses of white light. The cavity walls are lined with a reflective material in order to maximize the intensity of emitted light from the flash lamp. Two mirrors are positioned on the rear and front sides of the cylinder, the front mirror being partially transparent.

The principle of "stimulated emission," first discovered by Albert Einstein, can be summarized as follows: The flash lamp "pumps" white light into the cavity. The outer electrons of the metallic atoms suspended within the crystal lattice of the laser rod absorb some of the photons, creating a population inversion of excited electrons which have jumped to higher more unstable energy levels. When some of these electrons fall back into their ground state, they "spontaneously" emit photons of quantized light which travel back and forth between the two mirrors. As these photons pass through the doped laser rod, they prompt other excited electrons to fall back into their lowest energy state,

thereby releasing even more photons in the same direction and having the same polarization as the inciting photons. This cascading effect of "stimulated emission" generates a powerful pulse of nearly monochromatic linearly polarized light characteristic of lasers.

A diagram of a simple lidar set-up is illustrated in Figure 2.11. Lidars usually possess a bi-static design, meaning that the transmitter and the receiver are separated by some finite distance (Wandinger, 2005). Once the polarized beam of light exits the laser cylinder, it typically passes through several optical devices, such as a beam expander and collimator, which help to focus the beam and minimize the amount of divergence. The receiver consists of a reflecting Newtonian telescope which collects scattered light and projects it into a pinhole diaphragm through photo-multiplier tubes to the detector. The detector measures the intensity of the light and stores the information on a computer. Polarization lidars are equipped with an additional detector (Figure 2.12) and a beam splitter crystal. The crystal allows only the horizontally polarized light to pass, and

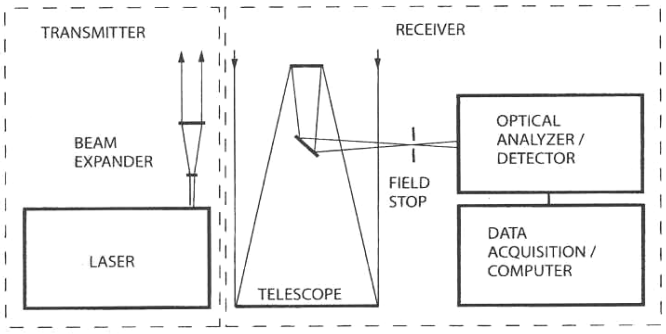


Figure 2.11: **A simple schematic of a lidar set-up.** A basic lidar system illustrating the transmitting laser box, the Newtonian telescope receiver, and the data acquisition system (Wandinger, 2005).

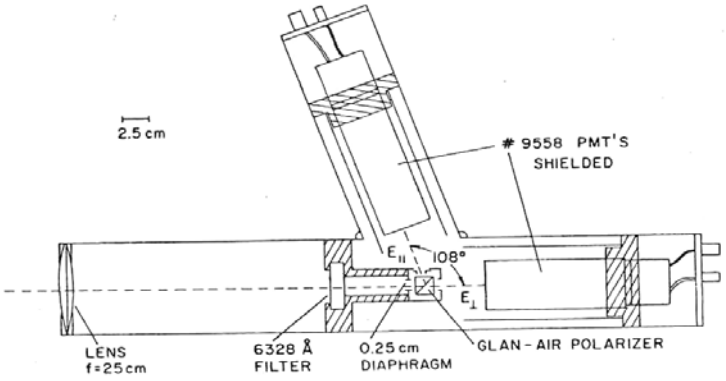


Figure 2.12: Schematic of a dual-channel polarization lidar. The receiver contains a glan-air polarizer that splits the incident beam into its parallel and perpendicular components, and diverts these signals to separate photo-detectors (Sassen, 2005).

reflects the vertically polarized light to a second detector.

The "vertical" resolution of the lidar is determined by the pulse length, h, which is computed by multiplying the pulse time, τ (not to be confused with optical depth), with the speed of light, c. The vertical depth of the instantaneous scattering volume is actually h/2, which is typically on the order of tens of meters. The maximum range of a lidar, R_{max}, is controlled by the pulse repetition frequency (PRF), defined as the number of pulses per second, and can be computed by the simple formula:

$$R_{max} = c/2PRF. \quad 2.13$$

The most basic form of the lidar equation can be expressed as (Wandinger, 2005)

$$P(R) = K*G(R)*\beta(R)*T(R). \quad 2.14$$

The term, P(R), represents the power received by the instrument originating from range, R. In the case of polarization lidar systems, P(R) can be broken down into the two constituent parallel, $P_{||}$, and perpendicular, $P_{\perp}$, polarization planes. The last two terms, β and T, represent the wavelength and polarization dependent backscattering and transmission terms, the two atmospheric unknowns. The backscatter coefficient, β, governs how much energy gets scattered back towards the receiver and depends upon the type, number density, and backscattering cross section of the particles. The transmission term, T, governs the exponential loss of light via extinction by molecules and aerosols along the path between the receiver and the target. T can be expressed as follows,

$$\exp\left(2 \int_0^R (R)\ dR \right), \qquad 2.15$$

which is similar to Equation 2.4 for atmospheric transmissivity. The σ term is the extinction coefficient, η represents the multiple scattering adjustment factor, and the expression must be multiplied by two to account for the total distance traveled by the beam between the target and the receiver.

The first two terms in Equation 2.14 are controlled operationally. More specifically, the performance factor, K, can be expressed as a product of the transmitted power, P_o, the instantaneous scattering volume, $c\tau/2$ described in the previous section, the physical area of the receiving antenna or telescope, A_t, and the system efficiency, η not to be confused with the multiple scattering adjustment in Equation 2.15. All lidars, to a certain degree, cope with a "blind-spot" directly above the receiver caused either by a

misalignment between the laser beam and the telescope's FOV, or by the instrument's optical limitations at being able to focus at close range. This effect, contained within the G(R) term, is explicitly expressed by the geometric overlap function O(R) divided by R^2. The overlap function will vary from instrument to instrument, and will depend on the beamwidth and divergence of the laser, the field of view of the telescope, and the distance between the transmitter and receiver.

Basic scattering rules dictate that the manner in which particles interact with radiation is governed by the size parameter, α, defined as

$$\alpha = 2\pi r/\lambda \qquad 2.16$$

where r is the particle radius and λ is the wavelength of incident light. For cases where $\alpha \leq 0.30$, Mie scattering theory can be simplified down to the Rayleigh approximation, where β is proportional to D^6 and λ^{-4}, and where extinction, σ, is proportional to D^3 and λ^{-1} (Lhermitte, 2002). When the particle sizes are much larger than incident wavelengths ($\alpha \geq 10$), the simple rules of geometric ray tracing apply. For such cases, both β and σ are proportional to the geometric cross-sectional area of the particles and are independent of wavelength. Between these extremes, when the particle size and incident wavelengths are of a similar size, β and σ can fluctuate wildly due to complicating effects from constructive and destructive interference. Figure 2.13 illustrates these behaviors. With the exception of the 10.6 μm CO_2 laser, most lidars operate between the 0.25 to 1.064 μm spectral range, placing them well within the Rayleigh approximation with regard to

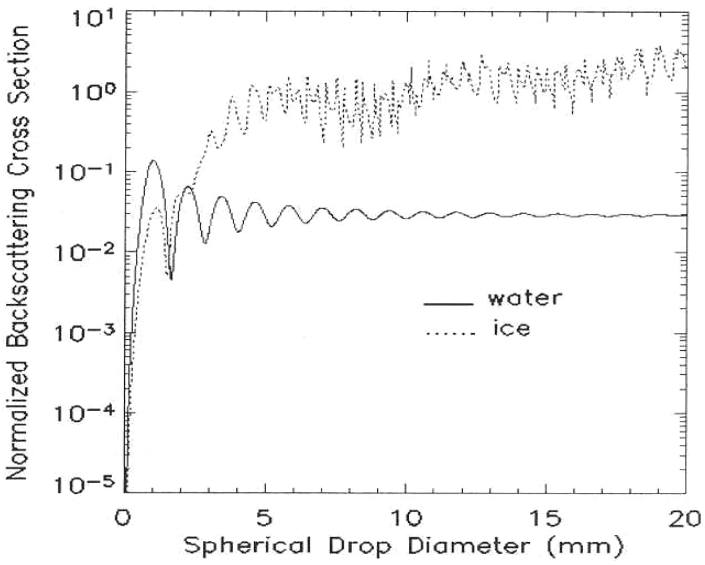


Figure 2.13: Normalized W-band radar backscattering cross sections of water and ice as a function of drop diameter. Backscattering cross sections for a W-band radar (λ = 0.32 cm) for ice (dotted) and liquid (solid) water spheres roughly demonstrating the Rayleigh approximation ($d < 0.5$ mm), Mie scattering ($0.5 \text{ mm} \leq d \leq 15$mm) and the geometric optics region ($d \geq 15$ mm) (Sassen et al., 2005).

molecular scattering, and within the geometric optics domain with regard to most cloud, precipitation and aerosol particles.

2.5 Polarization Lidar Technique

One of the most important advances in the realm of lidar research has been the development of the polarization technique. Lidars naturally emit linearly polarized light, which means that the orientation of the electric field oscillates in a fixed plane. When this light gets scattered, the orientation of the E-vector can change depending on the shape, size and the refractive index of the particles that interact with the beam (Sassen and Mace, 2002; Sassen, 2005). Spherical particles, like liquid cloud droplets, do not

modify the orientation of the E-vector in single scattering. Conversely, solid ice or dust particles can reorient the direction of the E-vector a great deal. As a result, the thermodynamic phase of a cloud can be distinguished unambiguously. Following the location and the evolution of the liquid and ice layers in real-time can provide much insight into the dynamical processes that are occurring within the cloud. This is especially useful for understanding precipitation or how embedded aerosol layers can influence cloud microphysical processes (Sassen, 1991, 2002b; Sassen et al., 2003).

The P(R) term in Equation 2.14 can be deconstructed into its parallel and perpendicular components, simply by specifying the polarization plane of the backscatter coefficient, $\beta_{\parallel}$ or $\beta_{\perp}$. The polarization lidar technique involves the calculation of the depolarization ratio, δ, which is defined as the ratio between $P_{\perp}/P_{\parallel}$. The term, δ, simplifies to $\beta_{\perp}/\beta_{\parallel}$, ignoring any polarization dependence in the transmission term (Sassen and Mace, 2002). The depolarization ratio is the key element underlying the polarization technique, and is found directly by taking the ratio of the two measured power terms.

Ray tracing models in Figure 2.14 illustrate how the orientation of the E-field vector can change when a photon interacts with a particle. In the case of spherical droplets, single scattered photons that are perfectly backscattered ($\theta = 180^{o}$) can follow either one of three paths; an axial reflection off the front or rear inner wall of the drop, or a circumnavigation around the peripheral edge. None of these three interactions alter the

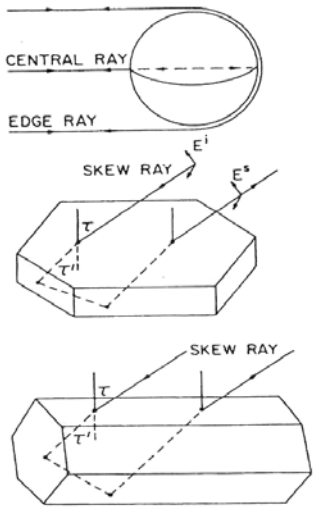


Figure 2.14: Three ray tracing models demonstrating changes to the polarization of a light beam. A schematic depiction of how cloud particle shape influences changes in polarization of an incident beam of light for a (top) spherical droplet (center) hexagonal plate and (bottom) a hexagonal column. The polarization of the incident and exiting beam are identical in the top scenario but internal refractions change the polarization of the exiting beam in the center and bottom (Sassen, 2005)**.**

incident orientation of the E-field, generating δ values that are essentially zero. However, as the beam penetrates deeper into a cloud, the probability of encountering multiple-scattered photons increases. Typically, δ values near the base of a water cloud are zero, but exhibit a sharp rise due to multiple scattering immediately before the signal becomes completely attenuated.

For the case of solid ice crystals, which come in a variety of habits (shapes) and orientations, photons can undergo a series of refractions and internal reflections, causing the orientation of the backscattered E-vector to bend. Ray tracing calculations and field experiments have shown that the amount of depolarization will depend upon the shape, orientation, and particle shape ratio defined as the ratio between the length of the crystal

and twice its base (Noel et al., 2002). An important exception to the rule is the case of flat plate crystals which fall with their basal faces nearly horizontal. The photons from zenith pointing lidars undergo specular reflections off the flat base of the plate which does not change the incident orientation of the E-field (Sassen, 2005). By pointing the laser a few degrees off zenith, the depolarization values will readjust back to levels more appropriate of ice crystals.

Depolarization due to pure molecular scattering tends to be less than 5%, while aerosols may exhibit a range in δ values depending on their size and composition. For example deliquesced haze particles and smoke exhibit very low δ values (Sassen et al., 2005), which compares with large irregularly shaped dust particles that can produce δ values exceeding 0.20 (Sassen et al., 2003).

2.6 Polarization Lidar Field Observations

The use of polarization lidars to study clouds began in earnest during the 1970s (Sassen, 1991). Field work validated by lab experiments demonstrated that δ values for spherically shaped cloud particles fell around zero, whereas irregularly shaped ice particles could be expected to generate large δ values. Cloud chamber experiments that can create ice crystals of a known shape demonstrated that the amount of depolarization was related to the habit (shape) and orientation of the crystal (Sassen, 1991). Other studies have confirmed a link between cirrus cloud temperature and δ values, which provide evidence for the discrimination of crystal habit using the polarization lidar technique (Sassen and Benson, 2001).

Figure 2.15 shows the range normalized backscattering lidar $P_{\perp}$ (white areas) and $P_{||}$ (hatched areas) profiles for a drizzle producing marine stratus cloud, a continental mixed phase cumulus cloud and an ice-phase cirrus cloud respectively (Sassen, 1991). The liquid marine stratus cloud exhibits low depolarization values in both the drizzle and cloud regions, typical of spherically shaped targets. The small δ-values that are observed, especially in the cloud base near 1.95 km, are due to multiple scattering effects. In the rain region of the continental cumulus cloud (center), depolarization ratios are predictably low. However, a sharp spike in both the backscattered power and in the depolarization ratio is observed near the 3.3 km level. This "bright band" anomaly is associated with the phase transition region known as the melting layer. The cause and position of the "bright band" near and below the 0^{o}C isotherm varies depending on the type and frequency of the

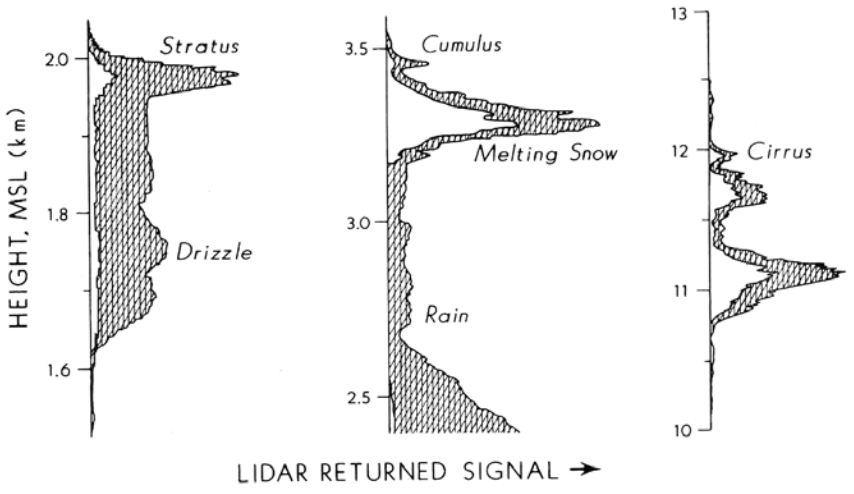


Figure 2.15: Three profiles of parallel and perpendicular backscatter. Profiles of returned power for the parallel (hatched) and perpendicular (clear) polarization lidar channels for a liquid dominated marine stratus cloud (left), the precipitation region below a mixed phase continental cumulus cloud (center), and a an ice phase cirrus cloud (right) (Sassen, 1991).

instrument. In this case, it is likely that snowflakes with large cross-sectional areas near the top of the layer are providing bright scattering targets for the lidar. As the flakes melt, they shrink and their number densities decrease due to an increase in their fall velocities, contributing to a reduction in the backscattered signal. The optically thin cirrus cloud comprised of ice crystals exhibits consistently high δ values throughout the entire layer, although δ values in the 11.5 – 12 km region appear to be larger than in the lower 10.7 – 11.3 region, which could reflect differences in the ice crystal shape between the colder cloud top and the warmer cloud base.

Another feature associated with the melting layer is the lidar "dark band" (Sassen et al., 2005). This feature is commonly observed where small ice fragments remain embedded in raindrops near the bottom of the melting layer. The ice nuclei absorb some of the light and block axial reflections off the back wall of the raindrop, thereby creating the characteristic dip in backscatter of the lidar "dark band."

An interest in cloud seeding experiments during the 1980s lead to research on mixed phase alto-cumulus clouds in mountainous regions. One such study used a polarization lidar, Ku-band radar and a dual channel microwave radiometer to monitor the structure of an orographic mixed phase cloud in Colorado before and during a frontal passage (Sassen, 1984). The combination of all three instruments provided a thorough assessment of the evolution of the internal structure of the cloud as it developed. Polarization lidar measurements observed layers of super-cooled liquid droplets consistently at temperatures well below -30^{o}C. During one twenty minute episode, the uniform super-cooled liquid layer, which was present during most of the time,

temporarily disappeared due to sedimenting ice particles from above. This lidar observation captured the well-known Bergeron-Findeisen process whereby differences in the equilibrium vapor pressure over liquid water and ice causes the ice particles to grow at the expense of the liquid particles, until the liquid droplets completely evaporate (Wallace and Hobbs, 1977). Also, extremely high δ-values below the liquid layer revealed the presence of rimed ice particles, likely created from the descending snowflakes through the super-cooled liquid layer.

Other polarization lidar studies have captured the so-called indirect aerosol effect. For months following the eruption of Mt. Pinatubo in 1991, the aerosol laden stratosphere provided a rich source of sulfates to the troposphere. One study, conducted during December of 1991, captured a tropopause fold induced by a strong jet stream, which allowed stratospheric aerosols to seed a cirrus cloud in the troposphere (Sassen et al., 1995). Lidar observations of unusually large haze particles and of unusually high concentrations of ice crystals near the tropopause fold revealed various heterogeneous cloud forming mechanisms at work associated with the volcanic aerosol.

More recently, polarization lidars have captured the abnormally active ice nucleating properties of desert dust. Unusually warm cirrus clouds over Utah with top temperatures exceeding -20^{o}C were observed in association with an Asian dust layer, during which a liquid altocumulus cloud became glaciated as it came in contact with the dust layer (Sassen, 2002b). Other observations associated with Saharan dust storms indicated unusually strong heterogeneous freezing processes occurring in clouds with top temperatures as warm as -8^{o}C (Sassen et al., 2003). The shape of desert dust mimics that

of ice crystals which makes them suitable ice nuclei. This is confirmed by the polarization lidar data, which shows relatively large δ-values ($\delta > 0.25$) associated with dust plumes (Sassen, 2005). Observations of the indirect aerosol effect could not be conducted remotely without the use of the polarization lidar technique.

2.7 Sensor Specifications: Polarization Cloud Lidar (PCL)

The data used to compare with the MODIS products is acquired by the Polarization Cloud Lidar (PCL), informally known as the "ruby lidar" because utilizes a ruby crystal to emit a ruby colored ($\lambda = 0.694$ μm) laser beam. The PCL is one of several manually operated surface-based instruments used to examine clouds and aerosols located at AFARS. PCL data has been collected regularly since February of 2004, and can be found, either in its raw form or as a plot of backscattered power or depolarization ratio, at *http://rainbow.gi.alaska.edu* (Sassen and Zhu, 2008a). Prior to 2004, the PCL operated at the University of Utah Facility for Atmospheric Remote Sensing (FARS) located in Salt Lake City, UT. The Utah data, collected between 1992 and 2002, is also available at the rainbow server.

The PCL shoots vertically polarized pulse of ruby colored (0.694 μm) light approximately once every ten seconds. The maximum energy and the pulse length are 1.5 Joules and 27 nanoseconds respectively. This translates to a sampling resolution (vertical bin size) of about 8.0 m. The data collection algorithm averages every six consecutive shots corresponding to a 1-minute average, which reduces the noise in the

signal. The instrument measures the backscattered signal up to an altitude of 14 km above mean sea level (MSL).

The PCL possesses two photo-detectors, one that measures the parallel component of the backscattered light and a second that measures the perpendicular component. The ratio of the latter to the former provides the depolarization ratio (δ-values). Applications of the depolarization ratio in discriminating cloud particle phase have already been discussed in the preceding section. The instrument is secured to a platform that allows the viewing angles to be adjusted manually by the operator. This is useful when trying to discriminate oriented plate crystals from liquid droplets, when the instrument needs to be tilted a few degrees off nadir to check for changes in the δ values. Instances in which the instrument is tilted away from zenith are denoted by hatch marks (x's) along the top of the plotting window in both the backscattered power and depolarization ratio images. Additional information regarding the PCL and about the AFARS in general can be viewed at: *http://www.gi.alaska.edu/AtmosSci/Research/afars_site.html* (Sassen and Zhu, 2008b).

The next chapter delves into the methodology of this research. This section covers the MODIS data extraction procedures, lidar data interpretation methods, and data interpolation using radiosonde profiles. In addition, a preview of the results is provided in order to demonstrate precisely how the MODIS and lidar datasets get evaluated in chapter four.

Chapter 3 Methods

3.1 Data Sources

The primary objective of this study is to evaluate the performance of the MODIS cloud mask, cloud phase, cloud top properties (temperature, pressure, altitude) and aerosol type against ground based measurements from a dual-channel polarization lidar (PCL). These datasets are extracted from the following MODIS products: MOD04, MOD06 and MOD35 (prefix 'MYD' for the Aqua products). On occasion, a true color image is created from a MOD02 level-1B radiance file using the ENVI (Environment for Visualizing Images) image processing and visualization software. The determination of cloud base and top altitudes, phase and aerosol layers was conducted through visual inspection of the lidar plots. The lidar data is collected over a 3.5-year period beginning in February of 2004 and ending August of 2007. Jiang Zhu manages the rainbow web server, and creates the lidar plots using an Interactive Data Language (IDL) code. IDL is a data analysis programming language with plotting capabilities. To facilitate the inspection of the lidar plots, the plotting codes were slightly modified in order to superimpose the times of the Aqua and Terra overpasses.

To obtain the concurrent MODIS data product, the precise date, start and end times of each lidar acquisitions, as well as the spatial coordinates of AFARS, are queried into the Level-1 and Atmosphere Archive and Distribution System (LAADS) website. All MODIS files that encompass AFARS during the specified lidar acquisition are returned. Each MODIS HDF file contains five minutes worth of data called a "granule."

The files are downloaded through an ftp server, and the relevant datasets are extracted in batch using IDL, which possesses a set of commands specifically designed to manipulate HDF formatted files.

In addition to the MODIS and lidar data, several other supplemental resources are used. Dr. Sassen, who operates the ruby lidar, keeps a log of his observations during each lidar acquisition. In the majority of the cases, Dr. Sassen records the precise lidar signal return times for the bases and tops of the various cloud and aerosol layers present during the MODIS overpass. He also documents the types of aerosols present, cloud cover conditions, presence of optical phenomenon or precipitation, and whether or not the lidar signal is being attenuated.

To help identify precise cloud base and cloud top altitudes, plots representing the profile of the lidar returned power averaged over a five minute period are generated for each overpass. A lidar backscatter profile represents the height distribution of the average range normalized ($1/R^2$) return signal. In an atmosphere devoid of clouds or aerosols, the profile of returned signal exhibits a smooth exponential decrease with height, associated with the exponential decrease in atmospheric density. The presence of a cloud or aerosol layer, however, creates a well-defined spike in the return signal, where the inflection points are assumed to represent the layer boundaries.

These plots are made using an IDL program created, again, by Jiang Zhu. The program allows the user to specify the start and end times over which the return signals are averaged. A five minute (± 2.5 minute) average of the MODIS overpass was deemed an acceptable length of time to minimize the noise in the profile. However, a small

handful of cases occur in which a satellite overpass took place within two and a half minutes of either the start or end times of the lidar acquisition. For these cases, the signal was averaged for the period beginning (ending) with the lidar acquisition and the two and a half minute period after (before) the satellite overpass.

In this study, cloud and aerosol boundaries and particle phase are determined by visual inspection of the lidar plots. However, automated computer algorithms can perform the same duties using vertical profiles of backscatter and depolarization (Wang and Sassen, 2001; Pal et al., 1992). However, the automated algorithms, which require a set of absolute thresholds of backscatter and depolarization, are not always reliable since the business of cloud detection can be extremely ambiguous.

One such program, called the Cloud Base Cloud Top (CBCT) algorithm, is used for this research. The CBCT calculates the base of the lowest cloud layer and the top of the highest cloud layer for the specified one minute average or PCL data overlapping the MODIS overpass time. In essence, the algorithm pinpoints the cloud boundaries by examining the slope of the lidar backscattered signal. If a cloud is present, the cloud base exhibits a sharp positive slope, larger than a pre-determined noise level, in the backscatter. If the cloud is optically thin, the top represents the point where the slope of the backscattered power falls back to the molecular signal.

Discrimination of clouds from precipitation, aerosols and noise can be difficult since these chosen thresholds may vary in time and space. Another difficulty arises when trying to distinguish between an actual cloud top and an *effective* cloud top caused by the rapid attenuation of the signal above the peak (Wang and Sassen, 2001). The two can

usually be distinguished by analyzing the behavior of the slope above the peak return; if the slope resembles a smooth exponentially decreasing molecular profile, one can assume that the lidar has penetrated completely through to the cloud top. Likewise, phase discrimination can also be conducted using depolarization thresholds from vertical profiles of depolarization ratio (Wang and Sassen, 2001), but again, one encounters the difficult task of choosing the right thresholds.

For this research, the CBCT outputs provide a supplemental source of cloud top altitude. The cloud top estimates for the 1-minute average during and the 1-minute averages immediately preceding and following the MODIS overpass are examined. These cloud boundaries are examined collectively along with the lidar images and the 5-minute backscatter profiles to estimate the final cloud and aerosol boundaries. CBCT commonly misses thin cirrus visible in the lidar data and documented in the field notes. At other times, the algorithm misidentifies thick aerosol layers as cloud layers.

In order to compare the MODIS cloud tops in hPa with the lidar cloud top in meters, the pressure (altitude) units need to match. The conversion of the MODIS top pressures to meters and the lidar top altitudes to hPa is accomplished using radiosonde profiles of height, pressure and temperature measured. Weather balloons are launched twice per day from approximately 800 stations, including Fairbanks, at 0000 and 1200 UTC. Most of the MODIS overpasses over Alaska occur between 2100 and 2300 UTC, several hours prior the 0000 UTC balloon sounding. This time discrepancy is expected to introduce some error in the cloud top interpolations since temperature and pressure fields are constantly changing in the atmosphere. However the magnitude of the error is

difficult to quantify and may vary depending on the meteorological conditions at the time and the delay between the satellite overpass and the radiosonde measurement. Atmospheric upper air soundings can be obtained from the University of Wyoming at *http://weather.uwyo.edu/upperair/sounding.html* (Atmospheric soundings, 2008).

3.2 MODIS Data Extraction Procedure

A minimum distance algorithm locates the pixel directly over the AFARS facility using the Pythagorean Theorem. Each rectangular MODIS granule is comprised of individual pixels that approximately reside in a Cartesian coordinate system with latitude representing the y-axis and longitude representing the x-axis. The AFARS site is located at 64.87 N, -147.83 W. To locate the pixel that encompasses the AFARS site, the following steps are executed in batch mode:

1. Open the *latitude* and *longitude* SDS's in the HDF file.
2. Place the *latitude* values into a 2-D variable called *Latitude*; repeat for longitude.
3. Create two new 2-D variables called *lat_diff_squared* and *long_diff_squared* where

   $$lat\_diff\_squared = (64.8689 - Latitude)^2,$$

   $$long\_diff\_squared = (-147.8281 - Longitude)^2.$$

4. Create another 2-D variable called *dist_squared* where

   $$dist\_squared = (lat\_diff\_squared + long\_diff\_squared)$$

5. Find the i^{th} and j^{th} coordinate that represents the minimum value of the *dist_squared* variable. This coordinate represents the location of the pixel directly above AFARS.

Once the correct indices are found, the corresponding i^{th} and j^{th} pixel values for the remaining SDS's are extracted and printed to a text file in batch mode.

Similarly, this procedure is used to define the 5 x 5 (25 pixel) box for the 1-km cloud mask (MOD35) product. Once identified, the i^{th} and j^{th} indices of the center pixel get stored in a 2-element variable named *index*, in which *i* is assigned to the first element, *index*[0], and *j* is assigned to the second element, *index*[1]. The variable *Box*, declared as a 2-column by 25-row matrix, stores the i^{th} and j^{th} indices for each of the twenty-five pixels. For example, the corresponding x and y-values (x, y) of the top left, center and bottom right pixels in variable *Box* would represent the coordinates (*index*[0] - 2, *index*[1] - 2), (*index*[0], *index*[1]) and (*index*[0] + 2, *index*[1] + 2) respectively. The pixel indices for the remaining twenty-three pixels can be similarly defined.

By convention, the first index of an IDL array is the zeroth element. Additionally IDL follows a "column major format" whereby the first dimension represents the column number. For example, *array*[3,2] represents the element in column 4, row 3 for the variable named *array*.

The following fields are routinely extracted for each HDF file: date, scan start time, latitude, longitude and sensor zenith angle. The scan start time is recorded at the onset of each new mirror scan about every three seconds. A new time gets stamped for every ten rows, two rows and one row of 1-km, 5-km and 10-km resolution data respectively. The units are given in seconds since January 1, 1993 at 00:00:00.0 UTC, and are then converted to UTC.

3.3 A Case Study

To illustrate precisely how the lidar and MODIS data are extracted, recorded and analyzed, an example from the July 24, 2006 lidar acquisition is presented here. This case captures a cirrus cloud, an attenuating liquid altocumulus and several layers of dense smoke.

According to the *AFARS* cloud description list, the lidar collection time on July 24, 2006 occurs from 2006 – 2319 UTC. These times and the AFARS latitude and longitude coordinates are input into the LAADS search query (*http://ladsweb.nascom .nasa.gov/data/search.html*; LAADS, 2008b) in order to locate the overlapping MODIS overpasses. For this case, the search engine returns four granules, two Terra overpasses at 2045 and 2220 UTC, and two Aqua overpasses at 2100 and 2235 UTC. Twelve files in total (4 granules x 3 files per granule) are downloaded through and ftp address, the relevant SDS's for the pixel over AFARS are extracted using IDL and printed to a text file, which are given in Table 3.1. Table 3.2 provides the first byte of the MODIS cloud mask. The first three rows represent the results of the AFARS centered pixel from the MOD06 5-km SDS. The next three represent the 1-km AFARS centered pixel from the MOD35 cloud mask product. The final three represent the mode or most common value of the twenty-five 1-km pixel box (5 x 5) centered over AFARS. The cloud mask pixel integers are listed in column seven. The subsequent columns provide the bit translation.

The cloud mask results are, for the most part, consistent. However, note the 5-km cloud mask pixel for the 2100 UTC Aqua overpass yields an erroneous *confident clear*

Table 3.1: **MODIS derived cloud and aerosol data from July 24, 2006.** The four cases comprise two Aqua (A) and two Terra (T) overpasses. The cloud top height values (column 10) are interpolated from the radiosonde profiles. Cloud product parameters are given at 5-km resolution and the aerosol product aerosol type is given at 10-km.

Date	Platform	UTC	Pixel Latitude	Pixel Longitude	Sensor Zenith Angle	Top Retrieval Method	Cloud Top Pressure (hPa)	Cloud Top Temp (^{o}C)	Cloud Top Height (m)	Cloud Mask	Cloud Phase	Aerosol Type
7/24/2006	T	2045	64.814	-147.79	39.58	3-CO2	285	-43.04	10500	249	2-ice	-9999
7/24/2006	A	2100	64.907	-147.81	51.39	4-CO2	225	-51.51	10800	255	1-water	2-sulfate
7/24/2006	T	2220	64.864	-147.86	42.37	1-CO2	315	-37.53	8800	249	2-ice	-9999
7/24/2006	A	2235	64.879	-147.83	24.54	2-CO2	275	-45.19	10000	249	2-ice	-9999

Table 3.2: **The first byte (Byte #0) translations of the MODIS cloud mask.** Outputs for the (a) '5-km point,' (b) '1-km point,' and (c) the '1-km box' methods. The 1-km data are obtained from the MOD35 file and the 5km from the MOD06 file.

(a) MOD06 '5 km point' Date	Platform	UTC	Pixel Latitude	Pixel Longitude	Zenith Angle	Cloud Mask	bit 0	bit 1-2	bit 3	bit 4	bit 5	bit 6-7
7/24/06	T	2045	64.814	-147.79	39.58	249	D	C-Clouds	Day	sunglint-n	sn/ice-n	land
7/24/06	A	2100	64.907	-147.81	51.39	255	D	C-Clear	Day	sunglint-n	sn/ice-n	land
7/24/06	T	2220	64.864	-147.86	42.37	249	D	C-Clouds	Day	sunglint-n	sn/ice-n	land
7/24/06	A	2235	64.879	-147.83	24.54	249	D	C-Clouds	Day	sunglint-n	sn/ice-n	land
(b) MOD35 '1 km point'												
7/24/06	T	2045	64.852	-147.84	0	249	D	C-Clouds	Day	sunglint-n	sn/ice-n	land
7/24/06	A	2100	64.853	-147.85	0	249	D	C-Clouds	Day	sunglint-n	sn/ice-n	land
7/24/06	T	2220	64.858	-147.84	0	249	D	C-Clouds	Day	sunglint-n	sn/ice-n	land
7/24/06	A	2235	64.858	-147.84	0	249	D	C-Clouds	Day	sunglint-n	sn/ice-n	land
(c) MOD35 '1 km box'												
7/24/06	T	2045	64.852	-147.84	0	249	D	C-Clouds	Day	sunglint-n	sn/ice-n	land
7/24/06	A	2100	64.853	-147.85	0	249	D	C-Clouds	Day	sunglint-n	sn/ice-n	land
7/24/06	T	2220	64.858	-147.84	0	249	D	C-Clouds	Day	sunglint-n	sn/ice-n	land
7/24/06	A	2235	64.858	-147.84	0	249	D	C-Clouds	Day	sunglint-n	sn/ice-n	land

value, likely due to the fact that the center pixel chosen from the 1-km cloud mask assigned to the 5-km cloud product is different than the 1-km pixel nearest to AFARS.

Next the lidar data gets evaluated. The backscattered power and corresponding depolarization ratio plots are given in Figure 3.1. The top depolarizing layer represents a cirrus cloud while the middle and lower layers represent smoke as recorded in Dr. Sassen's field notes. Note that the 2220 UTC Terra overpass occurs at the edge of a signal attenuating altocumulus cloud located at approximately 4 km above MSL. The corresponding range normalized five minute backscatter profiles coinciding with the four overpasses are given in Figure 3.2. Finally, the outputs from the CBCT algorithm are listed in Table 3.3. After examining these sources (Figure 3.1, Figure 3.2, Table 3.3), the cloud and aerosol top altitudes, and path integrated cloud phase, listed in Table 3.4, are determined for each case. Additional relevant information, such as the number of cloud layers, whether or not the beam was attenuated and aerosol type, are also noted.

First, note that cloud is clearly present above AFARS during each of the four overpasses. Three of the four contain only one cloud layer (2045, 2100, 2235 UTC) while one happens to intercept two layers (2220 UTC). For this case, note the "shadow" situated above the bottom layer, caused by the complete attenuation the lidar beam. The cloud top altitudes for the top layer and the cloud top and base boundaries are denoted by the red dots in the five minute backscatter profiles in Figure 3.2. To be thorough, the base and top altitudes of the remaining cloud and aerosol layers are recorded.

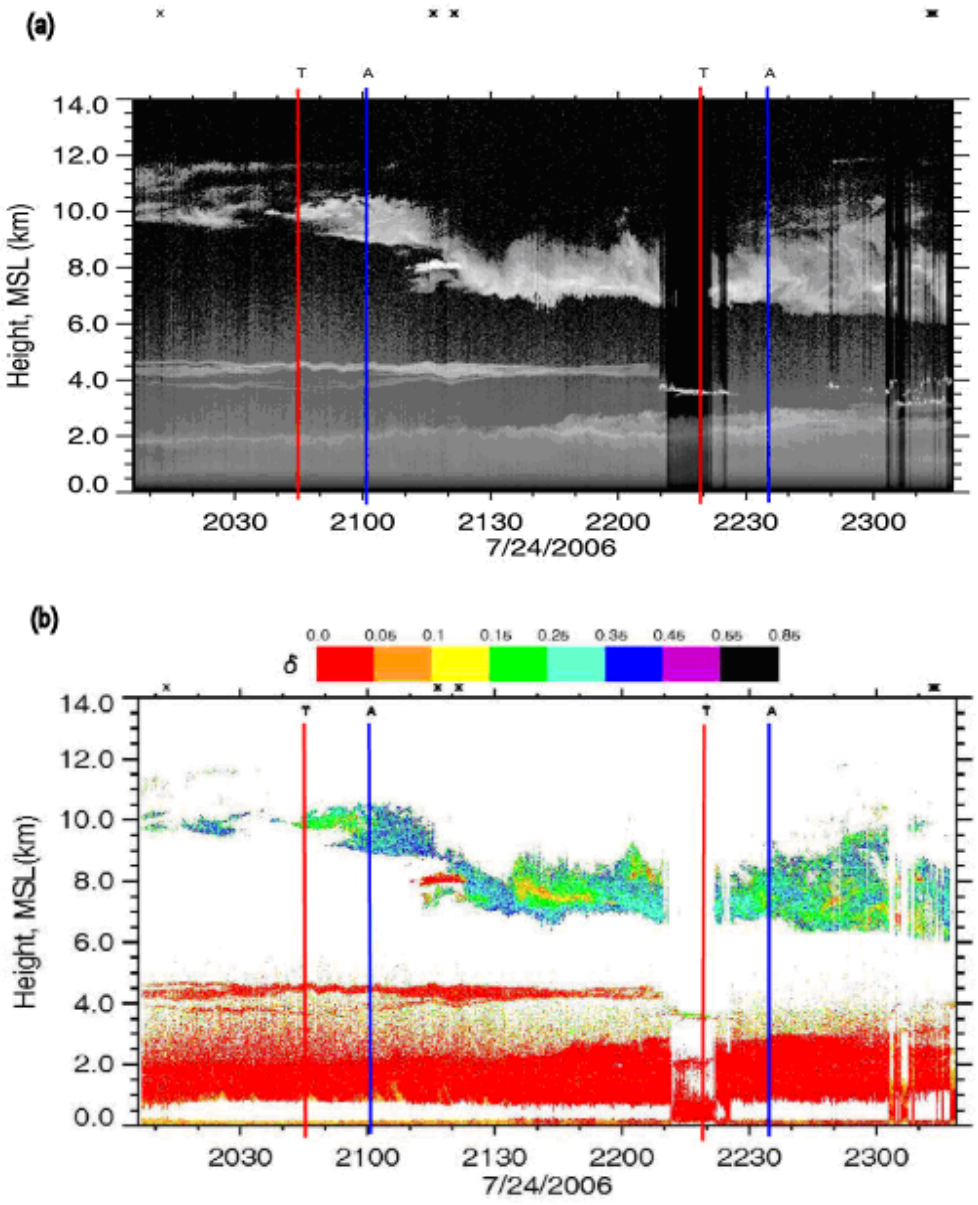


Figure 3.1: Lidar backscatter and δ-ratio plots for July 24, 2006. Vertical distribution of (a) lidar backscatter and (b) depolarization ratio above means sea level (MSL) as a function of time (UTC) for July 24, 2007. The 'T' and 'A' and the corresponding red and blue lines denote approximate Aqua and Terra overpasses. Hatch marks (x) indicate when the lidar was tilted away from zenith.

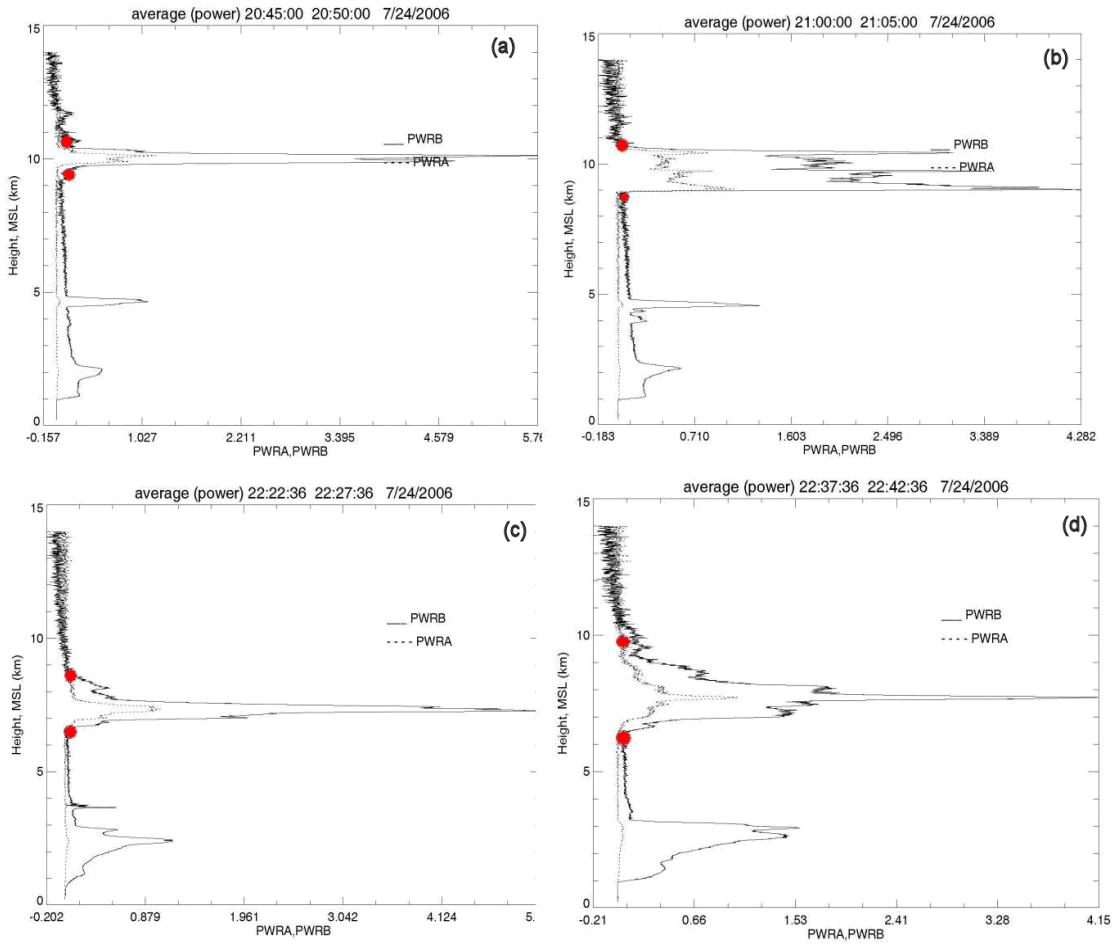


Figure 3.2: Mean five minute profiles of lidar backscatter. Vertical profiles of mean five-minute backscatter of the parallel (PWRB) and perpendicularly (PWRA) polarized channels during the (a) 20:45 UTC Terra (b) 21:00 UTC Aqua (c) 22:22 UTC Terra and (d) 22:37 UTC Aqua overpasses on July 24, 2006. Red dots denote visually estimated cloud top and base boundaries of the top cirrus layer.

Table 3.3: **The CBCT outputs for July 26, 2006**. Cloud base (Cb) and cloud top (Ct) outputs from the CBCT algorithm for the minute preceding (min-1), during (min) and after (min+1) the July 24, 2006 satellite overpasses.

Platform	(UTC)	row scan (UTC)	Cb at min-1 (km)	Cb at min (km)	Cb at min+1 (km)	Mean Cb (km)	Ct at min-1 (km)	Ct at min (km)	Ct at min+1 (km)	Mean Ct (km)
T	2045	20:47:28	5.8	5.77	5.73	5.77	11.83	11.78	11.9	11.8
A	2100	21:02:24	10.16	10.14	9.96	10.09	11.98	12.03	11.94	11.98
T	2220	22:25:04	4.85	4.85	4.83	4.84	9.98	8.99	9.98	9.65
A	2235	22:40:00	7.69	7.75	7.66	7.66	10.54	10.5	10.66	10.57

Table 3.4: **Cloud and aerosol data based on the lidar analysis for July 26, 2006**. Results of the lidar analysis related to the (a) cloud data and (b) aerosol data for the July 24, 2006 overpasses. 'L1' denotes layer 1 or top layer.

(a) Platform	UTC	Cloud Y/N	Cloud Mult. Cld. layers (Y/N)	Cloud # of cld. layers	Cloud Att-Lim (Y/N)	Cloud L1 Top (km)	Cloud L1 Base (km)	Cloud L1 thick. (km)	Cloud L2 Top (km)	Cloud L2 Base (km)	Cloud Path Phase
T	2045	Yes	No	1	No	10.5	9.6	0.9	---	---	Ice
A	2100	Yes	No	1	No	10.8	8.9	1.9	---	---	Ice
T	2220	Yes	Yes	2	Yes	8.8	6.8	2.0	Att_lim	3.8	Mixed
A	2235	Yes	No	1	No	10.0	6.5	3.5	---	---	Ice
(b) Platform	UTC	Aerosol (Y/N)	Aerosol Type	--- ---	--- ---	Aerosol L1 Top (km)	Aerosol L1 Base (km)	Aerosol L1 thick. (km)	Aerosol L2 Top (km)	Aerosol L2 Base (km)	Cld. & Aero (Y/N)
T	2045	Yes	smoke	---	---	4.9	4.3	0.6	2.3	1.7	Yes
A	2100	Yes	smoke	---	---	4.8	3.9	0.9	2.4	1.8	Yes
T	2220	Yes	smoke	---	---	2.9	1.5	1.4	---	---	Yes
A	2235	Yes	smoke	---	---	3	1	2	---	---	Yes

Next, cloud phase gets assessed. The high δ values ($\delta \geq 0.15$) within the top cloud layer indicate the presence of *ice* for cases one, two and four under column twelve in Table 3.4a. Due to the presence of both ice and an underlying liquid altocumulus layer, the 2220 UTC Terra overpass receives a *mixed* classification. However, note that, for this case, the actual scan time, below column three in Table 3.3 occurs at 22:25:04 UTC. It is not entirely obvious from the five minute profile (Figure 3.2c) that the altocumulus layer is present at the instant of the Terra overpass. Nevertheless, since the satellite overpass occurs at the edge of the altocumulus, and because the sensor zenith angle was relatively steep (42^{o}), one can judge that the probability that at least part of the pixel intercepted the altocumulus is high, which would have introduced a liquid signature yielding a *mixed* (*ice* + *liquid*) path integrated phase. However this case demonstrates the ambiguity sometimes inherent in this analysis, especially for the phase analysis.

To properly compare the MODIS cloud top estimates against the lidar the units must be consistent. Radiosonde data of temperature, pressure and altitude acquired nearest to the MODIS overpass is used to convert the units. For most cases, this coincides with the 0000Z (1500 UTC local) launching, since most MODIS overpasses occur between 2100 - 2300 UTC (1200 – 1400 UTC local), several hours prior to the balloon launching. Linear interpolation is used to match the following: lidar top altitude (km) to pressure (hPa); MODIS top pressure (hPa) to altitude (km); lidar top altitude (km) to temperature (^{o}C). The results for this case study are given in Tables 3.5 and 3.6.

Table 3.5: **MODIS and lidar derived cloud top estimates**. Estimates of pressure (P), height (Z) and temperature (T) for the four July 24, 2006 overpasses. Fields labeled with an asterisk (*) denote radiosonde interpolated values.

platform	UTC	MOD Top P (hPa)	Lid Top P* (hPa)	ΔP Lid_P - MOD_P	MOD Top Z* (m)	Lid Top Z (m)	ΔZ Lid_Z - MOD_Z	MOD Top T (^oC)	Lid Top T* (^oC)	ΔT Lid_T - MOD_T
T	2045	285	250.41	34.59	9646.4	10500	-853.6	-43.04	-49.42	6.38
A	2100	225	238.97	-13.97	11193.42	10800	393.42	-51.51	-51.65	0.14
T	2220	315	322.5	-7.5	8954.98	8800	154.98	-37.53	-36.32	1.21
A	2235	275	270.67	4.33	9893.14	10000	-106.86	-45.19	-45.56	0.37

Table 3.6: **MODIS and lidar derived cloud and aerosol information**. Information on cloud phase, aerosol type, and the relevant MODIS cloud mask ancillary inputs and lidar cloud cover parameters used in the data evaluation.

Platform	UTC	Day/Night	Snow (Y/N)	Mult. Cld Layers (Y/N)	Lidar Cld. (Y/N)	MOD06 cld. mask	Lid. Path Phase	MOD Phase	Lid. Aero Type	MOD04 Aero Type
T	2045	Day	No	No	Yes	C-Clouds	ice	Ice	smoke	-9999
A	2100	Day	No	No	Yes	C-Clear	ice	Water	smoke	-9999
T	2220	Day	No	Yes	Yes	C-Clouds	mixed	Ice	smoke	sulfate
A	2235	Day	No	No	Yes	C-Clouds	ice	Ice	smoke	-9999

Once the top properties are obtained, the MODIS value gets subtracted from the lidar value to obtain pressure, height and temperature errors (ΔP, ΔZ, ΔT).

The next four sections explain the various data extraction and analysis procedures by variable (cloud mask, cloud top properties, cloud phase, aerosol type). They convey, in greater detail, how the MODIS data are extracted, how the lidar data gets interpreted and how the two sets get compared.

3.4 Analyis Methods

3.4.1 Analysis Methods: Cloud Mask

The main byte (Byte #0) of the 1-km MODIS cloud mask is included with the 5-km cloud product (MOD06). The 5-km version is generated by assigning the center pixel of a 5 x 5 pixel box from the 1-km MOD35 product as the representative 5-km pixel in the MOD06 product. However, as mentioned earlier, this 1 km^2 region may not accurately represent the cloud cover conditions over the much larger 25 km^2 region. Therefore, three approaches are used to evaluate the cloud mask. The first compares the the 5-km cloud mask from the MOD06 product, for the pixel centered over AFARS, with the with the lidar result, dubbed '5 km point.' The second simply examines the 1-km result from the MOD35 product, dubbed '1 km point,' and the third compares the mode value of a 5 x 5 box of 1-km pixels from the MOD35 product, dubbed '1 km box.'.

The second and third bits of the MODIS cloud mask are translated into one of four categories: *confident clear*, *probably clear*, *probably cloudy*, *confident cloudy*. The origin and meaning of these categories are explained in the algorithm section. The

additional bits of the first byte of the cloud mask are also extracted. In every case, the first bit (bit 0) returned a value of '1', indicating that all cases were deemed *Determined*.

As described in the previous section, a cursory inspection of the lidar plots, five minute backscatter profiles and field notes is enough to determine if a cloud is present during the time of the MODIS overpass. So with respect to the cloud mask comparison, the lidar data can be considered unambiguous. The presence of cloud over AFARS is denoted with either a simple *Yes* for cloudy or a *No* for clear (column three Table 3.4a, column six Table 3.6). Also aerosol layers are disregarded. For example, if a layer of smoke is detected over AFARS but clouds are not, the lidar field-of-view is considered to be cloud-free.

The results of the cloud mask comparison are organized into a table format with the four MODIS cloud mask categories displayed as columns and the two lidar categories as rows. An example of such a table is presented below (Table 3.7). Each case falls within one of eight possible categories. Once all of the cases have been grouped, the accuracy of the MODIS data can be evaluated following two approaches. First, the

Table 3.7: **Summary of the MODIS cloud mask results**. The MODIS returns (*Conf. Clear*, *Prob. Clear*, *Prob. Clouds*, *Conf. Clouds*) representing all 549 sample cases are sorted by lidar cloud detection (*Cloudy*, *Clear*).

Cloud Mask SUMMARY						
Lidar Cloud	**MOD06 Cloud Mask** Conf.Clear	Prob. Clear	Prob. Clouds	Conf. Clouds	Total	**% Agreement *Clear*[1]**
Cloudy	142	10	31	293	476	83.56
Clear	55	6	1	11	73	***Cloud*[1]**
Total	197	16	32	304	549	68.07
% Agree.	**Clear[2] =**	28.64	**Clouds[2] =**	96.43		

probability that the MODIS cloud mask returns a cloudy (clear) result given that the Lidar observation denotes clouds (clear) is examined. These results are presented in the far right column under the *% Agreement* heading. The *clear*^{1} finding is determined by dividing the sum of the MODIS *conf. clear* and *prob. clear* cases (55 + 6 = 61) in the *clear* row under the *lidar cloud* column with the total number of lidar clear sky observations (55 + 6 + 1 + 11 = 73). Similarly, the *cloud*^{1} agreement is computed by summing the MODIS *prob. clouds* and *conf. clouds* cases (31 + 293 = 324) in the Lidar *cloud* row with the total number of lidar cloudy sky observations (142 + 10 + 31 + 293 = 476). These calculations yield an agreement between the MODIS and Lidar instruments of around 83.6% and 68.1% respectively for the *clear*^{1} and *cloud*^{1} analysis.

Alternatively, the MODIS cloud mask can be examined using a reverse approach: the probability that the lidar observation of clouds (clear) matches the MODIS result. These results are printed along the bottom row. In this case, *clear*^{2} is found by dividing the sum of the MODIS *conf. clear* and *prob. clear* cases in the Lidar *clear* row (55 + 6 = 61) by the sum of the MODIS *conf. clear* and *prob. clear* cases (197 + 16 = 213). Similarly, *cloud*^{2} is computed by dividing the sum of the MODIS *prob. clouds* and *conf. clouds* cases (31 + 293) in the Lidar *cloudy* row by the sum of the MODIS *prob. clouds* and *conf. clouds* cases (32 + 304). This approach yields dramatically different results than the former, where the *clear*^{2} (*cloud*^{2}) agreements have significantly worsened (improved) to roughly 24.0% (94.1%).

Recall that thin cirrus clouds, considered to be the most difficult type to detect from passive satellite sensors, comprise the majority of the cases in this sample. The

poor result associated with the *clear*^{2} probability, helps to reinforce this idea that MODIS has problems detecting thin cirrus clouds.

The results presented in Table 3.7 give the total number of cases. Using the ancillary inputs extracted from bits three and five of the cloud mask (refer to Table 2.2), the results can be further separated by the day versus night and the snow versus no snow cases in addition to platform (Aqua versus Terra). These results will be discussed in chapter four.

3.4.2 Analysis Methods: Cloud Top Properties

As discussed, determination of the bases and tops of the various cloud and aerosol layers is carried out primarily through visual inspection of the lidar plots, although, additional sources, such as the CBCT outputs and the five minute backscatter profiles, are consulted. Locating the bases and the tops of the layers visually from the plots is performed using two transparent rulers. The first connects the platform symbol (A, T) representing the approximate MODIS overpass time along the top of the plot with the x-axis, representing the time in UTC. The second is lined up perpendicular to the first, connecting the cloud or aerosol boundary (base or top), with the MSL height. This height gets recorded to the nearest tenth of a kilometer. Also, if the separation distance between layers is approximately less than 0.5 km, the layers are regarded as part of the same layer. This value represents about 25% of 2.1 km, the mean cloud layer thickness.

Once the base and top altitudes for a given layer are gathered from the plot, the corresponding five minute profiles and the CBCT outputs are reviewed. On occasion, the

original altitude estimates may change by a few tenths of a kilometer after consulting the five minute backscatter profiles. The CBCT estimates, on the other hand, are not given much weight. Recall that CBCT provides only the elevation of the lowest layer cloud base and the highest layer cloud top. Also, it is found that CBCT does a poor job of separating clouds from optically thick aerosols and of resolving the boundaries of optically thin cirrus clouds. The CBCT outputs are given in units of height above ground level (AGL), therefore, 0.20 km is added to convert the elevation to above MSL.

The bases and tops of each cloud and aerosol layer are documented, the existence of multiple cloud layers and the number of cloud layers are noted. As mentioned in the cloud top algorithm, multiple cloud layers can contaminate the cloud top estimate, especially if the top layer happens to be an optically thin cirrus cloud. Therefore, this analysis will evaluate the accuracy of the MODIS cloud top estimates based on single versus multiple cloud layers.

If it appears based on either the lidar plots, five minute profiles or from Dr. Sassen's field notes that the lidar pulse did not penetrate to the top of the highest cloud layer, it gets marked down as an attenuation-limited (*att-lim*) case. However, for a majority of the these cases, the cloud top altitude can still be inferred since, most of the time, the signal only gets blocked temporarily, such as the 2220 UTC Terra overpass on July 26, 2006 (Figure 3.1). Of the approximately 103 *att-lim* cases in the sample, cloud top estimates could not be inferred twenty-four times.

The MODIS derived cloud top pressure and temperature data are stored as the thirteenth and sixteenth SDS's of the 5-km cloud product (MOD06). During the

extraction, two subroutines convert the pixel numbers, using the appropriate *scale_factors* and *offsets*, to pressure (hPa) and temperature (^{o}C) values. In addition, the cloud top retrieval method, called *Cloud_Height_Method*, is extracted. As discussed, MODIS uses two techniques for estimating cloud top pressure: the CO_2 density slicing algorithm (CO_2), applied to clouds with tops above 700 hPa, and the 11-micron brightness temperature algorithm (BT_{11}), applied to low clouds assumed to behave as blackbodies. For this analysis, the cloud top errors are partitioned by top retrieval method.

As mentioned, to evaluate the MODIS data against the lidar estimates, the MODIS top pressures (hPa) must be converted to units of height (km) and temperature (^{o}C), and the lidar top heights converted to pressure, but the temperatures are given with the MOD06 product. These conversions are carried out using radiosonde profiles of altitude, pressure and temperature. The caveats associated with the sounding data have are discussed in section 3.1.

Deriving the MODIS cloud altitudes (lidar cloud pressure, temperature) from pressure (altitude) is conducted using a linear interpolation approach. The generic IDL procedure for matching the MODIS cloud top pressure to height using the sounding profile is summarized here. Note the variables *i* and *k* signify the i^{th} MODIS *case* and the k^{th} sounding *level* (altitude) respectively. The steps are as follows:

1. Read in the MODIS cloud top pressure for the i^{th} case, MOD_P[i].
2. Find the pressure levels on the sounding profile located directly below and above the

MODIS pressure (e.g. locate Sounding_P[k] and Sounding_P[k+1] where Sounding_P[k+1] $\leq$ MOD_P[i] $\leq$ Sounding_P[k]).

3. Match the two sounding pressure levels with the associated height (Z) levels (e.g. Sounding_Z[k+1] Sounding_P[k+1]; Sounding_Z[k] Sounding_P[k])
4. Compute the slope, M_{sound}, setting the sounding heights as the y-values and the sounding pressures as the x-values such that

$$M_{sound} = \frac{\text{Sounding\_Z}(k+1) - \text{Sounding\_Z}(k)}{\text{Sounding\_P}(k+1) - \text{Sounding\_P}(k)}. \quad (3.1)$$

5. Solve for the MODIS cloud top height, MOD_Z[i], using the slope, M_{sound}, calculated from step 4. A careful derivation reveals the solution for the MODIS cloud top altitude for the i^{th} case, MOD_Top_Z[i], is given by

$$\text{MOD\_Z}[i] = -M_{sound}(\text{Sounding\_P}[k+1] - \text{MOD\_P}[i]) + \text{Sounding\_Z}[k+1]. \quad (3.2)$$

Similar derivations provide the lidar cloud top pressure and temperature from the altitude estimate. These IDL procedures are automated so that the files can be processed in batch. For each case, the cloud top pressures, altitudes and temperatures corresponding to the MODIS and lidar instruments are printed as separate columns to a text file.

Next, assuming the lidar represents ground truth, the MODIS cloud top errors are computed and then evaluated against a number of conditions. The error is computed for by subtracting the lidar estimate from the MODIS value (e.g. ΔP = MOD_P[i] – Lid_P[i]).

Given the multitude of parameters, these errors (ΔP, ΔZ and ΔT) can be compared in a myriad of ways in order to pinpoint the strengths and weaknesses of the MODIS algorithm. For this research, the mean and standard deviations of the errors are assessed by platform (Aqua versus Terra), time of day (day versus night), snow cover (snow versus no snow), cloud layers (single versus multiple layers) and top retrieval method (CO_2 versus BT_{11}). Also, simple scatter-plots of the MODIS results versus lidar results are generated, and are separated and by top retrieval method. Additionally, scatter-plots of error versus sensor zenith angle and error versus the geometric thickness of the top cloud layer are produced. The former intends to determine if the steepness of the viewing angle affects the cloud top retrievals, and the latter to determine if the MODIS top estimates are correlated with cloud depth. The regression lines and correlation coefficients, R^2, are displayed with the plots.

The errors can also be represented in histogram form. The errors are sorted and counted into 100 hPa, 2000 m and 10^oC bins for pressure, height and temperature respectively. The series are also separated by top retrieval method.

Finally, the errors are plotted by month to determine if any seasonal patterns are evident. It is hypothesized that errors in the MODIS estimates will be more pronounced

in winter when sun angles are low, and when snow and surface inversions are likely to be present.

3.4.3 Analysis Methods: Cloud Phase

Plots representing the vertical distribution of depolarization ratio are used to visually evaluate the cloud thermodynamic phase corresponding to the MODIS overpass. Detailed discussions on the polarization lidar technique and a definition of δ, the depolarization ratio, are covered in chapter two. Recall different particle shapes generate markedly different signatures in δ due to the manner in which laser light gets scattered by the particles. Liquid cloud particles that possess a spherical shape do not change the orientation of E-vector during a single scattering event, whereas ice particles generally change the orientation of the E-vector a great deal, which is largely a function of the particle shape (Sassen, 1991). As a result, discrimination of cloud particle phase using a polarization lidar is somewhat straightforward provided that the signal does not get attenuated by an optically thick cloud. The notable exception to this rule is the case of flat hexagonal plate crystals, which fall with their flat surfaces oriented nearly perpendicular to the incident zenith-pointing laser. Such crystals do not change the polarity of the reflected light since the photons undergo a specular reflection. As discussed earlier, this problem can be resolved by tilting the laser slightly off nadir.

The MODIS cloud product (MOD06) provides three cloud phase SDS's: one for the day overpasses, one for the night, and a third, which combines the two. The tri-spectral brightness temperature difference (BDT) method works under all illumination

conditions while the visible near-infrared bands ratio (BR) method only works during day overpasses. Results of several individual cloud mask tests are also incorporated. The details of the phase algorithm are discussed in Section 2.3.3.1. The cumulative assessment (day and night) is represented as the 5-km SDS labeled *Cloud_Phase_Infrared*. Once the phase has been determined, one of seven integers representing five categories is assigned to each pixel as follows: *clear* = 0; *water* = 1, 5; *ice* = 2, 4; *mixed* = 3; *uncertain* = 6. As mentioned, the *uncertain* category gets grouped under *mixed*.

Evaluation of cloud phase using the lidar δ-plots (Figure 3.1b) is carried out by first, locating the satellite overpass time on the plot using the recorded scan times and the platform symbols. Next, a transparent ruler is placed on the image connecting the platform symbol along the top of the plot with the x-axis. If the transect intersects a cloud layer, the δ values are examined using the color-coded table as a guide. Cloud layers can be separated from aerosols based on their appearance (large backscatter returns, well-defined boundaries, texture) in the corresponding backscatter plot and are noted in Dr. Sassen's field notes. If the δ-values within the cloud layer are consistently greater than 0.15, the layer gets grouped under *ice*. Layers that exhibit low δ-values ($\delta \approx 0$) near cloud base get grouped under *liquid*. These layers also tend to be optically thick, generating noticeable 'shadows' caused by the attenuation of the signal above the layer.

Several criteria are used to identify cases of mixed phase clouds in the lidar data. Recall that, with respect to passive remote sensing, the radiation reaching the satellite is comprised of a mixture of the individual components along the sensor's line of sight. For

a situation in which an optically thin cirrus is situated above an optically thick stratus, the sensor will detect radiation originating from both sources in different proportions, even though the algorithm can only assume one emission source (cloud layer) in its field-of-view. Consequently, as discussed earlier, MODIS pixels that possess a liquid layer (e.g. marine stratus) underneath an ice layer (e.g. cirrus) commonly get classified into the *uncertain* or *mixed* categories (Platnick et al., 2003).

This logic must be considered when assessing the existence *mixed* phase cases in the lidar data in order to be consistent with the MODIS algorithm. When multiple discrete cloud or precipitation layers containing both liquid and ice phases are found to exist in the lidar plot during the MODIS overpass, the case receives a *mixed* label. Circumstances classified as *mixed* cases may include the following:

1. Cirrus over a low-level water cloud (e.g. cumulus, stratus)
2. Snow or ice virga falling out of a liquid stratus or altocumulus cloud
3. Any situation in which one or more liquid cloud layers is present concurrently with one or more ice layers (e.g. cirrus cloud, frozen precipitation, oriented plate crystals)

Discerning particle phase for mid-level clouds, which includes altostratus and altocumulus, is somewhat less standardized since mid-level clouds reside in the transition zone temperatures between 0^{o}C and -40^{o}C, the transition zone where both liquid and ice can be present. Ice dominated altostratus clouds are traditionally defined as cirrostratus clouds which have attained an optical depth of 3.0 sufficient to completely obscure the

solar disk (Sassen, 2002a). These clouds tend to exhibit δ-values commensurate of ice clouds, but are, by definition, attenuation-limited. Embedded layers of super-cooled liquid water are commonly seen in ice dominated altostratus clouds (Sassen, 1984), while water dominated altocumulus tend to exhibit δ-values, with practically zero depolarization near the base and a rapid increase in δ-values due to multiple scattering followed by total attenuation of the signal.

Mid-level cloud phase is determined from the δ-plots. If the δ-values are consistently larger than 0.15 throughout the layer, it gets grouped under *ice*. Altostratus that originate from a thickening cirrus layer are usually all ice. If large patches of orange and red indicating the presences of liquid droplets, or if the layer exhibits liquid like δ-values but is precipitating highly depolarizing snow or virga particles, the layer gets grouped under *mixed*. Most altocumulus cases, however, get classified as *liquid* because their δ-signatures are consistent with liquid clouds.

A handful of nimbostratus cases are included with the sample. For these, complete attenuation of the lidar signal above cloud base throughout most or the entire lidar acquisition period is observed. Even though it is impossible to discern what lies above, these cases receive either a *liquid* label if they are precipitating rain, or a *mixed* classification if the precipitation is frozen. None receive the *ice*-only tag. Nevertheless, one readily admits that being blind to what resides above the cloud base significantly increase the risk of misinterpreting cloud phase. However these cases represent only a small fraction of the entire sample.

In many situations, Dr. Sassen's field notes provide helpful sky and cloud cover observations that, at times, can be used to clarify or to confirm the phase in the lidar data. Observations of halos or coronas corroborate the presence of ice or liquid for those cases when the δ-images were difficult to interpret.

The phase product is compared much in the same way as the cloud mask: the results are counted and organized into a table similar to Table 3.7 in order to see how well the two datasets agree. An example is presented in chapter four.

3.4.4 Analysis Methods: Aerosol Type

The MODIS aerosol product (MOD04) includes the 10-km *Aerosol_Type_Land* SDS. Like the cloud phase, the aerosol types are stored as integers which translate to the following categories: *mixed* = 0; *dust* = 1; *sulfate* = 2; *smoke* = 3; *heavy_smoke* = 4. Selection of aerosol types based on the various seasonally and geographically defined aerosol models developed from the AERONET database is discussed in the section 2.3.4.1 devoted to the aerosol type algorithm.

Aerosol layers can, for the most part, be visually differentiated from cloud layers in the lidar plots (Wang and Sassen, 2001). However, the identification of aerosol type is best carried out by an experienced observer in the field, a person who can also synthesize other sources of information and can put observations into context with conditions on the ground. Dr. Sassen uses a combination of visual observation and real-time lidar data to document the bases, tops and types of the aerosol layers. Dr. Sassen records the layer boundaries within a minute of the Aqua or Terra overpass for most cases.

By far, smoke and Asian dust are the most frequent aerosol types observed during this study. Smoke from boreal forest fires is common over Alaska during the summer. Smoke may originate from Alaska or be transported from Siberia or Canada depending on the flow patterns (Tirichurapalli, 2006). In Alaska, the 2004 and 2005 summer fire seasons burned approximately 6.7 and 4.5 million acres respectively, ranking as the first and third worst fire seasons since 1956 respectively (Rozell, 2006). Also, frequent spring dust storms over the Gobi and Taklamakan deserts transport dust particles into Alaska between February and May (Tirichurapalli, 2006). As mentioned, dust has strong ice nucleating capabilities, and discriminating between dust layers and cirrus clouds can be difficult (Roskovensky and Liou, 2005).

In addition, Dr. Sassen uses the Naval Research Laboratory's (NRL) aerosol monitoring network to anticipate major aerosol events, such as Asian dust storms, before they reach Alaska. Other resources, such as the NOAA Hybrid Single-Particle Lagrangian Integrated Trajectory Model (HYSPLIT) can be used for backtrack analysis to determine the origin or to validate the source of an observed aerosol. The NRL's aerosol products can be found at *http://www.nrlmry.navy.mil/aerosol/* (Naval Research Laboratory, 2008). The HYSPLIT model can be accessed at *http://www.arl.noaa.gov/ready/hysplit4.html* (HYSPLIT Model, 2008).

The aerosol analysis for this research is divided into two sections: the first provides a climatological summary of aerosol observations derived from the field notes and from the lidar data for all 549 cases. The second compares the MODIS C004 and lidar derived aerosol types for the truncated February 2004 – June 2006 sample. In

addition, since aerosol layers can both resemble and help to form clouds, aerosol observations may come in handy at determining how efficient MODIS is at differentiating between cloud and aerosol layers, and vice versa.

3.4.5 Analysis Methods: Results by Optical Depth

3.4.5.1 Definition of Optical Depth

The term optical depth loosely describes the opaqueness of a medium or, how much light gets attenuated from the original beam via scattering and absorption as it propagates through a substance. In general, turbid media have a large optical depth, whereas transparent media have a small optical depth, depending on the wavelength. Optical depth ( ) can be expressed mathematically as (Liou, 2002)

$$= \quad s \tag{3.3}$$

where is the extinction coefficient, the combination of scattering plus absorption, is the density of the medium and s is the path length. is unit-less, and represents, in essence, the probability that a photon will interact with the media during its transit time.

Remote sensing techniques, such as the LIdar RADar (LIRAD) method, can be used to calculate the visible optical depth of cirrus clouds (Comstock and Sassen, 2001) however, such a method requires two sets of measurements taken concurrently, one from an active lidar and the other from a passive infrared radiometer. Unfortunately,

measurements of infrared downwelling radiances were not taken concurrently with the lidar measurements hence the LIRAD method could not be applied for this study.

3.4.5.2 A Proxy Measure of Optical Depth

Experienced field observations of cirrus clouds have led Sassen to devise an optical depth classification scheme based solely on the visual appearance of cirrus (Sassen and Cho, 1992). This scheme is given below in Table 3.8. Dr. Sassen's field cirrus cloud descriptions consistent with the naming scheme in Table 3.8, are used in conjunction with the lidar plots to classify the cases into different categories based on optical depth.

For this study, a slightly modified version of the naming scheme in Table 3.8 is used per Dr. Sassen's instructions. The categories and their respective optical depths are as follows: 1) "---" for clear or aerosol only, 2) "0.03-" for subvisual, 3) "0.03+" for thin, 4) "0.30-" for thinnish, 5) "0.30+" for opaque, 6) "3.0-" for opaquish and 7) "3.0+" for attenuation-limited. The descriptors (thinnish, opaquish, etc.) are used throughout the field notes to describe the visual appearance of the cirrus layers from the ground looking

Table 3.8: Cirrus optical depth classification scheme. Optical depth estimates of cirrus clouds based on their visual appearance. This table is reproduced from Table 2.1 from the book *Cirrus* (Sassen, 2002a).

Category	Tau Range	Description
Subvisual	<0.03	Invisible against the blue sky
Thin	0.03-0.30	Translucent, retains a bluish color
Opaque	0.30-3.0	Usually appears white
Altostratus	>3.0	Disk of sun becomes indistinct

towards nadir. However, they are not used for all cases. Important optical depths to keep in mind include the following: the boundary where the blue sky just barely appears to be obscured by a cirrus layer occurs for $\tau = 0.03$, the boundary where the sky loses its bluish appearance occurs for $\tau = 0.30$, and the boundary of complete attenuation in which the solar disk is no longer visible occurs for $\tau = 3.0$..

Since the MODIS cloud mask, cloud top property, and to some extent, cloud phase estimates are all dependent on optical depth, using it, even if it is a crude estimate, should prove to be useful. However, one readily admits, again, that the specter of subjectivity cannot be eliminated entirely from this process. First, the naming scheme, itself, is inherently intuitive since it is based solely on human observation. And although, for most cases, the cirrus layers are described using the terms defined in the naming scheme (subvisual, thinnish, opaque, etc.), in some cases, they are not. Also specific descriptions of cirrus clouds at precisely the time of the MODIS overpasses are not always present. Hence, it rests on the shoulders of the data analyst to synthesize both the field notes and the lidar plots to make a final judgment for many of these interpretations. Additionally, it is the intention of the analyst to determine the aggregate optical depth between the ground and the top of the highest cloud layer. For example, if two "thin" layers are present, the analyst may group such a case into "thinnish" depending on their appearance.

The next chapter, chapter four presents and discusses the results for this three and a half year study. All results pertaining to optical depth are given at the end of chapter four.

Chapter 4 Results

4.1 Sample Overview

4.1.1 MODIS Case Statistics

This dataset contains 549 individual MODIS overpasses spanning three and half years from February of 2004 to August of 2007. Most cases occur between the hours of 2000 - 2330 UTC (1100 – 1430 local) when the Terra and Aqua overpasses happen to take place within fifteen minutes of each other over Alaska. Less than 2% of the cases occur during the hours of 0520 - 0735 UTC (2020 – 2235 local), a time that coincides with additional Terra overpasses.

Because of Alaska's high latitude location and MODIS's wide swath width, a lidar collection, occurring over the course of several hours, would frequently capture two overpasses per platform. For example, referring to the case study presented in chapter three, the July 24, 2006 lidar collection (Figure 3.1) lasted a little over three hours from about 2005 to 2320 UTC. This collection captures both the 2045 and 2220 Terra and the 2100 and 2235 UTC Aqua overpasses respectively, yielding a total of four MODIS passes during a single lidar acquisition.

Figure 4.1 separates the cases into five basic groups based on cloud observations. The groups are as follows; L_{clr} (Lidar-clear) – M_{clr} (MODIS-clear) represents cases where both the MODIS and the Lidar instruments observe cloud-free conditions over AFARS; L_{cld} - M_{cld} are cases where both MODIS and the lidar observe cloud; L_{clr} – M_{cld} or *F-N*,

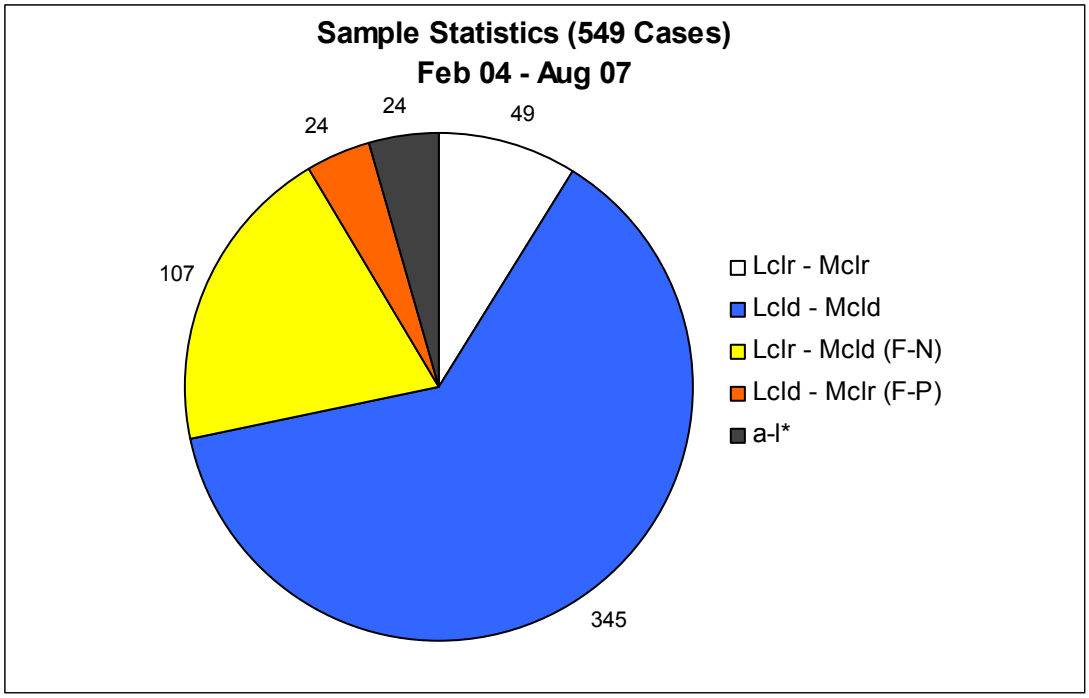


Figure 4.1: Sample breakdown sorted by cloud observations. White: L_{clr}-M_{clr}; Blue: L_{cld}-M_{cld}; Yellow: L_{cld}-M_{clr} (*F-N*); Orange: L_{clr} - M_{cld} (*F-P*); Gray: Attenuation-Limited.

which stands for 'false-negative,' signify cases when MODIS judges the pixel over AFARS to be clear while the lidar sees cloud; L_{cld} - M_{clr} or *F-P*, which stands for 'false-positive,' are cases where MODIS detects cloud when the lidar detects cloud-free; finally, *a-l** represents 'attenuation-limited' cases, cases in which the cloud above AFARS completely attenuated the laser before it reached the cloud top, and for which a cloud top altitude could not be inferred. Note the above categories may or may not contain aerosol layers.

The attenuation-limited cases are divided even further. Of the 549 cases in the sample, 103 (>18%) are documented to be attenuation-limited. However, the twenty-four cases in Figure 3.1, denoted by an asterisk, represent a special sub-category of

attenuation-limited cases for which no information above the laser beam could be inferred. These cases are omitted from the cloud top data analysis, but remain in the phase analysis since phase can still be estimated based the field notes and on the meteorological conditions at the time of the overpass. Typically these cases receive either a *liquid* or *mixed* phase classification. In fact, many of the *a-l** cases in Figure 4.1 are cases of precipitating stratiform clouds.

For the other attenuation-limited cases, cloud top information could still be deduced. For example, the 2220 UTC Terra overpass in Figure 4.1 occurs during a ten minute period when an alto-cumulus cloud with a base near 4 km passed over the laser. Before 2210 and after 2222 UTC, a cirrus layer is visible from 7 – 9 km in altitude. One can safely assume that this layer remains present during the twelve minute window even though it was within the shadow of the cumulus cloud. At other times, a thickening cirrus or altostratus cloud may attenuate the signal. Such clouds barely exceed the $\tau = 3$ threshold representative of an optically thick cloud, but these cases are still assigned a cloud top height since it is assumed that the actual top would fall reasonably close (within 1 – 2 km) of the apparent top in the lidar plot. However, one readily admits that such guesses are inherently speculative and thus, will likely introduce unknown amounts of error into the cloud top estimates.

4.1.2: Lidar and Surface Observations

Figure 4.2 catalogs the various cloud cover conditions as observed from the ground during each MODIS overpass. The ten categories are specified as follows:

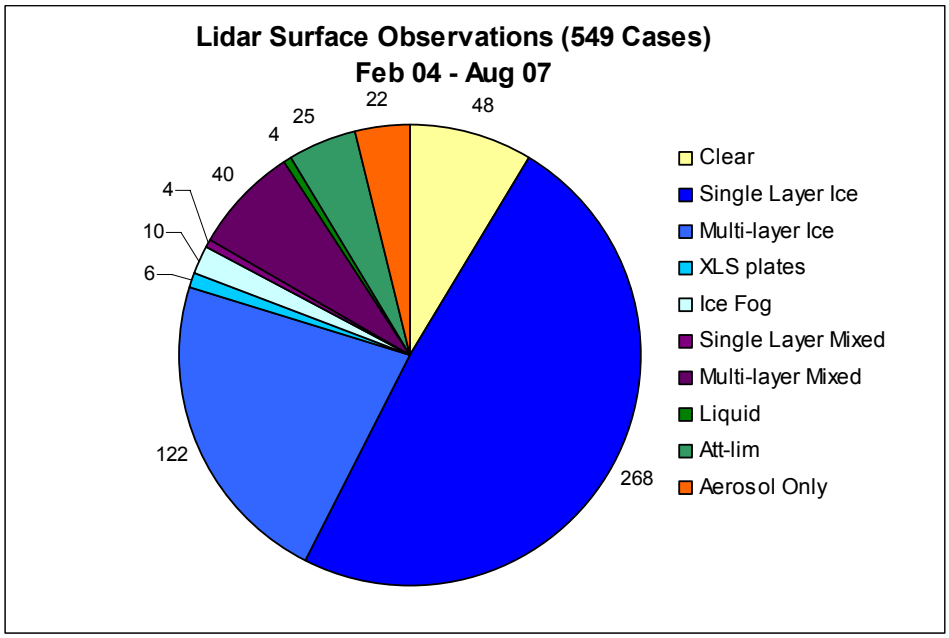


Figure 4.2: Field and lidar observations of clouds and aerosols. Cloud and aerosol cases observed at AFARS from February 2004 to August of 2007.

Clear *(clear)*: Cloud-free; may contain thin aerosol layers.

Single layer ice (*ice)*: One and only one ice cloud layer; may contain thin aerosol

Multi-layer ice *(ice)*: Two or more ice cloud layers, or one ice cloud layer and one oriented plate crystal (xls) layer; may contain thin aerosol.

XLS *(ice)*: Single or multiple layers of oriented plate crystals (XLS) only; may contain thin aerosol.

Ice fog *(ice, clear)*: Surface ice fog is prevalent; may contain additional very thin cloud layers or thin aerosol.

Single layer mixed *(mixed)*: One and only one discernable altostratus or

altocumulus layer. May contain virga (ice or liquid) and/or thin aerosol.

Multi-layer mixed *(mixed)*: Two or more discernable mixed cloud layers, or an ice cloud layer above one or more mixed or liquid cloud layers; may contain thin aerosol.

Liquid *(liquid)*: One or more liquid layers, usually low level broken cumulus or thick low level rain producing stratiform. Regions above can be inferred; may contain thin aerosol.

Att-lim *(mixed, liquid)*: Low or mid level attenuating stratiform clouds. Unlike *liquid*, regions above cannot be inferred; may contain thin aerosol.

Aerosol Only *(Clear)*: One or more thick aerosol layers. No cloud layers.

These categories are chosen based on several criteria. First and foremost, they are differentiated by phase since phase is the primary cloud variables being evaluated. Note, the italicized phase types (*ice*, *liquid*, *mixed*, *clear*) adjacent to each category name specifies under which group the given category is likely to fall. The above categories are also separated by number of cloud layers (single verse multiple). Per the chapter two discussion, MODIS top pressure estimates of thin cirrus tend to be more accurate for single layer cases since additional layers can skew the cloud top estimates towards a lower (higher) cloud top height (temperature). In addition, multiple layers of different phase clouds can also impede definitive cloud phase identification.

The oriented plate crystal caterogy (*XLS*) represents a unique type of stratum that belongs somewhere in between cloud and virga. These layers are comprised of

suspended plate crystals and are on the order of tens to several hundred meters thick. Because of their flat shape and orientation, they exhibit low depolarization values ($\delta \leq 0.05$) for the reasons discussed in section 2.4.2. When the beam is tilted off nadir, the amount of depolarization undergoes a sharp rise.

XLS layers are observed ninety-three times during all twelve months of the year at various heights in the atmosphere. Many appear to be associated with aerosol layers such as smoke or Asian dust, or embedded within cloud layers. Others exist as distinct layers and are considered to be subvisual cirrus layers; the tops and bases of which are recorded as any other cloud layer would.

During the winter months, ten *ice fog* cases are documented. Ice fog typically forms during clear cold conditions within strong surface inversions (Pruppacher and Klett, 1997). Exhaust plumes from cars, buildings and power plants can enhance or even cause ice fog to enshroud the city of Fairbanks for days at a time. Technically, true ice fog only forms when the temperature is at or below -40^{o}C, the threshold at which liquid droplets homogenously freeze (Pruppacher and Klett, 1997). However, ice fog in this sample is more loosely assigned to cases of regular fog composed of super-cooled liquid droplets or diamond dust (large suspended ice crystals) which can develop under similar meteorological situations but at warmer temperatures ($T \leq -15^{o}$C). In some cases, ice fog is treated as a distinct cloud layer with a defined cloud top, usually at or below 2 km, and a base at the surface (0 km). For other cases, it is not. The criteria under which ice fog is or is not treated as a cloud layer is somewhat arbitrary and depends on the strength of the backscatter in the lidar plot and highly on the observations in the field notes. Localized

ice fog observations originating from power plant plumes as recorded in the field notes are typically disregarded, while thick more widespread events are not.

One must bear in mind that the terms "thin" and "thick" with respect to cloud, fog and aerosol layers are somewhat arbitrary, since calculations of cloud and aerosol optical depths from the lidar backscatter profiles are not conducted. Although, as described in section 3.8, a crude measure of optical depth is applied to the data analysis using the classification scheme based on visual appearance developed by Sassen. The results based on this classification scheme are organized at the end of this chapter.

But admittedly, without definitive knowledge of optical depth, it is impossible to quantify the limits of the MODIS cloud and aerosol detection algorithms. For example, the *aerosol only* cases, represented by the orange class in Figure 4.2, are based solely on visual inspection of the lidar plots. If the aerosol layer appears "thick," producing a strong bright backscatter signal in the lidar plot, then it is grouped into *aerosol only*. If the aerosol layer appears "thin," then the case is classified under *clear*. For this analysis, all aerosol layers, regardless of type or opacity, are noted, and the top and base heights are documented when possible.

One might also notice that the twenty-five *att-lim* cases in Figure 4.2 do not match the twenty-four *a-l** cases in Figure 4.1, even though they represent the same scenario. This is due to the fact that one of the *a-l** also happened to be an *F-N* case, and was subsequently classified into the latter group.

4.1.3: Summary of Comparison by Cloud Type

A cursory comparison by cloud phase for the MODIS and lidar derived observations, presented below in Figure 4.3, reveals large discrepancies in both the sky cover conditions and the types of clouds between the two instruments. As mentioned, the sample is comprised mostly of thin cirrus clouds, clouds that enable the lidar to completely penetrate to the top of the highest cloud layer. Unfortunately, thin cirrus clouds also represent the most difficult type of cloud for MODIS to detect. When one also considers the northern locale of the AFARS research site, where snow, low sun angles and temperature inversions are common over half the year, it is safe to assume that this sample represents one of the more difficult ones with which to evaluate the MODIS derived cloud parameters. The large discrepancies observed in Figure 4.3 can be thought of as a consequence of these difficult conditions.

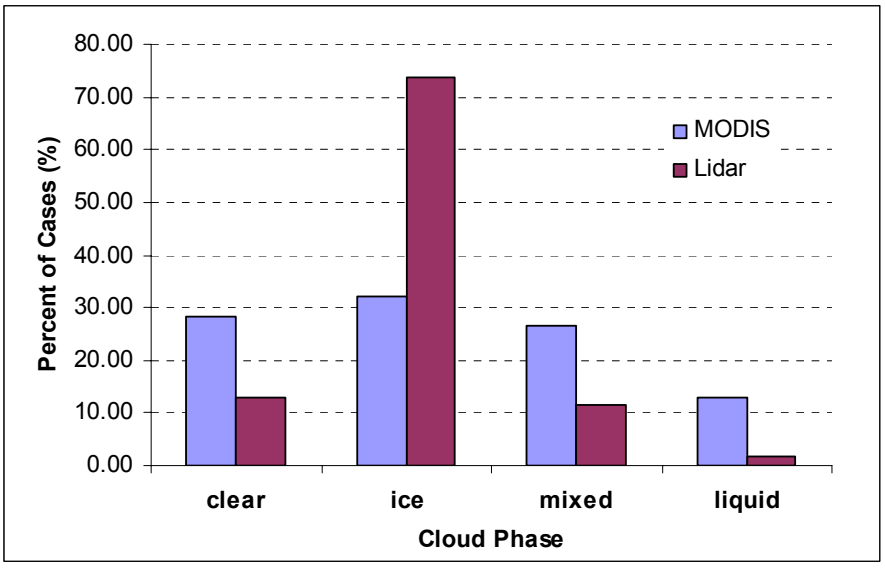


Figure 4.3: Overall percentage of cases by cloud phase sorted by instrument.

MODIS erroneously identifies *clear* conditions at nearly double the rate of the lidar. These cases are represented by the *F-N* cases in Figure 4.1. It is likely, given the problems of thin cirrus detection from satellites, many of these missed clouds are of thin *ice* clouds. This is partly supported by the observed disparity in the percentage of *ice* cloud cases between the two instruments; MODIS detects *ice* clouds approximately 32% of the time while, based on the lidar depolarization plots, *ice* clouds are detected from the surface at over double the rate, 74% of the time.

In addition it appears that MODIS overestimates the number of *mixed* and *liquid* phase clouds. In the case study presented in chapter three, the 2100 UTC Aqua overpass for the July 24, 2006 (Figure 3.1) classifies the pixel over AFARS as *water* (Table 3.1), whereas, based on the δ-plot in Figure 3.1b, the lidar observation is *ice* (Table 3.4). A more detailed analysis into the false identifications (F-N, F-P) in sections 4.5 and 4.6, will show that on occasion the MODIS algorithm can be fooled into identifying smoke as *liquid*, since the two have similar characteristics.

4.1.4 Case Summary by Month

Figures 4.4a and 4.4b give the monthly frequency of cases by platform (*Aqua*, *Terra*), time of day (*day*, *night*), snow cover (*snow*, *no snow*), MODIS cloud top retrieval method (BT_{11}, CO_2) and false identifications (F-N's and F-P's). The time of day and the snow cover cases plotted in Figure 4.4a are derived from the cloud mask bits extracted from the 5-km cloud mask SDS. Recall from chapter two that the MODIS cloud mask algorithm uses ancillary inputs, such as time of day, snow cover and surface type, to

decide which processing path (domain), and hence, cloud detection tests and corresponding threshold values are applied. The *day* versus *night* determination is based on the value of the solar zenith angle (θ) at the Earth's surface: if $\theta \leq 85^{o}$, the scene receives a *night* flag. The θ-value is predictable so one expects this flag to be accurate. Note, the majority of the *night* cases (black solid) in Figure 4.4a take place during November, December and January, during the weeks surrounding winter solstice. Even though the overpasses occur during the early afternoon, none of the documented cases in December receive a *day* flag due to the lowness of the midwinter sun in Fairbanks.

Recall that snow cover is determined either from the NSIDC snow cover maps or from the MODIS derived NSDI (Normalized Snow Difference Index). Unfortunately, the number of *snow* cases (white diamond with black dotted line in Figure 4.4a) during the winter months is significantly underestimated. Meteorological observations from the National Climatic Data Center (NCDC) confirm that a measurable snow depth was present in Fairbanks on all days spanning Nov, 1 through March, 31 for all three winter seasons (04-05, 05-06 and 06-07). Therefore, any *no snow* cases (green diamond and dotted line) recorded during these months are incorrect. This error may suggest an over reliance on the NDSI method, which uses the reflectance channels (VIS and NIR) to determine snow cover. Widespread areas of spruce forest and asphalt, both of which are dark targets, could drastically reduce the reflectance in the visible bands rendering the NDSI method ineffective at detecting bright snow. Also since it relies on the shortwave channels, the NSDI can only be applied to the *day* pixels. However, the fact that several

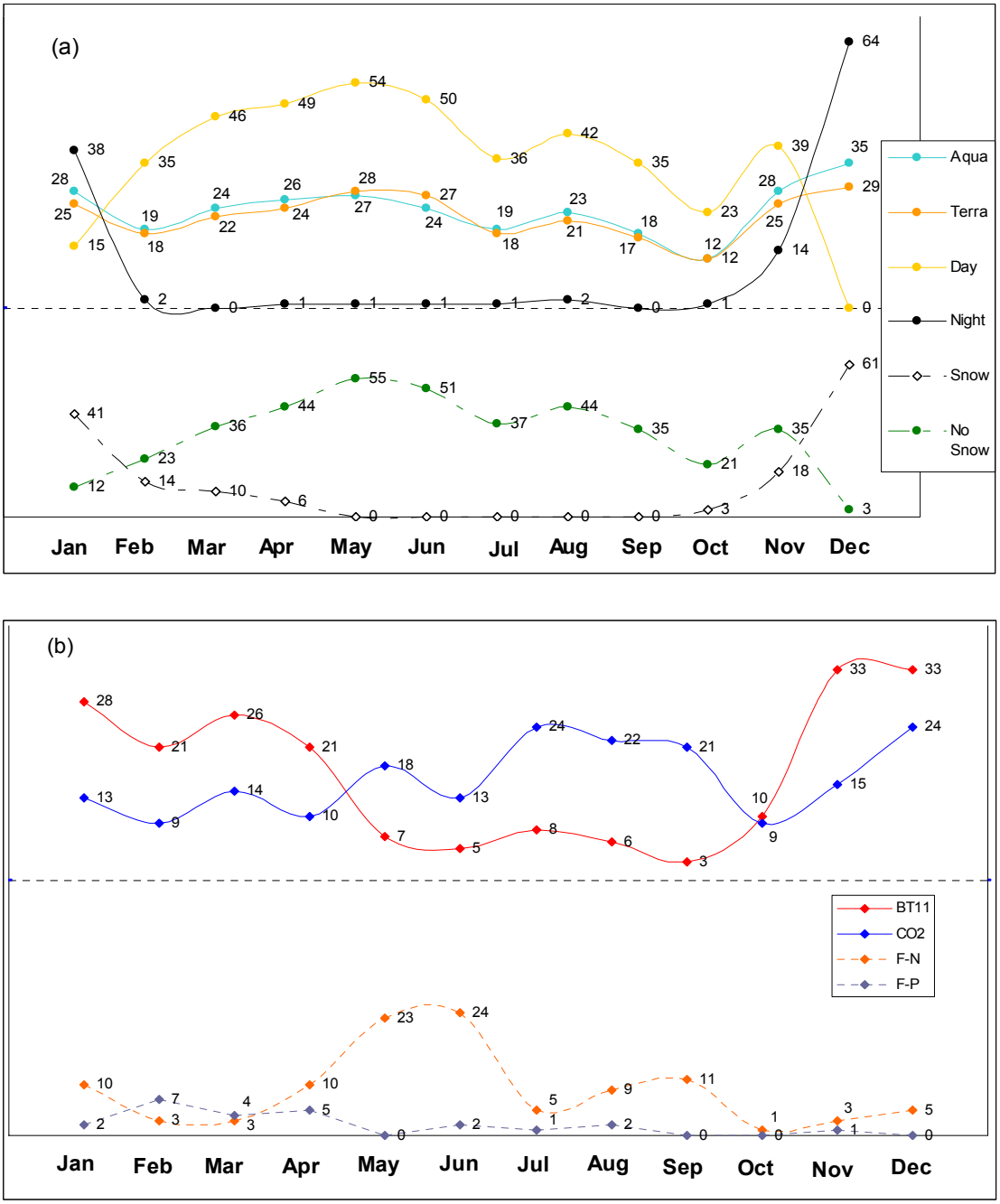


Figure 4.4: Ten MODIS overpass categories distributed by month. The cases are sorted by (a) platform, time of day, snow cover and (b) MODIS cloud top retrieval method, false negative (F-N) and false positive (F-P) cases. The day, night, snow and no snow flags are derived from the 5-km MODIS cloud mask. The top retrieval methods are from the MODIS cloud product.

no snow cases are identified in December, even though all of the cases receive a *night* flag, suggests that perhaps the NSIDC snow cover maps are inaccurate. One could find out how the snow cover determination is performed by extracting bits two and three in Byte #8 of the cloud mask, however such an analysis is not conducted for this study. At this point, it is unclear how choosing the wrong snow cover flag trickles down into the cloud mask and subsequent cloud product estimates.

Figure 4.4b gives the time series of the MODIS cloud top retrieval method and false cloud identification cases. Note the retrieval methods appear to exhibit a seasonal pattern. The frequency of CO_2 cases (solid blue) is largest during the summer months, from May to September, when the surface temperatures are warm, the ground is snow-free and the days are long. Conversely, the frequency of BT_{11} cases (solid red) is largest during the winter months, from November to April, when the days are short, surface temperatures are cold, inversions are likely, and the ground is snow covered. Recall that if the clear minus the cloudy sky radiances fall within the instrument noise level, the BT_{11} method is selected to compute the cloud top height. This occurs when the clouds are low ($P_c \leq 700$ hPa), or when cloudy and clear sky signals are difficult to separate, such as for the case of thin cirrus over a cold bright surface (Menzel et al., 2002).

The false identifications (F-N, F-P) are plotted in Figure 4.4b in conjunction with the top retrieval methods to determine if any seasonal pattern can be detected, and to see if the false identifications are in any way correlated to top retrieval method. A cursory examination of the F-N plot (dotted orange) seems to show that most of the cases seem to take place, with the exception of July, during the summer months from May to

September. The F-P series (dotted gray) is largest between February and April. Based on the cloud mask algorithm discussion, one could hypothesize that the times when MODIS would most likely misidentify cloud under clear skies (F-P) is during the winter months when the contrast in both temperature and reflectance between the ground and the cloud is small. Similarly, the times during which MODIS would most likely misidentify clear sky when a thin cirrus cloud is present (F-N) would be over a warm dark landscape. The time series of F-N and F-P in Figure 4.4b seem to support these statements.

4.2 MODIS Cloud Mask Results

4.2.1. Analysis Methods Revisited

The MODIS cloud mask was discussed at great length in section 2.3.1 Recall that the cloud mask computes a Q-score, the value of which can fall within one of four categories: *confident clear*, *probably clear*, *probably cloudy* and *confident cloudy*. For this study, the *confident clear* and *probably clear* categories as well as the *confident cloudy* and *probably cloudy* are combined into simply *clear* and *cloudy*, which facilitates the comparison with the lidar based determination which, itself, yields either *clear* or *cloudy*.

For reasons discussed in section 3.4, the accuracy of the MODIS cloud mask is assessed using three different methods. The first evaluates the cloud mask derived from the 5-km pixel derived from the cloud product (MOD06) situated directly over AFARS. This method will be dubbed '5 km point.' The second, identified as the '1 km point,' evaluates the pixel value obtained over AFARS from the actual 1-km MODIS cloud

mask file (MOD35). The third, labeled '1 km box,' compares the mode value of a 5 x 5 1-km box of pixels centered over AFARS.

The analysis methodology used to compare the MODIS cloud mask with the lidar data was explained in section 2.7.3, but will be repeated here for convenience with an example. Tables 4.1a-c summarize the performance of the cloud mask for all sample cases with Table 4.1a, 4.1b and 4.1c representing '5 km point,' '1 km point' and '1 km box' respectively. The results are counted using an IDL routine, which partitions the cases into one of four MODIS classes (*Conf. Clear*, *Prob. Clear*, *Prob. Clouds*, *Conf. Clouds*) and one of two lidar classes (*Cloudy*, *Clear*).

As discussed, the performance of the cloud mask can be approached in two ways. The first compares the MODIS result against the lidar. These calculations are given under $clear^{l}$ (turquoise) and $cloud^{l}$ (light green) in the far right column. In essence, this method attempts to answer the following: given a lidar result of clear (clouds), what is the probability that a MODIS result of clear (clouds) will match? These are computed by summing the number of *confident clear* and *crobably clear* (*confident clouds*, *probably clouds*) cases in the lidar *clear* (*cloudy*) row, and dividing by the total number of lidar *clear* (*cloudy*) cases. For example, referring to the '5 km point' method (Table 4.1a), the probability that the MODIS cloud mask algorithm returns either a *confident clear* or a *probably clear*, given a lidar result of *clear*, is [(55+6)/73]*100, or 83.56%.

The second compares the lidar result against the MODIS cloud mask finding, which in essence, is like attaching a reliability estimate to the MODIS cloud mask. These

Table 4.1: **MODIS cloud mask summary**. The results for the (a) '5 km point,' (b) '1 km point' and (c) '1 km box' methods are sorted by the lidar observation.

(a) **MOD06 '5 km point'**

Lidar Cloud	Conf. Clear	Prob. Clear	Prob. Clouds	Conf. Clouds	Total	% Correct *Clear*[1]
Cloudy	142	10	31	293	*476*	83.56
Clear	55	6	1	11	*73*	*Cloud*[1]
Total	197	16	32	304	549	68.07
% Corr.	*Clear*[2] =	28.64	*Cloud*[2] =	96.43		

(b) **MOD35 '1 km point'**

Lidar Cloud	Conf. Clear	Prob. Clear	Prob. Clouds	Conf. Clouds	Total	% Correct *Clear*[1]
Cloudy	145	5	30	295	*475*	71.23
Clear	50	2	5	16	*73*	*Cloud*[1]
Total	195	7	35	311	548	68.42
% Corr.	*Clear*[2] =	25.74	*Cloud*[2] =	93.93		

(c) **MOD35 '1 km box'**

Lidar Cloud	Conf. Clear	Prob. Clear	Prob. Clouds	Conf. Clouds	Total	% Correct *Clear*[1]
Cloudy	149	2	20	304	*475*	76.71
Clear	56	0	4	13	*73*	*Cloud*[1]
Total	205	2	24	317	548	68.21
% Corr.	*Clear*[2] =	27.05	*Cloud*[2] =	95.01		

calculations are delineated next to the *clear*[2] (yellow) and *cloud*[2] (tan) headings along the bottom row. One way to interpret these results is to consider the following: if MODIS determines that a given pixel is clear (cloudy), how accurate is this finding against the lidar, considered to be the ground truth? This probability is computed by summing the number of *confident clear* and *probably clear* (*confident clouds*, *probably clouds*) cases within the lidar *clear* (*cloudy*) row and dividing by the total number of the MODIS

confident clear and *probably clear* (*confident clouds*, *probably clouds*) cases. For example, again using '5 km point' (Table 4.1a), assuming that the MODIS cloud mask yields a result of either *probably* or *confident clear*, the probability that the sky is, in fact, clear is [(55 + 6)/(197 + 16)]*100 or 28.64%. Both evaluation methods are factual, but may yield a wide range of probabilities, which must be interpreted accordingly.

The cloud mask results are further conditionally evaluated by platform (*Aqua*, *Terra*), snow cover (*snow*, *no snow*) and time of day (*day*, *night*), the latter two fields being derived from the cloud mask bits. To conserve space, the results for all categories are condensed from twenty-seven tables similar to Table 4.1, into Figures 4.5a and 4.5b. The former summarizes the *clear-sky* probabilities and the latter summarizes *clouds*. The seven aforementioned groups are labeled along the x-axis. Black represents the '5 km point' method, red represents '1 km point,' and blue '1 km box.' Solid circles (diamonds) signify either the $clear^1$ or $cloud^1$ ($clear^2$, $cloud^2$) percentage assessments respectively. The hollow circles and diamonds signify the number of cases associated with each series, with the shape and color matching the shape and color of the respective category. The symbols are labeled with the percentage values or the case numbers for convenience.

To help clarify how the results in Table 4.1 are presented in Figure 4.5, note that the $clear^1$ (83.6, 71.2, 76.7) and $clear^2$ (28.6, 25.7, 27.1) percentages in Figure 4.5a above *All* coincide with the $clear^1$ and $clear^2$ percentages in Table 4.1 highlighted in turquoise and yellow. The same can be said for the $cloud^1$ and $cloud^2$ percentages in Figure 4.5b.

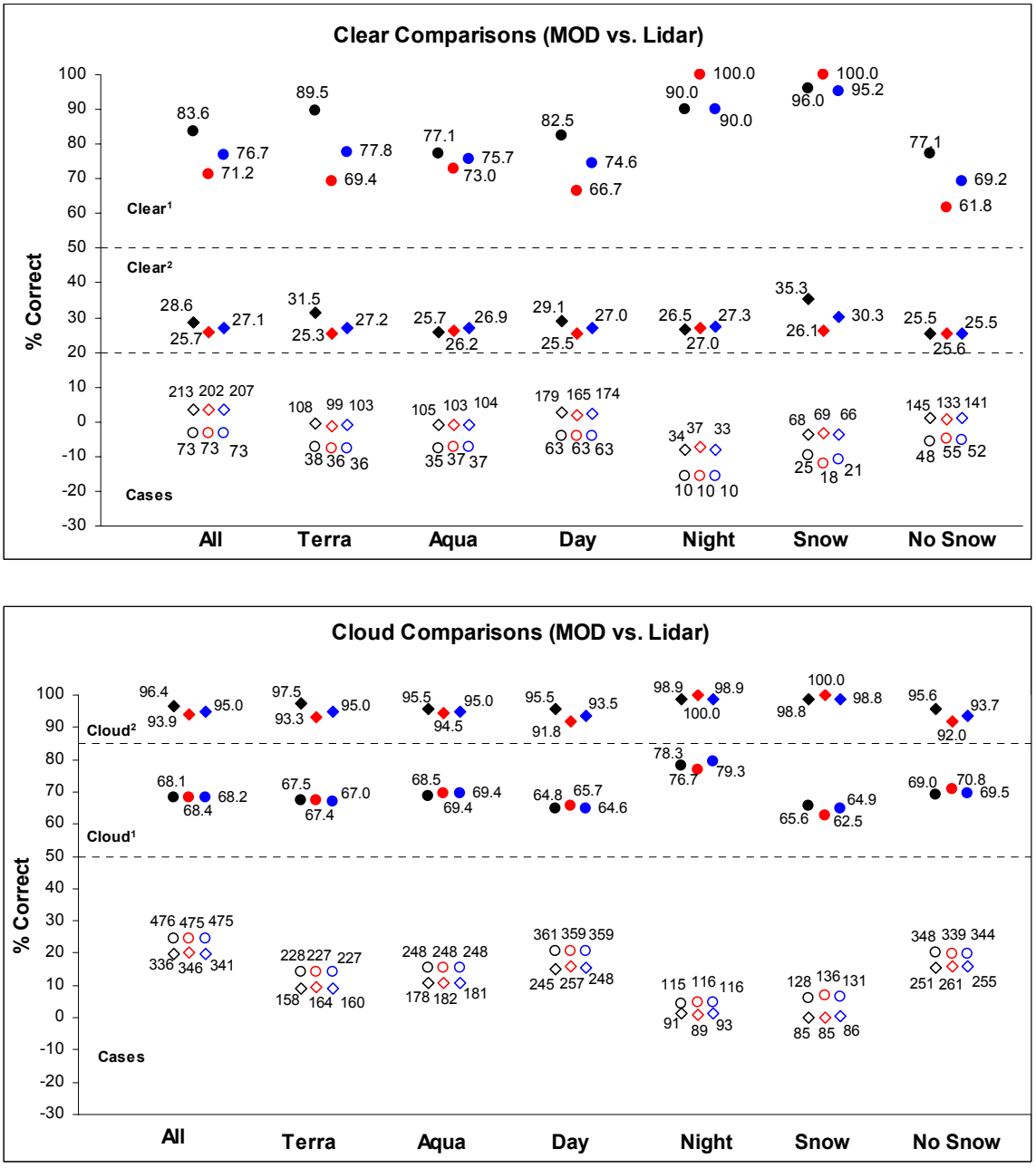


Figure 4.5: Summary of MODIS cloud mask agreements by analysis approach for the a) clear sky and (b) cloudy sky results sorted by the categories *All*, *Terra*, *Aqua*, *Day*, *Night*, *Snow* and *No Snow* are printed along the x-axis. Diamonds signify approach[2], circles signify approach[1]. Solid black, red and blue denote percent agreements for the '5 km point,' '1 km point' and '1 km box' methods respectively. Hollow symbols denote the corresponding number of cases.

Notice also that the case numbers associated with each category, denoted as hollow diamonds and circles, are provided along the bottom. The actual number of *clear*^{1} (73, 73, 73) and *clear*^{2} (213, 202, 207) cases are printed next to the symbols, which match the case numbers in Table 4.1, as do the *cloud*^{1} and *cloud*^{2} cases in Figure 4.5b.

4.2.2 Results by Method: '5 km point' vs. '1 km point' vs. '1 km box'

As Figure 4.5 illustrates, differences in the probabilities between the three methods ('5 km point,' '1 km point,' '1 km box') appear to be minimal. With the exception of the *night* and *no snow* categories, the '5 km point' surprisingly outperforms the other two by about 10-15% for *clear*^{1} and 1-4% for the *cloud*^{2}. Also '1 km box' appears to be slightly more accurate (by 1-8%) than '1 km point' for those same cases. With respect to the *clear*^{2} and *cloud*^{1} approaches, the variability between the three methods appears to be insignificant. Overall, '5 km point' seems to be the most accurate method. One should briefly mention that the number of 1-km cases derived from the cloud mask MOD35 files is one less, 548, than the number from the cloud product MOD06 files because it was found that the April 16, 2006, 2120 UTC Terra overpass returned an 'ND' signifying *not determined* for the cloud mask flag. Hence this case was omitted from the cloud mask analysis.

4.2.3 Cloud Mask Summary

The mean and standard deviations of percent agreement (correct) for the three methods are presented in Figure 4.6. The *cloud*^{2} examination consistently yields the best

agreement; assuming that the MODIS cloud mask detects cloud, the chances of this agreeing with the lidar ranges from 93.8% for *no snow* to 99.3% for *night*. Conversely, if the lidar has established that cloud is overhead, the probability that MODIS will agree (i.e. *cloud*^{1}) is much worse, with the low and high-end outcomes of 64.3% to 78.1% for the *snow* and *night* cases respectively.

With respect to the clear comparisons, the MODIS cloud mask seems to do a decent job with the *clear*^{1} approach; if clear conditions over AFARS are initially confirmed by the lidar, the probability that the cloud mask agrees ranges from 69.4% to 97.1% respectively for the *snow* and *no snow* cases, with a median value falling around

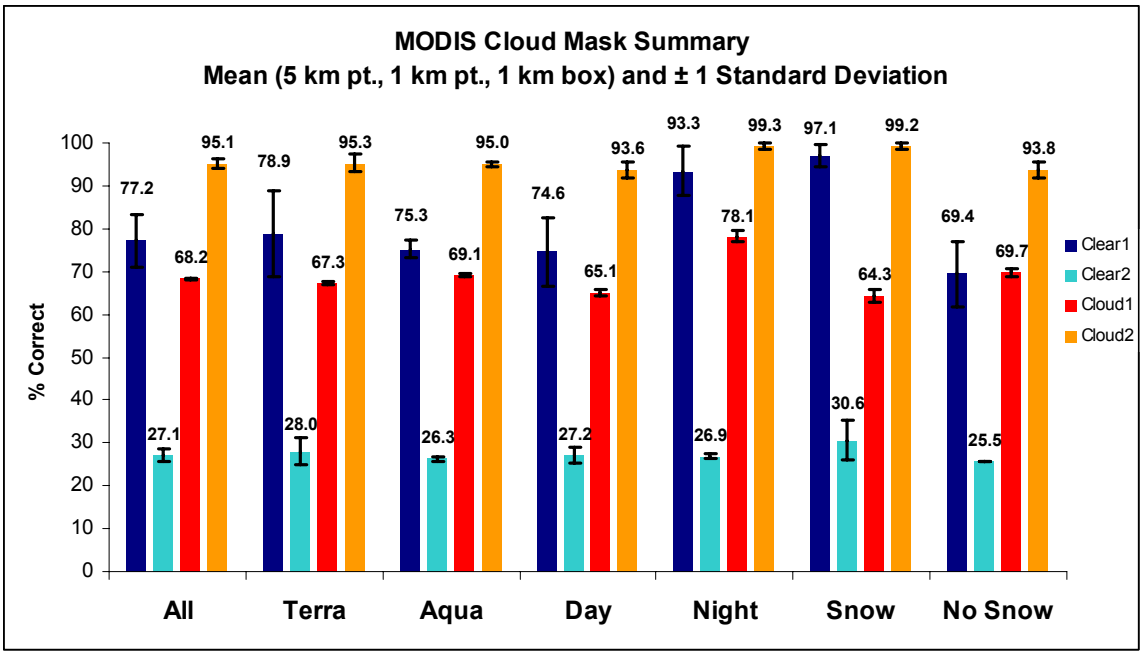


Figure 4.6: Summary of cloud mask accuracy scores. Each columns and bars represent the % correct (agreement) mean and standard deviations derived from the '5 km point,' '1 km point' and '1 km box' methods respectively for the *clear*^{1} (blue), *clear*^{2} (turquoise), *cloud*^{1} (red) and *cloud*^{2} (yellow) analysis.

77%. However note that the for many of the categories in Figure 4.5a, the *clear*^{1} group contains the smallest sample sizes, with *night* and *snow* containing only ten and less than twenty-six cases respectively. Therefore, one should view these numbers with some skepticism.

The worst results, by far, are associated with the *clear*^{2} approach. Given a MODIS cloud mask result of *probably* or *confident clear*, the chances that these results match a lidar *clear* finding falls under 30% for most categories. This finding is consistent with the large number of *F-N* (false negative) cases denoted in Figures 4.1 and 4.3, and the explanation provided in section 4.1.3.

Differences in the results by platform (*Aqua* vs. *Terra*) appear to be unremarkable. The results for all approaches (*clear*^{1}, *clear*^{2}, *cloud*^{1}, *cloud*^{2}) fall within 2-4% of one another for just about all categories. Terra slightly outperforms Aqua in the *clear*^{1} and *clear*^{2} approaches but also yields a relatively high amount of deviation between the three methods.

The accuracy of the MODIS cloud mask based on time of day (*day*, *night*) produces a surprising finding; *night* significantly outperforms *day* by about 19% for the *clear*^{1}, 13% for *cloud*^{1} and 6% for *cloud*^{2}. Recall that the number of cloud tests associated with the *night* domain is theoretically fewer that the number available to the *day* domain. *Day* should be more accurate than *night*. However, notice, again, the discrepancy in the sample sizes; there are three times as many *day* cases associated with the *cloud*^{1} and *cloud*^{2} analysis as there are *night* cases, and there are only ten *clear*^{1} cases associated with *night* (see Figure 4.5a).

With respect to the *cloud*^{1} and *cloud*^{2} analyses, one might expect that the cloud mask is more likely to misidentify clouds at night over the cold ground when inversions are more likely to form. But if this were the case, then one might also observe a similar result for the *snow* flagged cases. However, comparison between the scores for *snow* and *no snow* cases are mixed: *snow* outperforms *no snow* by about 6% for the *cloud*^{2} cases, but *no snow* is over 5% better than *snow* with respect to *cloud*^{1}. Also *snow* significantly outperforms *no snow* by about 28% in the *clear*^{1} category, and although 30.6% agreement in the *clear*^{2} cases can be considered a poor result, this finding is still 5% better than for the similar *no snow* calculation. This also represents the highest *clear*^{2} agreement for all the categories.

4.3 Cloud Top Property Results

4.3.1 Summary of Pressure, Height and Temperature Estimates

Omitting the twenty-four *a-l** in Figure 4.1, there exist 345 cases of documented overlapping cloud top heights between the satellite and the lidar. As described in chapter three, atmospheric profiles of temperature, pressure and altitude collected from the nearest radiosonde launchings are used to interpolate the corresponding pressure (hPa) and temperature (^{o}C) from the lidar cloud top height (m), and the corresponding height and temperature from the MODIS cloud top pressure. One readily acknowledges that the mean, median and max time delays between the MODIS overpasses and the radiosonde launchings are found to be 112, 110, and 245 minutes respectively. This discrepancy will introduce some unknown amount of error into the calculations, which will depend on the

delay in between the overpass and the radiosonde launching and on the weather conditions during this delay.

The cloud top height (pressure, temperature) derived for each lidar case is subtracted from the corresponding MODIS value yielding a ΔZ (ΔP, ΔT) error: the smaller the error (ΔZ, ΔP, $\Delta T \approx 0$), the greater the accuracy. These errors are evaluated against a number of conditions.

Figure 4.7 plots the raw MODIS derived cloud top parameter of pressure, height and temperature as a function of the corresponding lidar derived values. The blue dots represent values derived from the CO_2 method and red represents the BT_{11} method. Dashed lines signify ± 200 hPa, ± 2000 m and ± 10^{o}C respectively. Points residing within these arbitrarily chosen error ranges offer reasonable approximations, based on the range of modeled pressure error estimates presented in the ATBD (Menzel et al., 2002).

First, note the general tendency of the MODIS cloud top algorithms to underestimate (overestimate) the cloud top height (temperature, pressure). This finding is consistent with the discussion in section 2.3.2.2, which mentions that the CO_2 density slicing algorithm measures the "radiative center-of-mass" of the cloud, which, optically thin cirrus cloud, falls near the geometric center (Menzel et al., 2002). Theoretically, the magnitude of the cloud top error is inversely proportional to the optical thickness of the cloud and is also affected by the number and thickness of the underlying cloud layers.

Figure 4.7 also illustrates the large disparity in the cloud top estimates between the CO_2 (blue) and the BT_{11} (red) derived top retrieval methods. The largest clusters of acceptable estimates, those falling within the dashed lines, happen to be the cases derived

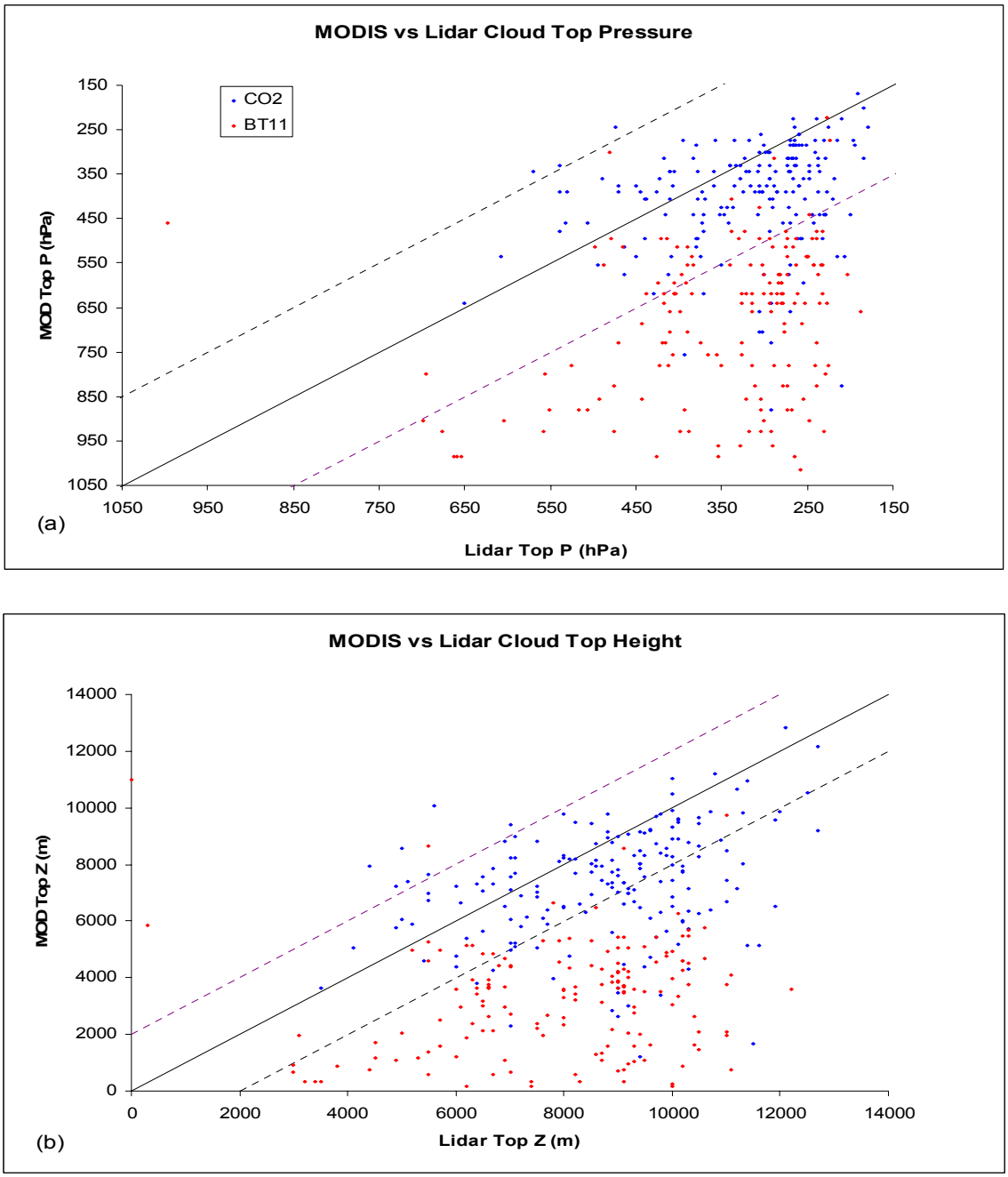


Figure 4.7: **Scatter-plots of MODIS cloud top parameter versus lidar parameter**. (a) MODIS cloud top pressure versus lidar cloud top pressure (hPa) and (b) MODIS cloud top height versus lidar cloud top height (m) representing the 345 $L_{cld} - M_{cld}$ cases in Figure 4.1. Blue and red dots denote CO_2 density slice and BT_{11} methods respectively. Solid lines represent "zero-error" while dashed lines denote 200 hPa and 2000 m.

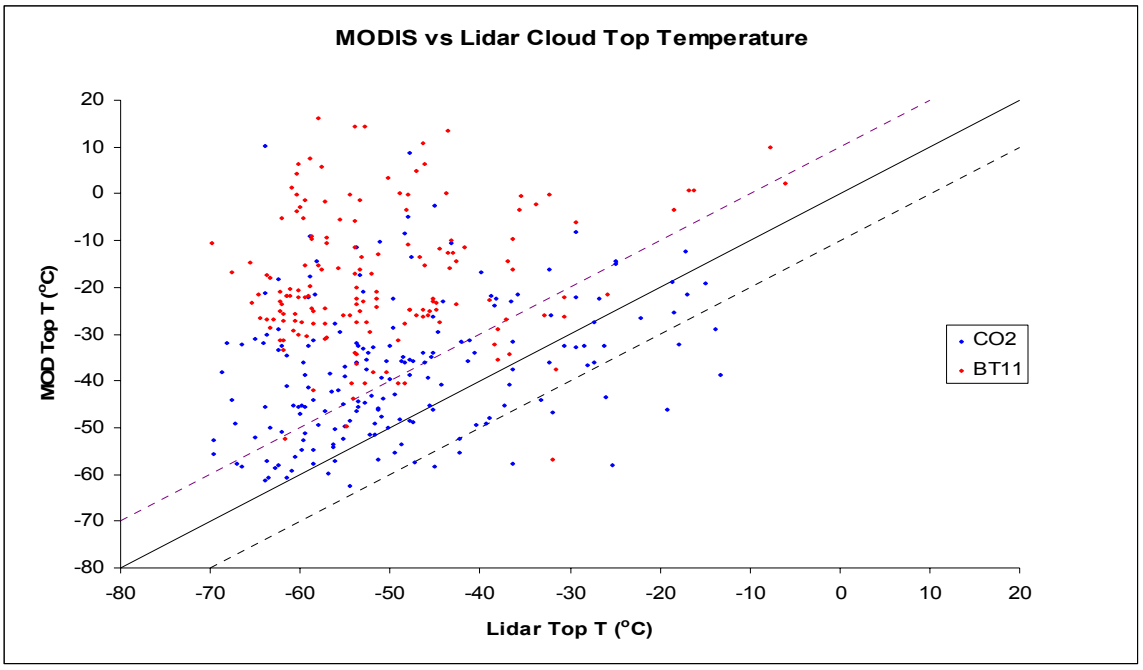


Figure 4.7 (cont'd): Corresponding scatter-plot of (c) MODIS cloud top temperature versus lidar top temperature (^{o}C). Dashed lines denote 10 ^{o}C.

from the CO_2 slicing algorithm (blue), while the least accurate (red) are derived from the BT_{11} method. Also note that with the exception of one or two cases, almost all BT_{11} estimates consistently fall above (below) the solid "zero-error" lines for the cloud top temperature (pressure, height) plot, while the CO_2 estimates generate a more skewed Gaussian appearance around the "zero-error" line.

These behaviors are best depicted in histogram form in Figures 4.8, 4.9 and 4.10. The bins represent different ranges in cloud top error (ΔP, ΔZ, $\Delta T = MODIS_{top} - Lidar_{top}$) estimates ranging from $\Delta P \leq -400$ to $\Delta P > 400$ hPa in 100 Pa intervals for pressure (Figure 4.8), $\Delta Z \leq -8000$ to $\Delta Z > 8000$ m in 2000 m intervals for height (Figure 4.9), and $\Delta T \leq -40$ to $\Delta T > 40$ ^{o}C in 10 ^{o}C intervals for temperature (Figure 4.10). Red, blue and

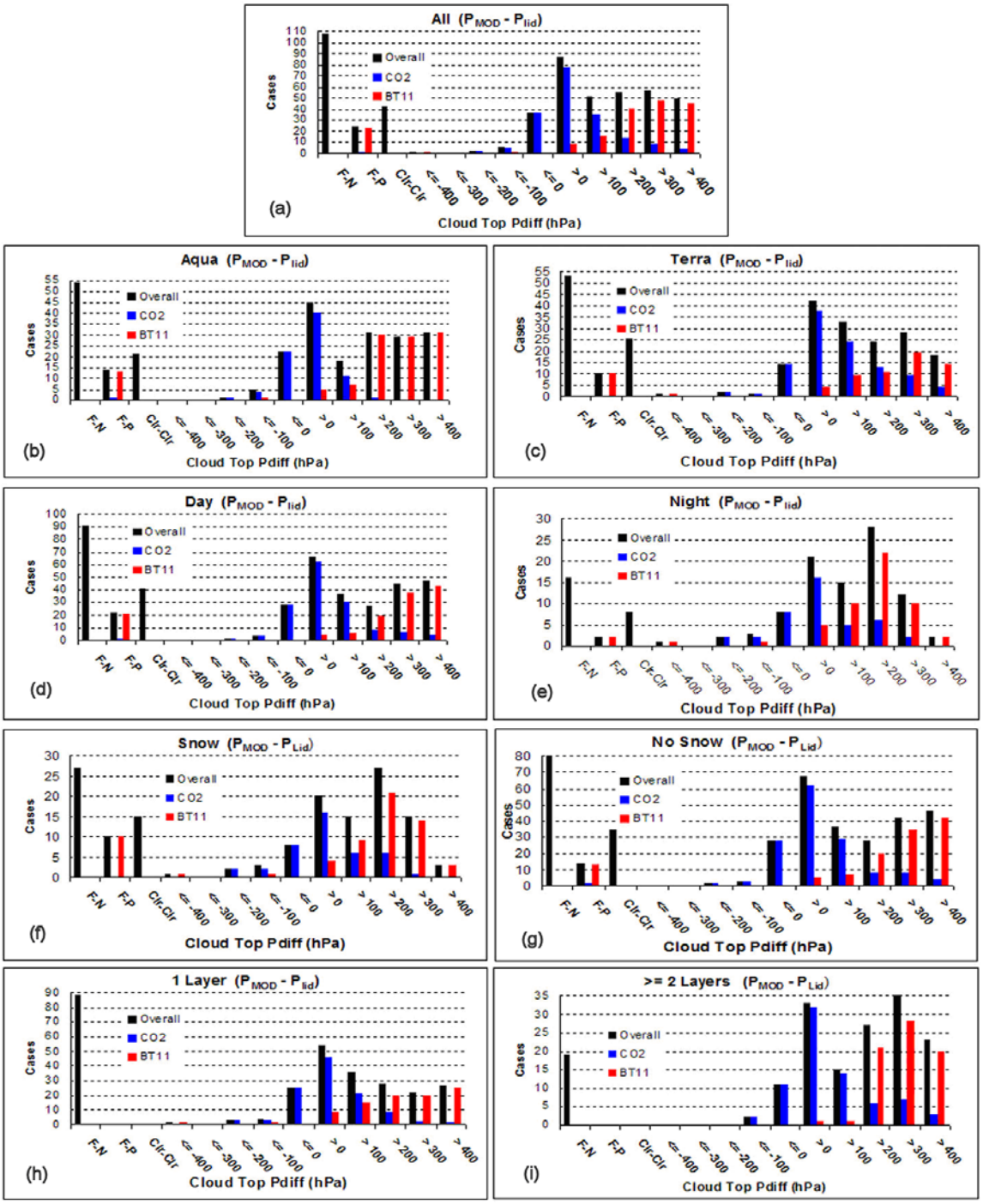


Figure 4.8: Histograms of cloud top pressure errors: ($\Delta P = P_{MOD} - P_{Lid}$). The categories are for (a) all, (b) Aqua, (c) Terra, (d) day, (e) night, (f) snow, (g) no snow, (h) single cloud layer, (i) multiple cloud layers. Blue, red and black denote CO_2, BT_{11} and all cases respectively; 'F-N'- false negatives; 'F-P'- false positives; 'Clr-Clr'- (L_{clr} - M_{clr}).

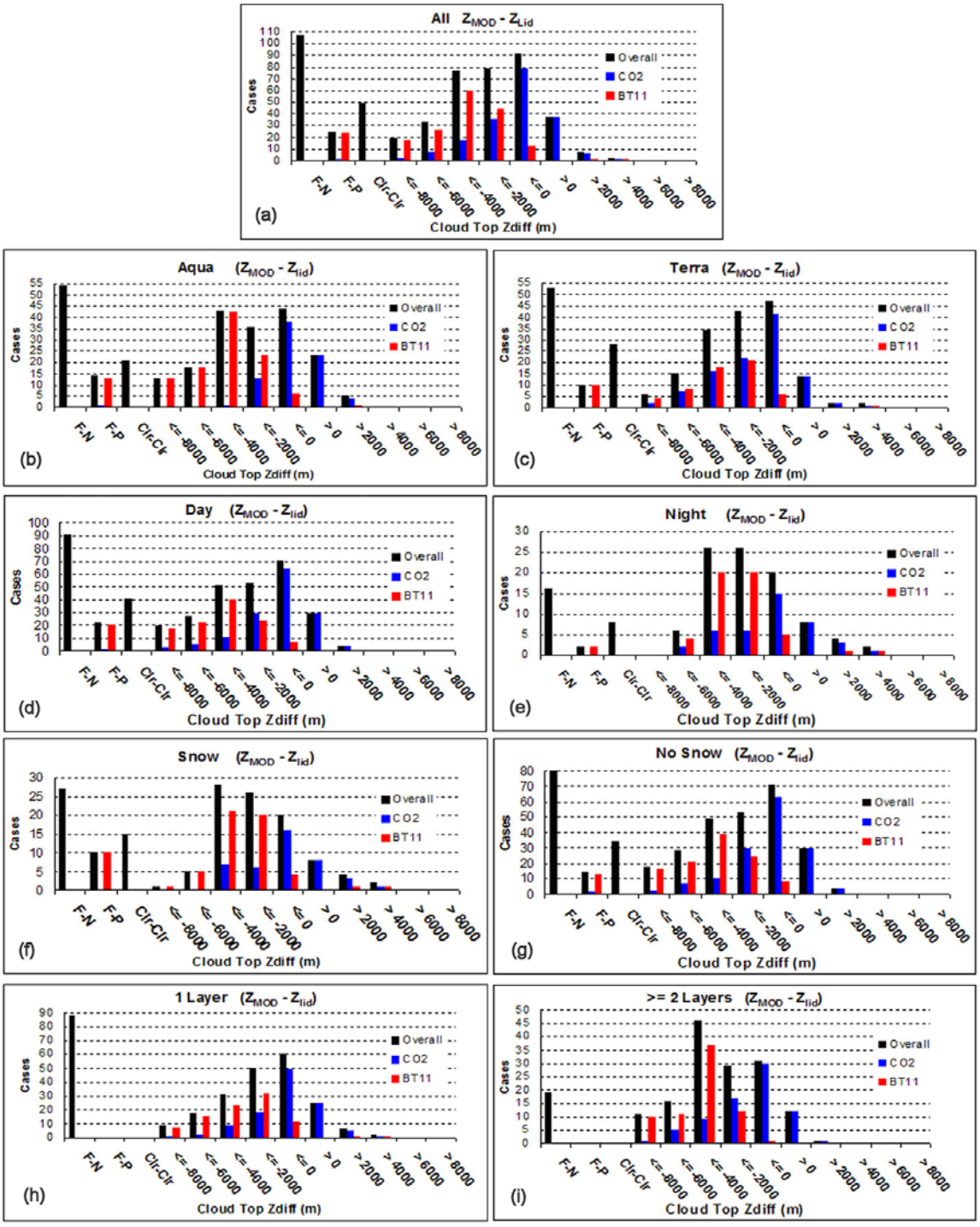


Figure 4.9: Histograms of cloud top height errors: ($Z = Z_{MOD} - Z_{Lid}$). The categories are (a) All, (b) Aqua, (c) Terra, (d) Day, (e) Night, (f) Snow, (g) No Snow, (h) single cloud layer and (i) multiple cloud layers. Categories, colors, and labels are analogous to those in Figure 4.8.

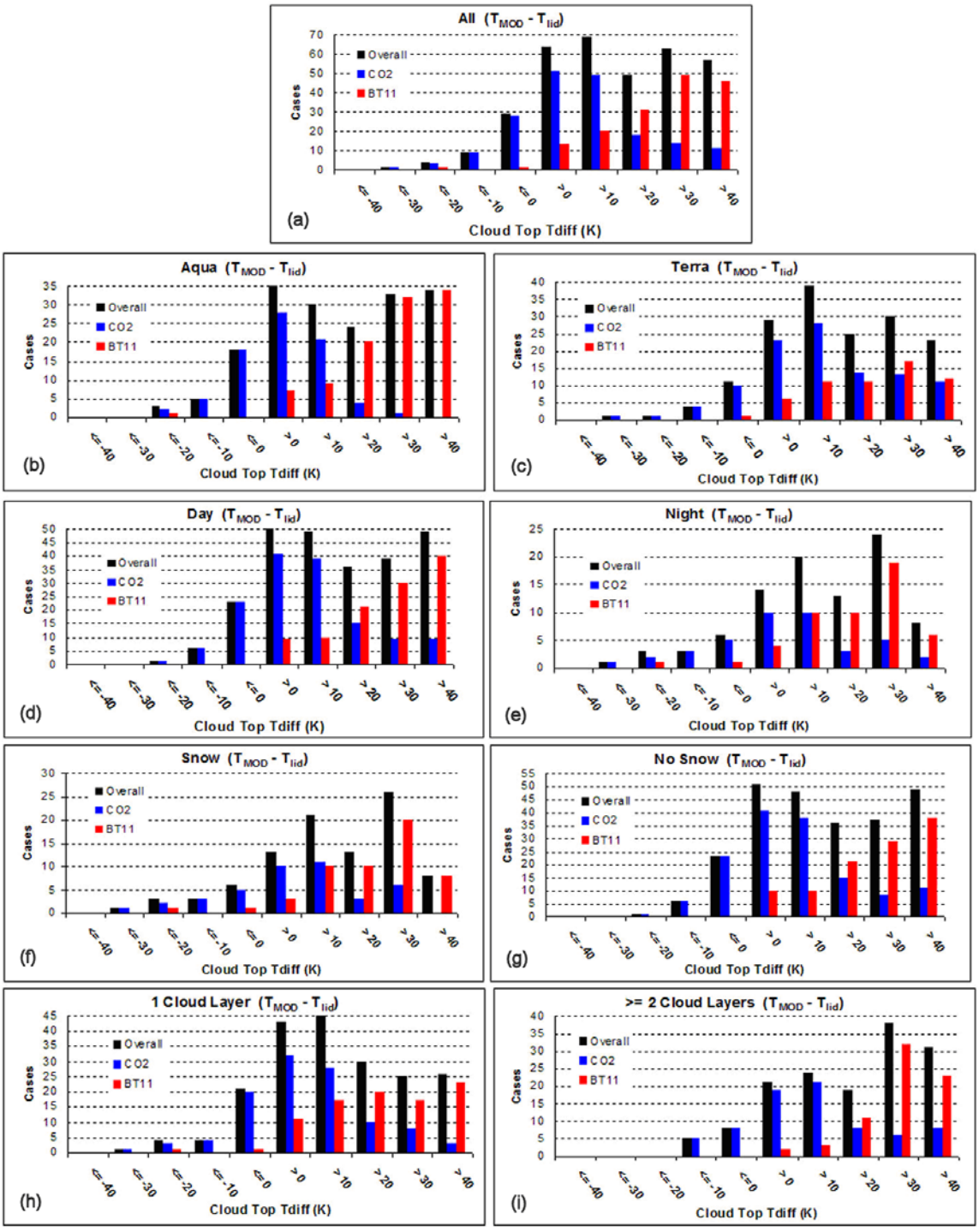


Figure 4.10: Histograms of cloud top temperature errors: ($T = T_{MOD} - T_{Lid}$). The categories are (a) All, (b) Aqua, (c) Terra, (d) Day, (e) Night, (f) Snow, (g) No Snow, (h) single cloud layer and (i) multiple cloud layers. Categories, colors and labels are analogous to those in Figure 4.8.

black represent the BT_{11}, CO_2 and both (BT_{11} + CO_2) respectively. Each set of figures contains nine plots. *All* provides the distribution for all cases. The others are separated by platform (*Aqua*, *Terra*), time of day (*day*, *night*), snow cover (*snow*, *no snow*), and cloud layers (*1 layer*, ≥ *2 layers*). The pressure and height plots also give the bins for the false-negative (*F-N*), false positive (*F-P*), and the $L_{clr} - M_{clr}$ (clr-clr) cases respectively.

As in Figure 4.7, the histograms demonstrate the tendency of the MODIS cloud top algorithm to significantly underestimate (overestimate) cloud top height (pressure, temperature). Similarly, the BT_{11} cases (red) contribute the largest source of error, while the CO_2 (blue) cases, although generally skew towards a lower (higher) cloud top height (pressure, temperature), exhibit a more symmetrical distribution around zero. Also note that the BT_{11} estimates cause the majority of F-P cases. Only one case of F-P originating from a CO_2 derived case is documented, which is associated with a *day-aqua-no snow* overpass.

Figure 4.11 gives the mean and 1 standard deviation of the cloud top errors for each of the nine categories (a-i) represented in the histogram sets (Figure 4.8, 4.9, 4.10). The categories are similarly color coded with blue red and black representing the CO_2, BT_{11} and all (CO_2 + BT_{11}) cases respectively. As in Figure 4.5, hollow symbols represent the number of cases corresponding to the above top retrieval method and the case numbers are labeled for convenience.

On first inspection, it seems that the MODIS cloud top height algorithm performs rather dismally. The average cloud top pressure difference for all 345 L_{cld}-M_{cld} cases is 196.8 hPa, corresponding to height and temperature errors of -3008 m and 21.3 K

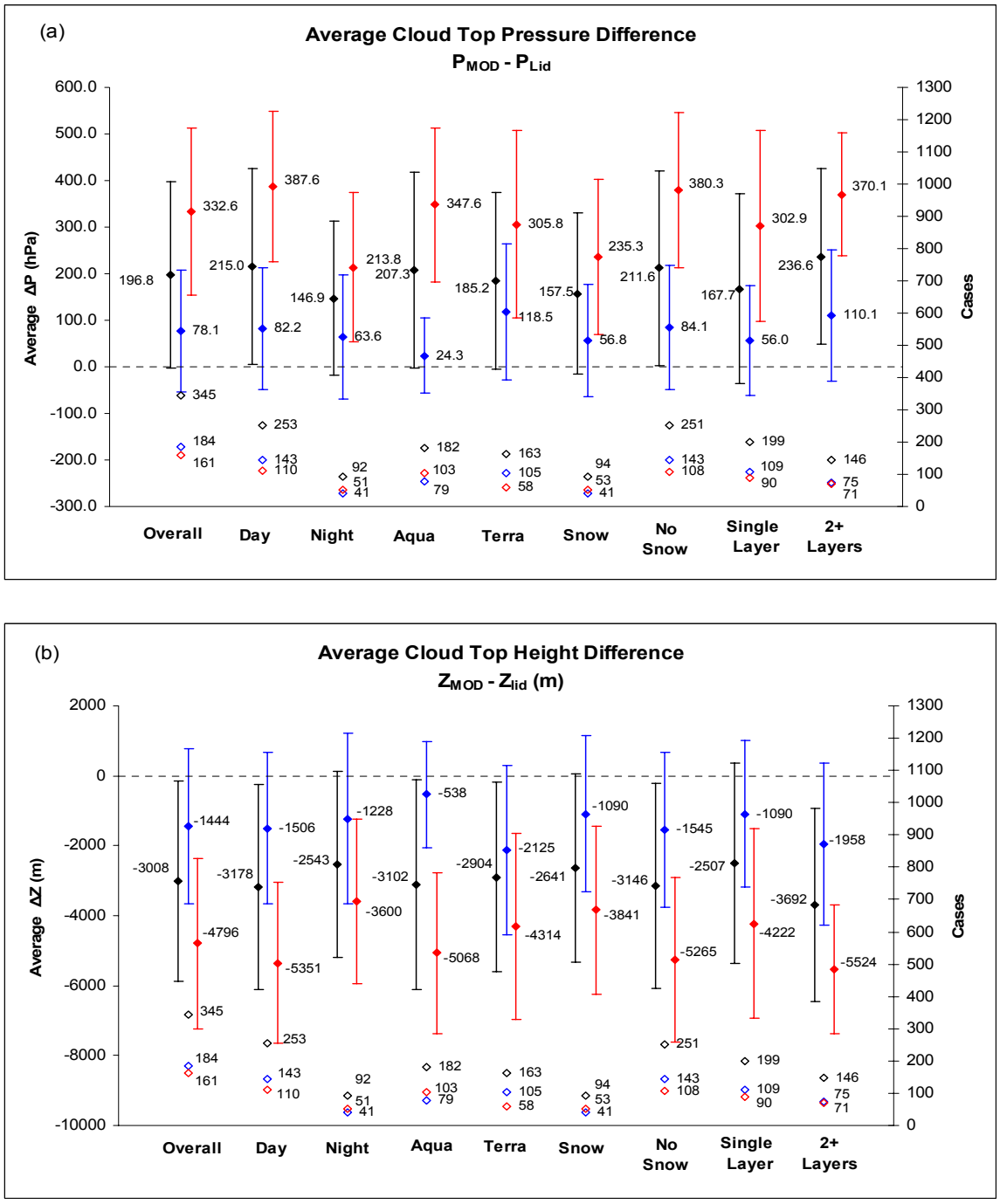


Figure 4.11: **Mean and 1 standard deviation of cloud top error by category**. Parameters include (a) pressure (hPa) and (b) height (m) error sorted by the categories listed along the x-axis. Blue, red and black denote CO_2, BT_{11} and both (CO_2 + BT_{11}) derived estimates. Hollow symbols denote associated case numbers. Values are printed next to their respective symbols for convenience.

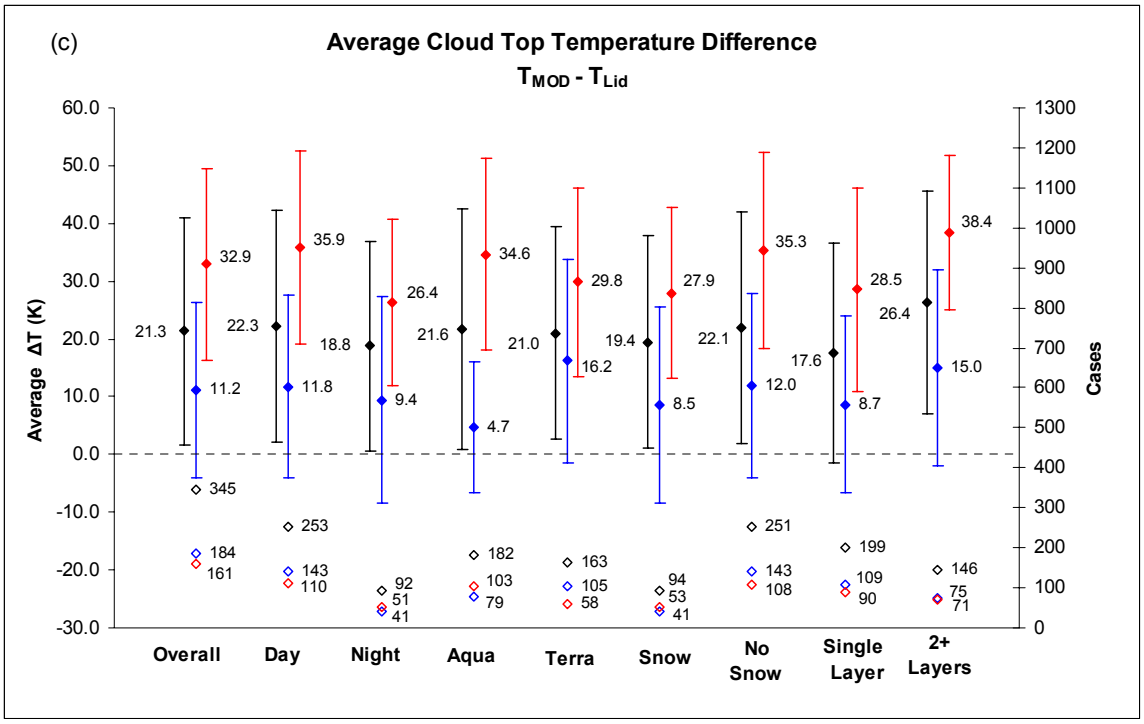


Figure 4.11 cont'd: Mean and ± 1 standard deviation of (c) cloud top temperature (K). Categories, colors and symbols are synonymous with those presented in Figure 4.11a and 4.11b.

respectively. However, it is clear that the source of the error is primarily weighted within the BT_{11} estimates, which, out of 161 cases, yield a deplorable -4796 m, 332.6 hPa and 32.9 K mean error for height, pressure and temperature. In the context of this heavily biased thin cirrus sample, the BT_{11} results can, for the most part, be discounted since this method only works for low optically thick clouds. It seems that the root cause of these large BT_{11} errors lie in MODIS's failure to detect thin cirrus layers while, simultaneously, detecting a nonexistent low cloud. Such a scenario would lead to the enormous disparities between the two estimates observed in the results.

In contrast, the CO_2 estimates for cloud top height, pressure and temperature fall much closer: -1444 m, 78.1 hPa and 11.2 K respectively. These findings are consistent with the ATBD model calculations and with earlier studies discussed in sections 2.3.2.2 and 2.3.2.3. The remaining discussion will focus solely on the CO_2 results.

Once again, similar to the cloud mask results, *night* on average, outperforms *day* by about 70 hPa, 600 m and 4 K, but note that the *day* sample is three times larger than *night*. When comparing the CO_2 derived heights, the differences between *night* and *day* are only separated by about 20 hPa, 270 m and 2 K respectively.

Conversely, the differences in the CO_2 derived errors between *Aqua* and *Terra* are substantial. Of the nine categories, *Aqua* yield the smallest amount of mean P, Z and T errors of 24.3 hPa, -538 m and 4.7 K, while for *Terra*, the corresponding values are four to five times greater: 118.5 hPa, -2125 m, and 16.2 K. It is unclear why Aqua's CO_2 estimates are substantially more accurate than Terra's even though the cases are, in general, evenly distributed throughout the year (Figure 4.4a). Also, one must keep in mind that both sensors presumably capture similar cloud cover conditions since their overpass times occur less than fifteen minutes apart. Worth noting, however, is that *Terra* contains a significantly higher percentage of CO_2 derived cases, 64.4%, to *Aqua's*, 43.4%, even though Aqua's sample size of 182 is nineteen more than Terra's 163. This finding seems somewhat paradoxical; Terra more often correctly chooses to use the CO_2 method, but Aqua's approximations are significantly more accurate. Such differences may lead one to speculate the existence of design or calibration inconsistencies between the two MODIS instruments.

With respect to snow cover, *snow* cases defeat *no snow* by 30 hPa, 500 m, and 3.5 K respectively. As with the cloud mask results, the fact that *snow* outperforms *no snow* seems unexpected given how snow camouflages clouds both visually and thermally. However differences in sample sizes may, yet again, play some role in the observed error disparity; *no snow* contains nearly three times as many CO_2 derived cases as *snow*.

One expected outcome is the link between the error magnitude and the number of cloud layers. Multiple cloud layers are observed over 42% of the time within the L_{cld} - M_{cld} sample. As discussed, multiple cloud layers will tend to overestimate (underestimate) cloud top pressure (height) for optically thin cloud layers. This idea seems to account for the findings presented in Figures 4.11. The average CO_2 derived ΔP, ΔZ and ΔT errors associated with the *single layer* are 56 hPa, -1090 m and 8.7 K respectively. For the *2+ layer* cases, the error magnitudes nearly double to 110.1 hPa, -1958 m and 15 K. Given that the sample is biased towards thin cirrus clouds, these findings are not unexpected.

Finally, Figure 4.12 gives the mean and ± 1 standard deviation of ΔP, ΔZ and ΔT by month to determine if the observed errors exhibit a seasonal correlation. The colors, symbols and error bars in Figure 4.12 are synonymous with those in Figure 4.11. Also, note the number of CO_2 and BT_{11} cases by month presented earlier in Figure 4.4b do not correspond with the values in Figure 4.12, since the latter omits *F-P* and *att-lim* cases.

Unfortunately, samples containing fewer than twenty cases occur eight of the twelve months for the BT_{11} and nine of the twelve for CO_2. Some samples even contain fewer than ten cases, making it difficult to draw any definitive statistically significant

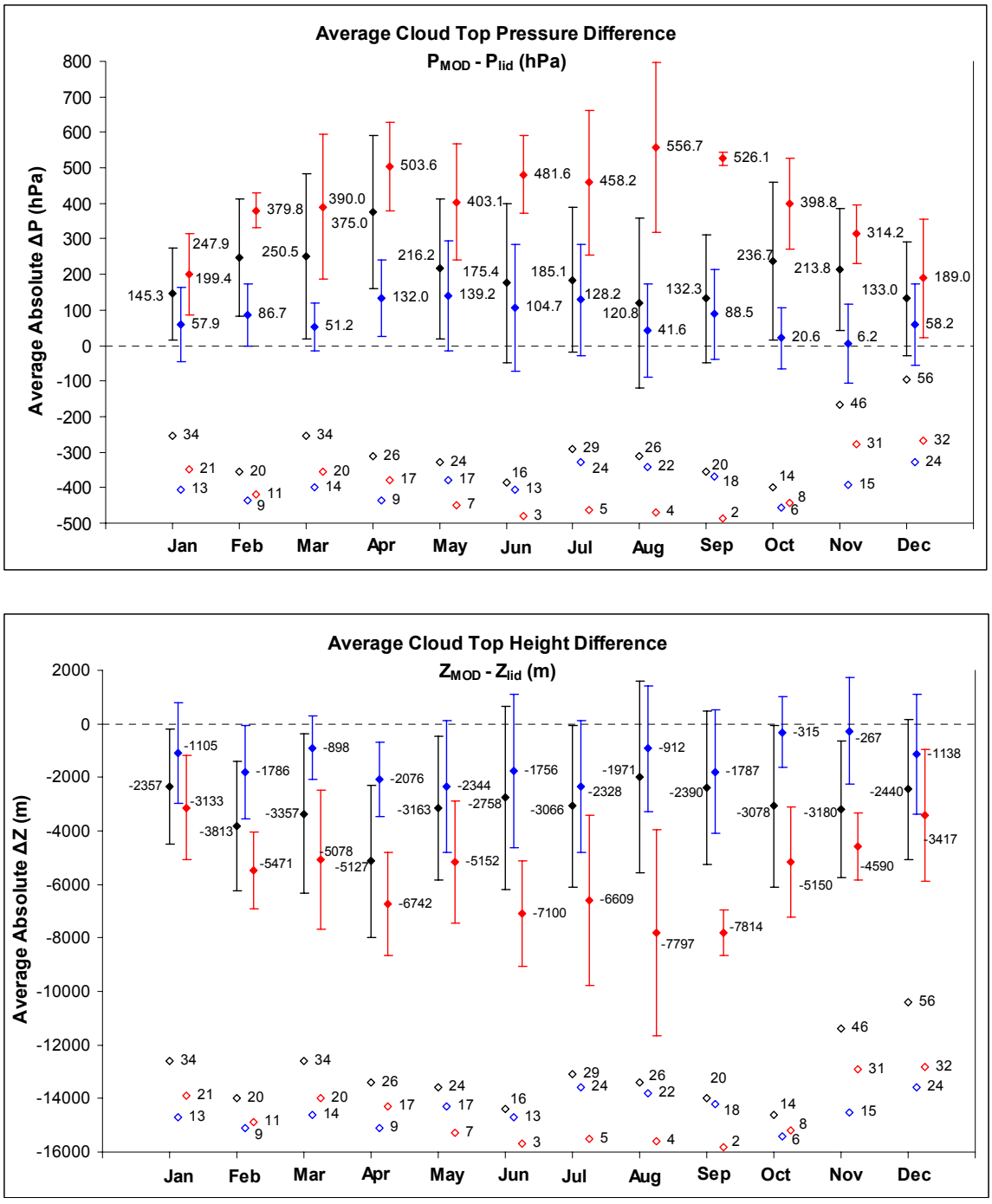


Figure 4.12: **Mean and 1 standard deviation of cloud top error by month**. Parameters include cloud top (a) pressure (hPa) and (b) height (m). Blue, red and black signify the CO_2, BT_{11} and both (CO_2 + BT_{11}) derived estimates respectively. Hollow symbols represent sample sizes associated with each group. The values are printed next to the symbols for convenience.

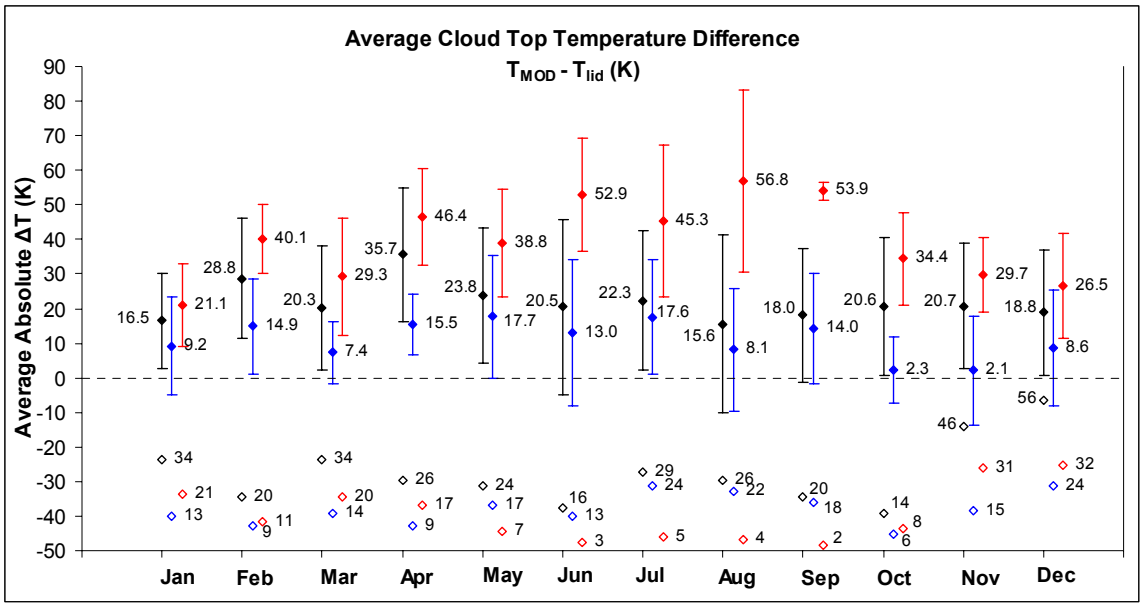


Figure 4.12 cont'd: Mean and ± 1 standard deviation of (c) cloud top temperature (K). Colors and symbols are analogous to those in Figure 4.12a and Figure 4.12b.

conclusions regarding the influence of the seasons on the satellite derived estimates of the cloud top parameters. But again, as discussed in section 4.1.4, the selection of cloud top retrieval method demonstrates a seasonal correlation; the majority of the retrievals during the warmer months (May – September) are carried out using the CO_2 method, while BT_{11} is preferred during the cold months (October to April). It is suggested that the cold snow covered ground increases the risk that MODIS falsely detects a low cloud, while the risk of this occurring during the warm snow free months is smaller for reasons described in section 4.1.4.

4.3.2 Sensor Zenith Angle Effects on Cloud Top Error

The sensor zenith angle is defined as the angle made with the zenith, the line orthogonal to the ground and the satellite platform, and the sideways "looking" direction of the sensor perpendicular to its flight trajectory. Although not specifically cited as a possible error source in the cloud top property ATBD, sensor zenith angle could potentially introduce error into the results. Steeper zenith angles significantly alter the viewing geometry between the sensor and the ground, and increase the path length through which the radiation reaching the sensor must travel. According to the ATBD, the measured radiances are adjusted for zenith angle effects (Menzel et al., 2002).

To find out if this is true, the absolute value of the cloud top pressure differences are plotted against the corresponding zenith angles in Figure 4.13. Figure 4.13a includes all 345 cases in the L_{cld}-M_{cld} sample, while Figure 4.13b separates the sample by top retrieval method. Linear regression lines have been added and the line equations and R^2 values are printed.

It appears, both visually from the random distribution of the points in the plot area and empirically from the small R^2 values, that zenith angle has no bearing whatsoever on the magnitude of the cloud top pressure differences between MODIS and the lidar. The regression lines for the BT_{11} cases and the overall sample both exhibit a slope that is slightly positive and approximately equal to one (1 hPa per 1^o in zenith), but the slope of the CO_2 regression line is approximately -0.8. The lack of correlation obviously negates the significance of these trends. Therefore, it appears that the MODIS algorithm does an adequate job at accounting for the zenith angle effects.

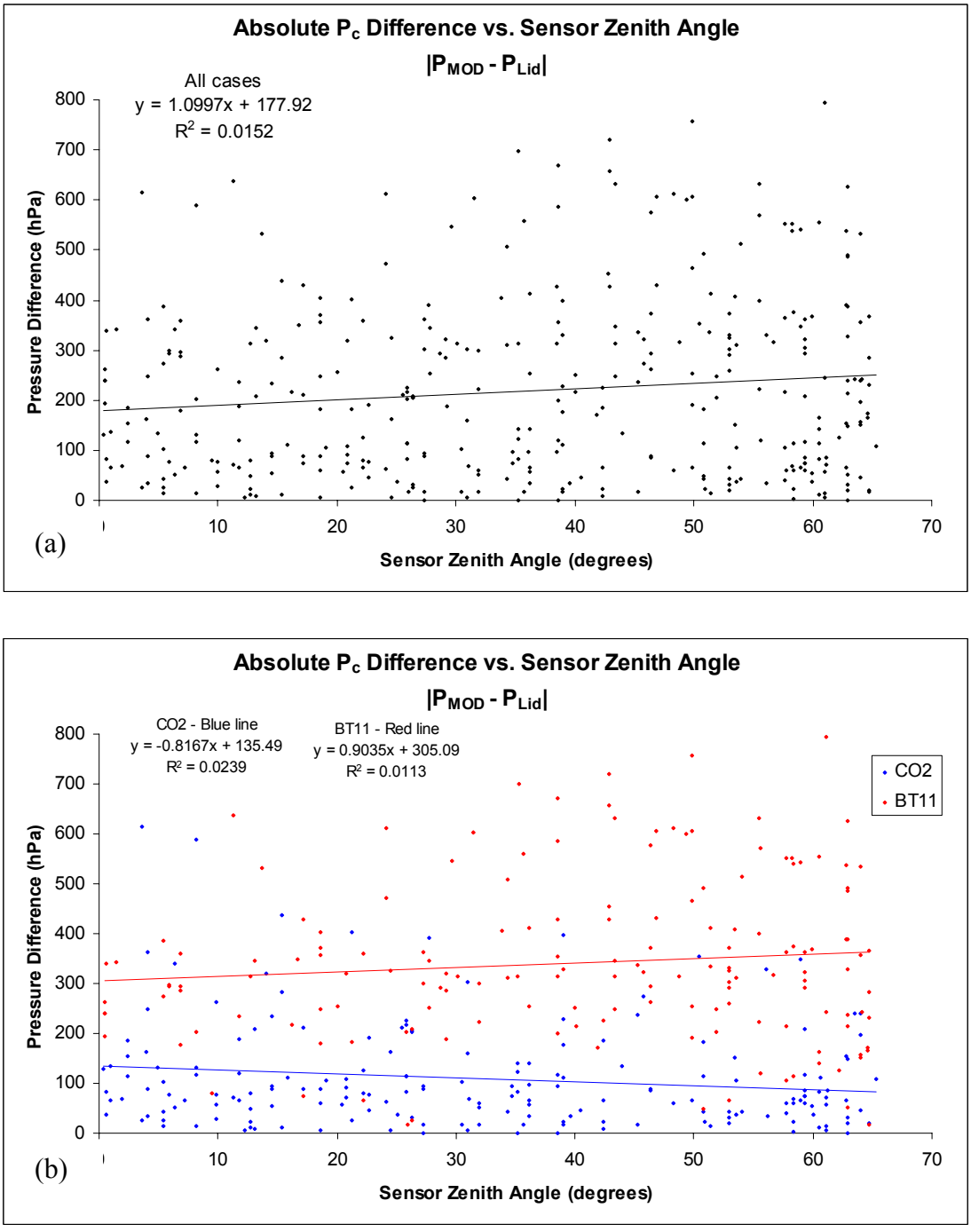


Figure 4.13: Absolute pressure error, |ΔP|, versus sensor zenith angle. Magnitude of the MODIS cloud top pressure error versus sensor zenith angle for (a) all $L_{cld} - M_{cld}$ cases and (b) cases separated by the (blue) CO_2 and (red) BT_{11} top retrieval methods.

4.3.3 Cloud Top Error and Top Layer Geometric Thickness

Recall that with respect to thin cirrus layers, the cloud top estimate originates from the "radiative center-of-mass," which, for optically thin clouds, falls closer to the geometric center. Since the majority of the clouds in this sample possess optical depths smaller than three ($\tau \leq 3$), one might surmise a link between pressure estimate and cloud geometric thickness exists for this biased sample. The correlation between the ΔP, the MODIS minus lidar pressure difference, and cloud geometric thickness could potentially go both ways. If the *geometric* thickness is directly proportional to *optical* thickness, ΔP would be negatively correlated. On the other hand, for optically thin vertically thick clouds, ΔP error may exhibit a positive correlation with thickness since the "radiative-center-of-mass" would conceivably fall further below the cloud top as the thickness increased.

During the analysis, the tops and bases of each cloud and aerosol layer was recorded and the vertical thickness of the top layer was computed. The absolute value of cloud top pressure error is plotted in Figure 4.14 as a function of the geometric thickness of the top cloud layer. Ideally one would prefer to plot the error estimate as a function of cloud *optical* thickness, since, as discussed in sections 2.3.2.2, pressure error and optical thickness are inversely proportional. Unfortunately, the relationship between cloud *optical* thickness and *geometric* thickness is tenuous at best, especially when dealing with cirrus clouds. In one sense, when the number density and particle size distribution of cloud particles is fixed, optical thickness is directly proportional to geometric thickness. However, the number density and particle size distribution can vary widely in space and

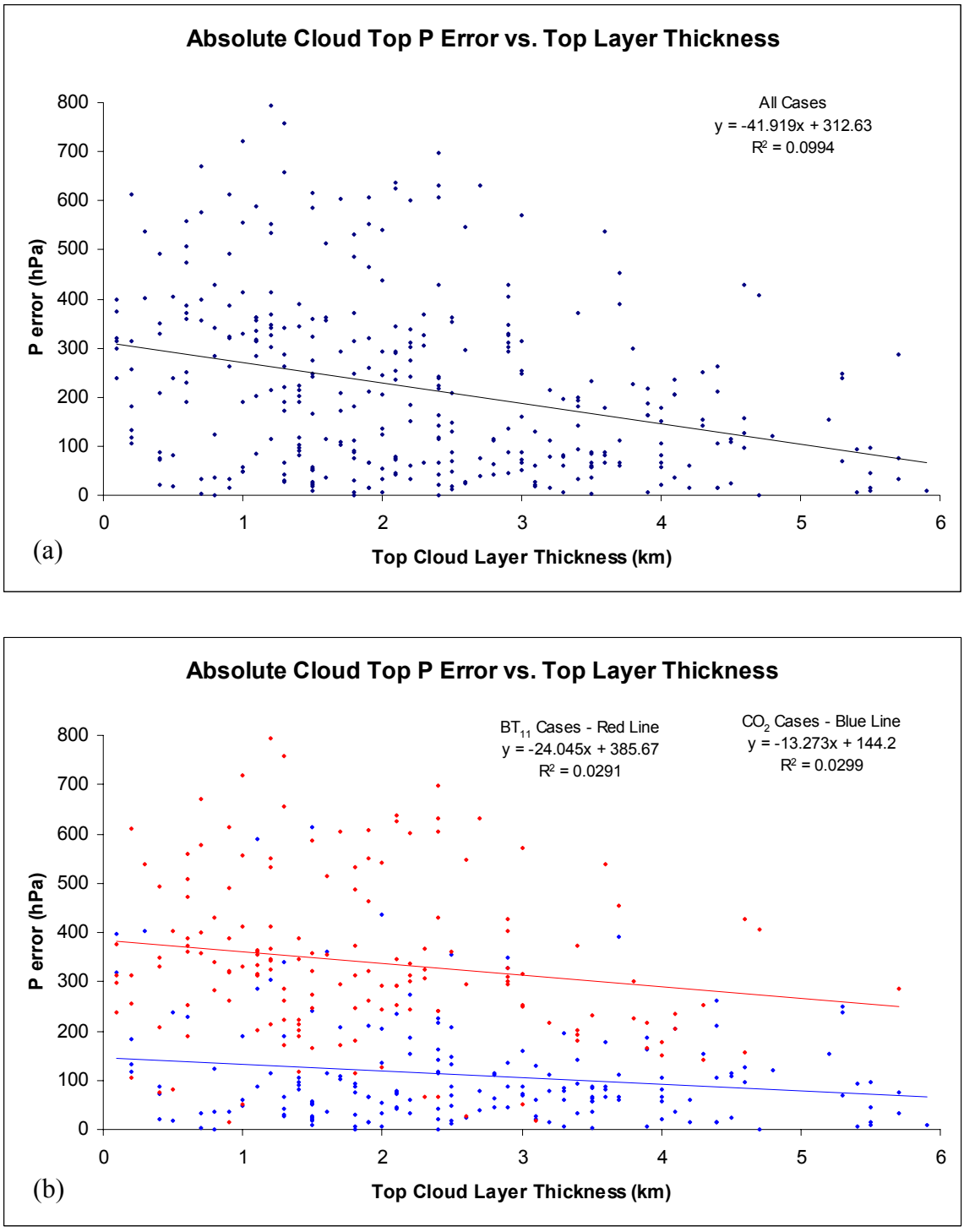


Figure 4.14: Absolute pressure error, |ΔP|, versus top layer geometric thickness. Magnitude of the MODIS cloud top pressure error versus top layer thickness for (a) all $L_{cld} - M_{cld}$ cases and (b) separated by the (blue) CO_2 and (red) BT_{11} retrieval methods.

time and from cloud to cloud. Some clouds are highly transparent but can be several kilometers thick, while others are opaque but may be less than a hundred meters thick.

For this particular sample, as the linear regression lines in Figure 4.14 demonstrate, there is a negative correlation between ΔP error and top geometric thickness, suggesting perhaps a crude correlation between *optical* thickness and *geometric* thickness. But again, the correlation R^2 values are extremely low. A somewhat rudimentary examination of cloud top error versus cloud optical depth is presented at the end of this chapter in section 4.7.

4.4 Cloud Phase Results

4.4.1 Cloud Phase "Sanity Check"

As discussed in section 2.3.3.1, accurate cloud phase identification is essential not only for understanding the how radiation interacts with cloud, but also for calculating cloud microphysical parameters such as optical thickness and particle size (King et al., 2004). The MODIS phase detection algorithm relies on the brightness temperature difference (BTD) and the bands ratio (BR) tests, both of which were discussed at great length in section 2.3.3. Additionally, results of several cloud detection tests implemented by the cloud mask specific to certain cloud types are considered. Also, according to King et al. (2004), a cloud top temperature "sanity check" is applied, whereby the threshold temperatures of 233 K (-40^{o}C) and 273 K (0^{o}C) are used to automatically designate pixels classified as *uncertain* or *mixed*, into either *liquid* or *ice*. The cloud top temperatures are obtained from the MODIS cloud product.

The top temperature "sanity check" for this study is presented in Figure 4.15. Recall, the *uncertain* classified pixels have been added to the *mixed* group as discussed in section 2.4.1. Figure 4.15 illustrates that not all pixels evidently passed the "sanity check." Although all *ice* phase pixels had cloud top temperatures correctly below 0°C and thus "pass" the test, eight *mixed* and two *liquid* cases have top temperatures below -40°C. These findings are inconsistent with the "sanity check," although these ten "failed" cases may not necessarily be incorrect in actuality. For example, a situation in which a thin cold cirrus layer with a measured temperature below -40°C situated above a low-

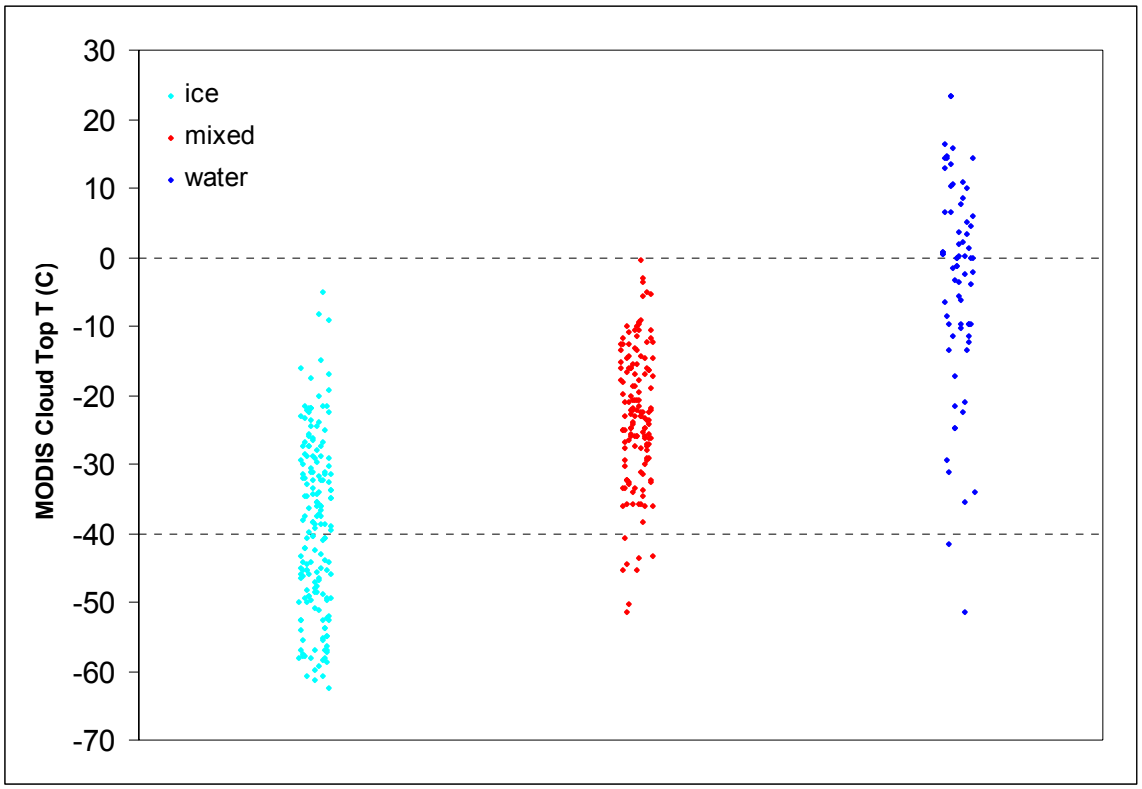


Figure 4.15: MODIS cloud top temperatures sorted by MODIS cloud type. The phase groups are color coded as follows: turquoise – ice; red – mixed; blue – liquid.

level water or mixed phase cloud could end up with a *mixed* or *liquid* classification depending on how much influence each layer has on the measured radiances. However, such a hypothetical does not abide by the rules of the "sanity check" according to King et al. (2004).

4.4.2 Summary of the MODIS Cloud Phase

Figure 4.3 already alluded to the fact that overall, the MODIS cloud types match poorly with the lidar observations. The total number of *ice* cases observed from the lidar data is over twice the number returned by MODIS, while MODIS detects *mixed* and *liquid* phase clouds much more frequently than the lidar observations. Also MODIS vastly overestimates the number of clear cases.

The detailed results of the phase detection are presented below. The analysis uses a methodology identical to the one applied for the cloud mask results in section 4.2. This methodology is examined using an example.

Table 4.2 summarizes the results of the cloud phase for all cases. Note the cloud phase for the July 28, 2006, 2350 UTC Aqua overpass is undetermined and was omitted, leaving the sample with one less, 548, than the actual 549 cases. The four MODIS phase types are listed along the top row with the *mixed* type being a combination of the *mixed* and *uncertain* classes. The left column lists the corresponding lidar phase types. As with the cloud mask analysis, the cloud phase evaluation can follow two approaches. The first assesses the probability that the MODIS phase determination matches the lidar. These

Table 4.2: **Summary of the MODIS cloud phase retrievals.** The MODIS results are listed by column and the lidar observations are listed by row.

All cases							
LIDAR Phase	MOD06 Cloud Phase Clear	Ice	Water	Mi+Unc Mixed	Total	MOD-LID % Agree	
Clear	48	3	11	9	71	*Clear*[1]	67.61
Ice	104	152	37	110	403	*Ice*[1]	37.72
Water	1	1	7	1	10	*Water*[1]	70.00
Mixed	3	20	16	25	64	*Mixed*[1]	39.06
Total	156	176	71	145	548	Overall	42.34
% Agree	*Clear*[2] 30.77	*Ice*[2] 86.36	*Water*[2] 9.86	*Mixed*[2] 17.24			

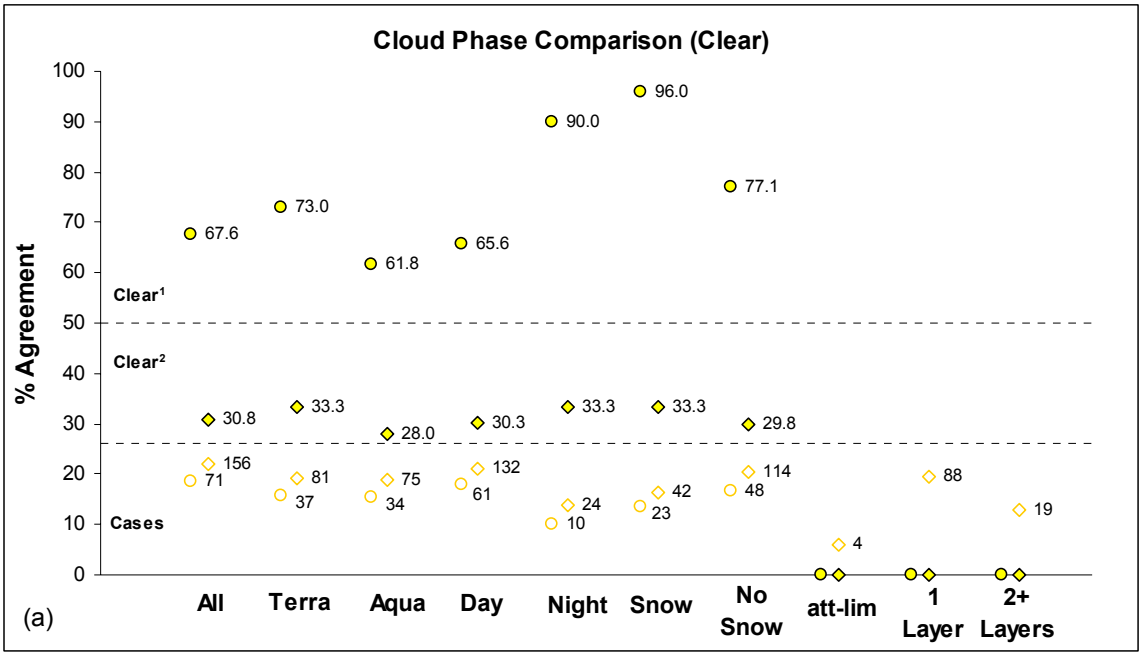
percentages are denoted by the superscript '1' along the far right column. The second assesses the probability that the lidar determination matches MODIS. These are denoted by the superscript '2' along the bottom row. For example, given a MODIS phase determination of *clear*, the probability that this agrees with the lidar ($clear^1$) is equal to [(48/71)*100] or 67.6%. Likewise, given a lidar determination of clear, the probability that this matches MODIS ($clear^2$) is [(48/156)*100] or 30.8%. Similar calculations are conducted for the *ice*, *water* and *mixed* cloud phases. An overall accuracy can be computed by summing the diagonal values (48 + 152 + 7 + 25) and dividing by, 548, the total number of cases; in this case, the overall phase accuracy is equal to 42.3%

Before continuing, it might be helpful to clarify the following issue; the $clear^1$ and $clear^2$ calculations associated with the cloud phase SDS will not necessarily agree with those associated with the cloud mask (Figure 4.5a). The reason for this inconsistency is due, in part, to the differences associated with how the cloud mask and the cloud phase algorithms ultimately determine the cloudiness of a pixel. Recall that the 5-km cloud

mask that comes with the MOD06 product is determined by simply assigning the center pixel value of the 25-pixel array from the 1-km MOD35 product. Theoretically, this pixel could deliver a *confident clear* result even if the remaining twenty-four pixels return *confident cloudy*. Such a finding obviously does not represent the sky conditions within the 25 km^2 area. The cloud phase algorithm, on the other hand, assesses all of the pixel values within the 25-pixel box; if four of the twenty-five return a *confident* or *probably cloudy* result, then the entire area gets designated as cloudy, and a cloud phase determination is performed. As is often the case, a given 5-km pixel from the MOD06 product may paradoxically end up with a *confident clear* cloud mask result, but still possess values cloud microphysical parameters. Likewise, the cloud mask could yield a *confident clouds* label, but possess no associated cloud parameters whatsoever.

The phase comparisons are summarized in Figures 4.16. The figures are separated by cloud type (*clear*, *ice*, *mixed*, *liquid*) and results are separated by the categories listed along the x-axis. Approach[1] percentages are represented by the circles while the diamonds represent approach[2]. The corresponding sample sizes, hollow symbols, stretch along the bottom. The hollow circles can also be thought of as the total number of observed lidar cases and the diamonds as the total number of MODIS cases. Results for the 103 attenuation limited cases are provided above the *att-lim* category, third from the right.

As with cloud mask results, the performance of the MODIS cloud phase algorithm is somewhat mixed and depends on the phase in question and on the approach used. Also, small sample sizes are, again, an issue for some of the groups.



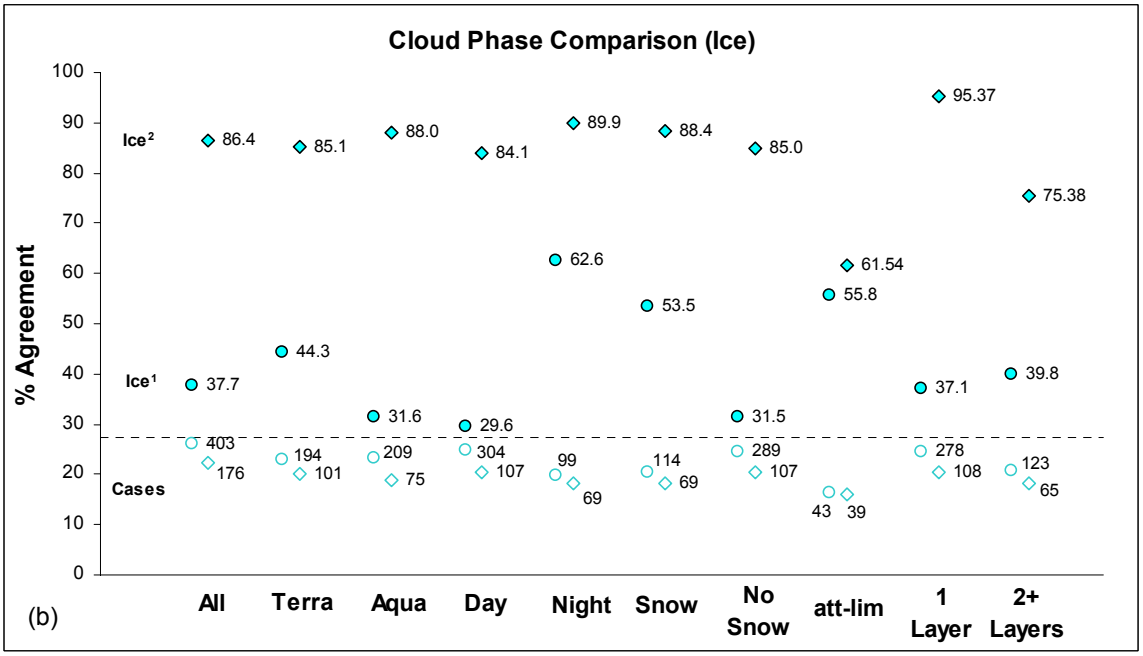
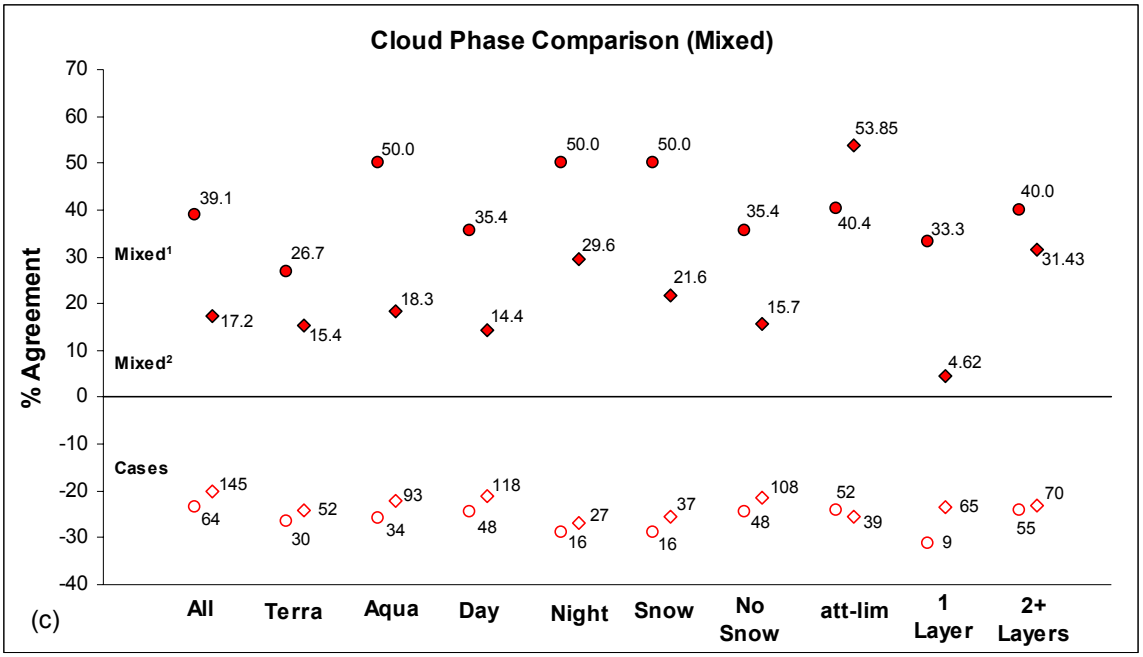


Figure 4.16: **Summary of MODIS cloud phase agreements**. Cloud phase results for the (a) clear sky (yellow) and (b) ice (turquoise) cloud cases. Approach[2] and approach[1] agreements are denoted with the diamonds and circles respectively. Hollow circles and diamonds represent the lidar and MODIS cases, respectively.



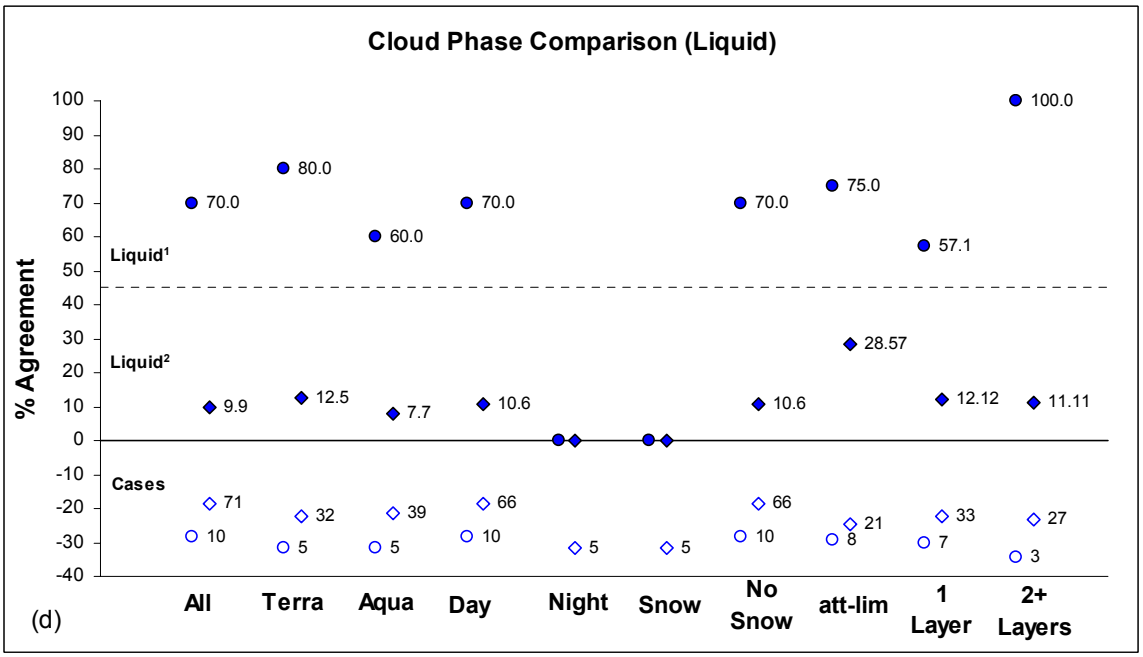


Figure 4.16 cont'd: Summary of cloud phase agreement for (c) the mixed (red) and (d) liquid (blue) cloud types. Categories and symbols are identical to those in Figure 4.16a and Figure 4.16b.

The clear comparisons (Figure 4.16a) closely mirror those presented in Figure 4.5a for the cloud mask. The *clear*^{1} percentages all exceed 60%, with the lowest and highest accuracies of 61.8% and 96% belonging to the *Aqua* and *snow* categories respectively. In addition, as with the cloud mask, *night* outperforms *day* and *snow* defeats *no snow*. Also note that the *att-lim*, *1 layer* and *2+ layer* categories can be ignored for this particular comparison since the number of clear lidar cases (hollow circles) is inherently zero. Similarly the *clear*^{2} values in Figure 4.16a are almost identical to the values in Figure 4.5a, all falling within the 25 – 35% range.

The ice cloud comparisons in Figure 4.16b reveal a wide disparity between *ice*^{1} and *ice*^{2}. Assuming an ice cloud is detected in the lidar beam, the probability that MODIS will also detect ice (*ice*^{1}) ranges from a dismal 29.6% for the 304 recorded *day* cases to 62.6% for the 99 recorded *night* cases, with the median falling around 38.8%. With respect to *ice*^{2}, assuming MODIS detects an ice cloud in the pixel over AFARS, the probability that the lidar agrees is significantly higher, ranging from 61.5% to 95.4% for the *att-lim* and *1 layer* categories respectively, with the median falling around 85.5%. In summary, if the MODIS phase algorithm assigns *ice* to a given pixel, there is a high chance that this assignment is, in fact, correct. However it is also clear that MODIS consistently underestimates the amount of ice cloud in its pixels represented by the low *ice*^{1} probabilities.

Identifying *mixed* clouds, either from the satellite or through visual inspection of the lidar δ-plot, is inherently difficult. The *mixed* cloud scenarios determined from the lidar, as discussed in section 3.6, encompass a wide diversity of cloud cover conditions.

For example, a layer of ice over a layer of liquid (e.g. cirrus over cumulus), a single layer altostratus, and ice virga precipitating from a super-cooled liquid cloud would all be categorized under *mixed*. From the satellite's viewpoint, these three scenarios may appear "spectrally" different. Consequently, one does not expect the agreements of the mixed comparisons to be all that remarkable.

Not surprisingly, the matches in Figure 4.16c are quite poor. *Mixed*^{1} percentages range from 26.7% for *Terra* to 50% shared between *Aqua*, *night* and *snow*. The median value is 39.5%. The variation in the *mixed*^{2} percentages is significant, with a low of 4.6% corresponding to *1 layer* to a high of 53.9% for *att-lim*. With respect to the former, sixty-one of the sixty-five MODIS *mixed* cases under the single cloud layer category are considered to be *ice* by the lidar. For the latter, recall that the *att-lim* cases are much more likely to be of *mixed* or *liquid* phase clouds and are least likely to be ice-only. Thus, it is not surprising that the best *mixed* agreements are associated with *att-lim*.

Finally, the *liquid* category represents the smallest sample of all cloud types. Only ten *liquid* cases in all are documented by the lidar, while MODIS, in contrast, detects liquid clouds seventy-one times. The results are presented in Figure 4.16d. The differences between *liquid*^{1} and *liquid*^{2} are stark. *Liquid*^{1} values range from 57.1% to 100%, with a median score of 70%. The worst and best results are associated with the *1-layer* and the *2+ layers* categories respectively. So given that the lidar detects liquid cloud, the probability that MODIS will agree is decent, however note the sample sizes (only ten lidar *liquid* cases) are much too small to be statistically significant.

In contrast, the *liquid*[2] comparisons are dismal, and demonstrate that the MODIS algorithm significantly overestimates liquid clouds at seven times the rate as the lidar. The range encompasses 0% on the low end to about 28.6% on the high end with a median value of 10.6%. The best value, 28.6%, is, again, associated with the *att-lim* category, the one most likely to contain *liquid* or *mixed* phase clouds.

It is possible that MODIS's consistent overestimation of *liquid* (*water*) clouds is correlated with its excessive use of the BT_{11} method for cloud top property retrievals. Of the seventy-one cases of MODIS derived *water* clouds, approximately 75% of those cases applied BT_{11} for their respective cloud top determinations, which compares with 63% (92/146) for the *mixed* phase clouds and 32% (56/176) for *ice* phase clouds. This implies that perhaps the MODIS algorithm is being tricked into falsely detecting low-level water clouds during cases of clear-sky or ice clouds.

4.5 Aerosol Type

4.5.1 Observed Aerosol Climatology at AFARS

Dr. Sassen chronicles the type and altitude of each aerosol layer during the lidar acquisition. The bottom and top boundaries of each layer (cloud and aerosol) are denoted by the return time of the lidar signal in micro-seconds. These times help to validate the boundary measurements in the lidar plots.

The observations are predominantly that of *smoke* or *Asian dust*, but in addition, four cases of *volcanic* aerosol, two cases of *Arctic haze*, two cases of *pollen* and four of *cottonwood* seed are also documented. However, according to Dr. Sassen, it would be

dubious to recognize the latter two biogenic types as separate categories since it is difficult to ascertain their distribution in the lower atmosphere from AFARS. Such cases are grouped into the category *unknown*, which includes weak aerosol events of unknown origin, such as particulates from biomass, building, power plant and automobile emissions. Local particulate pollution in Fairbanks is common especially in winter during inversion events, but such pollution could also get classified under the *Arctic haze* category and vice versa. Arctic haze is a documented phenomena that has been shown to be caused by factory emissions originating from northern Europe and Siberia (Raatz and Shaw, 1984). In spite of this ambiguity, documented observations of smoke and Asian dust are considered to be very reliable as is the volcanic aerosol event during February of 2006. The other categories (*unknown*, *Arctic Haze*) should be regarded with skepticism.

Surface aerosol observations covering all 549 cases are broken down by month in Figure 4.17. Column heights signify the total number of lidar acquisitions during that month; this number is written explicitly as the denominator of the bold italicized fraction printed above each column. The colored regions within each column signify the aerosol observations, the sum of which is written as the numerator of the fraction above the column. White areas signify aerosol free observations. Case numbers for the most common aerosols, *Asian dust* (yellow) and *smoke* (brown), are printed for convenience, with the dust stamped in italics within the respective yellow region and the latter printed above the shaded areas. For example, of the fifty lidar acquisitions collected during the month of April, forty-two contain aerosol including thirty-two *Asian dusts* and ten

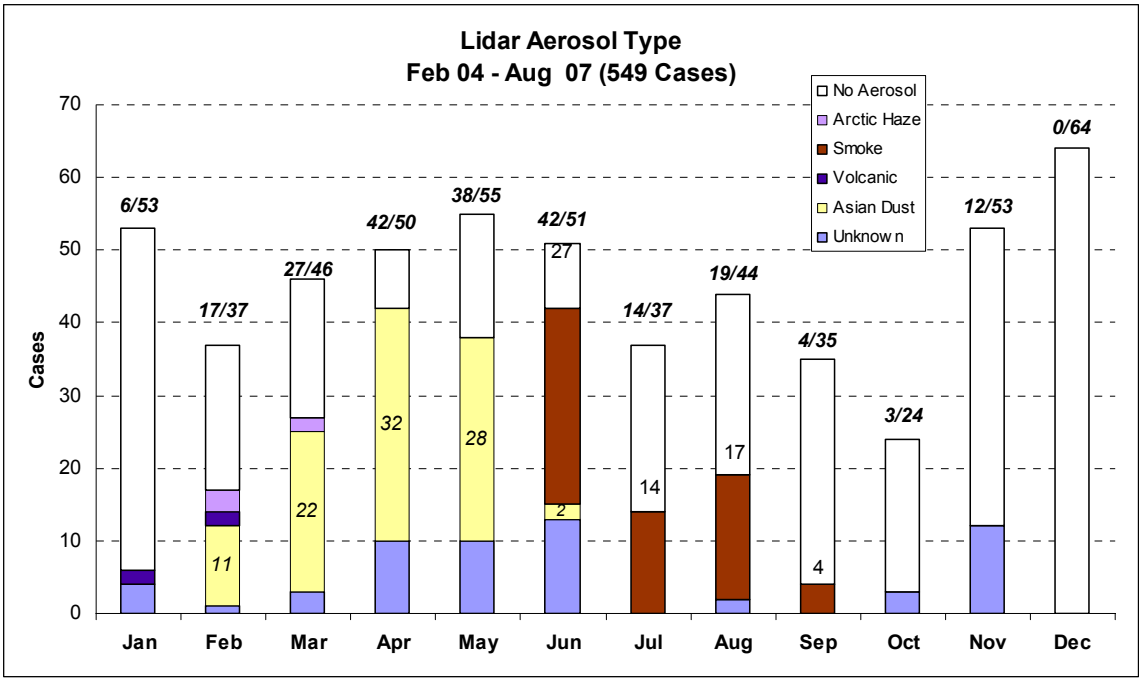


Figure 4.17: Monthly summary of surface aerosol observations. Surface observations of various aerosol types (color-coded) for the February 2004 to August of 2007 period are recorded in the field notes**.**

unknowns. For June, forty-two aerosol cases, comprised of twenty-seven *smoke*, two *Asian dust*, and thirteen *unknown* are observed out of fifty-one cases.

Note Figure 4.17 displays specifically what is recorded in the field notes. With the exception of December, aerosols are documented throughout the year, with the greatest number of cases being observed during the spring and summer months in association with the Asian dust and fire seasons. Four *volcanic* cases (dark purple) attributed to the eruption of Mt. Augustine in early 2006 are also noted, two each in January and February. Several *Arctic Haze* (light purple) cases are documented in February and March, although one suspects that the *unknown* (light blue) cases,

especially during the months of October through March, could also fall under *Arctic Haze*.

The general trends in Figure 4.17 are supported by other research focusing on long range aerosol transport over Alaska (Tiruchirapalli, 2006). A peak is usually observed during the spring and summer months in association with smoke and Asian dust. Also, keep in mind that, as with the cloud analysis, optical thickness is not computed. Consequently the performance of the MODIS aerosol detection algorithm cannot be tested quantitatively since thin translucent layers get grouped into the same hat as the thick ones.

4.5.2 Aerosol Comparison Summary: MODIS vs. Lidar

Table 4.3 lists the MODIS derived aerosol types by column and the lidar derived ones by row. The abbreviation *ND* represents "*N*one *D*etected." Additionally, the

Table 4.3: MODIS and lidar aerosol observation comparison. February 2004 to June 2006 comparison of aerosol observations between the MODIS derived aerosol types (top row) and the lidar observations (left column)**.**

	MODIS	Mixed	dust	sulfate	smoke	Heavy smoke	ND	No-file	Total
Lidar	unknown	1	1	0	11	0	16	6	35
	Asian Dust	7	0	0	13	0	48	3	71
	volcanic	0	0	0	0	0	1	3	4
	smoke	1	0	2	26	0	17	2	48
	Arctic Haze	0	0	0	0	0	2	0	2
	ND	1	0	0	7	3	94	76	181
	Total	10	1	2	57	3	178	90	341

MODIS *no-file* category represents cases in which the MOD04 aerosol product file was unavailable, a situation that commonly arises in winter if all of the pixels in the overpass receive a *night* flag. Recall that a *night* label gets attached if the solar zenith angle ( ) is greater than 85^{o}. Such conditions are common in Fairbanks during the afternoon wintertime MODIS overpasses. A *night* label automatically renders the aerosol algorithm inoperative, even if the sun happens to be above the horizon during the overpass. Note that *no-file* does not account for cases where files exist but where pixels may have still received a *night* flag; such cases would have been grouped under the none detected (*ND*) category. The *ND* group implies that the pixel was flagged as day, the aerosol algorithm is applied, and no aerosols are detected. At this time it is unknown how many of the *night* pixels fall under the *ND* heading.

It is evident from Table 4.3 that the aerosol type identifications between the two instruments do not strictly match. The MODIS product contains a *mixed* category, but it is unclear which aerosols *mixed* represents, as the five MODIS aerosol types along the top of Table 4.3 do not explicitly match with the aerosol models discussed in Kaufman and Tanré (1998) or Remer et al. (2006) (see Table 2.3).

The lidar and MODIS instruments simultaneously observe aerosol of some form or another over AFARS sixty-two times. This compares with 160, the total number of aerosol observations made by the lidar and seventy-three, the total number of MODIS aerosol type returns. Recall from the earlier discussion and from Figure 4.2 that surface observations of aerosols get documented regardless of thickness or cloud presence. Of the 549 cases represented in Figure 4.2, only twenty-two fall purely under *aerosol only*

defined as “one or more thick aerosol layers; no cloud layers”. This number is likely to be slightly fewer for this particular C004 sub-sample (341 cases). Nonetheless, one can expect that cloud contamination from above will impede aerosol detection from the satellite. For example, the July 24, 2006 lidar acquisition presented in Figure 3.1 contains two distinct layers of smoke, one situated near 4.5 km and the other, which roughly extends from the surface to about 2 km in altitude. From MODIS’s point-of-view, these layers are obstructed by a cirrus cloud. According to the C005 aerosol data (since C004 does not extend past June of 2006), the 2045 and 2220 UTC Terra overpasses as well as the 2235 UTC Aqua overpass do not detect aerosol (the 2100 Aqua UTC erroneously returns a *sulfate* identification) in the field-of-view, presumably, in part, because the cloud is blocking the view. Therefore, the strict standards applied to the lidar aerosol analysis will inevitably cause disagreements with the MODIS data, since lidar aerosol observations from the ground are usually less impeded by cloud cover.

The only match in Table 4.3 worth noting is associated with the smoke categories. Similar to the cloud mask and cloud phase analysis, the evaluation can be divided into the $smoke^1$ and $smoke^2$ approaches. The former is equal to the number of smoke matches divided by the total number of smoke cases identified by the lidar: [(26/48)]*100 = 54.2%. The latter divides by the number of MODIS cases: [(26/57+3)]*100 = 43.3%. Note the total number of MODIS smoke cases include fifty-seven *smoke* as well as three *heavy_smoke* cases. Also, peculiarly enough, it appears that the MODIS sensor detects a *dust* case even though, according to the Kaufman and Tanré (1996) “dust will be sensed

only over the ocean." Dust detection over land is only provided with the collection 005 version and higher, which is not used in this analysis.

The monthly distributions of aerosol cases corresponding to the February 2004 to June 2006 C004 MOD04 dataset are presented below in Figure 4.18 with the MODIS types given on top and the lidar types underneath. Figure 4.18 is analogous to Figure 4.17 with the column heights representing the total number of lidar acquisitions and the colored regions representing the fraction of aerosol cases by type, also denoted by the fraction printed above each column. Additionally, the numbers corresponding to the MODIS smoke cases and lidar smoke and Asian dust cases are printed for convenience.

To examine the performance of the MODIS aerosol identifications in Figure 4.18a, recall the limitations of the C004 aerosol data. First, the algorithm does not operate at night, nor does it operate over snow and ice covered surfaces since it relies on the *dark target* approach. This limits the aerosol detection to daytime overpasses from about the middle of April to the middle of October, and why there are no aerosol detected from October to February. Although in Fairbanks, snow usually remains on the ground through the end of April, it is mostly hidden under the dark spruce forest canopy, and any large piles still left in the urban areas have a low albedo due to the accumulation of soot and gravel. Therefore, old end-of-season snow-pack probably will not significantly influence the application of the "dark target" approach.

With the exception of one *smoke* case in March, the MODIS aerosol type cases span the months of April through September. The distribution in Figure 4.18a roughly

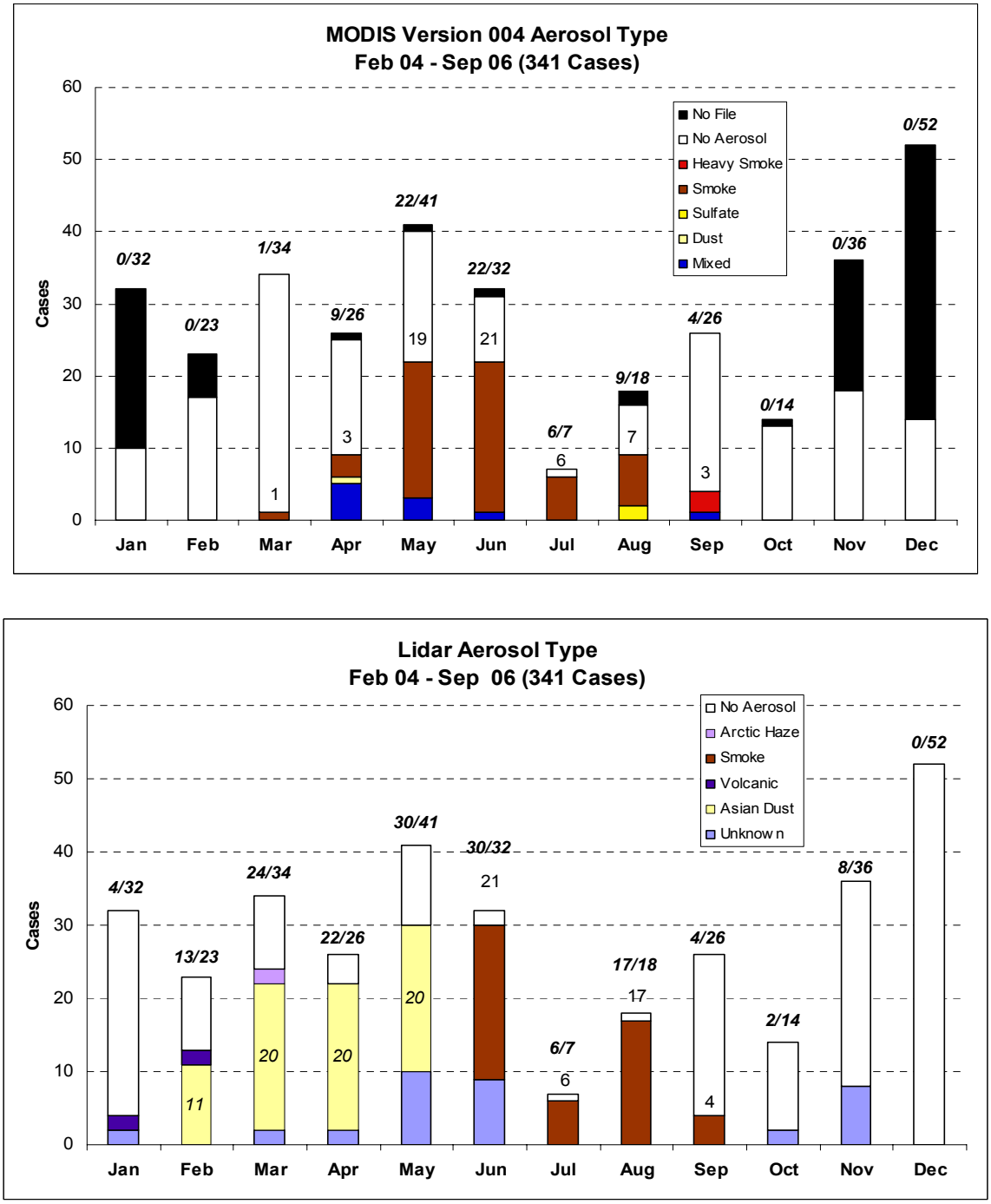


Figure 4.18: Monthly distributions of MODIS and lidar aerosol observations. Distribution of aerosols by month for the (a) MODIS derived and the (b) lidar and surface derived observations for the 341 MODIS collection 004 cases beginning in February of 2004 and ending in June of 2006.

mimics the one in Figure 4.18b, with the aerosol season peaking during the summer in association with fires season. However *Asian dust*, a significant source of aerosol from late winter through spring, does not manifest itself in the MODIS data. Many of the Asian dust events in April and May get misclassified as either *smoke* or *mixed* (Table 4.3), while many more in February and March go undetected, *ND*, or are misidentified as an ice cloud, which will be discussed later. In general, the frequency of *smoke* cases between MODIS and the lidar are pretty closely matched from June to September, though at the case-by-case level in Table 4.3, the match is somewhat more tenuous.

4.5.3 MODIS *Heavy_Smoke*: Three Case Studies

MODIS erroneously detects three *heavy_smoke* cases in September, one on September 7, 2004 at 2305 UTC, the second on September 2, 2005 at 2215 UTC and the third on September 11, 2005 at 2210. All three are detected from the Aqua platform. From the lidar plots (Figure 4.19), these appear to be elevated cirrus layers with δ-values exceeding 0.15. Note the 2210 UTC Aqua overpass on 9/11/2005 (Figure 4.19c) seems to have barely grazed what appears to be a cirrus cloud located near 11 km. The Aqua overpass occurs a little past 2212 UTC although the cloud does not appear until about 2215 UTC. According to both the field notes and the CBCT outputs, the overpass did clip the edge of some cirrus. Furthermore the cloud seems to be vertically sliced off in the lidar data lending evidence that the cloud was present during the Aqua overpass; ultimately for this case, the cloud top is recorded at 11.7 km, however it is possible that

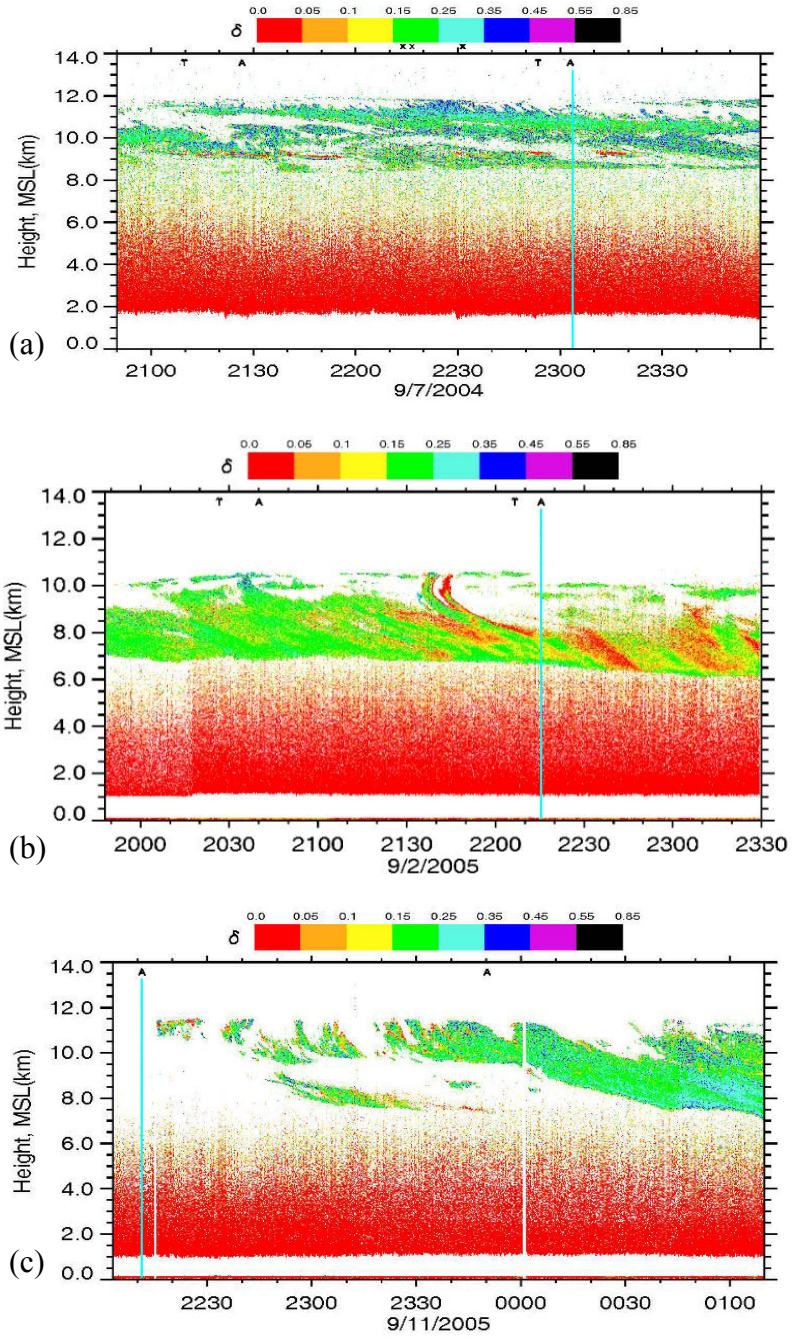
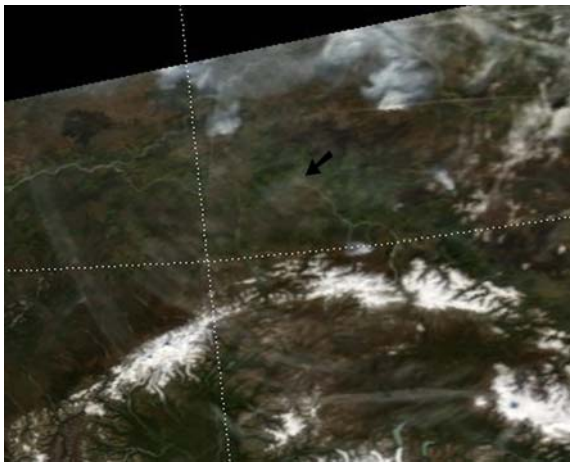
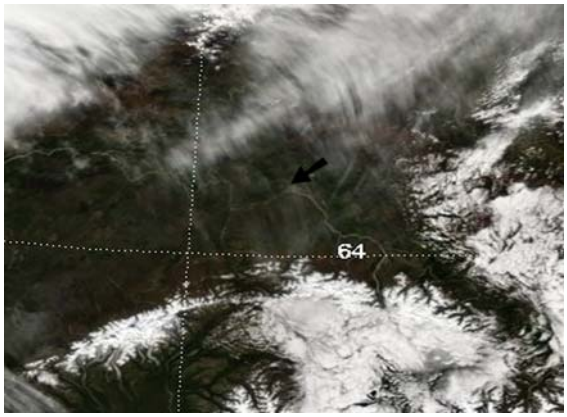


Figure 4.19: Lidar δ-plots corresponding to three MODIS *heavy_smoke* cases. Plots represent the (a) September 7, 2004, (b) September 2, 2005 and (c) September 11, 2005 lidar acquisitions. The turquoise lines denote approximate Aqua overpass times corresponding to the *heavy_smoke* cases in Figure 4.18a.

(a)



(b)



(c)

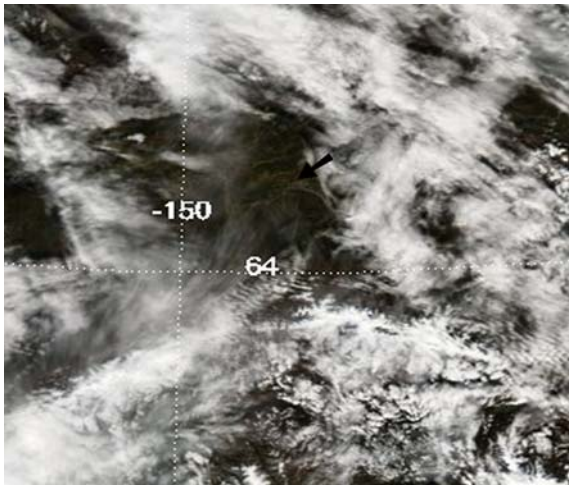


Figure 4.20: Aqua MODIS true color composite images for three *heavy_smoke* cases. Images for the (a) September 7, 2004 2305 UTC, the (b) September 2, 2005 2215 UTC and the (c) September 11, 2005 2210 UTC overpasses. The black arrows (center) denote Fairbanks. Images courtesy of http://modis-atmos.gsfc.nasa.gov.

MODIS could have missed the cloud altogether, depending on the viewing geometry.

The true-color MODIS images corresponding to the three Aqua overpasses are presented in Figure 4.20. The two 2005 cases do, in fact, appear to be that of cirrus cloud, and thus cases of mistaken identity (Figure 4.20b, 4.20c). However, the 2004 image in Figure 4.20a captures a large area of smoke burning several hundred miles north of Fairbanks along the Yukon River. It is possible that this plume, observed in the lidar and the field notes as cirrus, contains smoke transported by the westerly to northwesterly trajectories of the clouds and the smoke plume apparent in the image (4.20a). The 5-km MODIS cloud mask for these three cases all return *clear*, yielding three of the 107 documented false negative (*F-N*) cases in Figure 4.1. These F-N cases get scrutinized more closely in the next section.

4.6 False Negative (*F-N*) Analysis

False-negative (*F-N*) cases are instances in which the lidar unambiguously detects cloud situated over AFARS while MODIS simultaneously detects cloud-free conditions. One hundred and seven of the 549 cases, or 19.5% of the total sample, receive an *F-N* label, denoted earlier in Figure 4.1. The underlying causes resulting in such a high percentage of *F-N* identifications are discussed in sections 4.1.3 and 4.1.4. The primary explanation put forth is that the sample is overwhelmingly comprised of thin single layer cirrus clouds (Figure 4.2), the most difficult type to detect with a passive satellite instrument. This statement is circuitously supported in Figure 4.3, which shows that

MODIS detects fewer than one half the number of ice clouds observed by the lidar, while it detects double the number of clear-sky conditions compared with the lidar.

Additionally, when one ranks the top five number of *F-N* cases by month in Figure 4.4b, four of the five months (June, May, September, August) happen to reside within the warm, snow-free part of the year. It has been speculated that thin cirrus projected against a warm dark background are more likely to get missed by the satellite, than cirrus against a cold bright ground. Although a cloud would be better camouflaged against a cold reflective background, such a surface tends to lead to a higher frequency of false positive (F-P) detections and a higher rate of the BT_{11} retrievals during the cold months (Figure 4.4b). Therefore, it is likely that the seasonal difference in the *F-N* observations is more an artifact of MODIS coincidentally capturing a thin cirrus cloud when it believes that a low-level thick water cloud is present.

Figure 4.21 compares how the *F-N* cases are distributed among the various categories compared to the overall sample. The four categories are broken down by platform (*Aqua*, *Terra*), snow cover (*snow*, *no snow*), time of day (*day*, *night*) and season (*cold**, *warm**). The *cold** season is defined from October 1 to March 31 and the *warm** season from April 1 to September 31. The most striking difference between the *F-N* cases and the sample overall, is the distribution of cases by season. While overall the cases are pretty evenly distributed between the warm and cold months, the chances of an *F-N* occurring during a *warm** month is approximately three times greater than during a *cold** month, for reasons just discussed in the previous paragraph.

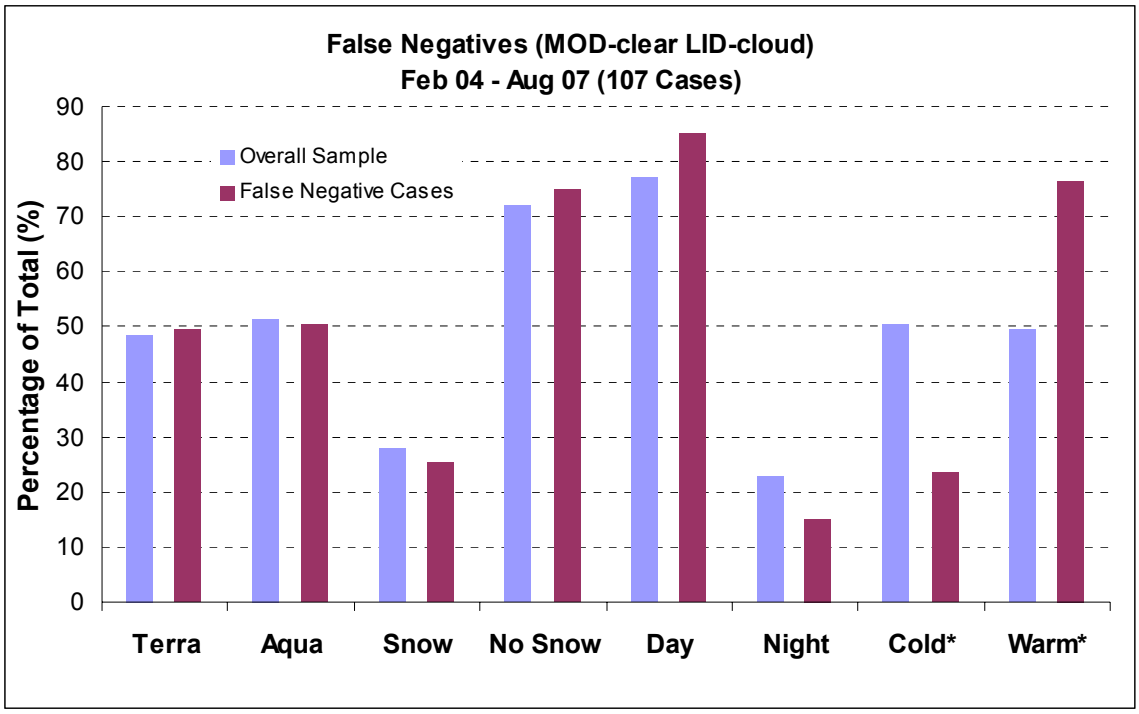


Figure 4.21: Distribution of false negative cases sorted by various categories. A plot depicting how the 107 *F-N* cases are distributed by category compared to the distribution of categories overall. The columns represent 'percent of sample' for the (maroon) F-N cases only and for the (blue-gray) sample overall. *Cold** spans the six months from October 1 to March 31 while *warm** encompasses April 1 to September 31.

Although the disparity is not quite so large, differences between the *day* and *night* categories are also significant. Whether or not this is due to the fact that daylight is an influencing factor or it is simply an artifact of having a greater amount of overlap between the *day* cases in the *warm** categories, is uncertain. One category implies that temperature is the influencing factor while the other implies that illumination conditions are most important. It is likely that both play some role based on the above discussion.

Snow cover also appears to have a small but meaningful influence on the *F-N* frequency. Again, *F-N's* are more likely to fall under the *no snow* category but, with the

exception of the six MODIS flagged *snow* cases in April (Figure 4.4a), all of the *warm** category cases are *no snow*.

It is apparent that MODIS frequently fails to detect clouds, but could this observation be a case of mistaken identity? Figure 4.22 depicts sixty-two of the 107 *F-N* cases sorted by lidar cloud phase for the shortened sample coinciding with the collection 004 MODIS aerosol data (February 2004 – June 2006). First, note that the overwhelming majority of *F-N*'s, fifty-nine of the sixty-two, fall under the MODIS *ice* category. Also, not one of the F-N cases fall under the ten lidar derived liquid cases. Within the lidar derived *ice* column, thirty-four of the forty-seven *F-N* cases or 72% (not counting the *No-*

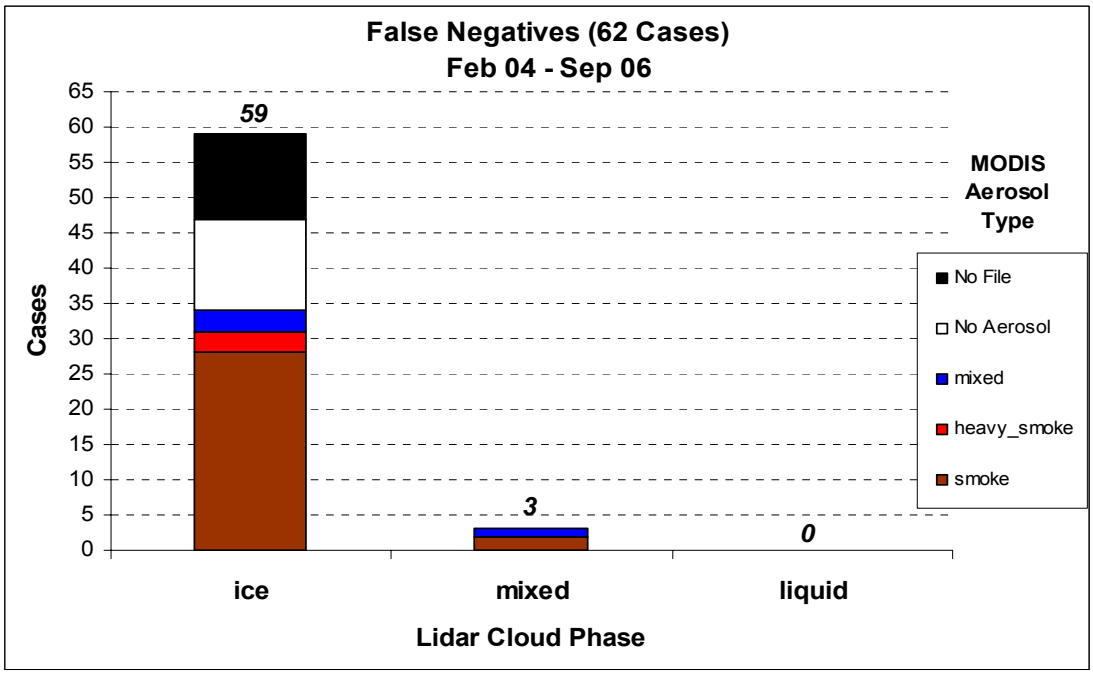


Figure 4.22: Distribution of *F-N* cases sorted by lidar phase and MODIS aerosol type. The sample encompasses the C004 MODIS aerosol product (February 2004 – June 2006). Numbers printed above the columns denote the total number of *F-N* cases corresponding to a given lidar cloud phase.

file cases) are determined by MODIS to have some form of aerosol be it *smoke*, *heavy_smoke* or *mixed*. The three *heavy_smoke* case studies are closely examined in section 4.5.3. Similarly, all three lidar *mixed* cloud categories coincide with two MODIS *smoke* and one *heavy_smoke*

In summary, it seems that MODIS regularly misidentifies *ice* and occasionally *mixed* clouds for aerosols. In one sense, this finding is encouraging since it shows that, for the majority of the *F-N* cases, the satellite is at least detecting something in its field-of-view, although it may have a hard time distinguishing aerosols from clouds, especially ice clouds. Perhaps better spectral fingerprinting tests can be developed to alleviate this problem.

The F-N cases for the entire dataset (549 cases) are presented in Figure 4.23. After re-examining all 107 *F-N* cases, the cases get binned under four hypothetical categroies possibly to blame for MODIS's cloud detection failures, based on the lidar observations. The two most obvious explanations are from *thin cirrus*, or from the overpass occurring at a moment when a *cloud edge* is in the vicinity of AFARS. These cases (62 + 12), when combined, represent the largest percentage of explanations falling within the realm of "reasonability." Some cases may even fall under both categories. For example, recall the controversy surrounding the 9/11/2005 2210 UTC Aqua overpass, a case in which it appears MODIS barely intercepts the edge of a thin cirrus cloud (Figure 4.19c).

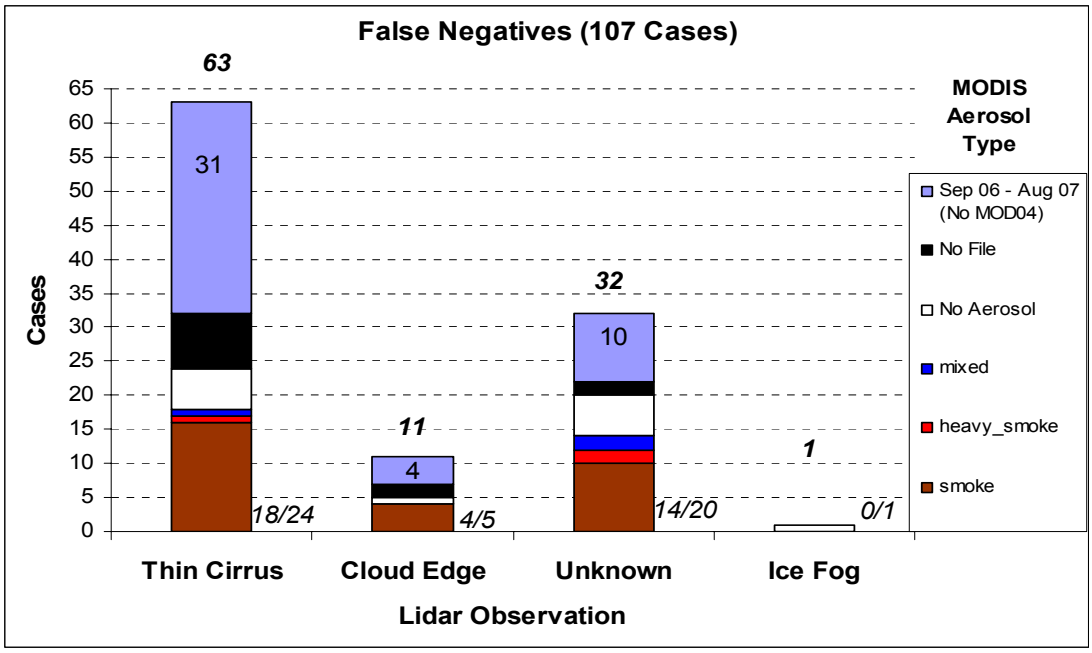


Figure 4.23: Breakdown of *F-N* cases based on hypothetical deductions from the lidar observations. Columns are divided up by MODIS aerosol type. Bold italicized numbers at the top signify total number of cases in the respective category. Regular numbers signify number of cases for which MODIS C004 data was unavailable. Italicized fractions along the bottom denote fraction of cases for which C004 data was available in which a MODIS aerosol type was returned.

The *unknown* group represents cases which occur neither near a cloud edge nor, as best as one can tell, over a thin cirrus layer. These are cases of relatively "thick" cloud layers that, for the most part, should have been detected by MODIS. Hence, the reasons for missing the cloud in the FOV are *unknown*. However, as Figure 4.23 illustrates, a very large percentage of the dataset for which MOD04 aerosol data is available (70%), yield an aerosol type in this category. This, again, supports the argument conveyed earlier, that MODIS appears to, at times, have a hard time distinguishing clouds from aerosols.

Lastly, note that one of the *F-N* cases occurs during one of the ten ice fog events. Although based on the lidar returns and the field notes, this 12/24/2004 case appears to be a relatively thick, deep layer of surface ice fog, expecting MODIS to detect shallow boundary layer ice fog may not be reasonable even though the chances that the BT_{11} method gets applied under these cold low-level inversion scenarios seems rather high (Figure 4.4b)

Analagous to Figure 4.22, Figure 4.23 provides the breakdown of *F-N* cases coinciding with the MODIS C004 aerosol data (February 2004 – June 2006). These cases include the colored stacks up through black. The remaining September 2006 to August 2007 cases are represented by the top blue-grey regions, of which there are thirty-one, four and ten cases for the *thin cirrus*, *cloud edge* and *unknown* categories respectively. The fractions written to the right of each column along the bottom, in italics, signify the number of cases in which an aerosol is observed out of the total number of cases represented in the C004 subcategory. For example, the fraction *18/24* associated with the *thin cirrus* group means that of the twenty-four cases in which a MOD04 aerosol file is available (omitting the *No-file* cases in black), MODIS detects aerosol eighteen times; the type is denoted by the color in the legend.

Assuming that all of the *F-N* cases within the *thin cirrus* or *cloud edge* categories are correctly accounted for, and that the aerosol detections documented within *unknown* are, in fact, the underlying causes of the *F-N* identifications, then very few unaccounted for *unknown* cases remain. There can be no more than eighteen (6-white, 2-black, 10-bluegrey) and no fewer than six (6-white) *unknown* cases, which completely lack a

reasonable explanation for why MODIS failed to see cloud. However, this analysis admittedly bumps into the major drawback of subjectivity. Since there are no quantitative definitions of "thin" and "thick," and since cloud optical thickness is not explicitly measured, it is impossible to know for certain if the cases belonging to the *unknown* category, or even within the *thin cirrus* or *cloud edge* categories, possessed optical thicknesses that were appropriately above or below MODIS's own cloud detection limits.

4.7 False Positive (*F-P*) Analysis

4.7.1 Summary of *F-P* Cases

False positives, or *F-P*s, are defined for cases in which the MODIS pixel over AFARS detects a cloud while simultaneously the lidar indicates cloud-free conditions. Twenty-four *F-P* cases are observed, significantly fewer than the 107 associated with the *F-N*s. Possible causes include mistaking an aerosol layer for a cloud, a cloud edge in the vicinity of AFARS and the presence of cold surface inversions. A discussion regarding the influence of inversions on cloud detection can be found at Liu et al., (2004).

All twenty-four *F-P* cases are closely examined. The results are organized below in Figure 4.24. Similar to Figure 4.23, possible explanations for the erroneous identifications are presented along the x-axis. Again, the terms "heavy" and "thin" with regards to aerosol optical thickness are subjective. Also, a *strong inversion* is defined as a temperature increase exceeding 5^{o}C over the first 500 m of the atmosphere ($T \geq 5^{o}$C per 500 m), and a *weak inversion* is defined as a temperature increase below 3^{o}C over the

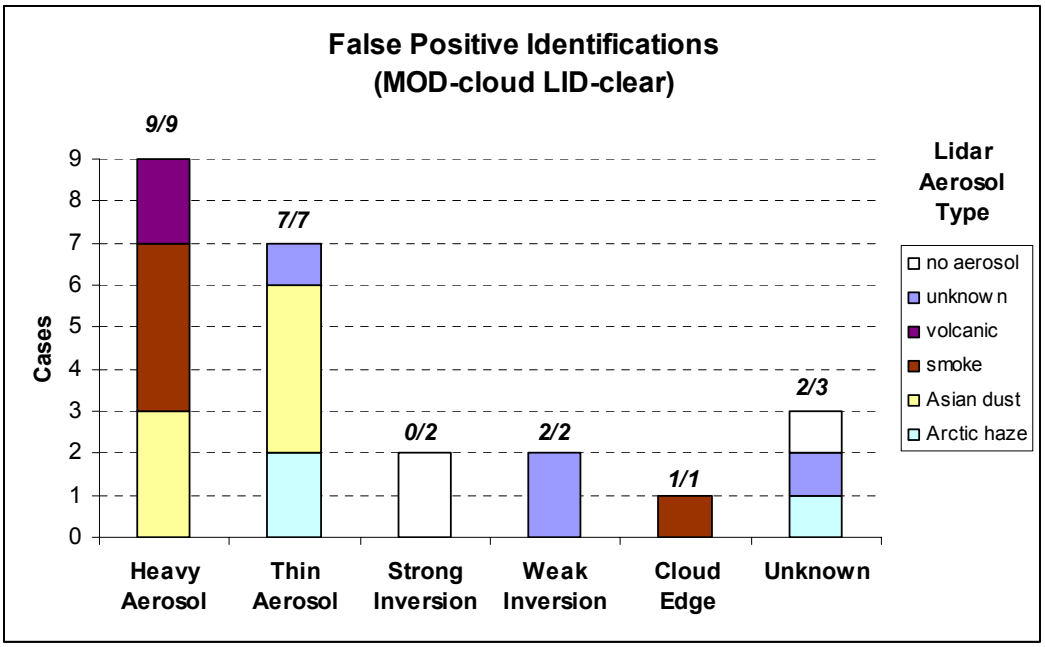


Figure 4.24: False positive cases sorted by hypothetical underlying deductions from the lidar and field observations. The columns are partitioned by lidar aerosol observations; the types are denoted by the colors in the legend.

first 500 m ($0^{o}C \leq T \leq 3^{o}C$ per 500 m). The *strong inversion* cases correspond to the 1/17/2005 2230 UTC and the 1/18/2006 0600 UTC Terra overpasses, and the *weak inversions* with the 2/20/2007 2130 UTC Aqua and the 3/01/2007 2245 UTC Terra overpasses.

Unknown is comprised of three cases which do not fit into the other five categories. Two of the three cases contain aerosol layers, which are deemed much too thin and amorphous to be grouped under *thin aerosol*. As in Figure 4.23, the fractions above the columns represent the fraction of aerosol observations for that category.

Numerous *F-P* cases coincide with lidar derived aerosol layers, again raising the possibility of mistaken identity. Nine of the twenty-four cases are deemed *heavy aerosol*,

of which, four are identified as *smoke* and two as *volcanic*. These cases will be elaborated on in the next two sections as case studies.

Both of the *strong inversion* cases are aerosol-free, which eliminates aerosol contamination a possible cause, leaving the inversion as the main culprit for the MODIS *F-P* identification. On the other hand, tenuously thin aerosol layers are documented for both *weak inversion* cases.

Figure 4.25 further sorts the *F-P* cases by MODIS derived cloud phase. Note that, unlike the *F-N* cases in Figure 4.22, the majority of the *F-P* cases correspond to the *liquid* (MODIS – water) and *mixed* categories. Only three are determined to be *ice* cases. All five *smoke* cases reside within the *liquid* category, which seems sensible considering that smoke releases vast amounts of water vapor into the air, and its low depolarization characteristics are similar to that of water droplets (Sassen et al., 2006). Also the single *cloud edge* case in Figure 4.24 happens to intercept a thin smoke layer, and it occurs near the edge of low-level attenuating cumulus clouds, hence the liquid classification in Figure 4.25.

Analagous to Figure 4.21, Figure 4.26 compares the distribution of *F-P* cases by category (e.g. platform, season, etc.) with the overall dataset. The most notable disparity between the F-P cases and the overall dataset is differences associated with the top retrieval methods. BT_{11} is selected twenty-three of the twenty-four *F-P* cases (96%), which compares to roughly 51% of the time overall. Also, F-P cases are also more likely occur for the *day* and *snow* flagged pixels during the *cold** season.

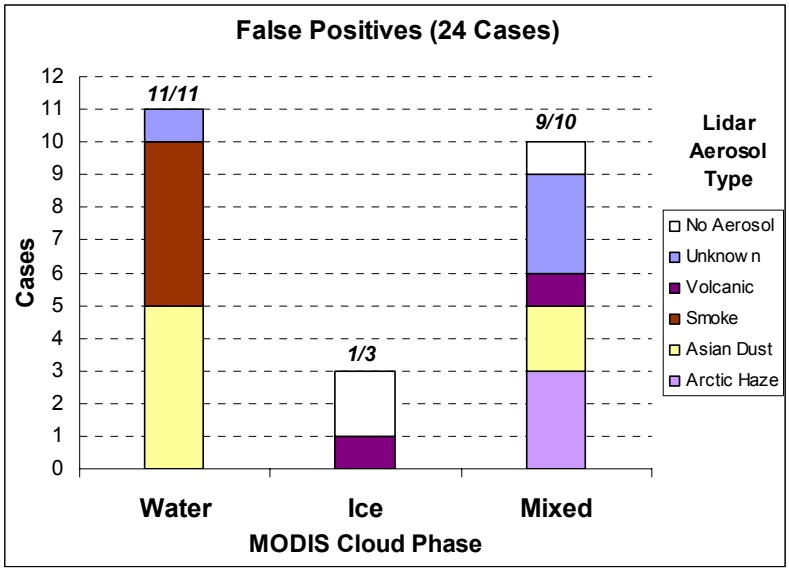


Figure 4.25: Distribution of *F-P* cases sorted by MODIS cloud types. The columns incorporate concurrent lidar derived aerosol observations separated by type.

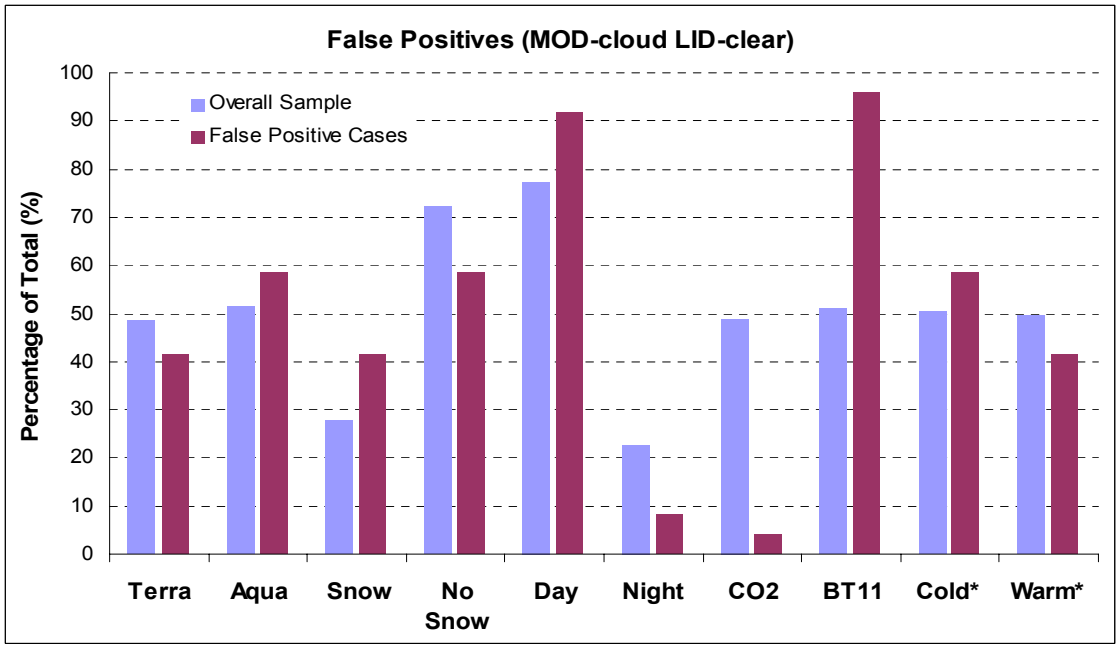


Figure 4.26: Distribution of false positive cases sorted by various categories. *F-P* cases sorted by platform, snow cover, time of day, top retrieval method, and season compared to the overall sample. *Cold** denotes October 1 to March 31; *warm** denotes April 1 to September 30.

4.7.2 A June 27, 2005 Case Study: Cloud as Smoke

Recall, four of the nine *heavy aerosol* cases in Figure 4.24 coincide with *smoke* aerosol based on the lidar and surface observations. This section will examine two of these cases: the 2135 and 2145 UTC Terra and Aqua overpasses for June 27, 2005.

The lidar plots of backscattered power and depolarization ratio corresponding to the aforementioned cases are given in Figures 4.27. Both indicate the presence of several contiguous layers, extending from the surface to roughly 5 km in altitude. The top layer becomes diffuse and eventually disappears a little after 2130 UTC. The layers exhibit very low depolarization values (Figure 4.27b) consistent with smoke or a water cloud, but in the case of a water cloud, the lidar tends not to penetrate much beyond a few hundred meters above cloud base. Additionally, the depolarization of liquid droplets exhibits a rapid increase above base due to multiple scattering effects. In this scenario, neither clue is observed. Therefore one can conclude that this layer is undoubtedly that of smoke.

Figures 4.29 and 4.30 are maps of cloud top pressure and aerosol optical depth coinciding with the Terra image in Figure 4.28. The cloud top pressures and the aerosol optical depths are given at 5 km and 10 km resolutions respectively. Note the narrow smoke plume visible in Figure 4.28 appears to coincide with the red tongue of apparent low clouds in Figure 4.29 and with the black region in 4.30. Note also, the approximate location of AFARS is marked by the small yellow star within these regions. If Figure 4.29 and Figure 4.30 are stacked on top of one another, the clear dark regions of the latter

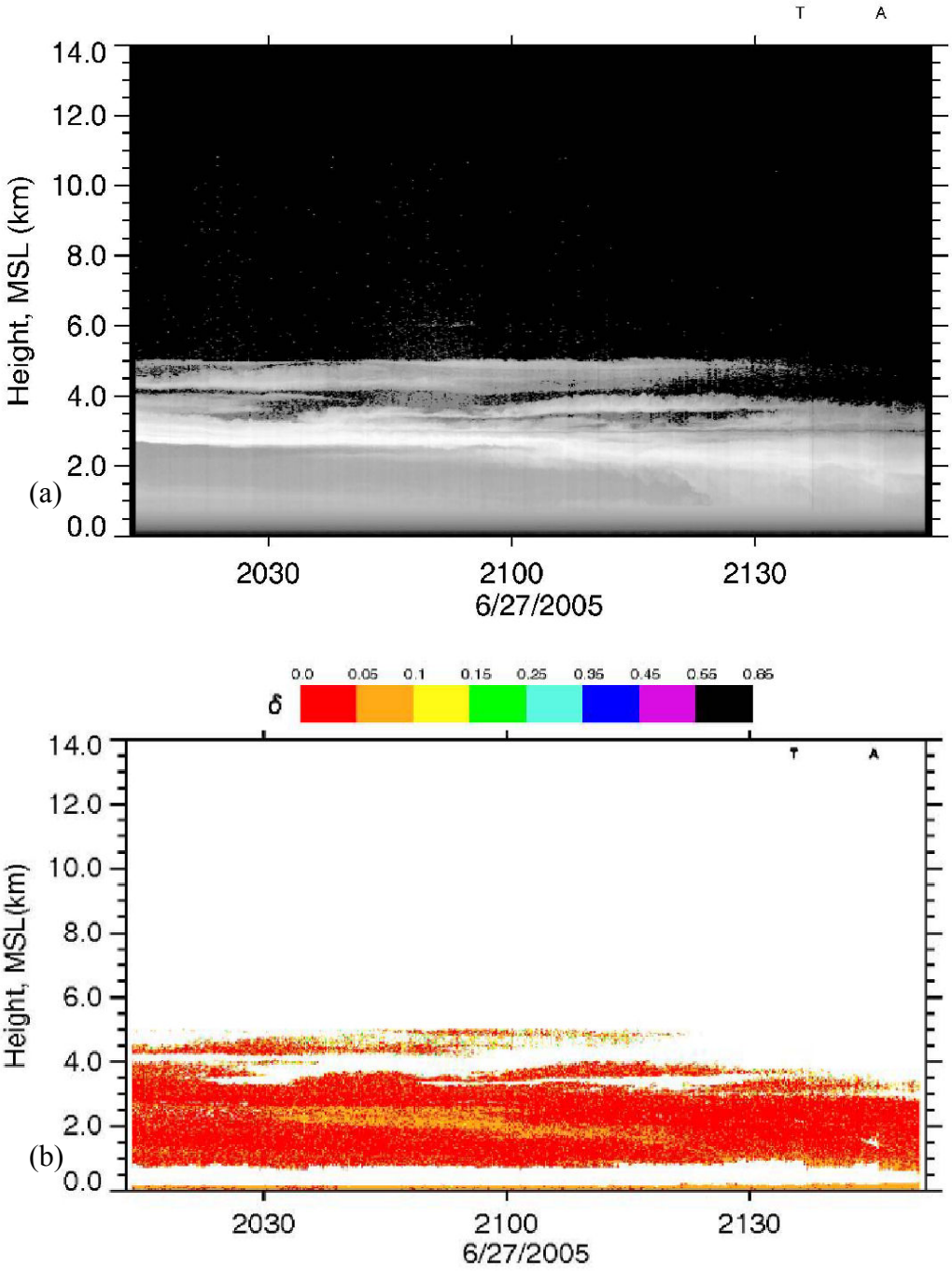


Figure 4.27: Lidar plots of an *F-P* smoke case study. The plots give (a) returned power and (b) depolarization ratio corresponding to the June 27, 2005 acquisition. The T and A along the top of the plot windows denote approximate overpass times for the Terra and Aqua platforms.

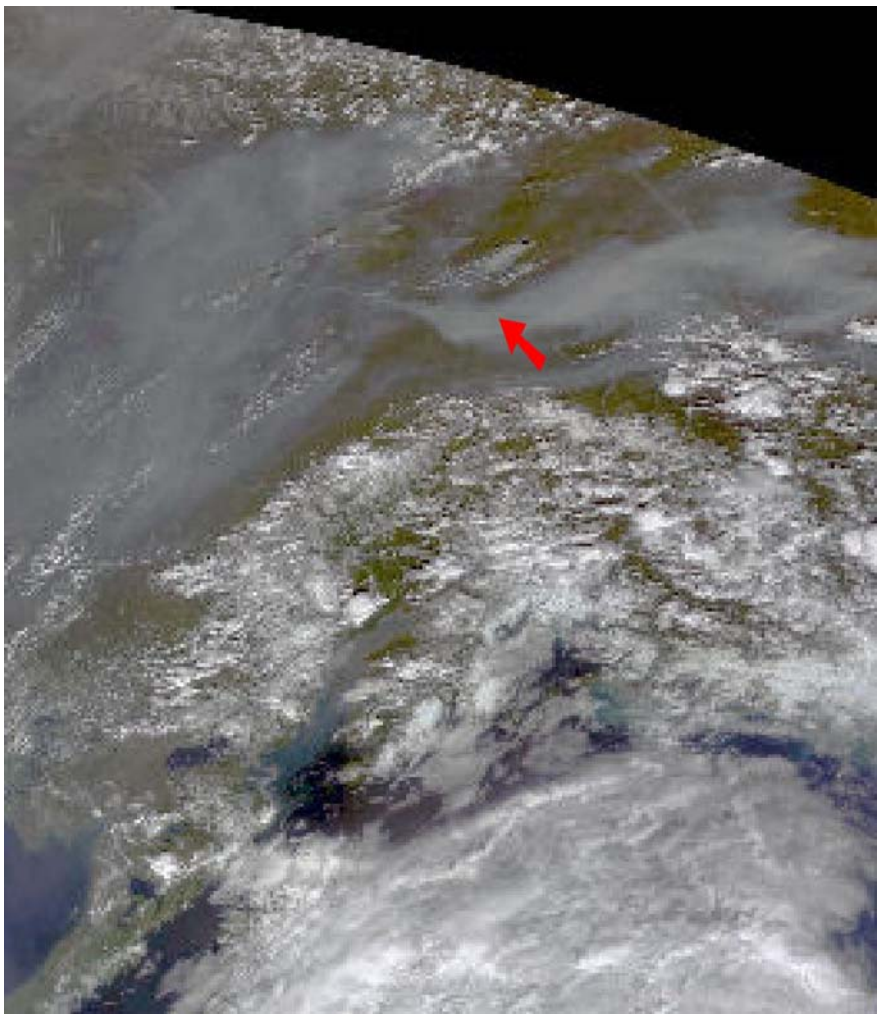


Figure 4.28: MODIS true color composite image of smoke over Alaska. The image corresponds to the 6/27/2005 2135 UTC Terra overpass. The red arrow indicates the approximate location of AFARS beneath the smoke plume**.** Image was generated from a MOD02 1-km geo-calibrated radiance channel file using ENVI.

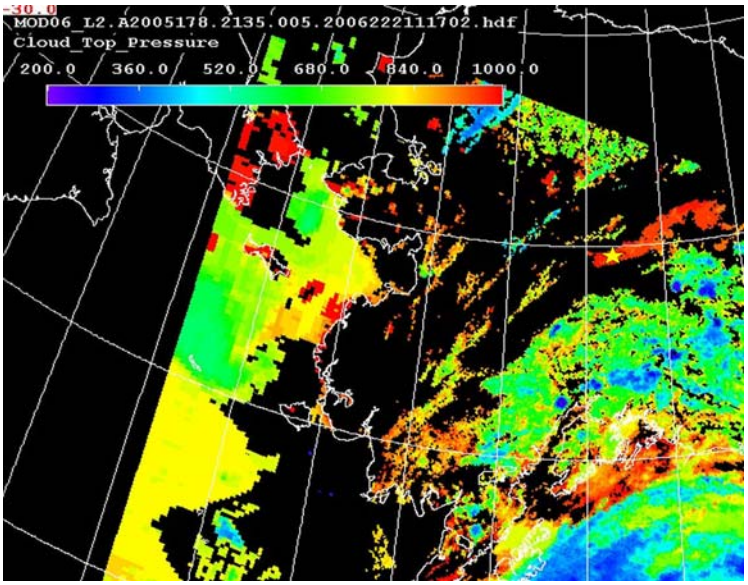


Figure 4.29: MODIS derived cloud top pressure of a smoke layer. This view corresponds to the 6/27/2005 2135 UTC Terra overpass (Figure 4.28) at 5km resolution. The yellow star in the center-right denotes the approximate location of AFARS. Cloud top pressures are given by the legend. Courtesy of *http://ladsweb.nascom.nasa.gov*.

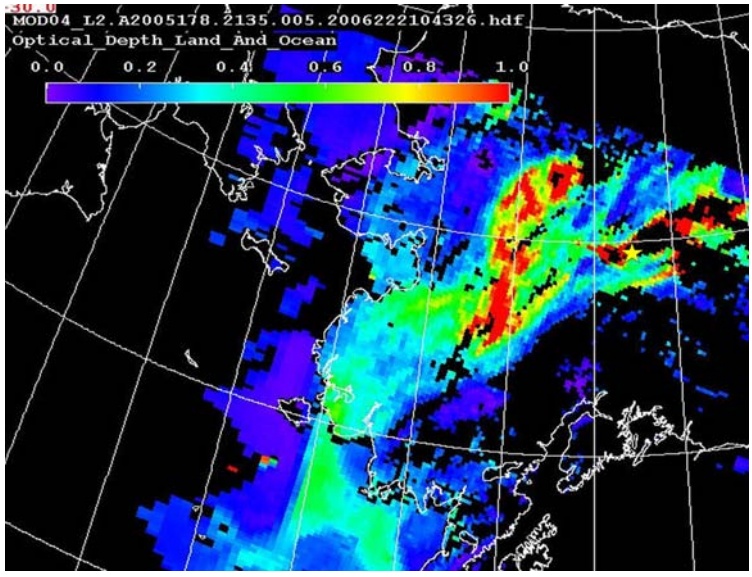


Figure 4.30: MODIS derived aerosol optical depth (AOD) of smoke. The AOD map corresponds to the 6/27/2005 2135 UTC Terra overpass at 10 km resolution. The yellow star, center-right, denotes AFARS. Courtesy of *http://ladsweb.nascom.nasa.gov*.

would almost fit precisely into the colored regions of the former, like a jigsaw puzzle. However, the fit is not perfect. Some overlap exists between the aerosol optical depths and cloud top properties over the same regions, demonstrating how the aerosol algorithm uses a cloud mask different from the main MODIS cloud mask used by the cloud products, discussed in section 2.3.4.1. Hence, it is possible for aerosol and cloud data to be retrieved concurrently from the same area. In fact, the C004 aerosol type determinations and the cloud product data for the June 27, 2005 2135 and 2145 UTC Aqua and Terra overpasses return *smoke* and cloud top property values simultaneously, illustrating that the MODIS cloud and aerosol detection are not mutually exclusive.

The cloud top properties over AFARS for the Aqua and Terra overpasses both yield a cloud top pressure, altitude and temperature of 985 hPa, 233.2 m and 23.4^{o}C respectively. However, the actual top of the smoke layer in the lidar plots falls near 5 km, which is not even close to the MODIS estimation. Furthermore, the MODIS cloud product yields a *liquid* cloud phase result, which is, at least, consistent with the cloud top temperature "sanity check" and with the low lidar depolarization values in Figure 4.27b.

4.7.3 A February 2, 2006 Case Study: Cloud as Volcanic Aerosol

Two *F-P* cases happen to coincide with a volcanic plume originating from Mt. Augustine, an active volcanic island situated on the southwestern side of Cook Inlet. Throughout January and early February of 2006, Mt. Augustine erupted numerous times, spewing ash and steam as high as 12 km into the atmosphere (Sassen et al., 2007). During one such event, the flow pattern was set up such that the ash plume got

transported directly over AFARS. The lidar was able to capture this plume as it passed over Fairbanks on February 2, 2006. Two MODIS overpasses occur during this event: a 2155 UTC Terra and a 2210 UTC Aqua overpass.

Figure 4.31 provides the lidar return power and depolarization ratio plots corresponding to the February 2 ash plume. The layer of volcanic aerosol is approximately 2 km thick and its top sits at approximately 3.5 km. Note the relatively high amounts of depolarization ($0.10 \leq \delta \leq 0.25$) denoted in Figure 4.31b, which is suggestive of small irregularly shaped particles consistent with volcanic ash particles (Sassen et al., 2007). The relatively smooth yet stratified texture of this layer (Figure 4.31a) is also indicative of an aerosol as opposed to the patchy sheared appearance of an ice cloud.

The February 2, 2005 MODIS Aqua true color composite image and the corresponding cloud top pressure map is provided in Figure 4.32 and Figure 4.33. The approximate location of AFARS is denoted by the yellow arrow in the former and the blue arrow in the latter. The counter-clockwise flow pattern advects the plume initially east across Prince William Sound in the lower right and then north over interior Alaska. Results from the Puff model, a Lagrangian computer model designed to track in real-time the trajectory of volcanic ash plumes, further validates the trajectory of Augustine's plume over Fairbanks (Sassen et al., 2007).

The MODIS cloud top data corresponding to the February 2, 2006 2155 and 2210 UTC Terra and Aqua overpasses indicate that the pixels over AFARS possess tops of 620 hPa, 3476 m, and -33.4^{o}C respectively. These values match well with 3900 m, the

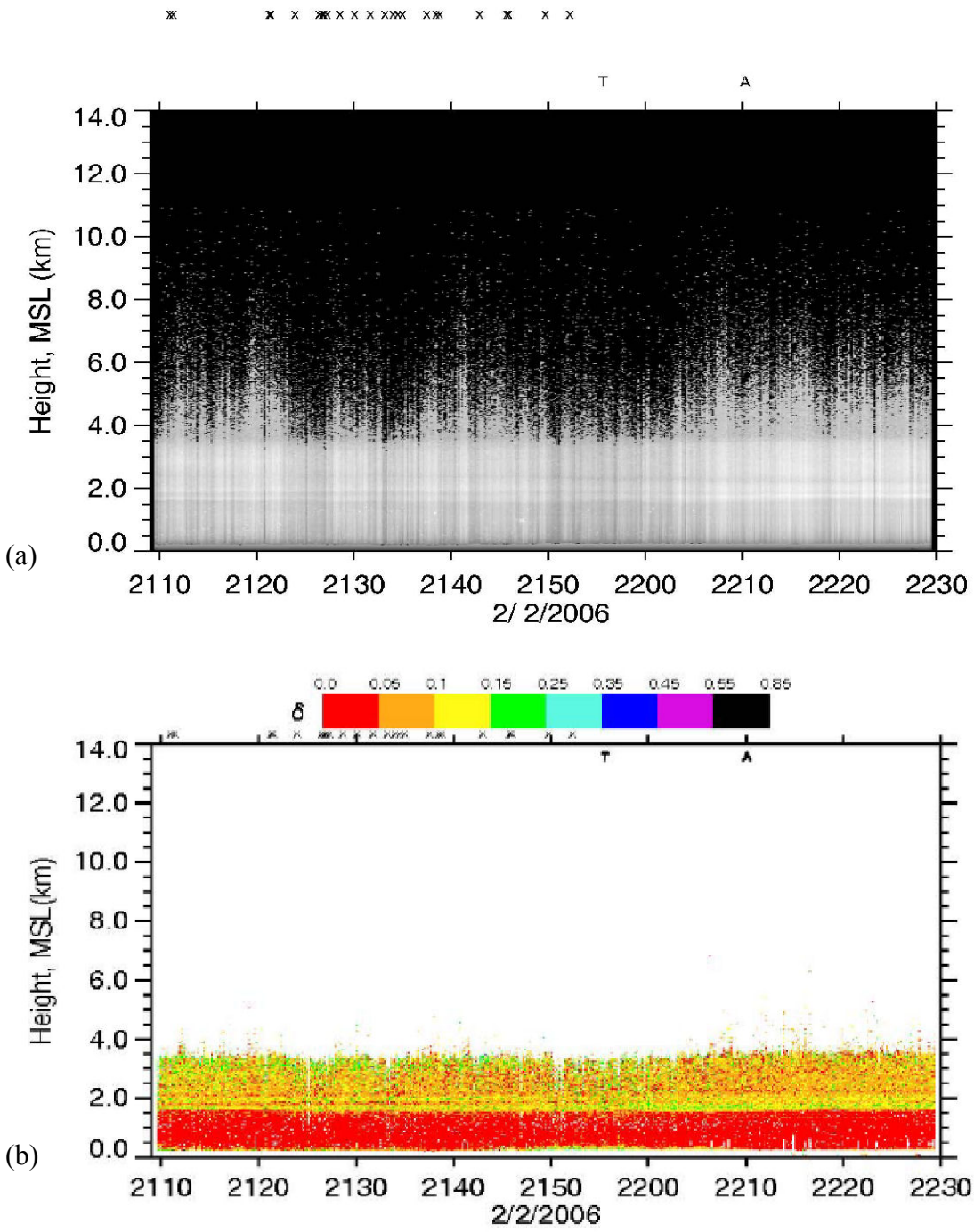


Figure 4.31: Lidar plots capturing Mt. Augustine's ash plume. Plots depicting the lidar (a) returned power and (b) depolarization ratios, which depict Augustine's ash plume over AFARS on February 2, 2005. The A and T symbols at the top of the plot areas denote the Aqua and Terra overpasses.

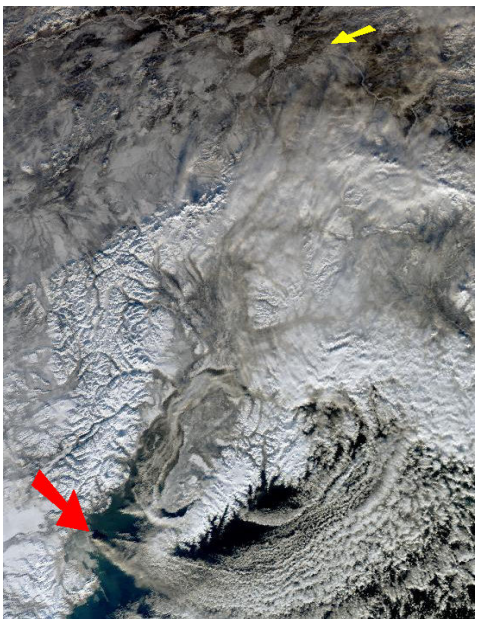


Figure 4.32: MODIS true color composite image of the Mt. Augustine eruption. The image covers south-central Alaska for the 2/2/2006 2210 UTC Aqua overpass. Red and yellow arrows denote approximate locations of Mt. Augustine and AFARS respectively.

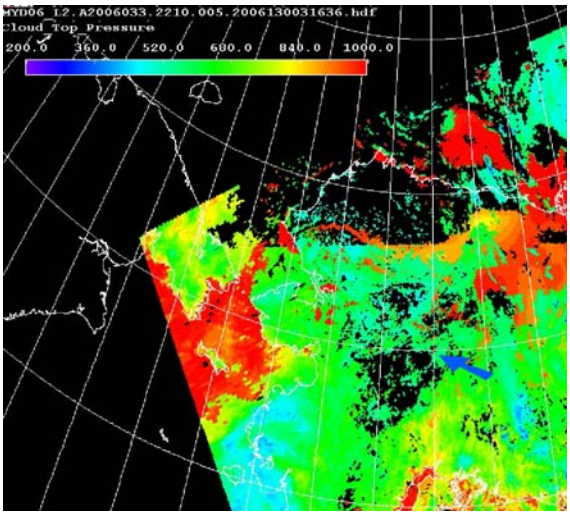


Figure 4.33: MODIS derived cloud top pressure of Mt. Augustine's plume. Map corresponding to the 2/2/2006 2210 UTC Aqua overpass. Pressures are color-coded in the legend. The blue arrow denotes the approximate location of AFARS. Courtesy of *http://ladsweb.nascom.nasa.gov*.

approximate aerosol top derived from the lidar plots, CBCT outputs and field notes.

Interestingly, the air and dewpoint temperature measured by the upper air profiler at 3710 m altitude, are -35.9^{o}C and -39.3^{o}C respectively which puts the relative humidity at 71% with respect to water, but a little over 100% with respect to ice, illustrating that conditions within the plume are favorable for possible heterogeneous ice nucleation.

Also, Terra yields an *ice* phase cloud while Aqua returns *mixed*. The moderately high depolarization values (0.10) associated with the ash plume in Figure 4.31b eliminates the possibility that the particles are either spherical or liquid. However, one can assume that the index of refraction for volcanic mineral particles at 8.55, 11.03 and 12.02 μm (the BTD bands) is different than that of pure ice, which begs the question; if the BTD algorithm detects an ice-like signature, how much water ice is present in this layer and to what degree are these ash particles behaving like CCN?

It is clear that even though the MODIS falsely identifies this layer as a cloud, it still detects and accurately estimates the location of this layer in the atmosphere above AFARS. However, it is difficult to discern the water content from strictly the lidar plots, and to know for certain the ratio of water to ash particles without in situ measurements, but perhaps the MODIS determination is more correct than previously thought.

4.8 Comparison by Optical Depth

The methodology followed to estimate cloud optical depth is presented in section 3.8. Recall that optical depth is not computed quantitatively, but rather estimated subjectively using the field note descriptions and the lidar plots. There are six groups of

optical depth, not including the clear-sky category, which range from "0.03-" for subvisual to "3.0+" for attenuation-limited. The specific categories are listed in section 3.8.2.

Figure 4.34 provides the breakdown of the sample by optical depth. The largest percentage of cases falls within the thin (τ = 0.03+) category, while the second largest are deemed to be attenuation-limited. The other cases are distributed rather evenly between the opaquish (τ = 3.0-), opaque (τ = 0.30+), thinnish (τ = 0.30-) and subvisual (τ = 0.03-) categories. Note, after re-examining all 549 cases, the number of attenuation- limited cases presented in Figure 4.34, ninety-three, is ten fewer than the number initially reported in section 4.1.1. Similarly, the seventy-two *no cloud* cases, which theoretically combines the forty-eight *clear* and twenty-two *aerosol only* categories in Figure 4.2, does not in fact, match the sum of these latter two groups. With respect to the attenuation-

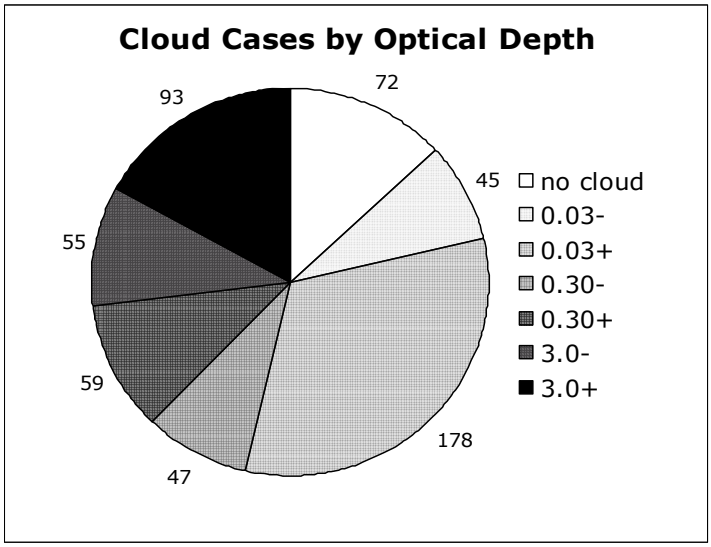


Figure 4.34: **Breakdown of sample by estimated cloud optical depth**. The cloud groups range from subvisual (0.03-) to attenuation-limited (3.0+).

limited cases, several borderline altostratus clouds initially deemed to be attenuation-limited are, after a second look, later considered to be "opaquish". With respect to the *no-cloud* cases, two of the ten *ice-fog* cases in Figure 4.2 are classified under *no-cloud* in Figure 4.34. These relatively minor inconsistencies are not expected to alter the results presented in the preceding sections in any significant way. But this illustrates how utilizing a relatively subjective methodology to analyze a dataset may lead to inconsistent results even when efforts are made to standardize the analysis.

The MODIS cloud mask and cloud top property errors, when sorted by estimated optical depth, tend to separate out in a predictable manner with the greatest amount of error associated with the thinnest cloud categories, and the least amount associated with the thickest. Note how the false identifications are generally distributed by estimated optical depth in Figure 4.35. First, MODIS incorrectly detects cloud (F-P) about one third of the time assuming cloud-free conditions in the lidar FOV. This is validated by the results given earlier in Figure 4.1 where the rate of F-P identification is approximately 32.8%. Second, the rate of F-N identifications initially starts around 70% for the "subvisual" category, but then drops rapidly to below 40% for "thin" and then fluctuates between 2 -7% for the remaining four categories. As expected and observed thus far, the thinnest clouds pose the greatest challenge to the MODIS sensor.

These results are also further validated by the MODIS cloud mask agreements. Figure 4.36 gives the summary of cloud mask agreements sorted by estimated optical depth, and is almost the inverse of Figure 4.35. Theoretically, the columns for each

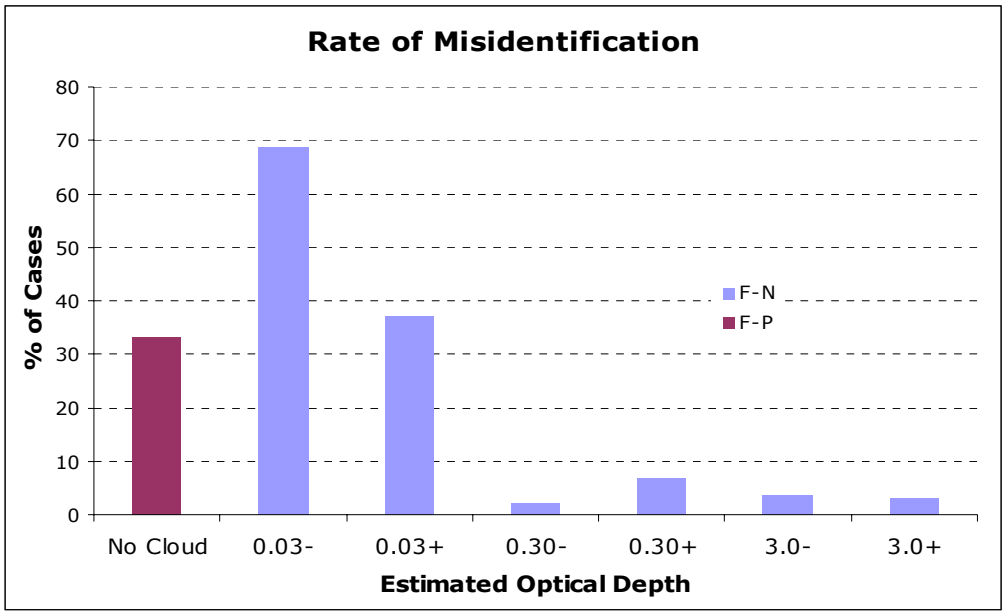


Figure 4.35: **Distribution of false identifications by estimated optical depth**. Cloud detections are based on the results of the MOD06 cloud top property dataset.

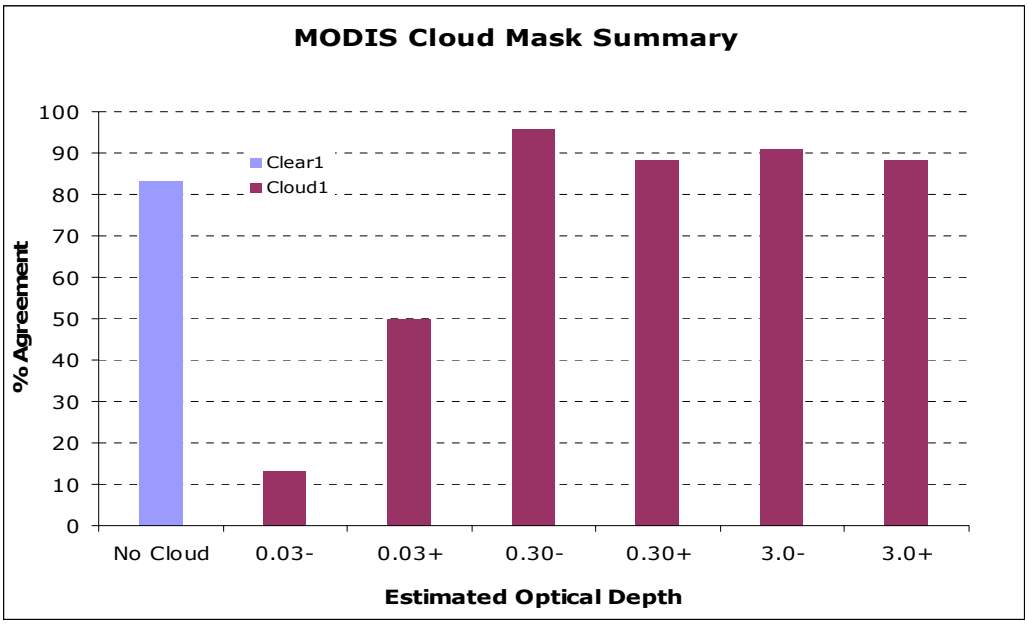


Figure 4.36: **Summary of MODIS cloud mask agreements by estimated optical depth**. This plot provides the *clear*[1] and *cloud*[1] results as *clear*[2] and *cloud*[2] are meaningless in the context of this analysis.

category in Figure 4.35 and in Figure 4.36 should equal 100% when combined, however, they do not for the same reasons described in section 4.4.2 regarding the inconsistencies between the number of *clear* cases returned by the MODIS cloud phase dataset versus the number returned by the cloud mask. With respect to the lidar *no clouds* category, the probability that MODIS agrees with the lidar ($clear^1$) exceeds 80%. As in Figure 4.35, the cloud detection agreements in Figure 4.36 become progressively better as estimated cloud optical depth increases from "subvisual" to "thinnish" and then subsequently levels off to around 90%. Note that the MODIS cloud mask agreements in Figure 4.36 are determined following the same procedures described in sections 3.4 and 4.2.2, using tables similar to Table 4.1. Also note that the approach^2 findings are lacking from Figure 4.36 since the lidar sky observations are, in essence, absolute, meaning that there are no lidar *cloudy* cases for the clear comparison, nor are there any lidar *clear* cases for the cloud comparison. Hence, $clear^2$ will always equal 100% for *no cloud* and 0% for the remaining cloud groups. Likewise, $cloud^2$ will always equal 0% for *no cloud* and 100% thereafter. Therefore, the approach^2 findings in this context are essentially meaningless.

The mean cloud top errors arranged by estimated optical depth are provided in Figure 4.37. Recall from the discussion in section 2.3.2.2 that for the case of single layer optically thin clouds, the CO_2 density slicing method will tend to estimate the cloud top near the cloud's "radiative center-of-mass" below the actual geometric top. Hence as optical thickness increases, one generally expects the error to decrease, which is precisely what is manifested in Figure 4.37. Mean maximum differences in cloud top pressure,

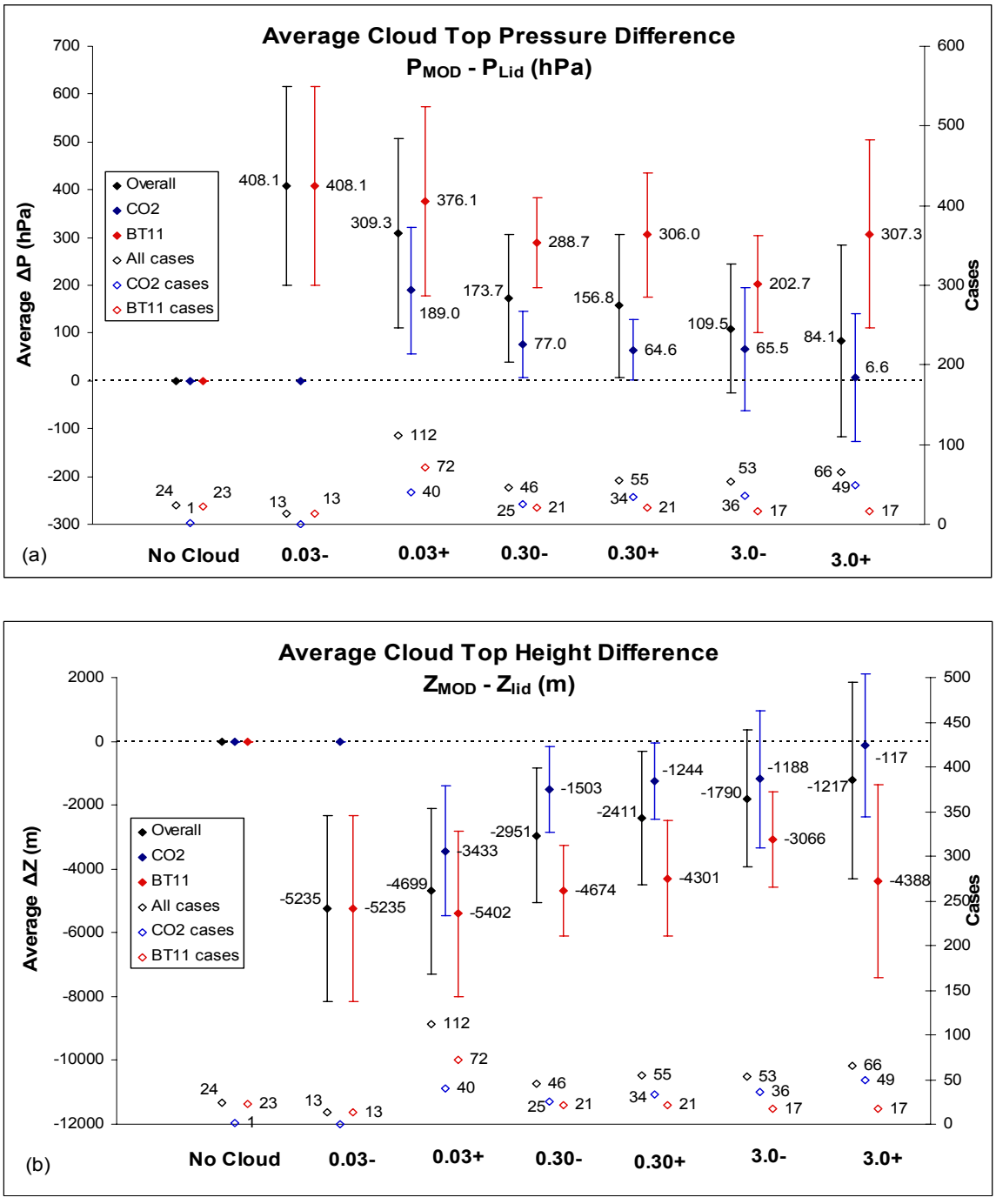


Figure 4.37: **Mean and 1 standard deviation of cloud top errors by optical depth**. Plots give the (a) cloud top pressure (hPa) and (b) cloud top height (m) errors. Red, blue and black diamonds denote BT_{11}, CO_2 and all (BT_{11} + CO_2) results respectively. Hollow symbols signify case numbers associated with each group. Case numbers and mean errors are printed next to their respective symbols for convenience.

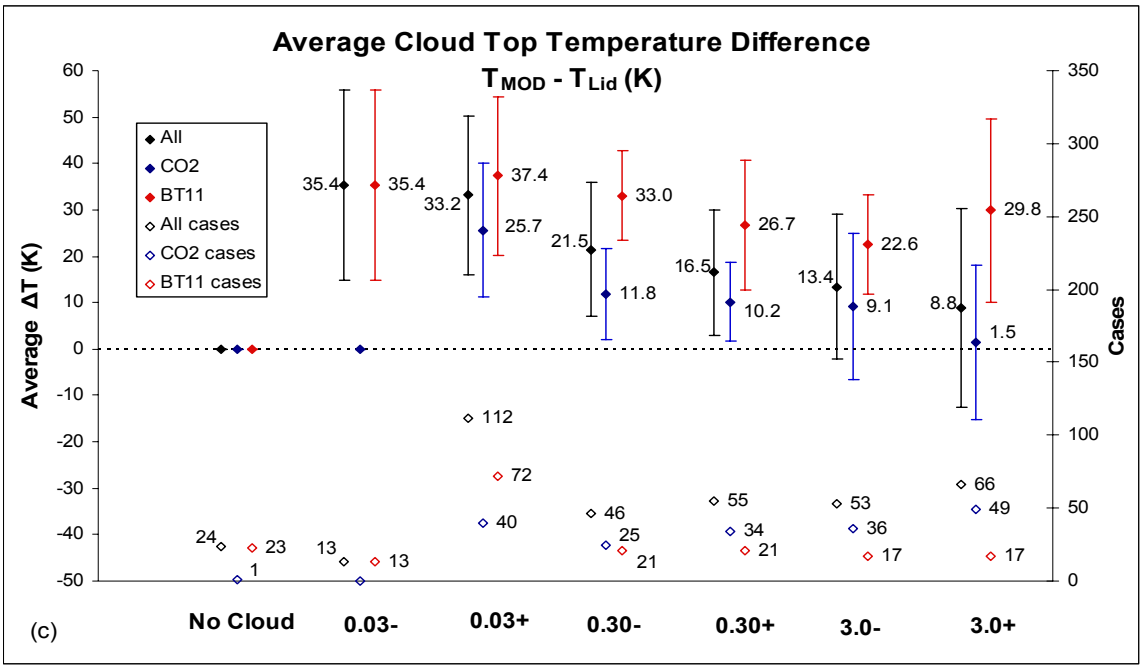


Figure 4.37 cont'd: Mean and 1 standard deviation of (c) cloud top temperature error (K) sorted by estimated cloud optical depth. Categories, colors and symbols are analogous to those presented in Figure 4.37a and 4.37b.

height and temperature between MODIS and the lidar are largest for the subvisual (0.03-) group and smallest for the attenuation-limited category (3.0+). Also note how the magnitude of the error is correlated with the frequency with which the CO_2 density slicing method gets applied; the CO_2 method is selected 0% of the time for the subvisual group containing the largest amount of error, but increases steadily to over 60% of the time for the opaque (0.30+) and opaquish (3.0-) groups, cases with the smallest associated error. Given how MODIS is more likely to see "thick" cirrus clouds, it follows then that the chances MODIS properly selects the CO_2 method increases for those high cirrus clouds that it can perceive (thick clouds) as opposed to those that it may miss altogether (thin clouds).

Mean and 1 standard deviations of cloud top error based on the number of cloud layers (single versus multiple) are plotted in Figure 4.38. Note that the diamonds and circles denote single and multi-layer (+2) groups respectively. As in previous assessments, small sample sizes for several sub-groups in Figure 4.38 unfortunately reduce the confidence levels attributed to the mean error calculations, resulting in unusually large or small deviations as is evident for several of the categories. However, some results can still be salvaged.

First, the single layer CO_2 derived errors are predictably smaller than those linked to the multi-layer calculations for the thin cloud categories (0.03+, 0.30-). Note CO_2 is not selected for any of the subvisual (0.03-) cases, since these clouds are, in all likelihood, too translucent for MODIS to detect. At the 0.30+ group and above, the differences in the errors become a bit more arbitrary. Recall from section 2.3.2.2 that the largest theoretical cloud top pressure difference ($P_{MOD} - P_{actual}$) of 220 hPa occurs for an optically thin ice cloud situated above a thick water cloud. In this sample, most of the non-attenuating multi-layer cases are comprised of two or more cirrus layers of unknown optical depth. Nonetheless, MODIS cloud top estimates for single layer cases are still more accurate for both the optically thin cloud cases (Figure 4.38) and for the sample overall (Figure 4.11).

Note that for the attenuation-limited category (3.0+), the mean CO_2 derived mean pressure error for single layer clouds actually goes negative (-63.8 hPa) while for the multi-layer clouds, it stays positive (80 hPa). Theory supported by the findings up to this

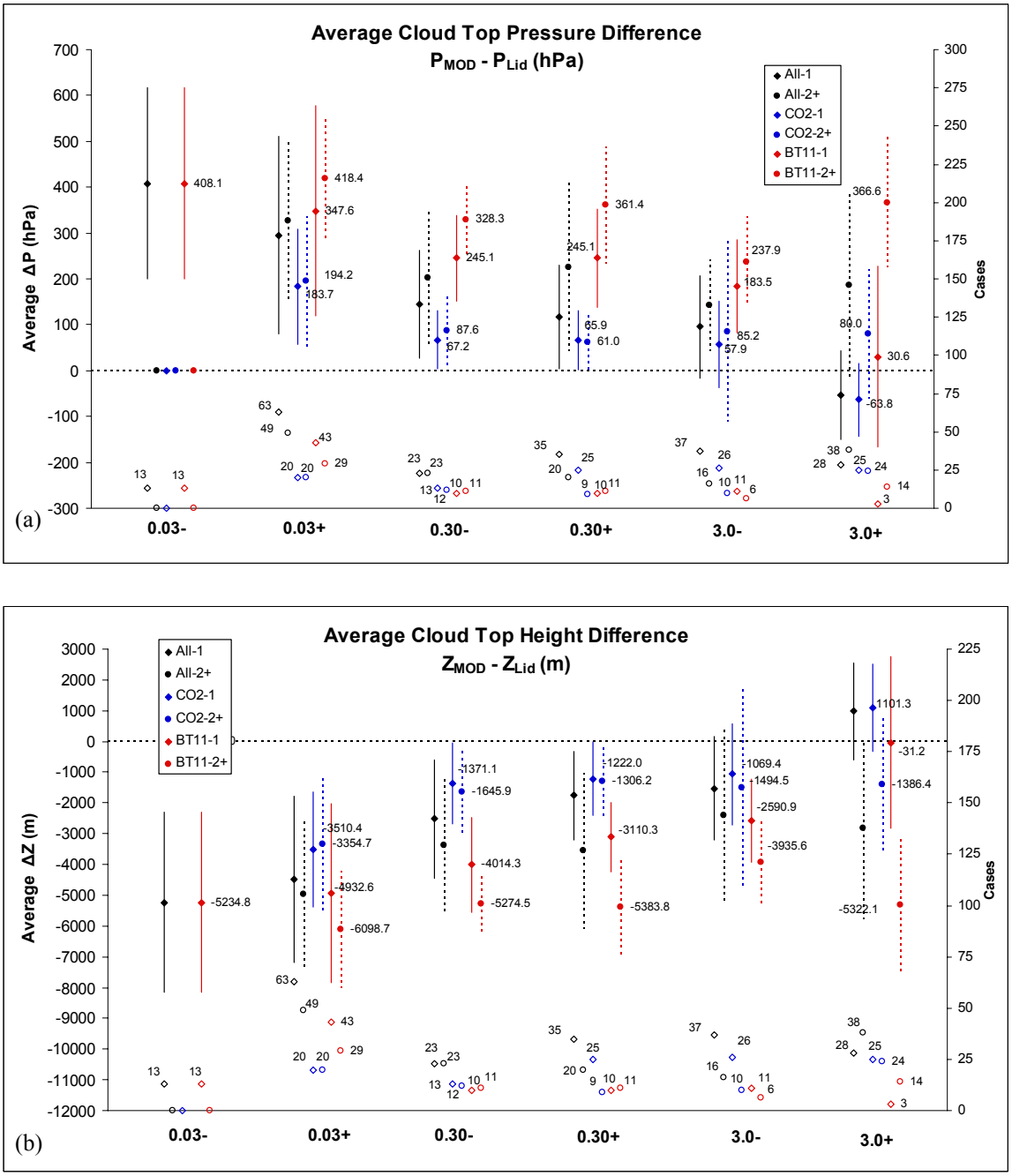


Figure 4.38: **A comparison of the mean and 1 standard deviaton of cloud top error sorted by cloud layers**. The plots give (a) cloud top pressure (hPa) and (b) cloud top altitude (m) for single layer cases (diamonds with solid error bars) and multi-layer cases (circles with dashed error bars) sorted by optical depth. Blue, red and black signify errors associated with the CO_2, BT_{11} and All (CO_2 + BT_{11}) cases respectively. The number of cases associated with each category is represented by the hollow symbols.

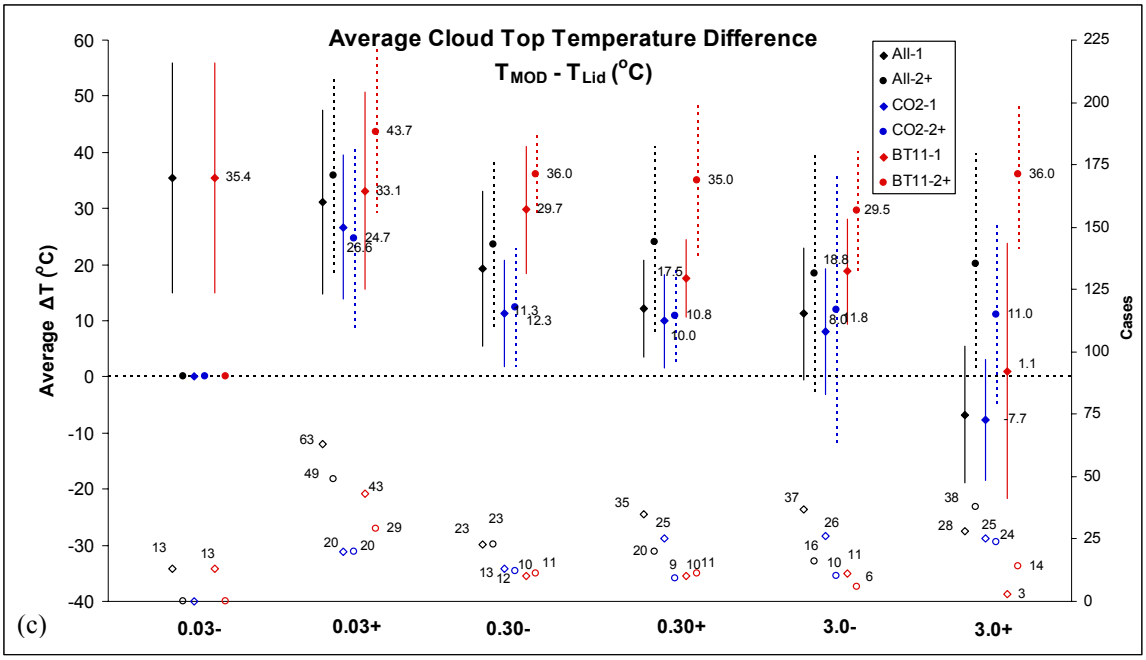


Figure 4.38 cont'd: Mean and 1 standard deviation of (c) cloud top temperature (°C) error for the single and mult-layer cloud cases sorted by optical depth. Categories, colors and symbols are analogous to those in Figure 4.38a and 4.38b.

point indicate that this negative mean pressure error is somewhat of an anomaly, but bear in mind that single layer attenuation-limited cases are likely to consist of thick altostratus ice clouds. Even though penetration of the lidar signal may not be complete, a cloud top altitude may still get documented. Such an underestimation in the cloud top height translates to an overestimation in pressure leading to a negative pressure difference as is seen for the single layer 3.0+ group in Figure 4.38a. In all likelihood, the MODIS cloud top height (pressure) is still being undervalued (overvalued) somewhat, but it is impossible to know definitively, without knowing the actual cloud top altitude.

On the other hand, multiple layer attenuation-limited cases are much more likely to consist of cumulus or alto-cumulus clouds below a cirrus layer. In this situation the

cirrus top can still be correctly delineated in the lidar data since cumuliform clouds only tend to block the lidar signal for short periods of time. Therefore, cloud top pressure differences are much more likely to remain predictably in positive territory, as is evident in Figure 4.38.

The distribution of cloud phase agreement sorted by estimated lidar derived cloud optical depth is depicted in Figure 4.39. Again the clear comparison (Figure 4.39a) uses the results derived from the cloud phase dataset, not those derived from the cloud mask. As with Figure 4.16, circles and diamonds denote approach^1 and approach^2 results respectively and the hollow symbols give the associated case numbers.

The clear results, Figure 4.39a, are consistent with earlier findings, and like the cloud mask analysis presented in Figure 4.36, the approach^2 results with respect to this comparison can be ignored. Of the seventy MODIS *clear* cases in the lidar *no cloud* group, the lidar data match 67.1% of the time. MODIS also returns a significant number of *clear* cases for the subvisual (0.03-) and thin (0.03+) groups, reiterating its aforementioned weakness at detecting thin clouds.

More specifically, its weakness at detecting thin *ice* clouds is best depicted in Figure 4.39b. MODIS *ice* matches with the lidar only 2.3% and 21.3% of the time for the thinnest cloud categories (0.03-, 0.03+) respectively. However, this agreement improves markedly for the thicker cirrus groups possessing an estimated optical depth of around 0.30 or higher. The *ice*^{1} agreements for these thicker groups range from 54.4% to 71.1%, and have relatively large samples. The *ice*^{2} agreements are consistently high in many

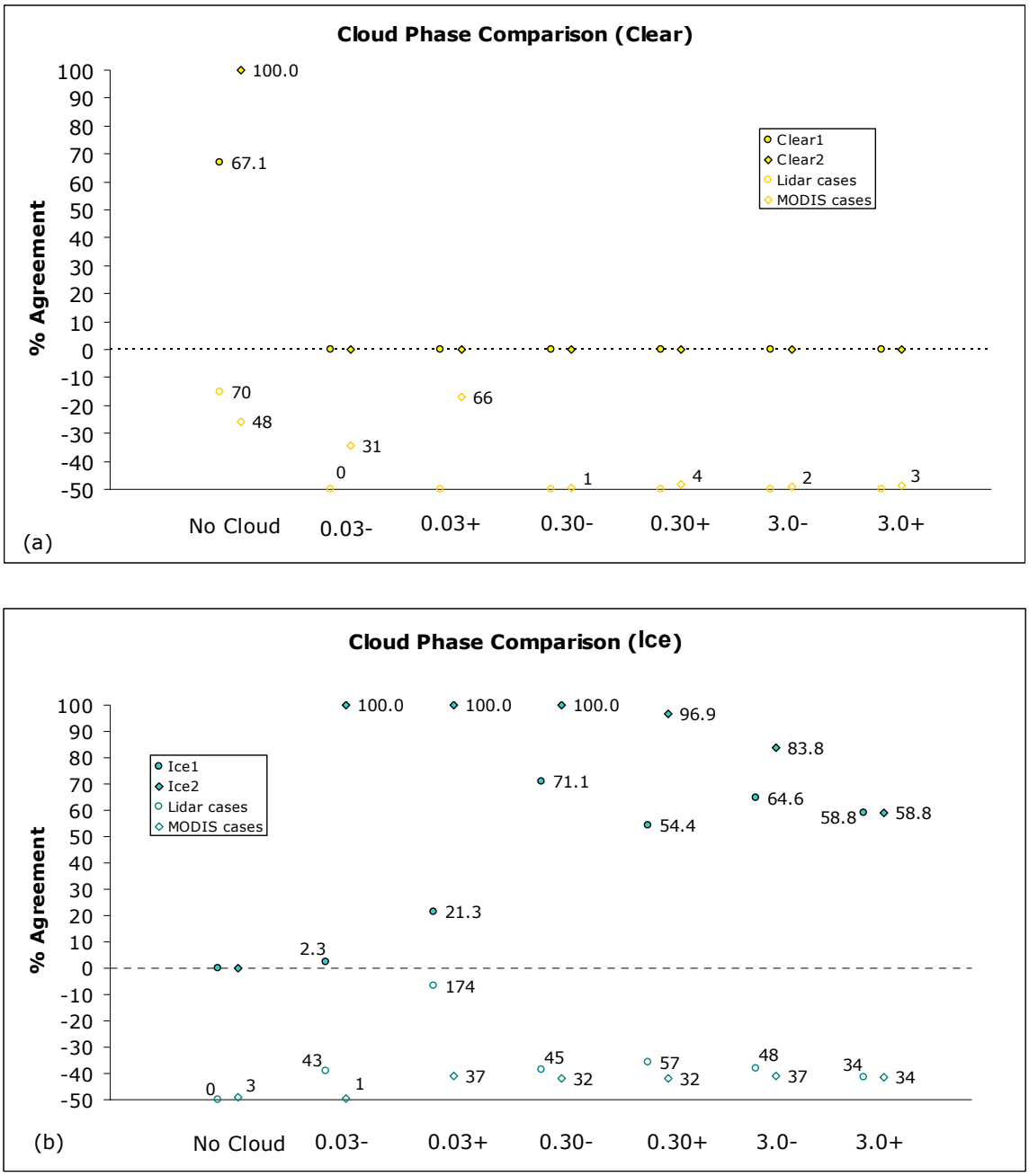
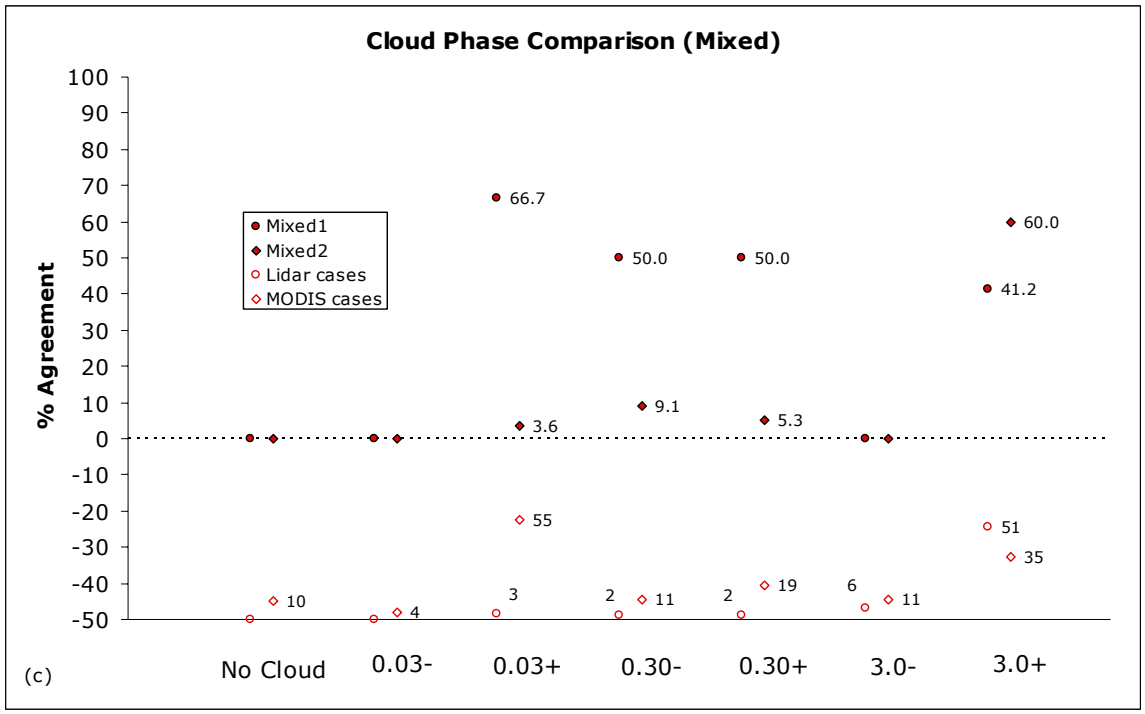


Figure 4.39: **MODIS cloud phase agreements sorted by estimated optical depth**. The comparisons for the (a) clear and (b) ice phase clouds are given. Solid circles and diamonds denote the approach[1] and approach[2] values respectively. Hollow symbols represent the respective case numbers. Percentages and case numbers are printed next to their respective symbols for convenience.



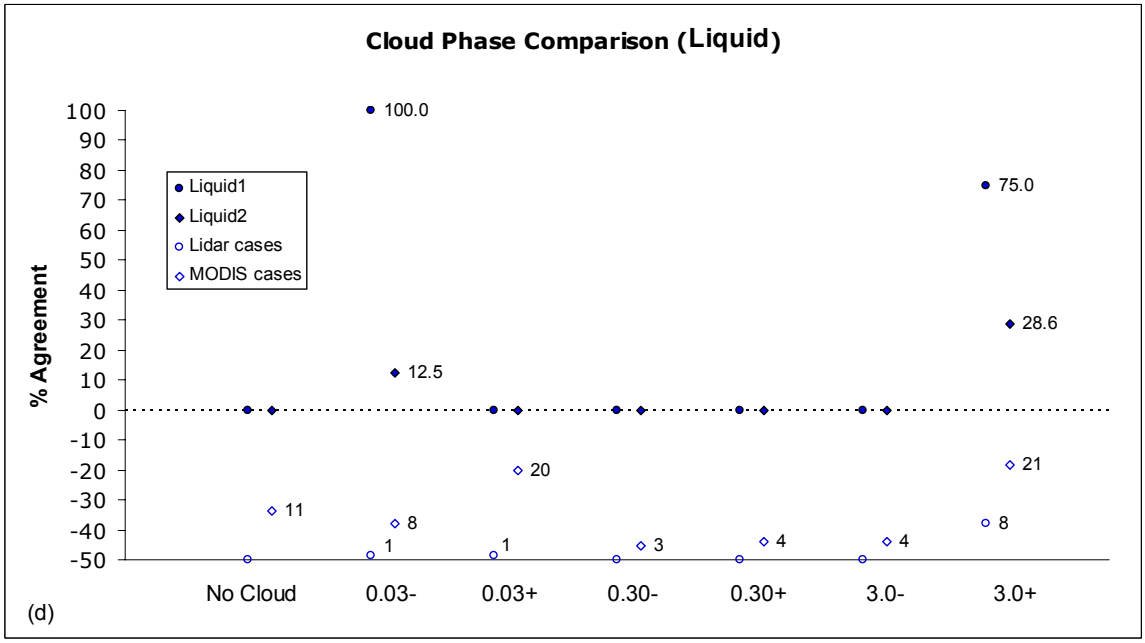


Figure 4.39 cont'd: Cloud phase agreements for the (c) mixed and (d) liquid phase clouds. Categories and symbols are synonymous with those depicted in Figure 4.39a and 4.39b.

cases even yielding perfect agreements for the "thin" and "thinnish" groups. But the agreements drop steadily as cloud optical thickness approaches the attenuation-limited threshold. This is likely an indication of an increase in the frequency of the *mixed* and *liquid* observations in the lidar data as optical depth increases.

As for the *mixed* and *liquid* results presented in Figure 4.39c and 4.39d, it is clear that MODIS frequently and inappropriately returns *mixed* and *liquid* phase types for many cases of thin *ice* clouds. Under most circumstances, optically thin *mixed* or *liquid* phase clouds do not exist, hence the lidar sample sizes associated with the approach[1] agreements are exceptionally small. Of the six optical depth groups, only the attenuation-limited group (3.0+) is worthy of attention, since it is the only category with any sizeable number of lidar *mixed* or *liquid* cloud observations.

The *mixed*[1] and *mixed*[2] agreements in Figure 4.39c are generally fair assuming optically thick clouds. Given a lidar determination of *mixed*, MODIS agrees 41.2% of the time, and given a MODIS result of *mixed*, the lidar agrees 60% of the time. Both agreements also have relatively large samples, fifty-one and thirty-five respectively. The *liquid* agreements seem to be a little less reliable. MODIS agrees with the lidar 75% of the time, however the lidar only documents a total of eight *liquid* phase cloud cases. Of the twenty-one MODIS *liquid* cases, the lidar only agrees 28.6% of the time. Keep in mind that many of these attenuation-limited cases are comprised of multi-layer ice over liquid scenarios. So in all likelihood, there is a high chance that cases of thin *ice* cloud over a *liquid* cumuliform cloud, which would be denoted as a *mixed* case in the lidar data, are getting classified as *liquid* by MODIS.

Overall, given the inherent drawbacks of this dataset and the complex surface cover of the field site, the MODIS cloud products seem to deliver what they promise. The important results will be summarized in the next and last chapter, chapter five.

Chapter 5 Conclusions and Future Work

Rising sea levels, more extreme weather events, and geographical shifts to plants, animals and diseases are expected to have major social and economic consequences in the coming century. If mankind is to adapt, it is necessary to anticipate how ever increasing greenhouse gas concentrations and their associated feedback mechanisms (e.g. water vapor feedback) will alter temperature and precipitation patterns at the local, regional and global scale. The climate system is complex: the variables influencing the climate are numerous and their coupled relationships, nonlinear. Global circulation models currently offer the best tool with which to simulate climate, but their capacity to predict future climate scenarios, although improving, remains in question.

Clouds are widely considered to be the source of greatest uncertainty in climate model predictions. One reason for this is due simply to a basic lack of knowledge of their distribution in space and time. Better global measurements would surely lead to improved GCM parameterizations and more reliable predictions, but expensive, small-scale, in-situ measurements alone, do not offer the spatial and temporal coverage of cloud and aerosol data needed parameterize the climate models.

Satellite sensors, such as MODIS, have the potential to alleviate this problem. MODIS atmosphere products provide a global inventory of the physical and radiative properties of clouds and aerosols every day at relatively high spatial resolutions. But, can the data be trusted? Inferring cloud top altitude, temperature, water content, phase and particle size distribution using passive remote sensing techniques is inherently difficult.

In response, a slew of validation studies from the ground, air and space are currently underway to pinpoint the weaknesses of the MODIS algorithms in hopes that they can be improved.

This thesis presents one such study. A long-term comparison of cloud and aerosol data between a ground based PCL and the MODIS sensor is performed. The comparison specifically examines the following data: the presence or absence of cloud, cloud top properties, cloud thermodynamic phase and aerosol type. The data analysis is further broken down by MODIS platform, snow cover conditions, time of day, month of the year, number of cloud layers, sensor zenith angle, and cloud optical and geometric thickness.

The sample is biased towards thin cirrus clouds, clouds that are particularly challenging to detect using space-borne passive sensors. Additionally, the training site, AFARS, is located in a sub-arctic mixed urban environment where darkness, low sun angles, snow cover and strong low-level temperature inversions are present during much of the year. Such conditions make cloud detection by MODIS an especially difficult undertaking, but nonetheless important considering the growing consensus that cirrus clouds augment the greenhouse effect. With respect to the cloud mask comparison, the lidar observations are considered to be the ground truth. However the protocols for deriving cloud top altitude, phase and optical depth are rather subjective since they rely primarily on visual inspection of the lidar plots and field notes. Nevertheless, the results presented in chapter four compared to other studies (Frey et al., 1999; Mahesh et al.,

2004; Liu et al., 2004; Mace et al., 2005) suggest that the methods described in chapter three are sufficiently standardized. The findings are briefly summarized below.

MODIS failed to detect cloud 107 of the 476 times (Fig. 4.1) when cloud was evident over AFARS (*F-Ns*), and the lidar -plots seem to indicate that most of these *F-N* cases are ice clouds (Fig. 4.3). This is supported further in section 4.7 where it is found that the vast majority of *F-Ns* fall within the subvisual (0.03-) and thin (0.03+) cloud categories (Fig. 4.34), but for clouds with an estimated optical depth larger than 0.30, MODIS rarely misses. It is also found based on the MODIS aerosol data that MODIS frequently misidentifies thin ice cloud as aerosol, usually smoke (Fig. 4.22).

MODIS falsely detects cloud twenty-four of the seventy-three times when cloud is absent over AFARS (F-P), many, of which, appear to involve cases of mistaken identity with thick aerosol events. Two such case studies are presented in sections 4.6.2 and 4.6.3. One involves a moderately depolarizing volcanic aerosol layer, which MODIS mistook for an ice cloud. In this situation, the MODIS cloud top estimate, 3476 m, fell within 0.5 km of the actual top. Upper air soundings collected two hours later, find this layer to be saturated with respect to ice, which may vindicate somewhat the MODIS assessment. The other case study involves a dense layer of smoke, which exhibits a weakly depolarizing signal. MODIS falsely identifies this layer as a low-level water cloud, but did not accurately calculate its top.

Results pertaining to the MODIS cloud mask are mixed. First, the variability in the performances between the three defined methods, '5 km point,' '1 km point' and '1 km box,' is small, although, on the whole, '5 km point' outperforms the other two by 7-

12% for the *clear*[1] category. The median probability that a MODIS cloud mask finding of clear (cloudy) agrees with the lidar is approximately 77% (68%) (Fig. 4.6). Conversely, the probability that a lidar clear (cloudy) agrees with MODIS is 27% (95%). The poorest agreement associated with the *clear*[2] category is yet another sign that MODIS misses a significant number of the thin ice clouds.

When sorted by estimated optical depth, the cloud mask agreements separate out in predictable ways, with the worst agreements associated with the thinnest cloud layers and the best, with the thickest (Fig. 4.36). Additionally the snow/ice background flag, extracted from the 5-km cloud mask bit, significantly undercounts the times when the ground is snow covered (Fig. 4.4a). However, it remains unclear how choosing the wrong cloud mask "domain," in this case, *no snow* over *snow*, contaminates other cloud product algorithms such as thermodynamic phase and cloud top properties.

The cloud top property (P, Z, T) estimates demonstrate an enormous disparity between the CO_2 derived tops and the BT_{11} derived ones (Fig. 4.7-4.11). Although both, as expected, underestimate (overestimate) the cloud top height (pressure, temperature), the average difference in cloud top height (pressure, temperature) associated with the CO_2 cases is -1444 m (78.1 hPa, 11.2 K), ballpark errors consistent with the ATBD calculations. Conversely, the average BT_{11} cloud top error falls around -4796 m (332.6 hPa, 32.9 K) (Fig. 4.11a-c), which is not even close. It is hypothesized that many of these large errors are attributed to MODIS falsely detecting the presence of a non-existent low-level water or mixed phase cloud while altogether missing a thin ice layer. Evidence for this hypothesis can be found in the cloud top error by optical depth analysis (Fig.

4.37). Both the errors and the percentage of cases using BT_{11} are high for the thin cloud categories, which MODIS is most likely to miss. But as optical depth increases, both the mean cloud top error and the frequency with which BT_{11} gets selected decrease. This is due in part to two reasons: first, MODIS is better able to estimate clouds that it can actually "see," as opposed to clouds that it cannot, and second, as optical depth increases, the "radiative center-of-mass" approaches the actual geometric top of the cloud.

Cloud top retrieval method is also found to follow a seasonal pattern (Fig. 4.4b) in which BT_{11} is preferred during the colder months (October – April), while CO_2 is the method of choice during the warm months (May – September). Recall, the BT_{11} method gets selected when the clear minus cloudy sky radiances fall within instrument noise levels, which is more likely to occur in winter when the visible and thermal contrast between cloud and the cold snow covered ground is poor. Conversely during summer, the contrast between cold bright cloud and the warm dark surface is significantly greater, which increases the chance that the algorithm correctly chooses the CO_2 method. In addition, twenty-three of the twenty-four F-P cases (96%) are associated with the BT_{11} method, which compares to 51%, the overall BT_{11} selection rate (Fig 4.26).

With respect to platform, Aqua's CO_2 derived estimates fall much closer to the lidar estimate than Terra's, however it is found that Terra is much more likely to choose the CO_2 method over BT_{11}, which, within the context of this cirrus-biased dataset, is the more appropriate of the two methods (Fig. 4.11). The age difference between the two sensors could be a factor.

Surprisingly, the average cloud top errors associated with the *night* and *snow* categories are smaller than those associated with *day* and *no snow* (Fig. 4.11). This could be attributed to differences in sample size with the *day* (*no snow*) cases exceeding *night* (*snow*) by nearly three to one. However, *day* and *no snow* do select the CO_2 method at a higher rate than do *night* and *snow*.

As expected, multiple cloud layers are found to influence the magnitude of the cloud top errors (Fig. 4.11), which is especially weighted towards the "thin" and "thinnish" cases (Fig. 4.38). The average CO_2 derived cloud top pressure error for cases containing more than one cloud layer, 15 hPa, was nearly double the error, 8.7 hPa, associated with single layer cases.

Cloud top error is not found to be correlated with sensor zenith angle (Fig. 4.13), indicating that the algorithms do properly account for changes in path length due to changes in viewing geometry. Also, a small negative correlation between cloud top error and the geometric thickness of the top layer is observed (Fig. 4.14) although the correlation coefficient is less than 0.10, and is even smaller when the results are parsed by top retrieval method. But, a correlation between cloud top error and estimated optical depth is found to exist (Fig. 4.37).

Results of the phase analysis are also mixed and are unfortunately marred by small sample sizes. In this study, MODIS frequently overestimates *clear*, *liquid* and *mixed* clouds while it underestimates *ice* clouds (Fig. 4.3). Eight cases of *mixed* cloud and two cases of *liquid* failed the cloud phase "sanity check" described by King et al.

(2004) (Fig. 4.15). The corresponding cloud top temperatures for these ten cases are all found to be colder than -40^{o}C, the cut-off for *ice* only.

Regarding the phase comparison, the *ice*^{2} analysis, by far, yields the best agreement; given a MODIS result of *ice*, the chance that the lidar agrees is 86.4% (Fig. 4.16b). Conversely, of 403 lidar *ice* cases, MODIS agrees 37.7% of the time. The other consistently high agreement, 70%, is associated with the *liquid*^{1} group, although, again, the sample size is abysmally small; only ten cases of *liquid* only exist from the lidar - plots, even though MODIS returns *liquid* seventy-one times (Fig. 4.16d). With exception to the *clear*^{1} results, the remaining agreements are consistently at or below 50%.

When sorted by estimated optical depth, the phase agreements exhibit some rather predictable patterns (Fig. 4.39). *Ice*^{1} agreements are poor for thin clouds, but then increase steadily as estimated optical depth increases. This implies that the *ice* agreements between MODIS and the lidar are relatively good so long as MODIS can "see" the cloud. With respect to the *mixed* and *liquid* phases, it is clear that MODIS does tend to overestimate the number of both cloud types during situations of thin cirrus cover for reasons already discussed. However, the *mixed* and *liquid* agreements for the attenuation-limited category (3.0+), the group most likely to contain lidar detected *mixed* and *liquid* clouds, is rather strong, with the exception of *liquid*^{2}.

In general the comparison by aerosol type is rather inadequate. To start, an older version of the MODIS aerosol data, C004, is used since the updated version, C005, seemed to be giving spurious results. Unlike the C005, the algorithm used to generate the C004 product does not detect dust over land. Hence one of the two major sources of

aerosol pollution over Alaska could not be properly evaluated. The other main aerosol, smoke, did appear to demonstrate some conformity. *Smoke*[1] and *smoke*[2] agreements fall around 54% and 42% respectively (Table 4.3).

The aerosol observations did prove useful when evaluating the false identifications (*F-N*, *F-P*). As discussed in great detail, MODIS frequently confuses clouds for aerosols and vice versa. The two most notable examples, mistaking volcanic aerosol for an *ice* cloud and smoke for a low-level *liquid* cloud, have been cited.

In summary, the ground based PCL has proven to be an excellent and inexpensive source of cloud and aerosol data with which to validate overlapping satellite measurements. It can detect the presence of clouds and aerosols unambiguously and pinpoint precisely their boundaries so long as the integrated optical depth is relatively low, less than three. In addition, information gleaned from its two polarization channels can infer the shape, size and phase of the particles within its field of view. However, such procedures have yet to be automated and visually interpreting the lidar data is inherently subjective and can be difficult in situations of multi-layer clouds or super-cooled mixed phase clouds.

Future studies would ideally employ an additional ground-based instrument to measure down-welling infrared radiation so that the LIRAD method can be used to obtain a much more precise measure of optical depth. This would help to define, more accurately, the detection limits of the satellite algorithms. Also, a site with a more homogeneous surface cover would help minimize the amount of error introduced by surface mixing. Isolating the signature of a translucent cirrus cloud is difficult over a

relatively large spatial area, containing a mixture of both rural and urban land cover types. It is not clear how much of a negative impact, if any, this may have had on the outcomes of this study. Considering all the complexities associated with measuring primarily thin cirrus clouds over a mixed sub-arctic environment, the overall performance of the MODIS cloud products, evaluated here, appear to have met the basic expectations outlined in the theoretical basis documents.

References

Ackerman, S., K. Strabala, P. Menzel, R. Frey, C. Moeller, L. Gumley, B. Baum, S. W. Seeman and H. Zhang, 2002: Discriminating clear-sky from cloud with MODIS: algorithm theoretical basis document (MOD35), version 4.0. *MODIS ATBD*, 112 pp.

Albrecht, B. A., 1989: Aerosols, cloud microphysics, and fractional cloudiness. *Science*, **245**, 1227-1230.

Atmospheric soundings, 2008: University of Wyoming, Department of Atmospheric Sciences, <http://weather.uwyo.edu/upperair/sounding.html>, last referenced, March 2008.

Barnes, W. L., T. S. Pagano and V. V. Salomonson, 1998: Prelaunch characteristics of the Moderate Resolution Imaging Spectroradiometer (MODIS) on EOS-AM1. *IEEE Transactions on Geoscience and Remote Sensing*, **36**, 1088-1100.

Baum, B. A. and B. A. Wielicki, 1994: Cirrus cloud retrieval using infrared sounding data: multilevel cloud errors. *Journal of Applied Meteorology*, **33**, 107-117.

Baum, B. A., P. F. Soulen, K. I. Strabala, M. D. King, S. A. Ackerman, P. Menzel and P. Yang, 2000: Remote sensing of cloud properties using MODIS airborne simulator imagery during SUCCESS 2. Cloud thermodynamic phase. *Journal of Geophysical Research*, **105**, 11,781-11,792.

Baum, B. A., R. A. Frey, G. G. Mace, M. K. Harkey and P. Yang, 2003: Nighttime multilayered cloud detection using MODIS and ARM date. *Journal of Applied Meteorology*, **42**, 905-919.

Berendes, T. A., D. A. Berendes, R. M. Welch, E. G. Dutton, T. Uttal and E. E. Clothiaux, 2004: Cloud cover comparisons of the MODIS daytime cloud mask with surface instruments at the North Slope of Alaska ARM site. *IEEE Transactions on Geoscience and Remote Sensing*, **42**, 2584-2593.

Bréon, F. M., D. Tanré and S. Generoso, 2002: Aerosol effect on cloud droplet size monitored from satellite. *Science*, **295**, 834-838.

Bowtie effect, 2008: Dundee satellite receiving station MODIS frequently asked question, <http://www.sat.dundee.ac.uk/modis-faq.html>, last referenced, March 2008.

Campbell, J. B., 2002: *Introduction to Remote Sensing, Third Edition*. The Guilford Press, 621 pp.

Cess, R. D., G. L. Potter, J. P. Blanchet, G. J. Boer, S. J. Ghan et al., 1989: Interpretation of cloud-climate feedback as produced by 14 atmospheric general circulation models. *Science*, **245**, 513-516.

Charlson, R. J., J. E. Lovelock, M. O. Andreae and S. G. Warren, 1987: Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *Nature*, **326**, 655-661.

Chou, M., K. Lee and P. Yang, 2002: Parameterization of shortwave cloud optical properties for a mixture of ice particle habits for use in atmospheric models. *Journal of Geophysical Research*, **107**(D21), 4600, doi:10.1029/2002JD002061, 2002.

Chylek, P., S. Robinson, M. K. Dubey, M. D. King, Q. Fu and W. B. Clodius, 2006: Comparison of near-infrared and thermal infrared cloud phase detections. *Journal of Geophysical Research*, **111**, D20203, doi:10.1029/2006JD007140, 2006.

Coakley, J. A., Jr., R. L. Bernstein and P. A. Durkee, 1987: Effect of ship-stack effluents on cloud reflectivity. *Science*, **237**, 1020-1022.

Collins, M. J., 2007: Advanced Remote Sensing: The Sensor Radiance Equation (Chapter 5 Notes for ENGO 435). University of Calgary, 17 pp. <http:// www.geomatics.ucalgary.ca/~collins/Courses/engo655/engo655_index.html>, last referenced, March 2008.

Comstock, J. M. and K. Sassen, 2001: Retrieval of cirrus cloud radiative and backscattering properties using combined lidar and infrared radiometer (LIRAD) measurements. *Journal of Atmospheric and Oceanic Technology*, **18**, 1658-1673.

Cox, S. K., 1971: Cirrus clouds and the climate. *Journal of the Atmospheric Sciences*, **28**, 1513-1515.

DeMott, P. J., K. Sassen, M. R. Poellot, D. Baumgardner, D. C. Rogers, S. D. Brooks, A. J. Prenni and S. M. Kreidenweis, 2003: African dust aerosols as atmospheric ice nuclei, *Geophysical Research Letters*, **30**, NO. 14, 1732, doi:10.1029/2003GL017410, 2003.

Dozier, J., 1990: Introduction to Environmental Optics: Applications in Physical Geography and Remote Sensing (Notes for Geography 102/202A). University of California, Santa Barbara, 67 pp.

Eloranta, E. W. and T. Uttal, 2006: Arctic observations with the University of Wisconsin high spectral resolution lidar. *12th AMS Conf. on Atmospheric Radiation*, 9-14 July 2006, Madison, WI.

Frey, R. A., B. A. Baum, W. P. Menzel, S. A. Ackerman, C. C. Moeller and J. D. Spinhirne, 1999: A comparison of cloud top heights computed from airborne lidar and MAS radiance data using CO_2 slicing. *Journal of Geophysical Research*, **104**, 24,547-24,555.

Fu, Q. and K. N. Liou, 1993: Parameterization of the radiative properties of cirrus clouds. *Journal of the Atmospheric Sciences*, **50**, 2008-2025

Gao, B., W. Han, S. Tsay and N. F. Larsen, 1998: Cloud detection over the Arctic region using airborne imaging spectrometer data during the daytime. *Journal of Applied Meteorology*, **37**, 1421-1429.

Groisman, P. Y., R. S. Bradley and B. Sun, 2000: The relationship of cloud cover to near-surface temperature and humidity: comparison of GCM simulations with empirical data. *Journal of Climate*, **13**, 1858-1878.

HDF file naming system, 2008: National Aeronautics and Space Administration, <http://modis-atmos.gsfc.nasa.gov/MOD06_L2/filename.html>, last referenced, March 2008.

Holz, R. E., S. Nasiri, R. E. Kuehn, R. Frey, B. Baum and E. W. Eloranta, 2001: High Spectral Resolution Lidar Validation of MODIS Derived Cloud Phase and Altitude. *11th AMS Conf. on Satellite Meteorology and Oceanography*, 14-18 October, 2001, Madison, WI.

Hsu, N. C., J. R. Herman and S. Tsay, 2003: Radiative impacts from biomass burning in the presence of clouds during boreal spring in southeast Asia. *Geophysical Research Letters*, **30**, NO. 5, 1224, doi:10.1029/2002GL016485, 2003.

HYSPLIT Model, 2008: National Oceanic and Atmospheric Administration Air Resources Laboratory, <http://www.arl.noaa.gov/ready/hysplit4.html>, last referenced, March 2008.

IPCC, 2007: Summary for Policymakers in *Climate Change 2007: The Physical 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.

Jayaweera, K. O. L. F., G. Wendler and T. Ohtake, 1975: Low cloud cover and the winter temperature of Fairbanks. *Climate of the Arctic*, *Proceedings*, 316-322.

Jensen, E. J. and O. B. Toon, 1992: The potential effects of volcanic aerosols on cirrus cloud microphysics. *Geophysical Research Letters*, **19**, 1759-1762.

Kaufman, Y. J. and D. Tanré, 1998: Algorithm for remote sensing of tropospheric aerosol from MODIS. *MODIS ATBD*, 85 pp.

Kerr, R. A., 1989: How to fix the clouds in greenhouse models. *Science*, **243**, 28-29.

King, M. D., S. Tsay, S. E. Platnick, M. Wang and K. N. Liou, 1997: Cloud retrieval algorithms for MODIS: optical thickness, effective particle radius, and thermodynamic phase, version 5. *MODIS ATBD*, 79 pp.

King, M. D., S. Platnick, C. C. Moeller, H. E. Revercomb and D. A. Chu, 2003: Remote sensing of smoke, land, and clouds from the NASA ER-2 during SAFARI 2000. *Journal of Geophysical Research*, **108**(D13), 8502, doi:10.1029/2002JD003207, 2003.

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. *Journal of Atmospheric and Oceanic Technology*, **21,** 857-875.

Koren, I., Y. J. Kaufman, L. A. Remer and J. V. Martins, 2004: Measurement of the effect of Amazon smoke on inhibition of cloud formation. *Science*, **303,** 1342-1345.

Kuang, Z. and Y. L. Yung, 2000: Reflectivity variations off the Peru coast: evidence forindirect effect of anthropogenic sulfate aerosols on clouds. *Geophysial Research Letters*, **27**, 2501-2504.

LAADS data product and image search information, 2008a: National Aeronautics and Space Administration, <http://ladsweb.nascom.nasa.gov/index.html>, last referenced, March 2008.

LAADS level-1 and atmosphere search query, 2008b: National Aeronautics and Space Administration, <http://ladsweb.nascom.nasa.gov/data/search.html>, last referenced, March 2008.

Lee, J., P. Yang, A. E. Dessler, B. A. Baum and S. Platnick, 2006: The influence of thermodynamic phase on the retrieval of mixed-phase cloud microphysical and optical properties in the visible and near-infrared region. *IEEE Geoscience and Remote Sensing Letters*, **3**, 287-291.

Lhermitte, R., 2002: *Centimeter and Millimeter Wavelength Radars in Meteorology*, Lhermitte Publications, 550 pp.

Li, J., W. P. Menzel, Z. Yang, R. A. Frey and S. A. Ackerman, 2003: High-spatial-resolution surface and cloud-type classification from MODIS multispectral band measurements. *Journal of Applied Meteorology*, **42**, 204-226

Lillesand, T. M. and R. W. Kiefer, 1994: *Remote Sensing and Image Interpretation, Third Edition*. John Wiley & Sons, Inc, 750 pp.

Liou, K. N., 2002: *An Introduction to Atmospheric Radiation, Second Edition*. Academic Press, An Imprint of Elsevier Science, 583 pp.

Liu, Y., J. R. Key, R. A. Frey, S. A. Ackerman and W. P. Menzel, 2004: Nighttime polar cloud detection with MODIS. *Remote Sensing of the Environment*, **92**, 181-194.

Lohmann, U., B. Kärcher and C. Timmreck, 2003: Impact of the Mount Pinatubo eruption on cirrus clouds formed by homogeneous freezing in the ECHAM4 GCM. *Journal of Geophysical Research*, **108**(D18), 4568, doi10.1029/ 2002JD003185, 2003.

Lyons, W. A., T. E. Nelson, E. R. Williams, J. A. Cramer and T. R. Turner, 1998: Enhanced positive cloud-to-ground lightning in thunderstorms ingesting smoke from fires. *Science*, **282**, 77-80.

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. *Journal of Applied Meteorology*, **44**, 221-240.

Mahesh, A., M. A. Gray, S. P. Palm, W. D. Hart and J. D. Spinhirne, 2004: Passive and active detection of clouds: comparison between MODIS and GLAS observations. *Geophysical Research Letters*, **31**, L04108, doi:10.1029/2003GL018859, 2004.

Menzel, W. P., B. A. Baum, K. I. Strabala and R. A. Frey, 2002: Cloud top properties and cloud phase algorithm theoretical basis document, version 6. *MODIS ATBD*, 60 pp.

Menzel, W. P., R. A. Frey, B. A. Baum and H. Zhang, 2006: Cloud top properties and cloud phase algorithm theoretical basis document, version 7. *MODIS ATBD*, 55 pp.

Mishchenko, M. I., W. B. Rossow, A. Macke and A. A. Lacis, 1996: Sensitivity of cirrus cloud albedo, bidirectional reflectance and optical thickness retrieval accuracy to ice particle shape. *Journal of Geophysical Research*, **101**, 16,973-16,985.

MODTRAN 4 software description and download information, 2008: Kirtland Air Force Base, <http://www.kirtland.af.mil/library/factsheets/factsheet.asp?id=7915>, last referenced, March 2008.

MODIS orbit specifications and spectral bands, 2008a: National Aeronautics and Space Administration, <http://modis.gsfc.nasa.gov/about/specifications.php>, last referenced, March 2008.

MODIS instrument design, 2008b: National Aeronautics and Space Administration, <http://modis.gsfc.nasa.gov/about>, last referenced, March 2008.

MODIS atmosphere scientific dataset variable and attribute list by product, 2008c: National Aeronautics and Space Administration, <http://modis-atmos.gsfc.nasa.gov/ specs/>, last referenced, March 2008.

MODIS EOS data product levels, 2008d: National Aeronautics and Space Administration, <http://ecsinfo.gsfc.nasa.gov/sec3/ProductLevels.html>, last referenced, March 2008.

MODIS geophysical data products, 2008e: National Aeronautics and Space Administration, <http://modis.gsfc.nasa.gov/data/dataprod/index.php>, last referenced, March 2008.

MODIS C005 atmosphere product updates, 2008f: National Aeronautic and Space Administration, <http://modis-atmos.gsfc.gov/products_C005update.html >, last referenced, March 2008.

MODIS cloud mask byte field interpretations, 2008g: National Aeronautic and Space Administration, <http://modis-atmos.gsfc.nasa.gov/MOD35_L2/format. html>, last referenced, March 2008.

MODIS cloud phase algorithm flow charts, 2008h: National Aeronautic and Space Administration, <http://modis-atmos.gsfc.nasa.gov/MOD06_L2/atbd.html>, last referenced, March 2008.

Myneni, R. B., S. Hoffman, Y. Knyazikhin, J. L. Privette, J. Glassy, Y. Tian, Y. Wang, X. Song, Y. Zhang, G. R. Smith, A. Lotsch, M. Friedl, J. T. Morisette, P. Votava, R. R. Nemani and S. W. Running, 2002: Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of Environment*, **83**, 214-231.

Naud, C. M., J. P. Muller, E. E. Clothiaux, B. A. Baum and W. P. Menzel, 2005: Intercomparison of multiple years of MODIS, MISR and radar cloud-top heights. *Annales Geophysicae*, **23**, 2415-2424.

Naval Research Laboratory/ Monterey aerosol page, 2008: Naval Research Laboratory, <http:// nrlmry.navy.mil/aerosol/>, last referenced, March 2008.

Noel, V., H. Chepfer, G. Ledanois, A. Delaval and P. H. Flamant, 2002: Classification of particle shape ratios in cirrus clouds based on the lidar depolarization ratio, *Applied Optics*, **41**, 4245-4257.

Noel, V., D. M. Winker, M. McGill and P. Lawson, 2004: Classification of particle shapes from lidar depolarization ratio in convective ice clouds compared to in situ observations during CRYSTAL-FACE. *Journal of Geophysical Research*, **109**, D24213, doi:10.1029/2004JD004883, 2004.

Pal, S. R., W. Steinbrecht and A. I. Carswell, 1992: Automated method for lidar determination of cloud-base height and vertical extent. *Applied Optics*, **31**, 1488-1494.

Platnick, S., M. D. King, S. A. Ackerman, W. P. Menzel, B. A. Baum, J. C. Riédi, and R. A. Frey, 2003: The MODIS cloud products: algorithms and examples from Terra. *IEEE Transactions on Geoscience and Remote Sensing*, **41,** 459-473.

Pruppacher, H. R. and J. D. Klett, 1997: *Microphysics of Clouds and Precipitation, Second Revised and Enlarged Edition with an Introduction to Cloud Chemistry and Cloud Electricity*. Kluwer Academic Publishers, 954 pp.

Raatz, W. and G. E. Shaw, 1984: Long range transport of pollution aerosols into the Alaskan Artic. *Journal of Climate and Applied Meteorology*, **23**, 1052-1064.

Ramanathan, V., R. D. Cess, E. F. Harrison, P. Minnis, B. R. Barkstrom, E. Ahmad and D. Hartmann, 1989: Cloud-radiative forcing and climate: results from the earth radiation budget experiment. *Science*, **243**, 57-63.

Remer, L. A., D. Tanré, Y. J. Kaufman, R. Levy and S. Mattoo, 2006: Algorithm for Remote sensing of tropospheric aerosol from MODIS: collection 005. *MODIS ATBD*, 88 pp.

Rogers, R. R. and M. K. Yau, 1989: *A Short Course in Cloud Physics, Third Edition*. Butterworth-Heinemann, An Imprint of Elsevier, 290 pp.

Rosenfeld, D., 2000: Suppression of rain and snow by urban and industrial air pollution. *Science*, **287**, 1793-1796.

Roskovensky, J. K. and K. N. Liou, 2005: Differentiating airborne dust from cirrus clouds using MODIS data. *Geophysical Research Letters*, **32**, L12809, doi:10.1029/2005GL022798, 2005

Rozell, N., 2006: 2005 fire season again oddly large. *Alaska Science Forum*, Article #1784, 3 pp. <http://www.gi.alaska.edu/ScienceForum/ASF17/1784.html>, last referenced, March 2008.

Sassen, K., 1984: Deep orographic cloud structure and composition derived from comprehensive remote sensing measurements. *Journal of Climate and Applied Meteorology*, **23**, 568-583.

Sassen, K., M. K. Griffin and G. C. Dodd, 1989: Optical scattering and microphysical properties of subvisual cirrus clouds, and climatic implications. *Journal of Applied Meteorology*, **28**, 91-98.

Sassen, K., 1991: The polarization lidar technique for cloud research: a review and current assessment. *Bulletin of the American Meteorological Society*, **72**, 1848-1866.

Sassen, K. and B. S. Cho, 1992: Subvisual-thin cirrus lidar dataset for satellite verification and climatological research. *Journal of Applied Meteorology*, **31**, 1275-1284.

Sassen, K., D. O'C. Starr, G. G. Mace, M. R. Poellot, S. H. Melfi, W. L. Eberhard, J. D. Spinhirne, E. W. Eloranta, D. E. Hagen and J. Hallet, 1995: The 5-6 December 1991 FIRE IFO II jet stream cirrus case study: possible influences of volcanic aerosols. *Journal of the Atmospheric Sciences*, **52**, 97-123.

Sassen, K. and S. Benson, 2001: A midlatitude cirrus cloud climatology from the facility for atmospheric remote sensing. Part II: microphysical properties derived from lidar depolarization. *Journal of the Atmospheric Sciences*, **58**, 2103-2112.

Sassen, K., 2002a: Cirrus: a modern perspective in *Cirrus*. D. Lynch et al., Eds., Oxford University Press, 11-40.

Sassen, K., 2002b: Indirect climate forcing over the western US from Asian dust storms. *Geophysical Research Letters*, **29**, NO.10, 10.1029/2001GL014051, 2002.

Sassen, K. and G. G. Mace, 2002: Ground-based remote sensing of cirrus clouds in *Cirrus*. D. Lynch et al., Eds., Oxford University Press, 168-195.

Sassen, K., P. J. DeMott, J. M. Prospero and M. R. Poellot, 2003: Saharan dust storms and indirect aerosol effects on clouds: CRYSTAL-FACE results. *Geophysical Research Letters*, **30**, NO.12, 1633, doi:10.1029/2003GL017371, 2003.

Sassen, K., 2005: Polarization in Lidar in *Lidar: range-resolved optical remote sensing of the atmosphere*. Claus Weitkamp et al., Springer Series in Optical Sciences, 19-40.

Sassen, K., J. R. Campbell, J. Zhu, P. Kollias, M. Shupe and C. Williams, 2005: Lidar and triple-wavelength doppler radar measurements of the melting layer: a revised model for dark- and brightband phenomena. *Journal of Applied Meteorology*, **44**, 301-311.

Sassen, K., and V. Khvorostyanov, 2006: Indirect cloud effects from Alaskan smoke: Evidence for ice formation below water saturation. *12th AMS Conf. on Cloud Physics*, 9-14 July 2006, Madison, WI.

Sassen, K., J. Zhu, P. W. Webley, K. Dean and P. Cobb, 2007: Volcanic ash plume identification using polarization lidar: Augustine eruption, Alaska. *GeophysicalResearch Letters*, **34,** L08803, doi: 10.1029/2006GL027237,2007.

Sassen, K. and J. Zhu, 2008a: Arctic Facility for Atmospheric Remote Sensing instrument datasets, <http://rainbow.gi.alaska.edu>, last referenced, March 2008.

Sassen, K. and J. Zhu, 2008b: Arctic Facility for Atmospheric Remote Sensing information and instrument design specifications, <http://www.gi.alaska.edu/ AtmosSci/Research/afars_site.html>, last referenced, March 2008.

Schott, J. R., 1997: *Remote Sensing, The Image Chain Approach*. Oxford University Press, 394 pp.

Senior, C. A. and J. F. B. Mitchell, 1993: Carbon dioxide and climate: the impact of cloud parameterization. *Journal of Climate*, **6**, 393-418.

Smith, W. L. and C. M. R. Platt, 1978: Comparison of satellite-deduced cloud heights with indications from radiosonde and ground-based laser measurements. *Journal of Applied Meteorology*, **17**, 1796-1802.

Smith, W. L., H. Revercomb, H. Huang, D. H. DeSlover, W. Feltz, L. Gumley and A. Collard, 1998: Infrared spectral absorption of nearly invisible cirrus clouds. *Geophysical Research Letters*, **25**, 1137-1140.

Spangenberg, D. A., P. Minnis, M. D. Shupe and M. R. Poellot, 2006: Retrieval of cloud phase over the Arctic using MODIS 6.7-12 μm data. *14th AMS Conf. on Satellite Meteorology and Oceanography*, 29 January – 2 February 2006, Atlanta GA.

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

Strabala, K. I., S. A. Ackerman and W. P. Menzel, 1994: Cloud properties inferred from 8-12-μm data. *Journal of Applied Meteorology*, **33**, 212-229.

Tiruchirapalli, R., 2006: Micro-pulse lidar observations of aerosols over the atmospheric radiation measurement site at Barrow. Masters Thesis, University of Alaska, Fairbanks,109 pp.

Twomey, S., 1977: The influence of pollution on the shortwave albedo of clouds. *Journal of the Atmospheric Sciences*, **34**, 1149-1152.

Wallace, J. M. and P. V. Hobbs, 1977: *Atmospheric Science, An Introductory Survey*. Academic Press, An Elsevier Science Imprint, 467 pp.

Wandinger, U., 2005: Introduction to Lidar in *Lidar: range-resolved optical remote sensing of the atmosphere*. Claus Weitkamp et al., Springer Series in Optical Sciences, 1-18.

Wang, Z. and K. Sassen, 2001: Cloud type and macrophysical property retrieval using multiple remote sensors. *Journal of Applied Meteorology*, **40**, 1665-1682.

Wendler, G. and P. Nicpon, 1975: Low-level temperature inversions in Fairbanks, central Alaska. *Monthly Weather Review*, **103**, 34-44.

Zhang, H. and W. P. Menzel, 2002: Improvement in thin cirrus retrievals using and emissivity-adjusted CO_2 slicing algorithm. *Journal of Geophysical Research*, **107**, (D17), 4327, doi:10.1029/2001JD001037, 2002.

Zhang, M., S. Klein, D. Randall, R. Cederwall and A. Del Genio, 2005a: Introduction to special section on toward reducing cloud-climate feedback uncertainties in atmospheric general circulation models. *Journal of Geophysical Research*, **110**, D15S01, doi:10.1029/2005JD005923.

Zhang, M. H., W. Y. Lin, S. A. Klein, J. T. Bacmeister, S. Bony, R. T. Cederwall, A. D. Del Genio, J. J. Hack, N. G. Loeb, U. Lohmann, P. Minnis, I. Musat, R. Pincus, P. Stier, M. J. Saurez, M. J. Webb, J. B. Wu, S. C. Xie, M.-S. Yao and J. H. Zhang, 2005b: Comparing clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements. *Journal of Geophysical Research*, **110**, D15S02, doi:10.1029/2004JD005021.