Study on optical properties of dust particles using Mie-scattering lidar: estimation of lidar ratios and improvement of cloud masking processes

(ミー散乱ライダーを用いたダスト粒子の光学特性に関する研究：
ライダー比の推定と雲マスクの改良)

Yoshitaka JIN

(神　慶孝)

A dissertation for the degree of Doctor of Science
Department of Earth and Environmental Sciences,
Graduate School of Environmental Studies, Nagoya University

(名古屋大学大学院環境学研究科地球環境科学専攻学位論文 博士 (理学))

2013

論文要旨

大気中に浮遊するダスト粒子は、太陽光の散乱と吸収を通じて直接的に、雲との相互作用を通じて間接的に、地球の放射収支に影響を与える。ダストの気候への影響を正しく評価するためには、光学特性 (消散係数など) の鉛直分布の情報が必要不可欠である。受動型センサによるダストの衛星観測では、光学特性の鉛直分布の情報は得られない。衛星 CALIPSO に搭載されたライダー (CALIOP) によって、2006 年 6 月から全球の雲とエアロゾルの鉛直分布を観測できるようなった。CALIOP の観測データから光学特性を導出するには、ライダー比 (消散係数と後方散乱係数の比) が必要となる。CALIOP などのミー散乱ライダーでは、ライダー比は未知数となる。そのため、CALIOP で検出された層に対して、エアロゾル光学モデルから得られたライダー比を適切に割り当てなければならない。この時、検出された層が雲かエアロゾルなのかを正しく判別する必要がある。雲とエアロゾルの誤判別は、ライダー比の誤った選択につながるため、光学特性を正確に推定できなくなる。

CALIOP データの雲を特定する雲マスクでは、ダストを雲と誤判別する場合がある。そのため、雲マスクによって特定された雲について、誤判別雲を識別する必要がある。また、光学特性を正確に推定するために、光学モデルから得られたライダー比を地上ライダー等で検証する必要がある。タクラマカン砂漠は主要なダストの発生源地域であるにも関わらず、これまでライダー比を測定した研究事例は無かった。本論文では、ミー散乱ライダーデータからダストの正確な光学特性の鉛直分布を推定するために、1) タクラマカン砂漠におけるダストのライダー比を地上・衛星ライダーの同期観測から推定すること、および、2) CALIOP の雲マスクに含まれる誤判別雲を判別分析によって識別することを目的とした。

タクラマカン砂漠におけるライダー比を推定するため、地上／衛星ライダーによるダストの同期観測が 2009 年 3 月に行われた。この時に同期測定された地上と衛星ライダーの信号から後方散乱係数を計算し、両者の差が最も小さくなる時のライダー比を推定値とした。推定されたライダー比は波長 532 nm に対して 42.0 sr、波長 1064 nm に対して 45.9 sr だった。532 nm のライダー比は、他のアジア域で測定された値よりも 3 ~ 24% 小さかった。ライダー比の推定誤差は、532 nm に対して 9.5%、1064 nm に対して 41.6% だった。エラー解析の結果、1064 nm のライダー比はライダー信号の校正の誤差に強く依存することがわかった。推定したライダー比を使って後方散乱係数を計算し、CALIOP level 2 プロダクトと比較した。その結果、

CALIOP level 2 の後方散乱係数は本研究による結果と比べて 21% 小さいことがわかった。この違いは、CALIOP level 2 の vertical feature mask (VFM) でダストを雲と誤判別し、ライダー比の選択が誤っていたことが原因であると示唆された。

CALIOP の雲マスクに含まれる、誤判別された雲を識別するために判別分析を行った。判別モデル (線形判別関数) を構築するために、雲と誤判別雲の学習データを決定した。この時、CloudSat および MODIS の雲マスクと相対湿度を用いた。構築した判別モデルを使って、CALIOP の雲マスクで検出された雲の再分類を行った。その結果、昼間と夜間、そして陸上と海上でも、誤判別雲の識別に成功した。一方、従来の判別モデルは昼間の学習データしか使用していないので、夜間の識別がうまくいかなかった。本論文の判別モデルは、91.7% の検出精度で誤判別雲を識別できた。雲マスクに含まれる誤判別雲の割合は北半球 (高度 1–7 km, 20°W–120°E, 0°–50°N) の夏季 (JJA) で 6% だった。誤判別雲の頻度はタクラマカン砂漠の上空で最も高かった。タクラマカン砂漠におけるダストのプロファイルのうち、34.6% が雲マスクで誤判別されたプロファイルだった。雲マスク改良後のダスト消散係数のプロファイルは改良前に比べて最大 2.6 倍大きかった。また、本論文で得られた消散係数は CALIOP level 3 データと比べて最大 2 倍大きかった。先行研究との比較の結果、CALIOP level 3 データの消散係数が負のバイアスをもつことがわかった。この負のバイアスは、VFM がダストを雲と誤判別していることが原因であると示唆された。

本論文の成果は、ミー散乱ライダーデータからダストの光学特性を正確に推定することに役立つ。また、データ同化などを通じて化学輸送モデルの精度向上への貢献が期待される。

Abstract

Mineral dust suspended in the atmosphere affects the earth's radiation budget directly through scattering and absorbing sunlight and indirectly through their interaction with clouds. To accurately predict the effect of dust on the climate system, information regarding its vertical distribution is needed. However, vertical information of dust cannot be derived from observations by most passive sensors or sampling instruments. The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument on the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite has enabled global observation of the vertical distribution of aerosols and clouds since June 2006. Lidar ratios (extinction-to-backscattering ratio) are needed to retrieve optical properties (e.g., extinction coefficient) from CALIOP signals, and are unknown in the analysis of Mie-scattering lidar as CALIOP.

If lidar ratios are not derived from observations, those from particle optical models are allocated on the basis of layer types (i.e., clouds, aerosols, and their sub-types). CALIOP cloud mask products, which determine whether the observed data are clouds or not, occasionally misclassify dense dust as clouds. Since the misclassification leads to the selection of incorrect lidar ratios of dust, the CALIOP cloud mask need to be improved. In addition, lidar ratios obtained by optical models need to be validated by ground-based lidar. Although the Taklimakan Desert is a major sources of the dust, no lidar ratio measurements have been conducted previously.

This thesis is designed for better estimation of the vertical profiles of dust optical properties by Mie-scattering lidar. The purposes of this thesis are: 1) to estimate the lidar ratio of dust originated from the Taklimakan Desert by simultaneous observations with ground-based lidar (Aksu-lidar) and CALIOP, and 2) to discriminate misclassified clouds from clouds in the CALIOP cloud mask with a discriminant analysis.

Simultaneous observations of dust by Aksu-lidar and CALIOP were conducted to estimate the lidar ratio in the Taklimakan Desert on March 23, 2009. The lidar ratio is estimated from a performance function that is defined as the square error of the backscattering coefficients between Aksu-lidar and CALIOP. If the performance function returns the minimum value, the selected lidar ratio is regarded as the optimal solution. The estimated lidar ratios at 532 and 1064 nm are 42.0 and 45.9 sr, respectively. The 532 nm lidar ratio is 3–24% smaller than those of Asian dust reported previously. The errors in the lidar ratio estimation are 9.52% at 532 nm and 41.6% at 1064 nm. The error analysis indicates that the estimation of the 1064 nm lidar ratio depended strongly on the calibration errors at 1064 nm. A comparison of our results to the CALIOP level 2 data shows that the backscattering coefficients of the CALIOP level 2 data are 21% smaller than those of this thesis

and that the smaller backscattering coefficients are caused by misclassification of dust as clouds in the vertical feature mask (VFM).

A discriminant analysis is conducted to detect misclassified clouds in the CALIOP cloud mask. A linear discriminant function (LDF) is used as the discrimination model. The training data are collected through tests with the CloudSat cloud mask, the Moderate Resolution Imaging Spectroradiometer (MODIS) cloud mask, and the relative humidity. Discrimination of dust from clouds is successful in cases over both land and water surfaces during the daytime and nighttime. In contrast, the discrimination model of previous studies was inadequate during the nighttime since training data were not collected during the nighttime.

The accuracy rate of the LDF classification is 91.7% for misclassified clouds. The estimated content rate of misclassified clouds in the CALIOP cloud mask is approximately 6% in the region of 20°W–120°E and 0°–50°N during June–August 2007. The cloud mask is most frequently misclassified in the Taklimakan Desert. The proportion of misclassified clouds to observed dust is ~34.6% (below 2 km) in the desert. Dust extinction coefficient estimated using the improved cloud mask is ~2.6 times larger than that estimated using the original cloud mask. A comparison of this thesis to the CALIOP level 3 products indicates that the extinction profile with the improved cloud mask is ~2 times larger than that of the CALIOP level 3. Aerosol optical depths (AOD) reported by previous studies indicate that AODs for the CALIOP level 3 have a negative bias. This thesis suggests that the negative bias of the CALIOP level 3 in the Taklimakan Desert is mainly caused by misclassification of dust as clouds in the VFM.

The results of this thesis are useful for correctly estimating dust optical properties from Mie-scattering lidar, and contribute to an improvement in the accuracy of chemical transport models.

Contents

1 Introduction

1.1 Aerosol particles in the global climate system and their observations by lidar instruments

Atmospheric aerosol particles (hereinafter referred to as aerosols) play an important role in the earth's climate system. Aerosols cool the atmosphere by scattering sunlight into space and heat the atmosphere through the absorption of solar and terrestrial radiation. Aerosols indirectly affect the earth's radiation budget through modifications of cloud microphysics (e.g., effective radius and number concentrations) by acting as cloud condensation and ice nuclei. According to the fifth assessment report (AR5) of the Intergovernmental Panel on Climate Change (IPCC), radiative forcing caused by aerosols is totally negative (i.e., cooling of the earth). However, the radiative forcing of tropospheric aerosols is not well understood due to aerosol characteristics as follows: the aerosols have short residence times (from a few days to a few weeks), different emissions from one region to another, and a variety of chemical compositions and size distributions (depending on their species). These characteristics lead to an inadequate understanding of aerosol global distribution [*Seinfeld and Pandis*, 1998]. Therefore, observational studies on tropospheric aerosols are necessary to improve our knowledge in the aerosol characteristics.

Mineral dust is an aerosol that has a large impact on scattering and absorption of solar radiation because of its high concentration [*Li et al.*, 1996]. *Tegen et al.* [1997] reported that the optical depth of dust accounts for 21–48% of the global aerosol optical depth (AOD). The radiative forcing (direct effect) of dust is estimated at -0.1 [± 0.2] W/m^2 in the AR5. Dust is emitted to the atmosphere by wind erosion in desert and arid regions [*Shao*, 2008]. The main source regions are located in the northern hemisphere and extend from the west coast of North Africa, the Middle East, Central and South Asia, to China [*Prospero et al.*, 2002]. It has been reported that dust emissions are from 1000 to 2150 Tg/yr since 2001 [*Zender et al.*, 2004]. Compared to other regions, the mixed layer height is higher in arid regions since clear skies and the low soil moisture contents allow higher surface temperatures, which can cause stronger sensible-heat fluxes to the atmosphere and deeper turbulent mixing [*Warner*, 2004]. As a result, dust layer heights can develop to altitudes of 6 km in great deserts [*Liu et al.*, 2008a]. Dust modifies radiative heating rates in the atmosphere depending on its vertical distribution [*Meloni et al.*, 2005; *Perrone et al.*, 2012], thereby changing atmospheric stability [*Zhao et al.*, 2004]. If dust reaches heights where the temperature is below 0°C, they can serve as ice nuclei. Several studies suggested that dust has the largest impact on cloud phases at

temperatures warmer than −40°C [*Wiacek and Peter*, 2009; *Wiacek et al.*, 2010; *Choi et al.*, 2010]. Dust is transported over long distances by large-scale winds. Dust layers emitted from arid regions in China have been observed in Korea, Japan, and even North American continent beyond the Pacific Ocean [*Uno et al.*, 2004; *Husar et al.*, 2001]. Dust layers from the Sahara Desert are transported to the Mediterranean Sea, the North Atlantic Ocean, and the Caribbean Sea [*Moulin et al.*, 1998; *Prospero and Carlson*, 1972]. Through these dust transport processes, optical properties of the atmosphere can be changed by mixing with anthropogenic particles [*Shibata and Yang*, 2010] and/or by acting as ice nuclei [*Sakai et al.*, 2003]. Thus, observations of dust optical properties are important to understand the effects of dust on the climate system. This thesis focuses on dust optical properties in the northern hemisphere.

To accurately predict the effects of dust on the climate system, information on its vertical distribution is essential [*Satheesh and Moorthy*, 2005]. However, vertical information cannot be derived from observations by most passive sensors (e.g., sun photometers) and sampling instruments (e.g., cascade impactors). For example, widely used satellite imagery has allowed the regional and global distribution of aerosols, but it cannot provide the vertical information. Lack of aerosol vertical information causes uncertainty for the solution of the radiative transfer equation. In contrast, LIght Detection and Ranging (lidar) is an active remote sensing instrument that emits laser pulses into the atmosphere and receives backscattered light from aerosols, clouds, and/or atmospheric molecules [*Weitkamp*, 2005]. Time intervals between emitting a laser pulse and receiving backscattered light depend on the distance from lidar to scatterers. Analyzing backscattered echo signals enable us to investigate the distance to scatterers. If laser pulses are emitted vertically, lidar can measure vertical distribution of scatterers. Therefore, lidar is an effective tool for the climate research, and this thesis uses it for the investigation of dust vertical distribution.

The Taklimakan Desert is one of the largest sources of Asian dust. It is located in the Tarim Basin, which is surrounded by high mountains exceeding 4000 m in altitude. Local circulations driven by the mountains can cause floating dust [*Kim et al.*, 2009]. Dust over the Taklimakan Desert floats in the atmosphere throughout the year due to the particular geography [*Liu et al.*, 2008b]. In addition, *Uno et al.* [2009] reported that dust originated from the Taklimakan Desert is transported over a circuit around the globe. Taklimakan dust can have significant effects on the climate in terms of its long residence time. However, lidar observations in the Taklimakan Desert were not conducted before 2001. By the Aeolian Dust Experiment on Climate impact (ADEC) project [*Mikami et al.*, 2006], ground-based lidar was set up at Aksu (hereinafter referred to as Aksu-lidar) located at the northern part of the desert. The project has carried out intensive

observations since the autumn of 2001 [*Tsunematsu et al.*, 2005a,b; *Tanaka et al.*, 2005; *Kai et al.*, 2008; *Kim and Kai*, 2007]. In this thesis, dust optical properties over the Taklimakan Desert are investigated by Aksu-lidar observations.

There have been few lidar observations in most arid regions due to accessibility and power limitations. However, Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) on board the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) spacecraft has enabled us to observe the global vertical distribution of aerosols and clouds since June 2006 [*Winker et al.*, 2007]. For example, *Liu et al.* [2008b] investigated the vertical profiles of dust over the Tibetan Plateau where it had been difficult to conduct lidar observations. Space-borne lidar can observe aerosols globally, but their signals have large amounts of noise and no temporal continuity at any fixed point. Conversely, ground-based lidar can measure high-quality (low noise) signals that have temporal continuity at fixed points. By making use of the advantages of each lidar instrument, scientific understanding on the role of aerosols in the climate system can be improved.

1.2 Importance of lidar ratios for Mie-scattering lidar

Aksu-lidar and CALIOP are Mie-scattering lidar systems. This type of lidar detects light backscattered at the same wavelength as the emitted laser. Mie-scattering lidar receives backscattered light from atmospheric particles (aerosols and clouds) and molecules simultaneously. Compared to High Spectral Resolution Lidar (HSRL), which receives backscattered light from atmospheric particles and molecules separately, Mie-scattering lidar has a simple optical system, and therefore it is widely used. For example, the National Institute for Environmental Studies (NIES) has used Mie-scattering lidar for a lidar network to observe Asian dust and air pollution aerosols over East Asia [*Sugimoto et al.*, 2011].

Mie-scattering lidar systems used in this thesis (specifically, Aksu-lidar and CALIOP) are equipped with a dual-wavelength laser (1064 nm and 532 nm). For each wavelength, the backscattering or extinction coefficients of observed particles can be retrieved from lidar signals. Analyzing the ratio of backscattering coefficients for aerosols at these two wavelengths (referred to as color ratio), information related to the aerosol size distribution is derived [*Sasano and Browell*, 1989]. In addition, the lidar systems measure the polarization components of backscattered light at 532 nm. The ratio of polarization signals (referred to as depolarization ratio) gives us a parameter for nonspherical particles (or particle size distribution with a mode radius smaller than 532 nm). The depolarization ratio, in turn, allows us to identify dust from other aerosol species in the lidar data empirically [*Sassen*, 2000]. In summary, optical properties derived by the lidar systems are backscattering or extinction coefficients, color ratios, and depolarization ratios.

To retrieve backscattering or extinction coefficients from signals measured by Mie-scattering lidar, an inverse problem must be solved. The lidar equation that describes signals measured by lidar includes backscattering and extinction coefficients (see Chapter 2.2). These are theoretically calculated for molecular components, but unknowns for particle components. The lidar equation cannot be solved since it contains two unknowns. To solve the equation, lidar ratios (extinction-to-backscattering ratios) should be introduced to reduce an unknown in the equation [*Fernald et al.*, 1972]. The lidar ratio ranges from 10 to 90 sr and depends on the size distribution, complex refractive indices and shapes of particles, the wavelength of incident light, and relative humidity [*Anderson et al.*, 2000].

The reported lidar ratios for dust over East Asia are 42-55 sr at 532 nm [*Liu et al.*, 2002; *Sakai et al.*, 2003; *Murayama et al.*, 2004; *Noh et al.*, 2007]. If incorrect lidar ratios are used, lidar signal attenuations are inaccurately corrected. Such inaccurate attenuation correction causes errors in the calculation of particle backscattering or extinction coefficients [*Sasano et al.*, 1985]. Therefore, it

is important to estimate correct lidar ratios. In contrast, there is no need to assume lidar ratios for HSRL and Raman scattering lidar systems since they can directly observe lidar ratios by measuring particle backscattering and particle extinction coefficients independently. However, much more studies have been conducted with Mie-scattering lidar, rather than those with the other kind of lidar systems. Therefore, the estimation of lidar ratios has significance for studies with Mie-scattering lidar.

This thesis attempts to estimate the lidar ratio of dust over the Taklimakan Desert by using Aksu-lidar and CALIOP signals, and is the first attempt to estimate the lidar ratio over this desert. The estimation can be achieved by assuming that Aksu-lidar and CALIOP observe the same atmosphere column (see Chapter 3.1). An intensive observation of dust by Aksu-lidar was conducted simultaneously with CALIOP in March 2009. Several studies have made simultaneous observations of aerosols by ground-based lidar and CALIOP. *Kim et al.* [2008] indicated that the extinction profiles of daytime CALIOP data differ substantially from those of ground-based lidar data due to the small signal-to-noise ratio (SNR) of CALIOP data. They derived lidar ratios from sun photometer data. *Hara et al.* [2011] investigated the temporal variations of the AOD of spherical particles in East Asia by using CALIOP and ground-based lidar data. In that study, the lidar ratio was set to 50 sr to calculate the AOD. However, the theoretical lidar ratio of polluted aerosols, which are the main constituents of spherical particles in East Asia, is approximately 70 sr [*Omar et al.*, 2009]. Unknown lidar ratios result in uncertainty in the calculation of optical properties. There have been few studies estimating lidar ratios from simultaneous observations.

1.3 Misclassification issues in cloud masking processes for the CALIPSO lidar

As explained above, it is necessary to select correct lidar ratios in order to accurately retrieve backscattering and extinction coefficients from lidar data. However, lidar ratios are rarely estimated by measurements, especially for CALIOP. In such cases, lidar ratios obtained from optical models incorporating the microphysics of particles (i.e., the size distribution, complex refractive index, and shape of particles) are used [e.g. *Omar et al.*, 2009]. To assume the microphysics of particles, lidar-observed layers must be classified into aerosols or clouds. The backscattering or extinction coefficients of aerosols are calculated after screening cloud layers and determining lidar ratios. There are two products for cloud identification (i.e., cloud masking) in CALIOP data: the vertical feature mask (VFM) developed by the CALIPSO science team [*Vaughan et al.*, 2009; *Liu et al.*, 2009] and the C2 cloud mask developed by *Hagihara et al.* [2010]. The latter is based on cloud masks used in ship-borne lidar and 95 GHz cloud radar observations in the western Pacific Ocean near Japan [*Okamoto et al.*, 2007]. The name "C2" means the cloud mask scheme for lidar-only ("C1" is for radar-only). These two cloud mask schemes (VFM and C2) are different in terms of the layer detection and discrimination of aerosols and clouds. This thesis uses the C2 cloud mask because of the better cloud detection than the VFM as described below.

The VFM discriminates between clouds and aerosols at five horizontal resolutions (333 m, 1 km, 5 km, 20 km, and 80 km) and provides the product mixed with all the resolutions. However, the mixture state of the resolutions makes analysis complicated. For example, researchers using cloud-resolving models are forced to estimate the distribution of cloud fraction on some horizontal scales that are larger than the reported resolution of the VFM for comparison. In addition, the VFM layer detection scheme makes false detections through its large horizontal averaging procedure and may misclassify noise or aerosols as clouds [*Marchand et al.*, 2008]. Dust layers are often misclassified as cloud layers since the optical properties of dust are similar to those of clouds. In the Taklimakan Desert during springtime, about 43% of dust layers are misclassified as cloud layers in the VFM [*Chen et al.*, 2010]. If such misclassified layers are used for the retrieval, the estimated optical properties of aerosols and clouds will have uncertainties due to the selection of incorrect lidar ratios.

The C2 cloud mask has a fixed horizontal resolution of 1.1 km. This scheme identifies clouds in regions with strong signals and spatial continuity, while the VFM attempts to detect clouds even for weak signal regions by the large horizontal average. In general, clouds have larger backscatter-

ing cross sections than aerosols. Clouds detected by weak signals may contain aerosols in the VFM scheme. Therefore, the C2 cloud mask produces less misclassification and identifies clouds more certainly. The zonal mean cloud fraction obtained from the VFM is 25% larger than those from the C2 cloud mask in low altitudes (below 2 km) due to the misclassification issues of the VFM [*Hagihara et al.*, 2010]. The C2 cloud mask was used for evaluating cloud microphysics from the Nonhydrostatic Icosahedral Atmospheric Model (NICAM) [*Hashino et al.*, 2013]. In addition, the C2 cloud mask scheme is planned to be used as a basis of a cloud mask for lidar that will be on board the Earth Clouds, Aerosols and Radiation Explorer (EarthCARE) satellite [*JAXA_EORC*, 2012]. However, aerosol layers with strong signals (e.g., dense dust layers) are occasionally misclassified as cloud layers in the C2 cloud mask. As a result, further improvements are needed to make a conclusive cloud mask product that has little misclassification.

Aerosols misclassified as clouds should be removed from cloud mask data for the reasons mentioned above. The vertical profiles of dust must be reevaluated after cloud mask improvements. Several studies attempted to solve the VFM misclassification issues using discrimination analysis, which is a method to predict the group of uncategorized data by a discriminant model (discriminant function) [*Johnson and Wichern*, 2007]. The model is constructed from training data consisting of two or more independent variables that the groups they belong to are already known. Several studies discriminated misclassified clouds (mainly dust layers) from VFM cloud data by incorporating lidar derived optical and geometrical variables in their discriminant models [*Xie et al.*, 2010; *Chen et al.*, 2010; *Naeger et al.*, 2013]. However, these studies only applied the discriminant analysis to the daytime data since the training data were manually determined by reference to satellite imagery, resulting in difficulties in the detection of misclassified clouds for the enormous CALIOP dataset.

In this thesis, cloud layers detected by the C2 cloud mask are reclassified into clouds or misclassified clouds by using discriminant analysis. The cloud masks from Moderate Resolution Imaging Spectroradiometer (MODIS) on board Aqua and cloud profiling radar (CPR) on board CloudSat are used to determine the training data used for the discriminant analysis (see Chapter 4.2). The application of the MODIS cloud mask to the discriminant analysis is the first approach.

1.4 Purposes of this thesis

This thesis is designed for better estimation of dust optical properties derived from Mie-scattering lidar. The purposes of this thesis are: 1) to estimate the lidar ratio of dust originated from the Taklimakan Desert by simultaneous observations with Aksu-lidar and CALIOP, and 2) to discriminate misclassified clouds from CALIOP cloud mask data by discriminant analysis. The first purpose is expected to lead to more accurate estimations of the optical properties of Taklimakan dust in the spring season. Using the estimated lidar ratio, the vertical profiles of dust optical properties will be retrieved from Aksu-lidar data. The retrieved backscattering coefficient will be compared to CALIOP level 2 products. The products are necessary to be validated since they are used in many studies, such as for comparison with numerical models [e.g. *Koffi et al.*, 2012]. The second purpose is expected to lead to improvements of the C2 cloud mask with little misclassification through the removal of dust, which is considered as the main source of misclassified clouds. The vertical profiles of dust extinction coefficients will be reevaluated after the cloud mask improvement. The studies in this thesis for solving the issues of lidar ratio selection and layer misclassification are expected to be applicable for better estimation of the vertical profiles of dust optical properties.

2 Description of the lidar systems and data analyzing procedures

In Chapter 1, the importance of lidar ratios in the analysis of Mie-scattering lidar data and the issues of lidar ratio selection caused by cloud mask misclassifications were pointed out. This chapter presents basic information on the Mie-scattering lidar used in this thesis for subsequent studies. First, the ground-based lidar system in the Taklimakan Desert and the CALIOP lidar system are described. Second, the retrieval procedures of optical properties from the lidar equation are described. The retrieval error caused by selecting an incorrect lidar ratio is shown.

2.1 Systems of ground/space-based lidar instruments

This section presents a description of the Aksu-lidar and CALIOP systems, which are used to investigate the optical properties of dust. An observation of dust by Aksu-lidar was conducted at the Aksu Water Balance Experimental Station of the Xinjiang Institute of Ecology and Geography of the Chinese Academy of Science from March 23 to 25 in 2009. The station is located in the northern part of the Taklimakan Desert (40.62°N, 80.83°E, 1028 m above sea level), as shown in Figure 2.1. The observational data are used to estimate the dust lidar ratio by comparison with the simultaneous CALIOP observation data.

A schematic diagram of the Aksu-lidar system is illustrated in Figure 2.2 and the specifications of that system are summarized in Table 2.1. The lidar consists of transmitter, receiver, and data processing systems. The transmitter system employs a Nd:YAG laser at the fundamental (1064 nm) and the second (532 nm) harmonic wavelengths. A linearly polarized and pulsed laser beam is collimated at 0.2 mrad by a beam expander. The pulse repetition rate is 10 Hz, and the energy per pulse is 300 mJ. The laser pulses are emitted into the atmosphere in the vertical direction. The laser axis is set to be coincident with the field of view of the receiver's telescope. The receiver system employs a Schmidt-Cassegrain telescope with a 355-mm diameter. The backscattered light collected by the telescope is separated into four optical paths through an optical system. The incident light is separated by a dichroic mirror that reflects the 532 nm wavelength light and transmits the 1064 nm wavelength light. Each component of the separated light passes through a narrow-band interference filter to transmit the light of the laser wavelength (i.e., the other wavelengths are excluded). The reflected light (532 nm) is divided into perpendicular and parallel components to the polarized plane of the emitted laser by a polarizer. The parallel component is further divided

into 90% and 10% strengths by a beam splitter. The amount of each divided light is controlled by a neutral density (ND) filter to reduce the background light. Since stronger lidar signals are measured at the lower altitudes, the 10% and the 90% components are used for the lower and the higher altitudes, respectively. Using the two components allows the dynamic range of observable signal strength to be extended.

All components of the separated light are detected by three photomultiplier tubes (PMTs) for 532 nm, and by an avalanche photodiode (APD) for 1064 nm. To avoid signal-induced noise caused by strong signals at lower altitudes, the detector that receives the 90% component is gated from the lidar up to 3 km (i.e., 20 μs) by a delay generator. The detected light is converted to an electric signal. A transient recorder that is a part of the data processing system digitizes the converted signal. All signals are digitized by a 12 bit A/D converter, and the 532 nm signals are also digitized by a discriminator that detects voltage pulses above a selected threshold (i.e., a photon counting system with a maximum count rate of 250 MHz). The transient recorder records 16000 signals per shot (i.e., 120 km height) at 50 ns interval (i.e., 7.5 m vertical resolution). To improve the SNR, the signals are averaged for every 3000 shots (5 minutes time resolution). The processed signals are stored by a computer.

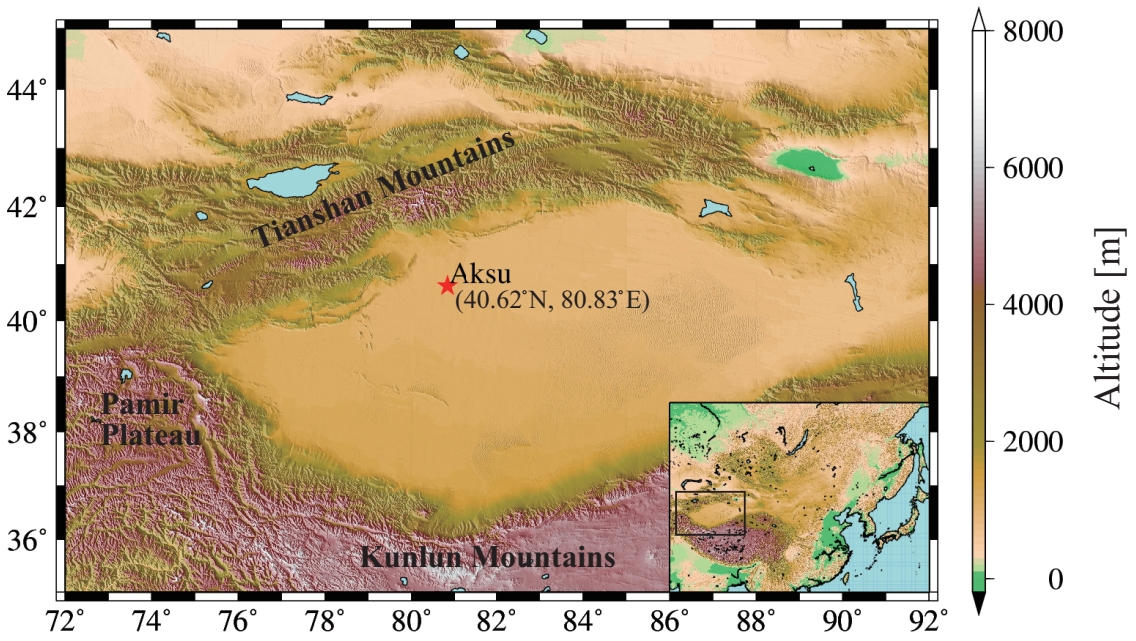

Figure 2.1: Location of the Taklimakan Desert and the observation site (Aksu). The Tarim Basin is surrounded by the high mountains above 5 km, which are the Tianshan Mountains in north, the Pamir Plateau in west and the Kunlun Mountains in south.

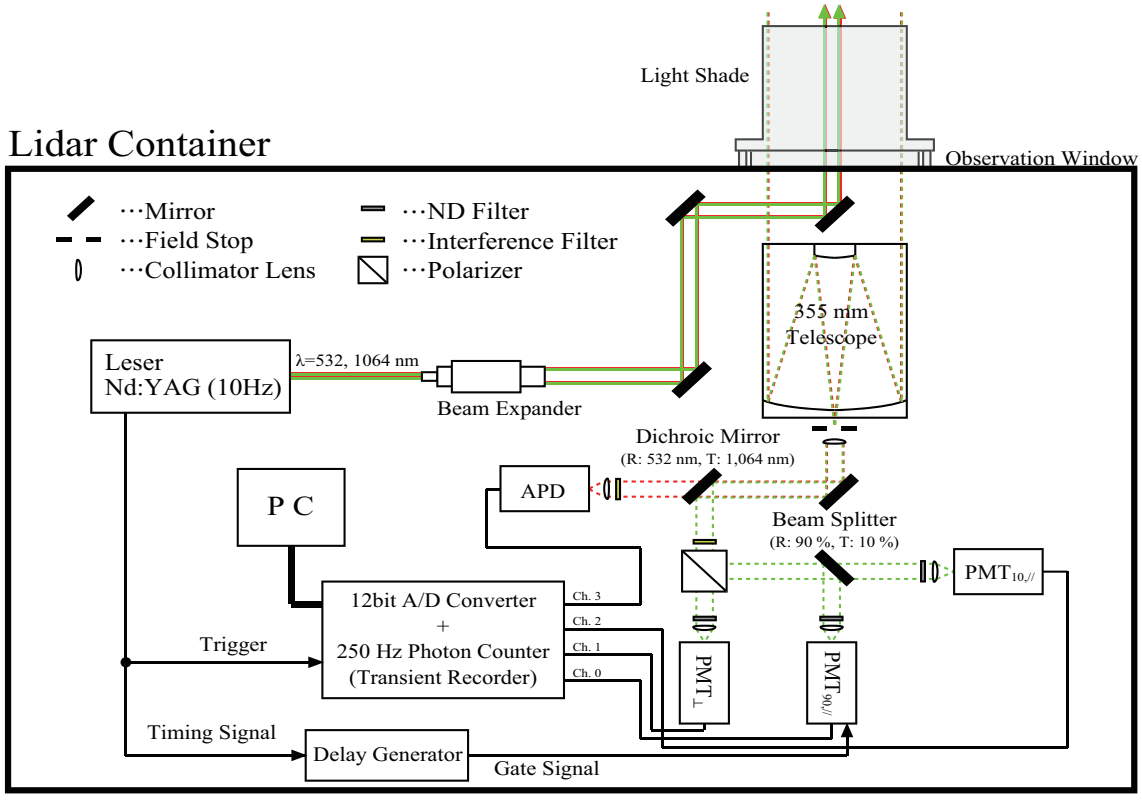

Figure 2.2: Schematic diagram of the Aksu-lidar system.

Table 2.1: Specification of the Aksu-lidar system.

Transmitter	
Laser	Nd:YAG Laser
Wavelength	532 nm, 1064 nm
Pulse Energy	300 mJ
Pulse Repetition Rate	10 Hz
Beam Divergence	0.2 mrad (after beam expander)
Receiver	
Telescope	Schmidt-Cassegrain
Diameter of Telescope	355 mm
Direction of Measurement	Zenith
Geometry	Coaxial
Field of View	2 mrad
Digitization	50 ns
532 nm channel:	
Detector	Photomultiplier (HAMAMATSU R3234)
Voltage	2000 V
$PMT_{90,//}$ Delay	20 μs
ND Filter	
$PMT_{\perp}$	1%
$PMT_{//}$	10%
1064 nm channel:	
Detector	Avalanche Photodiode (Licel APD-3.0)
Voltage	340 V

The CALIPSO satellite, on which the CALIOP is mounted, was launched in April 2006 as a part of the so-called A-train satellite constellation. The satellites in the constellation are positioned at an altitude of 705 km in a sun-synchronized polar orbit with an inclination of 98°. They travel 14.56 orbits per day and move 24.7° longitude westward per orbit at the equator. The CALIPSO mission is to improve our understanding of the role of aerosols and clouds in the climate system [*Winker et al.*, 2007]. In this thesis, the CALIOP data are used for the estimation of lidar ratios and investigations of the optical properties of dust. To accurately estimate the optical properties, misclassification in the CALIOP cloud mask is improved.

The CALIOP lidar measures backscattered signals at 1064 and 532 nm and polarization component at 532 nm in common with Aksu-lidar. The specification of the CALIOP system is summarized in Table 2.2. The laser pulse is collimated at 0.1 mrad by a beam expander, and is emitted to the atmosphere with 3 degree off-nadir angle (0.3 degree before November 28, 2007). The repetition rate of the laser pulse is 20.16 Hz, and the energy per pulse is 110 mJ. The diameter and the field of view of the telescope are 1 m and 0.13 mrad, respectively. To eliminate the background noise of 532 nm signals, an etalon filter with a narrow-band is used in combination with the interference filter. The CALIOP level 1B data (version 3), which are calibrated raw data, are used for the investigation of the optical properties. The CALIOP data are downloaded from Atmospheric Science Data Center (ASDC: https://eosweb.larc.nasa.gov). The horizontal and vertical resolutions of the level 1B data depend on the altitude and are summarized in Table 2.3. In the data analysis of simultaneous observations with Aksu-lidar, the vertical resolution of 532 nm signals below 8.2 km is changed into 60 m to reduce random noise. The vertical resolution of Aksu-lidar is adjusted in accordance with the CALIOP resolution.

Table 2.2: Specification of the CALIOP lidar system.

Transmitter	
Laser	Nd:YAG Laser
Wavelength	532 nm, 1064 nm
Pulse Energy	110 mJ
Pulse Repetition Rate	20.16 Hz
Beam Divergence	0.1 mrad (after beam expander)
Receiver	
Diameter of Telescope	1 meter
Direction of Measurement	off-nadir (3 degree)
Geometry	Coaxial
Field of View	0.13 mrad
Digitization	100 ns
532 nm channel:	
Detector	Photomultiplier
Etalon Passband	37 pm
Etalon Peak Transmission	85 %
Blocking Filter	770 pm
1064 nm channel:	
Detector	Avalanche Photodiode
Optical Passband	450 pm
Peak Transmission	84 %

Table 2.3: Horizontal and vertical resolutions of the CALIOP level 1B data.

Altitude Range [km]	Horizontal Resolution [km]	532 nm Vertial Resolution [m]	1064 nm Vertical Resolution [m]
30.1 to 40.0	5.0	300	—
20.2 to 30.1	1.67	180	180
8.2 to 20.2	1.0	60	60
-0.5 to 8.2	0.33	30	60
-2.0 to -0.5	0.33	300	300

2.2 Lidar equation

A signal derived from the Mie-scattering lidar is described by the lidar equation as follows:

$$P_\lambda(z) = \frac{K}{z^2} O(z)\beta_\lambda(z)T_\lambda^2(z), \tag{2.1}$$

where $P_\lambda(z)$ is the observed signal at an altitude z and a wavelength λ, K is the system constant depending on the lidar system (e.g., laser power), $O(z)$ is the beam overlap factor, $\beta_\lambda(z)$ is the backscattering coefficient of scatterers at λ, and $T_\lambda^2(z)$ is the atmospheric transmittance at λ between the lidar altitude z_0 and z, and is given by:

$$T_\lambda^2 = \exp\left\{-2\int_{z_0}^{z} \alpha_\lambda(z')dz'\right\}, \tag{2.2}$$

where $\alpha_\lambda(z)$ is the extinction coefficient of scatterers at λ. This term is based on the Lambert-Beer's law, and represents the laser attenuation by scatterers. The coefficient 2 denotes the two-way transmission path. The backscattering and extinction coefficients are further written as:

$$\beta_\lambda(z) = \beta_{1,\lambda}(z) + \beta_{2,\lambda}, \tag{2.3}$$

and

$$\alpha_\lambda(z) = \alpha_{1,\lambda}(z) + \alpha_{2,\lambda}, \tag{2.4}$$

where the subscripts 1 and 2 denote particles and the atmospheric molecules, respectively.

The backscattered light from the atmospheric molecules is due to Rayleigh scattering since the molecular size is much smaller than the laser wavelengths. The extinction cross section of the atmospheric molecules is theoretically calculated from the molecular number density of the standard atmosphere at the sea level (i.e., 15°C and 1 atm). The molecular number density is derived from U.S. Standard Atmosphere (1976). The extinction cross section, $\sigma_{2,\lambda}^e$, is written as:

$$\sigma_{2,\lambda}^e = \frac{24\pi^3\left(n_\lambda^2 - 1\right)^2}{\lambda^4 N^2\left(n_\lambda^2 + 2\right)^2}\left(\frac{3 + 6\delta_{2,\lambda}}{3 - 4\delta_{2,\lambda}}\right), \tag{2.5}$$

where $\delta_{2,\lambda}$ is the depolarization ratio for the atmospheric molecules, and the values of 1.441% at 532 nm and 1.400% at 1064 nm are used [*Collins and Russell*, 1976], and n and N are the refractive index and the molecular number density of the standard atmosphere. The refractive indices used in this thesis are 1.0002782 at 532 nm and 1.0002740 at 1064 nm [*Bucholtz*, 1995]. The molecular number density of the standard atmosphere is 2.54743×10^{25} /m^3. The extinction coefficient of the atmospheric molecules is calculated from vertical profiles of temperature. The profiles are derived from European Re-Analysis (ERA) interim data of the European Center for Medium-Range Weather Forecasts (ECMWF) [*Simmons et al.*, 2007]. The extinction coefficient at z is calculated by the following equation:

$$\alpha_{2,\lambda}(z) = N(z)\sigma^{e}_{2,\lambda} = \frac{N_A P(z)}{RT(z)}\sigma^{e}_{2,\lambda}, \tag{2.6}$$

where $N(z)$ is the molecular number density at z, N_A is the Avogadro's constant ($N_A = 6.02214 \times 10^{23}$ /mol), $P(z)$ is pressure at z, R is the gas constant ($R = 8.314472$ J/K·mol), and $T(z)$ is temperature at z.

The backscattering coefficient is calculated by the following equation:

$$\beta_{2,\lambda}(z) = \frac{\alpha_{2,\lambda}(z)}{S_2}, \tag{2.7}$$

where S_2 is the extinction-to-backscattering ratio of the atmospheric molecules and is approximated by $8\pi/3$ [*Collins and Russell*, 1976].

On the other hand, the backscattered light from particles is due to Mie scattering since the particle size is comparable to the laser wavelength. The backscattering and extinction coefficients of particles cannot be theoretically calculated since the vertical distribution of the number density, size distribution, complex refractive index, and shape of particles are unknown. The backscattering and extinction coefficients of particles are expressed as follows:

$$\beta_{1,\lambda}(z) = \int_0^\infty \frac{dn(r,z)}{dr}\sigma^{bk}_{1,\lambda}(r,z)dr, \tag{2.8}$$

and

$$\alpha_{1,\lambda}(z) = \int_0^\infty \frac{dn(r,z)}{dr} \sigma^e_{1,\lambda}(r,z) dr, \tag{2.9}$$

where $\sigma^{bk}_{1,\lambda}(r,z)$ and $\sigma^{bk}_{1,\lambda}(r,z)$ are the backscattering and extinction cross sections of particles, respectively, and $\frac{dn(r,z)}{dr}$ is the particle size distribution function. The cross sections are unknown parameter estimated by the size, complex refractive index, and shape of particles and the wavelength of incident light.

2.3 Retrieval of optical properties from lidar signals

As mentioned earlier, the backscattering and extinction coefficients of particles are unknown parameters. The lidar equation cannot be solved since there are two unknown parameters in the equation. Lidar ratios (S_1) are assumed to reduce the unknown parameters [*Fernald et al.*, 1972], and are expressed by:

$$S_{1,\lambda}(z) = \frac{\alpha_{1,\lambda}(z)}{\beta_{1,\lambda}(z)}. \tag{2.10}$$

This thesis assumes that lidar ratios are constant with altitude. In other words, the size distribution, complex refractive index, and shape of particles are not changing with altitude, and the variation of backscattering coefficients depends solely on the number density.

Fernald [1984] solved the lidar equation for backscattering coefficients in a simple scheme as follows:

$$\beta_\lambda(z_{i+1}) = \frac{X_\lambda(z_{i+1}) \exp\left[-A(z_i, z_{i+1})\right]}{\frac{X_\lambda(z_i)}{\beta_\lambda(z_i)} - S_{1,\lambda}\left\{X_\lambda(z_i) + X_\lambda(z_{i+1}) \exp\left[-A(z_i, z_{i+1})\right]\right\}\Delta z}, \tag{2.11}$$

where $A(z_i, z_{i+1}) = (S_{1,\lambda} - S_{2,\lambda})\left[\beta_\lambda(z_i) + \beta_\lambda(z_{i+1})\right]\Delta z$, and $X_\lambda(z_{i+1})$ is the range corrected signal ($= P_\lambda(z_{i+1}) z_{i+1}^2$). The backscattering coefficient and range corrected signal at a reference altitude z_i are used to calculate the backscattering coefficient at one step forward altitude z_{i+1} from the lidar. To calculate the backscattering coefficient at the next step altitude z_{i+2}, the backscattering coefficient and range corrected signal at z_{i+1} are used as those at the reference altitude. This scheme is designed for a forward inversion in which the backscattering coefficient is calculated from the near side to far side of the lidar.

The other scheme, the backward inversion is expressed as:

$$\beta_\lambda(z_{i-1}) = \frac{X_\lambda(z_{i-1}) \exp\left[+A(z_{i-1}, z_i)\right]}{\frac{X_\lambda(z_i)}{\beta_\lambda(z_i)} + S_{1,\lambda}\left\{X_\lambda(z_i) + X_\lambda(z_{i-1}) \exp\left[+A(z_{i-1}, z_i)\right]\right\}\Delta z}, \tag{2.12}$$

where $A(z_{i-1}, z_i) = (S_{1,\lambda} - S_{2,\lambda})\left[\beta_\lambda(z_{i-1}) + \beta_\lambda(z_i)\right]\Delta z$. In the backward inversion, the backscattering coefficient is calculated from the far side to near side of the lidar.

The backscattering coefficient at the reference altitude is necessary for the initial condition of the calculation both in the inversion methods. There are few aerosols above 30 km. The down-

ward calculation (i.e., the direction from space toward the earth) is started from 30 km, where the backscattering coefficient can be assumed to equal to the molecular backscattering coefficient. The downward calculation corresponds to the forward and backward inversions for CALIOP and Aksu-lidar, respectively. Aksu-lidar cannot observe the atmosphere at 30 km due to the laser attenuation and limited dynamic range. Tropopause regions generally have low concentration of aerosols, so the backscattering coefficient of particles is much smaller than that of the molecules. In this thesis, the backward inversion is started from 13 km that is referred to as a calibration altitude $z_{c,532}$ of Aksu-lidar. The calibration altitude $z_{c,1064}$ is set to 6 km since Aksu-lidar cannot detect 1064 nm signals above 6 km. When simultaneous observations are conducted, the particle backscattering coefficient at $z_{c,\lambda}$, which is an initial value of the inversion, is derived from the CALIOP data. If the particle backscattering coefficient at $z_{c,\lambda}$ cannot be derived, the forward inversion method should be applied to the retrieval. The reference height of the forward inversion corresponds to the nearest altitude from the lidar. From Equation (2.1), the backscattering coefficient at the nearest altitude z_1 is derived as follows:

$$\beta_\lambda(z_1) = \frac{X_\lambda(z_1)O(z_1)}{KT_\lambda^2(z_1)}, \tag{2.13}$$

where $O(z_1)$ is analytically estimated by the method of *Dho et al.* [1997], and $T_\lambda^2(z_1)$ is assumed to be 1.0 because of the small laser attenuation at the height. The system constant K is the unknown parameter in Equation (2.13). It can be estimated from the backscattering coefficient that is calculated by the backward inversion method in another time. Substituting the estimated system constant into Equation (2.13), the forward calculation can be started from the nearest height from the lidar.

The effect of incorrect lidar ratios on the retrieved optical properties is examined. A certain layer with the lidar ratio of 50 sr and a particle backscattering coefficient of 0.02 /km/sr is considered (i.e., the corresponding particle extinction coefficient is 1.0 /km). The attenuated backscattering coefficient $\beta'_\lambda(z)$ is calculated by the following equation:

$$\beta'_\lambda(z) = \frac{P_\lambda(z)z^2}{KO(z)} = \beta_\lambda(z)T_\lambda^2(z). \tag{2.14}$$

The attenuated backscattering coefficient is used for the inversion and corresponds to the range corrected signals in Equation (2.11) and (2.12). The backscattering coefficient is calculated by the

backward and forward inversion methods, selecting incorrect lidar ratios (30 and 70 sr). Figure 2.3 shows the attenuated backscattering coefficient and backscattering coefficients calculated by using the three lidar ratios (30, 50, and 70 sr). The starting point of the calculation for the backward and forward inversions is the farthest and nearest heights, respectively. As the range from the starting point increases, the backscattering coefficients by the incorrect lidar ratios diverge from the backscattering coefficient by the true lidar ratio (50 sr). Selecting the larger lidar ratio (70 sr) for the backward inversion results in the overestimate of the backscattering coefficients, whereas selecting the smaller lidar ratio (30 sr) results in the underestimate. In contrast to the backward inversion, the larger and smaller lidar ratios for the forward inversion cause the underestimate and overestimate, respectively. In addition, the results of the forward inversion have the larger margin of errors than those of the backward inversion.

The particle backscattering and particle extinction coefficients calculated by the inversion methods are integrated from the starting point to a certain height. The integrated values of the coefficients are compared to the true values, referring to an error analysis shown in the NIES website (http://www-lidar.nies.go.jp/AnalysisMethods/DataAnalysis1.html). The integrated extinction coefficient is hereafter referred to as an optical depth (τ). Here, the true particle backscattering coefficient of the layer is newly set to 0.1 /km/sr to examine the larger optical depth. Figure 2.4 shows relative errors of the optical depth and integrated backscattering coefficient (IBC) as a function of the true optical depth calculated by the true lidar ratio. The relative errors of the IBCs in the backward inversion increase with the true optical depth. The relative errors of the optical depth converge on the true optical depth at the larger values of the true optical depth. The relative errors are about $\pm 40\%$ at $\tau = 0.02$, but these are within $\pm 5\%$ at $\tau = 5$. In the forward inversion, the relative errors of the IBCs increase with the optical depth. The relative errors of the larger lidar ratio are overestimated, whereas those of the smaller lidar ratio are underestimated. The errors of the optical depth diverge rapidly with the true optical depth. The relative error of the smaller lidar ratio is about -100% at $\tau = 5$. The relative error of the larger lidar ratio exceeds 100% at $\tau = 0.4 \sim 0.5$, and it declines rapidly at the larger values of the true optical depth. This decrease indicates that the retrieved extinction and backscatter coefficients have negative values. Consequently, the forward inversion method has more possibilities to cause errors of the IBCs and the optical depth than the backward inversion method, especially in the case of optically thick layers. This is attributed to the fact that there is a negative sign in front of $S_{1,\lambda}$ in the denominator of Equation (2.11), so that the solution of the forward inversion method is unstable. In contrast, there is a positive sign in front of $S_{1,\lambda}$ in the denominator of Equation (2.12). This simulation experiment shows that it is important

to select the correct lidar ratio for reducing the retrieval errors, and the forward inversion method has the higher sensitivity to the lidar ratio compared to the backward inversion method.

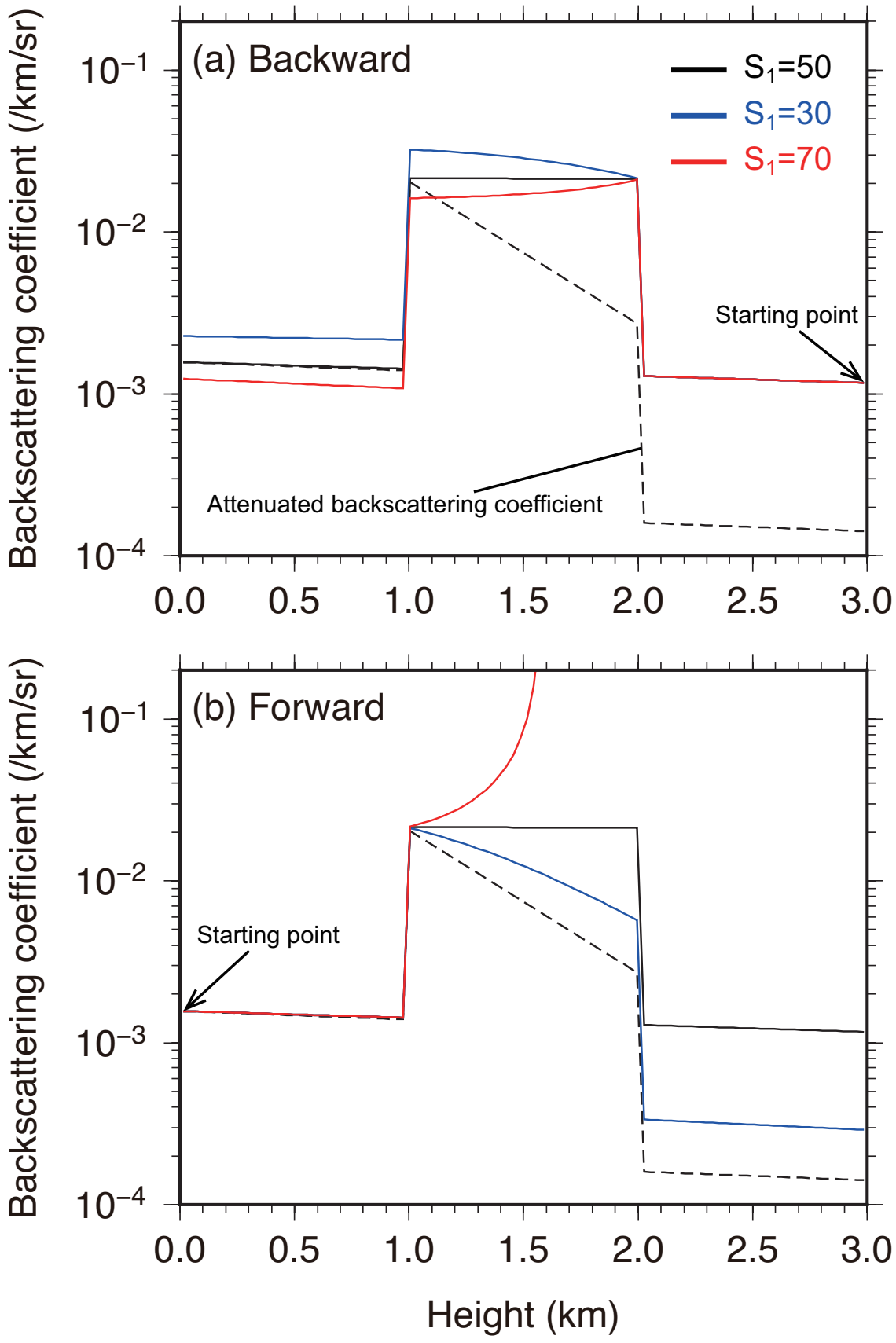

Figure 2.3: A demonstration of the calculation of (a) backward and (b) forward inversions with three lidar ratios of 30, 50(true), 70 sr.

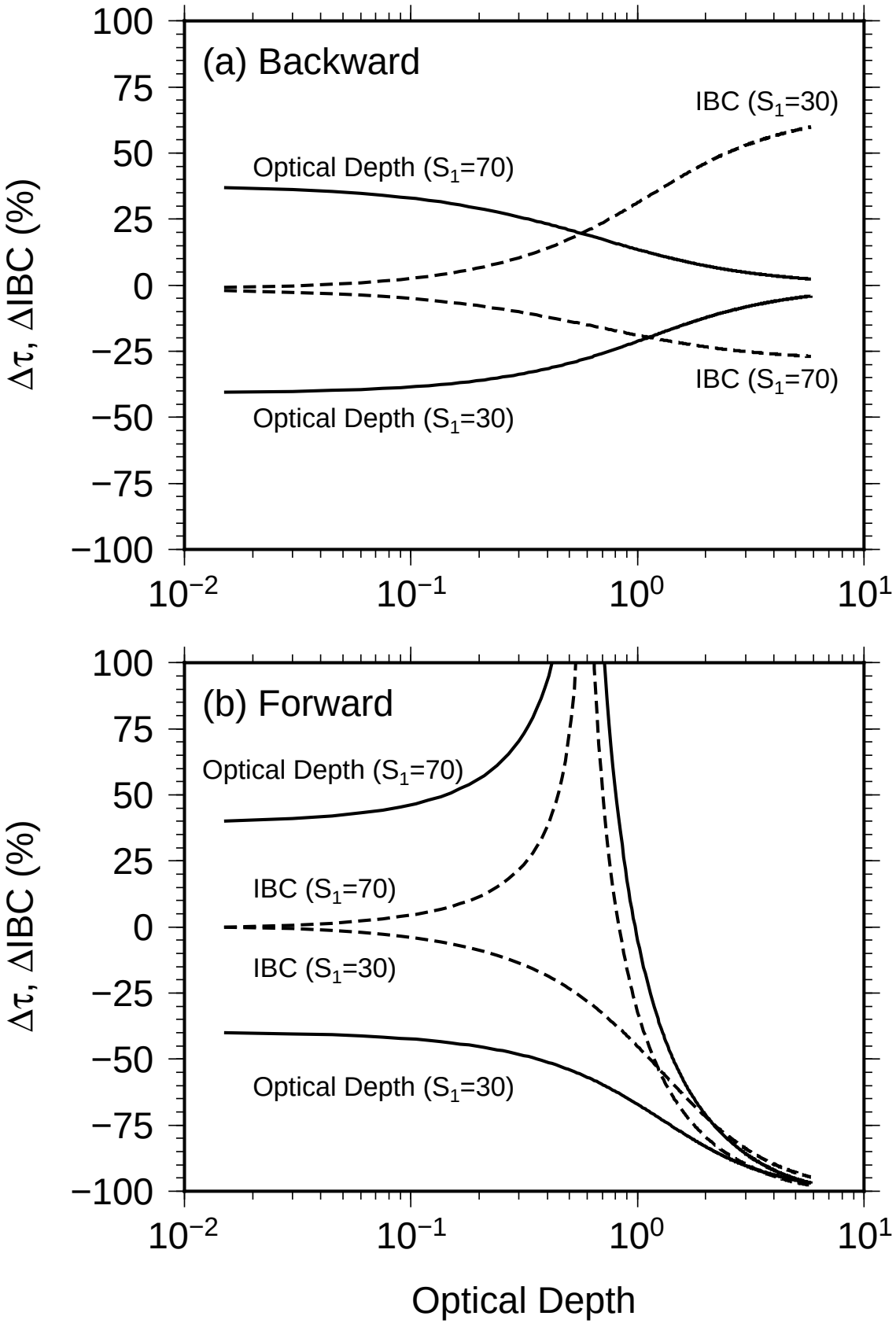

Figure 2.4: Relative errors of the optical depth and the integrated backscattering coefficient (IBC) as a function of the true optical depth (S_1 = 50 sr) for (a) backward and (b) forward inversions. The figures are depicted by reference to an error analysis shown in the NIES website (http://www-lidar.nies.go.jp/AnalysisMethods/DataAnalysis1.html).

The other optical properties of aerosols are also investigated in this thesis. The aerosol optical depth (AOD, $\tau_{a,\lambda}$) represents the attenuation rate of vertically incident light, and is defined as:

$$\tau_{a,\lambda} = \int_0^\infty \alpha_{a,\lambda}(z)dz, \tag{2.15}$$

where subscript a denotes aerosols. According to the Lambert-Beer's law, the incident light intensity decreases by 63.2% if AOD is 1.0.

The backscattering ratio is defined as the ratio of the total (particles plus molecules) backscattering coefficient to the molecular backscattering coefficient as follows:

$$R_\lambda(z) = \frac{\beta_{1,\lambda}(z) + \beta_{2,\lambda}(z)}{\beta_{2,\lambda}(z)}. \tag{2.16}$$

The value of $R_\lambda(z) - 1$ is proportional to the mixing ratio of the particles.

The depolarization ratio, which is measured at 532 nm in this thesis, is a parameter for non-spherical particles. It identifies dust from other aerosols and identifies ice crystals from water droplets empirically. If scatterers are uniformly spherical particles, the depolarization ratio is to be zero. The volume depolarization ratio is calculated from the ratio of the perpendicular component to parallel component with respect to the polarized plane of the emitted laser, and is given by:

$$\delta_v(z) = \frac{\beta_{1,\perp}(z) + \beta_{2,\perp}(z)}{\beta_{1,\parallel}(z) + \beta_{2,\parallel}(z)} = \frac{P_\perp(z)}{P_\parallel(z)}, \tag{2.17}$$

where subscripts $\perp$ and $\parallel$ denote perpendicular and parallel, respectively. If the particle backscattering coefficient is much larger than the molecular backscattering coefficient, the volume depolarization ratio is close to the particle depolarization ratio δ_p as the following equation [*Browell et al.*, 1990]:

$$\delta_p(z) = \frac{\beta_{1,\perp}(z)}{\beta_{1,\parallel}(z)} = \frac{\delta_v(z)\left[R(z) + R(z)\delta_2 - \delta_2\right] - \delta_2}{R(z) - 1 + R(z)\delta_2 - \delta_v(z)}. \tag{2.18}$$

The depolarization ratio of water droplets is approximately zero, but CALIOP observes the larger depolarization ratio for water droplets due to the multiple scattering caused by the large footprint of CALIPSO on the earth's surface [e.g. *Yoshida et al.*, 2010].

The color ratio χ is an indicator of the particle size, and is expressed by the ratio of the backscattering coefficients at two wavelengths as follows:

$$\chi(z) = \frac{\beta_{1,1064}(z) + \beta_{2,1064}(z)}{\beta_{1,532}(z) + \beta_{2,532}(z)}. \tag{2.19}$$

The wavelengths of 1064 and 532 nm are used to calculate the color ratio. Dust, sea-salt aerosols, and cloud particles have large size compared to the laser wavelengths, so that their color ratios are around 1.0. On the other hand, the color ratio of fine particles such as sulfate aerosols is smaller than 1.0. If the particle backscattering coefficient is much larger than the molecular backscattering coefficient, the color ratio is close to the particle color ratio χ_p as the following equation:

$$\chi_p(z) = \frac{\beta_{1,1064}(z)}{\beta_{1,532}(z)}. \tag{2.20}$$

2.4 Data processing of CALIOP products

In this section, a brief introduction of CALIOP products is presented. The optical parameters provided by the CALIOP level 1B data are attenuated backscattering coefficients at 532 and 1064 nm and perpendicular attenuated backscattering coefficients at 532 nm. The attenuated backscattering coefficient is given as:

$$\beta'_{S,\lambda}(z) = \beta_{S,\lambda}(z) T^2_{S,\lambda} = \beta_{S,\lambda}(z) \exp\left\{-2 \int_z^{z_s} \alpha_{S,\lambda}(z') dz'\right\}, \tag{2.21}$$

where the subscript S denotes the space-borne lidar and z_s is the altitude of the CALIPSO satellite. The attenuated backscattering coefficient at 532 nm is used for the C2 cloud mask, and those at 532 and 1064 nm are used for the estimation of the lidar ratio of dust. To improve the cloud mask, the volume depolarization ratio and attenuated color ratio are additionally used. The volume depolarization ratio at 532 nm is given as:

$$\delta_{S,v}(z) = \frac{\beta'_{S,\perp}(z)}{\beta'_S(z) - \beta'_{S,\perp}(z)} = \frac{\beta'_{S,\perp}(z)}{\beta'_{S,\parallel}(z)}, \tag{2.22}$$

where $\beta'_{S,\perp}(z)$ is the perpendicular attenuated backscattering coefficient. The attenuated color ratio is derived from the attenuated backscattering coefficients at 532 and 1064 nm as follows:

$$\chi'_S(z) = \frac{\beta'_{S,1064}(z)}{\beta'_{S,532}(z)}. \tag{2.23}$$

The CALIOP level 2 products provide the vertical feature mask (VFM), aerosol and cloud layer products, and aerosol and cloud profile products. The processing algorithms consist of the Selective Iterated BoundarY Locator (SIBYL), the Scene Classification Algorithms (SCA), and the Hybrid Extinction Retrieval Algorithms (HERA). The SIBYL algorithm detects aerosol and cloud layers, and determines the height of the layer base and top using attenuated backscattering coefficients at 532 nm at the five resolutions of 333 m, 1 km, 5 km, 20 km, and 80 km [*Vaughan et al.*, 2009]. The identified layers by SIBYL are classified into clouds and aerosols, and sub-types of them by SCA. The classification method in SCA utilizes a five-dimensional probability density function [*Liu et al.*, 2009, 2010]. The probability density functions (PDFs) of the layer-mean attenuated backscattering coefficient at 532 nm ($\overline{\beta'_{S,532}}$), the layer-integrated volume depolarization

ratio ($\overline{\delta_{S,v}}$), and the layer-integrated attenuated color ratio ($\overline{\chi'_S}$) are developed based on four months test data. These PDFs are binned by 20 altitudes from 0 to 20 km and 18 latitudes from 90°S to 90°N. Using PDFs and observed variables ($\overline{\beta'_{S,532}}$, $\overline{\delta_{S,v}}$, and $\overline{\chi'_S}$), a confidence function (f_{5D}) of the classification is calculated by the following equation:

$$f_{5D} = \frac{PDF_c\left(\overline{\beta'_{S,532}}, \overline{\delta_{S,v}}, \overline{\chi'_S}, z_m, lat\right) - PDF_a\left(\overline{\beta'_{S,532}}, \overline{\delta_{S,v}}, \overline{\chi'_S}, z_m, lat\right)}{PDF_c\left(\overline{\beta'_{S,532}}, \overline{\delta_{S,v}}, \overline{\chi'_S}, z_m, lat\right) + PDF_a\left(\overline{\beta'_{S,532}}, \overline{\delta_{S,v}}, \overline{\chi'_S}, z_m, lat\right)}, \tag{2.24}$$

where PDF is the five-dimensional probability function of the observed variables, subscripts a and c denote aerosols and clouds, respectively, z_m is the center altitude of the layer, and lat is the latitude. For example, if the confidence function is -1, the observed layer is identified as aerosols having the highest possibility. The cloud-aerosol discrimination (CAD) score reported in the products is the percentage of f_{5D}, and the score ranges from -100 to 100. The observed layer is classified as aerosols or clouds if the CAD score is negative or more than 0, respectively. The VFM mask, which is the data-masking product of observed signals, is based on the CAD score. It should be noted that the classification of 333 m resolution layers is not based on the CAD score. The 333 m resolution layers are automatically recognized as clouds. The layer detection of the 333 m resolution depends only on the backscattering intensity, causing the misclassification of aerosols that have the strong backscattering signals as clouds.

Optimal lidar ratios are selected after the layer classification. The lidar ratios are calculated for elevated layers if the observed backscattering signals can be assumed as the molecular backscattering coefficient just above and below the layer. The calculation method is described in Appendix A. If the molecular backscattering cannot be assumed just above and below the layer, the lidar ratio is given by modeled values. The aerosol optical model is based on cluster analysis of the Aerosol Robotic NETwork (AERONET) measurements [*Omar et al.*, 2009]. Aerosols are classified into six categories: desert dust, biomass burning, polluted continental, marine, polluted dust, and clean continental. The intended estimation errors is less than 30%. The retrieval of the backscattering or extinction coefficient profiles using the selected lidar ratios is implemented by HERA.

3 Estimation of the lidar ratio of the dust by the ground/space-based lidar instruments

The Taklimakan Desert is a large source of Asian dust, but the lidar ratio of dust over the desert has not been investigated in the past. In Chapter 2, issues caused by incorrectly selected lidar ratios for Mie-scattering lidar were shown by the simulation experiment using the backward and forward inversion methods. In order to solve the issues, this thesis attempts to estimate the lidar ratio of dust over the Taklimakan Desert by simultaneous observations with the ground/space-based lidars. This chapter presents the estimation method and results of the lidar ratio. Application of the estimated lidar ratio to the calculation of optical properties retrieved from ground-based lidar data is also shown.

3.1 Estimation method of lidar ratios using ground/space-based lidar data

The simultaneous observation by Aksu-lidar and CALIOP was conducted to estimate the dust lidar ratio on 23 March 2009. Figure 3.1 (a) illustrates a CALIPSO ground track and distance circles from the Aksu-lidar station. The ground track passed by the Aksu-lidar station at 20:57 UTC during the nighttime. The attenuated backscattering coefficients at 532 and 1064 nm wavelengths (CALIOP level 1B data) are used where the distance from the Aksu-lidar station is within 100 km. Figure 3.1 (b) shows the latitude-altitude cross section of attenuated backscattering coefficients at 532 nm used in this thesis. The discontinuity at the altitude of 8.2 km throughout the latitudes is caused by the different spatial resolutions between above and below the altitude as described in Chapter 2.1. From the surface level to 4.5 km, there is an aerosol layer recognized by weaker signals just above 4.5 km. High clouds cover the aerosol layer at 7.5–11.5 km. Since some of the clouds are optically thick, a part of the underlying aerosol layer is not visible due to the laser attenuation (e.g., 40.2°N–40.45°N).

To avoid noise contamination in averaging processes, profiles having a low SNR under optically thick clouds should be removed. First, backscatter profiles are horizontally averaged over 5 km (15 shots) to improve the SNR. Secondly, the cloud transmittance of each profile is calculated by the method of *Platt et al.* [1999], assuming that backscattering ratios just above and below clouds are 1.0. The calculation method is described in Appendix A. Finally, profiles whose cloud transmittance is less than 0.135 (corresponding to the optical depth of 1.0) are removed. The remainders of the profiles are further averaged. The averaged profiles at 532 and 1064 nm are

shown in Figure 3.1 (c). The altitudes just above and below the cloud are denoted as z_t and z_b, respectively. Profiles of range-corrected signals averaged over five minutes from 20:54 to 20:59 UTC are used as the Aksu-lidar data. The averaged profiles are shown in Figure 3.1 (d). The calibration altitudes of Aksu-lidar at 532 and 1064 nm are denoted as $z_{c,532}$ and $z_{c,1064}$, respectively. Clouds are not observed by Aksu-lidar at that time. The volume depolarization ratios derived from Aksu-lidar and CALIOP are shown in Figure 3.1 (e). The large depolarization ratio (~0.4) of the cloud indicates ice crystals. The depolarization ratio of the aerosol layer below 7 km is ~0.3. Therefore, the aerosol layer is identified as dust originated from the Taklimakan Desert.

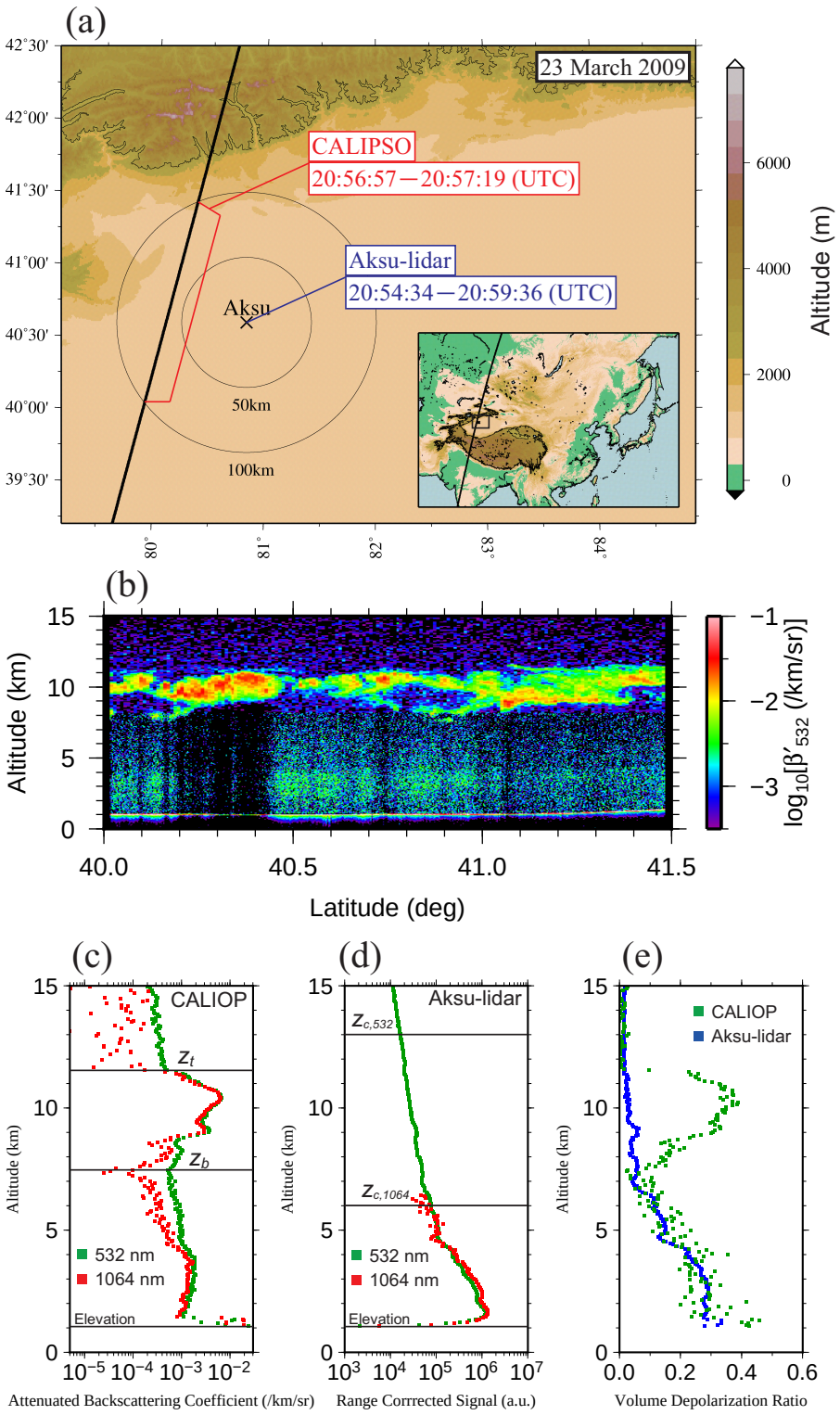

Figure 3.1: (a) CALIPSO ground track and distance circles from the Aksu-lidar station. CALIPSO passed within 100 km of the Aksu-lidar station from 20:56:57 to 20:57:19 (UTC). The ground-based lidar observations were conducted during that time. (b) Latitude-altitude cross section of CALIOP attenuated backscattering coefficients at 532 nm. Bottom panels are vertical profiles of (c) CALIOP averaged signals, (d) Aksu-lidar signals, and (e) volume depolarization ratios.

In order to estimate the lidar ratio of dust, two assumptions are made as follows: 1) the lidar ratio is constant with altitude, and 2) Aksu-lidar and CALIOP measure the same atmospheric column except for clouds. If attenuation corrections are precisely conducted, the backscattering coefficients retrieved from Aksu-lidar should agree with that from CALIOP. The lidar ratio is estimated by using a performance function that is defined as the square error of backscattering coefficients between Aksu-lidar and CALIOP. The performance function to estimate the lidar ratio was introduced by *Tao et al.* [2008a]. In this thesis, the performance function is modified to estimate the lidar ratio under following conditions: 1) clouds covers over aerosol layers, and 2) signals of ground-based lidar cannot be calibrated at an altitude of 30 km. The performance function Ψ_λ is described by following equations:

$$\Psi_{532}\left(S_{1,532}, R_{S,532}(z_b)\right) = \sum_{z=2km}^{z=7km} \left[\beta_{G,532}(z) - \beta_{S,532}(z)\right]^2, \tag{3.1}$$

$$\Psi_{1064}\left(S_{1,1064}, R_{G,1064}(z_{c,1064})\right) = \sum_{z=2km}^{z=6km} \left[\beta_{G,1064}(z) - \beta_{S,1064}(z)\right]^2, \tag{3.2}$$

where the subscripts G and S denote the ground-based and space-borne lidar systems, respectively. The backscattering coefficients of Aksu-lidar and CALIOP data are calculated by the backward and forward inversion methods, respectively. Therefore, their calculation directions are identical with each other (i.e., from the higher to lower altitudes). The integral ranges of the right side of the equations are set to the altitude ranges where the signals of dust are available both from the Aksu-lidar and CALIOP data. The 1064 nm signals of Aksu-lidar cannot be used above 6 km due to the laser attenuation.

The unknowns of the equations are: 1) the 532 nm lidar ratio, 2) the 532 nm backscattering ratio of CALIOP at z_b, 3) the 1064 nm lidar ratio, 4) the 1064 nm backscattering ratio of Aksu-lidar at $z_{c,1064}$. The 532 nm backscattering ratio of CALIOP at z_b is needed to calculate the cloud lidar ratio. The details of the calculation method of cloud lidar ratios are described in Appendix A. The 1064 nm backscattering ratio of Aksu-lidar at $z_{c,1064}$ is necessary for the particle backscattering coefficient of the initial value of the inversion as described in Chapter 2.3. The performance functions are calculated by changing the lidar ratios and the backscattering ratios. If the performance function returns a minimum value, the unknowns are regarded as the optimal solution. Therefore, the lidar ratio can be estimated.

3.2 Estimation results of the lidar ratio of dust over the Taklimakan Desert

Results of the lidar ratio estimated by the performance functions are shown in this section. Figure 3.2 shows the result of the performance functions, as a function of the lidar ratio and the backscattering ratio. The performance function returns small values if the difference of backscattering coefficients between Aksu-lidar and CALIOP is small. The deep blue regions in the figure represent the small values. The minimum point of the performance function represents the optimal solution of the lidar ratio and backscattering ratio. The estimated lidar ratios are 42.0 sr at 532 nm and 45.9 sr at 1064 nm. The estimated backscattering ratios are 1.48 at 532 nm and 9.92 at 1064 nm. In Figure 3.2 (a), the region of the local minimal values changes from the higher to lower lidar ratio as the backscattering ratio increases. This result denotes that the lidar ratio estimation at 532 nm is sensitive to the backscattering ratio of CALIOP at z_b. Therefore, it is influenced by the presence of clouds. In Figure 3.2 (b), the 1064 nm lidar ratio at the local minimal values increases with the 1064 nm backscattering ratios of Aksu-lidar at $z_{c,1064}$.

The decrease (532 nm) and increase (1064 nm) of the lidar ratio at the local minimal values are illustrated in Figure 3.3. The backscattering coefficients are calculated on the assumption of the constant lidar ratio and selected backscattering ratios. The constant lidar ratios at 532 and 1064 nm are 40 and 45 sr, respectively. The selected backscattering ratios at 532 nm are 1.0 and 2.0, and those at 1064 nm are 2.0 and 50.0.

In the 532 nm case of the smaller backscattering ratio ($R_{S,532}(z_b) = 1.0$), the backscattering coefficients of CALIOP are lower than those of Aksu-lidar. To minimize the performance function, the lidar ratio should be larger than the selected one. The reason is illustrated by the demonstration of the inversion shown in Figure 2.3. Again, the backward and forward inversions are applied to Aksu-lidar and CALIOP, respectively. The backscattering coefficients calculated by using the larger lidar ratio are underestimated in the backward inversion and are overestimated in the forward inversion. In the case of the larger backscattering ratio ($R_{S,532}(z_b) = 2.0$), the backscattering coefficients of CALIOP are larger than those of Aksu-lidar. The 532 nm lidar ratio should be smaller than the selected one to minimize the performance function. Therefore, the lidar ratio at the local minimal values of the performance function decreases with the backscattering ratio.

The 1064 nm backscattering coefficients of Aksu-lidar in Figure 3.3 (b) vary depending on the selected backscattering ratio. In the case of the smaller backscattering ratio ($R_{G,1064}(z_c) = 2.0$), the backscattering coefficients of Aksu-lidar are smaller than those of CALIOP. According to Figure 2.3, the lidar ratio should be smaller than the selected one to minimize the performance function. In the case of the larger backscattering ratio ($R_{G,1064}(z_c) = 50.0$), the lidar ratio should be larger

than the selected one in the same way. Therefore, the 1064 nm lidar ratio at the local minimal values of the performance function increases with the backscattering ratio.

The profiles of backscattering coefficients calculated by using the estimated lidar ratio are shown in Figure 3.4. Two profiles at each wavelength agree with each other except for the cloud. The random noise contamination of CALIOP decreases with altitude. In contrast, the noise contamination of Aksu-lidar increases with altitude. The direction of the increasing contamination corresponds to the direction of the laser propagation. A scatter plot of backscattering coefficients is shown in the lower left of Figure 3.4. The solid line in the scatter plot is the regression line. The correlation coefficients are 0.98 at 532 nm and 0.95 at 1064 nm. The profile agreement with the high correlation coefficient indicates that the estimated lidar ratio is an accurate value. In addition, it indicates that dust over the Taklimakan Desert distributes homogeneously since the distance between the CALIPSO ground track and Aksu-lidar station is more than 60 km (see Figure 3.1 (a)).

Estimation errors of the lidar ratio are investigated. Table 3.1 shows errors caused by the random noise and signal calibration. The estimation procedure of errors caused by the random noise is described in Appendix B. The estimated errors of the random noise are 7.2% at 532 nm and 21.1% at 1064 nm. The calibration error is examined by changing the calibration constant of CALIOP. The uncertainty of the 532 nm calibration constant is approximately 3.5% [*Hostetler et al.*, 2006]. The corresponding error of the lidar ratio is ~6.3%. The uncertainty of the 1064 nm calibration constant is approximately 5.0%. In the algorithm of the CALIPSO science team, the 1064 nm signal is calibrated by assuming that the color ratio of ice clouds is 1.0. The calibration constant depends on the latitude and existence or non-existence of solar insolation in this method [*Hunt et al.*, 2009]. Moreover, it is reported that the color ratio for ice clouds is less than 1.0 [*Bi et al.*, 2009; *Tao et al.*, 2008b]. According to *Okamoto et al.* [2010], the 1064 nm correction factor calculated by the water cloud calibration is approximately 0.85 at the northern mid-latitudes. The corresponding error of the lidar ratio is 35.8%. The total errors (root sum square) are 9.5% at 532 nm and 41.6% at 1064 nm. The other variation factors of the lidar ratio estimation are the relative humidity, internal or external mixtures of dust and other aerosols, and the variation of size distribution with altitude. The relative humidity at the simultaneous observation is less than 50% according to the ECMWF data. The dry air condition has small effects on the lidar ratio of dust [*Sakai et al.*, 2003].

The lidar ratio estimated by this thesis is compared with those reported in previous studies in Table 3.2. This is the first study of the estimation of the dust lidar ratio over the Taklimakan Desert.

The 532 nm lidar ratio in this thesis is 3 $\sim$ 24% smaller than those of Asian dust reported previously. Since the investigation of 1064 nm lidar ratios is rarely conducted, this result contributes to the reduction of the uncertainty of the 1064 nm lidar ratio.

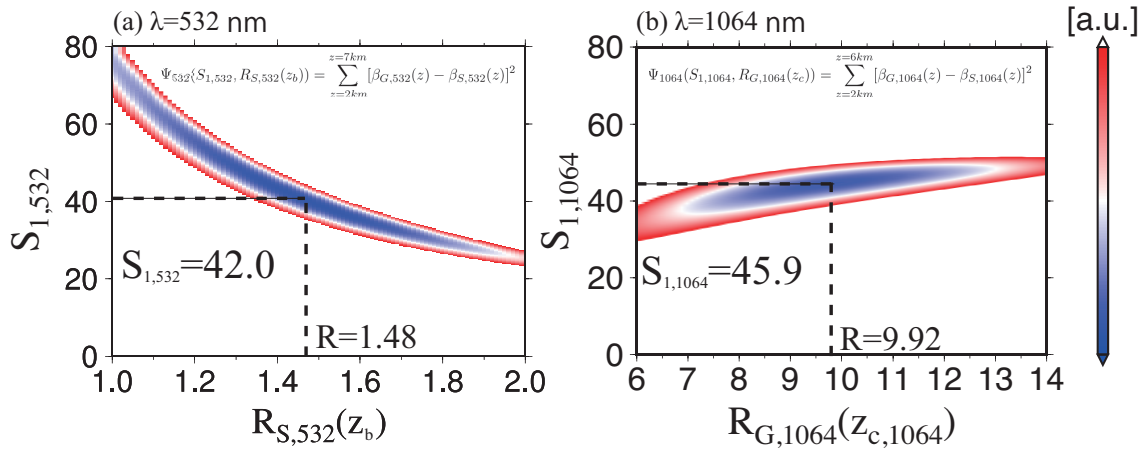

Figure 3.2: Performance function at (a) 532 nm and (b) 1064 nm as a function of the lidar ratio and backscattering ratio. The estimated lidar ratio and backscattering ratio are shown. The depicted color means the magnitude of the performance function (blue: small, red: large).

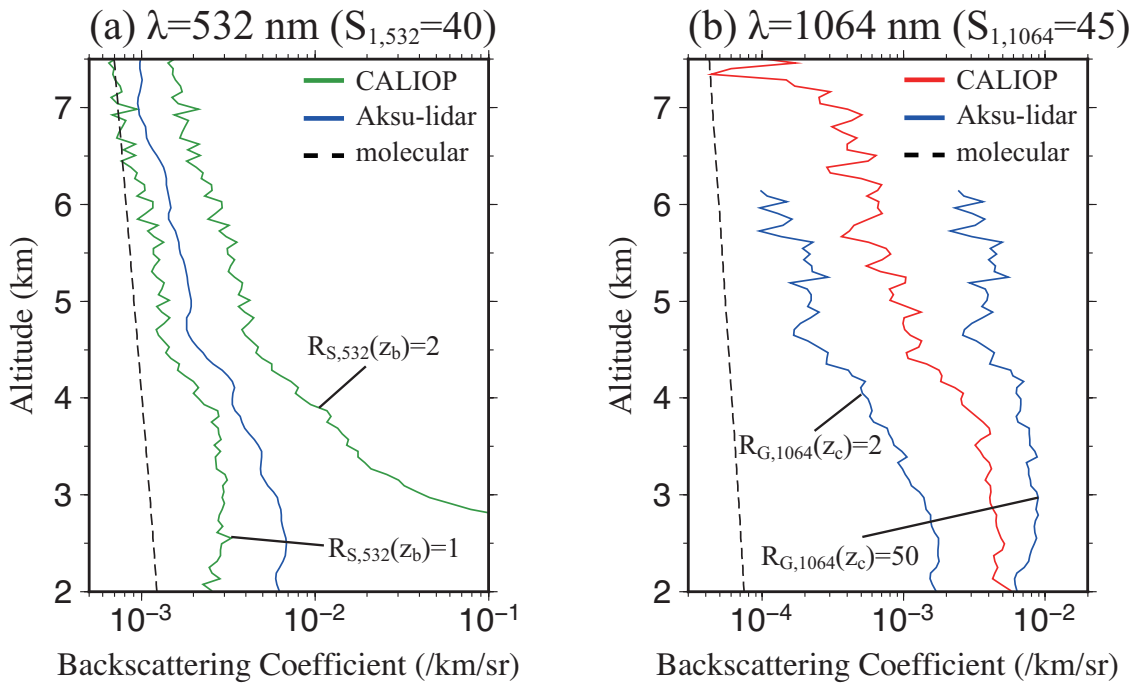

Figure 3.3: Backscattering coefficients at (a) 532 nm and (b) 1064 nm, calculated by changing the backscattering ratio.

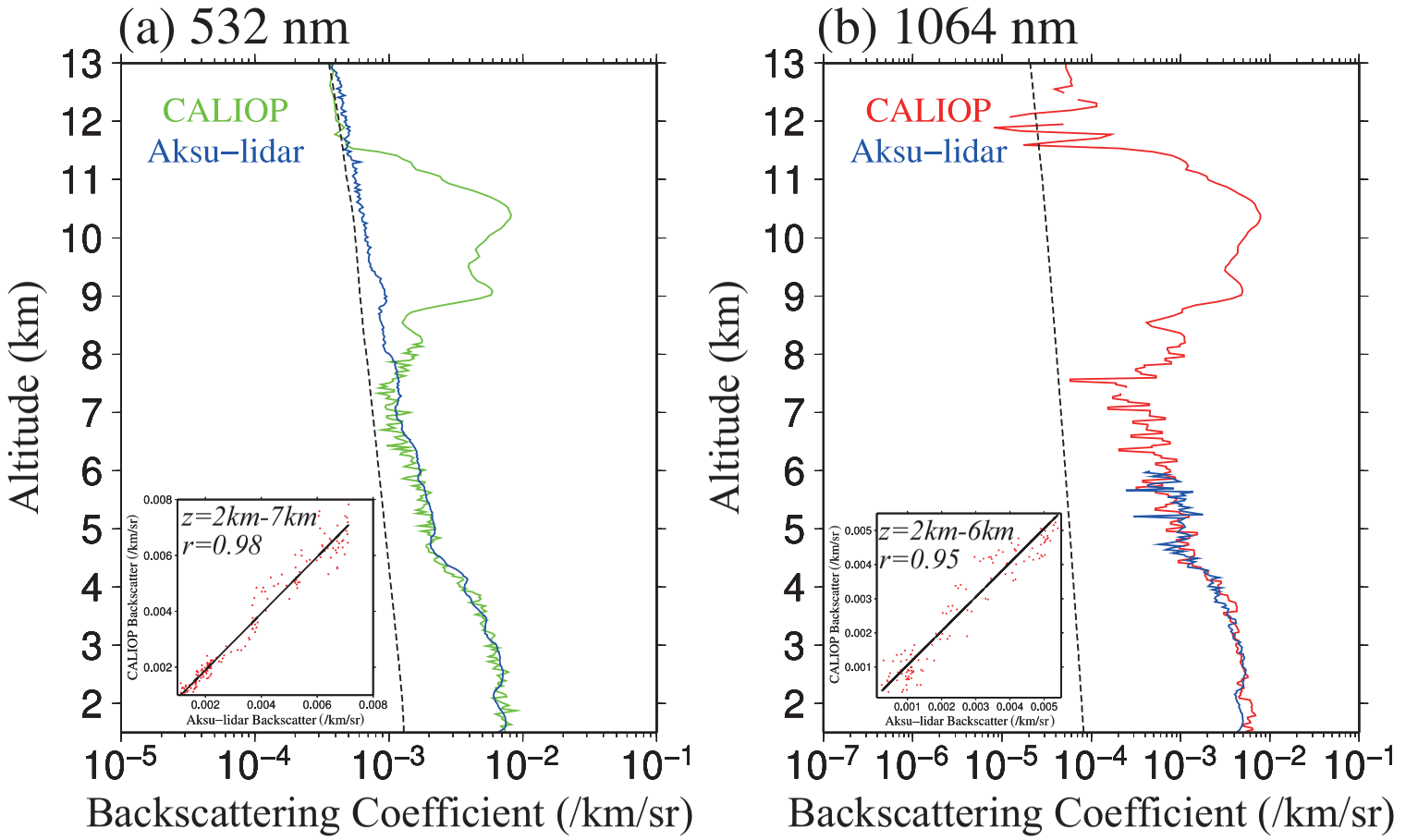

Figure 3.4: CALIOP and Aksu-lidar backscattering coefficients at (a) 532 nm and (b) 1064 nm, retrieved by using estimated lidar ratios. The dotted line indicates the molecular backscattering coefficients. The scatter plot between the backscattering coefficients is shown at the lower left of the figures.

Table 3.1: Estimated Error in $S_{1,532}$ and $S_{1,1064}$ in Parcent.

Error Source	$S_{1,532}$	$S_{1,1064}$
Random Noise	7.2	21.1
Calibration	6.3	35.8
Total (root sum square)	9.5	41.6

Table 3.2: Comparison of lidar ratios and color ratios observed in the present study with those of previous studies.

Lidar Ratio			
Dust Type	**532 nm**	**1064 nm**	**Reference**
Saharan Dust	26±4.8	35±18.3	Vaughan [2004]
	41±3.0	52±5.0	Liu et al. [2008]
Asian Dust	45.5±8.6		Noh et al. [2007]
	47±18		Sakai et al. [2003]
	42-55		Liu et al. [2002]
	43.1±7.0		Murayama et al. [2004]
	42.0±4.0	45.9±19.1	Present Study
Desert Dust	40±20	55±17	CALIOP (version 3)
Particle Color Ratio			
Dust Type	**1064 nm/532 nm**		**Reference**
Saharan Dust	0.86±0.01		Vaughan [2004]
	0.74±0.07		Liu et al. [2008]
Asian Dust	0.79±0.12		This Thesis

3.3 Dust optical properties from ground-based lidar data with the estimated lidar ratio

The lidar ratio estimated by the performance function is applied to the inversion of the Aksu-lidar data. Figure 3.5 shows time-altitude cross sections of backscattering ratios, volume depolarization ratios, and particle color ratios on 23-25 March 2009. The arrows in the top of the figures are the profile used in the simultaneous observation with CALIOP. The backscattering ratio shows no cloud by chance at only that time. Commonly, cloud backscattering ratios are much lager than aerosol one because of the larger particle size. There is a dust layer below 5 km and cloud layers above 5 km. The depolarization ratio and color ratio of dust are 0.25∼0.30 and 0.5∼1.0, respectively. In this case, a floating dust is observed throughout the observation period. Since the dust occurrence frequency over the Taklimakan Desert is close to 1.0 during springtime [*Liu et al.*, 2008b], dust can be observed at any time.

There are two-layered dust structures from 4 to 6 km altitudes at 13–22 LST on 23 and 24 March. The lower dust layer sinks to 1 km below the top most layers. Dust with small backscattering ratios ($R_{532} < 3.0$) extends to 7 km from 22 LST on 23 March to 09 LST on 24 March. A local circulation can cause the downward and upward movements of the dust layer height during the afternoon and morning. *Kim et al.* [2009] investigated the influence of meteorological conditions on the variation of dust layer heights at Aksu in April 2002 by a numerical simulation and lidar observations. They suggested that dust layer heights are decreased during the daytime by descending flows from the Tianshan Mountains, and it rises rapidly due to surface convergence in the evening.

Clouds are observed at the height where weak dust layers exist. For example, they are seen at 02 and 08 LST on 24 March, 00–12 LST on 25 March (Figure 3.5 (c)). It can be expected that the dust aerosols are mixed with the clouds and/or serve as the cloud nuclei. For the clouds at 02 and 08 LST on 24 March, and 09 LST on 25 March, the cloud top temperatures are from −25°C to −30°C. The cloud top temperature is derived from the ECMWF-interim data. These clouds would contain ice crystals because of the high depolarization ratio. There is a low depolarization (near zero) of the clouds formed at the top of the boundary layer (∼4.5 km) at 09–12 LST on 25 March. The cloud top temperature is about −10°C, and plate-like ice crystals (2D-plate) or water droplets can be expected for these clouds. The 2D-plate ice is frequently confirmed at this temperature according to lidar and in-situ observations [e.g. *Yoshida et al.*, 2010].

The particle color ratio of dust decreases rapidly from 0.7 to 0.5 after 10 LST on 25 March. The decrease of the color ratio indicates that the size of observed aerosol particles becomes smaller.

The possible causes of the decrease are: 1) the concentration of large particles decreases due to the dry deposition, and 2) fine particles are transported from the surrounding regions. The deposition of large particles is reasonable since the 1064 nm backscattering ratio decreases after 10 LST on 25 March. However, there is a small difference of the 532 nm backscattering ratios before and after 10 LST on 25 March although the coarse particles may settle. This can be explained by advections of fine particles that are less sensitive to the backscattering at 1064 nm than that at 532 nm.

Using the backscattering ratio and volume depolarization ratio, the particle depolarization ratio of dust is investigated. Figure 3.6 shows a scatter plot of the volume depolarization ratio versus the 532 nm backscattering ratio. The depicted theoretical lines ($\delta_p = 0.25, 0.3, 0.35, 0.4$) are calculated by Equation (2.18). The particle depolarization ratio of the relatively dense dust ($R_{532} > 4.0$) decreases gradually through the observation from 0.38 to 0.3. Since the external or internal mixture with dust and spherical particles can cause the decrease of the particle depolarization ratio, the spherical particles may be advected from the surrounding regions.

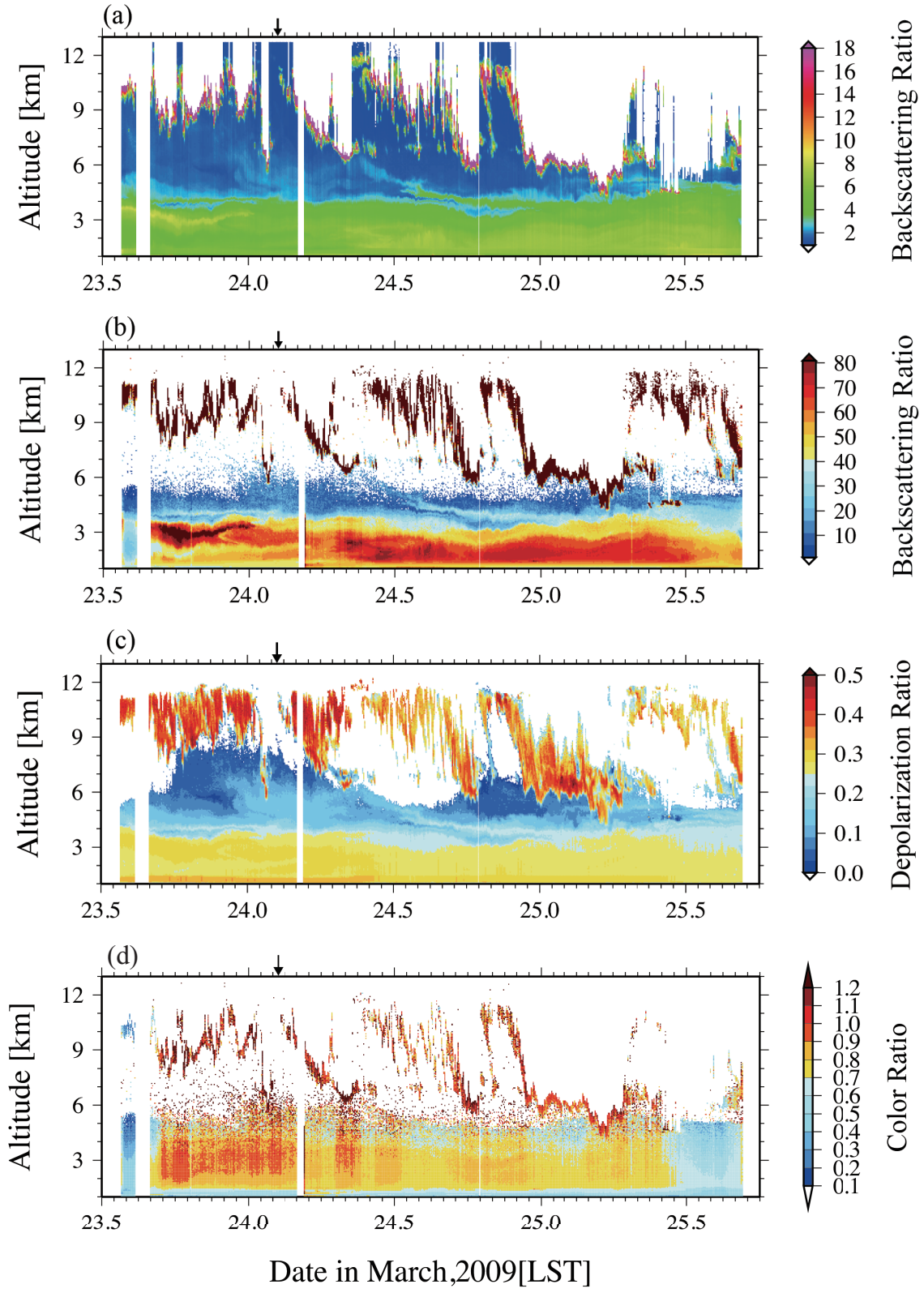

Figure 3.5: Time-altitude cross section of backscattering ratios at (a) 532 nm and (b) 1064 nm, (c) volume depolarization ratios at 532 nm, and (d) particle color ratios observed by Aksu-lidar in March 2009. The backscattering ratio is masked in the white color if the optical depth is larger than 1.0.

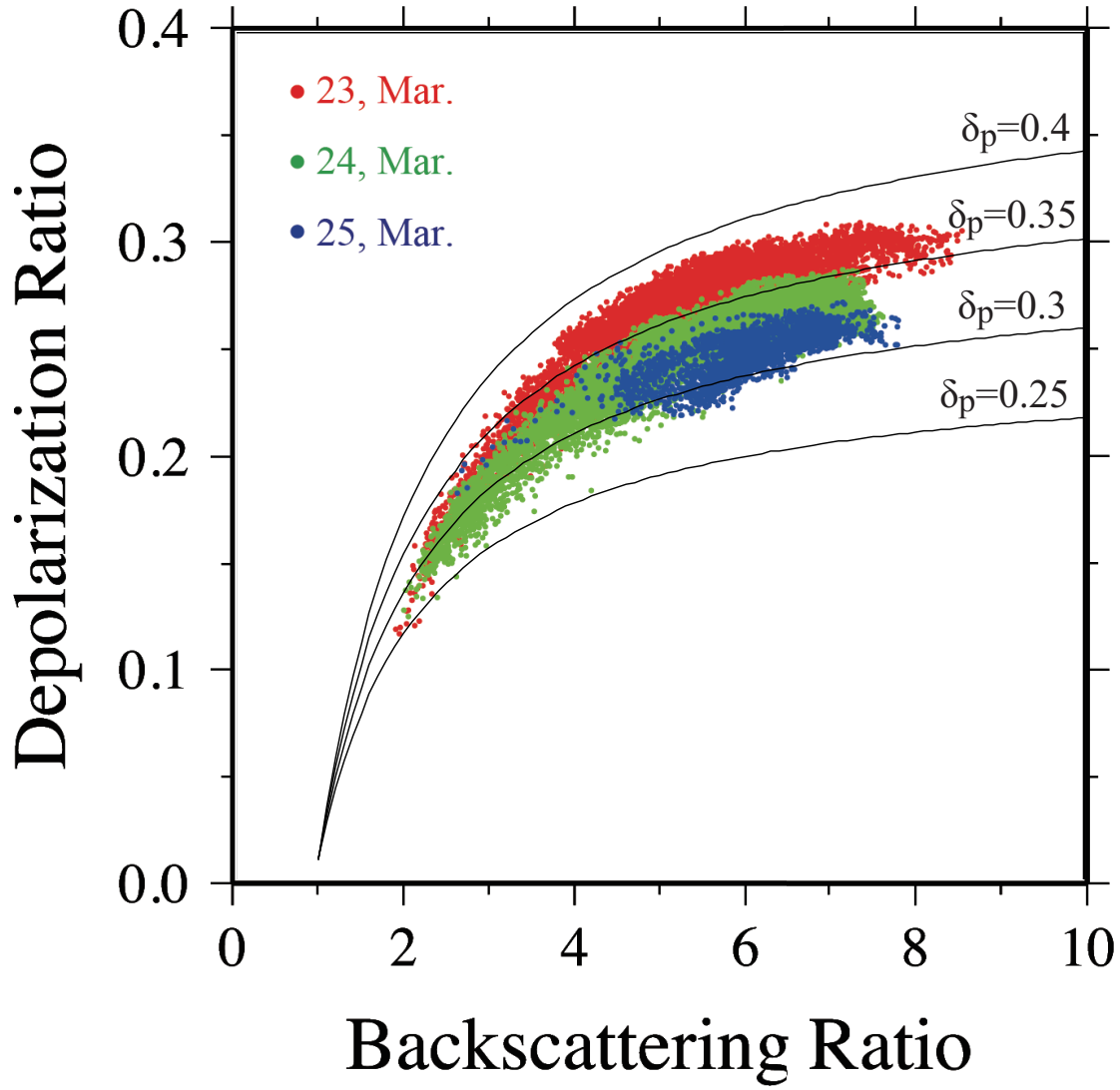

Figure 3.6: Scatter plot of the volume depolarization ratio versus the backscattering ratio. The theoretical lines of δ_p from 0.25 to 0.4 are depicted.

3.4 Comparison with CALIOP level 2 products

The backscattering coefficients from the simultaneous observation by Aksu-lidar and CALIOP are compared with the CALIOP level 2 products. Figure 3.7 (a) shows the time-altitude cross section of the 532 nm particle backscattering coefficients of aerosols on 23 March 2009, derived from the CALIOP profile products. The broken line of the red rectangle is the region where the profiles are used in the simultaneous observation. There are missing data in the red rectangle region since the missing areas (black color) are classified as clouds. The time-altitude cross section of the vertical feature mask (VFM) is shown in Figure 3.7 (b). Most of the data in the red rectangle is identified as clouds, followed by aerosols. As shown in Figure 3.1 (b), there are no cloud layers below 7 km. Dust is presumably misclassified as clouds in the VFM result. The misclassification appears mainly at higher than 40°N. The probability density functions in the VFM are binned by 10° latitude as described in Chapter 2.4. This thesis suggests that the latitudes of higher than 40°N have an issue in the discrimination between clouds and aerosols.

The aerosol sub-type arrayed in the Tarim Basin is desert dust. The lidar ratio selected from the aerosol model is 40 sr at 532 nm. The selected lidar ratio is little different from the lidar ratio estimated by this thesis. However, the averaged values of aerosol backscattering coefficients in the red rectangle region are smaller than those estimated by this thesis as shown in Figure 3.7 (c). The averaged profile of the CALIOP level 2 data is 21% smaller than our results. The difference of the backscattering coefficients is large below 3 km. It is caused by the underestimates of the attenuation corrections above 3 km. The lidar ratio of clouds is selected approximately 20 sr from the theoretical studies [*Pinnick et al.*, 1983]. The misclassified clouds are assigned to the cloud lidar ratio. As demonstrated in Figure 2.3, the smaller lidar ratio causes the underestimates of backscattering coefficients in the forward inversion used in the CALIOP analysis. Since the cloud lidar ratio ($\approx$ 20 sr) is smaller than the dust lidar ratio ($>$ 40 sr), the backscattering coefficient of dust misclassified as clouds is expected to be smaller. This thesis indicates that the accuracy of the cloud-aerosol discrimination is important to select the appropriate lidar ratio below the cloud layers.

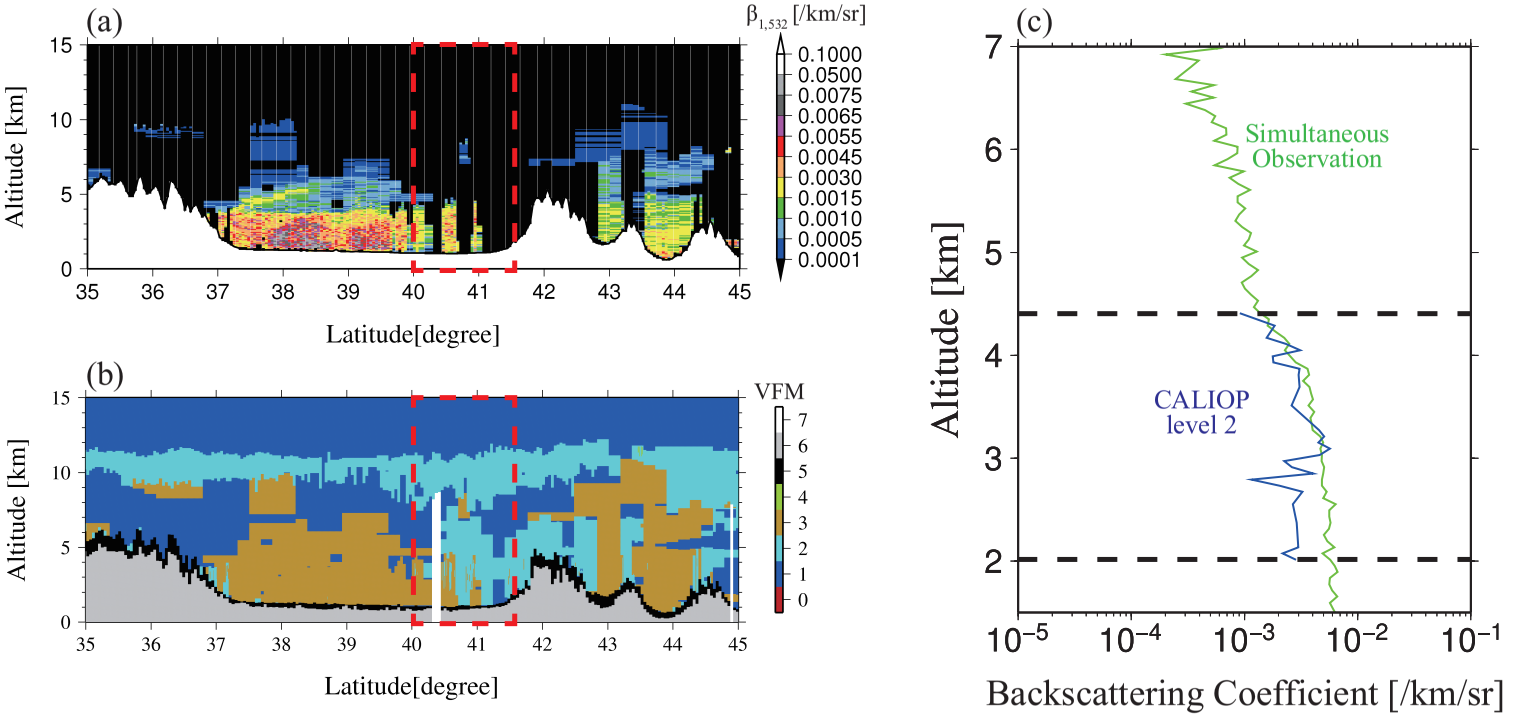

Figure 3.7: Results of the CALIOP level 2 products over the Taklimakan Desert on 23 March 2009. Latitude-altitude cross sections of (a) CALIOP level 2 particle backscattering coefficients and (b) CALIOP level 2 vertical feature mask (VFM), whose color bar indicates the feature type (0: invalid, 1: clear air, 2: cloud, 3: aerosol, 4: stratospheric feature, 5: surface, 6: subsurface, 7: no signals). (c) Profiles of the backscattering coefficient of the CALIOP level 2 data (blue) and our results (green).

4 Improvement of CALIOP cloud masks by a discriminant analysis

The lidar ratio of aerosol or cloud layers observed by CALIOP is rarely estimated by the method shown in Chapter 3.2. If the lidar ratio cannot be estimated by measurements, the modeled value is used according to the layer types (i.e., clouds, aerosols, and their sub-types). The misclassification in CALIOP cloud masks leads to the selection of the incorrect lidar ratio of aerosols. This thesis tries to solve the misclassification issues by a discriminant analysis. This chapter presents a new determination method of training data and the applicability of the discrimination model. The profiles of dust extinction coefficients after the improvement of cloud masks are also shown.

4.1 Description of cloud mask products used in this thesis

Prior to the introduction of the discriminant analysis, the descriptions of CALIOP cloud mask, CloudSat cloud mask, and MODIS cloud mask products are presented. The CALIOP cloud mask to be improved is the C2 cloud mask. The CloudSat (C1) cloud mask and MODIS cloud mask are used to determine misclassified clouds in the C2 cloud mask. This section describes the schemes of the C1, C2, and MODIS cloud masks.

4.1.1 C1 and C2 cloud masks

The C1 and C2 cloud masks are included in a CloudSat-CALIPSO merged dataset developed by *Hagihara et al.* [2010]. These cloud masks are derived from Cloud Profiling Radar (CPR) and CALIOP. The cloud mask schemes are based on cloud masks used in observations by a ship-borne 95 GHz cloud radar and a lidar in the western Pacific Ocean near Japan [*Okamoto et al.*, 2007]. The 94 GHz CPR is onboard CloudSat, which was launched simultaneously with CALIPSO as a part of the A-train constellation. CloudSat and CALIPSO had passed the same ground track within 15 seconds until CloudSat dropped out of the orbit in April 2011.

The C1 cloud mask scheme uses a cloud mask product of the CPR level 2B GEOPROF data. The CloudSat data are downloaded from the data processing center (http://www.cloudsat.cira.colostate.edu/index.php). The CPR cloud mask product is constructed by the signal-to-noise ratio (SNR), spatial continuity test, and horizontal averaging [*Marchand et al.*, 2008]. The product consists of a confidence level ranging between 0 and 40. If the confidence level is higher than 20, the estimated false detection is $< 5\%$. An observed bin (i.e., a data point at an altitude at a

certain time) whose confidence level of the cloud mask is higher than 20 is identified as a cloud bin in this thesis. In a provisional cloud mask, cloud bins are labeled as 1, and others are labeled as 0. The weighted average of the provisional cloud mask is vertically calculated. Each profile has the same altitude registration with respect to the geoid data. The altitude registrations from 0 to 20 km include 83 vertical bins with a vertical resolution of 240 m. The averaged cloud mask has fractional values ranging from 0 to 1. If the averaged cloud mask of a bin is more than 0.5, the bin is identified as cloud bins in this thesis. In the C1 cloud mask, the identified cloud bin is labeled as 1, and others are labeled as 0. The lower bins below the height of 1 km above the surface elevation are eliminated since surface clutter is dominated.

The C2 cloud mask scheme uses 532 nm attenuated backscattering coefficients $\beta'_{S,532}$ of the CALIOP Level 1B data (version 3). To detect a cloud bin in a lidar-observed bin, two criterions are applied to the attenuated backscattering coefficient.

The first criterion is constructed by threshold values of the attenuated backscattering coefficient at 532 nm. The threshold values β_{th} are determined by a background noise and the molecular signals. The first criterion is passed if the attenuated backscattering coefficients exceed the threshold values. Figure 4.1 (a) is an example of the threshold values at a snap shot from *Hagihara et al.* [2010]. The attenuated backscattering coefficients in Figure 4.1 (a) exceed the threshold values from 10.5 to 12.5 km where there is a cirrus layer. There are strong attenuated backscattering coefficients due to an aerosol layer below 5 km, but most of them are less than the threshold values. This example demonstrates the successful discrimination of the cloud and the aerosol layers.

The second criterion is constructed by a spatial continuity test using the surrounding bins. Figure 4.1 (b) shows a graphical representation of the second criterion. The test considers a sliding data window centered at the bin where the first criterion has been passed. The window sizes are 5×5 bins (horizontal by vertical) at altitudes < 5 km and are 9×9 bins at altitudes > 5 km. The selection of the larger window size at altitudes > 5 km is to reduce noise contamination and to detect weak signals due to thin cirrus clouds. If more than half of bins in the window pass the first criterion, the target bin (center bin) pass the second criterion. A bin that has passed the second criterion is identified as the cloud bin. According to the window size at altitude < 5 km, this scheme cannot detect clouds having a horizontal dimension < 0.8 km and a vertical dimension < 75 m. In a provisional cloud mask, cloud bins are labeled as 1, and others are labeled as 0.

Results of the provisional cloud mask have the original resolution of CALIOP as shown in Table 2.3. To compare the cloud masking results with the C1 cloud mask, the cloud fraction is calculated from the provisional cloud mask within ±0.55 km along track from the center of the

CloudSat footprint. In addition, the weighted average of cloud fractions is vertically calculated within the vertical range of the CPR resolution (240 m). The averaged cloud fraction is defined as the ratio of the number of cloud bins to the number of data bins in the CPR grid box. Since the provisional cloud mask is 0 (no cloud) or 1 (cloud), the averaged cloud fraction ranges from 0 to 1 at the common grid with the C1 cloud mask. In this thesis, if the averaged cloud fraction is more than 0.5, the bin is identified as the cloud bin at the CPR grid box. In the C2 cloud mask, identified cloud bins are labeled as 1, and others are labeled as 0. The C2 cloud mask has the horizontal resolution of 1.1 km and the vertical resolution of 240 m. These resolutions are identical to the resolutions of the C1 cloud mask. Further information of the C1 and C2 cloud mask schemes are described in *Hagihara et al.* [2010].

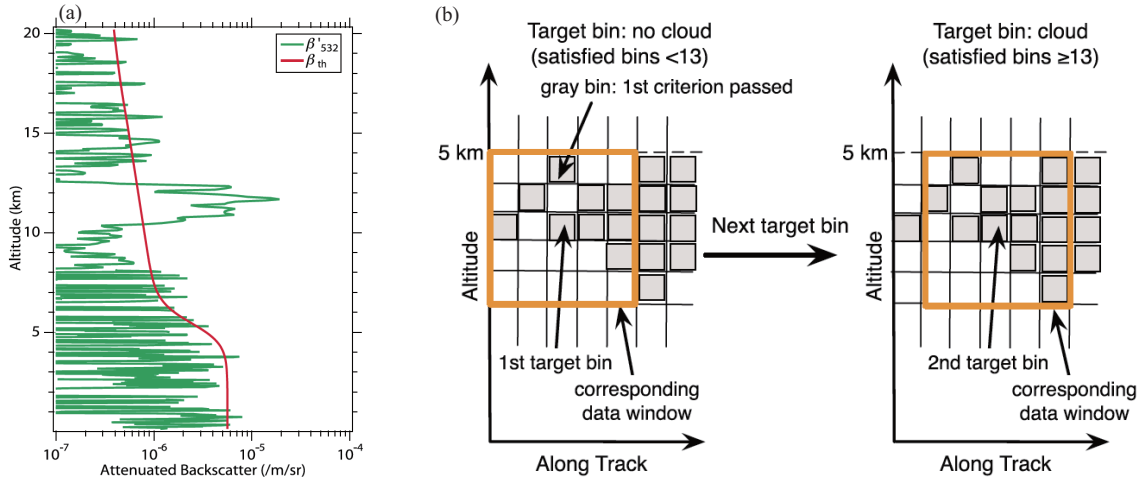

Figure 4.1: (a) A profile of the attenuated backscattering coefficient at 532 nm from CALIOP (green) and the threshold profile β_{th} used in the C2 scheme as the first criterion (red) on 8 October 2006 (~20°N latitude and ~27°W longitude). The resolutions are identical to those of the original CALIOP level 1B. (b) A graphical representation of the second criterion (i.e., spatial continuity test) used in the C2 scheme. Although the (left) first target bin is not considered to be cloud because less than half of the bins located in the (right) corresponding data window satisfied the first criterion, the second target bin is considered to be cloud. The resolutions are identical to those of the original CALIOP level 1B; the data window sizes are 5×5 bins at altitude < 5 km and 9×9 bins at altitude > 5 km (from *Hagihara et al.* [2010]).

4.1.2 MODIS cloud mask (MOD35) in MODIS level 2 products

MODIS measures radiance at 36 spectral bands including the infrared and visible bands with spatial resolutions from 0.25 to 1 km. The algorithm of the MODIS cloud mask for collection 5 uses several cloud detection tests with the spectral band data and surface temperatures of the Global Data Assimilation System (GDAS) model output [*Frey et al.*, 2008]. The individual cloud test threshold is based on the radiative transfer theory. Each test gives a confident level that is an indicator of clear skies. The confident level ranges in value from 0 (very low confidence of clear or high confidence of cloud) to 1 (high confidence of clear). The thresholds of tests are summarized in Appendix C. These detecting tests are grouped into five categories in accordance with similar cloud conditions. As described by *Ackerman et al.* [1998], the final cloud mask confidence (Q) is determined by the following equation:

$$Q = \sqrt[N]{\prod_{j=1}^{N} G_j}, \tag{4.1}$$

where

$$G_j = \min\left(F_{i,j}\right)_{i=1,m}, \tag{4.2}$$

where $F_{i,j}$ is the confidence level of each test, subscripts i and j are the test index and group index, respectively, m is the number of tests for a given group, and N is the number of groups (= 5). The term G_j denotes the minimum confidence for each group. The following four confident levels are determined for the output of the cloud mask: 1) confident clear ($Q > 0.99$), 2) probably clear ($0.99 > Q > 0.95$), 3) uncertain/probably cloudy ($0.95 > Q > 0.66$), 4) cloudy ($Q > 0.66$). This is a clear-sky conservative approach in the sense of detecting cloud pixels as many as possible by each test. This thesis relies on only the confident clear scene, and does not use the other scenes. To compare the cloud mask results with the CloudSat and CALIOP cloud masks, the cloud mask data of MODIS on board the Aqua satellite, which is part of the A-Train constellation, are used. The CloudSat MODIS-AUX product is the MODIS cloud mask data collocated in the CloudSat footprint. The product consists of the 1-km MODIS resolution data at a 3-pixel (across-track) by 5-pixel (along-track) grid. In this thesis, 9 pixels (i.e., a 3- by 3-pixel grid) from the pixel nearest to the CloudSat footprint are used to determine whether the misclassified cloud is included in the

CALIOP cloud mask. If all the pixels are labeled as confident clear, the MODIS cloud mask of the nearest pixel is regarded as the clear sky scene.

4.2 Discriminant analysis for detecting misclassified clouds in CALIOP cloud masks

This section describes the new method to discriminate misclassified clouds from the C2 cloud mask. Discriminant problems of two groups are solved by a discriminant analysis. As described in Chapter 2.4, the VFM cloud mask was developed based on the probability density function (PDF) algorithm [*Liu et al.*, 2004, 2009, 2010]. The observed layers are classified into clouds or aerosols depending on the possibility in the pre-established PDF. Basically, higher dimensional data make better classification, but the VFM mask has a limitation in regard to the classification of dust and clouds. *Chen et al.* [2010] incorporated brightness temperature differences from Imaging Infrared Radiometer (IIR) onboard CALIPSO into a discrimination model. Dust layers were successfully discriminated from cloud layers by the method. The BTD method using MODIS was also developed by the other studies [*Xie et al.*, 2010; *Naeger et al.*, 2013]. *Ma et al.* [2011] presented successful discrimination between dust and clouds by using a nonlinear support vector machine (SVM). The discriminant model was made based on a small number of training data at the limited area, and therefore the applicability of the model to the large amount of the CALIOP data is unclear. In addition, most of the studies introduced above selected the training data by the visual determination from satellite imagery, resulting in a lack of the objectivity and repeatability. Moreover, since the satellite imagery is available only in the daytime, training data cannot be derived from the nighttime data. Therefore, the true PDF of a statistical population cannot be obtained.

This thesis makes a discriminant analysis to discriminate misclassified clouds form the C2 cloud mask. Cloud layers detected by the C2 cloud mask are reclassified into clouds and misclassified clouds. Figure 4.2 illustrates the flow chart of the analysis. The reclassification in this method consists of two steps: 1) to make the discrimination model from training data and 2) to calculate the discriminant function for new cloud data. In the first step, the training data are determined by using the CloudSat cloud mask, MODIS cloud mask, and relative humidity (RH) from the ECMWF data. Here, the training data consist of two groups: 1) clouds and 2) misclassified clouds, and are extracted during a year from June 2006 to May 2007. Next, a linear discriminant function (LDF) is made as the discriminant model based on the training data. In the second step, the new cloud data detected by the CALIOP cloud mask are reclassified into clouds and misclassified clouds by the LDF test.

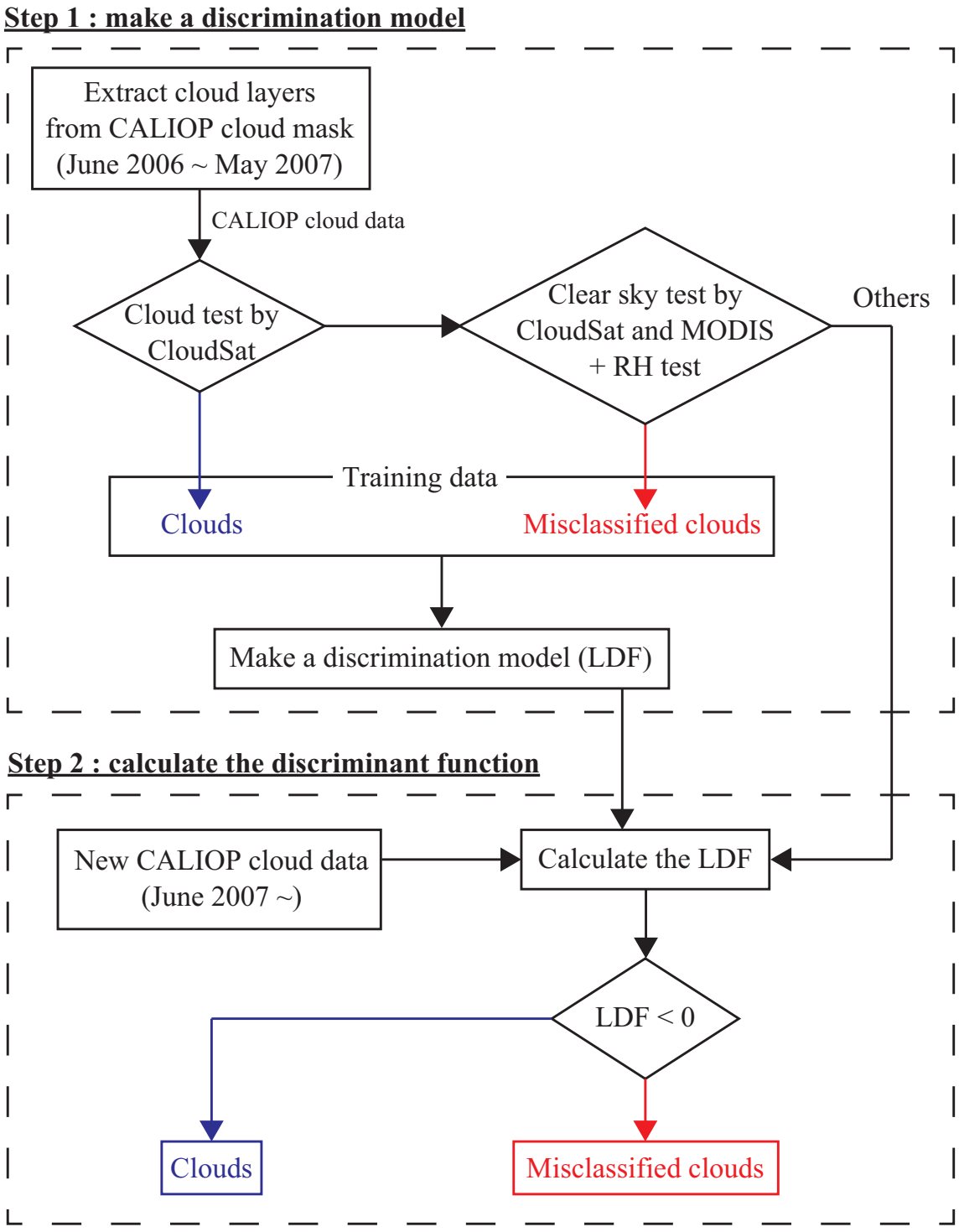

Figure 4.2: A flow chart of the discriminant analysis for the C2 cloud mask.

The LDF is derived from the ratio of the multi-dimensional PDFs of the two groups at the minimum misclassification rate [*Johnson and Wichern*, 2007]. The derivation of LDF is described in Appendix D. The LDF is described by variables x and coefficients C as follows:

$$LDF = C_0 + \sum_{i=1}^{k} C_i x_i \begin{cases} \geq 0 \Longrightarrow G_1 \\ < 0 \Longrightarrow G_2 \end{cases}, \tag{4.3}$$

where subscript i denotes the index of the variables, k is the dimension number of the variables, and G_1 and G_2 are the group 1 and 2, respectively. The superscript of the coefficients denotes the different types of the involved variables. The coefficients are umambiguously calculated by the averages and variance-covariance matrices of the PDFs of the two groups. If the LDF score is greater than 0 (below 0), the observed data are classified as G_1 (G_2). Here, the group 1 and 2 are clouds and misclassified clouds, respectively. This thesis uses five variables from CALIOP and IIR: 1) layer-mean attenuated backscattering coefficients ($\log_{10}\left(\overline{\beta'_{S,532}}\right)$, 2) layer-integrated volume depolarization ratios ($\overline{\delta_{S,v}}$), 3) layer-integrated attenuated color ratios ($\overline{\chi'_S}$), 4) layer top altitudes (z_{lt}), and 5) brightness temperature differences between 10.6 and 12.05 μm (BTD). The CALIOP and IIR data are averaged to have the same spatial resolutions as the C2 cloud mask prior to the calculation of the variables. The vertical and horizontal resolutions are 240 m and 1.1 km, respectively (the vertical resolution is only for the CALIOP data). The LDF score in Equation (4.3) is calculated for the observed data after the determination of the coefficients. The PDFs of the variables are determined by the training data.

The determination procedure of the training data is described below. The CALIOP cloud mask used in this section is the C2 cloud mask. The training data of clouds and misclassified clouds are referred to as cloud data and misclassified cloud data, respectively. The conclusive clouds and certainly misclassified clouds should be selected as the training data. The cloud data are determined by a cloud test using the C1 cloud mask. If the C1 cloud mask detects cloud bins within a cloud layer detected by the C2 cloud mask, the cloud layer is regarded as the cloud data. The misclassified cloud data are determined by a clear sky test using the C1 cloud mask, MODIS cloud mask, and relative humidity. The C1 and MODIS cloud masks are used to detect clear sky conditions. The clear sky in the C1 cloud mask means that the C1 cloud mask does not detect any cloud bins through the profile. If both the C1 and MODIS cloud masks return the clear sky condition, cloud layers detected by the C2 cloud mask are possibly misclassified. These cloud layers are regarded as a first candidate of misclassified clouds (hereafter referred to as FCMC).

Figure 4.3 shows the latitude-altitude cross section of the sample size (N) of FCMC data collected from June 2006 to May 2007. The altitude of FCMC data is derived from the layer top altitude. The high latitude regions ($> 50°$) are excluded from the analysis due to two reasons: 1) the cloud identification in the polar regions is difficult for MODIS [*Ackerman et al.*, 2008], and 2) optically thick aerosols are rarely observed at the high latitudes. The FCMC data below the height of 1 km above the surface elevation are excluded since surface clutter would affect the C1 cloud mask.

There are large amounts of sample sizes above 8 km for all seasons in Figure 4.3. These FCMC data may be considered as optically thin clouds that the CloudSat and MODIS cloud masks cannot detect. *Stephens et al.* [2002] reported that cirrus clouds having an optical thickness of $0.1 \sim 0.3$ are not detectable with CloudSat. *Ackerman et al.* [2008] showed that the limitation of the cloud optical thickness in the MODIS cloud mask is approximately 0.4. On the other hand, CALIOP can detect clouds having optical thickness of 0.01 [*McGill et al.*, 2007]. Another notable point is that sample sizes in the northern hemisphere are larger than those in the southern hemisphere below 8 km. The larger sample size is clearly recognized in the March–May (MAM) and June–August (JJA) seasons, when dust is frequently observed in the northern hemisphere. In this thesis, the southern hemisphere is excluded from the discriminant analysis because of the small sample size.

Optically thin clouds should be removed from FCMC data. A relative humidity test is additionally examined. The ECMWF data interpolated to each CloudSat bin are provided as an ancillary data of the standard CloudSat products. The relative humidity is calculated from the water vapor density and temperature data. If the temperature is below 0°C, the relative humidity with respect to ice is calculated. Figure 4.4 shows the vertical profiles of the sample size of FCMC data. The samples are the integrated value from 0° to 50°N. Three cases of the profiles have relative humidity below thresholds of 50%, 30%, and 10%.

The sample sizes decrease with the lower threshold of relative humidity. These thresholds result in the smaller sample sizes above 7 km than those below 3 km. It indicates that FCMCs with the lower relative humidity have the smaller influence from thin clouds compared to FCMC without the relative humidity test. To use the large sample size with minimum contamination from thin clouds, this thesis selects the profile whose relative humidity is less than 30%. In addition, the samples above 7 km in the selected profile are excluded from the FCMC data. The content rate of thin clouds is calculated by dividing the sample size at 7–8 km by the sample size at 1–7 km. The calculated content rate is approximately 4%, and the result has small effect on the form of the PDF.

To make misclassified cloud data more strictly, the horizontal distribution of FCMC data is

examined. The FCMC sample sizes for the four cases are shown in Figure 4.5. The conditions of sample sizes are the same as those in Figure 4.4, but the samples above 7 km are excluded. The samples of FCMC decrease with the lower threshold of the relative humidity. The samples are mainly distributed over arid regions in all the figures, and therefore dust is considered as the main source of misclassified clouds. This thesis restricts FCMC data to the dust region denoted by the red rectangle in Figure 4.5 (c). The FCMC data are finally selected as misclassified cloud data (i.e., 1 km $< z_{lt} <$ 7 km, RH $<$ 30%, 20°W $<$ longitude $<$ 120°E, and 0° $<$ latitude $<$ 50°N).

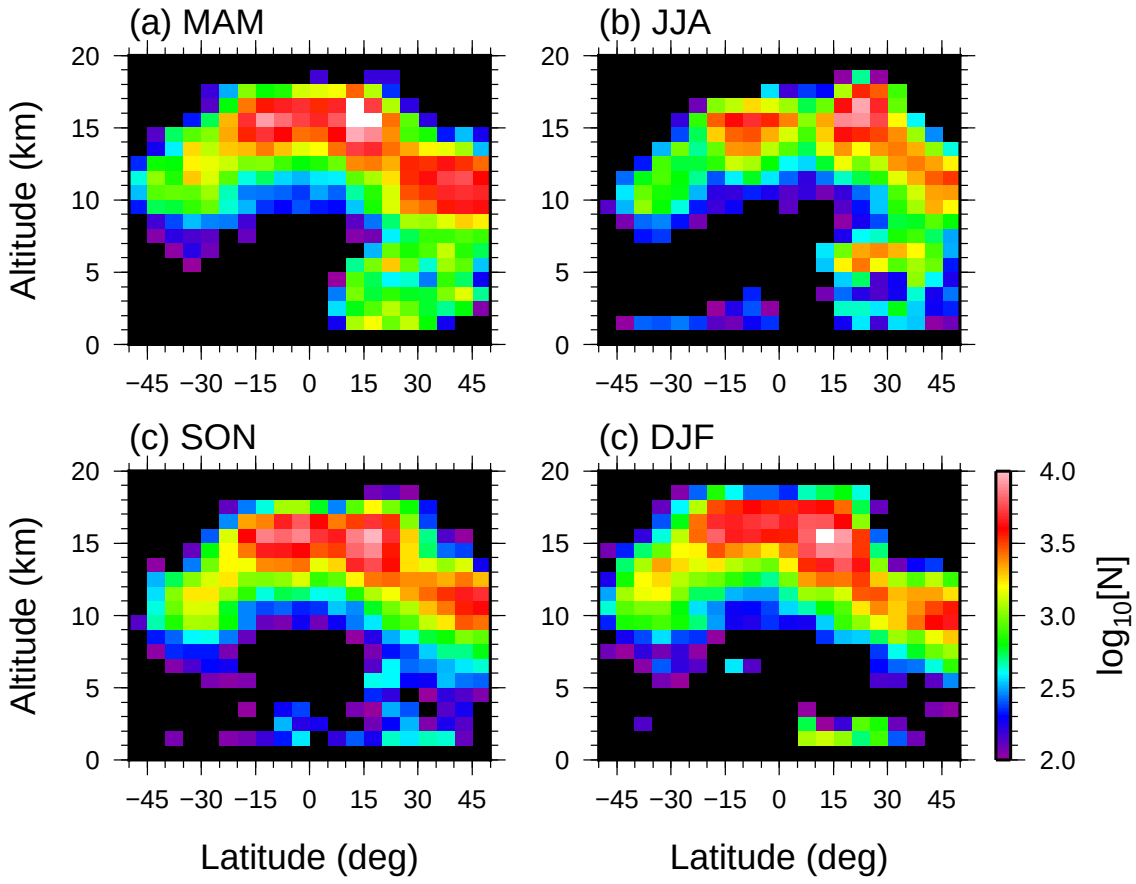

Figure 4.3: Latitude-altitude cross sections (5° × 1 km) of logarithmic sample number of the first candidate of misclassified clouds (FCMC) identified by the C2 cloud mask scheme during (a) March–May (MAM), (b) June–August (JJA), (c) September–November (SON), and (d) December–February (DJF) from June 2006 to May 2007. The high-latitude regions (> 50°) and altitudes lower than 1 km are not shown.

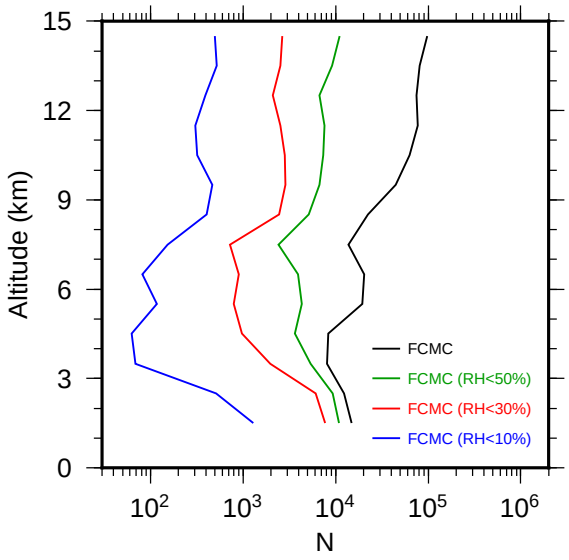

Figure 4.4: Vertical profiles of the sample number of FCMC. The green, red, and blue lines are additionally tested results by the relative humidity of 50%, 30%, and 10%, respectively. The samples are collected at each 1 km in altitude.

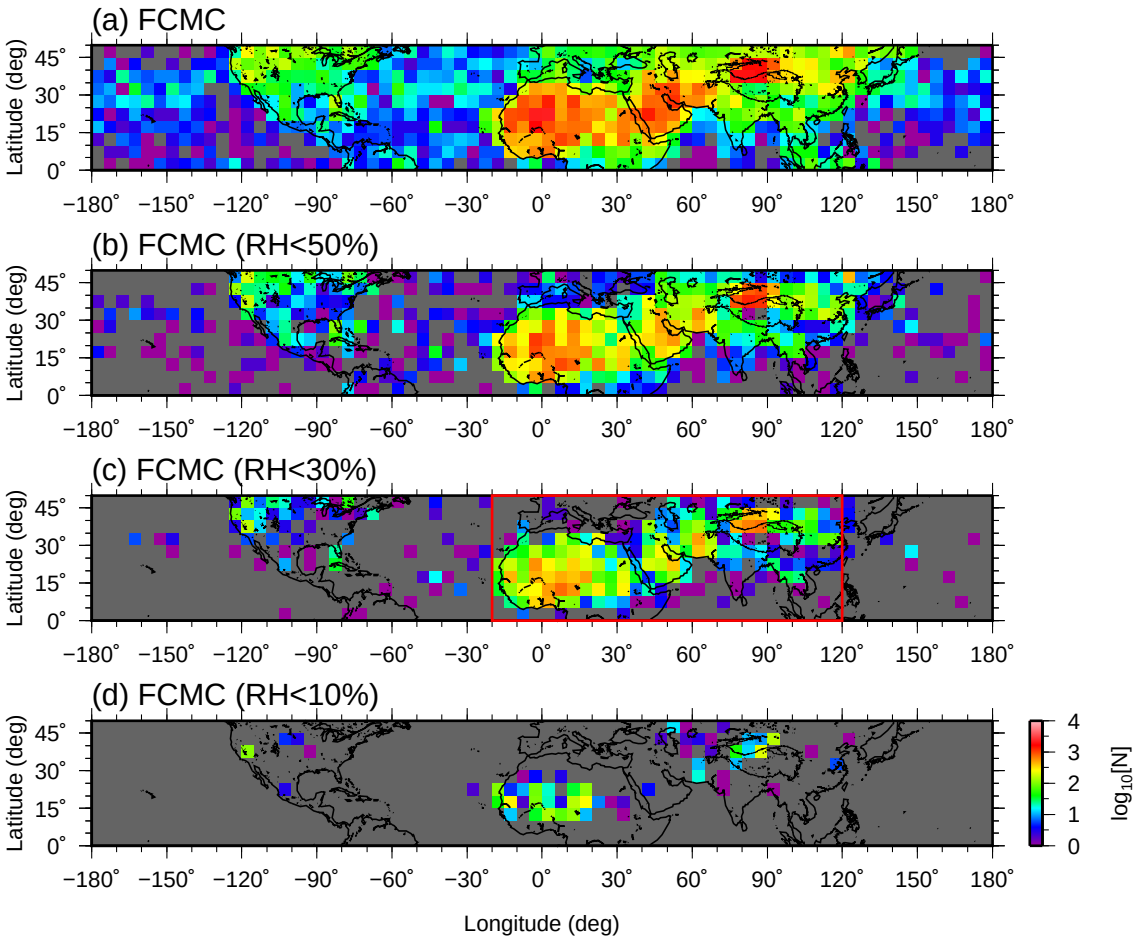

Figure 4.5: Longitude-latitude cross sections (5°×5°) of the logarithmic sample number of FCMC. The conditions of relative humidity are same as those in Figure 4.4. The sample number at each pixel is the integral value from 1 km to 7 km. The red rectangle in Panel (c) is the region used for the training data.

The collected training data is used to make the discrimination model (i.e., LDF). Table 4.1 shows the sample size of clouds, misclassified clouds (MC), and the others collected from June 2006 to May 2007. The others mean cloud layers detected by the C2 cloud mask, but not including clouds and misclassified clouds. The sample size of misclassified clouds is much smaller than that of clouds and the others. It should be noted that misclassified clouds would be partially included in the others. Although the sample size of misclassified clouds is relatively small, it is enough to form the PDF.

Figure 4.6 shows the PDF of the five variables ($\log_{10}\left(\overline{\beta'_{S,532}}\right)$, $\overline{\delta_{S,v}}$, $\overline{\chi'_S}$, z_{lt}, and BTD) incorporated in LDF. The variables of misclassified clouds are smaller than those of clouds except for the variable $\overline{\delta_{S,v}}$. The PDF of the variable $\overline{\delta_{S,v}}$ of clouds has two peaks at 0 and 0.2 (blue line in Figure 4.6 (b)). The peak at 0 would attribute to plate-like ice crystals (2D-plate), and the other would be caused by water droplets. The cloud and misclassified cloud data have the positive and the negative values of BTD, respectively. It can be explained by their inverse spectral dependence of refractive indices [*Gu et al.*, 2003; *Zhang et al.*, 2006].

The dotted lines in the figure denote the misclassified cloud data (MC′) collected without using the MODIS cloud mask. The notable difference between the MC and the MC′ is seen in the PDF of $\log_{10}\left(\overline{\beta'_{S,532}}\right)$. The PDF of the MC′ clearly includes the clouds since there is a peak at -1.0 of $\log_{10}\left(\overline{\beta'_{S,532}}\right)$. Therefore, the MODIS cloud mask is needed to collect the training data of certainly misclassified clouds. The PDFs of the others are similar to those of clouds. The PDF of the variable $\log_{10}\left(\overline{\beta'_{S,532}}\right)$ of the others has a peak at -2.1, which may be caused by misclassified clouds. Among the five variables, the variable $\log_{10}\left(\overline{\beta'_{S,532}}\right)$ has the smallest overlap between clouds and misclassified clouds. In contrast, the variable $\overline{\delta_{S,v}}$ has the largest overlap. The LDF is made based on the PDFs of clouds and misclassified clouds with minimum misclassification. The averages and common variance-covariance matrices of the five variables are summarized in Appendix E.

The coefficients of LDF in Equation (4.3) are estimated from the PDF results. The estimated coefficients and normalized coefficients by the standard deviations are shown in Table 4.2. Since the ranges of the values are different among the variables, the normalization is needed to understand the contribution of the variables to the LDF score. According to the normalized coefficients, the variables $\log_{10}\left(\overline{\beta'_{S,532}}\right)$ and z_{lt} have a large effect on the LDF score, whereas the variable $\overline{\delta_{S,v}}$ has small effect. These results are consistent with the PDF results. The negative value of the variable $\overline{\delta_{S,v}}$ reflects the smaller PDF of clouds as described above.

Table 4.1: Sample size of clouds, misclassified clouds (MC), and the others collected from June 2006 to May 2007. The others mean cloud layers detected by the C2 cloud mask, but not including clouds and misclassified clouds.

Sample	N ($\times 10^2$)	Percentage (%)
Clouds	19010	59.3
MC	181.04	0.57
Others	12841	40.1
Total	32032	

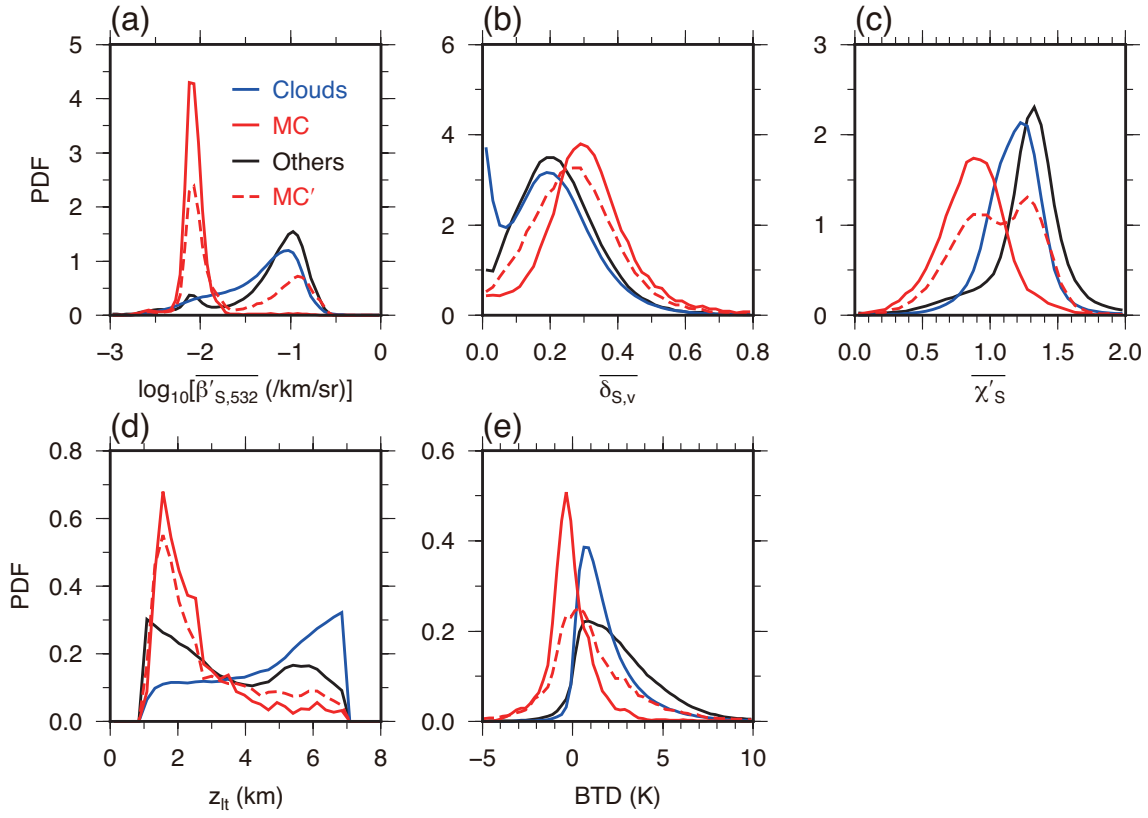

Figure 4.6: Probability density functions (PDF) of (a) layer-mean attenuated backscattering coefficients, (b) layer-integrated volume depolarization ratios, (c) layer-integrated total color ratios, (d) layer top altitudes, and (e) brightness temperature differences between 10.6 and 12.05 μm (BTD). The term MC denotes misclassified clouds, and MC′ denotes misclassified clouds determined without the MODIS cloud mask.

Table 4.2: Coefficients of the linear discriminant function (LDF) and values normalized by the standard deviations. The coefficients C1, C2, C3, C4, and C5 are for $\log_{10}\left(\overline{\beta'_{S,532}}\right)$, $\overline{\delta_{S,v}}$, $\overline{\chi'_S}$, z_{lt}, and BTD, respectively.

Coefficient	C_0	C_1	C_2	C_3	C_4 (km^{-1})	C_5 (K^{-1})
Value	−0.6654	4.9686	−2.8791	4.5227	1.3460	0.4775
Value*SD		2.1165	−0.4035	1.2746	2.2959	0.8372

4.3 Reclassification of cloud layers in CALIOP cloud masks

The constructed LDF is applied to new CALIOP cloud data that are independent from the training data. Two demonstrations and statistical results of the discriminant analysis for detecting misclassified clouds in the C2 cloud mask are presented in this section. First, a case of dust diffused over East Asia during the daytime on 31 March 2007 is shown in Figure 4.7. The sand storms were widely observed from 30 to 31 March at the ground stations in Mongolia and the northwest of China. From the Yellow Sea to the Chinese Continent, dust is recognized from the desert sand colors in the MODIS true color image. The negative BTD values over the Chinese Continent also confirm the existence of dust.

Along with the CloudSat/CALIPSO ground track (gray line in Figure 4.7 (a) and (b)), clouds are recognized in the latitudes of 38°N–39.5°N and the north of 42°N. The latitude-altitude cross section of the attenuated backscattering coefficient at 532 nm (Figure 4.7 (c)) shows that there are strong signals ($\beta'_{S,532} > 0.01$ /km/sr) throughout the latitudes. A data mask makes based on the C2 cloud mask is shown in Figure 4.7 (d). Dust regions labeled in the figure are determined by the volume depolarization ratio and attenuated color ratio. The determination method of dust is described in Appendix F. The detected clouds embedded in the dust layers below the altitudes of 5 km are considered as misclassified clouds from the MODIS images. The relative humidity (contours in Figure 4.7 (d)) also shows the dry air in the regions of misclassified clouds.

The result of the LDF score for the cloud layers is shown in Figure 4.8 (a). The positive (red) and negative (blue) values indicate clouds and misclassified clouds, respectively. The misclassified regions are successfully detected by LDF even though the BTD values differ depending on the underlying surfaces (i.e., the Chinese Continent and the Yellow Sea). This demonstration indicates that the method incorporating the BTD for the discrimination of dust from clouds is effective even for dust above water surfaces. The corrected data mask after the calculation of LDF is shown in Figure 4.8 (b). If the layer-integrated volume depolarization ratio of misclassified clouds (negative LDF) is larger than 0.06, misclassified clouds are identified as dust. The corrected data mask agrees with the results in the MODIS images.

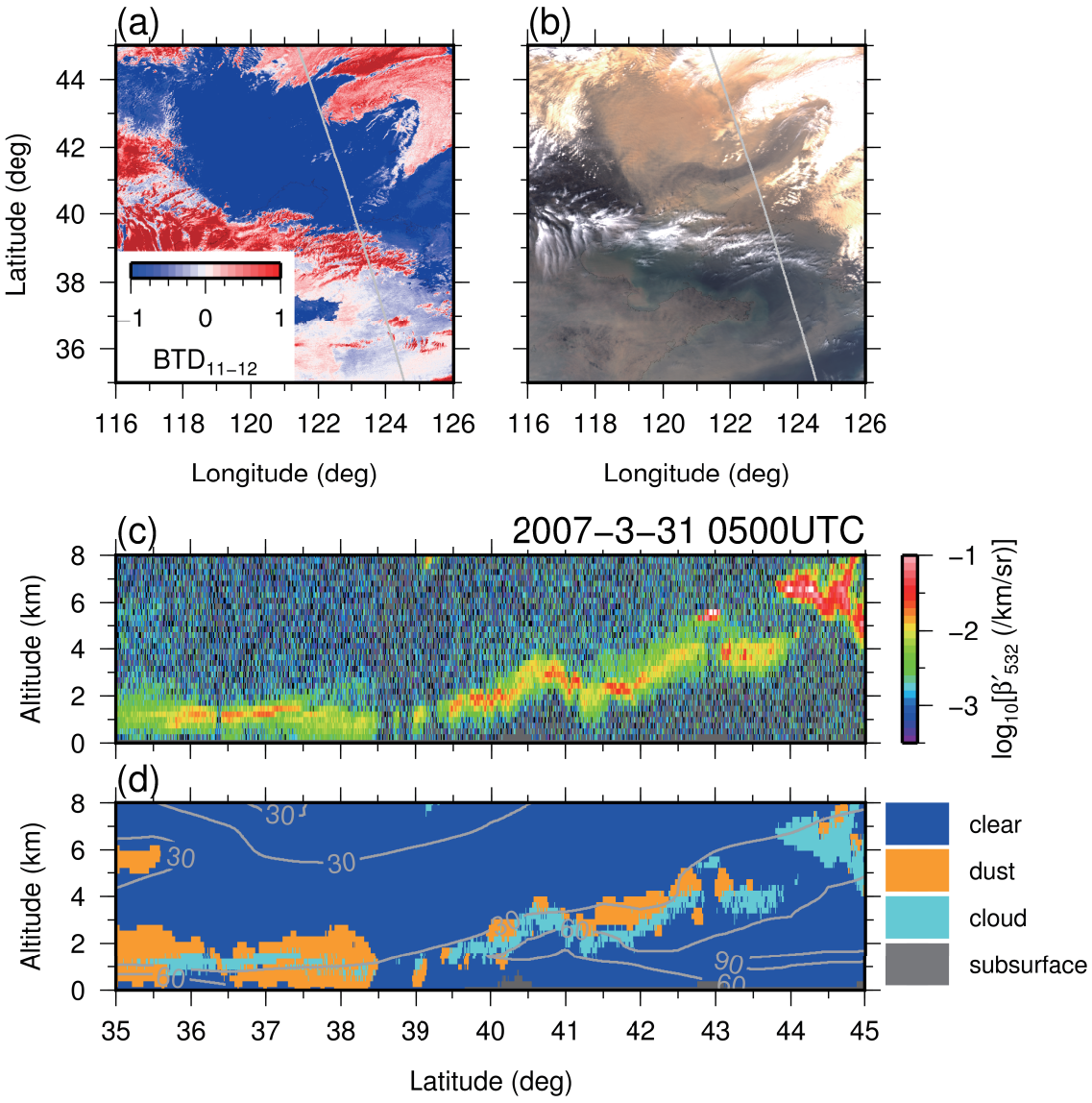

Figure 4.7: An example of dust misclassification as clouds during the daytime on 31 March 2007. Panels (a) and (b) are the brightness temperature difference (BTD) and true color image, respectively, derived from the MODIS/Aqua at 0500 UTC. Latitude-altitude cross sections of (c) attenuated backscattering coefficients at 532 nm, (d) data mask based on the C2 cloud mask. The contour lines in Panel (d) denote the relative humidity. The data mask classifies the observed bins into four categories: clear, dust, cloud, and subsurface.

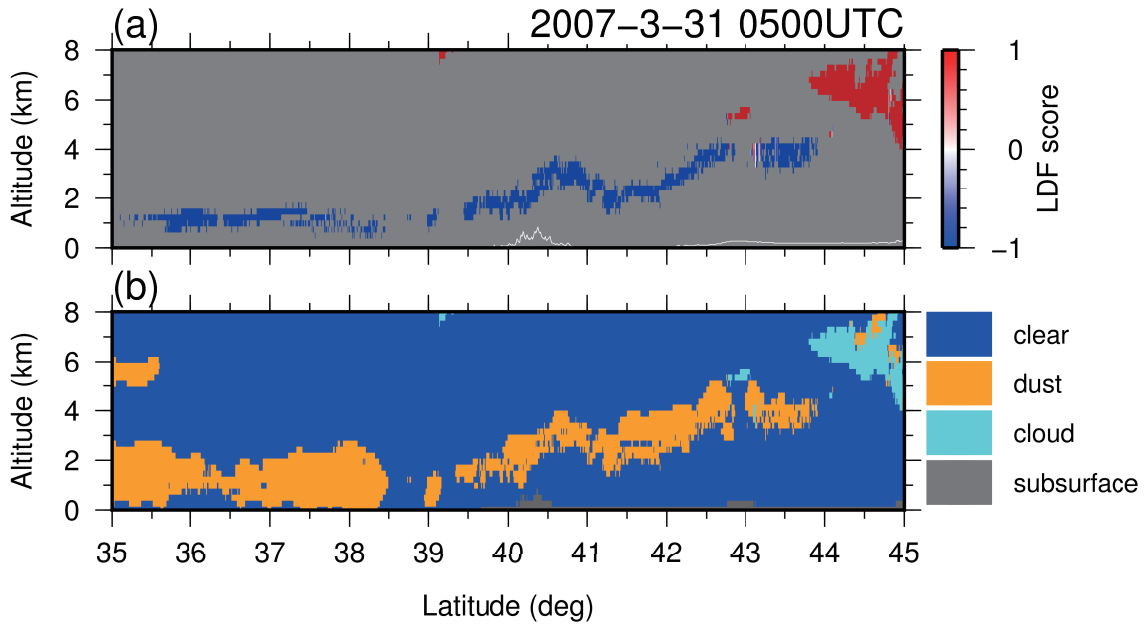

Figure 4.8: A demonstration of the refinement of the C2 cloud mask in the case of Figure 4.7. (a) Linear discriminant function (LDF) scores for the detected clouds and (b) modified data mask after the cloud screening.

The other case of misclassified clouds during the nighttime on 22 June 2007 is shown in Figure 4.9. The MODIS cloud mask (Figure 4.9 (b)) shows clear sky conditions over the west of North Africa. Although the cloudy conditions are reported in the latitudes from 10°N to 12°N along with the ground track, dust would exist there because of the negative BTD values. In Figure 4.9 (c), there is a layer having large attenuated backscattering coefficients from 3 to 6 km through the latitudes. The data mask shows that clouds are embedded in the dust layer. According to the MODIS results and low relative humidity in the cloud regions, clouds above 3 km at the latitudes from 7.5°N to 16.5°N are possibly misclassified.

The result of the LDF score is shown in Figure 4.10 (a). The possible misclassified regions are successfully reclassified as misclassified clouds that have the negative LDF scores. The corrected data mask also agrees with the BTD values of MODIS. For comparison of the LDF applicability, LDF results developed by *Chen et al.* [2010] (referred to as LDF′ hereafter) are shown in Figure 4.10 (c). Most of the possible misclassified regions have the positive values of LDF′. The corresponding data mask is rarely different from the data mask before the LDF correction. It indicates that LDF′ is less effective than LDF in this thesis. The reason of the false detection by LDF′ is that the training data during the nighttime are not collected in *Chen et al.* [2010], resulting in the uncertainty of true PDFs which are needed for making the greatest discrimination model. The two demonstrations of the discriminant analysis result in the successful reclassification of dust over land and sea during the daytime and nighttime.

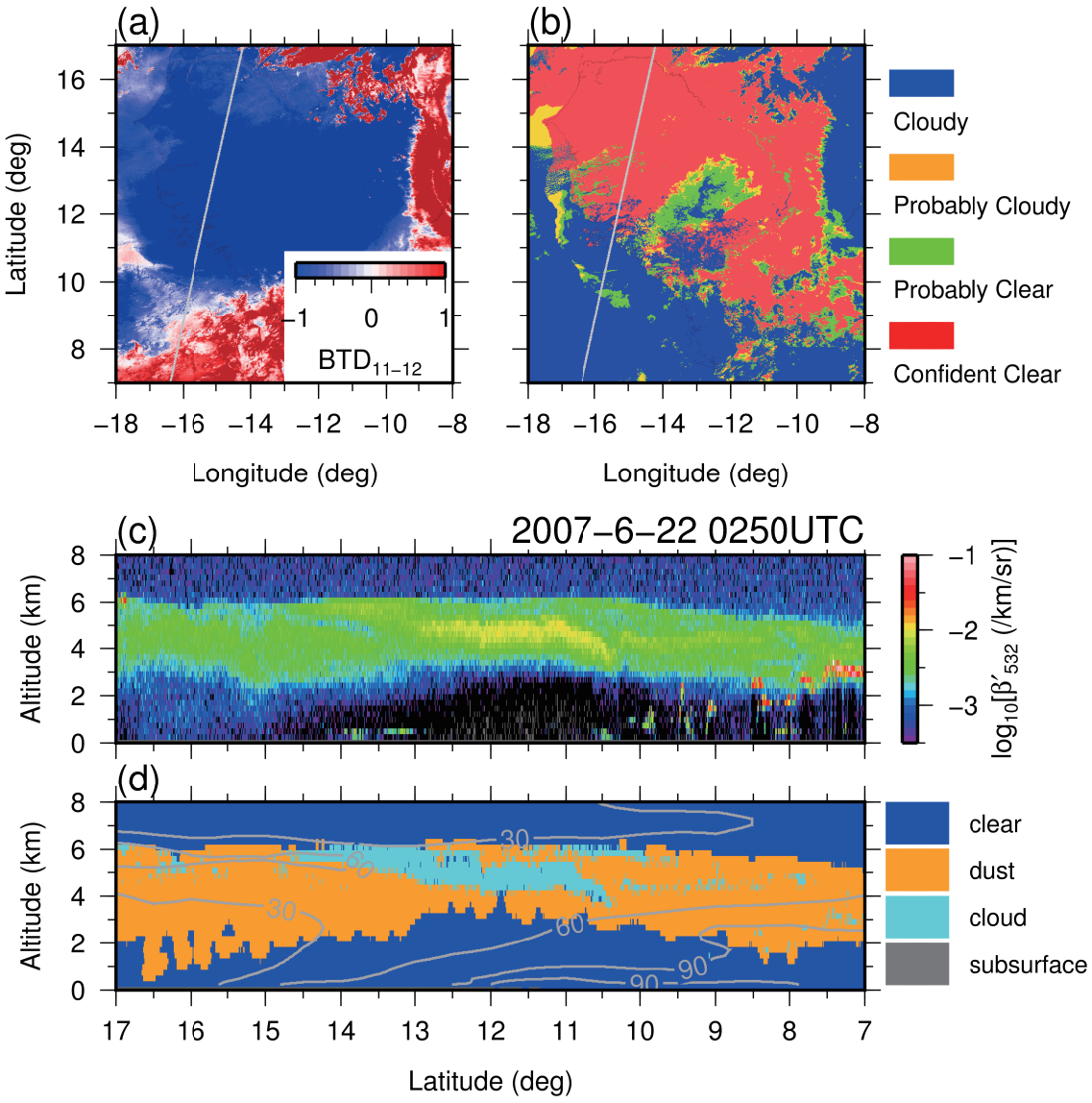

Figure 4.9: An example of dust misclassification as clouds during the nighttime on 22 June 2007. The figures are the same as Figure 4.7, but the MODIS cloud mask is shown in Panel (b) instead of the true color image.

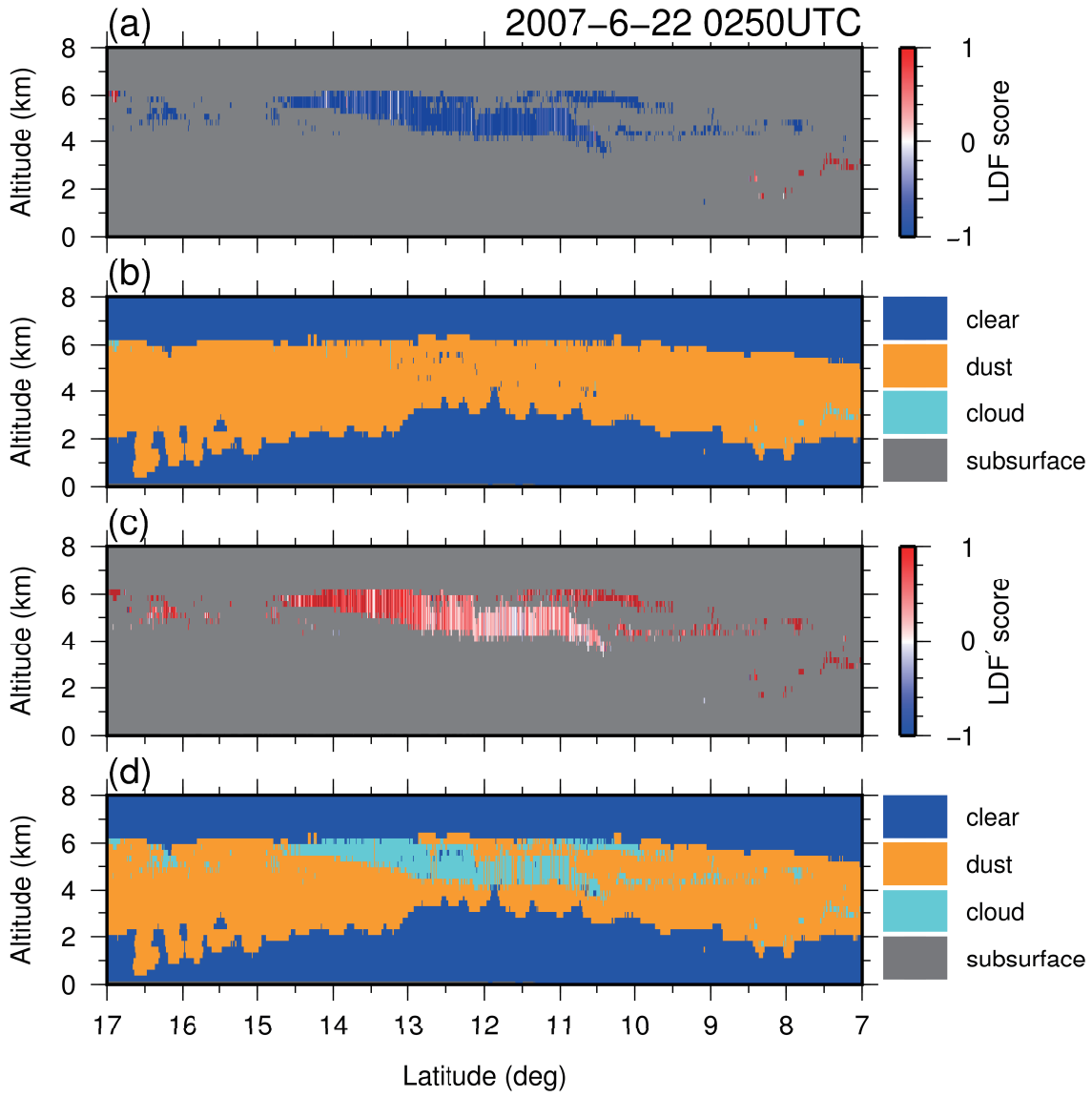

Figure 4.10: Demonstrations of the refinement of the C2 cloud mask in the case of Figure 4.9. (a) Linear discriminant function (LDF) scores by the method of this thesis, (b) modified data mask after the cloud screening from the results of Panel (a), (c) LDF′ scores by the method of *Chen et al.* [2010], and (d) modified data mask after cloud screening from the results of Panel (c).

In order to investigate the LDF accuracy, a statistical analysis is conducted. Table 4.3 shows reclassification results of cloud layers in the C2 cloud mask during June–August 2007. The cloud layers are divided into clouds, misclassified clouds, and the others data by the same procedure as the collection of training data. The accuracy rates of the LDF classification are 96.5% and 91.7% for clouds and misclassified clouds, respectively. The others data are reclassified as misclassified clouds by 12.1%. The number of clouds classified as misclassified clouds by LDF (LDF < 0) accounts for 25.8% of total misclassified clouds. On the other hand, the number of misclassified clouds classified as clouds by LDF (LDF > 0) accounts for 0.4% of total clouds. The estimated content rate of misclassified clouds in the C2 cloud mask is approximately 6%. Here, the misclassification by LDF (i.e., clouds but having negative LDF scores) is excluded.

The accuracy rates of the LDF classification for different variables are calculated to derive optimal LDF. Table 4.4 shows the accuracy rates of optimal variables at the number of variables changing from 1 to 5. If only $\log_{10}\left(\overline{\beta'_{S,532}}\right)$ is used for making LDF, the accuracy rates are 87.9% and 95.5% for clouds and misclassified clouds, respectively. The accuracy rates increase with the additional variables. The highest accuracy rates are derived from the four variables except for $\overline{\delta_{S,v}}$. It indicates that $\overline{\delta_{S,v}}$ is not the effective parameter for the LDF classification. The possible reason is that the depolarization ratio of clouds differs depending on the cloud particle types. Randomly oriented ice crystals (3D-ice) have a larger depolarization ratio (~0.5) than misclassified clouds, but plate-like ice crystals (2D-plate) and water droplets have the lower depolarization ratio than misclassified clouds. Therefore, it is difficult to discriminate misclassified clouds from all the cloud particle types. The LDF coefficients that exclude $\overline{\delta_{S,v}}$ are presented in Appendix G. The exclusion of the depolarization ratio enables us to use LDF for cloud layers even after a change of the view angle of CALIPSO from 0.3° to 3.0° on 28 November 2007. The view angle change decreases (increases) the observation frequency of 2D-plate (3D-ice) [*Hu et al.*, 2009], resulting in the change of the depolarization ratio.

Table 4.3: Summary of the accuracy rate of linear discriminant function (LDF). The reclassification of the cloud layers in the C2 cloud mask is performed during June–August 2007. The cloud layers are divided into clouds, misclassified clouds (MC), and the others data by the same procedure as the collection of training data, and are compared to the classification results by LDF.

LDF Classification	Clouds	MC	Others	Sum
$LDF > 0$	444009 (96.5%)	275 (8.3%)	317396 (87.9%)	761680
$LDF < 0$	16260 (3.5%)	3031 (91.7%)	43835 (12.1%)	63126
Total	460269	3306	361231	824806

Table 4.4: Accuracy rates of optimal variables at the number of variables changing from 1 to 5. Clouds and misclassified clouds (MC) data are determined by the same procedure as the collection of training data.

Variables	Accuracy rate (%) Clouds	MC
$\log_{10}\left(\overline{\beta'_{S,532}}\right)$	87.9	95.5
$\log_{10}\left(\overline{\beta'_{S,532}}\right)$, z_{lt}	93.1	91.4
$\log_{10}\left(\overline{\beta'_{S,532}}\right)$, z_{lt}, BTD	95.8	92.0
$\log_{10}\left(\overline{\beta'_{S,532}}\right)$, $\overline{\chi'_S}$, z_{lt}, BTD	96.6	92.1
$\log_{10}\left(\overline{\beta'_{S,532}}\right)$, $\overline{\delta_{S,v}}$, $\overline{\chi'_S}$, z_{lt}, BTD	96.5	91.7

4.4 Effect of misclassifications in CALIOP cloud masks on vertical profiles of dust extinction coefficients

Dust extinction coefficients are retrieved after screening of clouds detected by the cloud mask in the CALIOP data analysis. If dust in practice is misclassified as clouds, the misclassified regions are eliminated from the retrieval, causing the sampling bias. Since dust misclassified as clouds is expected to have a large extinction coefficient, the retrieved dust extinction coefficient without the misclassified regions result in smaller. The dust extinction coefficient should be reevaluated after the improvement of the CALIOP cloud mask.

In this thesis, dust extinction coefficients are retrieved by the following procedure. First, the nighttime CALIOP signals are averaged to have the vertical and horizontal resolutions as the C2 cloud mask. Secondly, dust is detected by the method described in Appendix F. Thirdly, the average of signals are calculated for every five profiles. The horizontal resolution of the averaged profiles results in 5.5 km. The averaged profiles are used in the retrieval only if dust bins are detected in all the five profiles. The signals below the top of clouds are excluded. The sky conditions of the averaged profiles corresponds to a combination of the "cloud-free" and "above cloud". Finally, the retrieval of the dust extinction coefficient is performed.

To eliminate signals below the top of clouds, two cloud masks are used for the cloud detection. One of the cloud masks is the C2 cloud mask, and the other is a C2′ cloud mask that is improved by LDF. The clouds detected by the C2′ cloud mask have the positive LDF values. If clouds detected by the C2 cloud mask have the negative LDF values and the volume depolarization ratio of larger than 0.06, the clouds are regarded as dust in the C2′ cloud mask. The LDF calculation is also applied to clouds below 1 km from the ground level. The dust extinction profiles using the two cloud masks are compared each other.

The lidar ratio used in the retrieval is 40 sr at 532 nm. This value is identical to an initial lidar ratio in the CALIOP products. Since the solution of the forward inversion is unstable as shown in Chapter 2.3, the lidar ratio should be adjusted to retrieve the true extinction coefficients. If the retrieved extinction coefficients are diverged in the positive direction, the lidar ratio is reduced, and the retrieval is restarted using the reduced lidar ratio. The divergence in the positive direction is detected by testing if the successive extinction coefficients are not converging [*Young and Vaughan*, 2009].

The profiles of dust extinction coefficients are calculated at arid regions. Figure 4.11 (a) shows the frequency of misclassified clouds (negative LDF values) in the $5^\circ \times 5^\circ$ grid scale from March

2007 to August 2007. The misclassification number is the largest at the Taklimakan Desert, followed by the west of North Africa. The extinction profiles are investigated in three regions denoted by the red rectangles: 1) the Sahara Desert (5°W–5°E and 15°N–25°N), 2) the Middle East (37.5°E–55°E and 17.5°N–32.5°N), and 3) the Taklimakan Desert (75°E–90°E and 35°N–42.5°N).

The results of extinction profiles are shown in Figure 4.11 (b)–(d). In all the regions, the extinction coefficients using the C2′ cloud mask are clearly larger than those using the C2 cloud mask especially below 3 km from the ground level. The maximum ratios of the C2′ to C2 profiles are factors of 1.55, 1.77, and 2.6 in the Sahara Desert, the Middle East, and the Taklimakan Desert, respectively. Therefore the effect of cloud misclassifications on the extinction profile is the largest in the Taklimakan Desert. The fraction of the misclassified cloud profiles in the calculated extinction profiles is 34.6% (below 2 km) in the Taklimakan Desert. The errors caused by the misclassification of clouds as dust are ~3.0% below 2 km.

The retrieved extinction profile is compared to the CALIOP level 3 products. The CALIOP level 3 products provide the monthly averaged data that are applied to several quality checks before the averaging of the level 2 products [*Winker et al.*, 2013]. The combined (cloud-free + above cloud) sky condition and nighttime product is used for the CALIOP level 3. Only the extinction profiles classified as dust are used. Figure 4.12 shows the extinction profiles from the C2′ cloud mask and the CALIOP level 3 in the Taklimakan Desert. The extinction profile of the C2′ cloud mask below 2 km is approximately 2 times larger than that of the CALIOP level 3. The extinction profile of the CALIOP level 3 is based on VFM. In order to confirm the correctness of the retrieval process in this thesis, the retrieval of the extinction profile using the VFM cloud mask is performed. The extinction profile using the VFM cloud mask (dotted line in the figure) agrees with that of the CALIOP level 3 even though the finally selected lidar ratios in the CALIOP level 3 may be different from those in this thesis. Therefore the difference of the C2′ and the CALIOP level 3 profiles is mainly due to the difference of the cloud screening processes.

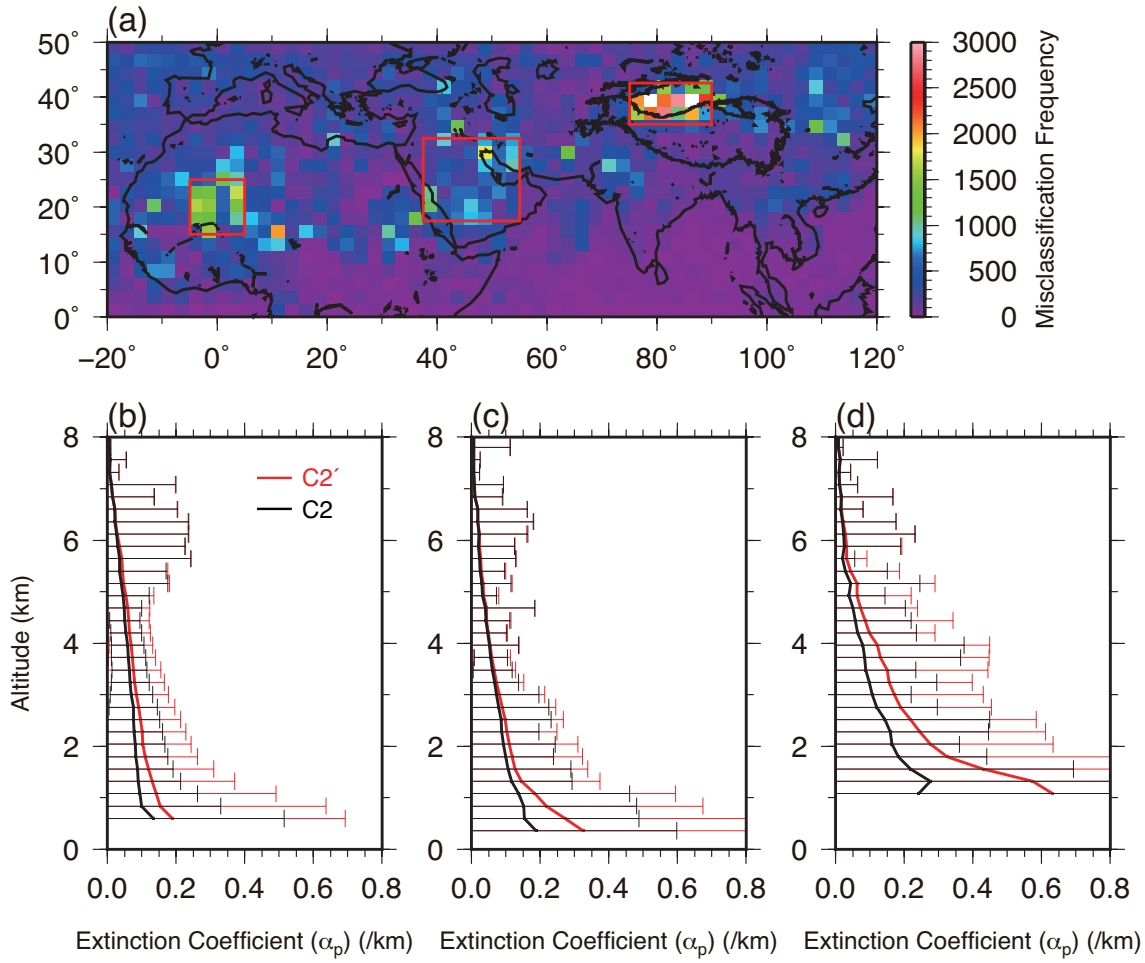

Figure 4.11: (a) A longitude-latitude cross section of the misclassification frequency (negative linear discriminant function (LDF) scores) and (b)–(d) profiles of dust extinction coefficients before (C2) and after (C2′) the refinement of the C2 cloud mask in the regions of the Sahara Desert, the Middle East, and the Taklimakan Desert denoted by the red rectangles in Panel (a). The error bar denotes the standard deviation.

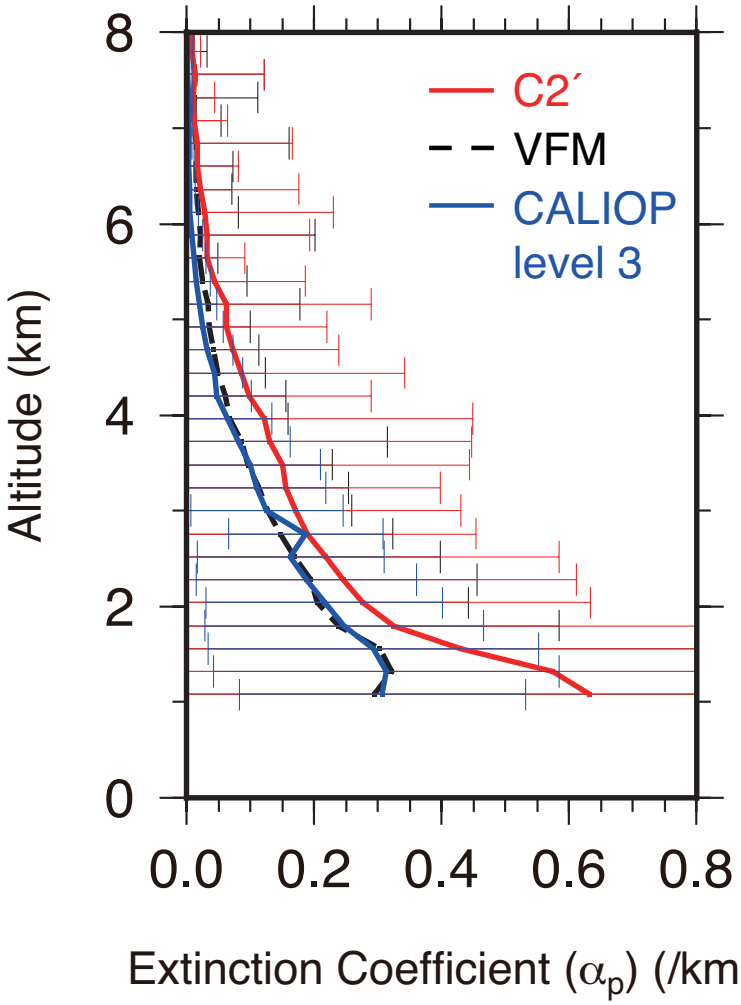

Figure 4.12: Profiles of dust extinction coefficients for the C2′, VFM cloud masks, and CALIOP level 3 products. The error bar denotes the standard deviation.

Aerosol optical depth (AOD) calculated from the extinction profiles using the C2′ cloud mask and the CALIOP level 3 are 1.03 and 0.63, respectively. *Ma et al.* [2013] revealed that AOD of the CALIOP level 3 aerosol products are ~0.25 smaller than that of MODIS in Northwest China during the dust season in 2007. The AOD reported at the center of the Taklimakan Desert (Tazhong, 83.7°E and 39°N) are approximately 0.7 during March–August from 2004 to 2008 [*Che et al.*, 2013]. The morning and evening AOD at Tazhong was ~0.1 larger than daytime AOD. AOD of the CALIOP level 3 has the smaller values (negative bias) compared to reported AOD although AOD of our study may be too large.

The misclassification of dust as clouds can affect the extinction profiles since the misclassified regions have the strong backscattering. An example of the misclassification in the VFM cloud mask on 18 May 2007 is shown in Figure 4.13. *Yumimoto et al.* [2009] analyzed this dust event and revealed that a long-range transportation of dust from the Taklimakan Desert to North American Continent. The data mask using the C2′ cloud mask (Figure 4.13 (b)) shows correct discriminations between dust and clouds even at the strong signal regions from 38°N to 40.5°N. On the other hand, the data mask using the VFM cloud mask (Figure 4.13 (c)) shows the significant misclassification of dust as clouds through the latitudes.

The horizontal resolutions of the cloud layers in VFM are shown in Figure 4.13 (d). Most of the misclassified regions from 38°N to 40.5°N have the resolution of 333 m. The 333 m resolution is not applied to the cloud-aerosol discrimination (CAD) test in the CALIOP products (i.e., the layers detected at 333 m resolution are automatically classified as clouds). Therefore, the misclassifications at the 333 m resolution can cause smaller extinction coefficients. The dust regions at higher than 41°N are mostly misclassified as clouds. As pointed out in Chapter 3.4, dust in VFM is frequently misclassified at higher than 40°N. Clouds embedded in the dust layer are shown at higher than 42.5°N from the C2′ cloud mask (Figure 4.13 (b)). The top of the clouds in the VFM cloud mask is at higher than 6 km in altitude. The dust above the clouds is misclassified as clouds at the 1 km and coarser resolutions. The CALIOP products identify these mixed layers (aerosols + clouds) as clouds to avoid cloud contamination of the aerosol dataset [*Winker et al.*, 2009], but the misclassified dust in the mixed layers can cause the negative bias in the dust extinction coefficients.

Schuster et al. [2012] suggested that the negative bias of AOD from the CALIOP products is caused by the small values of the dust lidar ratio in the Sahara regions. However, smaller AOD from the CALIOP products in the Taklimakan Desert is unclear whether it is caused by the small lidar ratio or not due to an extreme lack of the observation of the lidar ratio. The dust lidar ratio estimated by the simultaneous observation with the ground-based lidar is 42.0 sr at 532 nm in the Taklimakan

Desert (Chapter 3). The smaller AOD is not explained even if the lidar ratio of 42.0 sr is used in the retrieval using the VFM cloud mask. This thesis suggests that the dust misclassifications as clouds in the C2 and VFM cloud masks have a significant effect on the retrieved extinction coefficients especially in the Taklimakan Desert and suggests that cloud layers at 333 m resolution in the VFM cloud mask should be reclassified into cloud or aerosol layers to avoid the negative bias of the dust extinction coefficient.

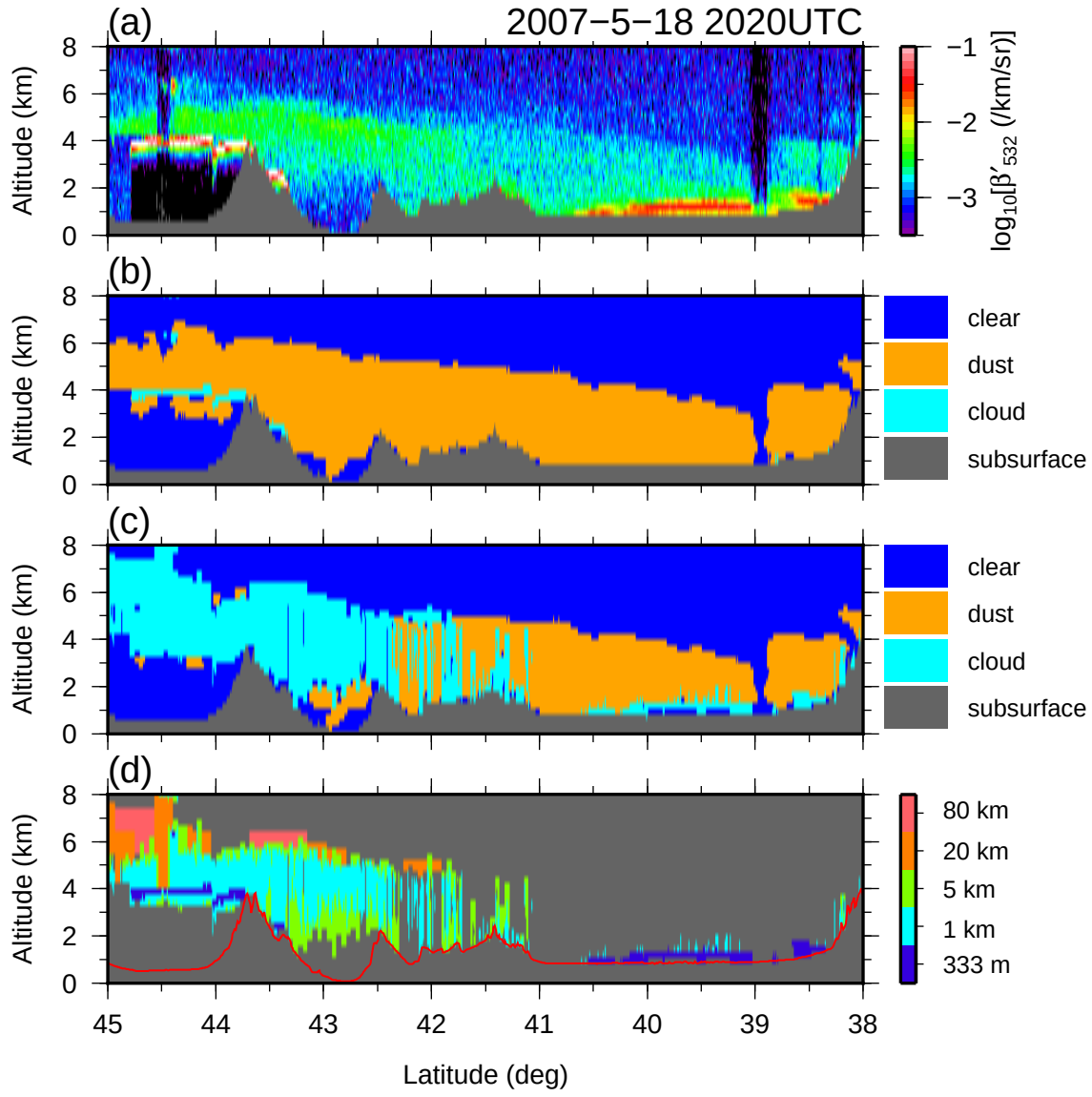

Figure 4.13: An example of the dust misclassifications as clouds in the VFM cloud mask. Latitude-altitude cross sections of (a) attenuated backscattering coefficients at 532 nm, (b) data mask based on the C2′ cloud mask, (c) data mask based on the VFM cloud mask, and (d) horizontal resolution in the VFM cloud mask.

5 Conclusion

This thesis presents a study for better estimation of the vertical profiles of dust optical properties by Mie-scattering lidar.

First, the extinction-to-backscattering ratio (lidar ratio) of dust is estimated in the Taklimakan Desert, where the lidar ratio has not been investigated. This thesis estimates the lidar ratio by simultaneous observations with Aksu-lidar and CALIOP on 23 March 2009. The estimated lidar ratios at 532 and 1064 nm are 42.0 and 45.9 sr, respectively. The 532 nm lidar ratio is 3–24% smaller than those of Asian dust reported previously. The errors in the lidar ratio estimation are 9.5% at 532 nm and 41.6% at 1064 nm. Error analysis shows that the estimation of the 1064 nm lidar ratio strongly depends on signal calibration errors at 1064 nm.

Using the estimated lidar ratios, dust optical properties are retrieved from the Aksu-lidar data. The particle depolarization ratio of dust is from 0.3 to 0.38, and the particle color ratio of dust is from 0.5 to 1.0. The small depolarization and color ratios on March 25 suggest that spherical and fine particles are advected from the surrounding regions (and/or that coarse particles settle due to the dry deposition). A comparison of Aksu-lidar to CALIOP level 2 data shows that the backscattering coefficients of the CALIOP level 2 data are 21% smaller than those of this thesis. The smaller backscattering coefficients are caused by the misclassification of dust as clouds in the VFM.

Secondly, the C2 cloud mask is improved by discriminant analysis. This thesis tries to discriminate misclassified clouds (mainly dust) from clouds detected by the C2 cloud mask. LDF is used as the discriminant model. The training data are collected by tests with the C1 cloud mask, MODIS cloud mask, and relative humidity. Five variables $\log_{10}\left(\overline{\beta'_{S,532}}\right)$, $\overline{\delta_{S,v}}$, $\overline{\chi'_S}$, z_{lt}, and BTD are used in LDF.

The reclassifications of the C2 cloud mask are demonstrated in two cases over land and water surfaces during day and night. The discriminations of dust from clouds are successfully performed in every case, even though the BTD varies depending on surface characterization [*Ackerman*, 1997]. The discrimination model of *Chen et al.* [2010] worked inadequately during the nighttime since the training data were not collected during the nighttime due to satellite imagery limitations. The LDF of this thesis makes based on training data collected during both night and day, resulting in the better discrimination even during the nighttime.

The accuracy rates of the LDF classification are 96.5% and 91.7% for clouds and misclassified clouds, respectively. The estimated content rate of misclassified clouds in the C2 cloud mask is ap-

proximately 6% in the region of 20°W–120°E and 0°–50°N during June–August 2007. An analysis changing the combinations of variables in the LDF reveals that the depolarization ratio $\overline{\delta_{S,v}}$ is not effective for discrimination and can be excluded from LDF. The exclusion of the depolarization ratio enables us to use modified LDF (i.e., using remained four variables) for clouds even after view angle of CALIPSO changed from 0.3° to 3.0° on 28 November 2007.

Vertical profiles of dust extinction coefficients before and after the improvement of the C2 cloud mask are investigated. Extinction coefficients using the improved cloud mask (C2′) are clearly larger than those using the C2 cloud mask, especially below 3 km from ground level. The maximum ratios of the C2′ to C2 profiles of extinction coefficients are factors of 1.55, 1.77, and 2.6 in the Sahara Desert, the Middle East, and the Taklimakan Desert, respectively. The fraction of misclassified clouds in observed dust is 34.6% (below 2 km) in the Taklimakan Desert.

A comparison of our resutls to CALIOP level 3 products indicates that the extinction profile using the C2′ cloud mask is ~2 times larger than that of the CALIOP level 3 in the Taklimakan Desert. The difference between the C2′ and the CALIOP level 3 profiles is mainly due to the different cloud screening processes. AODs calculated from the extinction profiles using the C2′ cloud mask and the CALIOP level 3 are 1.03 and 0.63, respectively. AOD reported by previous studies using the MODIS and the ground-based measurements in the Taklimakan Desert indicates that AOD of the CALIOP level 3 has negative bias.

This thesis suggests that the negative bias of AOD in the Taklimakan Desert is mainly caused by misclassifications of dust as clouds in the VFM cloud mask. This thesis also suggests that cloud layers at 333 m resolution in the VFM cloud mask should be reclassified into cloud or aerosol layers to avoid the negative bias of dust extinction coefficients. In addition, measurements of dust lidar ratios are needed to reduce uncertainty in the extinction retrieval. Although the Taklimakan Desert is the one of the major sources of dust, long-term measurements of dust lidar ratios have not been conducted. This thesis provides information on dust lidar ratios in the Taklimakan Desert. Furthermore, the improvement of the CALIOP cloud mask reduces uncertainty in the retrieval of dust extinction coefficients. These results would contribute to improvements in accuracy of chemical transport models.

Further studies will be continued to refine the vertical profiles of dust optical properties. Optically thin clouds that are not detected by the C2 cloud mask should be detected in order to avoid cloud contamination in the aerosol dataset. Optically thin clouds tend to be misclassified as dust due to the large depolarization ratio. The high spectral resolution lidar that will be on board the

EarthCARE satellite can measure extinction and backscattering coefficients independently (i.e., lidar ratios can be derived). Since dust lidar ratios are two times larger than cloud lidar ratios [*Sakai et al.*, 2003], discrimination between dust and clouds would improve. In addition, the cloud profiling radar (CPR) that will be on board the EarthCARE satellite is more sensitive than the CPR on board CloudSat, and therefore misclassified clouds would be more detectable.

Acknowledgments

The author would like to express his greatest appreciation to Prof. Kenji Kai (Nagoya University, Japan) for providing the research environment and supporting this thesis. The author also extends sincere gratitude to Prof. Hajime Okamoto (Kyushu University, Japan) and Dr. Yuichiro Hagihara (Kyushu University) for providing many accurate comments and the CloudSat-CALIPSO merged dataset. The author is also grateful to Prof. Takashi Shibata (Nagoya University) whose schooling made enormous contributions to this work, and to Assoc. Prof. Kazuo Osada (Nagoya University) and Assoc. Prof. Kengo Sudo (Nagoya University) for providing valuable advice to this thesis.

The author also appreciates Dr. Nobuo Sugimoto (National Institute for Environmental Studies, Japan), Dr. Atsushi Shimizu (National Institute for Environmental Studies), Dr. Tomoaki Nishizawa (National Institute for Environmental Studies), Dr. Tomohiro Nagai (Meteorological Research Institute, Japan), and Dr. Tetsu Sakai (Meteorological Research Institute) for providing useful comments regarding lidar data analysis. Moreover, the author is sincerely grateful to Prof. Zhou Hongfei (Xinjiang Institute of Ecology and Geography, China) for providing data on the ground-based lidar located at Aksu.

The CALIPSO data were obtained from the Atmospheric Science Data Center (ASDC) at the NASA Langley Research Center. The CloudSat and the ancillary data were obtained from the CloudSat Data Processing Center. The MODIS data were obtained from the Goddard Space Flight Center. The present thesis was supported by a Grant-Aid for Science Research from the Japan Society for the Promotion of Science (JSPS) Fellows.

Finally, special thanks are extended to Mr. Kaya Kanemaru (Nagoya University) and Mr. Takashi Sekiya (Nagoya University), who gave the author their encouragement and special advice. The author would like to thank all members of Kai Laboratory and his colleagues. The author is especially grateful to Ms. Wu Jing for her English writing instructions, and Ms. Akiko Nakashima who provided help and support with enormous paperwork related to the author. The author also expresses gratitude to his family for supplying mental and emotional support. This work would not have been performed without the assistance of every person mentioned above.

References

Ackerman, A. S., 1997: Remote sensing aerosols using satellite infrared observations. *J. Geophys. Res.*, **102**, 17069–17079.

Ackerman, A. S., Holz, R. E., Frey, R., Eloranta, E. W., Maddux, B. C., and McGill, M., 2008: Cloud detection with modis. part ii: Validation. *J. Atmos. Ocean. Tech.*, **25**, 1073–1086.

Ackerman, A. S., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., and Gumley, L. E., 1998: Discriminating clear sky from clouds with modis. *J. Geophys. Res.*, **103**, 32141–32157.

Anderson, T., Masonis, S., Covert, D., and Charlson, R., 2000: In situ measurement of the aerosol extinction-to-backscatter ratio at a polluted continental site. *J. Geophys. Res.*, **105**, 26907–26915.

Bi, L., Yang, P., Lattawar, G. W., Baum, B. A., Hu, Y. X., Winker, D. M., Brock, R. S., and Lu, J. Q., 2009: Simulation of the color ratio associated with the backscattering of radiation by ice particles at the wavelengths of 0.532 and 1.064 μm. *J. Geophys. Res.*, **114**, D00H08.

Bohren, C. F., and Huffman, D. R., 1983: *Absorption and Scattering of Light by Small Particles*. Willy, 530 pp.

Browell, E. V., Butler, C. F., Ismail, S., Robinette, P. A., Carter, A. F., Higdon, N. S., Toon, O. B., Schoeberl, M. R., and Tuck, A. F., 1990: Airborne lidar observations in the wintertime arctic stratosphere: Polar stratospheric clouds. *Geophys. Res. Lett.*, **17**, 385–388.

Bucholtz, A., 1995: Rayleigh-scattering calculations for the terrestrial atmosphere. *Appl. Opt.*, **34**, 2765–2773.

Che, H., Wang, Y., Sun, J., Zhang, X., Zhang, X., and Guo, J., 2013: Variation of aerosol optical properties over the taklimakan desert in china. *Aerosol and Air Quality Research*, **13**, 777–785.

Chen, B., Huang, J., Minnis, P., Hu, Y., Yi, Y., Liu, Z., Zhang, D., and Wang, X., 2010: Detection of dust aerosol by combining calipso active lidar and passive iir measurements. *Atmos. Chem. Phys.*, **10**, 4241–4251.

Choi, Y.-S., Lindzen, R. S., Ho, C.-H., and Kim, J., 2010: Space observations of cold-cloud phase change. *Proc. Natl. Acad. Sci.*, **107(25)**, 11211–6.

Collins, R. T. H., and Russell, P. B., 1976: Lidar measurement of particles and gases by elastic backscattering and differential absorption. *Topics in Applied Physics*, **14**, 71–151.

Dho, S. W., Park, Y. J., and Kong, H. J., 1997: Application of geometrical form factor in differential absorption lidar measurement. *Opt. Rev.*, **4**, 521–526.

Fernald, F. G., 1984: Analysis of atmospheric lidar observations: some comments. *Appl. Opt.*, **23**, 652–653.

Fernald, F. G., Herman, B., and Reagan, J., 1972: Determination of aerosol height distributions by lidar. *J. Appl. Meteor.*, **11**, 482–489.

Frey, A. R., Ackerman, S. A., Liu, Y., Strabala, K. I., Zhang, H., Key, J. R., and Wang, X., 2008: Cloud detection with modis. part i: Improvements in the modis cloud mask for collection 5. *J. Atmos. Ocean. Tech.*, **25**, 1057–1072.

Gu, Y., Rose, W. I., and Bluth, G. J. S., 2003: Retrieval of mass and sizes of particles in sandstroms using two modis ir bands: A case study of april 7, 2001 sandstorm in china. *Geophys. Res. Lett.*, **30(15)**, 1805, doi:10.1029/2003GL017405.

Hagihara, Y., Okamoto, H., and Yoshida, R., 2010: Development of a combined cloudsat-calipso cloud mask to show global cloud distribution. *J. Geophys. Res.*, **115**, D00H33.

Hara, Y., Uno, I., Shimizu, A., Sugimoto, N., Matsui, I., Yumimoto, K., Kurokawa, J., Ohara, T., and Liu, Z., 2011: Seasonal characteristics of spherical aerosol distribution in eastern asia: Integrated analysis using ground/space-based lidars and a chemical transport model. *Scientific Online Letters on the Atmosphere (SOLA)*, **7**, 121–124.

Hashino, T., Satoh, M., Hagihara, Y., Kubota, T., Matsui, T., Nasuno, T., and Okamoto, H., 2013: Evaluating cloud microphysics from nicam against cloudsat and calipso. *J. Geophys. Res.*, **118**, 7273–7292.

Hostetler, M., Liu, Z., Reagan, J., Vaughan, M., Winker, D., Osborn, M., Hunt, W., Powell, K., and Trepte, C., 2006: Cloud-aerosol lidar infrared pathfinder satellite observations caliop algorithm theoretical basis document. available on line at http://www-calipso.larc.nasa.gov/resources/pdfs/PC-SCI-201v1.0.pdf.

Hu, Y., Winker, D., Vaughan, M., Lin, B., Omar, A., Trepte, C., Flittner, D., Yang, P., Nasiri, S. L., Baum, B., Sun, W., Liu, Z., Wang, Z., Young, S., Stamnes, K., Huang, J., Kuehn, R., and Holz, R., 2009: Calipso/caliop cloud phase discrimination algorithm. *J. Atmos. Ocean. Tech.*, **26**.

Hunt, W. H., Winker, D. M., Vaughan, M. A., Powell, K. A., Lucker, P. L., and Weimer, C., 2009: Calipso lidar description and performance assessment. *J. Atmos. Ocean. Tech.*, **26**, doi:10.1175/2009JTECHA1223.1.

Husar, R. B., Tratt, D. M., Schichtel, B. A., Falke, S. R., Li, F., Jaffe, D., Gasso, S., Gill, T., Laulainen, N. S., Lu, F., Reheis, M. C., Chun, Y., Westphal, D., Holben, B. N., Gueymard, C., McKendry, I., Kuring, N., Feldman, G. C., McClain, C., Frouin, R. J., Merrill, J., DuBois, D., Vignola, F., Murayama, T., Nickovic, S., Wilson, W. E., Sassen, K., Sugimoto, N., and Malm, W. C., 2001: Asian dust event of april 1998. *J. Geophys. Res.*, **106**, 18317–18330.

JAXA_EORC, 2012: Level 2 Algorithm Theoretical Basis Document (ATBD), available on line at http://www.eorc.jaxa.jp/EARTHCARE/document/reference/dev/NDX-110018_EarthCARE_L2_ATBD.pdf.

Johnson, R., and Wichern, D., 2007: *Applied Multivariate Statistical Analysis*. Pearson, 6th edn., 800 pp.

Kai, K., Nagata, Y., Tsunematsu, N., Matsumura, T., Kim, H.-S., Matsumoto, T., Shunjun, H., Zhou, H., Abo, M., and Nagai, T., 2008: The structure of the dust layer over the taklimakan desert during the dust storm in april 2002 as observed using depolarization lidar. *J. Meteor. Soc. Japan*, **86(1)**, 1–16.

Kim, H.-S., and Kai, K., 2007: Recent dust outbreaks in the taklimakan desert and their relation to surface wind and land surface condition. *Scientific Online Letters on the Atmosphere (SOLA)*, **3**, 69–72.

Kim, H.-S., Nagata, Y., and Kai, K., 2009: Variation of dust layer height in the northern taklimakan desert in april 2002. *Atmos. Env.*, **43**, 557–567.

Kim, S.-W., Berthier, S., Raut, J., Chazette, P., Dulac, F., and Yoon, S., 2008: Validation of aerosol and cloud layer structures from the space-borne lidar caliop using a ground-based lidar in seoul, korea. *Atmos. Chem. Phys.*, **8**, 3705–3720.

Koffi, B., Schulz, M., Bréon, F., Griesfeller, J., Winker, D., Balkanski, Y., Bauer, S., Berntsen, T., Chin, M., Collins, W., Dentener, F., Diehl, T., Easter, R., Ghan, S., Ginoux, P., Gong, S., Horowitz, L., Iversen, T., Kirkevåg, A., Koch, D., Krol, M., Myhre, G., Stier, P., and Takemura, T., 2012: Application of the caliop layer product to evaluate the vertical distribution of aerosols estimated by global models: Aerocom phase i results. *J. Geophys. Res.*, **117**, D10201.

Li, X., Maring, H., Savole, D., Voss, K., and Prospero, M., 1996: Dominance of mineral dust in aerosol light-scattering in the north atlantic trade winds. *Nature*, **380**, 416–419.

Liu, D., Wang, Z., Liu, D., winker, D., and Trepte, C., 2008a: A height resolved global view of dust aerosols from the first year calipso lidar measurements. *J. Geophys. Res.*, **113**, D16214, doi:10.1029/2007JD009776.

Liu, Z., Kuehn, R., Vaughan, M., Winker, D., Omar, A., Powell, K., Trepte, C., Hu, Y., and Hostetler, C., 2010: The calipso cloud and aerosol discrimination: Version 3 algorithm and test results. *Proc. 25th ILRC*, **2**, 1245–1248.

Liu, Z., Liu, D., Huang, J., Vaughan, M., Uno, I., Sugimoto, N., Kittaka, C., Trepte, C., Hostetler, C., and Winker, D., 2008b: Airborne dust distributions over the tibetan plateau and surrounding areas derived from the first year of calipso lidar observations. *Atmos. Chem. Phys.*, **8**, 5045–5060.

Liu, Z., Sugimoto, N., and Murayama, T., 2002: Extiction-to-backscattering ratio of asian dust observed by high-spectral-resolution lidar and raman lidar. *J. Atmos. Ocean. Tech.*, **20(10)**, 1388–1402.

Liu, Z., Vaughan, M., Winker, D., Hostetler, C., Poole, L., Hlavka, D., Hart, W., and McGill, M., 2004: Use of probability distribution functions for discriminating between cloud and aerosol in lidar backscatter data. *J. Geophys. Res.*, **109**, D15202.

Liu, Z., Vaughan, M., Winker, D., Kittaka, C., Getzewich, B., Kuehn, R., Omar, A., Powell, K., Trepte, C., and Hostetler, C., 2009: The calipso lidar cloud and aerosol discrimination: Version 2 algorithm and initial assessment of performance. *J. Atmos. Ocean. Tech.*, **26**, 1198–1213.

Ma, X., Bartlett, K., Harmon, K., and Yu, F., 2013: Comparison of aod between calipso and modis: significant difference over major dust and biomass burning regions. *Atmos. Means. Tech.*, **6**, 2391–2401.

Ma, Y., Gong, W., Wang, P., and Hu, X., 2011: New dust aerosol identification method for space-borne lidar measurements. *J. Quant. Spectrosc. Radiat. Transfer*, **112**, 338–345.

Marchand, R., Mace, G. G., Ackerman, T., and Stephens, G., 2008: Hydrometeor detection using cloudsat-an earth-orbiting 94-ghz cloud radar. *J. Atmos. Ocean. Tech.*, **25**, 519–533.

McGill, M. J., Vaughan, M. A., Trepte, C. R., Hart, W. D., Hlavka, D. L., Winker, D. M., and Kuehn, R., 2007: Airborne validation of spatial properties measured by the calipso lidar. *J. Geophys. Res.*, **112**, D20201.

Meloni, D., Sarra, A., Iorio, T., and Fiocco, G., 2005: Influence of the vertical profile of saharan dust on the visible direct radiative forcing. *J. Quant. Spectrosc. Radiat. Transfer*, **93**, 397–413.

Mikami, M., Shi, G., Uno, I., Yabuki, S., Iwasaka, Y., Yasui, M., Aoki, T., Tanaka, T., Kurosaki, Y., Masuda, K., Uchiyama, A., Matsuki, A., Sakai, T., Takemi, T., Nakawo, M., Seino, N., Ishizuka, M., Satake, S., Fujita, K., Hara, Y., Kai, K., Kanayama, S., Hayashi, M., Du, M., Kanai, Y., Zhang, Y. Y. X., Shen, Z., Zhou, H., Abe, O., Nagai, T., Tsutsumi, Y., Chiba, M., and Suzuki, J., 2006: Aeolian dust experiment on climate impact: An overview of japan-china joint project adec. *Global and Planetary Change*, **52**, 142–172.

Moulin, C., Lambert, C., Dayan, U., Masson, V., Ramonet, M., Bousquet, P., Legrand, M., Balkanski, Y., Guelle, W., Marticorena, B., Bergametti, G., and Dulac, F., 1998: Satellite climatology of african dust transport in the mediterranean atmosphere. *J. Geophys. Res.*, **20**, 13137–13144.

Murayama, T., Müller, D., Wada, K., Shimizu, A., Sekiguchi, M., and Tsukamoto, T., 2004: Characterization of asian dust and siberian smoke with multi-wavelength raman lidar over tokyo, japan in spring 2003. *Geophys. Res. Lett.*, **31**, L23103, doi:10.1029/2004GL021105.

Naeger, A. R., Christopher, S. A., Ferrare, R., and Liu, Z., 2013: A new technique using infrared satellite measurements to improve the accuracy of the calipso cloud-aerosol discrimination method. *IEEE Trans. Geosci. Rem. Sens.*, **51**, 642–653.

Noh, Y. M., Kim, Y. J., Choi, B. C., and Murayama, T., 2007: Aerosol lidar ratio characteristics measured by a multi-wavelength raman lidar system at anmyeon island, korea. *Atmos. Res.*, **86**, 76–87, doi:10.1016/j.atmosres.2007.03.006.

Okamoto, H., Kumaoka, N., Nishizawa, T., Sugimoto, N., and Hagihara, Y., 2010: Calibration of 1064 nm channel and retrieval of aerosol extinction from calop. *Proc. 25th ILRC*, **1**, 636–639.

Okamoto, H., Nishizawa, T., Takemura, T., Kumagai, H., Kuroiwa, H., Sugimoto, N., Matsui, I., Shimizu, A., Emori, S., Kamei, A., and Nakajima, T., 2007: Vertical cloud structure observed from shipborne radar and lidar: Midlatitude case study during the mr01/k02 cruise of the research vessel mirai. *J. Geophys. Res.*, **112**, D08216.

Omar, A., Winker, D., Kittaka, C., Vaughan, M., Liu, Z., Hu, Y., Trepte, C., Rogers, R., Ferrare, R., Lee, K., Kuehn, R., and Hostetler, C., 2009: The calipso automated aerosol classification and lidar ratio selection algorithm. *J. Atmos. Ocean. Tech.*, **26**, 1994–2014.

Perrone, M., Tafuro, A., and Kinne, S., 2012: Dust layer effects on the atmospheric radiative budget and heating rate profiles. *Atmos. Env.*, **59**, 344–354.

Pinnick, R. G., Jennings, S. G., Chylek, P., Ham, C., and Grandy, W. T., 1983: Backscatter and extinction in water clouds. *J. Geophys. Res.*, **88**, 6787–6796.

Platt, C. M. R., 1973: Lidar and radiometric observations of cirrus clouds. *J. Atmos. Sci.*, **30**, 1191–1204.

Platt, C. M. R., Winker, D. M., Vaughan, M. A., and Miller, S. D., 1999: Backscattering-to-extinction ratios in the top layers of tropical mesoscale convective systems and in isolated cirrus from lite observations. *Amer. Meteor. Soc.*, **38**, 1330–1345.

Prospero, J., and Carlson, T., 1972: Vertical and areal distribution of saharan dust over the western equatorial north atlantic ocean. *J. Geophys. Res.*, **77**, 5255–5265.

Prospero, J., Ginoux, P., Torres, O., Nicholson, S., and Gill, T., 2002: Environmental characterization of global source of atmospheric soil dust identified with the nimbus 7 total ozone mapping spectrometer (toms) absorbing aerosol product. *Rev. Geophys.*, **40**, doi:10.1029/2000RG000095.

Reba, M. N. M., Rocadenbosch, F., Sicard, M., noz, C. M., and Tomás, S., 2007: Piece-wise variance method for signal-to-noise ratio estimation in elastic/raman lidar signals. *Proc. IEEE*, 3158–3161, doi:10.1109/IGARSS.2007.4423515.

Sakai, T., shibata, T., Iwasaka, Y., Nagai, T., Nakazato, M., Matsumura, T., Ichiki, A., Kim, Y. S., Tamura, K., Troshkin, D., and Hamdi, S., 2003: Ice clouds and asian dust studied with lidar measurements of particle extinction-to-backscattering ratio, particle depolarization, and water-vapor mixing ratio over tsukuba. *Appl. Opt.*, **42(36)**, 7103–7116.

Sasano, Y., and Browell, E., 1989: Light scattering characteristics of various aerosol types derived from multiple wavelength lidar observations. *Appl. Opt.*, **28**, 1670–1679.

Sasano, Y., Browell, E., and Ismail, S., 1985: Error caused by using a constant extinction/backscattering ratio in the lidar solution. *Appl. Opt.*, **24**, 3929–3932.

Sassen, K., 2000: Lidar backscatter depolarization technique for cloud and aerosol research. *Light Scattering by Particles and Surfaces: Theory, Measurements, and Applications edited by Mishchenko M., J. Hovenier, and L. Travis*, Academic Press, chap. 14, 393–416.

Satheesh, S. K., and Moorthy, K. K., 2005: Radiative effects of natural aerosols: A review. *Atmos. Env.*, **39**, 2089–2110.

Schuster, G., Vaughan, M., MacDonnell, D., Su, W., Winker, D., Dubovik, O., Lapyonok, T., and Trepte, C., 2012: Comparison of calipso aerosol optical depth retrievals to aeronet measurements, and climatology for the lidar ratio of dust. *Atmos. Chem. Phys.*, **12**, 7431–7452.

Seinfeld, J., and Pandis, S., 1998: *Atmospheric Chemistry and Physics*. Wiley, 1st edn., 1326 pp.

Shao, Y., 2008: *Physics and Modelling of Wind Erosion*. Springer, 2nd revised and expanded edn., 452 pp.

Shibata, T., and Yang, Y., 2010: Decrease in depolarization of dust over populated areas of eastern asia observed by the space-borne lidar caliop. *Earozoru Kenkyu*, **25**, 62–76.

Simmons, A., Uppala, S., Dee, D., and Kobayashi, S., 2007: Era-interim: New ecmwf reanalysis products from 1989 onwards. ECMWF Newsletter 110, European Centre for Medium-Range Weather Forecasts (ECMWF), Winter 2006/2007.

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., Illingworth, A. J., O'Connor, E. J., Rossow, W. B., Durden, S. L., Miller, S. D., Austin, R. T., Benedetti, A., and Mitrescu, C., 2002: The cloudsat mission and the a-train. *Bull. Amer. Meteor. Soc.*, **83**, 1771–1790.

Sugimoto, N., Hara, Y., Shimizu, A., Yumimoto, K., Uno, I., and Nishikawa, M., 2011: Comparison of surface observations and a regional dust transport model assimilated with lidar network data in asian dust event of march 29 to april 2, 2007. *Scientific Online Letters on the Atmosphere (SOLA)*, **7A**, 13–16.

Tanaka, T., Kurosaki, Y., Chiba, M., Matsumura, T., Nagai, T., Yamazaki, A., Tsunematsu, A. U. N., and Kai, K., 2005: Possible transcontinental dust transport from north africa and the middle east to east asia. *Atmos. Env.*, **39**, 3901–3909.

Tao, Z., Liu, Z., Wu, D., McCormick, M. P., and Su, J., 2008a: Determination of aerosol extinction-to-backscattering ratios from simultaneous ground-based and spaceborne lidar measurements. *Opt. Lett.*, **33**, 2986–2988.

Tao, Z., McCormick, M. P., Wu, D., Liu, Z., and Vaughan, M. A., 2008b: Measurement of cirrus cloud backscatter color ratio with a two-wavelength lidar. *Appl. Opt.*, **47**, 1478–1485.

Tegen, I., Hollring, P., Chin, M., Fung, I., Jacob, D., and Penner, J., 1997: Contribution of different aerosol species to the global aerosol extinction optical thickness: Estimates from model results. *J. Geophys. Res.*, **102**, 23895–23915.

Tsunematsu, N., Kai, K., and Matsumoto, T., 2005a: The influence of synoptic-scale air flow and local circulation of the dust layer height in the north of the taklimakan desert. *Water, Air, and Soil Pollution Focus*, **5**, 175–193.

Tsunematsu, N., Sato, T., Kimura, F., Kai, K., Kurosaki, Y., Nagai, T., Zhou, H., and Mikami, M., 2005b: Extensive dust outbreaks following the morning inversion breakup in the taklimakan desert. *J. Geophys. Res.*, **110**, D21207.

Uno, I., Eguchi, K., Yumimoto, K., Takemura, T., Shimizu, A., Uematsu, M., Liu, Z., Wang., Z., Hara., Y., and Sugimoto, N., 2009: Asian dust transported one full circuit around the globe. *Nature Geoscience*, **2**, 557–560.

Uno, I., Satake, S., Carmichael, G. R., Tang, Y., Wang, Z., Takemura, T., Sugimoto, N., Shimizu, A., Murayama, T., Cahill, T. A., Cliff, S., Uematsu, M., Ohta, S., Quinn, P. K., and Bates, T. S., 2004: Numerical study of asian dust transport during the springtime of 2001 simulated with the chemical weather forecasting system (cfors) model. *J. Geophys. Res.*, **109**, D19S24.

Vaughan, M., Powell, A., Kuehn, R. E., Young, S. A., Winker, D. M., Hostetler, C. A., Hunt, W. H., Liu, Z., McGill, M. J., and Getzewich, B. J., 2009: Fully automated detection of cloud and aerosol layers in the calipso lidar measurements. *J. Atmos. Ocean. Tech.*, **26**, 2034–2050.

Warner, T., 2004: *Desert Meteorology*. Cambridge University Press, 595 pp.

Weitkamp, C., 2005: *Lidar: Range-Resolved Optical Remote Sensing of the Atmosphere*. Springer, 455 pp.

Wiacek, A., and Peter, T., 2009: On the availability of uncoated mineral dust ice nuclei in cold cloud regions. *Geophys. Res. Lett.*, **36**, L17801.

Wiacek, A., Peter, T., and Lohmann, U., 2010: The potential influence of asian and african mineral dust on ice, mixed-phase and liquid water clouds. *Atmos. Chem. Phys.*, **10**, 8649–8667.

Winker, D. M., Hunt, W. H., and McGill, M. J., 2007: Initial performance assessment of caliop. *Geophys. Res. Lett.*, **34**, L19803.

Winker, D. M., Tackett, J., Getzewich, B., Liu, Z., Vaughan, M., and Rogers, R., 2013: The global 3-d distribution of tropospheric aerosols as characterized by caliop. *Atmos. Chem. Phys.*, **13**, 3345–3361.

Winker, D. M., Vaughan, M., Omar, A., Hu, Y., Powell, K., Liu, Z., Hunt, W., and Young, S., 2009: Overview of the calipso mission and caliop data processing algorthms. *J. Atmos. Ocean. Tech.*, **26**, 2310–2323.

Xie, Y., Qu, J. J., and Xiong, X., 2010: Improving the calipso vfm product with aqua modis measurements. *Rem. Sen. Lett.*, **1**, 195–203.

Yoshida, R., Okamoto, H., Hagihara, Y., and Ishimoto, H., 2010: Global analysis of cloud phase and ice crystal orientation from cloud-aerosol lidar and infrared pathfinder satellite observation (calipso) data using attenuated backscattering and depolarization ratio. *J. Geophys. Res.*, **115**, D00H32.

Young, S., and Vaughan, M., 2009: The retrieval of profiles of particulate extinction from cloud-aerosol lidar infrared pathfinder satellite observations (calipso) data: Algorithm description. *J. Atmos. Ocean. Tech.*, **26**, 1105–1119.

Yumimoto, K., Eguchi, K., Takemura, T., Liu, Z., Shimizu, A., and Sugimoto, N., 2009: An elevated large-scale dust veil from the taklimakan desert: Intercontinental transport and three-dimensional structure as captured by calipso and regional and global models. *Atmos. Chem. Phys.*, **9**, 8545–8558.

Zender, C., Miller, R., and Tegen, I., 2004: Quantifying mineral dust mass budgets: terminology, constraints, and current estimates. *Eos Trans. Amer. Geophys. Union*, **85**, 509–512.

Zhang, P., Lu, N., Hu, X., and Dong, C., 2006: Identification and physical retrieval of dust storm using three modis thermal ir channels. *Global Planetary Change*, **52**, 197–206.

Zhao, C., Ruby, L., and Hagos, S., 2004: Radiative impact of mineral dust on monsoon precipitation variability over west africa. *Eos Trans. Amer. Geophys. Union*, **11**, 1879–1893.

Appendix

A Calculation of lidar ratios for elevated cloud layers

According to *Platt* [1973], a cloud lidar ratio $S_{c,\lambda}$ is given as:

$$S_{c,\lambda} = \frac{1 - T_{c,\lambda}^2}{2\gamma'_\lambda}, \tag{A.1}$$

where γ'_λ is the value of the integral of attenuated backscattering coefficients from the cloud base altitude to the cloud top altitude. The cloud transmittance $T_{c,\lambda}^2$ is estimated by a method of *Platt et al.* [1999] and is given as:

$$T_{c,\lambda}^2 = \frac{R'_\lambda(z_b)R_\lambda(z_t)}{R'_\lambda(z_t)R_\lambda(z_b)}, \tag{A.2}$$

where z_t is the altitude just above the cloud top altitude and $R'_\lambda(z)$ is the attenuated backscattering ratio that is derived from observed attenuated backscattering coefficients and the molecular backscattering coefficient as follows:

$$R'_\lambda(z) = \frac{\beta'_\lambda(z)}{\beta_{2,\lambda}(z)} = \frac{\beta_\lambda(z)T_\lambda^2(z)}{\beta_{2,\lambda}(z)}. \tag{A.3}$$

In Equation (A.2), the term $R_\lambda(z_t)$ is derived from calculated backscattering coefficients assuming a lidar ratio, but the term $R_\lambda(z_b)$ cannot be derived unless the cloud lidar ratio is known. Since the CALIOP backscattering ratio $R_{S,532}(z_b)$ is unknown, we must determine $R_{S,532}(z_b)$ in order to calculate the cloud transmittance, resulting in the cloud lidar ratio at 532 nm. The cloud transmittance at 1064 nm is assumed to be the same as that at 532 nm since the extinction coefficient is wavelength-independent for large particles with respect to the laser wavelengths [*Bohren and Huffman*, 1983]. The cloud lidar ratio at 1064 nm is calculated by the cloud transmittance at 532 nm and the integrated attenuated backscattering coefficient at 1064 nm.

B Analysis of errors caused by random noise in the lidar ratio estimation

An error analysis is conducted to investigate the effects of random errors in the lidar ratio estimation from the simultaneous observation by Aksu-lidar and CALIOP. First, lidar signals with the arbitrary backscattering ratios shown in Figure B.1 (a) and (b) are constructed. This thesis presupposes a thin cloud layer above 7.5 km and a dust layer below 7.5 km. The dust lidar ratios at 532 and 1064 nm are assumed to be 40 and 45 sr, respectively. Using the backscattering and extinction backscatter coefficients of the atmospheric molecules, attenuated backscattering coefficients are virtually calculated. A stochastic noise generated by normal Gaussian white noise is added to the attenuated backscattering coefficient. The stochastic noise is generated based on the standard deviation calculated from the signal-to-noise ratio (SNR), which is calculated at each altitude by a smoothing method described by *Reba et al.* [2007]. An example of the attenuated backscattering coefficients contaminated by this random noise is shown in Figure B.1 (c) and (d). The lidar ratios at 532 and 1064 nm are then estimated using the performance function in Equations (3.1) and (3.2). The estimation is conducted in 1000 cases changing the stochastic noise. The statistical results of the error are 7.2% at 532 nm and 21.1% at 1064 nm, respectively.

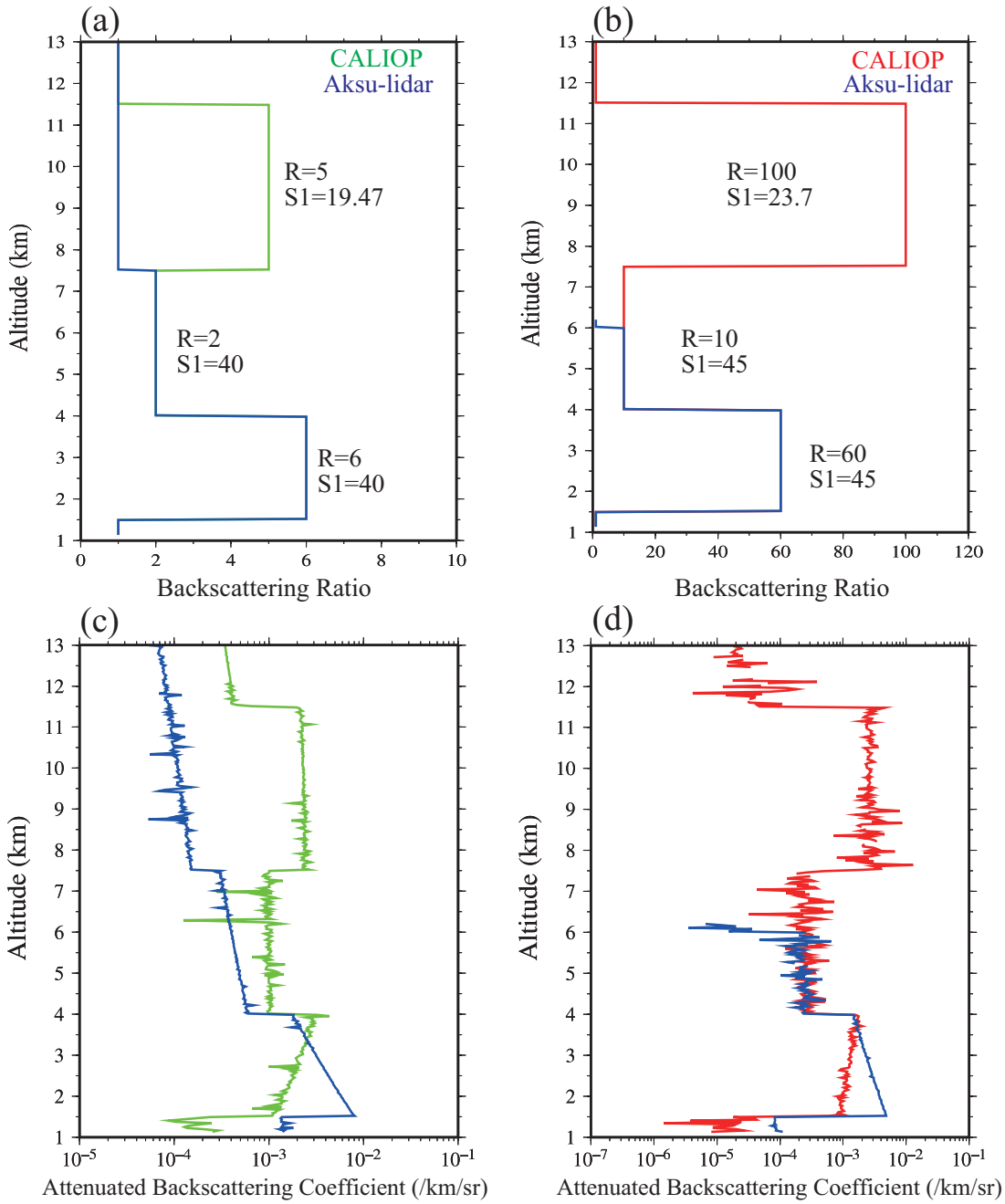

Figure B.1: An example of simulated signals in the case of the simultaneous lidar observation by Aksu-lidar and CALIOP. Panels (a) and (b) illustrate the vertical profiles of the arbitrary backscattering ratios at 532 and 1064 nm wavelengths, respectively. Panels (c) and (d) show the attenuated backscattering coefficients, which are added the random noise in accordance with the signal-to-noise ratio (SNR). The blue, green, and red lines denote Aksu-lidar, CALIOP at 532 nm, and CALIOP at 1064 nm, respectively.

C Tables of cloud test thresholds in the MODIS cloud mask

Table C.1: holds of the infrared (IR) cloud test used in the MODIS cloud mask and the scene types that are applied to the test (from *Frey et al.* [2008]). The term LI denotes the linear interpolation, and the term VZA denotes the viewing zenith angle.

MODIS IR cloud test thresholds		
IR test	Thresholds for confidence limits (0.0, 0.5, 1.0, or low, middle, high)	Scenes
11 μm (freezing test)	267, 270, 273 K	All ocean
13.9 μm	222, 224, 226 K	All nonpolar
6.7 μm	215, 220, 225 K	All except Antarctic night
Surface temperature−11 μm	6 K modified by VZA and 11 − 12-μm BTD +1 K/−2 K low/high	Day, night deep ocean
Surface temperature−11 μm	Same as above, but base threshold=10 K	Day, night shallow ocean
Surface temperature−11 μm	12 K modified by VZA and 11 − 12-μm BTD ±2 K low/high	Nonarid night land
Surface temperature−11 μm	20 K modified by VZA and 11 − 12-μm BTD ±2 K low/high	Arid, semiarid night land
11 − 12-μm BTD	Function of VZA and 11-μm BT Confidence limits vary by scene type, latitude	All except Antarctica
11, 12, 8.6-μm (trispectral test)	8.6 − 11-μm BTD threshold based on 11 − 12-μm BTD, ±0.5 K low/high	All ocean
11 − 3.9-μm BTD	−14, −12, −10 K	Nonarid day land
11 − 3.9-μm BTD	−20, −18, −16 K	Arid, semiarid day land 11-μm BT≤320 K
11 − 3.9-μm BTD	10, 7, 4 K 14, 10, 6 K for elevations>2000 m	Day snow/ice
11 − 3.9-μm BTD	11 − 12-μm> +1.00 : −2.0, −2.5, −3.0 K 11 − 12-μm> −1.00 : +5.0, +4.5, +4.0 K −1 ≤ 11 − 12-μm BTD≤ +1 LI between −2.5 and +4.5 K, ±0.5 K low/high	Night land
11 − 3.9-μm BTD	0.7, 0.6, 0.5 K	Night snow/ice
11 − 3.9-μm BTD	−10, −8, −6 K	Day ocean
11 − 3.9-μm BTD	1.25, 1.00, −1.00 K	Night ocean
11 − 3.9-μm BTD	11-μm BT<235 K: −0.1, −0.2, −0.3 K 11-μm BT>265 K: +1.1, +1.0, +0.0 K 235 K≤ 11-μm BT≤265 K LI between −0.2 and +1.0 K, ±0.1 K low/high	Polar night land, snow/ice
11 − 3.9-μm BTD	11-μm BT<230 K: −17.5, −14.5, −11.5 K 11-μm BT>245 K: −10, −7, −4 K 230 K≤ 11-μm BT≤245 K LI between −14.5 and −7 K, ±3 K low/high	Polar day, snow/ice 11-μm BT>230 K
3.9 − 12-μm BTD	15 K, 10 K, 5 K	Night land
3.9 − 12-μm BTD	4.5 K, 4.0 K, 3.5 K	Night snow
3.9 − 12-μm BTD	11-μm BT<235 K: 4.5, 4.0, 3.5 K 11-μm BT>265 K: 2.5, 2.0, 1.5 K 235 K≤ 11-μm BT≤265 K LI between 4.0 and 2.0 K, ±0.5 K low/high	Polar night, snow/ice (Elevation≤2000 m)
7.2 − 11-μm BTD	−8 K, −10 K, −11 K	Night land, 11 − 3.9$\mu m \leq -2$ K

Table C.1: (Continued)

MODIS IR cloud test thresholds		
IR test	Thresholds for confidence limits (0.0, 0.5, 1.0, or low, middle, high)	Scenes
7.2 – 11-μm BTD	11-μm BT<220 K: −1, 0, 1 K 220 K<11-μm BT<245 K LI between 0 K and −4.5 K, ±1 K low/high 245 K≤11-μm BT<255 K LI between −4.5 K and −10.5 K, ±1 K low/high 255 K≤11-μm BT≤265 K LI between −10.5 K and −20 K, ±1 K low/high 11-μm BT>265 K: −21, −20, −19 K	Polar night land, night snow
7.2 – 11-μm BTD	11-μm BT<265 K: 0, 1, 2 K 220 K≤11-μm BT<245 K LI between 1 K and −7 K, ±1 K low/high 245 K≤11-μm BT<255 K LI between −7 K and −16.5 K, ±1 K low/high 255 K≤11-μm BT<265 K LI between −16.5 K and −20 K, ±1 K low/high 11-μm BT>265 K: −21, −20, −19 K	Night ice
8.6 – 7.2-μm BTD	16, 17, 18 K	Night ocean 11 μm BT≥280 K for polar night ocean
11-μm BT variability No. surrounding pixel BTs minus center pixel BT≤0.5 K	3, 6, 7	Night ocean

Table C.2: Thresholds of the visible (VIS) and near-infrared (NIR) cloud tests used in the MODIS cloud mask and the scene types that are applied to the tests (from *Frey et al.* [2008]). The abbreviations are identical to those used in Table C.1, but the term REF denotes the reflectance.

MODIS VIS/NIR cloud test thresholds		
VIS/NIR test	Thresholds for confidence limits (0.0, 0.5, 1.0, or low, middle, high)	Scenes
0.86-μm REF	Aqua 0.065, 0.045, 0.030 Terra 0.055, 0.040, 0.030	Nonglint unfrozen day ocean
0.86-μm REF	0.34, 0.30, 0.26	Arid, semiarid day land (no snow)
0.86-μm REF	Glint angle 0°-10°: 0.115, 0.105, 0.095 Glint angle 10°-20° LI between 0.105 and 0.075, ±0.01 low/high Glint angle 20°-36° LI between 0.075 and 0.045, ±0.01 low/high (Aqua) LI between 0.075 and 0.040, ±0.01 low/high (Terra)	Sun glint
0.66-μm REF	0.22, 0.18, 0.14	Nonsnow, nonarid day land
1.38-μm REF	0.040, 0.035, 0.030	Day scenes except snow/ice Elevation≤2000 m
1.38-μm REF	0.0600, 0.0525, 0.0450	Day snow/ice Elevation≤2000 m
0.86/0.66-μm REF ratio	0.95, 0.90, 0.85	Nonglint unfrozen day ocean
0.86/0.66-μm REF ratio	1.05, 1.00, 0.095	Sun glint
0.86/0.66-μm REF ratio [Modified Global Environment Monitoring Index (GEMI)]	1.85, 1.90, 1.95	Noncoastal, nonarid day land (no snow)

Table C.3: Thresholds of the clear-sky restoral (CSR) test used in the MODIS cloud mask and the scene types that are applied to the test (from *Frey et al.* [2008]). The abbreviations are identical to those used in Tables C.1 and C.2.

MODIS clear-sky restoral test thresholds		
test	Threshold	Scenes
11-μm BT	Aqua: >295.0 K, >300.0 K, >305.0 K Terra: >292.5 K, >297.5 K, >302.5 K Assign probably cloudy, probably clear, clear (thresholds adjusted for elevation)	Day land, no snow, original confidence≤0.95 No IR cloud tests positive
11-μm BT	Aqua: >290.0 K, >295.0 K, >305.0 K Terra: >287.5 K, >292.5 K, >302.5 K Assign probably cloudy, probably clear, clear (thresholds adjusted for elevation)	Nonvegetated land, no snow, original confidence≤0.95, no IR cloud tests positive
0.55/1.24-μm, 3.9 – 3.9-μm BTD, 3.9 – 11-μm BTD	0.55/1.24 μm>3.0 and 3.7 – 3.9-μm BTD<11 K and 3.9 – 11-μm BTD<15 K, assign probably clear	Day land, no snow, original confidence ≤0.95, no IR cloud tests positive, above CSR tests negative
11-μm BT	>287.5 K, >292.5 K, >297.5 K Assign probably cloudy, probably clear, clear (thresholds adjusted for elevation)	Nonpolar night land, no snow, original confidence≤0.95 no high- or middle-cloud tests positive
Normalized difference vegetation index	NDVI≤ −0.18 or NDVI≥ 0.40 Assign clear	Coast and shallow water No high- or middle-cloud tests positive
3.9 – 11-μm BTD	≥13 K, assign probably cloudy	Sun glint, original confidence<0.95, no high-middle-, or surface temperature cloud tests positive
0.895/0.935-μm	>3.0, assign probably clear	Same as above, positive 3.7 – 11-μm BTD test
0.86-μm REF	σ mean<0.001 over 3×3 pixel region, assign probably clear	Same as above, positive 3.7 – 11-μm BTD test
11-μm BT	8 surrounding pixels BTs minus center pixel BT≤0.5 K, assign probably clear	Ocean, 0.66<original confidence≤0.95
11-μm BT	8 surrounding pixels BTs minus center pixel BT≤0.5 K, assign probably cloudy	Ocean, 0.05<original confidence≤0.66
6.7 – 11-μm BTD	> +10 K, assign clear	Night snow
13.3 – 11-μm BTD	> +3 K, assign clear	Polar night snow
7.2 – 11-μm BTD	> +5 K, assign clear	Polar night snow

D Equations of the discriminant function

In this section, a derivation of the equation of the discriminant function used in this thesis is described. Here, two groups of classification categories are defined: clouds (G_1) and misclassified clouds (G_2). The probability density function (PDF) of each group is assumed to be a multivariate normal distribution as follows:

$$f_i(\mathbf{x}) = \frac{1}{\left(\sqrt{2\pi}\right)^k \sqrt{|\mathbf{S_i}|}} \exp\left[-\frac{1}{2}(\mathbf{x} - \overline{\mathbf{x}}_\mathbf{i})^T \mathbf{S_i^{-1}} (\mathbf{x} - \overline{\mathbf{x}}_\mathbf{i})\right], \tag{D.1}$$

where i is the group number (1 or 2), k is the number of variables, $\mathbf{x}$ is the variable from observations, $\overline{\mathbf{x}}_\mathbf{i}$ is the sample average of the variables in the i group, and $\mathbf{S}_i$ is the variance-covariance matrix of the i group. The superscript T denotes a transposed matrix. If observed values $\mathbf{x}$ exist within a dot product space (R^i), the values that are attributed to group i are identified. R^1 and R^2 are mutually exclusive events, and the summation of these regions corresponds to the total region R^t. The total misclassification rate (P_M) is written as:

$$P_M = \frac{1}{2}\left[\int_{R^2} f_1(\mathbf{x})\mathbf{dx} + \int_{R^1} f_2(\mathbf{x})\mathbf{dx}\right] = \frac{1}{2}\left\{1 + \int_{R^1} \left[f_2(\mathbf{x}) - f_1(\mathbf{x})\right]\mathbf{dx}\right\}. \tag{D.2}$$

The term P_M will be minimum if R^1 is set to $f_2(\mathbf{x}) - f_1(\mathbf{x}) \leq 0$. The discriminant equation is then written by the logarithm of the ratio of f_1 to f_2:

$$LDF(\mathbf{x}) = \ln\frac{f_1(\mathbf{x})}{f_2(\mathbf{x})} = \left(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2\right)^T \mathbf{S^{-1}}\mathbf{x} - \frac{1}{2}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)^T \mathbf{S^{-1}}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) \begin{cases} \geq 0 \Longrightarrow G_1 \\ < 0 \Longrightarrow G_2 \end{cases}, \tag{D.3}$$

where $\mathbf{S}$ is the common covariance-variance matrix given as:

$$\mathbf{S} = \frac{1}{n_1 + n_2 - 2}\left[(n_1 - 1)\mathbf{S_1} + (n_2 - 1)\mathbf{S_2}\right], \tag{D.4}$$

where n_1 and n_2 are the sample sizes of G_1 and G_2, respectively. If $LDF(\mathbf{x})$ returns a positive value (negative value), the observed object is classified as G_1 (G_2).The second term in Equation (D.3) corresponds to C_0 in Equation (4.3). The coefficient of $\mathbf{x}$ in the first term in Equation (D.3) corresponds to C_i in Equation (4.3).

E Table of the average and variance-covariance of the variables

Table E.1: The average of the variables of clouds and misclassified clouds (MC), and their common variance-covariance matrix.

Statistic		$\log_{10}\left(\overline{\beta'_{S,532}}\right)$	$\overline{\delta_{S,v}}$	$\overline{\chi'_S}$	z_{lt}	BTD
Average	MC	−2.0655	0.3156	0.8746	2.5080	−0.0869
	Clouds	−1.3532	0.2005	1.1892	4.6452	1.8485
S	$\log_{10}\left(\overline{\beta'_{S,532}}\right)$	0.1815	0.0018	0.0490	−0.2796	−0.0618
	$\overline{\delta_{S,v}}$	0.0018	0.0196	0.0055	−0.0506	−0.0507
	$\overline{\chi'_S}$	0.0490	0.0055	0.0794	−0.1850	−0.0489
	z_{lt}	−0.2796	−0.0506	−0.1850	2.9094	0.6311
	BTD	−0.0618	−0.0507	−0.0489	0.6311	3.0745

F Determination method of dust in CALIOP data

The purpose of the method described here is to determine the presence of dust in the CALIOP data whose horizontal and vertical resolutions are 1.1 km and 240 m, respectively. The vertical feature mask (VFM) provides cloud detection at the five horizontal resolutions of 333 m, 1 km, 5 km, 20 km, and 80 km. The presence of dust is determined after clouds are screened at each horizontal resolution. Random noise is removed from weak signals by the averaging processes.

On the other hand, since cloud detection in the C2 cloud mask is only implemented at the 1.1 km horizontal resolution, aerosol signals remaining after cloud screening are noisy. In order to detect dust at the horizontal resolution, a different method from the VFM should be developed. In this thesis, dust is detected by thresholds and continuity tests that were used to remove the noise effect in the previous studies [*Marchand et al.*, 2008; *Hagihara et al.*, 2010].

The volume depolarization ratio and attenuated color ratio are used in the determination method of dust. The detection procedure consists of four steps including two threshold tests and two continuity tests, and is illustrated by a demonstration in Figure F.1. The ultimate goal of this method is to detect the dust as shown in Figure F.1 (d). First, a threshold test of the volume depolarization ratio is conducted. If the volume depolarization ratio of an observed bin is within a range from 0.06 to 0.35, the result of the threshold test (TT1) is 1. If the C2 cloud mask detects the cloud in the observed bin, the TT1 of the bin is forced to be 0. Figure F.1 (e) shows the results of the TT1 in the data window that is denoted by the red rectangle in Figures F.1 (a)–(d).

Secondly, a continuity test of the depolarization ratio is conducted by using the data window. The data window sizes are 15×3 bins (horizontal by vertical) and correspond to 16.5 km (horizontal) and 720 m (vertical). The continuity test (CT1) is defined as the ratio of the summation of the TT1 in a data window to the total number of bins (= 45). Figure F.1 (f) shows the results of the CT1 in the data window. The CT1 at each bin in the data window is calculated in the data window centered at each bin. The CT1 of the target bin, which is the center of the data window in the figure, is 0.44 (20/45).

Thirdly, a threshold test of the attenuated color ratio is conducted in the same manner as the volume depolarization ratio. If the attenuated color ratio of an observed bin is within a range from 0.3 to 1.3, the result of the threshold test (TT2) is 1. If the C2 cloud mask detects clouds in the observed bin, the TT2 of the bin is forced to be 0. Figure F.1 (g) shows the results of the TT2 in the data window. Finally, a continuity test of the attenuated color ratio is conducted. The continuity test (CT2) is defined as the ratio of the summation of the TT2 in a data window to the total number of bins. Here, the bins whose CT1 are less than 0.5 are excluded from the summation of the TT2.

Figure F.1 (h) shows the results of the CT2 in the data window. The CT2 of the target bin is 0.38 (17/45).

Dust is detected by using the results of CT2. Bins whose CT2 are more than 0.25 are regarded as dust bins in this thesis. The threshold values in the TT1 and the TT2 are determined by the reference to a probability density function (PDF) reported by *Liu et al.* [2008b]. The data window size and the threshold values in the CT1 (0.5) and the CT2 (0.25) are determined by a trial and error step.

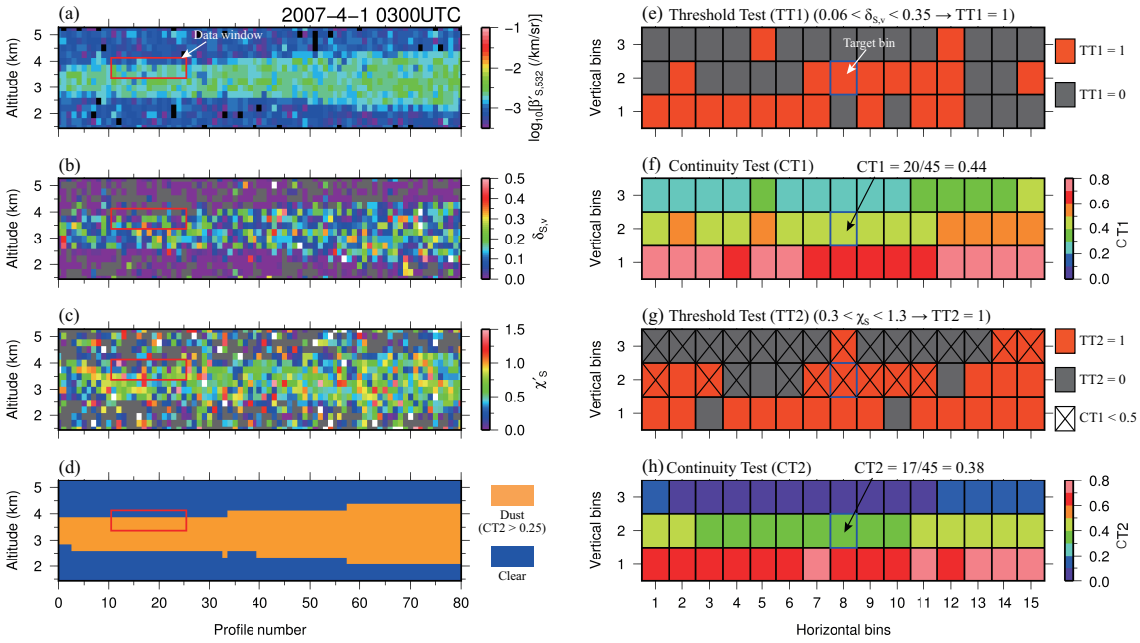

Figure F.1: A demonstration of the dust detection procedure in the CALIOP observation on April 1, 2007. Time-height cross sections of (a) the attenuated backscattering coefficient, (b) volume depolarization ratio, (c) attenuated color ratio, and (d) dust mask. The red rectangle denotes an example of the sliding data window. Panels (e), (f), (g), and (h) are the results of the threshold test (TT1), continuity test (CT1), threshold test (TT2), and continuity test (CT2) in the data window.

G Coefficients of LDF derived from the four variables

Table G.1: Coefficients of the linear discriminant function (LDF). The coefficients C1, C2, C3, and C4 are for $\log_{10}\left(\overline{\beta'_{S,532}}\right)$, $\overline{\chi'_S}$, z_{lt}, and BTD, respectively.

Coefficient	C_0	C_1	C_2	C_3 (km^{-1})	C_4 (K^{-1})
Value	−1.3117	5.0528	4.3918	1.3874	0.5160