CA Algorithm Theoretical Basis Document

NMSC/SCI/ATBD/CA, Issue 1, rev.4

26 December 2012

REPORT SIGNATURE TABLE

Function	Name	Signature	Date
Prepared by	Yong-Sang Choi, Heaje Cho		26 December 2012
Reviewed by	Yong-Sang Choi		26 December 2012
Authorised by	NMSC		26 December 2012

DOCUMENT CHANGE RECORD

Version	Date	Pages	Changes
Version5	2012.12.26	-	-Nothing has changed for contents besides ATBD form.

Table of contents

List of Tables

List of Figures

List of Acronyms

ARM	Atmospheric Radiation Measurement
CA	Cloud Amount
CF	Cloud Fraction
CMDPS	COMS Meteorological Data Processing System
COLL	Collocation module
COMS	Communication, Ocean, and Meteorological Satellite
CTP	Cloud Top Pressure
GTS	Global Telecommunication System
IPCC	Intergovernmental Panel on Climate Change
ISCCP	International Satellite Cloud Climatology Project
KMA	Korea Meteorological Administration
MOD06	MODIS Cloud product
MODIS	Moderate Resolution Imaging Spectroradiometer
RMSE	Root Mean Square Error
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SGP	Southern Great Plains
TSI	Total Sky Imager
VAM	Validation Module

1. Overview

Cloud amount (CA) is a meteorological factor which deeply influences our daily life in areas such as sailing, aviation, agriculture, and outdoor activities. Cloudiness in an atmosphere adjusts the intensity of incoming solar energy. This in turn regulates the global climate system. It is also a key parameter in adjusting the emission for terrestrial radiation at the top of the atmosphere. CA is one of the essential objects of weather forecasting, so it has long been derived by ground measurements. Due to the importance of CA in daily life over the Korean Peninsula, it has become an object of attention to the general public. This includes the CA of surrounding countries and bodies of water (Oh et al. 2005).

In order to acquire high quality data, The digitized retrieval of ground observation CA is requested through satellites because of the limitations of ground observation. This section will explain the CA algorithm for satellite observation CA and eye measurement CA including the steps and processes involved.

2. Background and purpose

The algorithm extracts CA viewed by satellites using pixel information for the existence or nonexistence of clouds retrieved by scene analysis. When compared with other satellites, satellite observation CA extracts data through averaging Moderate Resolution Imaging Spectroradiometer (MODIS) 5x5 data, Spinning Enhanced Visible and Infrared Imager (SEVIRI) 3x3 data, and (IPCC) 1x1 data. Satellite observation CA is unaffected by the location and size of clouds, but these factors significantly affect ground observation CA. Therefore, to retrieve both satellite observation CA from COMS Meteorological Data Processing System (CMDPS) and ground observation CA generates the best data. CA is an essential part of the equation to predict the maximum and minimum temperature. However, this currently relies on the ground CA observations of human observers. It has the limitations of objective reliability with time restrictions.

Satellite observation CA from Communication, Ocean, and Meteorological Satellite (COMS) and ground observation CA can contribute to quantitative weather forecasting. It can also be utilized in a digital forecast by providing information with specific number and graphic form in time and space. This product can also be utilized for the study of an indirect impact of aerosol, along with effective

particle radius of clouds.

3. Algorithm

3.1 Theoretical background and basis

Ground-observed CA is to be retrieved using inversion of method proposed by Kassianov et al. (2005). Figure 1 illustrates the difference between ground-observed CA and satellite observation CA. Figures 1(a) and (b) show real cloud distribution. Figures 1(c) and (d) show what can be observed from the ground.

When many clouds exist at the zenith of ground observations without many clouds around them, ground-observed CA can be overestimated. While clouds do not exist at the zenith of ground observations with many clouds around them, ground-observed CA can be underestimated.

The differences between satellite-observed CA and ground-observed CA are caused by the following:

First, there is a difference of observation conditions. Satellite observations detect the upper level of clouds from outside of the Earth's atmosphere. Ground-observed CA detects the lower level of clouds from the ground. Also, Satellite data has the value of pixels, but ground-observed data must consider the curvature of the earth.

Second, ground-observed CA is affected by the horizontal and vertical structures of clouds, while satellite observation does not consider these factors. To retrieve satellite-observed data, this difference must be corrected. Therefore, we need to convert the results of satellite observations to ground-observed CA. The physical law can be used to convert by the method shown in Kassianov et al. (2005) and Oh et al. (2006).

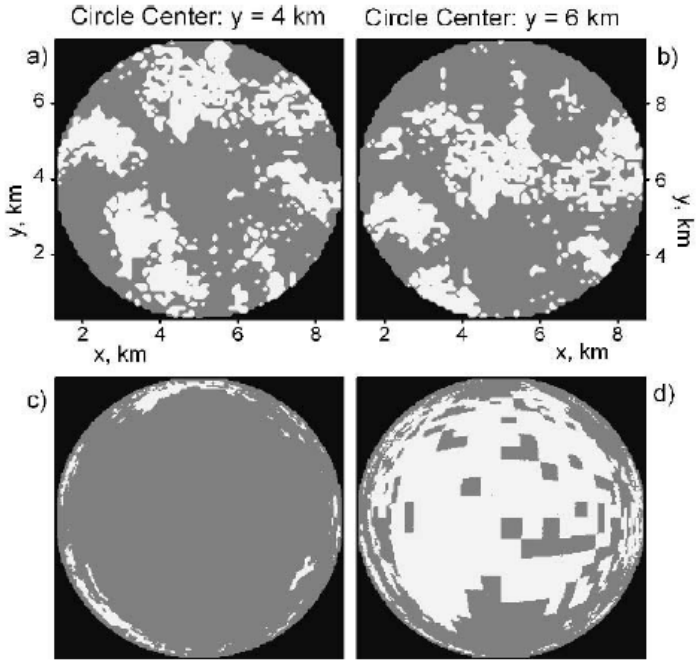


Fig. 1. Images of (a), (b)nadir-view cloud fraction and (c), (d)fractional sky cover (Kassianov et al., 2005)

3.2 Retrieval method

3.2.1 Look up table (LUT) retrieval method

With the observer at the center, we made a look up table (LUT) for weighting values depending on the location of clouds. The weighting value is defined by the angle ($\frac{\alpha}{\alpha^*}$) between the cloud and observer locations for maximum zenith angle in view of the observer. This weight value must be applied to each pixel from the zenith to $\alpha=\alpha^*$.For example, in the case of a cloud located over the zenith, the weighting value becomes 0 because it does not need to consider the weighting value due to the cloud's location. However, in the case of a cloud located around the edge, the same cloud tends to be more overestimated than those located at the zenith. In the α^* of the maximum angle increasing the weighting value to 1.

In the case of several pixel groups being selected, the most consistent results with ground-observed CA investigate the sensitivity and have MODIS CA and ground-observed CA of the Southern Great Plains (SGP) region of the Atmospheric Radiation Measurement (ARM). In this process, ground observers used a cloud base height of 2km on the assumption that they looked at cloud base.

The distance up to the edge of the clouds from the zenith of ground observers becomes about 12km, either direction to them at the center becomes about 24km. The size of one pixel is 5km in the case of MODIS. The edge seen by the observers needsfive pixels, one pixel in the middle, and two pixels each direction. Thus, the observation area located at a certain point is 5x5 pixels.

When it is taken to 50km the maximum distance visible to an observer, increasing pixel size up to 15x15 and calculated for correlation coefficient and Root Mean Square Error (RMSE) with ground observation data.Considering both the time for retrieving the ground-observed CA and consistency with ground observation data, 5x5 appears to be the most suitable.

Table 1. Correlation coefficients and RMSE of sensitivity test to the cloud weight part.

		$R_{5\times5}$	$R_{7\times7}$	$R_{9\times9}$	$R_{11\times11}$	$R_{13\times13}$	$R_{15\times15}$
SGP	Correlation	0.94	0.94	0.93	0.92	0.90	0.90
	RMSE	13.34	13.57	14.00	14.88	16.16	17.25

According to the sensitivity results for cloud weight value, COMS is 3km per pixel. The lookup table to apply the weight value due to cloud position is a pixel group of 7x7 for ground-observed CA. The retrieval method is as follows:

```
program makelut
!-----------------------------------------
!!Description :
!   This subroutine is matrix to give cloud weight,
!     == >   Calculate matrix
!              from formula of EVGUENI KASSIANOV et. al.
!              IF fof Maximum degree(80)=distance (D) and    a degree=d
!               tan a = d/H
!               # according to resolution, H is changed
!                  because tan80 = D/H, D is related to resolution
!              Thus, a=tan-1(d/H)
```

```
!                cloud weight= a/80
!-----------------------------------------
implicit none
  integer, parameter :: grid=3, pi=3.1452, res=3
  real, parameter :: H=1.6
  integer :: m, n
  real, dimension(7,7) :: mata, matrix

open (1, file='lut_matrix.txt', status='unknown')
  DO n= 1, grid*2+1 ; DO m= 1, grid*2+1
mata(m, n)= (grid+1-m)**2 + (grid+1-n)**2
mata(m,n)= ((mata(m,n))*(res))**(0.5)/H
          matrix(m,n)= ATAN(mata(m,n))
matrix(m,n)= ((matrix(m,n)) * 180.) / pi
matrix(m,n)= matrix(m,n)/80.

          IF (matrix(m,n)>=1.) THEN
matrix(m,n) = 1.
          ELSEIF (matrix(m,n)<1.) THEN
               matrix(m,n) = matrix(m,n)
          END IF

    END DO ; END DO

write (1,*) matrix

end program makelut
```

3.2.2 Methodology

This algorithm retrieves satellite-observed CA of an interim product that converts from % unit to ratio including pixels of clouds from the results of cloud detection. In other words, cloud pixels under confidence level of some clear are 1. After summing up to convert into 0 in clear pixels will divide

into the number of total pixels.

The weighted value of ground-observed CA retrieved for position information of clouds from the results of cloud detection and correcting differences that occur due to the geometric structure of clouds.

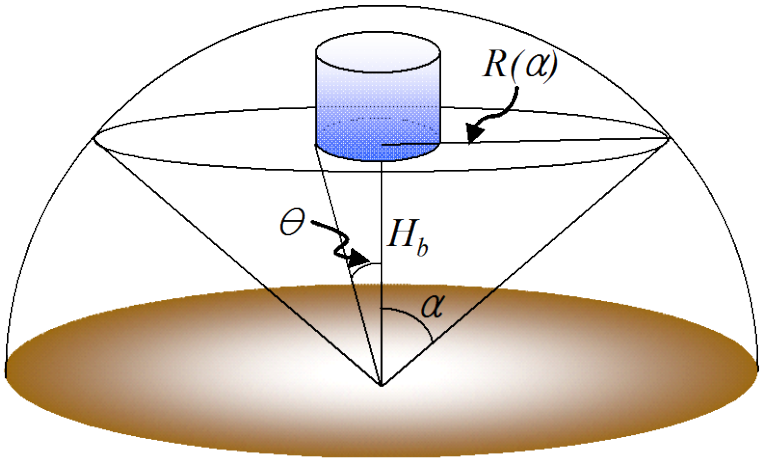


Fig. 2.Illustration of the parameters.

First, the correction for cloud position information is performed through the following process:In the case of clouds located over the zenith, the weight value need not be considered according to cloud position. The weight value of clouds are 0. However in the case of clouds located at theedge(i.e., α=α*= 80°), the same clouds tended to be overestimated than the zenith and the weight value becomes 1. Therefore, in cases of clouds locatedbetween the zenith and edge, the weight value isbetween 0 and 1.

After deciding the presence of clouds based on cloud detection information for basic input into the algorithm, the cloud is assigned the weight value depending on ground observers and the position of cloud based on the pixels.

In order to obtain the weight value depending on geometric information of clouds, it considers the horizontal area and vertical height of clouds through the following process: The CA(Nnadir(α)) observed from satellites is defined as equation (1) for horizontal area that is occupied by clouds for the area of a circle with a radius of R(α)(Figure 2). Ground-observedCA(Nhemisph(α)) is dependent upon the zenith angle from the area occupied by the cloud for the solid angle.

$$N_{nadir}(\alpha) = \frac{S_{cld,nadir}(\alpha)}{S_{nadir}(\alpha)} \quad (\%) \qquad (1)$$

$$N_{hemisph}(\alpha) = \frac{S_{cld,hemisph}(\alpha)}{S_{hemisph}(\alpha)} \quad (\%) \qquad (2)$$

Where $S_{cld,nadir}(\alpha)$ = horizontal cloud area (km), $S_{nadir}(\alpha)$ = total area of the circle (km), $S_{cld,hemisph}(\alpha)$ = fraction of the solid angle filled by clouds (sr), $S_{hemisph}(\alpha)$ = observed solid angle with cone zenith angle 2α(sr) are defined. The relationship between CA and ground-observed CA can be expressed as equation (3).

$$N_{hemisph}(\alpha) = N_{nadir}(\alpha)\left[1 - \gamma \frac{\alpha}{\alpha *}\right]^{-1} \qquad (3)$$

Where $\frac{\alpha}{\alpha *}$ = cloud weight, γ = cloud aspect ratio, and α^*, in this algorithm the observer supposed 80°to be the maximum zenith angle.

In order to retrieve the cloud aspect ratio the horizontal and vertical size must be considered.First of all, the horizontal area(D) is calculated considering N_{nadir} retrieved into pixel unit and area of pixels. When considering the vertical height (H) of clouds, after dividing into low/medium/high clouds, Cloud Top Pressure (CTP) is used to determine the geometric thickness.

Cloud type is based on ISCCP, which is divided into low, medium, and high clouds depending on the CTP. This criteria is listed in Table 2. Bottom cloud height is determined depending on the criteria listed in Table 2. The cloud geometrical thickness is obtained by subtracting the cloud bottom height from the cloud top height. Thus, in low, medium, and high clouds, median CTP pressures of 300, 500, and 850 hPa are applied, respectively. The geometrical thickness for horizontal area of the retrieved cloud is defined by the cloud aspect ratio.

The algorithm consists of three parts retrieved by ground observed CA: Weighting value of the cloud, cloud-observed CA, and cloud aspect ratio.Ground observed CA is retrieved in cloud/clear conditions of averaging the surrounding 7x7 pixels converting by the method of Kassianov et al.

(2005). More detail contents can be found in pages 2-3 of the appendix.

Table 2. Cloud top pressure and cloud bottom height corresponding to cloud type.

Level	Genera	Cloud top pressure (hPa)	Cloud bottom height (km)
High	Cirrus, cirrocumulus, cirrostratus	50-440	8
Middle	Altocumulus, altostratus, nimbostratus	440-680	4
Low	Cumulus, stratocumulus, stratus cumulonimbus	680-1000	1

3.3 Retrieval process

3.3.1 CA retrieval

The retrieval process of this algorithm is presented in Figure 3. The necessary input data to estimate Satellite-observed CA and ground-observed CA are divided into basic input data and cloud analysis data.

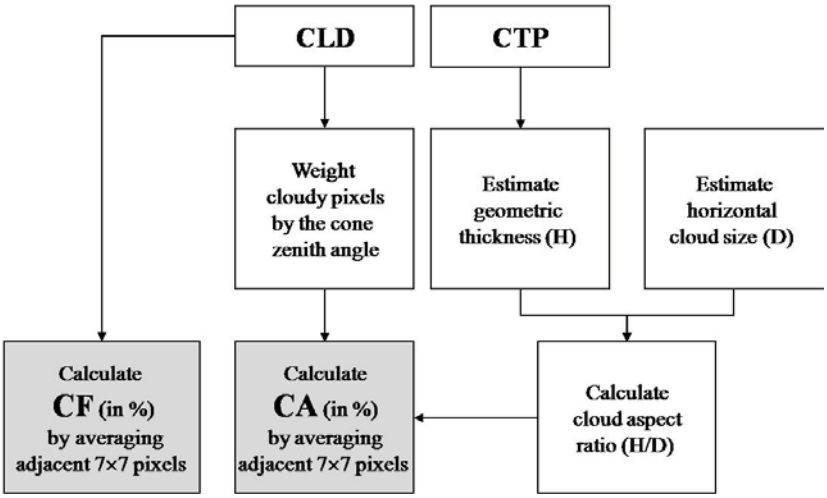
The basic input data is the weighting value information of clouds depending on cloud detection, scene analysis data, and the position of the observer.

Cloud analysis data needs CTP data.

After deciding the presence of clouds based on cloud detection information, the cloud is assigned the weight value depending on ground observers and the position of cloud by pixel detection. CA of satellite (N_{nadir}) is calculated on the average cloud/clear pixel of the surrounding 7x7. We calculate the cloud aspect ratio(γ) considering the horizontal and vertical size of cloud.

After inputting the weighted value of the clouds, cloud-observed CA, and cloud aspect ratio, ground-observed CA is retrieved.The final product for ground-observed CA of cloud/clear condition was generated by averaging the surrounding 7x7 pixels and converting by the Kassianov et al. (2005) method.

Fig. 3. Flowchart of Cloud Fraction (CF) and Cloud Amount (CA) algorithm



In order to determine satellite-observed CA, the test deciding clear sky confidence level including scene analysis has to be preceded. If clear confidence level is too high, the number of pixels classified according to clear pixel aredecreased and the CA is increased.If clear confidence level is too low, the number of pixels classified according to clear pixel are increased and the CA is a decreased. The case analysis is compared with visible image data from the COMS algorithm experiment test in order to decide a suitable clear confidence level need. The confidence level will be within 95%.

3.3.2. QC Flag

The position information of clouds and the horizontal and vertical size of clouds influence retrievalof ground-observed CA. Therefore, it presents each QC flag for two processes, which is shown in table 3. First, it presents up to 96-240 for the ratio of weight value (i.e., weight_rate=sum of cloudy pixel weight / total weight in 7x7) with the flag for position information of cloud which derived the cloud pixels for total weighting value in the 7x7 matrix. It is designed to provide with QC up to 1-10 for cloud aspect ratio(H/D).

Table 3.QC Flag.

CLA - CA		
bit	**Bit Interpretation**	**Field Description**
8~5 (Pixel weights in terms of the cone zenith angle.) $\frac{\sum^{M} W_{ij}}{\sum^{N} W_{ij}}$ **unavail => 0**	240 224 208 192 176 160 144 128 112 96	0~0.1 0.1~0.2 0.2~0.3 0.3~0.4 0.4~0.5 0.5~0.6 0.6~0.7 0.7~0.8 0.8~0.9 0.9~1
4~1 (Estimated Cloud aspect ratio) $\frac{H}{D}$ **unavail => 0**	10 9 8 7 6 5 4 3 2 1	0~0.1 0.1~0.2 0.2~0.3 0.3~0.4 0.4~0.5 0.5~0.6 0.6~0.7 0.7~0.8 0.8~0.9 0.9~1

3.4 Validation

3.4.1 Validation method

Satellite observation CA and ground-observed CA retrieved from COMS are performed in two ways. First is operational validation by CMDPS.Second, it has itsown validation to see the effectiveness of the algorithm This method needs the averaging process 30kmx30km(6x6 pixels) to adjust the spatial resolution.

In the case of ground-observed CA, after it adjusts time information between location information of ground-observed CA and of satellite observation CA, validating the effectiveness of product on R, Bias, and RMSE between the two products.

3.4.2 Validation data

(1) CMDPS Validation (COLL/VAM)

The data used to validate CMDPS satellite-observed used the data of CA MODIS Terra and Aqua and ground-observed CA validated from November 1 to November 5 using Ground and GTS data. To determine the weakness of the algorithm, we calculated the statistic (Correlation, BIAS, RMSE) value dividing into each latitude (the equator : below latitude 30°, middle latitude: the South-North: 30-60°). Also, in CA, unlike in other products, the validation performed included clear sky.

(2) Developer validation

In order to validate satellite observation CA, MODIS cloud product (MOD06) with spatial resolution(270x406 pixels) of 5x5 km is needed. In the case of ground-observed CA, we used ground observation data in the SGP of the Atmospheric Radiation Measurement (ARM) program (available online at http://www.arm.gov). In the case of ARM data, the total sky imager(TSI) was used for measuring sky cover. It can derive more objective ground observation CA. We validated the effectiveness based on data from August to December 2002 in the SGP site located mid-latitude similar to the location of our country.

We used it to validate ground-observed CA at KMA, in Korea for a one month period in December 2007.

3.4.3 Temporal and spatial collocation method

(1) CMDPS Validation (COLL/VAM)

To validate thesatellite observation CA algorithm, we excluded the validation in high latitude regions above 60° north and south.

We excluded the validation that is represented the difference more than 1-standard deviation of 5x5 pixels in MODIS to validate in the case of homogeneous position for temporal and spatial collocation

Ground data signifies the ground-observed CA at KMA. GTS are compared with cloud amount entering within ±0.1°the observed data every hour on the hour from ground observation data. In the case of the ground-observed CA, data more than 20 % are assumed to be navigation and cloud detection errors and are compared with pixels within 20% for line equivalent of 1:1.

(2) Developer validation

We corresponded with temporal and spatial collocation to validate ground-observed CA retrieved through this algorithm and ground satellite CA in the SGP site. We compared average ground-observed CA retrieved through satellite observation CA and the algorithm to enter within ±30 minute, ±0.1 degreefrom the location of ground observation.

3.4.4 Validation result analysis

(1) CMDPS validation (COLL/VAM)

Table 4. Validation results of CF and CA

	Reference	Time	Region	R	Bias	RMSE
CF	MOD	11/1~11/5	Global	0.85	-3.57	21.68
	MYD			0.81	-3.83	22.30
	MOD		Low	0.85	-5.50	21.48
	MYD			0.81	-5.84	23.56
	MOD		Mid	0.77	-1.52	21.68
	MYD			0.78	-1.86	20.90
CA	GTS	11/1~11/5	Global	0.95	0.18	0.81
	GROUND			0.75	0.61	1.95
	GTS		Low	0.95	0.23	0.83
	GROUND			-	-	-
	GTS		Mid	0.91	0.19	0.78
	GROUND			0.75	0.61	1.95

Table 4 shows the results in comparison with CA derived from satellite observation CA of MODIS, Ground, and GTS. First, it shows high correlation coefficients over 0.8 in all regions if we look at the results of satellite observation CA. When we look at the results of validation by latitude, the difference of latitude did not influence the results of validation and it was the same as the ground-observed CA.

The results show better correlation coefficients for GTS observation results than ground. In the case of ground, as the results observed by ground-observed CA, the objective reliability must be given attention to decline further than GTS.

(2) Developer validation

Figure 4.shows the relationship between ground observation CA of SGP site and satellite observation CA derived from MODIS(a), and CA converted through the algorithm, respectively. According to Berendes et al.(2004), a disparity of within ±20 % represents “good agreement" between Satellite and Ground observation CA.

As shown in Figure 4, the results show a correlation coefficients of 0.96 between ground and satellite observation CA within ±20 % and represent a correlation coefficients 0.98 of Ground-observed CA converted through ground observation CA and the algorithm.

When the overlapped large circles in Figure 4(b) compared with ground observation CA, the percentage of the improved case for CA retrieved through the algorithm was found to be 53%.

Fig. 4. Validation results of cloud fraction(a) and sky cover(b) for the SGP site. Larger circles in b represent improved$N_{hemisph}$values compared with N_{nadir} in the site.

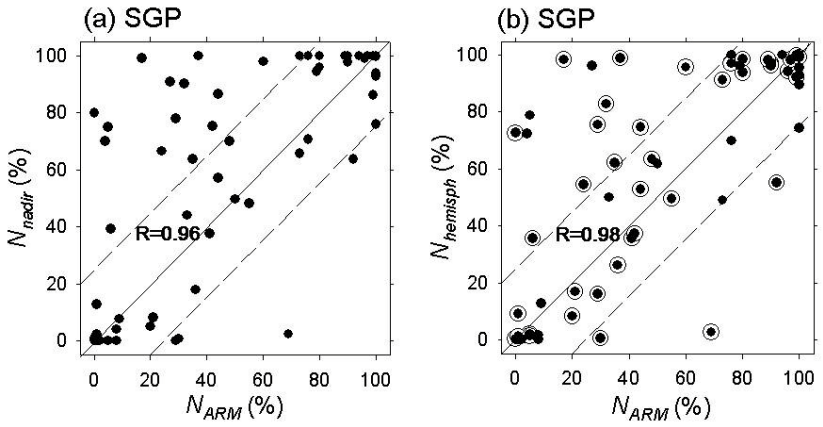


Figure 5 shows the difference between CA converted for ground and satellite observation CA, and ground observation CA and algorithm. The result is a relatively underestimated CA in ground observation. Thin clouds are widely distributed in mid-latitude region. ground-observed CA at the

ground seem to not exactly measure for thin cirrus.To convert ground-observed CA for algorithm from results of Fig. 4 and Fig. 5, the correlation coefficients with ground observation CA is high and the difference is decreased with real ground observation CA. CA retrieved through this algorithm and ground-observed CA for one month, December at KMA, the correlation coefficient and bias are 0.61, 0.57, and RMSE denotes 3.17. This results shows that this algorithm is valid on the Korean peninsula for smaller than RMSE (0.84) in the SGP region.

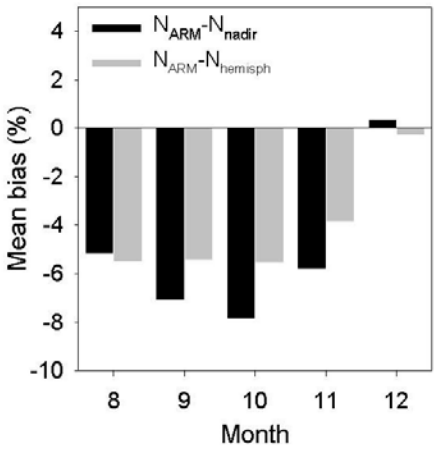


Fig. 5.monthly mean bias between ground measured cloud fraction and MODIS retrieved cloud fraction (algorithm-retrieved sky cover).

4. Analysis method of retrieval results

The CA indicates the percentage of precision and accuracy are 10.

Table 5.Detailed Output data for the CA algorithm.

OUTPUT DATA								
Parameter	**Mnemonic**	**Units**	**Min**	**Max**	**Prec**	**Acc**	**Res**	**To**
Mean cloud vertical size	mean_cld_vert						pixel	MCV
Mean cloud horizontal size	mean_cloud_hori						pixel	MCH
Cloud aspect ratio	cld_asp_ratio		0	10				CASR
Ratio of cone zenith angles	pos_weight		0	1				RA
Cloud Fraction	Cloud_fraction	%	0	100			pixel	CF

Sky cover	sky_cover	%	0	100	10	10	20km	CA

Prec: Precision, Acc: Accuracy, Res: Resolution

5. Problems and possibilities for improvement

In this algorithm, on account of cloud detection and CTP results used as input data, the accuracy of this algorithm depends on the accuracy of this input data. In general, in the case of ground observation, it does not measure thin cirrus at high altitudes and can affect ground-observed CA.

Table 6.Quality test result for the CA algorithm.

Quality flag			
Parameter	bit	Value	Meaning
sky cover	7	from 0 up to100 ; step: 1	undefined
define illumination and viewing conditions	3	0	undefined
		1	night
		2	twilight
		3	day
		4	sunglint
describe the quality of the processing itself	2	0	non processed
		1	good quality
		2	poor quality
		3	bad quality

6. References

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

Kassianov, E. I., Long, C. N. and Ovtchinnikov, M. 2005: Cloud sky cover versus cloud fraction: whole-sky simulations and observations, Journal of Applied Meteorology, 44, 86-98.

Oh, H. R., Choi, Y. S., Ho, C. H. and Ahn, M. H., 2006: Development of sky cover retrieval algorithm for the Communication, Ocean and Meteorological Satellite (COMS) imagery, Journal of Korean Meteorology Society, 42, 389-396.

7. Appendix

JOURNAL OF THE KOREAN METEOROLOGICAL SOCIETY, 42, 6, 2006, p.

Development of Sky Cover Retrieval Algorithm for the Communication, Ocean and Meteorological Satellite (COMS) Imagery

Oh, Hye-Ryun[1], Yong-Sang Choi[1] and Chang-Hoi Ho[1] and Myoung-Hwan Ahn[2]

[1]*School of Earth and Environmental Sciences*
Seoul National University, Seoul, Korea
[2]*Meteorological Research Institute*
Korea Meteorological Administration, Seoul, Korea

(Manuscript received 31 October 2006; in final form 13 December 2006)

Abstract

We have developed a retrieval algorithm to estimate the sky cover of clouds in the celestial sphere centered at a ground observer. In this algorithm, the sky cover is obtained from the satellite-observed (i.e., nadir-viewing) cloud fraction through two procedures: assigning the weights of the ratio of the solid angle filled by clouds to the entire solid angle and estimating the cloud aspect ratio, i.e., the geometric thickness vs. the horizontal cloud size. The sky cover retrieved from the cloud fraction of the Moderate Resolution Imaging Spectroradiometer is validated by ground-observed sky cover at two sites of the Atmospheric Radiation Measurement Program, the tropical western Pacific (12.42°S and 130.89°E) and the southern Great Plains (36.60°N and 97.48°W). The validation shows that the present algorithm provides a successful conversion into sky cover, particularly in the latter site.

Key words: sky cover, cloud fraction, COMS, MODIS, ARM

1. Introduction

Cloudiness deeply influences our daily life such as in sailing, aviation, agriculture, and outdoor activities. In meteorology, cloudiness determines the magnitude of incoming solar energy which regulates the entire climate system over the globe. It is a key parameter in controlling the emission of terrestrial radiation at the top of the atmosphere. In addition, cloudiness is one of the essential objects of weather forecasting, so it has been measured routinely by ground observers. As the importance of cloudiness increases in the daily life, in Korea, the general public requests the digitized values of cloudiness nation-wide and in the neighboring oceans.

A longer than two-decade record of the nadir-viewing cloud fraction, i.e., the ratio of the total area of cloud cross-section in a horizontal plane to the total horizontal area, has been archived with an aid of many satellite programs, e.g., the International Satellite Cloud Climatology Project (ISCCP) (Schiffer and Rossow, 1983). Recently, cloud detection skill has been dramatically improved by the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument, particularly for thin cirrus (Ackerman *et al.*, 1998). Over the regions of Asia and western Pacific, a series of satellites including the *Japanese* Multifunctional Transport Satellite (MTSAT) and the *Chinese* Feng Yun have monitored cloudiness. Furthermore, Korea will launch the first Korean geostationary satellite, the Communication, Ocean and Meteorological Satellite (COMS), in 2008 (Ahn *et al.*, 2005). The COMS will provide detailed cloud information in real time, predominantly for the cases of severe weather over the Korean peninsula and adjacent regions every 10 minutes or even shorter time intervals.

The cloudiness observation by human observers (hereafter sky cover) is, in fact, fairly subjective in

Corresponding Author: Chang-Hoi Ho, Climate Physics Laboratory, School of Earth and Environmental Sciences, Seoul National University, Seoul 151-742, Korea.
Phone : +82-2-880-8861, Fax : +82-2-876-6795
E-mail: hoch@cpl.snu.ac.kr

a local area because it represents the fraction of the solid angle of clouds estimated by humans. It also has a limitation, in that it is susceptible to daytime measurement only. For this reason, the foregoing satellite remote-sensing of cloudiness, with continuous and objective measurements for a vast area, has been generally employed instead of sky cover. However, the cloud fraction observable from satellites reflects the view from space, which differs from the view of humans on the ground. Strictly, the general public is interested in the amount of sky cover, not the satellite-viewing (i.e., nadir-viewing) cloud fraction. Despite of this necessity, a satellite algorithm for calculating the ground-observed sky cover has not been developed yet.

This study attempts to develop the sky cover algorithm, which will be adapted to the COMS, and validate the resultant algorithm. For this purpose, the MODIS cloud product is used as input data for the algorithm in substitution for the COMS data. The ground-observed sky cover from the Atmospheric Radiation Measurement (ARM) program is employed to validate the algorithm. The algorithm overview and data utilized are described in sections 2 and 3, respectively. The validation results of the sky cover with the ARM data are discussed in section 4. Finally, a summary is given in section 5.

2. Algorithm overview

Prior to describing the sky cover algorithm, we discuss the reasons for the discrepancy between satellite-observed cloud fraction and ground-observed sky cover. Kassianov *et al.* (2005) pointed out two causes: the location of the observer and cloud geometry. First, the effect of the curvature of the celestial sphere is included in the ground-observed sky cover. In general, the value of sky cover is larger when clouds locate at the zenith when compared to the same amount of clouds distributed along the horizon. Second, ground observations detect both the bottom and the side of clouds, while satellites do cloud top only. In other words, the ground-observed sky cover is a variable dependent on the cloud geometry, while satellite-observed cloud fraction is obtained from the nadir-viewing alone.

Allowing for the above-mentioned reasons, Kassianov *et al.* (2005) suggested an equation converting from the ground-observed sky cover to the satellite-observed cloud fraction based on Monte Carlo simulations. The input data is the ground-observed hemispherical measurements of cloud amount. Their approach aims at validating the satellite-observed cloud fraction with ground-observed sky cover. Since we attempt conversely to estimate ground-observed sky cover from satellite-observed cloud fraction, we modify their equation into its inverted form,

$$N_{ground}(\alpha) = \frac{N_{nadir}(\alpha)}{[1-\gamma(\alpha)\dfrac{\alpha}{\alpha_{max}}]} \qquad (1)$$

where α is the cone zenith angle between the zenith and the cloud. This value varies from 0 to a maximum value, α_{max} of 80°. $N_{ground}(\alpha)$ is the ground-observed sky cover, $N_{nadir}(\alpha)$ the nadir-viewed cloud fraction, and $\gamma(\alpha)$ the cloud aspect ratio, i.e., the ratio of the vertical to the horizontal cloud sizes. α/α_{max} indicates the position of a cloud, which is proportional to the distance from the zenith. Thus we can readily construct a weighting matrix, in which the weighting value (α/α_{max}) is zero at the center (α=0°) and unity at the edge ($\alpha=\alpha_{max}$=80°).

The sky cover algorithm is constructed based on Eq. (1) and its flow chart is presented in Fig. 1. The retrieval of sky cover in the algorithm is performed for each pixel by using the information (e.g., the positions of clouds and their aspect ratios) of the surrounding 5×5 pixels. As shown in the figure, the algorithm consists of three parts: the calculation of $N_{nadir}(\alpha)$, the estimation of $\gamma(\alpha)$, and the combination of a lookup table of α/α_{max}. First, $N_{nadir}(\alpha)$ for each pixel is calculated by simply averaging over adjacent 5×5 pixels from cloud masking information. Note that there is not a falloff in spatial resolution through this process for the COMS. Second, $\gamma(\alpha)$ is calculated by the ratio of the vertical cloud size to the horizontal cloud size for each pixel. Third, the pre-calculated lookup table of the weighting values

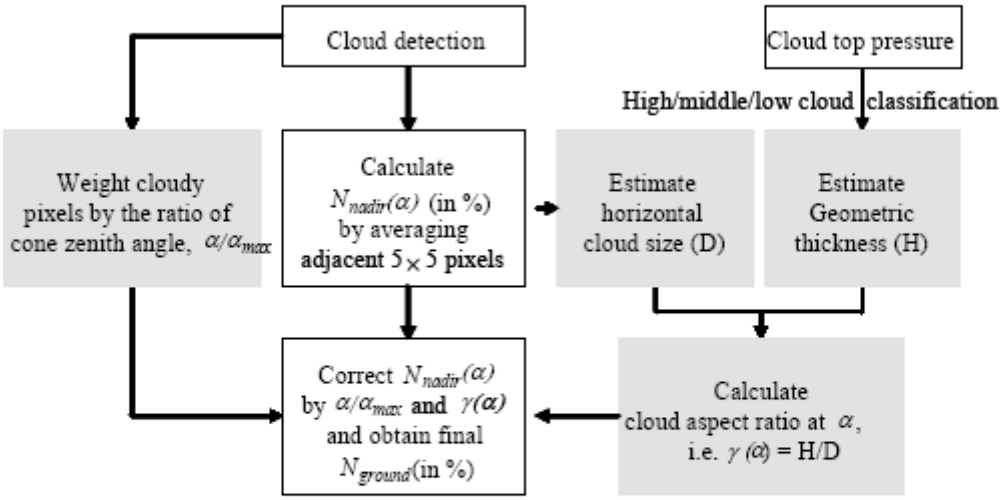


Fig. 1. Flow chart of the sky cover algorithm.

(α/α_{max}) in the 5×5 matrix are multiplied by the corresponding 25 $\gamma(\alpha)$s. After determining all required parameters in Eq. (1), $N_{ground}(\alpha)$ can be obtained from the equation. For the particular case when cloud exists at the zenith only, $N_{ground}(\alpha)$ is equal to $N_{nadir}(\alpha)$; the weight is zero at the zenith.

For estimating $\gamma(\alpha)$, the cloud sizes in the vertical and horizontal directions cannot be directly measured by remote sensing. Here, we have applied two approaches to obtain its value. First, the cloud sizes are derived using satellite-retrieved products such as cloud optical depth (τ_e), effective particle radius (r_e), and cloud top pressure (CTP). In detail, we calculate the liquid water path (LWP) and liquid water content (LWC), respectively, by using the typical formula ($1.5\tau_e r_e$) and the climatological look-up table introduced by Nakajima and Nakajima (1995). By taking LWP/LWC, the cloud geometrical thickness is determined. As an alternative method for calculating $\gamma(\alpha)$, we use CTP only. It is noted that the cloud size in the horizontal direction can be simply estimated by cloud fraction, regardless of cloud type. Here, CTP is used to resolve cloud type (e.g., high, middle, and low cloud) based on ISCCP. The climatological values of cloud height at the top and bottom are well established for various clouds (Table 1; see Hahn *et al.* (2001) for details). The cloud geometrical thickness is obtained by subtracting the cloud bottom height from the cloud top height.

3. Data

Two sets of data are analyzed in the present study: the MODIS cloud product (MOD06) and the ARM ground observations. MOD06 contains τ_e, r_e, CTP, and cloud fraction (Platnick *et al.*, 2003). In MOD06, the cloud fraction is calculated from 5 km×5 km cloud mask pixel groupings. It is noted that MODIS data provide bands for cloud detection including thin cirrus cloud (Choi *et al.*, 2005), which can fulfill the high accuracy for cloud fraction required in this study. While the COMS data is not available yet, MOD06

Table 1. Cloud top pressure and cloud bottom height corresponding to cloud type.

Level	Genera	Cloud top pressure (hPa)	Cloud bottom height (km)
High	Cirrus, cirrocumulus, cirrostratus	50-440	8
Middle	Altocumulus, altostratus, nimbostratus	440-680	4
Low	Cumulus, stratocumulus, stratus cumulonimbus	680-1000	2

is used as a surrogate for the COMS data.

We use ground observation data taken from two ARM program sites in the tropical western pacific (TWP, 12.42°S and 130.89°E) and the southern Great Plains (SGP, 36.60°N and 97.48°W). The TWP and SGP may represent a tropical marine and mid-latitude continental region, respectively. The instrument for measuring sky cover is the total sky imager (TSI). The TSI provides the time series of hemispheric sky images during the daylight hours and the retrieval of fractional sky cover for the periods when the solar elevation is greater than 10 degrees (available online at http://www.arm.gov). The data period at the TWP site is the first fifteen days of each month from August through December 2002. At the SGP site, the data for first fifteen days in both March and June are additionally obtained. These two ground observations are used to validate the retrieved sky cover.

4. Results

Returning to the issue of estimating $\gamma(\alpha)$, we find which of the methods, either using τ_e and r_e or CTP, is appropriate for the sky cover algorithm. For this purpose, we have compared the observed and the retrieved sky covers using both the τ_e/r_e method and the CTP method. Table 2 shows the correlation coefficient between the two methods and the root-mean-square error (RMSE) at the two sites. The RMSE is calculated to examine systematic and random errors. The correlation coefficients denote significantly high values, larger than 0.9, at the two sites indicating that the present retrieval algorithm is successful. The RMSEs for the two methods, respectively, have values 10-14 at the TWP site and approximate 8 at the SGP site.

Table 2. Correlation coefficients and RMSEs between sky covers from the CTP method and the τ_e/r_e method at SGP and TWP sites.

		CTP method	τ_e/r_e method
TWP	Correlation	0.94	0.90
	RMSE	10.74	13.56
SGP	Correlation	0.98	0.99
	RMSE	8.04	8.11

Based on the above results from the available data sets, it is suggested that the CTP method would be better than the τ_e/r_e method. Besides, Platnick *et al.* (2003) documented that two variables of τ_e and r_e contain errors which may be related to thin cirrus clouds. Furthermore, the τ_e/r_e method takes much time to calculate the geometrical thickness of the cloud because of its steps being more complicated than those of the CTP method. Because of these reasons, we selected the CTP method (see Fig. 1).

In the next two sub-sections, we first introduce the result of a sensitivity test examining the size of the weight matrix in the sky cover algorithm illustrated in Fig. 1. In other word, it explains why we select 5×5 pixels for retrieving sky cover. Using the optimal matrix size given by the sensitivity test, the sky cover is obtained and validated against the ground-observed data from the two sites.

a. Sensitivity test

The sensitivity test for the cloud weight is performed to derive the best agreement with the ground-observed sky cover in the ARM site. In the procedure, we have assumed that a ground observer can inspect the cloud base height only. Since the cloud base height is generally 2 km for low-level clouds, the maximal observable lateral size of the clouds (i.e. the case of $\alpha = \alpha_{max}$) is about 25 km. Considering one pixel in MODIS cloud fraction corresponds to 5 km, the distance 25 km in all directions fits to 5×5 pixels. For the sensitivity test of pixel numbers, the maximum matrix size is chosen to be 15×15 pixels due to the following two reasons. First, a matrix size of 15×15 pixels is adequate when the maximum sight distance for which observation can take place is 50 km. Second, longer time is required to calculate the sky cover on pixels and in the buffer zone on the edge as the matrix size increases. Therefore, the sky cover is retrieved as the matrix size increases with sizes of 5×5, 7×7, 9×9, 11×11, 13×13, and 15×15 pixels.

The correlation coefficients and RMSEs for various sizes with the ARM data are, respectively, calculated and compared with one other. The results are

given in Table 3. As seen in the table, the coefficients have values between 0.94 and 0.85 indicating quite high correspondence. Similarly, RMSEs show a relatively large difference from 18.15 to 11.81 depending on matrix size. These imply that RMSEs decrease when the size is reduced. As a summary, a 5×5 matrix shows the best performance, so it is selected in this study.

b. Validation

A validation of the sky cover retrieved through the algorithm is recorded in Fig. 2. In the TWP site, scatter plots of the ARM observation data vs. both the

Table 3. Correlation coefficients and RMSE of sensitivity test to the cloud weight part in the sky cover algorithm according to $R_{i\times i}$. The symbol "i" means the number of pixels, which is matrix size.

		$R_{5\times5}$	$R_{7\times7}$	$R_{9\times9}$	$R_{11\times11}$	$R_{13\times13}$	$R_{15\times15}$
TWP	Correlation	0.94	0.93	0.91	0.89	0.87	0.85
	RMSE	11.81	12.78	14.01	15.49	16.82	18.15
SGP	Correlation	0.94	0.94	0.93	0.92	0.90	0.90
	RMSE	13.34	13.57	14.00	14.88	16.16	17.25

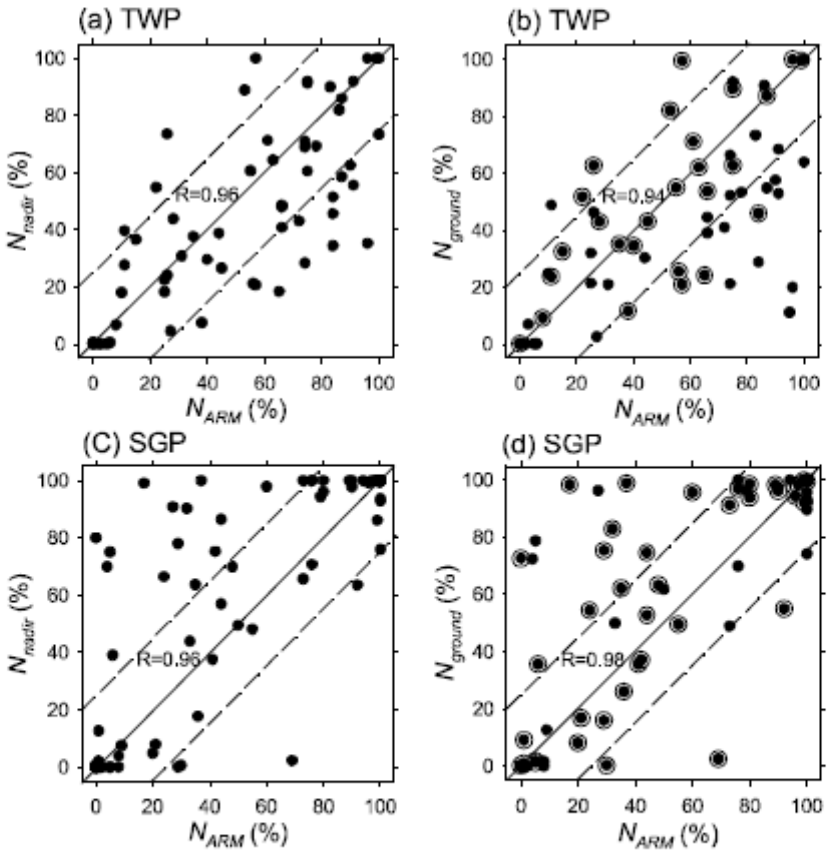


Fig. 2. Validation results of cloud fraction (a and c) and sky cover (b and d) for TWP (upper panel) and SGP sites (lower panel). R represents the correlation coefficient. Larger circles both b and d represent improved N_{ground} values compared with N_{nadir} in each site.

MODIS cloud fraction and the retrieved sky cover are presented in Figs. 2a and 2b, respectively. The same scatter plots for the SGP site are shown in Figs. 2c and 2d. In the figures, the solid lines denote the one-by-one line, along which the cloud fraction and sky cover have the same percentages as the ground observation data. The dashed lines represent that the differences between cloud fraction and ground observation or sky cover and ground observation fall within ±20%. According to Berendes *et al.* (2004), a disparity of within ±20% represents "good agreement" between the two retrieval methods. The results clearly indicate that the sky cover algorithm produces very reliable sky cover retrievals compared to the ARM data. In particular, in Figs. 2b and 2d, large open circles show the improved cases when compared with the cloud fraction of MOD06. The percentages of the improved cases for the SGP and TWP sites are found to be 53% and 46%, respectively.

We regard cases which have a discrepancy of within ±20% between cloud fraction and ARM data as reasonable. Under such considerations we calculate the correlation coefficients and RMSEs for the cases of good agreement between the ARM observation and the MODIS cloud fraction. The results show a strong agreement as a correlation coefficient of over 0.94 within ±20% range in both sites (Fig. 2). RMSEs are, respectively, 11.35 and 7.66 for sky cover in opposition to 9.51 and 11.65 for the cloud fraction from the TWP and SGP sites. This means that the sky cover algorithm may be more suitable for the SGP site compared with the TWP site.

As shown in Figs. 2b and 2d, 17 (21) cases are improved through the algorithm while 20 (19) cases are not improved for the TWP (SGP) site. Therefore, case studies are performed for which unimproved cases are selected in both sites and a scene analysis is performed in the following.

Figures 3a and 3b (3c and 3d) represent images which are recorded at 0210 UTC 10 November 2002

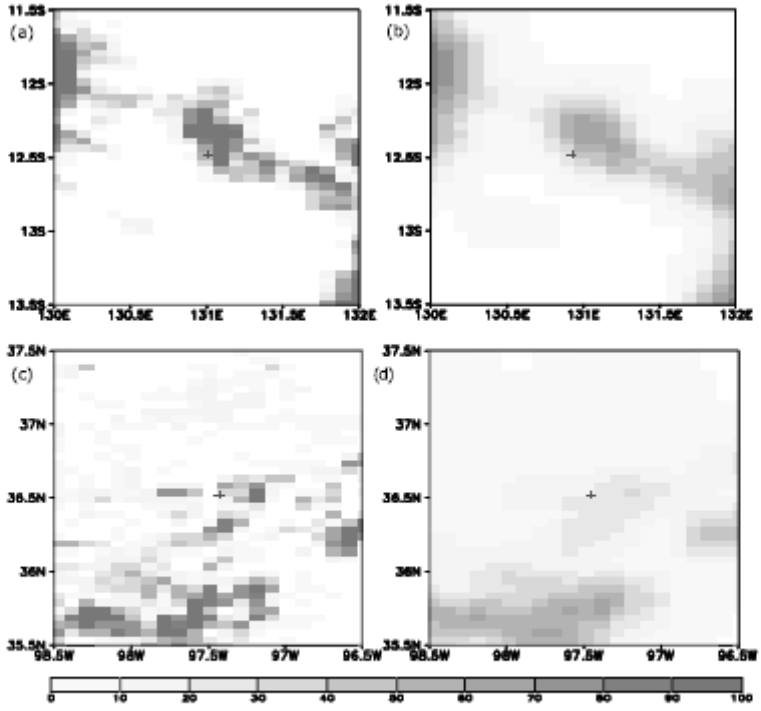


Fig. 3. MODIS cloud fraction (a and c) and algorithm-retrieved sky cover images (b and d) of a 100 km radius centering around geographical position of the TWP and SGP sites. These scenes respectively recorded at 0120 UTC 10 November 2002 (1725UTC 1 September 2002) at TWP (SGP) sites. The each site is outlined as a cross in the figures.

(1725 UTC 1 September 2002) at the TWP (SGP) site. Each figure is a shaded image in a box which size is 200 km × 200 km. The locations of TWP and SGP sites are depicted as a red cross within the figure. Figures 3a and 3c represent the images of the MODIS cloud fraction, and Figures 3b and 3d the algorithm-retrieved sky cover images in the TWP and SGP sites, respectively. The results show a distinct feature from the scene analysis; most of the cloud for bad retrieval cases is scattered and located on a boundary between cloudy and clear regions. Therefore, the reason for the bad retrieval result is considered to be connected with the sky cover algorithm. We infer that particularly the correction by the position of clouds becomes inaccurate for such scattered clouds because many pixels are not fully overcastted by them.

Figures 4a and 4b show, respectively, the monthly mean bias for the TWP and SGP sites. In the each figure, black bars (gray bars) represent bias between N_{ARM} and N_{nadir} (N_{ARM} and N_{ground}). Above all, there is positive bias in the TWP site, and a negative bias in the SGP site. In other words, there is an underestimated MODIS cloud fraction in the TWP site although the SGP site shows the opposite trend. To find out the reason for the feature, we analyze ARM ground observation data in detail (figure not shown). The result shows that opaque cloud occupies a relatively high percentage for the TWP site. By contrast, there is a greater proportion of relatively thin cloud at the SGP site. Also it is found that the monthly mean bias for the TWP site is larger than for the SGP site. The larger bias at the TWP site than at the SGP site may be attributed to the cloud aspect ratio of the algorithm. The frequent convective clouds at the TWP site cause larger mean cloud aspect ratio. This plays a role in decreasing sensitivity to the position of clouds (i.e., α/α_{max}) when the cloud fraction is fixed in Eq. (1). That is, the sensitivity to the position of clouds for the TWP site where the amount of convective clouds is greater becomes less than that for the SGP. Therefore, dull correction of cloud fraction through the position of clouds for the TWP site results in the larger monthly mean bias.

5. Summary

The sky cover retrieval algorithm is newly developed to be employed in the COMS scheduled to be launched in 2008. Using this algorithm, the amount of sky cover is retrieved from the MODIS cloud fraction. The algorithm has been developed based on the position, fraction, and aspect ratio of clouds. To form the optimum algorithm, a sensitivity test to the size of a weighting matrix for the position of clouds is firstly performed. It is found that a 5×5 weighting matrix is optimal. Next, the sky cover retrieved through the algorithm is compared with ARM ground-based observation data. The result shows that the

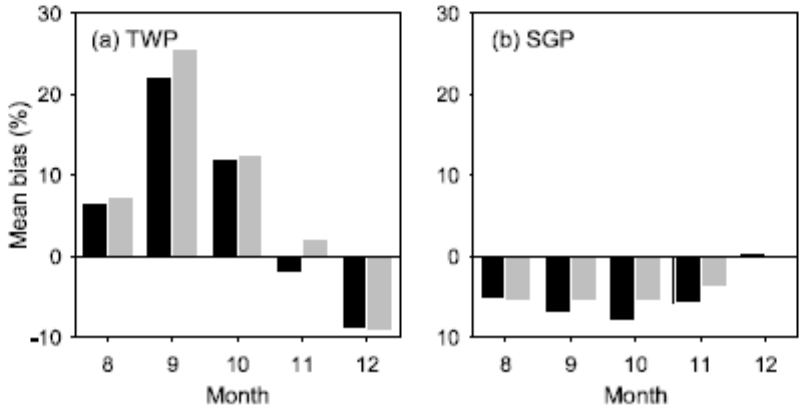


Fig. 4. Monthly mean bias in TWP site (a) and SGP site (b). Black and gray bars represent $N_{ARM} - N_{nadir}$ and $N_{ARM} - N_{ground}$, respectively.

present algorithm for converting into sky cover is relatively more successful in SGP sites. The reason for the poor retrieval cases is examined through scene analysis and monthly mean bias. In these cases, most retrieval positions are located on the edge of the cloud and are scattered from the scene analysis. In addition, the monthly mean bias result shows that relatively large mean cloud aspect ratio at the TWP site causes retrieval errors due to the algorithm insensitivity to the position of clouds. There remains further study to be done on the validation of this method using MTSAT-1R and visual sky cover of the KMA over Korea.

Acknowledgments. This study was undertaken for the COMS project funded by the Korean Meteorological Administration. The first and second authors are also supported by the BK21 project of the Korean government. The authors are grateful to two anonymous reviewers for constructive comments. The MODIS data were provided by the Earth Observing System Data and Information System, Distributed Active Archive Center, at Goddard Space Flight Center, which archives, manages, and distributes these data.

REFERENCES

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

Ahn, M.-H., E.-J. Seo, C.-Y. Chung, B.-J. Sohn, M.-S. Suh, and M. Oh, 2005: Introduction to the COMS meteorological data processing system, *Int. Sym. Remote Sens. 2005*, Jeju, Korea.

Berendes, T. A., D. A. Berendes, R. M. Welch, Member, IEEE, E. G. Dutton, T. Uttal, and E. E. Clothiaux, 2004: Cloud cover comparisons of the MODIS daytime cloud mask with surface instruments at the north slope of Alaska ARM site, *IEEE Trans. Geosci. Remote Sens.*, **42**, 2584-2593.

Choi, Y.-S., C.-H. Ho, M.-H. Ahn, and Y.-S. Kim, 2005: Enhancement of the consistency of MODIS thin cirrus with cloud phase by adding 1.6-μm reflectance, *Int. J. Remote Sens.*, **26**, 4669-4680.

Hahn, C. J., W. B. Rossow, and S. G. Warren, 2001: ISCCP cloud properties associated with standard cloud types identified in individual surface observations, *J. Clim.*, **14**, 11-28.

Kassianov, E. I., C. N. Long, and M. Ovtchinnikov, 2005: Cloud sky cover versus cloud fraction: Whole-sky simulations and observations, *J. Appli. Meteor.*, **44**, 86-98.

Nakajima, T. Y., and T. Nakajima, 1995: Wide-area determination of cloud microphysical properties from NOAA AVHRR measurements for FIRE and ASTEDX regions (1995), *J .Atmos. Sci.*, **52**, 4043-4059.

Platnick, S., M. D. King, S. A. Ackerman, W. P. Menzel, B. A. Baum, J. C. Riedi, and R. A. Frey, 2003: The MODIS cloud products: Algorithms and examples from Terra, *IEEE Trans. Geosci. Remote Sens.*, **41**, 459-473.

Schiffer, R. A., and W. B. Rossow, 1983: The International Satellite Cloud Climatology Project (ISCCP): The first project of the World Climate Research Programme, *Bull. Amer. Meteor. Soc.*, **64**, 779-784.