Investigation of Cloud Coverage Statistics in Asia Using NOAA AVHRR Time Series

Wataru Takeuchi

Institute of Industrial Science, University of Tokyo, Japan
Ce-508, 6-1, Komaba 4-chome, Meguro, Tokyo, 153-8505 Japan

Abstract

In order to compute cloud coverage statistics over Asian region, an operational scheme for masking cloud-contaminated pixels in Advanced Very High Resolution Radiometer (AVHRR) daytime data was developed, evaluated and presented. Dynamic threshold was used with channel 1, 2 and 3 to automatically create a cloud mask for a single image. Then the 10-day cloud coverage imagery was generated over the whole Asian region along with cloud-free composite imagery. Finally the monthly based statistics were computed based on the derived cloud coverage imagery in terms of land cover and country. As a result, it was found that 20-day is required to acquire the cloud free data over the entire Asian region using NOAA AVHRR. The 10-day cloud coverage and cloud-free composite imagery derived in this research is available via the web-site http://webpanda.iis.u-tokyo.ac.jp/CloudCover/.

Key words: Snow and ice, cloud shadow, short wave infrared, landcover

1. Introduction

1.1 Needs for cloud coverage statistics

In Asian region cloud cover is known to be extensive, and therefore presents a significant limitation to the availability of remote sensing data for operational monitoring of both land and sea areas. Cloud coverage statistics exist from traditional meteorological observations, mainly land based stations. Much effort have been focused on cloud detection or classification, however, cloud cover statistics are very rare (Tag et al., 2000). (Jorgensen et al., 2000) computed 20-years of cloud coverage over Denmark using 5500 scenes of Landsat TM and 1600 scenes of NOAA AVHRR, and concluded that the overall coverage ranges from 60-77% and the monthly mean average value of it is about 60 %. There has been not attempt, to authors' knowledge, to utilize satellite sensor data to derive cloud cover statistics in Asian region, the spatio-temporal patterns of cloud coverage is one of the most important aspects in land cover monitoring.

One of the most important sources of meteorological and climate data are satellite imagers. Imagers are defined as instruments that measure reflected solar and thermal radiation and are distinguished from sounder instruments by having fewer spectral channels and high spatial resolution (Simpson et al., 1996). Currently, the operational imager data on polar orbiting satellites, Advanced Very High Resolution Radiometer (AVHRR), provides five channels (0.63, 0.86, 3.75, 10.8, and 12 μm) with a global resolution of limited coverage at a resolution of 1 km over 25-years of history.

Much of the knowledge gained from the various studies into cloud detection was brought together within the International Satellite Cloud Climatology Project (ISCCP), which had as one of its goals the task of deriving automatically cloud information from AVHRR and other imagery. A number of researchers have addressed cloud classification from a variety of perspectives (Rossow et al., 1993). However, some cloud detection methods developed in Europe or USA may not be suitable for Asian region.

* Corresponding author: wataru@iis.u-tokyo.ac.jp
Phone: +81-3-5452-6407 Fax : +81-3-5452-6408

1.2 Objectives of this study

In order to compute cloud coverage statistics, an operational scheme for masking cloud-contaminated pixels in AVHRR daytime data was developed, evaluated and presented. Dynamic threshold was used with channel 1, 2 and 3 to automatically create a cloud mask for a single image. Then the 10-day cloud coverage imagery was generated over the whole Asian region along with cloud-free composite imagery. Finally the monthly based compiling was conducted based on the derived cloud coverage imagery in terms of land cover and country.

2. Methodology

2.1 Pre-processing of NOAA AVHRR data

A total of 4061 AVHRR scenes were acquired during the day-time from 2001 May to 2004 April received at Institute of Industrial Science, University of Tokyo in Japan and Asian Institute of Technology in Thailand. They were calibrated to the top-of-atmosphere (TOA) reflectances (channels 1 and 2) and brightness temperatures (channel 3-5), using a software package for noaa data analysis (PaNDA) (PaNDA, 1998). The images were geometrically corrected with ground control points and referenced to the Latitude- Longitude grid map projection with nearest neighbor re- sampling. After the screening of these inappropriate images with geometric correction errors bigger than 1 pixel, 3356 images were selected for analysis. Those images were spatially subset as rectangular area from S10 to N60 degree in latitude and from 70 to 150 degree in longitude. The processing is conducted by the web-based service available at the University of Tokyo (WebPaNDA, https://webpanda.iis.u-tokyo.ac.jp).

2.2 Automatic detection of clouds

First of all, the spectral characteristics of channel 1, 2 and 3 are investigated over fine and cloudy pixels. Figure 1 shows spectral characteristics of five types of clouds including think cloud, thick cloud, optical cirrus cloud, sea ice and land ice and snow pixels. Both the reflectance of channel 1 and 2 show highest values in thick clouds followed by land ice and snow, sea ice, thin clouds and optical cirrus. It implies that a single threshold can not differentiate between land ice and snow from clouds. Brightness temperature of channel 4 is characterized by optical cirrus, land ice and snow, thin clouds, sea ice and thick clouds in a descending order. This channel is not enough to distinguish between land ice and snow with clouds. In contrast to that, short wave infrared reflective channel 3 shows the highest reflectivity in thick cloud, thin cloud, optical cirrus, land ice and snow and sea ice. It was found that channel 3 is very effective to distinguish land ice and snow from clouds.

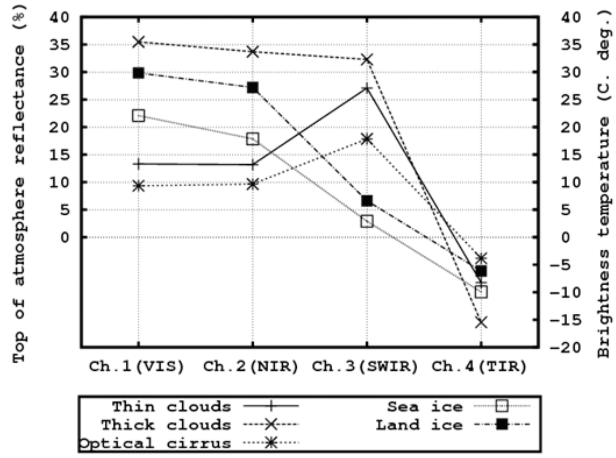


Figure 1. Spectral characteristics of five types of clouds, ice and snow.

Figure 2 shows histograms of channel 1, 2 and 3 and fine pixels are represented by solid line and cloudy pixels by dashed line based on the investigation of spectral characteristics of clouds.

As is shown in Figure 2-(a), (b) and (c), the threshold reflectance values dividing fine and cloud pixels are 15%, 17% and 25% in channel 1, 2 and 3, respectively. The ratio of channel 2 to channel 1 distributes as shown in Figure 2-(d) and high histogram values around 1.0 is classified as cloudy pixels. The simple threshold should not be suitable for each criterion because the border between fine cloudy pixels has overlapping area and they depends on the climatic conditions or coverage or region of the scene.

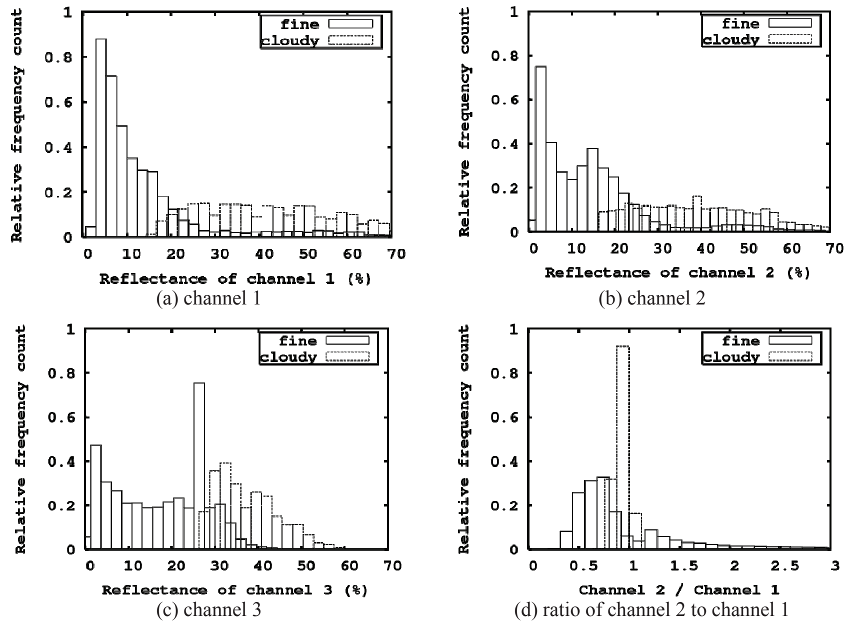
In this study, dynamic threshold technique is applied to the above mentioned four criterion (Vittorio et al. (2002)). The method is based on the simple idea that the threshold is defined by the inflection point in the overlapping area between fine and cloudy pixels on the corresponding histogram. The automatic inflection point searching is carried out on the histogram curve of each scene using initial values as shown in Equation (1), (2), (3) and (4). The pixels is classified as cloudy when it meet all the criteria shown in Equation (1) - (4) and others are as fine.

$$12.0\ (\%) < \text{Channel 1} \quad (1)$$

$$17.0\ (\%) < \text{Channel 2} \quad (2)$$

$$27.0\ (\%) < \text{Channel 3} \quad (3)$$

$$0.7 < \text{Channel 2} / \text{Channel 1} < 1.2 \quad (4)$$



(a) channel 1 (b) channel 2

(c) channel 3 (d) ratio of channel 2 to channel 1

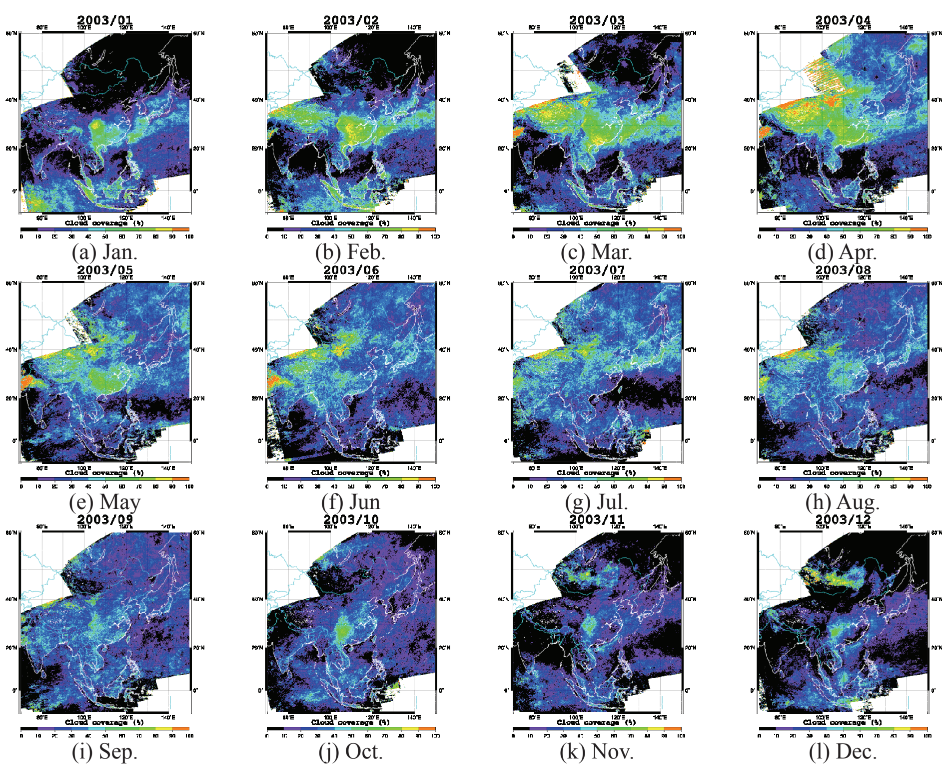
Figure 2. Histogram of NOAA AVHRR reflective channels over fine and cloudy pixels.

3. Results And Discussions

Figure 3 shows maps of monthly mean cloud coverage over Asia obtained from NOAA AVHRR imagery in 2003. The color contour represents the cloud coverage of that pixel and black area out of coverage contains no data available for further processing. The followings are investigations on the statistics of cloud coverage in terms of land cover types and countries.

3.1 Land cover based statistics of cloud coverage

First of all, land cover based statistics are computed to evaluate the derived cloud coverage. 1-km land cover map is used as a reference. It is a MODIS derived land cover product conducted by the Boston University's group and including 16 classes; evergreen needleleaf forest, evergreen broadleaf forest, deciduous needleleaf forest, deciduous broadleaf forest, mixed forest, bush, forested savanna, savanna,



(a) Jan. (b) Feb. (c) Mar. (d) Apr.

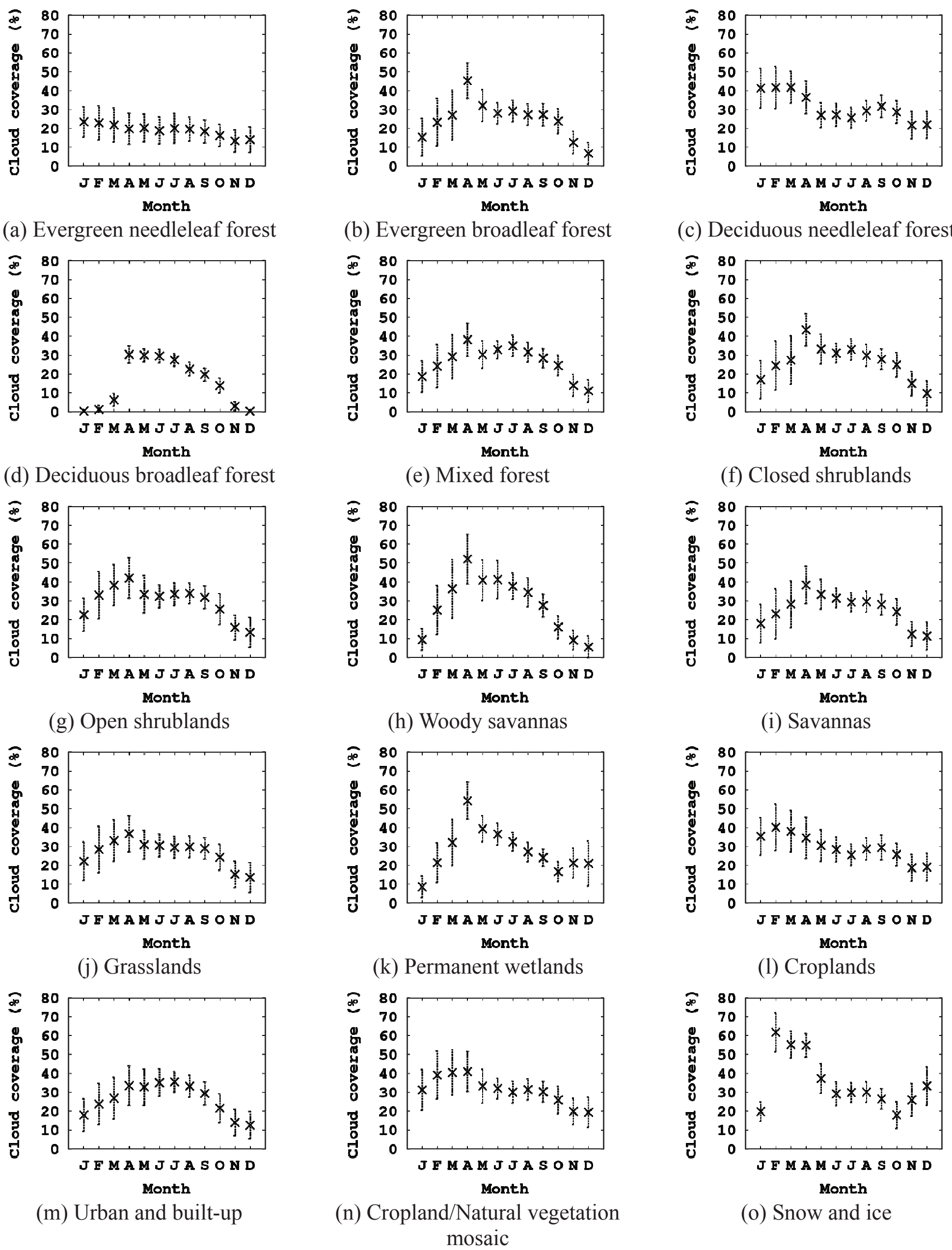
(e) May (f) Jun (g) Jul. (h) Aug.

(i) Sep. (j) Oct. (k) Nov. (l) Dec.

Figure 3. Monthly mean cloud coverage in Asia obtained from NOAA AVHRR imagery.

grassland, wetland, cropland, urban, cropland and natural vegetation mosaic, snow and ice. Monthly mean and standard deviation are calculated within 16 types of classes. Figure 3 shows the land cover based statistics of monthly mean cloud coverage. The error bars shown in the figure represents the standard deviation centered by mean values.

The point is that over all categories the cloud coverage is higher in summer from June to August and lower in winter from November to February. This is because most of Asian region belongs to monsoon region and cloud contamination is highly affected by it. In snow and ice category, the cloud coverage from February to April exceeds 50%. Since evergreen broadleaf forest distributes mostly in tropical area, the over all cloud coverage on that is ranging from 25 to 45 % that are relatively higher than those of other categories. In bare soil, miss identification of clouds seems to occur from April to August because of the following reasons; there is less rainfall and clouds in summer at bare soil category which is mainly composed of desert area in west China, (2) apparent miss identification of clouds from fine pixels are confirmed by a visual interpretation. The reason of miss identification is caused by the following three factors; (1) most of bare soil area distributes far from the Tokyo and Bangkok ground station and data acquisition frequency is very low, (2) much noise is included in the imagery originated from the receiving failure, (3) sensor scan angle is higher and off nadir. The maximum cloud coverage of all categories are



(a) Evergreen needleleaf forest (b) Evergreen broadleaf forest (c) Deciduous needleleaf forest
(d) Deciduous broadleaf forest (e) Mixed forest (f) Closed shrublands
(g) Open shrublands (h) Woody savannas (i) Savannas
(j) Grasslands (k) Permanent wetlands (l) Croplands
(m) Urban and built-up (n) Cropland/Natural vegetation mosaic (o) Snow and ice

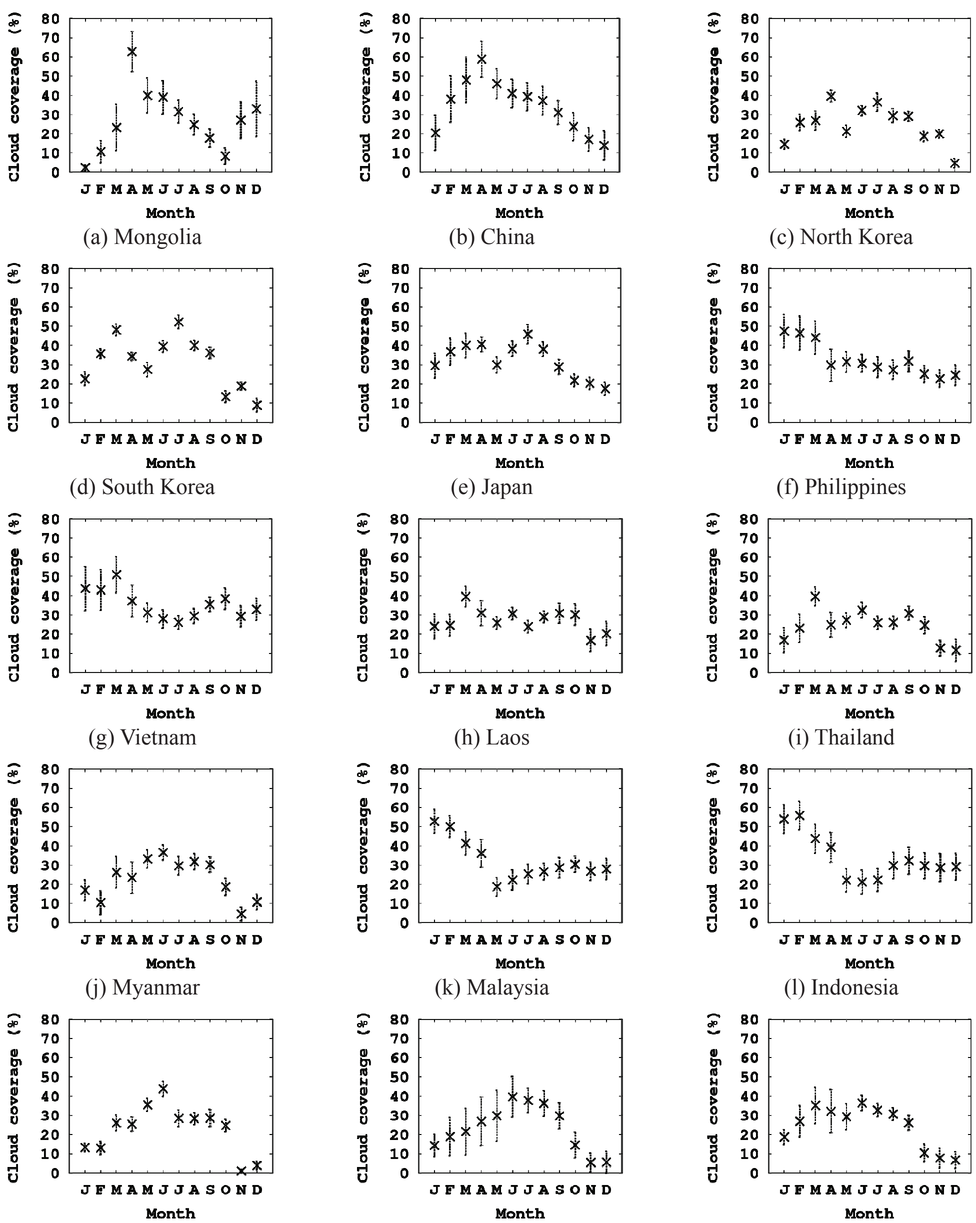
Figure 4. Land cover based statistics of monthly mean cloud coverage in Asia.

41, 61, 55, 54, 40, 41, 39, 34, 32, 27, 33 (%) from January to December, respectively. It statistically indicates that at least 20-day, 61(%) of 1-month, is required to obtain completely cloud free data.

3.2 Country cover based statistics of cloud coverage

Then, country cover based statistics are computed to evaluate the derived cloud coverage. The area of processing covers 15 countries including Mongolia, China, North Korea, South Korea, Japan, Philippines, Vietnam, Laos, Thailand, Malaysia, Indonesia, Myanmar, Bangladesh, India and Nepal. Monthly mean and standard deviation are calculated within 15 countries. Figure 4 shows the country-based statistics of monthly mean cloud coverage. The error bars shown in the figure represents the standard deviation centered by mean values.

In southern part of China, the cloud coverage from February to April is around 70 %, which is much higher values than other countries. This is due to the multiple factors such as high mountains, humidity and winds. In south- east countries such as Thailand, Myanmar and Bangladesh, cloud coverage is relatively low during dry season from November to February whereas it is comparatively high during rainy season from June to September. It is surprising that the maximum cloud coverage in these countries is at most 50 %



(a) Mongolia (b) China (c) North Korea

(d) South Korea (e) Japan (f) Philippines

(g) Vietnam (h) Laos (i) Thailand

(j) Myanmar (k) Malaysia (l) Indonesia

(m) Bangladesh (n) India (o) Nepal

Figure 5. Country based statistics of monthly mean cloud coverage in Asia.

and is lower than that of East Asian countries such as Japan or South Korea. The reason is that the rainfall timing in rainy season is given by a shower in the morning or early evening and it gets fine during the day-time when satellite passes. On the other hand, the rainfall annual cycle in southern hemisphere including Malaysia or Indonesia is quite the reverse with that of countries in northern hemisphere.

4. Concluding Remarks

In this study, an operational scheme for masking cloud-contaminated pixels in Advanced Very High Resolution Radiometer (AVHRR) daytime data was developed, evaluated and presented in order to compute cloud coverage statistics over Asian region. Land cover and country based statistics computation showed that cloud coverage in northern hemisphere over a range of categories is higher in summer from June to August and lower in winter from November to February affected by monsoon. The cloud coverage trend in southern hemisphere is in vice versa. It is noted that in bare soil miss identification of clouds seems to occur caused by the lack of coverage, noise included in the imagery originated from the receiving failure and higher sensor scan angle. It statistically indicates that at least 20-day, 61(%) of 1-month, is required to obtain completely cloud free data. In future work, this type of analysis will be applied to other applications, such as derived cloud and aerosol properties. The goal of these studies will be the same. They will attempt to quantify the impact of the additional spectral and spatial information offered by MODIS or VIIRS on the accuracy of the products and on the continuity with the AVHRR data record.

Acknowledgment

This research is partially supported by JSPS (Grant No. 19569002). The author would like to thank Dr. Vivarad Phonekeo for providing NOAA AVHRR data received at Asian Institute of Technology (AIT) in Bangkok, Thailand.

References

Jorgensen, P. V., 2000. Determination of cloud cover- age over Denmark using Landsat MSS/TMM and NOAA-AVHRR. Int. J. Remote Sens., 21(17), 3363-3368.

PaNDA committee Eds., 1998. Package for NOAA Data Analysis (PaNDA) manual.

Rossow, W. B, and Garder, L. C., 1993. Cloud Detection Using Satellite Measurements of Infrared and Visible Radiances for ISCCP. J. Climate, 6(12), 2341-2369.

Simpson, J. J., and Gobat, J. I., 1996. Improved cloud detection for daytime AVHRR scenes over land. Remote Sens. Environ., 55, 21-49.

Tag, P. M., Bankert, R. L., and Brody, R., 2000. An AVHRR multiple cloud-type classification package. J. Ap- plied Meteorology, 39, 125-134.

Vittorio, A. V. Di, and Emery, W. J., 2002. An automated, dynamic threshold cloud-masking algorithm for daytime AVHRR images over land. IEEE Trans. Geosci. Remote Sens., 40(8), 1682-1694.