Study on Noise Extraction and Correction of Time Series Satellite Profile Data Using Parallel Distributed System

Jonggeol Park¹, Masaki Hanada², Hidehiro Kanemitsu³ and Eiji Nunohiro⁴

¹,²,⁴Tokyo University of Information Sciences
4-1 Onaridai, Wakaba-ku, Chiba, 265-8501, Japan
³Global Education Center, Waseda University
1-104 Totukamachi, Shinjuku-ku, Tokyo 169-8050, Japan
E-mail: ¹amon@rsch.tuis.ac.jp, ²hanada@rsch.tuis.ac.jp, ³kanemih@gmail.com, ⁴nunohiro@rsch.tuis.ac.jp

Abstract  Satellite data has often been used to investigate damage from natural disasters (e.g., flood, drought, forest fire and volcano eruption) close to real-time. In particular, Terra/MODIS and Aqua/MODIS have been widely used for retrieving information on natural disasters and land cover changes because of their advantage of good ground resolution (maximum 250 m) and daily availability (full daily coverage of the Earth). However, the observation of areas covered by clouds using optical sensors does not work. Although the composite method is widely used to reduce the influence of clouds, clouds are likely to remain in some compositing periods. This paper presents the White Object Index (WOI), used for estimating the cloud fraction of a pixel in satellite imagery. The WOI was obtained from differences in reflectance of visible and short-wave infrared spectral bands using a mixture model. The efficiency of the WOI was tested by comparing the maximum value composite (MVC) method with MOD35 using time series data. As a result, WOI was confirmed to be an efficient tool for estimating the influence of clouds within a pixel.

Keywords: cloud, white object index (WOI), MODIS, gradient method, SIDAS

1. Introduction

The Moderate Resolution Imaging Spectroradiometer (MODIS), mounted on the Terra (EOS AM-1) and Aqua (EOS PM-1) satellites, operated by the National Aeronautics and Space Administration (NASA) as part of the Earth Observation System (EOS) program, uses three spatial resolutions (250 m, 500 m, and 1 km) and 36 bands distributed across the visible and thermal infrared spectral regions. They demonstrate remarkably improved performance in terms of spatial resolution, spectral resolution, brightness correction and satellite attitude information compared to the Advanced Very High Resolution Radiometer (AVHRR). A number of studies have been conducted to analyze land cover and land use changes over time using time series data from the continuous, high-frequency Earth observations [1-7].

The data obtained from an optical sensor is affected by various obstacles to land cover observation such as clouds and haze, and such effects should be removed. Because of low temperatures and the absence of chlorophyll, clouds have characteristics substantially different from those of plants that have a high normalized difference vegetation index (NDVI). Generally, this problem has been addressed by compositing images after selecting the most cloud-free day within a certain period based on NDVI and land surface temperature (LST) data [8,9].

Representative of such procedures is the composite method, in which the satellite data considered to be the least influenced by clouds during a particular period is selected and used as the representative value. This method presupposes that the NDVI and LST do not change during a particular period. In the northern hemisphere, considering that May and June are periods during which the LST and plant photosynthetic activities increase, the pixel dataset of a day nearest to the end of the month is selected as the representative value. By contrast, since the LST and NDVI values gradually fall during the period of autumn leaf coloring and fall, the dataset at the beginning of the month is selected as the representative value. In both cases, even though there is a pixel consisting of cloud-free data in the middle of the compositing period, the pixel of a different day is more likely to be selected. In addition, even if a pixel is affected...
by clouds, it cannot be distinguished by means of NDVI and LST under certain conditions.

MOD35 is the MODIS cloud mask product that provides a probabilistic calculation of the effects of clouds on a per-pixel basis [10]. However, there can be a discrepancy between the MOD35-based cloud mask image and the cloud distribution of the transmitted data. This phenomenon is observed frequently in low-NDVI forest areas, urban areas and deserts in autumn and winter. Cloud-free data may be masked, or inversely, cloud pixels may not be masked.

Against this background, we propose the White Object Index (WOI) to indicate the per-pixel cloud fraction based on a relatively simple computation. In this study, cloud-free images will be generated after detecting cloud-free pixels using WOI from MODIS data received in real time. Then the existing clear image database will be updated using the generated cloud-free images. Once a clear image database is updated, existing data can always be compared with the data received in real time whenever disaster monitoring is necessary. The accuracy of the WOI values will be verified by simulating the relationship between clouds and WOI values when clouds are inserted into a cloud-free pixel. Also, the reliability and applicability of the WOI will be evaluated by comparing MOD35 cloud mask images and the WOI image.

2. Data and System

2.1 MODIS MOD09A1 data

The MODIS Surface Reflectance products provide an estimate of the surface spectral reflectance as it would be measured at ground level in the absence of atmospheric scattering or absorption. MYD09A1 provides channels 1-7 at 500-meter resolution in an 8-day gridded level-3 product in the Sinusoidal projection. Sea Surface Temperature (SST) is the water temperature close to the ocean's surface. SST anomalies show La Niña, El Niño and Normal Conditions in the Pacific Ocean. El Niño is defined by prolonged differences in Pacific Ocean surface temperatures when compared with average values. The sign of an El Niño in the sea surface temperature pattern is when warm water spreads from the west Pacific and the Indian Ocean to the east Pacific. In addition, SST is used to check for the distribution of chlorophyll-a, flow of Kuroshio and Oyashio, for determining fishing grounds.

Land Surface Temperature (LST) is the land temperature estimated by satellite. LST changes depending on geographical features such as latitude, solar altitude and moisture content of the surface, and the type of land cover. Therefore, it is used for the investigation of heat island phenomenon in urban areas, desertification and land cover change.

| Table 1 Relationship between the feature of the wavelength and the resolution of the bands |
|---------------------------------------------|-----------------|------------------|------------------|
| **Band** | **Wavelength (nm)** | **Resolutions** | **Key Use** |
| 1   | 620–670 | 250m, 500m | Absolute Land Cover change, Vegetation Chlorophyll |
| 2   | 841–876 | 250m, 500m | Cloud Amount, Vegetation Land Cover change |
| 3   | 459–479 | 500m | Soil/ Vegetation Differences |
| 4   | 545–565 | 500m | Green Vegetation |
| 5   | 1230–1250 | 500m | Leaf/Canopy Differences |
| 6   | 1628–1652 | 500m | Snow/Cloud Differences |
| 7   | 2105–2155 | 500m | Cloud Properties, Land Properties |

2.2 Overview of satellite image data analysis system

Tokyo University of Information Sciences receives MODIS data which is one of the sensors equipped by NASA’s Terra and Aqua satellites, and researches the analysis on change of environment as part of the academic frontier project. MODIS is a multiple band radiometric sensor instrument aboard Terra (NASA launched on December 18, 1999, at local time 10:30AM) and Aqua (NASA launched on May 4, 2002, at local time 1:30PM) satellites for environmental monitoring of a large region of the earth covering a width of 2,330 km. It measures 36 spectral bands between 0.405 and 14.385 µm with three spatial resolutions (250m, 500m, 1,000m). For the information infrastructure of the frontier research project, we are currently developing a satellite image data analysis system (SIDAS) [11,12] to be implemented as a web system. SIDAS has a parallel distributed system configuration using multiple PC clusters, a database for MODIS data and applications to analyze the environmental changes. SIDAS is used to view the research results and MODIS data for public use. Figure 1 shows the configuration of SIDAS. We used the stream data processing model for data processing in SIDAS. In this model, data processing scenarios are registered in the system beforehand and the system receives and processes
huge satellite data in real-time according to the predefined process scenario. The processing flow of SIDAS1.1 is shown as follows in Fig. 1:

1. The Web server accepts the user’s request and submits it for processing to the application server.
2. The application server transfers the MODIS data used during processing from the MODIS database to the local disk in the application server.
3. The application server executes an application for the request and returns the results to Web server.

In this research, we use the following standard MODIS data products: MOD02 (MODIS calibrated radiances product spatial resolution), MOD03 (geological location product), MOD11 (land surface temperature), MOD13 (vegetation index), MOD28 (Sea Surface Temperature), MOD35 (cloud mask).

### 3. Methodology

#### 3.1 Ground truth data

Remote sensing is based on the measurement of reflected or emitted radiation from different bodies. Objects having different surface features reflect or absorb the sun’s radiation in different ways. The reflectance properties of an object depend on the particular element and its physical and chemical state, its surface roughness as well as its geometric circumstances. These differences make it possible to identify different earth surface features or materials by analyzing their spectral reflectance patterns or spectral signatures.

When using a mixture model to estimate the mixture ratio of a particular element, the characteristics of other elements should be known in addition to those of the element being evaluated. In this study, white objects (clouds and snow/ice) and nonwhite objects (e.g., vegetation, bare soil and deserts) present on the land surface were set as the elements included in a pixel. Although the area of interest in this study is land, sea was added as a land cover type to enable the estimation of the clouds over the sea. Specifically, the MOD09 surface spectral reflectance data received and processed by the Tokyo University of Information Sciences has seven land cover types: (1) cloud (summer, autumn and winter locations); (2) snow/ice (winter locations from Hokkaido in Japan, China and Russia and ice in the Sea of Okhotsk); (3) vegetation (summer, autumn, and winter locations including Aomori in Kanto Region and Fukuoka in Japan, Taebaek mountain range and Mount Baekdu in Korea, and Shenyang in China along with their respective surrounding areas); (4) urban (winter locations in Tokyo, Osaka, Nagoya, and Fukuoka in Japan and Seoul and Daegu in Korea); (5) desert (summer and winter locations in the arid areas around Hohhot in China); (6) bare soil (winter around Chiba in Japan, and Gimpo Plain in Korea); (7) Clean water (the Pacific Ocean). Graphs in Fig. 2 represent the spectral changes from the visible region to the short-wave infrared region (MODIS band 1 through band 7). Vegetation was classified into intense photosynthetic activity (summer), leaf coloring (autumn), and leafless canopy (winter) periods. Winter data was chosen for the urban and bare soil areas to exclude the influence of vegetation. The summer landcover in urban areas was set to be a combination of winter landcover and vegetation. The reflectance spectra of these land cover types were obtained by averaging the values of over 100 pixels each after extracting the areas from their respective MOD09A1 images.

![Fig. 1 Overview of SIDAS](image)

**Fig. 1** Overview of SIDAS

**Fig. 2** Reflectance spectra of the seven landcover types (Where B1 is Band1, B2 is Band2, B3 is Band3, B4 is Band4, B5 is Band5, B6 is Band6 and B7 is Band7)

Figure 2 shows the spectral signature of various natural features (Soil, Vegetation, Snow, Cloud and Sea Water). The spectral reflectance curve of healthy green vegetation has a significant minimum reflectance in the visible portion of the electromagnetic spectrum resulting from the pigments in plant leaves. Reflectance increases dramatically in the near infrared (b2) range. The spectral reflectance curve of bare soil is considerably less variable.

Journal of Signal Processing, Vol. 21, No. 6, November 2017
The reflectance curve is affected by moisture content, soil texture, surface roughness, presence of iron oxide and organic matter. These factors are less dominant than the absorbance features observed in vegetation reflectance spectra. The water curve is characterized by a high absorption at near infrared wavelengths range and beyond. Cloud and snow have a high solar reflectivity at visible wavelengths (b3, b4 and b1) compared to that of most surface features.

3.2 Mixture model

In the mixture model, the fraction of white objects within a pixel observed at wavelength $\lambda$ can be obtained using Eq. (1).

$$ R(\lambda) = a \times W(\lambda) + b \times V(\lambda) + c \times S(\lambda) $$

where $R$ stands for the reflectance of an object observed from a satellite and $W$, $V$ and $S$ represent the reflectance of white objects, vegetation, and urban/desert/bare soil areas, respectively. $a$, $b$ and $c$ denote their respective fractions (%) within the pixel. The white object fractions $a$ can be obtained by substituting the terms for $V$ and $S$ with $N$ in Eq. (1) to obtain Eq. (2).

$$ R(\lambda) = a \times W(\lambda) + (100-a) \times N(\lambda) $$

where $N$ stands for noise, which represents the composite reflectance of all nonwhite elements such as vegetation, urban area, desert, bare soil and sea water. The cloud fraction ($a$) in Eq. (2) was calculated using the gradient method. The gradient method can be applied only when the changed amount of the white objects between two points (wavelengths) $|W(\lambda_1) - W(\lambda_2)|$ of the observed spectrum ($\lambda_1 < \lambda < \lambda_2$) is significantly larger than that of the noise components $|N(\lambda_1) - N(\lambda_2)|$. The difference in reflectance between the two points ($\lambda_1$ and $\lambda_2$) can be calculated using Eq. (3).

$$ R(\lambda_1) - R(\lambda_2) = a \times W(\lambda_1) - W(\lambda_2) + (100-a) \times (N(\lambda_1) - N(\lambda_2)) $$

From this, the white object fraction ($a$) can be calculated using Eq. (4).

$$ a = \frac{R(\lambda_1) - R(\lambda_2)}{W(\lambda_1) - W(\lambda_2)} - (100-a) \times \frac{N(\lambda_1) - N(\lambda_2)}{W(\lambda_1) - W(\lambda_2)} $$

Eqs. (5) and (6) define the necessary conditions for applying the gradient method to Eq. (4). If the value of denominator $(W(\lambda_1) - W(\lambda_2))$ is much larger than the numerator $(N(\lambda_1) - N(\lambda_2))$ in Eq. (5), the right term of Eq. (4) will be close to 0.

$$ |N(\lambda_1) - N(\lambda_2)| << |W(\lambda_1) - W(\lambda_2)| $$

$$ |N(\lambda_1) - N(\lambda_2)| / |W(\lambda_1) - W(\lambda_2)| \approx 0 $$

If the value of Eq. (6) is close to 0, the second term in the right-hand side of Eq. (4) becomes 0, and Eq. (4) can be simplified to Eq. (7). If the value of Eq. (6) is much larger than 0, then the accuracy of the mixture model is lower. The accuracy of the mixture model increases as the value nears 0.

$$ a \approx \frac{R(\lambda_1) - R(\lambda_2)}{W(\lambda_1) - W(\lambda_2)} $$

In order to obtain an accurate white object fraction within a pixel of MODIS data using the gradient method, it is necessary to determine the spectra of $\lambda_1$ and $\lambda_2$, to meet the condition defined by Eq. (6). B3 and B7 in the visible and short-wave infrared regions, respectively, are considered suitable because they show the largest reflectance difference among the white object bands. However, simply substituting the reflectance values of B3 and B7 into Eq. (6) can yield a result that does not meet the necessary conditions for Eq. (6), for example 0.68 for the desert cover which has the largest gradient among the noise components as expressed below.

$$ \frac{|Desert(B_3) - Desert(B_7)|}{|Cloud(B_3) - Cloud(B_7)|} = 0.68 >> 0 $$

Thus, it is necessary to adjust the solution of Eq. (8) towards 0 by applying the coefficient $\alpha$ to the terms for $\lambda_2$ in Eq. (6) as expressed below.

$$ \frac{|N(\lambda_1) - \alpha \times N(\lambda_2)|}{|W(\lambda_1) - \alpha \times W(\lambda_2)|} $$

Using Eq. (9), the $\alpha$ value is obtained in the range 0 to 10 going up by 0.1 towards minimizing the least square error.

$$ D(\alpha) = \sum_{N=water} \left( \frac{|N(\lambda_1) - \alpha \times N(\lambda_2)|}{|W(\lambda_1) - \alpha \times W(\lambda_2)|} \right)^2 $$
The closer the value of \( D(\alpha) \) is to 0, the smaller the error in estimating the white object fraction within a pixel.

### 3.3 Investigation of \( \lambda_1 \) and \( \lambda_2 \)

In Eq. (5), the \( \lambda_1 \) term plays the role of maintaining the difference between the values of white objects and noise components to be as large as possible, and the \( \lambda_2 \) term and the coefficient \( \alpha \) determine the value of the noise components as close to 0 as possible without changing the value of the white objects (Eq. 9). Therefore, a visible spectrum band with strong white object characteristics is used for the \( \lambda_1 \) term, and a short-wave infrared spectrum band is used for the \( \lambda_2 \) term.

For example, visible bands \( B_3, B_4, \) and \( B_1 \), or their combinations (e.g., \( B_3 + B_1, B_4 + B_1, B_3 + B_4 + B_1 \)) can be considered for the \( \lambda_1 \) term, and \( B_6 \) or \( B_7 \) can be considered for the \( \lambda_2 \) term, which yields a total of 21 possible combinations. The \( \alpha \) value for each of these combinations is calculated using Eq. (9), and the wavelength combination leading to the smallest \( D(\alpha) \) value is determined accordingly. Table 2 shows the calculation results of some of these combinations.

| \( \lambda_1 \) | \( \lambda_2 \) | \( \alpha \) | \( D(\alpha) \) |
|---|---|---|---|
| \( B_3 + B_4 + B_1 \) | \( B_7 \) | 1.5 | 26 |
| \( B_3 + B_1 \) | \( B_7 \) | 0.5 | 27 |
| \( B_1 \) | \( B_7 \) | 0.6 | 37 |
| \( B_4 \) | \( B_7 \) | 0.6 | 37 |
| \( B_3 \) | \( B_7 \) | 0.3 | 57 |
| \( B_3 + B_4 + B_1 \) | \( B_6 \) | 1.3 | 75 |
| \( B_4 + B_1 \) | \( B_6 \) | 1.0 | 153 |

Substituting \( B_1 \) and \( B_7 \) into \( \lambda_1 \) and \( \lambda_2 \) results in the \( \alpha \) and \( D(\alpha) \) values of 0.6 and 37, respectively. The largest \( \alpha \) and smallest \( D(\alpha) \) among the 21 combinations, 1.5 and 26, respectively, were obtained by substituting \( B_3 + B_4 + B_1 \) and \( B_7 \), respectively, using Eq. (10).

\[
\alpha \approx \left| \frac{R(B_1) + R(B_3) + R(B_4) - 1.5 \times R(B_7)}{W(B_1) + W(B_3) + W(B_4) - 1.5 \times W(B_7)} \right| \quad (10)
\]

The absolute value \( |R(B_1) + R(B_3) + R(B_4) - 1.5 \times R(B_7)| \) is calculated for each of the five landcover types in the standard spectra shown in Fig. 2. While absolute values for cloud and snow/ice are as high as 210 and 257; those of vegetation, desert, bare soil and sea water are close to 0 at 2, 2, 2, and 4, respectively. As shown in Eq. (10), the gradient method uses computation with absolute values. However, given that the purpose of this study is to use the index from White Object Index (WOI) to determine the cloud fraction, it is not necessary to use absolute values. Moreover, the value of \( |W(B_1) + W(B_3) + W(B_4) - 1.5 \times W(B_7)| \), the denominator of the right-hand side of Eq. (10), is 210 in the standard spectrum (Fig. 2). From this value, the cloud fraction within a pixel (Eq. 10) can be rewritten to Eq. (11).

\[
WOI = \frac{(R(B_1) + R(B_3) + R(B_4) - 1.5 \times R(B_7))}{210} \quad (11)
\]

Since a standard spectrum does not consider the influence of clouds, the accuracy of the cloud fraction estimation increases as the WOI value gets closer to 0. While the estimation errors for desert, bare soil or vegetation covers are as low as 1%, errors in the urban cover are about 5%.

### 4. Results

#### 4.1 MODIS data analysis in SIDAS

SIDAS has developed a data stream function to execute the reception, analysis and web data release of satellite data in real time. This is processed using an application server and a Web server. NDVI, SST and LST are single band data and thus the image data is normally shown as a gray scale image.

![Fig. 3 Example of display for RGB image data](image)

MODIS data is analyzed beforehand and stored into the database. For example, image data was created using MODIS data from beginning of March 2010 to end of February 2011, and the image data was stored in the database. The color image data from the Tohoku region,
Japan, was created using radiance data (B1: Red, b4: Green and b3: Blue) of MOD02 (Fig. 3), NDVI, LST and SST.

4.2 Cloud free image

Time series cloud-free images were generated using WOI and visible data thresholding. Figure 4 shows the results of cloud-free image generation using 3 images taken on 6 days between May 1 and 10, 2003 (May 1, 4, 7, 8, 9, and 10) in Oita prefecture in Japan. Images (a), (c) and (e) are true-color images from May 1, 5, and 10, respectively.

During these 10 days, the NDVI increased from 0.75 to 0.8, and the peak LST was 25 °C, 7 °C and 6 °C higher than May 1 and 10, respectively. WOI values showed small differences of 2.8, 2.9 and 2.5. As a result of WOI compositing ((d) and (f)), the cloud-free pixels were updated without changing the cloud cover in the true-color images. The time series cloud-free image generated on May 5 (Fig. 4(d)) shows an unnatural boundary line (II) due to the change in surface reflectance that occurred during the short period of 5 days.

The difference in reflectance between the two images (May 1 and 5) could be broken down into 3% in the visible region and 17% in the near infrared region. This reflectance change is assumed to be due to the difference in the Sun altitude because of the observation time difference of the two images, and multiple scatterings of clouds and aerosols.

5. Discussion

Optical sensors have been widely used for Earth observations during the past 45 years since Landsat was launched in 1972. However, clouds blocking the view of the Earth surface have always posed problems in using optical sensors for disaster or vegetation monitoring. Therefore, the Earth’s surface is often observed from satellites between 10:00 and 11:00 a.m., when clouds appear less frequently. However, images captured via high-frequency low-resolution observation sensors, widely used for large-scale environmental monitoring (NOAA/AVHRR, Terra/MODIS and Suomi NPP/VIIRS), cannot be free of clouds because of the sheer width of the observed area. In order to mitigate the cloud effects, many researchers have used composite methods (e.g., MVC, MaxT, MinB) for single image generation from multiple images taken from a certain period of time. Currently available composite techniques resort to the method of selecting the day with the least probability of clouds in each compositing period, which does not give information on the actual presence or absence of clouds. Cloud fraction can be estimated with the WOI proposed in this study, and cloud-free images can be generated based on the estimated cloud coverage, greatly facilitating disaster or vegetation monitoring.

For example, paddy monitoring can be made easier. Rice farming is one of the most important economic activities in Asia. It is necessary to determine the optimal periods for planting and harvesting as well as estimating crop area and expected yield. However, cropland area estimation is not easy using only NDVI or LST used in existing research approaches, because neither can distinguish paddies well from fields. It is necessary to detect water during the rice sprout planting period, but existing composite methods cannot easily detect water because they have difficulty distinguishing water from clouds due to low NDVI and...
LST values. This study proposes a simple method of water and cloud distinction by using visible and short-wave infrared regions. WOI can also be efficiently applied to disaster monitoring (e.g., floods, forest fires) because damaged areas can be easily identified by comparing pre- and post-disaster images. Using the WOI threshold value proposed in this study, cloud-free pixels are constantly updated to enable the generation of the latest cloud-free image, thus making it possible to automatically identify the areas damaged by natural disasters, such as yellow dust, flood and forest fire, by comparing the images received in real time from satellites with those generated before the disaster.

Time series NDVI is often used when investigating land cover or change using vegetation phenology. Methods based on line interpolation [13,14], fast Fourier transformation (FFT) [15,16], and wavelet [17-19] have been generally used to correct pixels containing clouds included in a time series NDVI profile. However, these correction methods have limitations in minimizing cloud effects because the input data for profile correction are also used by data affected by clouds. For example, FFT correction only uses low-frequency regions by removing high-frequency regions after decomposing the previous input data into frequency components. However, since it is a Fourier series expansion using the previous data including cloud pixels, cloud effects are still present in low frequency regions. The residual cloud effects in low frequency regions can be mitigated by removing the pixels containing clouds by using the WOI and then expanding the Fourier series.

6. Conclusion

This study proposes a mixture model to calculate the white object fraction within a pixel, to characterize features such as clouds. Seven land cover types were selected from MODIS data as end-members necessary for the mixture model. They were classified into white objects (clouds and snow) and nonwhite objects (vegetation, urban, desert, bare soil, and sea water). White objects are characterized by high reflectance in the visible region and low reflectance in the short-wave infrared region. Inversely, nonwhite objects are characterized by low reflectance in the visible region and higher reflectance in the short-wave infrared region than in the visible region. The mixture model was solved using the gradient method. To enhance the efficiency of the gradient method, it is necessary to make the upper-right term (nonwhite object value) in addition to the targeted left-hand term (white object value) of the formula to be as close to zero as possible. Therefore, we substituted the visible bands (B3 + B4 + B1) into the left-hand term so that the white objects could have the largest value, and we substituted the short-wave band (B7) in the right-hand term with a smaller value than the visible bands. The coefficient of the right-hand term was calculated by the least squares error method, and its value was set at 1.5.

According to the findings of this study, it is recommended to keep the error threshold under 2% for forest areas and 5% for urban areas, when generating cloud-free images using WOI values. The difference of 4% in error threshold between forest and urban areas was found to be consistent with the simulation results. We investigated the problems involved in the time series cloud-free image generation using NDVI and LST, which are most frequently used for conventional cloud removal methods. In methods based on NDVI and LST, values lower than the maximum value in the compositing period are judged to be owing to cloud effects. As a result, no accurate cloud-free images can be generated when the NDVI value is low or when LST varies greatly.

For the purpose of this study, a linear model of land surface and cloud spectra was considered to calculate the cloud fraction of satellite imagery. However, clouds are greatly influenced by spatial heterogeneity and values observed from satellites can vary, depending on the spatial distribution pattern in the same surface area. Therefore, there is a need to conduct an in-depth study of the cases where the hypothesis of the linear mixed-order model is supported and not supported by calculating the 3D radiative transfer.

Acknowledgment

This research was supported in part by the Environment Research and Technology Development Fund (1-1405) of Ministry of Environment, Japan; and by a MEXT Japan grant-in-aid for scientific research (No. 26350403) and JSPS Grant-in-Aid for Scientific Research (No. 25340014).

References

[1] J. J. Simpson, A.B. Schmidt and A. Harris: Improved cloud detection in along track scanning radiometer (ATSR) data over the ocean, Remote Sensing of Environment, Vol. 65, No. 1, pp.1-24, 1998.
[2] X. Zhang, M. A. Friedl, C. B. Schaaf, A. H. Strahler, J. C. F. Hodges, F. Gao, B. C. Reed and A. Huete: Monitoring vegetation phenology using MODIS, Remote Sensing of Environment, Vol. 84, No. 3, pp.471-475, 2003.
[3] J. Suming and S. A. Sader: MODIS time-series imagery for forest disturbance detection and quanti-
ification of patch size effects, Remote Sensing of Environment, Vol. 99, No. 4, pp. 462-470, 2005.

[4] S. L. Ross, F. K. Joseph, E. Jayantha, G. L. John and L. D. Worthy: Land-cover change detection using multi-temporal MODIS NDVI data, Remote Sensing of Environment, Vol.105, No. 2, pp.142-154, http://dx.doi.org/10.1016/j.rse.2006.06.018, 2006.

[5] A. Langner, J. Miettinen and F. Siegert: Land cover change 2002-2005 in Borneo and the role of fire derived from MODIS imagery, Global Change Biology, Vol. 13, No. 11, pp.2329-2340, 2007.

[6] G. Sangram, A. F. Mark, T. Bin, Z. Xiaoyang and V. Manish: Land surface phenology from MODIS: Characterization of the collection 5 global land cover dynamics product, Remote Sensing of Environment, Vol. 114, No. 8, pp. 1805-1816, http://dx.doi.org/10.1016/j.rse.2010.04.005, 2010.

[7] D. Frantz, A. Röder, T. Udethoven and M. Schmidt: Forest disturbance mapping using dense synthetic Landsat/MODIS time-series and permutation-based disturbance index detection, Remote Sensing, Vol. 8, No. 4, pp. 277-298, doi:10.3390/rs8040277, 2016.

[8] B. N. Holben: Characteristics of maximum-value composite images from temporal AVHRR data, International Journal of Remote Sensing, Vol. 7, No. 11, pp. 1417-1434, 1986.

[9] R. Taddei: Maximum value interpolated (MVI): A maximum value composite method improvement in vegetation index profiles analysis, International Journal of Remote Sensing, Vol. 18, No. 11, pp. 2365-2370, http://dx.doi.org/10.1080/014311697217657, 1997.

[10] S. Ackerman, K. Strabala, P. Menzel, R. Frey, C. Moeller, L. Gumley, B. Baum, C. Schaaf and G. Riggs: Discriminating clear-sky from cloud with MODIS algorithm, Theoretical basis document (MOD35) 1997, http://modisatmos.gsfc.nasa.gov/docs/atbd_mod06_old.pdf (assessed on 14 Fab. 2017).

[11] E. Nunohiro, J. G. Park and K. J. Mackin: Development of satellite data analysis system, Proceedings of the 5th International Conference on Information, pp.143-146, 2009.

[12] A. Nakamura, J. G. Park, K. J. Mackin and E. Nunohiro: Development and evaluation of satellite image data analysis infrastructure, The Sixteenth International Symposium on Artificial Life and Robotics, Proceeding Index OS18-3, 2011.

[13] R. Yao, L. Wang, X. Huang, X. Guo, Z. Niu and H. Liu, Investigation of urbanization effects on land surface phenology in northeast China during 2001-2015, Remote Sensing, Vol. 9, No. 1, pp.66-81, 2017.

[14] Z. Luo and S. Yu: Spatiotemporal variability of land surface phenology in China from 2001–2014, Remote Sensing, Vol. 9, No. 1, pp.65-83, 2017.

[15] W. Verhoef, M. Menenti and S. Azzali: Cover a colour composite of NOAA-AVHRR-NDVI based on time series analysis (1981-1992), International Journal of Remote Sensing, Vol. 17, No. 2, pp. 231-235, 1996.

[16] G. J. Roerink, M. Menenti and W. Verhoef: Reconstructing cloud free NDVI composites using Fourier analysis of time series, International Journal of Remote Sensing, Vol. 21, No. 9, pp. 1911-1917, 2000.

[17] X. Lu, R. Liu, J. Liu and S. Liang: Removal of noise by wavelet method to generate high quality temporal data of terrestrial MODIS products, Photogrammetric Engineering & Remote Sensing, Vol. 73, No. 10, pp. 1129-1139, 2007.

[18] B. Martinez and M. A. Gilabert: Vegetation dynamics from NDVI time series analysis using the wavelet transform, Remote Sensing of Environment, Vol. 113, No. 9, pp.1823-1842, 2009.

[19] R. Quiroz, C. Yarlequé, A. Posadas, V. Mares and W. W. Immerzeel: Improving daily rainfall estimation from NDVI using a wavelet transform, Environmental Modelling & Software, Vol. 26, No. 2, pp. 201-209, 2011.

**Jonggeol Park** received his Ph.D. degree from Chiba University, Chiba, Japan, in 2000. He is an Associate Professor in the Department of Information Systems, Tokyo University of Information Sciences. His research interests include algorithm development of machine learning and Image Processing.

**Masaki Hanada** received a B.E. degree in resources engineering from Waseda University in 1996, and M.S. and D.S. degrees in global information and telecommunication studies from Waseda University in 2003 and 2007, respectively. He is an Associate Professor in the Department of Information Systems, Tokyo University of Information Sciences. His research interests include traffic control and resource management in communication networks.
**Hidehiro Kanemitsu** received his B.S. degree in science from Waseda University, Japan, and his M.S. and D.S. degrees in global information and tele-communication studies from Waseda University, Japan. His research interests include parallel and distributed computing, grid, peer-to-peer computing, and web service technology. He is currently an Assistant Professor at the Global Education Center, Waseda University, Japan. He is a member of IEICE and IEEE.

**Eiji Nunohiro** is a Professor in the Department of Informatics, Tokyo University of Information Sciences. His main works are research for program-training training support system, data analysis system in information security and compiler and parallel algorithm in high performance computing. He is a member of Information Processing Society of Japan and Japanese Society for Information System in Education.

(Received May 15, 2017; revised August 18, 2017)