ACCURACY IMPROVEMENT OF LAND COVER TYPE CLASSIFICATION USING NIGHTTIME AVHRR DATA

Chada Narongrit*, Mitsuharu Tokunaga, Shunji Murai, Kaew Nualchawee, Apisit Eiumnoh, Suphat Vongvisessomjai

Abstract: The new method developed in this study was the use of nighttime AVHRR thermal data as an additional channel in classifying land cover types with respect to improve classification accuracy. The NDVI, ordinary brightness temperatures and land surface temperature, were examined using two-classification algorithms namely Maximum Likelihood and Decision Tree. Various band inputs were comparatively assessed for the classification accuracy. The results of classification in both cool and hot seasons showed that overall accuracy obtained by using combination of day and nighttime data was better than that obtained by using only daytime data. Using a combination of three bands including NDVI, daytime LST, and nighttime LST gave the best accuracy. Using three bands and only daytime data yielded a lower accuracy compared to when using only two bands, but with day and nighttime data. Although an overall accuracy of using only daytime and using both day and nighttime data was not remarkably different, the accuracies when adding a nighttime data were notably increased such as in a case of forest and built up classes. The nighttime data can well classify (1) forest from densely grown crops, (2) deciduous forest in hot season from sparsely growing crops and harvested agricultures, and (3) built up from harvested land. The results also indicated that the proposed approach using nighttime data was effective as a new classification method in vegetation environment associated with landscape. In addition, this study suggested an integrated approach involving day and nighttime data to monitor urbanization and heat island.

INTRODUCTION

Recently, a numbers of practical satellites used to monitor land cover information around the globe have been increased. The use of high-resolution satellite data derived from Landsat and SPOT have suffered from the infrequent coverage, high data volume, and high costs. Thus, recent efforts have been directed toward the use of low-resolution data obtained by AVHRR sensor on board of the NOAA satellite. The AVHRR provides coarser spatial resolution data (1.1 km at nadir) but its advantage is a much better temporal resolution with daily coverage (Cihlar, et al., 1996). For such coarser AVHRR data, a main problem is pixels containing a mixture of land cover types. The supplementation of high-resolution image is generally used to overcome this problem. However, using such approach is quite expensive. Thus, the methodological challenge is to extract meaningful land cover information by using only AVHRR data. The AVHRR sensor has five spectral channels, which can provide useful data for studying land cover features. Channels 1 and 2 are effective in land cover identification by adding spectral albedo information. Channels 3, 4, and 5 are effective in temperature information, which depends on land covering (Cihlar, et al., 1996). Initially, NDVI, computed as the per-pixel difference between reflectances of channels 2 and 1 divided by the sum of these two channels, was used for classifying land cover types because NDVI values provide information on the seasonal vegetated dynamics of the surface covers. Subsequently, brightness temperatures of channels 3, 4, and 5 were used. More recent
researches use a combination or ratio of NDVI and thermal brightness temperatures. Most of the work uses only thermal data in daytime; while the AVHRR sensor provides daily useful data both in day and night times. However, the accuracy and number of land cover classes achieved from using only daytime data were limited.

The land cover types and land characteristics can be computed from the temperature differences over time. This thermal behavior of land surfaces can be described using the thermal inertia property, a physical variable describing the impedance to variations of temperature. For a given heat transfer, high thermal inertia values lead to small changes in temperature (Xue and Cracknell, 1995). Water is more resistant to the temperature change over time compared to other surface types. As the fractional vegetation cover increases, surface temperatures in daytime and temperature difference between daytime and nighttime decrease (Lambin and Ehrlich, 1997). The water surface shows the least temperature difference, and the least vegetated covered land with a dry surface shows the highest temperature difference. This is because the heavily vegetated surfaces are associated with a large latent heat flux. Hence, these surfaces are cooler than those with less vegetation.

This study did not intend to produce a definitive land cover map, but rather to compare the information contents of NDVI and thermal bands. Therefore, the research attempted to determine a usefulness of nighttime and daytime data combination, which yielded the greatest potential for land cover type classification. The new classification method was approached using nighttime AVHRR data as an additional band in order to improve classification accuracy of land cover types derived from the AVHRR coarse resolution data.

Figure 1 (Left) The Study Area in a Color Composite Image, R : G : B = NDVI : Day LST : Night LST (Top Right) A Profile Plot of NDVI and Different Temperatures (Bottom Right) of Day and Night Temperatures Crossing from West to East (Smoothed by 3x3 Window)
STUDY AREA

The study area of this research is Thailand, a country situated in the South East Asia between 5 to 20 degree North latitude and 97 to 105 East longitudes. There are three seasons namely wet (June-October), cool (November-February), and hot (March-May). The topography is plain in the central part, hilly and mountainous in the northern part, and plateau in the northeastern part. In Figure 1 (left), red color represents forest and active agricultural lands; green color represents inactive agricultural lands; and blue color represents water bodies. Figure 1 (right) shows profiles of NDVI and LST. Figure 2 shows the existing map of land uses of Thailand in 1995 (OAE, personal contact).

DEFINITIONS OF CLASS

In this study, the land cover types were broadly classified into seven classes, which were defined as follows:

1. Water bodies (Wt): This class, such as dam and reservoir, were defined as the lands that are covered by water bodies.

2. Built up areas (Ba): The urban lands with the least vegetation covered that are used for residential and industrial activities.

3. Field crops (Cr): The vegetated areas that are used to grow annual crops such as cassava, sugar cane, maize, and mixed field crops.

4. Paddy fields (Pa, Ph, Pu): The vegetated areas that are used for cultivating rice crops. This class was divided into three subclasses: lowland active paddy (Pa), lowland harvested paddy (Ph), and upland paddy (Pu).

5. Orchards (Or): The vegetated lands are used for growing fruit trees.

6. Para rubber (Rb): The lands are used for growing Para rubber tree.

7. Forests (Fe, Fd): This class was divided into two subclasses including (1) deciduous forest (Fd), the land in which most area is covered by deciduous forest trees; however, in some places, this kind of lands are mixed with the sparse evergreen trees, dry dipterocarp trees, and plantation forest trees, and (2) evergreen forest (Fe), the land with vegetated areas permanently covered by high density of evergreen forest trees.

REMOTE SENSING DATA PROCESSING

The AVHRR images derived from the NOAA-14 passed approximately at 2 AM and 2 PM were selected from the least cloud dates in December 1997 (cool season) and March 1998 (hot season). The procedures of AVHRR data processing consisted of a series of the following steps:

Image Pre-processing: The images were pre-
processed in order to convert the raw digital data into albedo reflectance (percent) of the AVHRR channels 1 and 2, and into thermal brightness temperature (°K) of the AVHRR channels 3, 4 and 5. Then, the images were corrected for geo-reference.

**NDVI Computation:** The AVHRR first two channels were transformed into the NDVI using the following equation:

\[
\text{NDVI} = \frac{(\rho_2 - \rho_1)}{(\rho_2 + \rho_1)}
\]  

(1)

Where \( \rho_1 \) and \( \rho_2 \) are visible channel 1 and near infrared channel 2 reflectances, respectively.

**Composite Image Generation:** The composite images of AVHRR channels 3, 4, 5, and NDVI were generated using Maximum values to reduce cloud pixels in the operational images. In this study, the AVHRR data were selected from the least cloud cover. Thus, the 3 day-continuous composite images were acceptable. Only the NDVI images were derived from daytime data; while others were derived from both daytime and nighttime data.

**LST Image Generation:** To correct the atmospheric effect on thermal channels 4 and 5, Land Surface Temperature (LST) images (°C) were generated by using the split window technique. In this study, three different LST models, proposed by Price (1984), Shambare and Mlindelwa (1998) and Kerr et al. (1992), were tested to select the best model for estimating LST in Thailand. The analysis found that LST data computed from each algorithm showed a high correlation (R=0.7) with actual ground temperatures measured at the provincial meteorological stations. However, the original models gave overestimation of LST. Therefore, the models were adjusted before they were used to generate LST images. The adjusted model used in this study was expressed in the following equation.

\[
\text{LST} = T_4 + 1.11(T_4 - T_5) - 273
\]  

(2)

**Cloud Masking:** Although the cloud effect was firstly screened in the step of three-day Maximum composites, the composites were still contaminated by some clouds. Hence, the remaining cloud pixels were masked by simple cloud detection algorithms including (1) LST pixels with temperature lower than 13°C and (2) pixels with temperature difference between day and night was less than 0°C and greater than 20°C.

**Scale Data:** All corrected composite images were re-scaled from 0 to 100 by using the following equation.

\[
\frac{(\text{Actual value} - \text{Minimum})}{(\text{Maximum} - \text{Minimum})} \times 100
\]  

(3)

**PROPOSED INDICES**

In Figure 1 (right), the NDVI showed negative relation with LST. From the fact that it is sometime difficult to use only NDVI to specify land cover types, which have a similar range of NDVI, therefore, a ratio of NDVI and LST can be an used as an alternative and effective choice to specify such land cover types, which also possess a different range of LST. In this study, three proposed indices using arctangent of NDVI/LST ratio (eq. 4, 5 and 6) were computed to transform the two-dimensional biophysical data into a quantitative, continuous, and linear variable (Lambin and Ehrlich, 1997). The index values ranged from 0 to 1. The minimum value was non-vegetated area; whereas the maximum value was dense forest.

\[
\text{Arctangent} (\text{NDVI/Day LST}) \quad (4)
\]

\[
\text{Arctangent} (\text{NDVI/Night LST}) \quad (5)
\]

\[
\text{Arctangent} (\text{NDVI/Dif LST}) \quad (6)
\]

**CLASSIFICATION METHOD**

The two-classification method namely Maximum Likelihood and Decision Rule were examined.

**Maximum Likelihood Classification**

The transected training sites represented region of interests (ROIs) was randomly sampled based on the overlaid digitized land use map. This map was supplemented by recent land cover types interpreted from NDVI and temperature profiles taken from the images. At least two training sites and more than 150 pixels per sites were labeled as ten classes.
These classes were defined as follows: Water bodies (Wt), Built up (Ba), Field crops (Cr), Upland paddy (Pu), Lowland harvested paddy (Hv), Lowland active paddy (Ac), Orchards (Or), Para rubber (Rb), Deciduous forest (Fd), and Evergreen forest (Fe).

**Decision Tree Classification**

The proposed indices of each land cover types in Figure 3 were used for the Decision Tree algorithms in Figure 4.

The resulting classified images were overlaid with the land use map in 1995 of Thailand (Figure 2) classified from Landsat-TM data. This was done in order to compute their overall accuracy.

**ACCURACY ASSESSMENTS**

Two accuracies were assessed: (1) overall accuracy calculated by summing the number of pixels that are correctly classified and then divided by the total number of pixels, and (2) individual accuracy calculated by dividing the pixel counts in each ground truth column by the total number of pixels in a given ground truth class.

![Figure 3 Proposed Indices for Each Land Cover Type in Cool Season (Left) and Hot Season (Right)](image-url)
RESULTS AND DISCUSSIONS

Land Cover Types Obtained from the Maximum Likelihood Method

By using a Maximum Likelihood Method, Figure 5 and Figure 6 show the classified images (the first three from left) in cool season and hot season, respectively. In cool season images, the overall accuracy assessed by the existing land use maps in 1995 was 80.64% for only daytime LST (LSTd) and 83.78% for a combination of LSTd and nighttime LST (LSTn). In hot season images, the overall accuracy was 75.35% for only LSTd and 79.40% for a combination of both LSTs. It can be seen that using a combination of LSTd and LSTn yielded a respective 3% and 4% increase of overall accuracy in cool and hot season images when compared to using only LSTd. By using only daytime data of NDVI and LSTd, the built up areas (Ba) and harvested paddy (Ph) can not be separated. The Ph area in the northeast region was incorrectly classified as Ba, especially in hot season. In addition, an evergreen forest (Fe) in hot season was misclassified as orchard (Or) and active paddy (Pp). These misclassifications could be corrected by using both day and nighttime data. Because field crop (Cr) in hot season was in early planting or in harvesting period, thus it

Figure 4  Decision Tree Algorithms in Cool Season (above) and Hot Season (below)

Figure 5  Classified Images in Cool Season : The first three were obtained from Maximum Likelihood and the last was obtained from Decision Algorithm
was incorrectly classified as harvested paddy (Ph) and deciduous forest (Fd). The analysis between LST and two environmental factors namely NDVI and elevation showed that LSTd was stronger correlated with NDVI than it was with elevation; while LSTn was fairly correlated with both NDVI and elevation (Chada et al. 2000). In the case of Fd, its LSTd and NDVI were similar to those of Cr, Ba, and Ph. However, its LSTn was remarkably lower than that of Cr, Ba, and Ph. This is because Fd situated on high elevation. Including channel 3 in the classifications gave a good accuracy. Figure 7 shows accuracy comparison of input bands in hot and cool seasons obtained by Maximum Likelihood method. The accuracy of three bands of daytime (NDVI, band 3, and LST) was 81.61% and 77.87% in cool and hot seasons, respectively. These numbers were lower than those obtained from the two bands of day and nighttime combinations such as NDVI and nighttime band 3 (82.43% in cool season and 78.96% in hot season), NDVI and nighttime LST (83.62% in cool season and 79.85% in hot season). This indicated that the accuracies were improved by including of nighttime data, but not by adding band numbers. Figure 8 shows a comparison of accuracy of the seven main land cover types over-
laid on the existing land use map in 1995. In both cool and hot seasons, the accuracies of built up and forest were significantly increased when using a combination of NDVI and LSTn. The best accuracy was obtained in paddy class. From the results, it can be explained that (1) temperature in forest is not fluctuated as in agricultural lands. The temperature of the agricultural lands does not only relate to amount of NDVI but also relate to soil water regime and land conditions, (2) the forest situates in high lands, where temperature in nighttime decreases even though deciduous forest was falling in hot season, (3) temperature between day and night times of built up areas was considerably different, and (4) the typical agricultural fields in Thailand, except paddy fields, are smaller than that of each AVHRR pixel. Thus, most pixels were contained with a mixture of land cover classes. When consider paddy class with a less unmixed-AVHRR pixel, the accuracy was better than that of other agricultural land uses with a higher proportion of mixed pixel such as orchards, para rubber, and field crops.

Land Cover Types Obtained from the Decision Tree Algorithms

(1) Cool Season Images

First, the land cover types were classified into six groups by using the index of atan(NDVI/LSTd). Then, the members in each group were sub-classified by LSTn. The classification algorithm in Figure 4 (upper) yielded the classified images as shown in the last image of Figure 5. In cool season, filed crops were cultivated, so its NDVI was close to active paddy but its LSTn was different. Among the vegetation areas of field crops, active paddy, orchard, para rubber and forest, empirical classifi-
cation proved that LSTn could better separate these classes than LSTd because these land cover types have overlapping LSTd values. The overall accuracy of the seven land cover types, comparatively computed from the land use map in 1995, was 84.53%.

(2) Hot Season Images

Although the Pa, Rb and Fe showed different atan(NDVI/LSTd) values, the empirical classifications showed that there was an overlapping of such values of these classes. Thus, all classes were first classified by index of atanNDVI/LSTd into three groups including (1) Wt, (2) Ba, Cr, Ph, Pu and Fd, and (3) Pa, Or, and Fe. These three groups were then classified by LSTn. The classification algorithm in Figure 4 (lower) yielded the classified images as shown in the last image of Figure 6. The overall accuracy of seven land cover types was 80.30%.

CONCLUSIONS

The new classification method using a combination between day and nighttime data was effective to classify land cover types by improving classification accuracy. The indices developed from the ratio between NDVI and temperature can be effectively used to separate the land cover types, which were difficult to separate when using only NDVI. The best classification accuracy was obtained from three-band combination of NDVI, LSTd and LSTn when using Maximum Likelihood and Decision Tree classification methods. The incorporation of nighttime data yielded consistent results both in cool and hot seasons.

Since spatial LST in nighttime strongly correlated with elevation, the use of LSTn can reduce confusion when classifying some land cover types, which have similar NDVI but situated in different elevation landscapes. The results showed that the nighttime data can improve the accuracies of paddy in lowland, crop in upland, and forest in highland.

It was found that the classification accuracies in hot season were lower than those found in cool season, for all cases of band inputs. This is because the characteristics of NDVI and temperature of many classes such as built up, harvested paddy, deciduous forest and field crops were similar in hot season. This could cause confusion in classifying these classes, resulting in lower accuracies compared to those of cool season.

It is noted that spatial variation of land surface temperature was not only related to land cover types but also related to other variables such as in-situ land use activities, soil moisture, local weather, elevation, and location. Thus, these variables should be recognized when using thermal data, particularly on studying land cover types at regional scale.

ACKNOWLEDGMENT

The authors wish to express appreciation to the director of ACRoRS, Dr. Kiyoshi HONDA, for contributing the AVHRR images and to the senior researcher of ACRoRS, Dr. Surat Lertlum, for a helpful guidance in NOAA-AVHRR processing. In addition, the authors would like to thank the officers of the Meteorological Department, especially Mr. Somchai Baimoung, for providing the ground temperature data. Finally the authors sincere thank to Dr. Supan Karnchanasutham for providing the updated land use map of Thailand.

(受付日2000.7.11，受理日2002.5.8)

REFERENCES


