

ANALYSIS OF THE LAND COVER AND ITS CHANGE OVER YELLOW RIVER BASIN USING SATELLITE DATA

M. Matsuoka ^a, T. Hayasaka ^a, Y. Fukushima ^a, Y. Honda ^b

^a Research Institute for Humanity and Nature, 335 Takashima-cho, Kyoto, Japan

- (matsuoka, hayasaka, yoshi)@chikyu.ac.jp

^b Center for Environmental Remote Sensing, Chiba University, 1-33 Yayoi-cho, Chiba, Japan

- yhonda@cr.chiba-u.ac.jp

KEY WORDS: Land cover, Classification, Yellow River basin, Terra/MODIS

ABSTRACT:

Land cover classification over Yellow River basin, China, by means of simple decision tree classification method using multi-temporal metrics derived from MODIS data is shown in this paper. Since the land cover is one of the input parameters for the hydrological model, the critical feature of classification method is to control the classification result using threshold values. The tentative result shows good correlations with existing digital land use map, though the small overestimation or underestimation are recognized in several categories and the improvement is necessary in discriminating the single cropped agricultural field from natural grassland. The adjustments of the threshold values established in the decision tree would be done by the aid of higher resolution satellite data.

1. INTRODUCTION

1.1 Background

The development of water resource management model in the Yellow River basin, China, is important because the stream of the river had not reached to the Bohai Sea for many days in a year since 1970s as a result of climate change and human activities. The land cover and its change reflect the interactions between human activities, such as agriculture and urbanization, and nature as dynamics of water and energy, therefore, it is one of the essential parameters for the development of the model. The purpose of this study is to analyze the land cover and its change for 20 years since 1980s over Yellow River basin using satellite data.

1.2 Outline of this study

Three kinds of sensor, Terra/MODIS, NOAA/AVHRR, and Landsat/TM or ETM+, are utilized in this study according to its spatial and temporal resolutions. The outline of this study, comprised of four frames, is shown in Figure 1. MODIS onboard Terra, which is new sensor operated since 1999, has two 250 meters bands in visible and near infrared. MODIS are viewing the entire Earth's surface every 1 to 2 days (Conboy). The land cover classification map is generated from the time series MODIS data in 2002 by means of the supervised decision tree method. AVHRR is operated over twenty years from the end of 1970s (Kidwell, 1998), and precious and long term seamless data are archived. The spectral and temporal characteristics of AVHRR data in 2002 are compared by each land cover categories of classification map, and thereafter the changes of the land cover since 1982 are detected using time series of AVHRR data. TM and ETM+ onboard Landsat series have much higher spatial resolution than AVHRR and MODIS, though the observation frequency

is low as once per 16 days (Sheffner, 1999) and the scheduled acquisition of the data is restricted by the cloud cover. TM series are mainly utilized for the validation and tuning of the land cover classification and change detection, and the detailed analysis intended for the drastically changed area. Land cover classification using MODIS data (number 1 in the figure) is the subject of this paper.

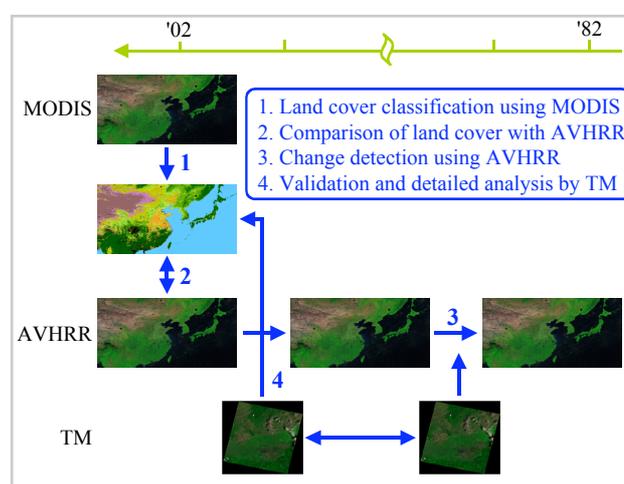


Figure 1. Outline of this study

2. METHODOLOGY

2.1 Overview

Land cover classification on the scale from regional to global has been implemented with the use of AVHRR data by means of a variety of classification method as clustering

(Pan et al., 2003; USGS - NASA Distributed Active Archive Center, 2004a), maximum likelihood classification (Liu et al., 2003), decision tree classification (Townshend et al., 1999), and hybrid algorithm of decision tree and neural network (Friedl et al., 2002). Running et al. proposed a simple new logic for classifying global vegetation based on the structural characteristics of vegetation (Running et al., 1995). Since the classification result in this study is the input parameter for the hydrological model, the critical feature of the classification method is to control the classification result using simple parameters.

2.2 Data used

MODIS products are being used by scientists from a variety of disciplines, including oceanography, biology, and atmospheric science. Most users will obtain data products by ordering them through an ordering system (Conboy). The product named "Surface Reflectance 8-Day L3 Global 250m" (abbreviated as MOD09Q1) is used in this study. MODIS 250 m Surface Reflectance is a two-band product computed from the MODIS Level 1B land bands 1 and 2 (centered at 648 nm and 858 nm, respectively). The product is an estimate of the surface spectral reflectance for each band as it would be measured at ground level if there were no atmospheric scattering or absorption (USGS - NASA Distributed Active Archive Center, 2004b). The 45 periods of MOD09Q1 products in the whole of 2002 are re-projected to the Equirectangular (latitude/longitude) projection by means of MODIS Reprojection Tool (USGS - NASA Distributed Active Archive Center, 2004c). The spatial coverage is from 50 degrees north to 20 degrees north in latitude, and 90 degrees east to 150 degrees east in longitude, and the resolution is 7.5 arc seconds. The following multi-temporal metrics are derived from the time series:

1. Annual maximum NDVI (NDVI_ann_max)
2. Annual minimum NDVI (NDVI_ann_min)
3. Annual amplitude of NDVI (NDVI_ann_amp)
4. Annual minimum band 1 reflectance (Ref1_ann_min)
5. Annual minimum band 2 reflectance (Ref2_ann_min)
6. April minimum NDVI (NDVI_apr_min)
7. June maximum NDVI (NDVI_jun_max)
8. August minimum NDVI (NDVI_aug_min)

Annual maximum/minimum of NDVI/reflectance (1,2,4, and 5) are the "quasi" maximum/minimum value, that is, the second maximum/minimum value in the 22 periods from April 23 to October 8 is selected for the purpose of avoiding the snow and erroneous data. Annual amplitude of NDVI is derived from $NDVI_{ann_max} - NDVI_{ann_min}$. April minimum, June maximum, August minimum NDVI is pure maximum or minimum value in the period from Apr. 7 to May 9, Jun. 2 to Jul. 4, and Aug. 5 to Sep. 6, respectively.

2.3 Algorithm

The supervised decision tree classification method is selected in this study for the reason of the easiness in controlling of the classification result. The decision tree classification algorithms has significant potential for land cover mapping problems (Friedl et al., 1997), and its performance is acceptably good in comparison with that of

other classifiers, except with high-dimensional data (Pal et al., 2003). The scheme of the classification method in this study is shown in Figure 2. The eight threshold values are arranged in the each steps of the decision tree, and the result of the land cover classification is tuned by these threshold values. Ref2_ann_min is applied to extract water area since the spectral reflectance of water in near-infrared wavelength band is much lower than that of land surface. NDVI_ann_max represent the most active status of the vegetation in the year, hence it is applied to categorize the non-vegetation and less-vegetation from other vegetated land. NDVI_ann_min is applied to the vegetated area for the discrimination between evergreen and deciduous vegetation. Ref1_ann_min is also applied to the vegetated area to discriminate tree type and grass type vegetation based on the rough assumption that the tree type land surface is generally "darker" in visible wavelength as against grass type land. The time series of NDVI_apr_min, NDVI_jun_max, and NDVI_aug_min are jointly used to extract double cropped agricultural area especially in the downstream of Yellow River basin. The first cropping season is from middle of February to end of May, and second season is from July to middle of November in this region. This phenological characteristic is unique compared with other natural grasslands or single cropped agricultural fields. Therefore, if NDVI_apr_min is greater than NDVI_jun_max and NDVI_jun_max is less than NDVI_aug_min, that pixel is categorized as double cropped agricultural field.

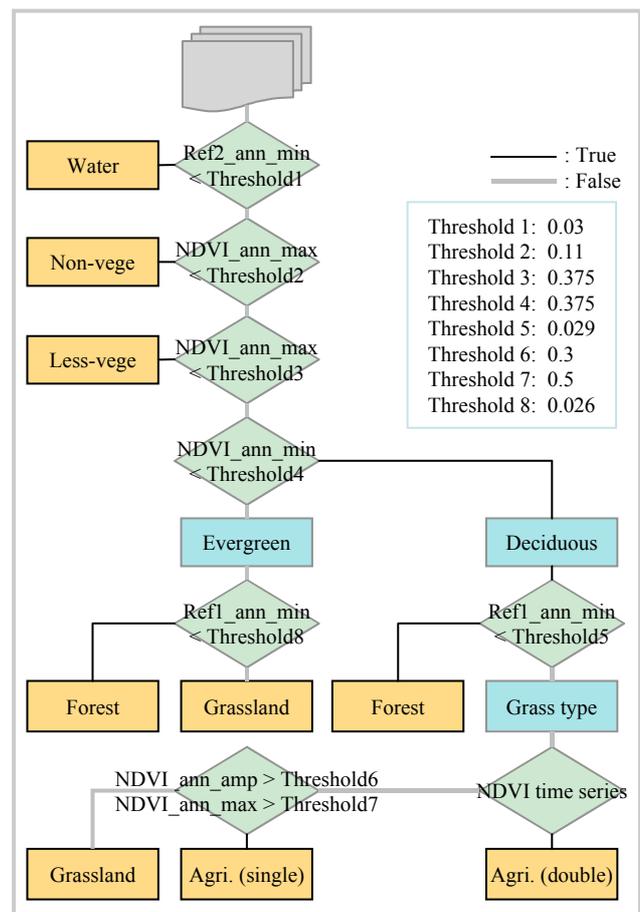


Figure 2. Flow of the land cover classification

NDVI_ann_amp and NDVI_ann_max are applied to deciduous grass type vegetation land to separate single cropped agricultural fields from natural grassland. These two categories have quite similar spectral and phenological characteristics. Therefore, this criterion (if NDVI_ann_amp is higher than threshold and NDVI_ann_max is higher than threshold, then it is categorized as single cropped agricultural field) has no supportive evidence and it is derived fully empirically by trial and error approach. Nine land cover categories (water, non-vegetation, less-vegetation, evergreen forest, evergreen grassland, deciduous forest, deciduous grassland, single cropped agricultural field, and double cropped agricultural field) are established from the viewpoints of the parameterization of the water balance model used in this study, and the accuracy of the land cover classification. In general, parameterization of the land surface properties in the physical model is expected to be difficult, and the accuracy of the land cover classification will be deteriorate, along with the increase in the class categories. Therefore minimum number of the categories is established in this stage, and category might be added hierarchically with a focus on outcomes from the model.

3. RESULTS AND DISCUSSION

The land cover classification map is shown in Figure 3. The threshold values are determined manually using image processing software. According to the visual interpretation, non-vegetation, less-vegetation, evergreen forest, and double cropped agricultural field have relatively high agreements with existing land cover maps and satellite based classification maps, though the categories of land cover are not fully corresponded each other. These categories have distinct characteristics in NDVI or reflectance. On the other hand, single cropped agricultural field is not sufficiently discriminated from grassland, and it is overestimated especially in the northern region from eastern Siberia to northeastern China, and southern region around Xizang (Tibet), Sichuan, and Qinghai provinces. In these regions, the multi-temporal metrics of grassland and agricultural field used in this study show the indistinctive characteristics as the result of the similarity in the spectral reflectance in two bands, thus NDVI, and vegetation phenology.

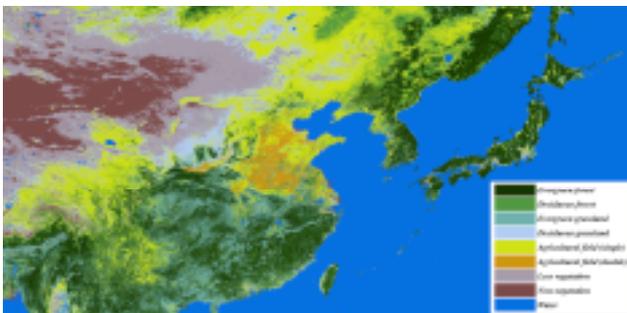


Figure 3. Land cover classification map (legend from top to bottom: evergreen forest, deciduous forest, evergreen grassland, deciduous grassland, single cropped agricultural field, double cropped agricultural field, less-vegetation, non-vegetation, and water)

Areas of each land cover category are compared with the existing land use map “1 KM Grid Data of landuse and landcover of China” (abbreviated as CASW data). This map is about 1 km grid data of land use and land cover of China in 1996. The value of a grid in every land use layer is a percent. The percent is about the real area of the land use type in this grid over 1 km² area and it is from the statistic of the area attribute of vector data of land use of China at 1:100000 (CASW, 2001). The areas of five land use categories (forest, agricultural field, grassland, less-vegetation, and non-vegetation) are figured out for each province by aggregation of the grids and appropriate categories of each map. The results are shown in Figure 4.

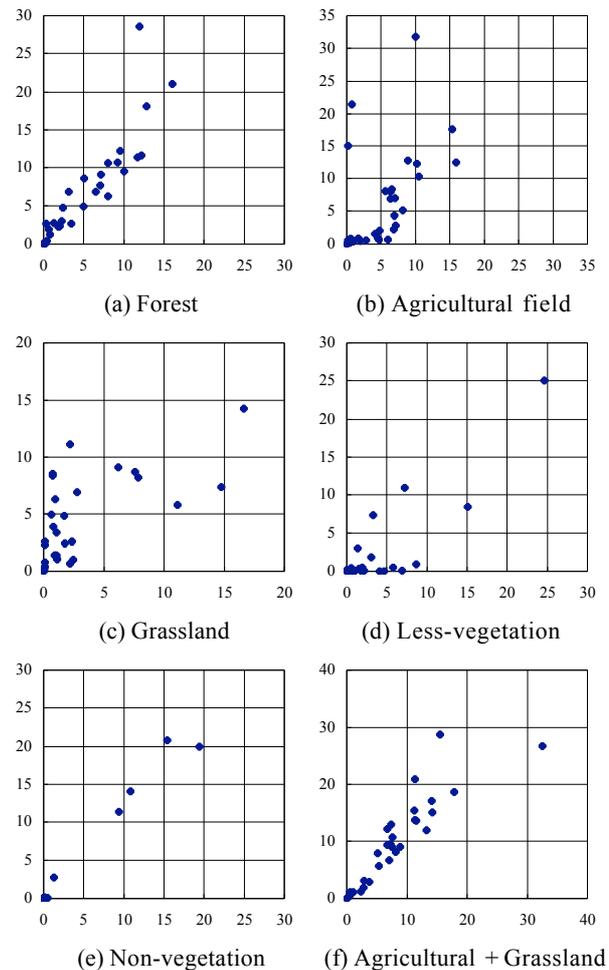


Figure 4. Comparison of land use area with CASW data (horizontal: CASW, vertical: this study, unit: 10000km²)

The good correlations are shown in forest and non-vegetation, though classification result has the tendencies to overestimate relative to the CASW data. In case of agricultural field, most provinces where there are small areas of agricultural field are resulted in underestimations. In contrast, grassland shows countertrend, that is, provinces that have small grassland areas have trends of overestimation, but in provinces with larger grassland,

underestimations. The comparison of total area of agricultural field and grassland is shown in Figure 4 (f). Total area has much better agreement than individual comparison. This might be because of the difficulty in discriminating the single crop agricultural field from natural grassland mentioned above. To overcome this insufficiency of performance, it might be given some indication of the utilization of additional information such as other temporal metrics, fine resolution satellite data, and digital elevation model. The tuning of the threshold values will be implemented by means of TM and ETM+ data. These higher resolution image are geometrically overlaid to the MODIS, and finer distribution of land use is interpreted manually and semi-automatically to get the appropriate threshold values.

4. CONCLUSIONS

The land cover classification over Yellow River basin by means of simple decision tree classification method using MODIS data is shown in this paper. This method is designed to control the classification result by tuning of threshold values, and to be applied to the multi-temporal metrics that represent the phenological characteristics of the vegetation. The classification result shows basically good agreements by comparison with the existing land cover map, though a certain degrees of overestimation or underestimation are recognized. The tuning of the threshold values would be implemented by the aid of fine resolution satellite data and existing maps. The additional data to get around the difficulty in discrimination of single cropped agricultural field from natural vegetation should be investigated.

References from Journals:

Friedl, M. A., et al., 1997. Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, 61, pp. 399-409.

Friedl, M. A., et al., 2002. Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, 83, pp. 287-302.

Liu, J. Y., et al., 2003. Land-cover classification of China: integrated analysis of AVHRR imagery and geophysical data. *International Journal of Remote Sensing*, 24(12), pp. 2485-2500.

Pal, M., et al., 2003. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment*, 86, pp. 554-565.

Pan, Y., et al., 2003. An integrate classification of vegetation in China based on NOAA AVHRR and vegetation-climate indices of the Holdridge life zone. *International Journal of Remote Sensing*, 24(5), pp. 1009-1027.

Running, S. W., et al., 1995. A remote sensing based vegetation classification logic for global land cover analysis. *Remote Sensing of Environment*, 51, pp. 39-48.

References from websites:

CASW, 2001. 1 KM Grid Data of landuse and landcover of China. <http://www.casw.com.cn/echannel/aboutus/eintroduction.asp> (accessed 30 Apr. 2004).

Conboy. MODIS home page. <http://modis.gsfc.nasa.gov/> (accessed 30 Apr. 2004).

Kidwell, K. B., 1998. NOAA polar orbiter data user's guide. <http://www2.ncdc.noaa.gov/docs/podug/> (accessed 30 Apr. 2004).

Sheffner, E., 1999. Welcome to the Landsat program. <http://geo.arc.nasa.gov/sge/landsat/landsat.html> (accessed 30 Apr. 2004).

Townshend, J., et al. 1999. MODIS Enhanced Land Cover and Land Cover Change Product Algorithm Theoretical Basis Documents (ATBD). <http://modis.umiacs.umd.edu/main.htm> (accessed 30 Apr. 2004).

USGS - NASA Distributed Active Archive Center, 2004a. Global land cover characterization. <http://edcdaac.usgs.gov/glcc/glcc.asp> (accessed 30 Apr. 2004).

USGS - NASA Distributed Active Archive Center, 2004b. MODIS data products. <http://edcdaac.usgs.gov/modis/dataproducts.asp> (accessed 30 Apr. 2004).

USGS - NASA Distributed Active Archive Center, 2004c. MODIS reprojection tool distribution page. <http://edcdaac.usgs.gov/landdaac/tools/modis/index.asp> (accessed 30 Apr. 2004).