Classification of vegetation in Japan using MODIS by machine learning method

Younghwan Kim , Jonggeol Park, and Icho Asanuma
Tokyo University of Information Sciences, 4-1, Onaridai, Wakaba-k, Chiba 265-8501, Japan,
Email:h16001yk@edu.tuis.ac.jp

KEY WORDS: pixel-based, object-based, machine learning, classification,

ABSTRACT: The classification of land use and land cover(LULC) from remotely sensed imagery can be conducted with two general image analysis approaches: i) pixel-based classifications, and ii) object-based classifications. While pixel-based analysis has long been the mainstay approach for classifying remotely sensed imagery. However, object-based analysis has become increasingly commonplace over the last years. Topography and geology are factors to characterize the distribution of natural vegetation. Topographic contour (elevation, slope, slop direction) determines the living conditions of plants such as soil moisture, sunlight, and windiness. The similar topographic conditions exhibit the similar distribution of vegetation unless natural disturbances or artificial disturbances being occurred. A vegetation map of Japan was developed using an object-based segmentation and pixel-based approach, included in machine learning package, with topographic information and climate information (monthly average temperature, monthly average precipitation) with MODIS remotely sensed data. The results of four methods were compared: i) object-based considering climate information ii) pixel-based considering climate information iii) object-based not considering climate information, and iv) pixel-based not considering climate information. Through the comparison, the object-based classification is more effective to produce a vegetation map than the pixel-based classification. In addition, the classification accuracy with considering climate information was higher than not considered. In a study on areas with elongated topographical features like Japan, it is necessary to consider horizontal distribution due to the latitude of vegetation and vertical distribution due to altitude.

1.Introduction

American satellite Landsat was launched, the observation method has changed to the site center to the remote sensing which can observe a wide area by the uniform precision at the same time. Due to the development of sensor technology and the appearance of high resolution satellites and the diversification of bands, the fields that can be applied also increased. In a pixel-based segmentation approach using high-resolution satellite imagery, it often happens that an area is not recognized as a single object because multiple types of land cover are detected within the area. The method which is called the object-oriented classification which cancels a problem of a classifying process of a pixel base is being popular. An object-based approach divides an image into small segments called image objects and is capable of discerning objects as accurately as visual interpretation by experts. Therefore, the research report for which object-oriented classification was used is increasing.

(Murakami, 2010) examined the relationship between parameters required for segmentation and classification accuracy for object-based forest phase classification using IKONOS pan sharpen image. (Yamamoto, 2011) examined parameter setting and feature quantity by object-oriented classification for plantation, one of the representative agricultural imprisonment in Indonesia, using Quick Bird's pan sharpen image. (Zhu, 2014) used low-resolution images and high-resolution images in combination and made a forest classification using a new object-based image analysis method in the Ogawa area of Sekimoto-cho, Ibaraki-shi. (Kamagata, 2006) clarified the effectiveness of the object-oriented classification method in the mixed area of features such as satoyama landscape area using IKONOS data. In the object-oriented classification, information such as shape and texture represented by an aggregate of adjacent pixels is easily reflected in classification, and it is possible to generate statistics such as mean and standard deviation for each object. Also, it is difficult to generate minute misclassification as seen in the pixel-based classification and it is easy to obtain classification results close to visual interpretation (Yamamoto, 2011).

In previous studies, cases of research on municipal scales with high resolution images have been reported, but there are very few research reports on wide scale. Also, there are no reports of vegetation classification taking climate information into consideration. Therefore, in this research, we aim to develop vegetation classification method by climate information using low resolution satellite data MODIS.

2. Data

We used SRTM90 for topographical analysis in this research. MCD43A4 which is surface reflectance data was used to extract vegetation phenology information. It was used 3 years from 2002 to 2004. In order to divide the area by climate information, the average amount of precipitation and average temperature data of the Geographical Survey Institute's average year mesh was used. This data is an estimate calculation of the average value for each 1 km mesh (tertiary mesh) from observation values in the past 30 years.

3.Methodology

Figure 1 is the flow chart of this research. First, determine relative elevation and slope from DEM data. Calculate duration of on-set and off-set in contact with reflectance 0.35 and 0.55 from MODIS data. Determine the size of the object using relative elevation, slop, and phenology information. Next is changed from pixel based (objective, precipitation, average temperature, MODIS) to object based using finally determined object information. Subsequently, zoning is performed using the precipitation amount and the average temperature data to divide the region by the presence or absence of snow. The zoned information is superimposed on the satellite data, and the area is divided. Finally, vegetation classification for each area is done using training data extracted from the vegetation map. The training data used only the area with land coverage of 80% or more.

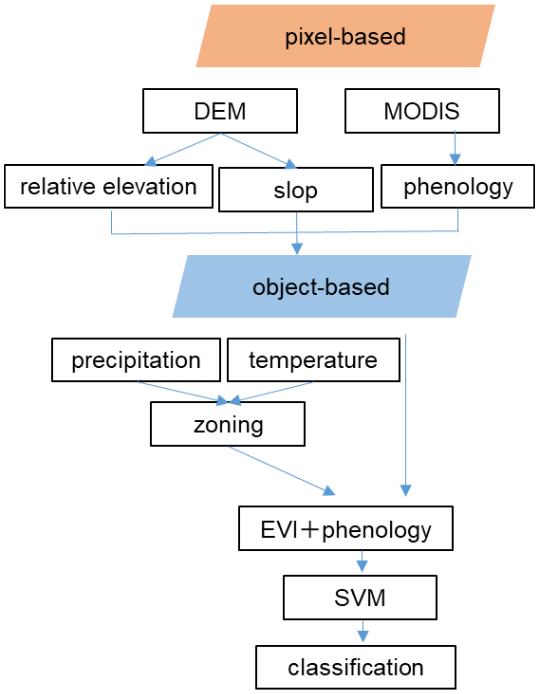


Figure1. Research flow chart

3.1 Conceptual Diagram of Relative Elevation

The relative elevation between a given location and the river was calculated as the difference in DEM between the pixel of the location and the pixel adjacent to the river that is determined by tracing the water flow from the location's pixel. Flat areas can be detected by calculating the slope angle of each pixel using a 3×3 array of DEM data. Figure 2 shows an example of calculating relative elevation. Figures 2a and 2b show DEM and river data, respectively. Figure 2c shows the direction of water flow from each pixel. Water flows from the target pixel to one of its surrounding eight pixels that has the largest elevation difference in comparison with the target pixel. Figure 2d shows relative elevation. In this case, water starts from DEM(3,1) through DEM(2,2) then to DEM(3,3), and finally reaches River(3,3). Thus, the relative elevation between DEM(3,1) and the river is calculated as: DEM(3,1) − DEM(3,3) = 8.

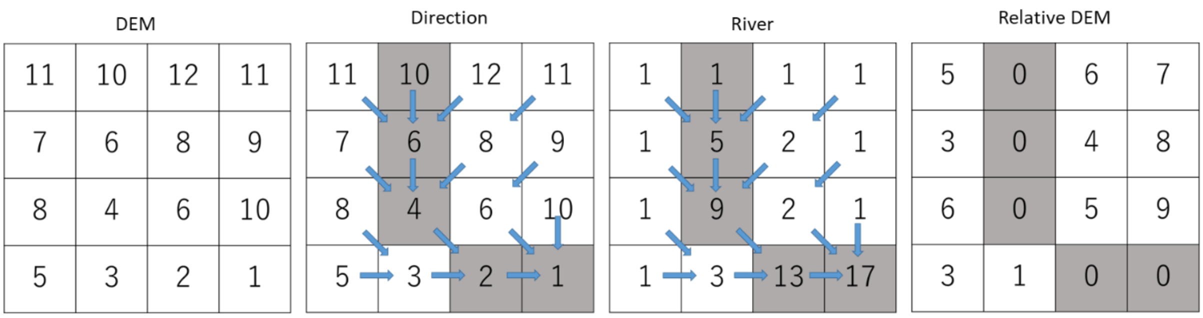
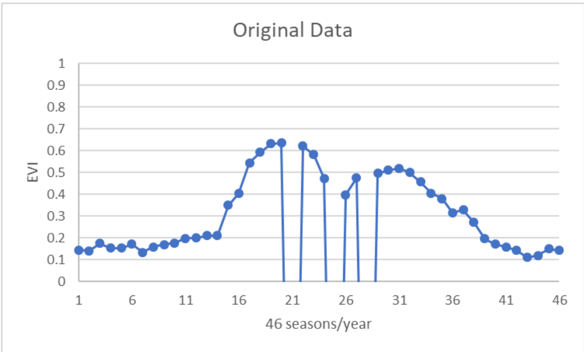


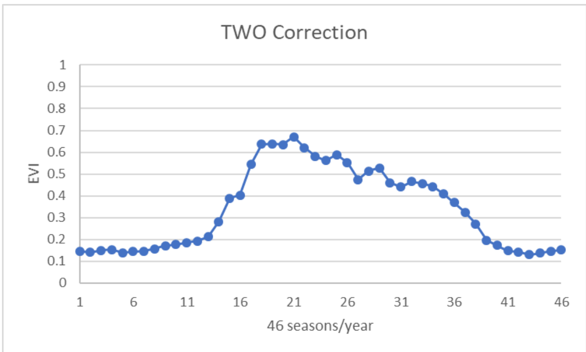
Figure2. Calculation of relative elevation

3.2 Profile Correction

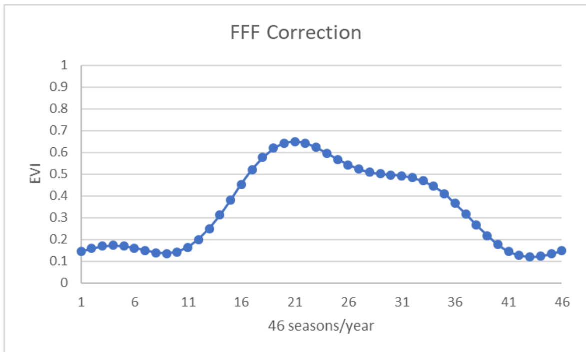
Time series EVI data was used to extract vegetation phenology information. However, even corrected products contain cloud effects and noise. Therefore, profile correction is necessary. Figure 2 is the original profile of time series EVI. As shown in Figure 2-a, data loss and noise are seen in several sections. Therefore, in this research, two-stage profile correction was performed to minimize the influence of noise. In the first stage, the pixels supposed to have clouds were removed by TemperalWindowOperation method and linear interpolation was performed (Park, 1999). However, it is unlikely that vegetation phenology repeatedly increases and decreases like in from 25th to32nd in Figure 2-b. Therefore, we made two stages of correction with FFT. A high frequency range including noise and clouds was removed with a low pass filter.



2-a



2-b



2-c

Figure2. Profile correction

3.3 Conceptual Diagram of Segment

Object oriented classification is used for high resolution data analysis. The reason is that the size of the object can be decided while visually checking the terrain feature on the ground. However, since MODIS data has low spatial resolution, it is difficult to read the MODIS data by visual inspection. Therefore, we thought that vegetation can be divided by using terrain, slope direction, and phenology information that are closely related to vegetation distribution. Figure 3 shows segment results for each process. A color image is the vegetation map of the Ministry of the Environment, and black and white image is the segment image. Step 1 is an image of the result using relative elevation data. The vegetation map is divided by several terrain parameters. The result of segmenting by relative altitude is divided in boundary like the vegetation map. The next step 2 is a result image using the slope direction information. It can be confirmed that it is divided finer than step 1. Step 3 is the result using duration 0.35 and 0.55. Although it is difficult to confirm by visual interpretation, it can be confirmed that an object is generated with certain specific phenology information. For the object value generated here, the average value of the pixels included in the range was used as a representative value.

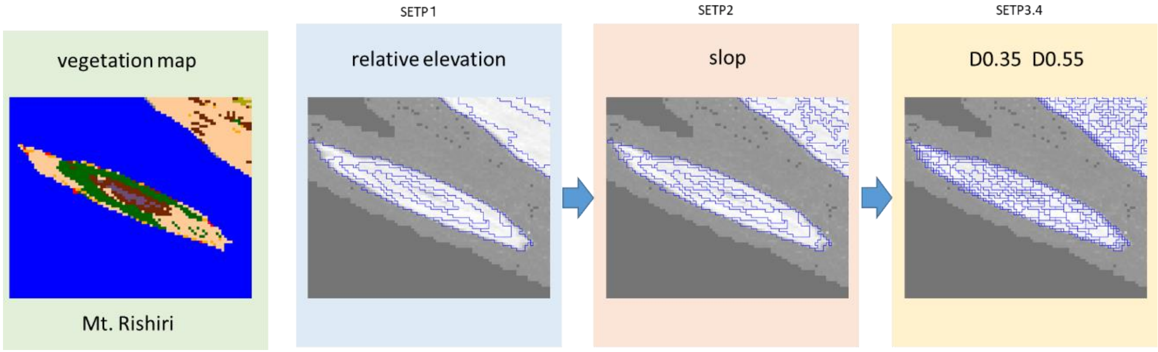


Figure3. Segmentation of the mountain area of Mt. Rishiri in Hokkaido

3.4 Comparison of Profiles of Japan Sea side Climate Vegetation and Pacific Climate Vegetation

Figure 4 compares the time series EVI of the Japan Sea side climate and the Pacific climate vegetation. Blue shows deciduous forest of climate of Japan Sea side, red is evergreen forest of climate of Japan Sea side and green is evergreen forest of Pacific climate. Evergreen forests where the frequency of snowfall is low are less likely to change year-round like the green profile. However, it is difficult to distinguish between fallen leaves and evergreen forests where snow falls from December to March like the Japan Sea side climatic zone, as the profile change amounts are similar。 If the whole country is treated uniformly without dividing it in the climate zone, the evergreen forest of the Japan Sea side climate is likely to be misclassified as deciduous forest.

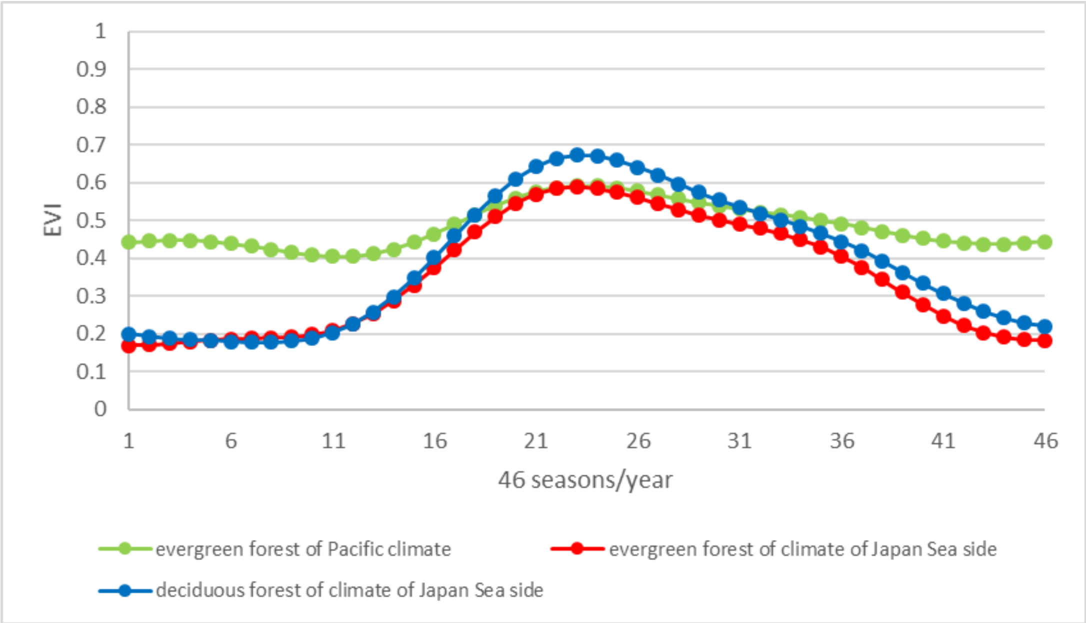


Figure4. Comparison of vegetation profiles of each climate zone

4 .Result

As the classification method, SVM (Support Vector Machine) which is supervised machine learning was used。

Figure 5 shows the Asahikawa area around Hokkaido. From the left, it is object, pixel, vegetation chart. Many mixed forests are distributed in the image of the result of object and pixel on the image of vegetation map. Moreover, it can be confirmed that salt-and-pepper noise seen in pixel-based is decreased in object-based. And it is surrounded by a circle of green that is different from the vegetation map. This is considered to be the result of extracting agricultural land data adjacent to deciduous broad-leaved trees when preparing training data. For this reason, a crop area that should not be detected should be detected. In the resulting image is around Rishiri, it can be confirmed that the result as well as deciduous around Asahikawa were misclassified as farmland. In order to solve such problems, it is necessary to consider adjacent classes around the training data. This is the result by using the information of the vegetation map of 1/50000. However, there was a place where ground-truth of corrected reflectance and vegetation do not match. In the future, it is necessary to use vegetation map and google earth together and to create training data.

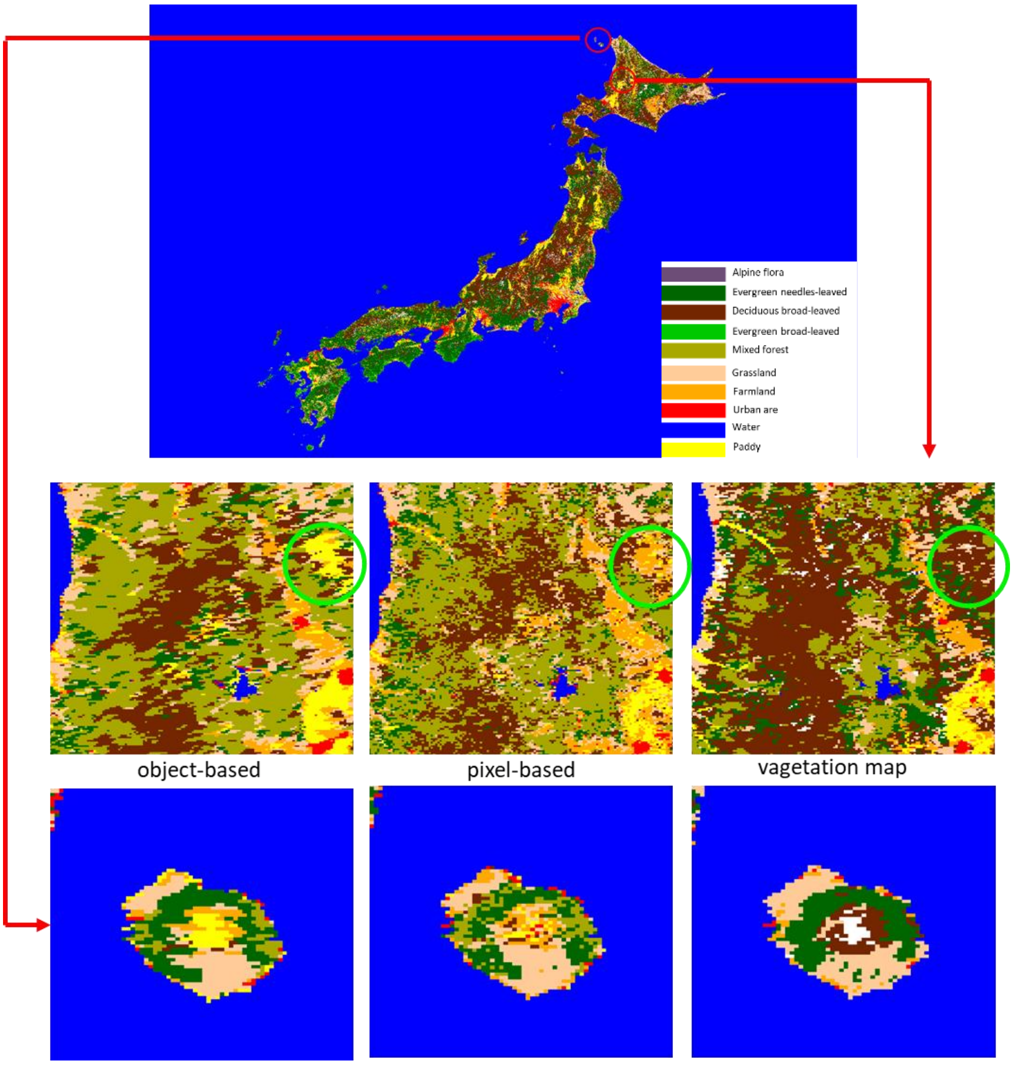


Figure5. Comparison of object based classification and pixel based classification

5. Conclusion

In this research, we did vegetation classification in Japan using low-resolution satellite MODIS. First, relative elevation and inclination were obtained from DEM data. We determined the size of the object using relative elevation, slope, MODIS phenology information. Next, it changed to object based from pixel based (precipitation, average temperature, MODIS) using object information finally decided. After that, zoning was carried out using precipitation and average temperature data, and the area by the presence or absence of snow was divided. The zoned information was superimposed on the satellite data, and the area was divided. Finally, vegetation classification for each area was done using training data extracted from the vegetation map. As a result, the accuracy of the object based method was 82.69% and the pixel based was 79.54%. Also, the salt-and-pepper noise seen in pixel-based was reduced in object-based. As a result, by using topographical information and phenology information, we found that object-based in low-resolution satellite data set is effective. However, the object based method using the high-resolution satellite was not suitable for human activity areas with complicated shape. When classifying vegetation in urban areas and cultivated areas, it is necessary to use a satellite such as LANDSAT which has higher resolution than MODIS.

Reference

Murakami, 2010. Examination of the optimum segmentation in the object-based image analysis for forest stand type classification. JSPRS, 49(3), pp. 159-165

Yamamoto, 2011. Feature Abstraction of Plantation Farms in Indonesia by Object-based Classification. 27(4), pp.137-148

Zhu, 2014. Development of Object-based Forest Type Classification Methodology. Journal of The Remote Sensing Society of Japan , 34(5), pp.341-355

Kamagata,2006. A new method of vegetation mapping by object-based classification using high resolution satellite data. JSPRS, 45(1), pp.43-49

Park, 1999. A proposal of the Temporal Window Operation(TWO) method to remove high-frequency noises in AVHRR NDVI tome series data. JSPRS, 38(5), pp.36-47