韓國林學會誌 83(4) : 441-449. 1994
Jour. Korean For. Soc. 83(4) : 441-449. 1994

# Vegetation Cover Type Mapping Over The Korean Peninsula Using Multitemporal AVHRR Data[1]

Kyu-Sung Lee[2]

# 時系列 AVHRR 衛星資料를 이용한 한반도 植生分布 區分[1]

李 奎 成[2]

## ABSTRACT

The two reflective channels(red and near infrared spectrum) of advanced very high resolution radiometer (AVHRR) data were used to classify primary vegetation cover types in the Korean Peninsula. From the NOAA-11 satellite data archive of 1991, 27 daytime scenes of relatively minimum cloud coverage were obtained. After the initial radiometric calibration, normalized difference vegetation index(NDVI) was calculated for each of the 27 data sets. Four or five daily NDVI data were then overlaid for each of the six months starting from February to November and the maximum value of NDVI was retained for every pixel location to make a monthly composite. The six bands of monthly NDVI composite were nearly cloud free and used for the computer classification of vegetation cover. Based on the temporal signatures of different vegetation cover types, which were generated by an unsupervised block clustering algorithm, every pixel was classified into one of the six cover type categories. The classification result was evaluated by both qualitative interpretation and quantitative comparison with existing forest statistics. Considering frequent data acquisition, low data cost and volume, and large area coverage, it is believed that AVHRR data are effective for vegetation cover type mapping at regional scale.

Key words : vegetation mapping, remote sensing, AVHRR, vegetation index

## 要 約

본 연구의 目的은 현재 한국에서 자료획득이 비교적 용이한 AVHRR 衛星資料를 이용하여, 한반도 全地域을 대상으로 植物의 時期別 變化類型을 분석하고 이를 응용하여 主要植生의 분포를 구분하고자 한다. 1991년 1년동안 NOAA-11 위성에서 受信된 AVHRR 자료중 비교적 雲量이 적은 날을 택하여 총 27일분의 日別映像資料를 추출하였다. 일별영상자료는 먼저 光學的 補正을 마친 후, 赤色파장대 및 近赤外線파장대에서의 反射特性을 조합한 植生指數(NDVI-Normalized Difference Vegetation Index)로 변환되었다. 구름으로 덮혀있는 지역의 식생지수는 식물이 존재하는 지역보다 상대적으로 낮은 값을 나타내므로, 구름제거를 위하여 4-5개의 日別植生指數資料를 중첩한 뒤 각 畵素지점의 식생지수중 최대치를 선택함으로써 구름의 영향이 최소화된 月別植生指數資料가 산출되었다. 월별식생지수자료는 식물 생장의 年中變化를 비교 분석하기에 용이하도록 非生長期間까지 포함하여 2월, 3월,

[1] 接受 1994年 3月 7日 Received on March 7, 1994.
[2] Presented at the IUFRO Conference on "Advancement in Forest Inventory and Forest Management Sciences", Seoul, Korea, 20-25 September 1993.
[3] 林業研究院 Forestry Research Institute(Presently with Inha University, College of Engineering)

5월, 8월, 9월, 그리고 11월까지 6개가 산출되었다. 식생별로 相異한 계절별 잎의 발달상태에 따라, 6개의 月別植生指數資料에 나타나는 식생지수의 변화특성을 이용하여 植生分類를 실시하였다. 사용된 자료의 광학적 解像力을 고려하여 분류집단은 침엽수림, 활엽수림, 침활혼효림, 농지, 초지관목림, 그리고 도시지역으로 구분하였다. 컴퓨터분류방식은 植生指數의 변화유형이 비슷한 집단끼리 스스로 糾合되게 하는 無監督類集分類法(unsupervised clustering)을 채택하였다. 컴퓨터분류 결과를 기존의 山林資源調査資料와 비교한 결과 상당히 근접한 통계치를 보여주었고, 산림지역內에서도 침엽수림, 활엽수림, 혼효림의 구분 또한 만족할만한 결과를 나타내고 있다. 넓은 지역을 대상으로 필요한 映像資料를 비교적 신속하고 용이하게 受信할 수 있고, 他 위성자료에 비교하여 자료의 양이나 가격 측면에서 유리한 AVHRR자료는 한반도 규모에 상응하는 넓은 지역의 식생현황을 주기적으로 모니터링하기에 적합한 위성자료로 판단된다.

## INTRODUCTION

During the past two decades, satellite remote sensor data have been widely used to derive information related to bio-physical conditions of various features on the earth surface. In particular, the two earth observation satellite programs(Landsat and SPOT) have provided enormous volume of imagery data for vegetation monitoring over large geographic area. Although the two data types have provided relatively fine spatial resolution, their usage for the cover type mapping has been limited to the area of approximately $50^2$-$200^2$ $km^2$ size that is approximately equivalent to the area covered by one scene of Landsat or SPOT data. If the area to be mapped is larger than the above size, several scenes of imagery should be stitched together. Processing several scenes of Landsat and SPOT data may cause problems related to the data acquisition, data calibration, and data processing. Due to the orbit cycle of 16 days for Landsat and 26 days for SPOT and their relative narrow swath widths, it is often difficult to obtain multiple scenes for the contiguous area at the same day without cloud cover. Even if one can obtain necessary data sets collected at different months or dates, the analyst should concern the temporal variations in solar illumination, growing stage of plants, and atmospheric conditions.

As an alternative to those fine resolution remote sensor data, NOAA Advanced Very High Resolution Radiometer(AVHRR) data have been used to derive vegetation information at continental and global scales. Vegetation index that can be calculated from the red and near infrared spectral channels of AVHRR data has been a valuable source of information for vegetation monitoring over large geographic areas(Eidenshink, 1992 ; Loveland et al., 1991 ; Townshend and Justice, 1986). Due to short repetition orbit cycle and large area coverage of AVHRR, they are also very effective to analyze the temporal changes of vegetation dynamics over continental scale. However, since AVHRR data has relatively coarse spatial resolution, they have been rarely used for cover type mapping. In recent years, there have been a few attempts to classify vegetation cover types using multitemporal AVHRR data(Derrien et al., 1992 ; Zhu and Evans, 1992).

The objective of this study is to define the potential of AVHRR data for vegetation cover type mapping at a regional scale($10^3$-$10^6$ $km^2$), in which it is often difficult to achieve such information by any other methods. This study also intends to describe pertinent methodology for processing multitemporal AVHRR data to map over the region where prevailing cloud cover can be often obstacle in obtaining necessary data. Regional-scale vegetation cover information that is regularly provided by multitemporal AVHRR data can be applicable to both resource management and environmental monitoring purposes.

## STUDY AREA AND DATA USED

Due to the geographical location in a temperate climate zone, the Korean Peninsula has diverse vegetation types with high seasonal variations in canopy development. Although satellite remote sensor data had been introduced and tested for a few applications in Korea, their primary uses are very site-specific and still remain as experimental stage. Under the current political situation between the south and the north, the vegetation-related information over the northern part of the peninsula is generally not available or outdated. There have been no attempts to produce a vegetation cover type map over the whole area of the Korean Peninsula using remote sensor data, probably due to the difficulty of obtaining necessary data. According to the climate data collected by the Korean Meteorological Service, the number of clear day (having less than 25% of cloud coverage) in Korea is about 90 days per year. Prevailing cloud cover, large data volume and cost, and lack of processing facilities would be the main reasons to restrict the vegetation cover mapping over the peninsula by using remote sensor data.

Considering large area coverage, short orbit cycle, and relatively low data volume and data cost, AVHBR data can be a very attractive alternative to derive vegetation information on large geographic areas. In addition, the establishment of receiving station for the NOAA satellites does not cost as much as for the Landsat and the SPOT. There are currently a few receiving stations to collect AVHRR data in Korea and their primary applications are oceanographic uses, such as deriving sea surface temperature.

In this study, AVHRR data were obtained from a single satellite (NOAA-11) among the several NOAA satellites to maintain consistency in sensor characteristics. Table 1 shows brief characteristics of AVHRR data obtained from the NOAA-11 satellite.

**Table 1**. Characteristics of the NOAA-11 AVHRR data.

| | |
|---|---|
| Launch | September, 1988 |
| Orbital Period | 102 min(14.1 orbits/day) daylight ascending, nighttime descending |
| Scan Angle from Nadir | ±55.4° |
| Orbit Altitude | 833 km |
| Coverage | every 12 hours |
| IFOV at Nadir | 1.1 km |
| Swath Width | approximately 2,400 km |
| Spectral Wavelength | Channel-1 0.58-0.68 μm |
| | 2 0.72-1.10 μm |
| | 3 3.55-3.93 μm |
| | 4 10.3-11.3 μm |
| | 5 11.5-12.5 μm |
| Gray Level | 10 bit |

The NOAA-11 satellite passes over the Korean Peninsula at around 2:00 pm everyday. After reviewing the cloud coverage from the weather data of 1991 and displaying the actual imagery on a display monitor, 27 daily AVHRR scenes ranging from February to November that had relatively small cloud covers were obtained. There were four or five daily AVHRR scenes for a month, which enable to make a cloud-free monthly composite. It was originally intended to build monthly composite for every month. However, only six monthly composites were generated due to the persistent cloud cover during winter and summer months.

## METHODS AND PROCEDURES

The classification scheme applied in this study was based upon an unsupervised clustering algorithm. In general, clustering process evaluates every pixel according to their spectral reflectances from multispectral bands and classifies them into a group of similar spectral values. In this classification, however, temporal signatures generated from the six monthly NDVI composites were used instead of those spectral signatures.

### Preprocessing

Each pixel of the raw AVHRR data is represented by a 10-bit depth digital number (DN) values ranging from 0 to 1023. The DN values of original AVHRR data were converted to spectral reflectance as to be compared with other AVHRR data collected at different date. It is often important to maintain the consistence of reflectance for processing multitemporal data sets. This radiometric calibration of converting the original DN values to spectral reflectance was performed from the following equation for the two channels of visible and near infrared spectrum.

$$A_i = \alpha_i \ DN_i + \beta_i \qquad (1)$$

where $A_i$ = % albedo of channel i

$\alpha_i, \beta_i$ = slope and gain coefficients for channel i

### Geometric Registration

The 27 daily AVHRR data sets were registered on a common map coordinate system to ensure that the pixel-by-pixel comparison could be performed on the same geographic location. Initially, the raw images were registered to a reference image by the image-to-image registration method. During the image-to-image registration, the nearest neighbor resampling method was used to preserve the original reflectance values. After the final classification, the classified image was then registered on a rectangular coordinate of Transverse Mercator projection with a pixel resolution of 1km(Lee, 1992).

### Maximum Value Monthly NDVI Composite

Vegetation index techniques have been developed to assess the relative amount of green vegetation from remote sensing data. In addition, the vegetation index calculation is known to be very effective for normalizing the different illumination effects

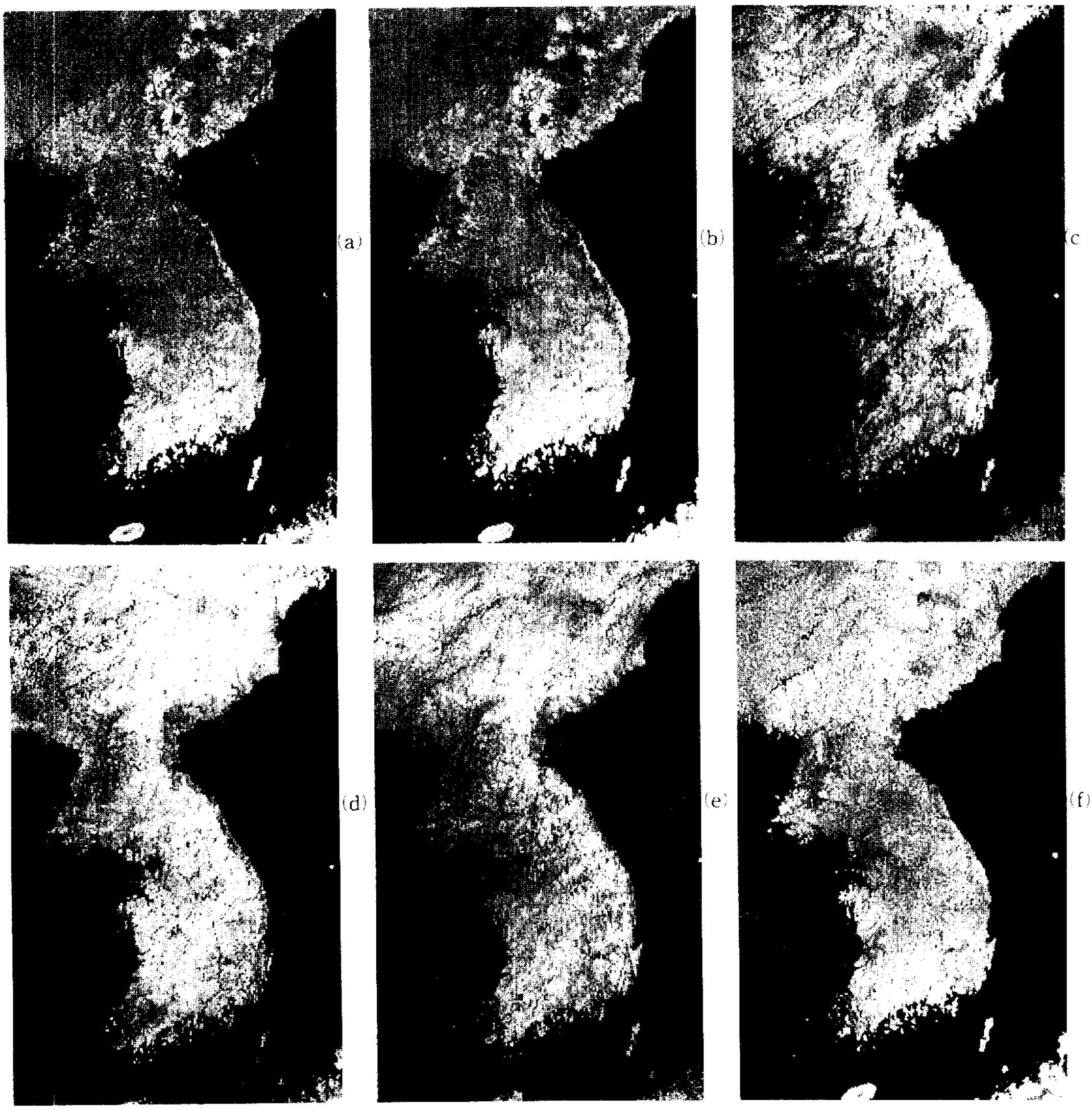

**Fig. 1.** Six monthly composites of NDVI data : (a) February, (b) March, (c) May, (d) August, (e) September, and (f) November. The gray tone is proportional to the values of NDVI in which lighter tone indicates more green vegetation.

caused by solar angle and atmospheric condition. Although there are several methods of calculating vegetation index using two spectral bands of red and near infrared wavelength, the normalized difference vegetation index(NDVI) has been most widely used in many fields of applied remote sensing community. It is calculated by dividing the difference of two spectral reflectance values($ch_2$-$ch_1$) by the sum of two spectral reflectance values($ch_2+ch_1$). Theoretically, NDVI value ranges from -1.0 to 1.0, in which the maximum value 1.0 suggests the most green vegetation. NDVI was calculated for each of the 27 daily data sets.

As stated earlier, the monthly composites of NDVI were produced to obtain cloud free imagery over whole area of the Korean Peninsula. Four or five daily NDVI layers were overlaid and the highest NDVI value was retained for every pixel location to produce a monthly composite. While the NDVI values for the pixels covered by cloud are approximately zero or lower, the NDVI at the cloud free area would be higher than zero. By comparing four or five pixels on the same ground location, the maximum value was retained and the resultant imagery would have nearly cloud free. The maximum-value composite technique reported by Holben (1986) is known to be very effective to produce a cloud free NDVI imagery. This technique also minimizes the effects of different sun angles, shadows, and atmospheric conditions existing among multiple data sets. Six bands of monthly NDVI composite were produced for February, March, May, August, September, and November(Figure 1).

**Computer-Aided Classification**

From the images in Figure 1, we can observe the seasonal differences of canopy development. Since vegetation index is correlated to the relative amount of green foliage, the winter images show the coniferous forests with light tone. Once the leaves come out in spring and summer, the deciduous forest has lighter gray level than the coniferous forest while the non-vegetated areas such as urban areas appear dark throughout the year.

Six bands of monthly NDVI composite were classified by an unsupervised block-clustering scheme. Because of coarse spatial resolution and lack of reference data to delineate reliable training fields, it was believed that unsupervised classification scheme was more suitable than supervised scheme. Assuming that developmental stage of plant canopy from early months of spring to late fall varies by different vegetation types, temporal variation of the NDVI values would have distinct patterns that can be separable each other by clustering.

The vegetation type and growing season are known to be rather different by latitudinal locations along the Korean Peninsula. For instance, certain deciduous species have one month of time interval for coming into leaves in spring by the geographic location. Such temporal variation can also occur in fall when leaves fall. Although there could be more detailed vegetation classes by certain species composition, tree size, and stand density, only six major cover types were defined for the classification(Table 2).

Since the type and growing pattern of forest vegetation are rather distinct by geographic region within the peninsula, four separate blocks were selected from the southern coastal area to the high mountains of northern region. Within each block, approximately five to ten clusters were generated by a clustering algorithm of the iterative self organizing data analysis technique(Tou and Gonzales, 1974). Total 31 clusters were generated from the four blocks and the statistics(mean and covariance matrix) for each

**Table 2.** Cover type classes defined for the computer classification.

| Classes | Descriptions |
|---|---|
| Conifers | mostly pines with some spruce, fir, and larch |
| Hardwood | mixed deciduous hardwoods(oaks, maples, birches, ashes, etc ···) |
| Mixed | mixed hardwood and conifers |
| Agriculture | rice pads and other agricultural crop |
| Grass | grass and shrub on natural conditions |
| Urban | urban and other industrial areas, bare soil and rock |

cluster were obtained for classifying the whole area. Once the classification of the whole area using the 31 cluster statistics was done, each of the 31 spectral classes was evaluated and designated to one of the six information classes previously defined.

**Evaluation on the Performance of Computer Classification**

One of the weaknesses of computer aided classification of remote sensor data is lack of an ideal method to evaluate the accuracy of classification result. Contingency matrix, that can be produced by comparing the classification result to known ground truth, has been most widely used although the selection of reference samples is still subjective to the analyst's experience. It was difficult to obtain reliable reference data that true ground features can be identified on a scale comparable to the resolution of AVHRR data. This is particularly true for the northern portion of the peninsula where there is almost no information for reference. Furthermore, the field checking was technically difficult because of the 1 $km^2$ ground resolution of AVHRR data. For these reasons, the classified result could not be evaluated by a contingency matrix.

The classification results were evaluated by both quantitative comparison of the acreage for each cover type between the classified result and existing forest inventory statistics and visual interpretation of the classified map. There are detailed forest inventory data for the southern half of the peninsula, which have been conducted by the Forestry Research Institute(1991) while the forest statistics of north Korea were obtained from very limited literature published in a third country(Mueller, 1987). The classified map was also qualitatively evaluated by visual interpretation.

## RESULTS AND DISCUSSIONS

Each of the 31 spectral classes was carefully evaluated by its statistics and limited reference data and merged into one of six cover type classes listed above. The final classification result is shown in Figure 2. In overall, the classification result matches well with the general land cover pattern in which most agriculture lands are located along the west coast of the peninsula and most forests occur over the mountainous areas along the east coast and northern peninsula. The classified map also shows well the major urban areas of Seoul, Pyongyang, Pusan, and Teagu. Overall size and distribution of forest lands appear reasonable.

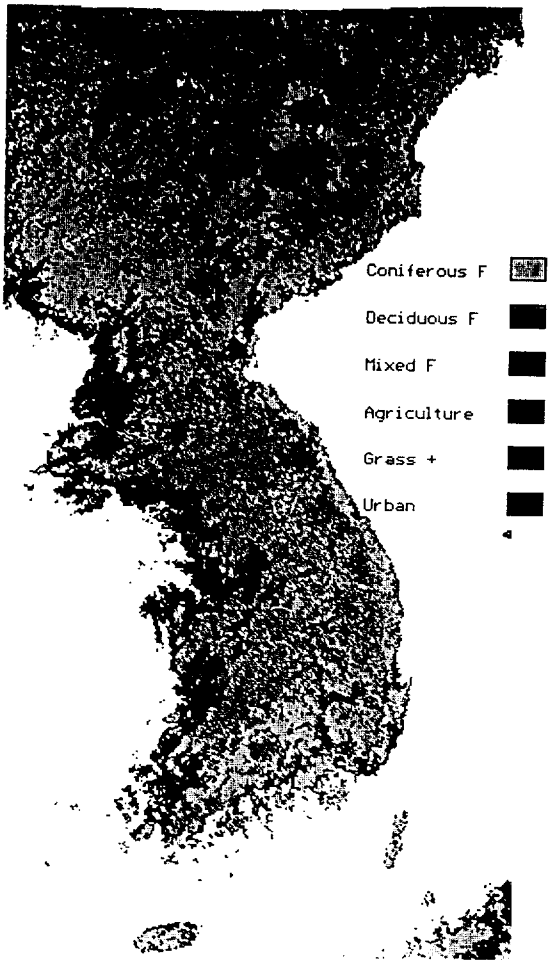


Fig. 2. Computer classified map of primary vegetation cover types over the Korean Peninsula using the six monthly NDVI data derived from multitemporal AVHRR data set.

The acreage of forest cover type obtained from the computer classification were compared to the existing forest statistics(Table 3). The overall statistics of forest cover types over the whole peninsular are very similar between the classification result and the reference data. The reference data of the South

**Table 3**. Comparison of total areas between the computer classification and existing forest statistics.

(Unit : km²)

| Region / Type | Whole Peninsula classified | Whole Peninsula reference | South Korea classified | South Korea reference | North Korea classified | North Korea reference |
|---|---|---|---|---|---|---|
| Conifers | 64261(28.9) | 73742(33.3) | 30465(30.4) | 29842(30.1) | 33824(27.7) | 43900(35.9) |
| Hardwood | 58354(26.2) | 52195(23.5) | 20959(20.9) | 15515(15.6) | 37497(30.7) | 36600(29.9) |
| Mixed | 28378(12.8) | 25600(11.5) | 13363(13.3) | 17400(17.5) | 15063(12.3) | 8200(6.7) |
| sub-total | 150993(67.9) | 151537(68.4) | 64787(64.6) | 62837(63.3) | 86384(70.7) | 88700(73.3) |
| Cropland | 59464(26.7) | - | 30187(30.1) | - | 29358(24.0) | - |
| Grass | 3800(1.7) | - | 292(0.3) | - | 3512(2.9) | - |
| Urban | 8221(3.7) | - | 4953(5.0) | - | 3005(2.4) | - |
| sub-total | 71485(32.1) | 70163(31.6) | 35432(35.4) | 36463(36.7) | 35875(29.3) | 33700(26.7) |
| Total | 222478(100) | 221700(100) | 100219(100) | 99300(100) | 122259(100) | 122400(100) |

* values within parentheses are percentile.

Korea are considered to have acceptable level of accuracy since the data have been acquired from the official forest inventory using aerial photographs and extensive field survey. The size of coniferous forest classified from the AVHRR data was almost the same as the existing inventory data although there was some misclassification to discriminate between deciduous and mixed forests. For the North Korea side, the area of coniferous forests from the classification is underestimated while the mixed forest shows overestimation. This misclassification may be resulted from the fact that there are substantial areas of deciduous conifer forest occupied by Korean larch(*Larix gmelinii var. principis-rupprechtii*). Considering the annual crown development of larch trees, the misclassification between coniferous and mixed forests seems reasonable.

Although the total area comparison is often used as a rough evaluation on the performance of computer classification, this method does not provide the source of classification error. Even though the size of coniferous forest is the same between the classification result and the reference data, some pixels of truly deciduous forest may be misclassified into coniferous forest and about the same number of truly coniferous pixels may be misclassified into deciduous forest. In such case, the total area comparison does not reveal the errors related to omission and commission. For this reason, it is necessary to conduct more comprehensive evaluation on the classification accuracy by collecting reliable ground truth data for each cover type.

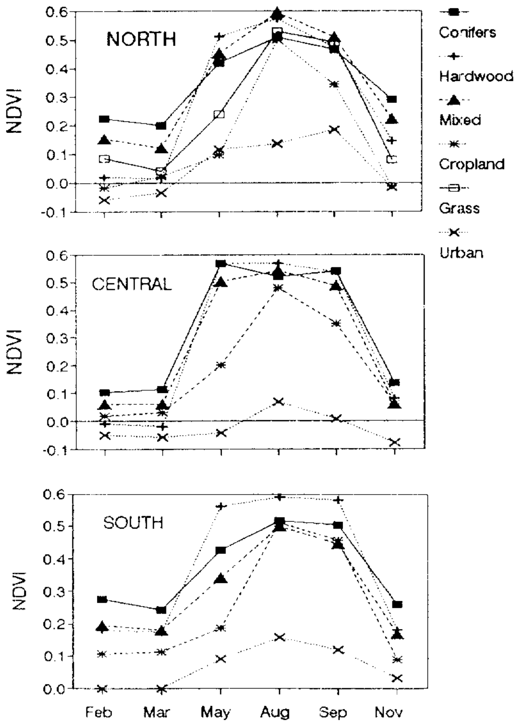


**Fig. 3.** Temporal profiles of the NDVI values from the six primary cover types, which were obtained at three different locations along the Korean Peninsula.

Figure 3 shows the temporal profile of NDVI values obtained at three different sites from the north to the south. The overall patterns of temporal profiles for the six cover type classes are similar at the three sites. As expected, the NDVI profile of coniferous forest has a maximum on the months of

February, March, and November when no green foliage remains for other plants. The mixed forest having substantial number of conifer trees has second highest NDVI values during the leaf-off season. Hardwood forest is distinguished by the high NDVI values from May to September. Two herbaceous classes of crop and grass have relatively low NDVI values during growing season. As can be seen from the figure 2, most grass lands are in high elevation areas of the North Korea and have slightly higher NDVI value than the cropland. The NDVI profiles of urban area having very minimal vegetation show the lowest NDVI values throughout all season. Temporal profile of the NDVI for a vegetation class might vary by several factors of geographic location, species composition, stand density, and tree size. From the figure 3, absolute magnitudes of the NDVI for a class are different each other among the three locations. This may suggest that the current vegetation cover classes used for this study can be further divided into more detailed classes of certain species groups.

## CONCLUSIONS

Vegetation indices derived from the two spectral bands of red and near infrared in AVHRR data have been effectively used to study vegetation features over very large geographic area. Due to frequent coverage of AVHRR data over the same area, the vegetation index could be obtained with regular time interval. Once the multiple sets of vegetation index images were overlaid on the same geographic basis, they can be used to produce a map showing the distribution of vegetation covers as well as to monitor temporal changes of vegetation pattern. The six monthly composites of the normalized difference vegetation index (NDVI) data used for this study were effective to classify the six primary cover type classes in the Korean Peninsula. From the quantitative and qualitative evaluations on the performance of computer classification, the resultant cover type map was useful to assess the spatial distribution of primary vegetation cover in the Korean Peninsular. Further studies are needed to verify the classification accuracy with more quantitative methods and to separate more detailed vegetation classes that can be useful for both resource management and environmental applications.

## ACKNOWLEDGMENT

The author would like to thank the Department of Oceanography at Seoul National University and the Oceanography Division at the National Fisheries Research and Development Agency in Korea for helping him to obtain the AVHRR data from their archived data tapes.

## LITERATURE CITED

Derrien, M., B. Farki, H. Legleau and A. Sairouni, 1992. Vegetation cover mapping over France using NOAA-11/AVHRR. *International Journal of Remote Sensing*, 13(9) : 1787-1795.

Eidenshink, J.C., 1992. The 1990 conterminous U.S. AVHRR data set. *Photogrammetric Engineering and Remote Sensing*, 58(6) : 809-813.

Forestry Administration, 1992. *Statistical Yearbook of Forestry*, Forestry Administration, Seoul, Korea.

Holben, B.N., 1986. Characteristics of maximum-value composite image from temporal AVHRR data. *International Journal of Remote Sensing*, 7(11) : 1417-1434.

Lee, K.S., 1992. A Cartesian coordinates system to cover Korean Peninsula as a single coordinate zone. *Journal of the Korean Society of Remote Sensing*, 8(2) : 93-104.

Loveland, T.R., J.W. Merchant, D.O. Ohlen, and J.F. Brown. 1991, Development of a land-cover characteristics database for the conterminous U.S. *Photogrammetric Engineering and Remote Sensing*, 57(11) : 1453-1463.

Mueller, F.B., 1987. Forestry in North Korea. *Sozialisticsche Forstwirtschaft*, March 1987. -translated by Y. Ha to Korean and published on *Wood and Culture*, Feb. 1993, pp.42-49.

Tou, J. and R. Gonzales, 1974. *Pattern Recognition Principles*, Addition-Wesley, Reading, Massachusetts, USA.

Townshend J.R.G. and C.O. Justice, 1986. Analysis of the dynamics of African vegetation using the normalized difference vegetation index. *International Journal of Remote Sensing*, 7(11) : 1435-1445.

Zhu, Z. and D.L. Evans, 1992. Mapping midsouth forest distributions. *Journal of Forestry*, 90(12) : 27-30.