Paper presented at the 22nd Asian Conference on Remote Sensing, 5 - 9 November 2001, Singapore.

APPLICATION OF LINEAR MIXTURE MODEL TO TIME SERIES AVHRR NDVI DATA

Lin ZHU and Ryutaro TATEISHI
Center for Environmental Remote Sensing (CEReS)
Chiba University
1-33 Yayoi-cho, Inage-ku, Chiba, 263-8522
Tel: (81)-43-290-3850 Fax: (81)-43-290-3857
E-mail: zhulin@ceres.cr.chiba-u.ac.jp
JAPAN

KEY WORDS: Linear Mixture Model, Monthly Composite NDVI

ABSTRACT: This research explores the possibility of using linear mixture model to generate fraction images from 1km time series NOAA AVHRR monthly composite NDVI data. The percentages of forest, grassland and farmland in the pixels are determined by applying the Constrained Least Squares (CLS) method over the study area in the northeast region of China. The validation of the model for AVHRR monthly composite NDVI data is performed by comparing the resulting fraction images with the classification results derived from coincident multi-temporal Landsat TM data and NOAA AVHRR monthly composite NDVI data using conventional methods. The results show that linear unmixing techniques, with improved estimates of endmember values and more sophisticated methods to fully utilize the included information, can have great potential when applied to coarse spatial time series AVHRR NDVI data for global studies.

1. INTRODUCTION

Study on geosphere-biophere-atmosphere interaction and analysis of global climate change depend on reliable and unambiguous definition of the existing terrestrial vegetation. Vegetation characteristics, including land cover and phenology, affect processes such as water cycling, absorption and reemission of solar radiation, momentum transfer, carbon cycling, and latent and sensible heat fluxes (Maselli et al., 1998). Variations in the composition and distribution of vegetation represent one of the main sources of systematic change on local, regional, or global scale, and the ability to detect these variations using remotely sensed data is of utmost importance for both environmental researches and management activities (Hall et al., 1995). Traditionally, vegetation monitoring by remotely sensed data has been carried out using vegetation indices, which are mainly derived from mathematical transformations of reflectance data of red (R) and near-infrared (NIR) channels. One of the most widely used indices is the well-known normalized difference vegetation index (NDVI).

Vegetation monitoring demands high temporal frequency information to follow the rapid vegetation phenological change. This can be realized by using satellite data such as the National Oceanic and Atmospheric Administration Advanced Very High Resolution Radiometer (NOAA AVHRR) data for its daily coverage and synoptic overview. NOAA AVHRR data have been widely utilized for global, continental and regional land use and land cover study. However, one major problem in using NOAA AVHRR data is its volume and noise. Fortunately, these can be solved through monthly maximum composite processing. The use of monthly composite data represents a compromise between temporal frequency and the need for cloud-free data (Moody and Strahler, 1994). It also provides a means to reduce data volume while maintaining vegetation phenological information. Actually, most vegetation mapping applications at broad spatial scales have been based on time series NDVI data (Tucker et al. 1985, Loveland et al. 1991), it is also indicated that global and regional land cover classification derived from remotely sensed data have to date generally relied on a single year time series to characterize phenology of the vegetation (Defries et al., 2000). Another major problem when using NOAA AVHRR data is its poor spatial resolution. Usually, the large ground area of each AVHRR pixel is a mixture of several land cover types. Conventional classification approaches that categorize each pixel into a discrete vegetation type do not fully utilize the rich information content of the data to describe gradients and mosaics in the landscape, and also variation in vegetation characteristics within a cover type is obscured (Defries et al., 2000). Efforts to address the problem of mixed pixels are of increasing importance as emphasis is being placed in providing global-scale monitoring.

A number of researches have been taken to depict subpixel heterogeneity in land cover from remotely sensed data. Linear mixture modeling to deconvolve proportional cover based on spectral reflectance of endmembers or pure pixels is a main one. In general, linear mixture techniques have been mainly applied at local or regional scales over limited areas using single date or multi-temporal spectral reflectance data. Clearly, there is a need for operational methods permitting the application of linear mixture model directly at NDVI level (i.e., the mixed NDVI is the sum of component NDVI weighted by the component surface proportions), especially for global and regional studies. Only a few researches have addressed this issue in view of improving crop monitoring and yield prediction through estimation of pure crop NDVI profiles (Kerdiles and Grondona, 1995), and almost no research has addressed application of linear mixture model for NOAA AVHRR monthly composite NDVI data, neither at local scales, nor regional and global scales. Therefore, a research in this area should be meaningful challenge.

In this study, the linear mixture model is applied to time series 1km AVHRR monthly composite NDVI data ranging from April to November of 1992 for generating fraction images. The percentages of forest, grassland and farmland in the pixels are determined by applying the Constrained Least Squares (CLS) method (Shimabukuro and Smith, 1991) over a typical agricultural area in the northeast region of China. Multi-temporal Landsat TM data and NOAA AVHRR monthly composite NDVI data are classified using maximum likelihood and minimum distance classification methods, respectively. The validation of linear mixture model for AVHRR monthly composite NDVI data is performed by comparing the resulting fraction images with the classification results derived from TM and AVHRR data using conventional methods.

2. LINEAR MIXTURE MODEL

Linear mixture model assumes that each image pixel contains information about the proportion and the spectral response of each component within the ground resolution unit, and the response of each pixel is taken as linear combination of the responses of each component in the mixed target. Actually this assumption should be true only for original channels. However, recent studies have demonstrated that a linear combination of NDVI values implies only very minor inaccuracies (Kerdiles and Grondona, 1995). In this study, the linear mixture model can be formulated as:

$$tNDVI_i = \sum pNDVI_{ij} f_j + e_i \tag{1}$$

where $tNDVI_i$ is the time series monthly composite NDVI value for a pixel in period band i, $pNDVI_{ij}$ is the pure monthly composite NDVI value of component j in period band i, f_j is the proportion of component j in the pixel, and e_i is the error term for period band i.

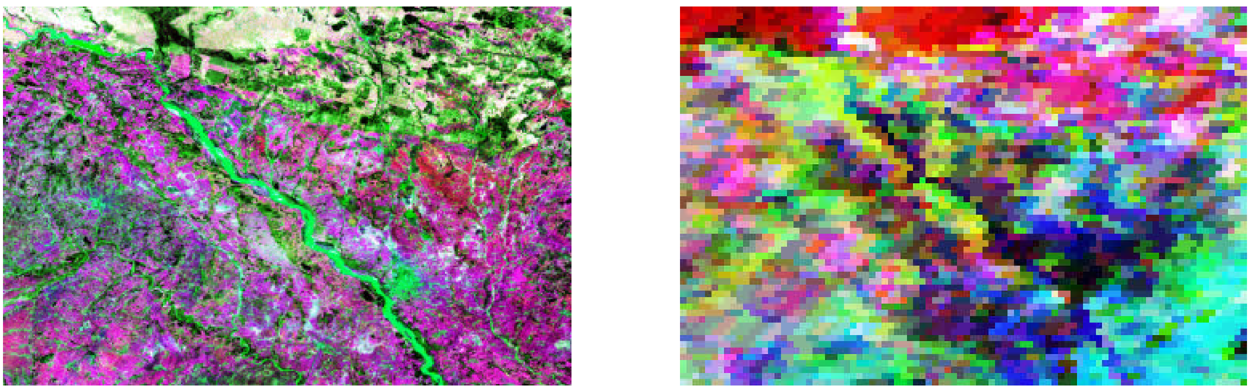
Subjected to the constrains

$$1 \geq f_j \geq 0 \tag{2}$$

and

$$\sum f_j = 1 \tag{3}$$

since the sum of the proportions for any resolution element must be one and the proportion values must be nonnegative. The Constrained Least Squares method is applied to estimate the proportion of each component in a pixel by minimizing the sum of squares of the errors.



(a) (b)

Figure 1. Color composite images of the study area. (a) Landsat TM image (band 5 of May=red, band 3 of September=green, and band 3 of May=blue). (b) NOAA AVHRR monthly composite NDVI image (April=red, July=green, and October=blue).

3. EXPERIMENT

3.1 Study Area

The study area, 90 x 60 km, with geographical coordinates from $42^{0}15'$N to $42^{0}45'$N in latitude and from $122^{0}0'$E to $122^{0}45'$E in longitude, is located in a typical agricultural region of Liaoning province in the northeast part of China.

The main cover types are forests, pastures and crops. The most common crops are corn and paddy, as well as s oybean, wheat, etc. Forests consist predomiantly of pine and poplar. The main crop season is from April to October, including the driest month July. The information collected by the ground truth survey is useful for the classification of multi-temporal Landsat TM and NOAA AVHRR monthly composite NDVI data and the application of linear mixture model.

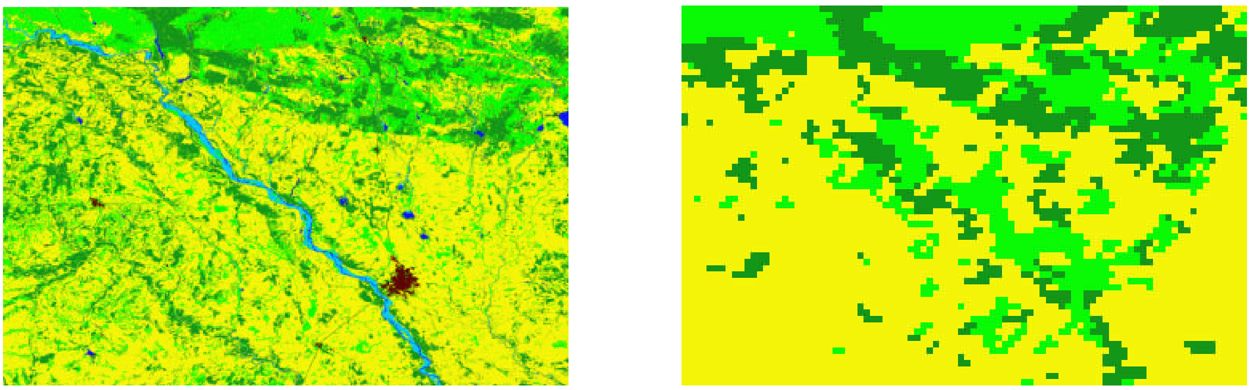
3.2 Data

Two Landsat TM images are used as the basis of endmember collection in the linear mixture model, and in validation of the resulting fraction images. The multitemporal Landsat TM images collected on 19 May 1994 and 8 September 1994 are selected mainly because they are the most cloud-free scenes in the crop season. A scene of 2700 x 1800 pixels is extracted from each of both Landsat TM images and georeferenced by nearest neighbor resampling algorithm, with resampling size of 25m pixel yielding RMS error of less than 1 pixel. All channels of TM data are used except the thermal channel 6.

The NOAA AVHRR data, already in form of NDVI, are produced by the Global Land 1km AVHRR Dataset Program at USGS EROS Data Center. The NOAA AVHRR 1km 10-day composite NDVI data spanning April 1992 through March 1993 is based on the Interrupted Goode Homolosine map projection. In order to march different purposes, it is transformed to latitude/longitude projection (Plate Carree Projection) with 30-seconds resolution. Then, the dataset is recomposed to monthly composite NDVI dataset based on maximum value compositing method, in which the NDVI is examined pixel by pixel during the compositing process to determine the maximum value. The retention of the highest NDVI value reduces the number of cloud-contaminated pixels and selects the pixels nearest to nadir (Holben, 1986). The monthly composite NDVI data of the study area, image size of 90 lines by 60 samples, is cut from this dataset. In this study, only 8 months time series monthly composite NDVI data ranging from April to November 1992 are used in order to include full growing season and reduce the effect of snow. Figure 1 shows the color composite images of multi-temporal Landsat TM and NOAA AVHRR monthly composite NDVI data.

3.3 Approach

Generation of a reference land use map is carried out by using a supervised maximum likelihood classification procedure to multi-temporal Landsat TM data. Eight land use classes are deemed sufficient to cover almost all the variability of the study area (Forest, Corn, Paddy, Grassland, bare land, Urban, Water and Seasonal water). The overall classification accuracy is estimated to exceed 85 percent (Zhu and Tateishi, 2000). Also, three land use types, namely forest, grassland and farmland, account for 97.4 percent of the total study area.



(a) (b)

Figure 2. Classified images derived from (a) multitemporal Landsat TM data. (b) NOAA AVHRR monthly composite NDVI data.

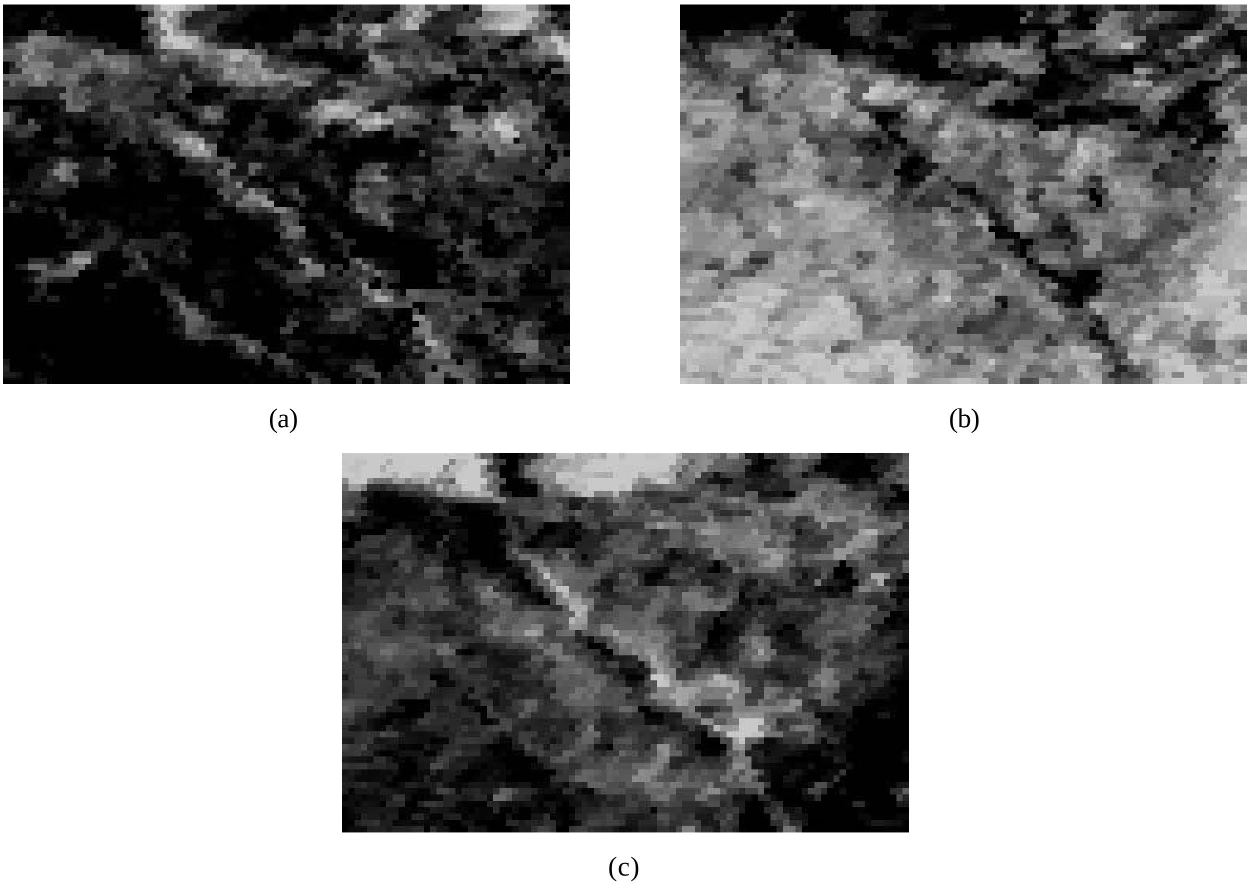
The 8 months AVHRR monthly composite NDVI data are also classified by using a conventional minimum distance classification procedure in order to provide a comparison with the fraction images resulting from linear mixture model.

For application of the linear mixture model, the classified Landsat TM image is filtered by using a 5 x 5 majority filter in order to reclassify isolated pixels belonging to one land use type to the same land use type as the majority of their neighbors in the image. Then, the filtered image is degraded to the spatial resolution of AVHRR data to make it easy for endmember collection. Assuming three components within a pixel, the CLS method is applied to the AVHRR monthly composite NDVI data. The derived fraction images are scaled to integer values from 0 to 200.

4. RESULTS AND DISCUSSION

Figure 2 shows the classified images derived from multitemporal Landsat TM and NOAA AVHRR monthly composite NDVI data using conventional classification methods. Figure 3 shows the faction images derived from the linear mixture model. Comparing to the TM classification result Figure 2(a), in general, there is a good agreement both for the results derived by using minimum distance classification and the linear mixture modeling. Actually, we can discover visually that the fraction images are more detailed, and can reflect the land cover reality more well than the result derived from the conventional method. The advantage of the fraction images is that they contain physical information, i.e., amount of each component within a pixel. Note the lower-left part of Figure 2(a) and 3(c), there is some grassland, but not revealed in Figure 2(b).

Based on a simple check on the land cover area estimates of the study area derived by using the above-mentioned methods, as shown in Table 1, the linear mixture model can identify the forest well. The area percentage estimation of grassland seems not so good. This can partly be attributed to the influence of other land cover types such as urban, water and seasonal water, which are not included in the mixture components. Moreover, the endmember collection has a significant effect on the final fraction images. Although the results derived from the conventional method are acceptable, the limitation of these methods is obvious and inevitable.



(a) (b) (c)

Figure 3. Fraction images derived from NOAA AVHRR monthly composite NDVI data. (a) forest, (b) farmland, and (c) grassland.

Table 1. Comparison of the land cover area estimates

		Percentage area under land cover type		
Satellite data	**Classification method**	**Forest**	**Farmland**	**Grassland**
Landsat TM	Maximum likelihood	20.97	56.59	19.86
AVHRR NDVI	Linear mixture	18.05	51.88	30.07
AVHRR NDVI	Minimum distance	17.53	62.06	20.41

5. CONCLUSION

This study confirmed the potential of the linear mixture model applied to time series NOAA AVHRR monthly composite NDVI data for the estimation of ground cover proportions at local scale. The mixed pixel problem is frequently encountered when using coarse spatial remotely sensed data. The possibility of deriving area percentage estimation directly from NOAA AVHRR NDVI data will undoubtedly permit a more precise vegetation classification and monitoring.

Despite many error sources that are partly due to NOAA AVHRR characteristics, the results obtained for the main land cover types of the study area are relatively acceptable comparing with multitemporal Landsat TM estimates. Inclusion of other land cover components would provide additional information and possibly more accurate results. Improvement of endmember estimation and development of more sophisticated methods to fully utilize the included information should be also addressed. Additionally, further quantitative assessment of the pixel proportions is required to fully interpret the results from mixture models. The research has shown that linear mixture techniques can have great potential when applied to coarse spatial time series NOAA AVHRR NDVI data for global studies.

References

Defries, R. S., Hansen, M. C., and Townshend, J. R. G., 2000. Global continuous fields of vegetation characteristics: a linear mixture model applied to multi-year 8 km AVHRR data. International Journal of Remote Sensing, 21(6&7), pp. 1389-1414.

Hall, F. G., Townshend, J. R., and Engman, E. T., 1995. Status of remote sensing algorithms for estimation of land surface state parameters. Remote Sensing of Environment, 51(1), pp. 138-156.

Holben, B.N., and Shimabukuro, Y. E., 1993. Linear mixing model applied to coarse spatial resolution data from multispectral satellite sensors. International Journal of Remote Sensing, 14(11), pp. 2231-2240.

Kerdiles, H., and Grondona, M. O., 1995. NOAA-AVHRR NDVI decomposition and subpixel classification using linear mixing in the Argentinean Pampa. International Journal of Remote Sensing, 16(7), pp. 1303-1325.

Loverland, D. T. R., Merchant, J. W., Ohlean, D. O., and Brown, J., 1991. Development of a land cover database for conterminous U. S. Photogrammetric Engineering and Remote Sensing, 57, pp. 1453-1463.

Maselli, F., Gilabert, M. A., and Conese, C., 1998. Integration of high and low resolution NDVI data for monitoring vegetation in Mediterranean environments. Remote Sensing of Environment, 63, pp. 208-218.

Moody, A., and Strahler, A. H., 1994. Characteristics of composited AVHRR data and problems in their classification. International Journal of Remote Sensing, 15(17), pp. 3473-3491.

Shimabukuro, Y. E., and Smith, J. A., 1991. The least squares mixing models to generate fraction images derived from remote sensing multispectral data. I.E.E.E. Transactions on Geoscience and Remote Sensing, 29(1), pp. 16-20.

Tucker, C. J., Townshend, J. R., and Goff, T. E., 1985. African land cover classification using satellite data. Science, 227, pp. 369-375.

Zhu, L., and Tateishi, R., 2000. Mapping of agricultural area using multitemporal remote sensing images. Journal of the Japan Society of Photogrammetry and Remote Sensing, 39, pp. 21-32.