PAPER • **OPEN ACCESS**

Linear Spectral Mixture Analysis Land Cover for Assessment Level Subpixel: A Case Study of Tasikmalaya City Area Based on Landsat Imagery

To cite this article: R Ridwana *et al* 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **286** 012042

IOP Conf. Series: Earth and Environmental Science **286** (2019) 012042 doi:10.1088/1755-1315/286/1/012042

Linear Spectral Mixture Analysis Land Cover for Assessment Level Subpixel: A Case Study of Tasikmalaya City Area Based on Landsat Imagery

R Ridwana[1*], P Danoedoro[2], S Herumurti[2], S Himayah[1], M Ihsan[1], R Arrasyid[1], and F Urfan[1]

[1]Department of Geography Education, Universitas Pendidikan Indonesia, Jl. Dr. Setiabudi 229, Bandung 40154, Indonesia

[2]Cartography and Remote Sensing, Geography Faculty, Gadjah Mada University, Yogyakarta 55281, Indonesia

[*]rikiridwana@upi.edu

Abstract. Land cover in urban areas can be detected through surveys or using high-resolution imagery with a better accuracy. Especially if the need related to land cover information regionally in the period of the nineties that require the availability of data and unavailability of high resolution images. Therefore, images with intermediate spatial resolution are still required. However, the use of medium-resolution images such as Landsat is constrained by the presence of mixed pixels that cause land cover in urban areas to vary. The mixed pixel will be the source of error in the multispectral classification process, so it takes analysis up to the subpixel level. The need for information up to the subpixel level for ground cover detection can be obtained through the Linear Spectral Mixture Analysis method, where one pixel in the Landsat image in this study will be separated into four endmember, ie vegetation, impervious surface, bare soil, and water. These four endmembers are assumed to represent linear combinations of land coverings contained in urban areas in the form of proportions in each pixel. The results show that the endmember can be well separated, whereas the RMS error of 1994th is 0.013 with an accuracy of 94.44%.

1. Introduction

Land cover changes in urban areas is a process that is not simple, the process of change among which is not free from the influence of physical factors and human factors. On the one side, the rapid expansion of cities is usually closely related to socioeconomic factors. On the other side, the urbanization process has a significant impact on the economy of the people in the area. In addition, states that changes in land cover are the main components of global change that can have a greater impact than climate change [1].

Spectral characteristics of urban land cover are more difficult to identify than rural areas. As a result of significant differences in physical processes and local climate variations in urban areas [2,3]. Several indicators are commonly used to determine the characteristics of the urban environment including vegetation density and cover, percent of built land surface area, sky-view factor, and structure and composition of buildings [4]. The significant difference between spectral reflection from the surface of built land, soil and rocks, but these differences will be difficult to recognize with the limitations of spatial resolution from intermediate resolution sensors such as Landsat TM or ETM+.

Mixed pixels in remote sensing data is one source of errors in accuracy assessment results of conventional classification algorithms such as Minimum Distance Algorithm for Average, Parallelepiped Algorithm, Maximum Possible Algorithm, and Nearest Neighbor Algorithm can only classify one pixel into one class only. The information contained in the subpixels on the spectral mixture of different land cover cannot be obtained from the classification algorithm. Therefore the conventional classification of mixed pixels can eliminate information, decrease the accuracy of classification results, and reduce the quality of modeling in several applications [5].

Overcoming the problem of the above-mentioned mixed pixels, there are several methods that can be used, one of them with the Linear Spectral Mixture Analysis method. This method is an approach with subpixel analysis that is able to extract the available fraction information in a single pixel, so that it can be an alternative solution to classify one pixel especially when applied to medium spatial resolution images in this case Landsat and heterogeneous urban areas. Although the application of this method has been done quite a lot in developed countries, but in Indonesia the application has not been widely implemented. This study will examine the accuracy level of Linear Spectral Mixture method in detecting land cover in urban areas in Tasikmalaya City and surrounding areas from Landsat imagery.

2. Methods

This resaerch uses a multitemporal Landsat Image digital analysis approach to detect land cover changes in urban areas at the subpixel level by using Linear Spectral Mixture method. An ideal change detection procedure should use data from the same sensor system, both in terms of spatial resolution, viewing geometry, number of channels, radiometric resistance, and recording time[6,7]. Various image processing techniques are applied to obtain better results, such as atmospheric correction, Minimum Noise Fraction, and Pixel Purity Index. In addition, detection of land cover changes is also done by using Change Detection analysis that will provide information on changes in fractions of each endmember (vegetation, impervious surface, bare soil, and water) at the subpixel level.

This research is located in the City of Tasikmalaya and Surroundings. The city of Tasikmalaya is geographic located at 108°08'38’’ - 108°24’02" East Longitude and 7°10' - 7°26’32" South Latitude with an area of 18,385 Ha (183, 85 Km^2). The data dimension of this study area is determined at 700 x 700 pixels where 1 pixel represents (30 x 30) square meters, so the total area of research is 44,100 hectares (441 Km^2), covering the City of Tasikmalaya and surrounding areas which are expected to represent the land cover fraction, vegetation, impervious surface, bare soil, and water. The research location can be seen in Figure 1 below.

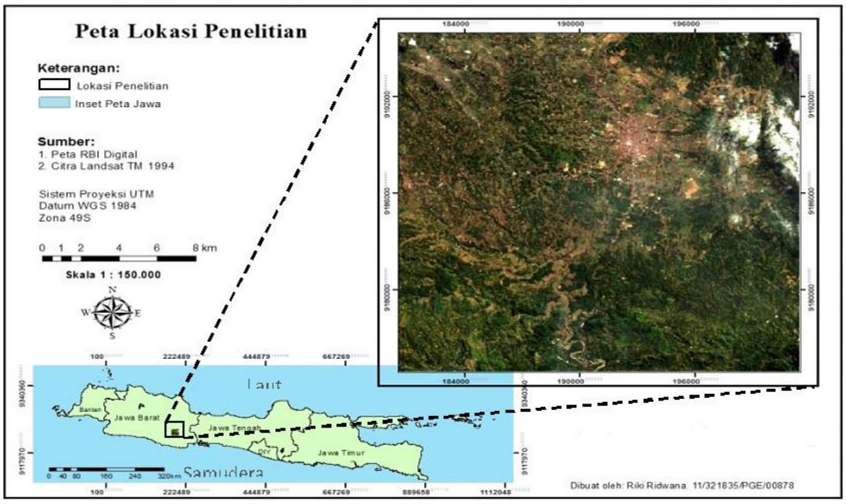


Figure 1. The research location

3. Results And Discussion

3.1 Atmospheric Correction

The atmospheric correction used in this study utilizes the DOS (Dark Object Substraction) facility available in Selva's IDRISI software. Atmospheric correction is carried out to convert digital landsat data to surface reflectance values as a prerequisite for the soft classification process of Linear Spectral Mixture Analysis. This Dark Object Substraction model has the assumption that the effect of the atmosphere is relatively uniform on the reflected energy value of the visible channel. One way to reduce atmospheric effects on images is to use reflected values that are close to zero in an area, such as dark cloud shadows and clear deep seas.

Processing using a DOS model requires the same estimation of digital values (Dn haze). The location of the Dn haze score collection on the scene of this research image was taken on the clear deep sea. The Dn Haze values of each channel are shown in Table 1. The results of atmospheric correction using the DOS model are shown in Table 2. While the image of the atmospheric correction using this model is shown in Figure 2 and Figure 4. The table shows that atmospheric correction uses the Dark Object model. Substraction (DOS), produces a reflectance value that corresponds to the theory of the range 0.0 - 1.0. Although visually the image that has changed its pixel value to reflectance value does not show differences (Figures 2 and 4), but the difference is very noticeable from the histogram shown in Figures 3 and 5.

Table 1. The Dn Haze value for each channel uses the Dark Object Substraction (DOS) correction model

Band	Landsat TM 1994			Landsat ETM+ 2003		
	Digital Values		*Dn Haze*	Digital Values		*Dn Haze*
	Mean	Std.Deviasi	Mean – 2.(Std.Deviasi)	Mean	Std.Deviasi	Mean – 2.(Std.Deviasi)
1	54,87	1.12	52,63	78,06	1,57	74,92
2	17	0	17	43,66	1,29	41,08
3	12,62	0,74	11,14	32,73	2,05	28,63
4	7,32	0,51	6,35	13,53	0,91	11,71
5	5,50	0,75	4	15,46	1,59	12,28
7	3,87	0,83	2,21	14	1,60	10,8

Table 2. Minimum and Maximum Reflectant Values Each Channel of Atmospheric Correction Results Using DOS Correction Model

Band	Landsat TM 1994				Landsat ETM+ 2003			
	Digital Values		Reflectance Values		Digital Values		Reflectance Values	
	Min	Max	Min	Max	Min	Max	Min	Max
1	0	255	0.0	1.0	0	255	0.0	1.0
2	0	255	0.0	1.0	0	255	0.0	1.0
3	0	255	0.0	1.0	0	255	0.0	1.0
4	0	255	0.0	1.0	0	255	0.0	1.0
5	0	255	0.0	1.0	0	255	0.0	1.0
7	0	255	0.0	1.0	0	255	0.0	1.0

Figure 2. Composite 1994 Landsat TM Image 321 Before Atmospheric Correction (Left) and After Atmospheric Correction (Right)

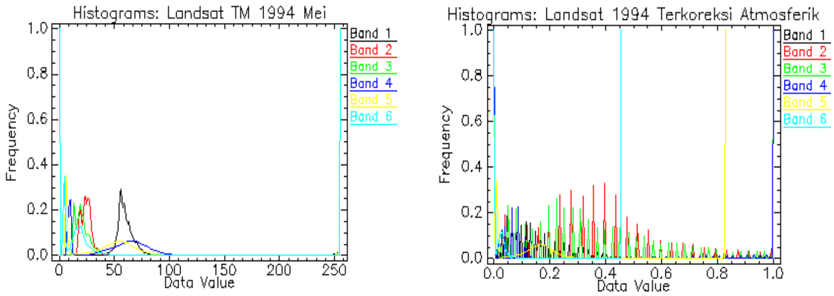


Figure 3. Histogram of Reflectance Value on TM Landsat Image 1994 Before Atmospheric Correction (Left) and After Atmospheric Correction (Right)

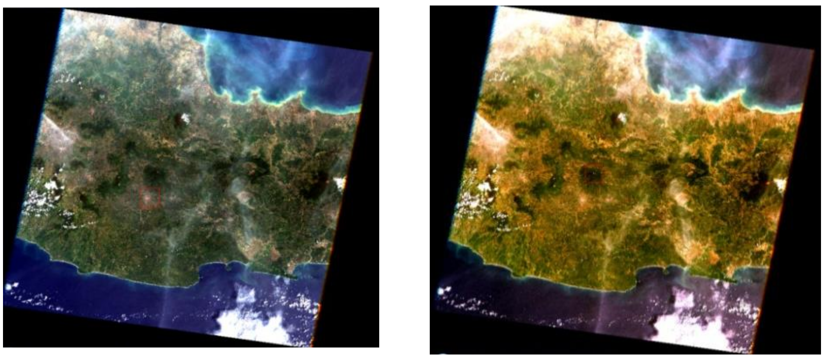


Figure 4. Landsat ETM+ Year 2003 Composite 321 Before Atmospheric Correction (Left) and After Atmospheric Correction (Right)

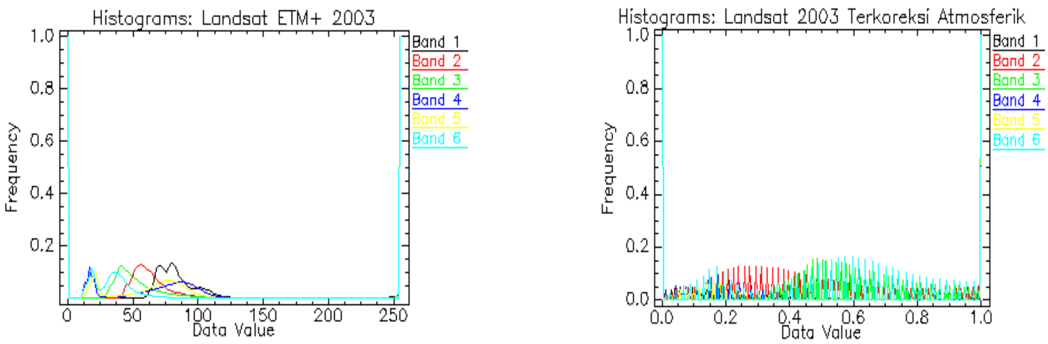


Figure 5. Histogram of Reflectance Value on Landsat ETM+ Image of 2003 Before Atmospheric Correction (Left) and After Atmospheric Correction (Right)

IOP Conf. Series: Earth and Environmental Science **286** (2019) 012042 doi:10.1088/1755-1315/286/1/012042

3.2 Geometric Corrections
Landsat TM 1994 and ETM+ 2003 imagery are 1G level, that is they have experienced systematic geometric corrections. Thus errors caused by sensors can be overcome properly. But the image that has been systematically corrected still needs to be done again by geometric corrections that are non-systematic, such as the shifting of images caused by the earth's rotational motion or variations in sensor heights. Geometric correction in the study was carried out by image to image registration between 1994 Landsat TM images and 2003 ETM+ corrected using 20 GCP (Ground Control Points). Geometric correction at this stage produces an RMS Error of 0.21 which means that the image has a shift of 6.3 meters (0.21 x spatial resolution of the image) to the reality in the field.

3.3 Minimum Noise Fraction
MNF transformations were applied to Landsat TM images in 1994, information was concentrated on the initial MNF optimally from all channels used. Can be seen in figure 6. MNF one has the highest eigenvalue of 17.6. Meanwhile the eigenvalues experience a decrease in the quality of information at MNF 5 with an eigen value of 3.0. Like with MNF 6 the quality of information is very low because noise is concentrated on the component which has the lowest eigenvalue of 1.6. Visually the results of MNF Landsat TM in 1994 can be seen in Figure 8.

MNF transformations applied to Landsat ETM + 2003 images concentration of information are at the initial MNF. This is indicated by the eigenvalue as shown in Figure 7, where the highest MNF of eigenvalues is 10.2. Then the next MNF experiences a decrease in eigenvalues, which means information is decreasing. Proven by the lowest eigenvalue there is at the last MNF that is 2.2. Visually the results of MNF transformation in Landsat ETM + 2003 images can be seen in Figure 9.

MNF	Eigenvalue	Percent
1	17.5917	43.52%
2	9.1977	66.28%
3	5.3887	79.61%
4	3.6658	88.68%
5	3.0196	96.15%
6	1.5559	100.00%

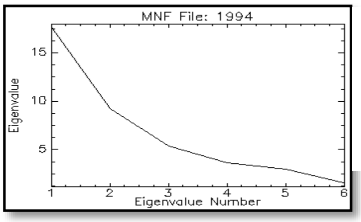


Figure 6.
MNF Result Eigenvector Values (left) and 1994 1994 MNF Citra Eigenvector Value Plots (right)

MNF	Eigenvalue	Percent
1	10.1559	33.76%
2	6.2453	54.52%
3	4.1516	68.33%
4	3.9902	81.59%
5	3.2981	92.56%
6	2.2390	100.00%

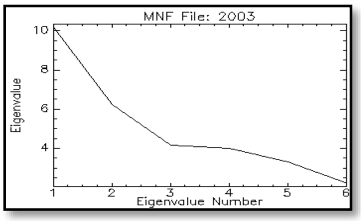


Figure 7.
MNF Result Eigenvector Values (left) and 2003 MNF Image Eigenvector Value Plots (right)

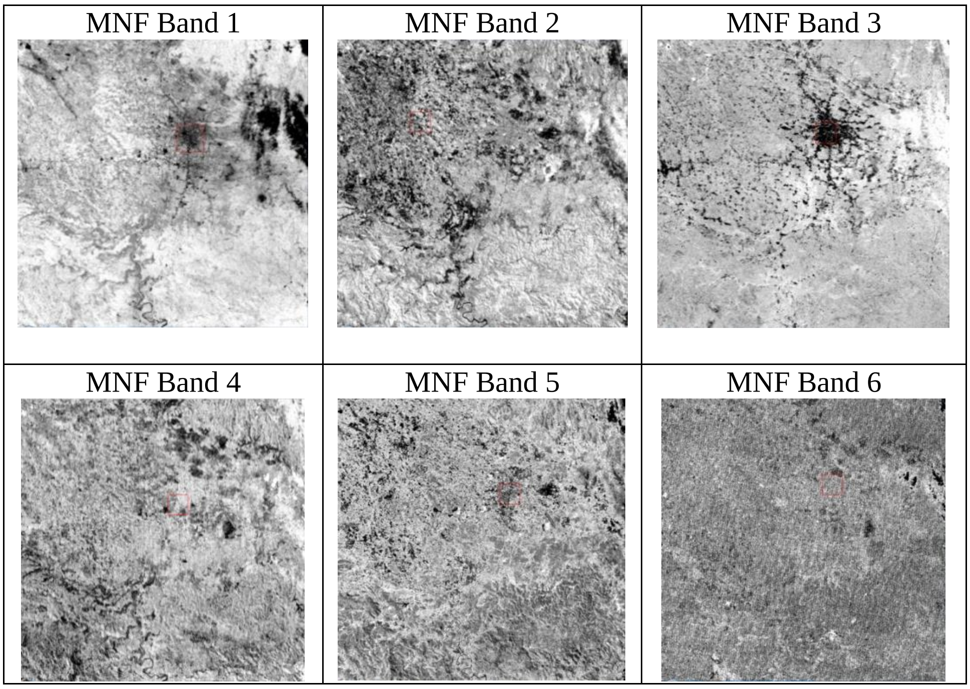


Figure 8. Information Quality Degradation Result of MNF Transformation on Landsat TM 1994 Image

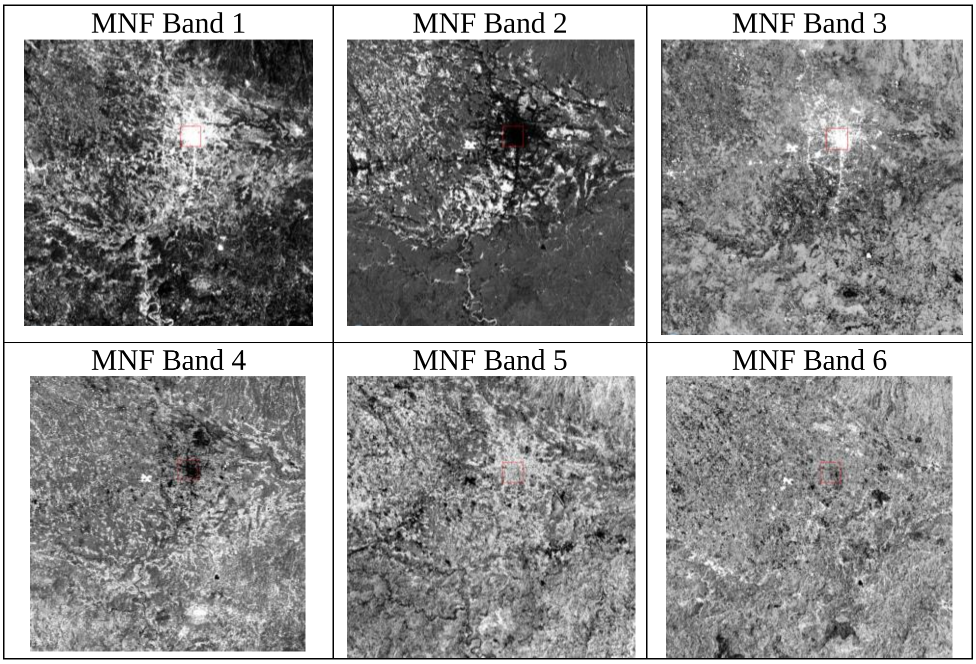


Figure 9. Information Quality Degradation of MNF Transformation Results on Landsat ETM + 2003 Imagery

3.4 Pixel Purity Index

Pixel Purity Index (PPI) is used to find the purest pixel among all mixed pixels in an image [10]. The PPI process uses MNF image input that is free of noise. The number of iterations in this study was determined as 104. The iteration was chosen referring to previous studies which concluded that the number of iterations that gave maximum results was 104 and 105 [10]. The PPI process also uses thresholds. Referring to Mitchell's (2007) study the threshold used is that the MNF input is 2. The PPI process results can be seen in Figures 10 and 11, showing the total number of extreme pixels found in the PPI process which is a function of the number of iterations used . Image in 1994 the total number of pure pixels found was $3x10^5$ pixels, while for 2003 the total number of pure pixels found was greater that was $4x10^5$ pixels.

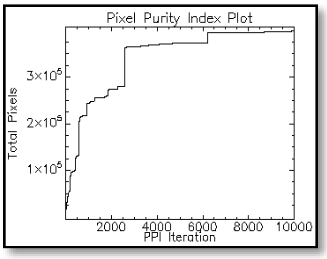
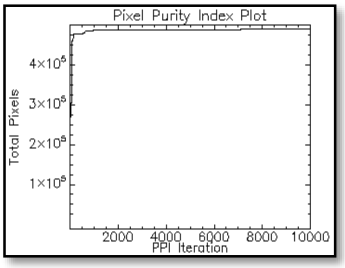


Figure 10. (top left) Plots of Pixel Purity Index with Iteration Amount104 (bottom left) Landsat ETM + 2003 imagery PPI results with 104 Bright Pixel Iterations (White) Are Pure Pixels



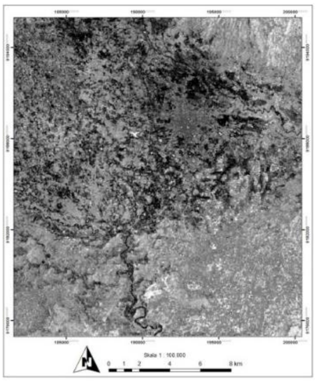


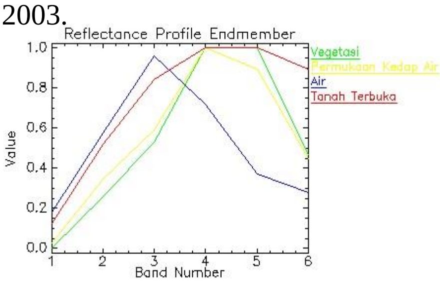
Figure 11. (top left) Plot of Iteration Number of Pixel Purity Index104 (bottom left) TM 1994 Landsat Image PPI Results with 104 Bright Iterations (White) Are Pure Pixels

3.5 n-D Visualizer

n-D Visualizer designed to receive input region of interest (ROI) which contains the purest spectrum of pixels in one image, and it is possible to separate pure pixels into each endmember. Visualization results via n-D Visualizer can be plotted on the reflection curve (z profile) with Landsat TM / ETM + image reflectance input input to see the spectral reflection of each pixel on each endmember. After being plotted on a curve the reflection chart is then made a new class and averaged into one spectral reflection value for each endmember. This spectral reflection is used as input in the classification process using a linear spectral mixture method.

3.6 Endmember

The most decisive process in the Linear Spectral Mixture Analysis model is endmember. This is because endmember is an object or material that will be separated in the pixel image. The success of the Linear Spectral Mixture model is strongly influenced by this endmember determination. This research utilizes 6 channels, namely channels 1 through 5 plus channel 7. The number of endmember itself is determined by four types of endmember, namely vegetation, impervious surface, bare soil, and water. The following below is a picture that shows the average reflection results of the four endmember land cover in 1994 and 2003.



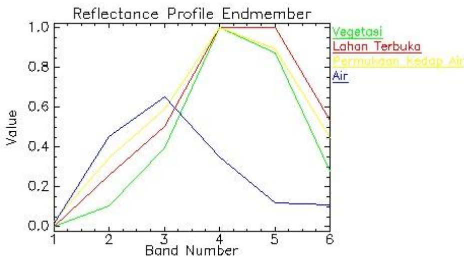


Figure 12. Endmember Reflection Curve Vegetation, impervious Surface, Bare Soil, and Water imagery 1994 (left) and 2003 (right)

3.7 Linear Spectral Mixture Analysis

Linear Spectral Mixture Analysis is used to overcome problems found in mixed pixels.Using the LSMA approach can develop fractions through unmixing the multispectral image based on selected endmembers [9, 10]. This approach separates linearly each endmember in a pixel based on the assumption that, a pixel consists of a linear combination of spectral reflectances of each endmember that is in the pixel. Processing Linear Spectral Mixture Analysis produces the fraction image of each endmember land cover (vegetation, watertight surface, open soil, and water) and accompanied by an RMS Error image. The fraction image shows the percentage of endmember in each pixel using the gray level. The higher the percentage of an endmember, then indicated by the hue that is getting closer to white, and vice versa. The results of the Linear Spectral Mixed Analysis process in the form of fractions from each endmember along with the RMS Error can be seen in Figures 13 to 17 for 1994, while for 2003 can be seen in Figures 18 to 22. The endmember fraction values that are input in the Mixed Analysis process Linear Spectral is in the range of 0-1 where the value is the percentage of an endmember in one pixel. In other words, a value of 0 means there is 0% of certain endmember in one pixel, which means that the endmember in question does not exist at all in the pixel. So is the case for endmember which is worth more than 1.

Table 3. Endmember Abundance Value LSMA Classification Results on the 1994 Landsat Image

Endmember	Minimum	Maximum	Average	Standard Deviation
Vegetaion	-5.715897	1.529173	0.582147	0.438726
Impervious surface	-0.925292	9.714310	0.346260	0.487136
Bare Soil	-4.965809	3.394641	-0.174264	0.450384
Water	-1.016518	2.882548	0.204753	0.276664
RMS Error	0.000022	0.144384	0.011910	0.009937

Table 4. Endmember Abundance Value LSMA Classification Results on the 2003 Landsat Image

Endmember	Minimum	Maxsimum	Average	Standard Deviation
Vegetaion	-0.314326	1.351876	0.687762	0.323244
Impervious surface	-0.331535	1.066640	0.313510	0.312493
Bare Soil	-0.825249	1.621522	-0.114500	0.237515
Water	-1.363634	2.090995	0.279264	0.282010
RMS Error	0,000000	0,397505	0,023684	0,017847

However, the statistical value shows both the image of the endmember fraction in 1994 (table 3.3) and the image of the endmember faction in 2003 (table 3.4) not in the range of values 0-1. land which not only includes input endmember which consists of vegetation, watertight surface, open land, and water, but there are objects or other land cover which are then classified into land cover adjusting to the input endmember.

After iterating several times in the input endmember, finally the results of the endmember fraction (vegetation, watertight surface, open soil, and water) were obtained with the lowest average RMS Error value for 1994 valued at 0.011 and for 2003 which was 0 , 02 (table 3.3 and table 3.4). The RMS error statistics result means the Linear Spectral Mixture Analysis method is able to classify the land cover properly. The fraction value in each endmember in one pixel if added together produces a value equal to one, this means meeting the boundary function of the spectral separation model linearly.

IOP Conf. Series: Earth and Environmental Science **286** (2019) 012042 doi:10.1088/1755-1315/286/1/012042

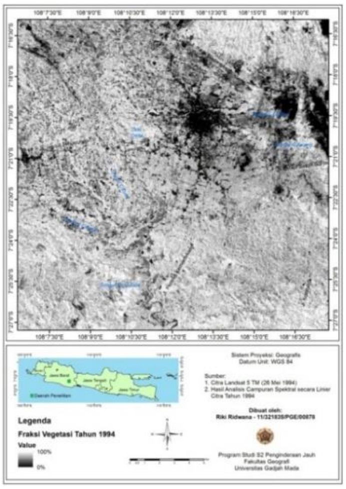


Figure 13. Vegetation Fraction Imagery in 1994

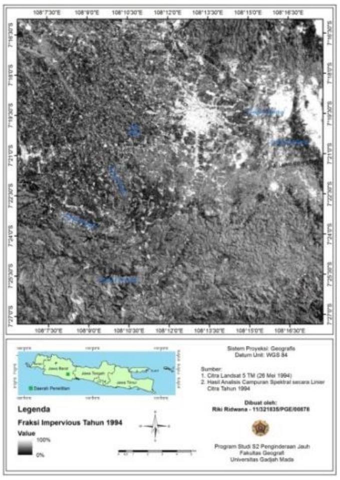


Figure 14. Impervious Surface Fraction Imagery in 1994

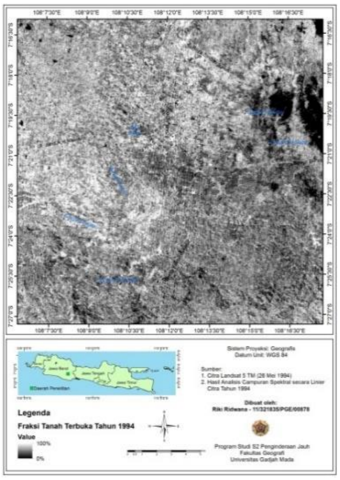


Figure 15. Bare Soil Fraction Imagery in 1994

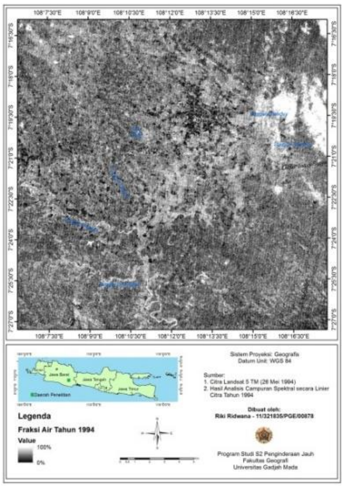


Figure 16. Water Fraction Image in 1994

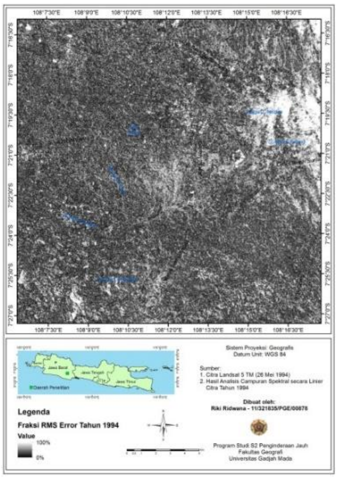


Figure 17. RMS Error Faction Image in 1994

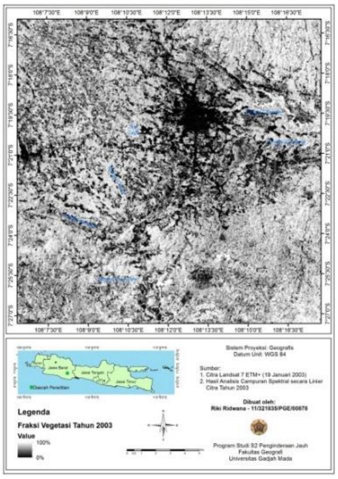


Figure 18. Vegetation Fraction Imagery in 2003

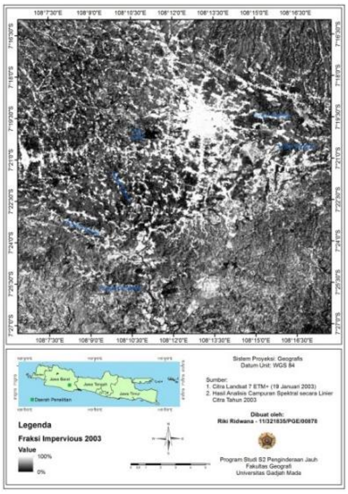


Figure 19. Impervious Surface Fraction Imagery in 2003

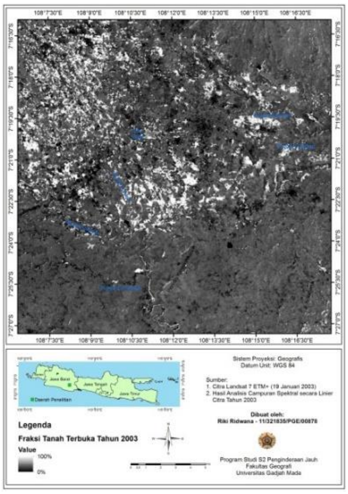


Figure 20. Bare Soil Fraction Imagery in 2003

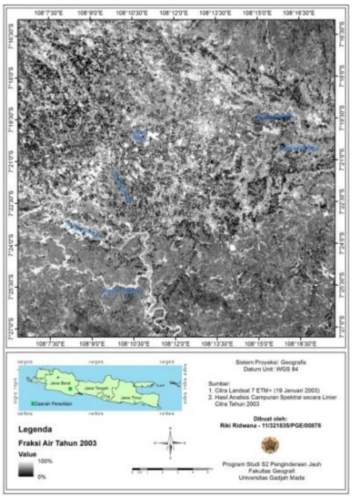


Figure 21. Water Fraction Image in 2003

Regardless of the reliability it has, it is necessary to know the weaknesses in this LSMA. The weakness in question is to be bound to the number of endmember which is input linearly, while in the field the existence of land cover is not only limited to vegetation, watertight surfaces, open soils, and water. So that the mixed pixels where the land cover is the most heterogeneous say urban area, has the highest error value. In addition to the processing of Linear Spectral Mixture that has been done, the reflected water values which have similarities to the open soil endmember are classified into bare soils (Figure 20), as well as cloud objects that have similar values to endmember impervious surfaces are classified into the cover impervious surface area (Figure 14). This is another weakness of LSMA which has actually been minimized through accurate endmember determination.

3.8 Accuracy Test
The process of testing the image accuracy of LSMA in 1994 on the measurement data of field samples and interviews of local residents, obtained an average RMS Error for the entire land cover of 94.44%. As for the process of testing the accuracy of image accuracy in Linear Spectral Analysis in 2003 on DigitalGlobe images with a spatial resolution of about 1.1, obtaining an average RMS Error for the entire land cover is 97.80% in other words this Linear Spectral Mixture method has a level high accuracy.

4. Conclusion
LSMA can be used as an alternative choice in detecting land cover up to subpixel level. Without using high-resolution imagery heterogeneous urban conditions can be well classified using Landsat imagery, because in one pixel Landsat covering 900 m^2 can be obtained quantitatively information on four types of objects at once (vegetation, impervious surfaces, bare soil and water). This fact is based on the level of accuracy of the LSMA method in this study capable of achieving average accuracy of each pixel by 94.44% for 1994 imagery (RMSE 0.013) and 97.80% for 2003 images (RMSE 0.021).

References
[1] Jensen, J. R., 2005, Introductory Digital Image Processing: A Remote Sensing Perspective. 3rdEdition, New Jersey, USA: Pearson Prentice Hall, Inc.
[2] Nurlina, A., 2008, Linear Spectral Mixture Analysis untuk Kajian Perubahan Tutupan Lahan Di Daerah Perkotaan Menggunakan Data Satelit Landsat Multitemporal, Tesis, Program Studi Penginderaan Jauh, Program Pascasarjana Fakultas Geografi, UGM, Yogyakarta.
[3] Pascari, M, R., 2013, Linear Spectral Mixture Analysis untuk (LSMA) untuk Tutupan Lahan Menggunakan Citra Landsat ETM+ di Yogyakarta dan Sekitarnya, Skripsi, Program Studi Kartografi dan Penginderaan Jauh, Fakultas Geografi, UGM, Yogyakarta
[4] Matthias. B., Martin, H., 2003, *Proceeding of the SPIE 10th International Symposium on Remote Sensin*g, 8 – 12 September 2003, Barcelona.
[5] Mynt. S. W., 2005, Quantifying Impervious Surface, Vegetation, Soil, and Shade in the Phoenix Metropolitan Area Using Multiple Endmember Spectral Mixture Anlysis.
[6] Jensen, J. R., 1986, Introductory Digital Image Processing: A Remote Sensing Perspectiv, New Jersey: Pearson Prentice Hall.
[7] Lillesand, T. M., R. W. Kiefer., and J. W. Chipman., 2004.Remote Sensing and Image Interpretation.5th edition, New York: John Wiley and Sons, Inc.
[8] Rahman, Abdur., 2010, Analisis Campuran Spektral Secara Linear (LSMA) Citra Terra Modis untuk Kajian Estimasi Limpasan Permukaan, Tesis, Program Studi Penginderaan Jauh, Program Pascasarjana Fakultas Geografi, UGM, Yogyakarta.
[9] Lu. D., Batistell. M., Moran. E., 2004, Multitemporal Spectral Mixture Analysis for Amazonian Land-Cover Change Detection, *Can. J. Remote Sensing*, Vol. 30, No. 1, pp. 87-100.
[10] Vohland. M., Stoffles. J., Hou. C., Schuler. G, 2007, Remote Sensing Techniques for Forest Parameter Assesment: Multispectral Classification and Liniear Spectral Mixture Analysis, Silvia Fennica, *Research Article* 41 (3): 441 – 456.