LAND-COVER CLASSIFICATION FOR EAST SUEZ CANAL REGION USING HYPERSPECTRAL EO-1 DATA

Nagwan M. Afify,[a]*Abdel Aziz S. Sheta,[b] Sayed M. Arafat,[a] Afify A. Afify,[c] Mohammed S. Abd-Elwahed2[b] and Adel S. El-Beltagy[b]

Keywords: Hyperion Data, Land cover, Support Vector Machine, Remote Sensing and GIS.

Earth Observation (EO-1) data provides a highest spectral resolution to get spectral information of Earth's Surface targets within 242 spectral bands at 30 m spatial resolution. In this context, the main objective of this paper is to produce a land cover map using hyperspectral data acquired by EO-1 Hyperion instrument over one test site. Atmospheric correction on the hyperspectral data was performed using ENVI's Fast Line-of-sight Atmospheric Analysis of Spectral Hyper-cubes (FLAASH) module. Support Vector Machine (SVM) classification was implemented on the dominant elements to produce a land cover map for test site. SVM is carried out in this research to deal with the multi-class issue of Hyperion data. Classification using the kernel functions in classification made the classifier robust against the outliers. The Land Cover Classification System (LCCS) was used to know the land cover classes. The result showed high accuracy for land cover map with machine learning classifier like SVM using hyperspectral remote sensing data. The overall classification accuracy obtained was 97.85.

* Corresponding Authors
Fax: +20 2 6225800
Email: nagwanafify6@gmail.com
[a] Authority for Remote Sensing and Space Sciences
[b] Soil Science Department, Faculty of Agriculture, Ain - Shames University, Egypt
[c] Soil Water and Environment Research Institute, Agricultural Research Centre

INTRODUCTION

Land cover (LC) is very important in a lot of natural resource applications. At local and regional scale, knowledge of LC forms a basic dimension of recourses available to any political unit.[1] On a wide scale, LC information is of main importance in determining the broad patterns of climate and vegetation which form the environmental framework for human activities. Furthermore, LC maps are also a valuable contribution in the development of maintain policies particularly for ecologically protected areas and the restoration of native environments, as well as the monitoring of desertification and land degradation in regions.[2] Remote sensing has been appropriate source for LC thematic mapping.[3] Accordingly image classification[4] is the most widely used for this purpose it is the most frequently applied approach in developing land use and land cover spatial distribution maps.[5] An overview of different remote sensing classification techniques has been published.[6]

Recently, progresses in sensor technologies have directed to the launch of hyperspectral remote sensing systems. Hyperspectral sensors are able to register reflected light from land surface elements in many narrow continuous spectral bands from the visible to the shortwave infrared parts of the electromagnetic spectrum.[7] this allows hyperspectral systems to provide spectral information useful for many applications but not limited to land cover maps.[8]

The launch of the Hyperion space-borne hyperspectral sensor in 2000 under NASA's New Millennium Program on-board the Earth Observer-1 (EO-1) satellite platform.[9] Hyperion acquiring spectral information of Earth's surface objects in 242 spectral bands and at a spatial resolution of 30 m. The Hyperion sensor has two spectrometers one in the visible and near-infrared (VNIR) (bands 8–57, region 427–925 nm) and one in the shortwave infrared (SWIR) region (bands 77–224, region 912–2395 nm). The swath width of Hyperion is 7.6 km across-track, and approximately 53.6 or 80.4 km along-track.

The potential of Hyperion imagery for land cover mapping has been verified by many investigators.[10] In this context, the main objective of this study is to produce a land cover map for east Suez Canal region using hyperspectral data acquired by the EO-1 Hyperion instrument with the Support Vector Machine (SVM) classification techniques. Another goal is to use Land Cover Classification System (LCCS) to define the land cover classes.

Site selection and data source

The study area was located in El Qantra-sharq District of Ismailia Governorate east of Suez Canal covers almost 34165 feddans. It represents the new reclaimed land for the agriculture land use. The coordinates of the upper left corner are 30° 38′ 20″ N and 32° 23′ 20″ E, while the lower right corner coordinates are 30° 30′ 0″ N and 32° 28′ 20″ E (Figure 1). According to Ismailia weather station (624400), it is the nearest recording station to the study area. The climate of the study area is aridic regime, which is characterized by a short winter season and a long hot summer. The temperature sometimes varied widely through different periods of the year, as the minimum mean was 8.4 °C in January, while the maximum mean was 35.6 °C during July. The normal values of the monthly rainfall show that the average of annual rainfall was approximately 25 mm / year. The relative humidity is higher in winter than in summer, it attains a minimum average of 52.2% in May and a maximum average of 66.5% in August. The relief of the area is variable, with the average altitude varying from 0 to

200 m above sea level. Geologically, the area is belong to Quaternary deposits were divided into Holocene and Pleistocene. The vegetation cover is limited to citrus tree, clover and hordeum.

The Hyperion imagery of site selection was acquired on January 9, 2016. The imagery was received from NASA Earth Observer (EO)-1 Hyperion sensor, record as a full long scene (185-km strip) and at level 1 (L1GST) processing. This processing level product is a geo-tiff image format, and is already radiometric corrected, geometrically resampled, and registered to a geographic map projection with elevation correction applied.[9]

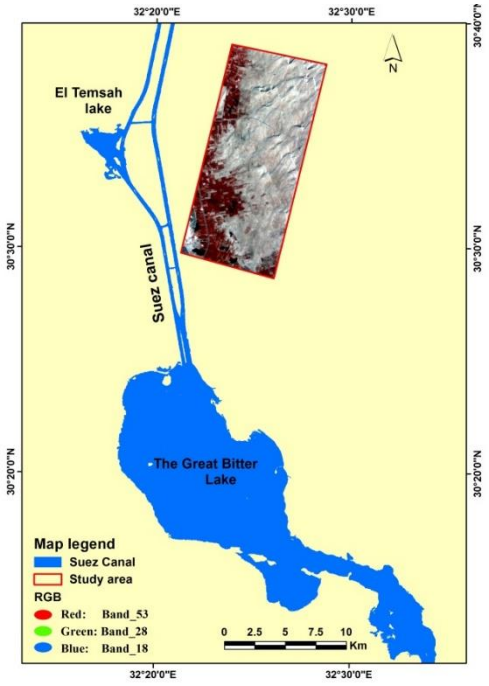

Figure 1. Location of the study area

METHODS

Summary of the methodology adopted in the study is explained in (Figure 2) and the image processing details are given in the following sections.

Linear interpolation of sensor

All pre-processing of the Hyperion imagery was carried out using ENVI (5.1). The first step in the pre-processing involved the linear interpolation of all the sensor detectors on a pixel by pixel, spectrum by spectrum, and band by band basis to a common set of wavelengths, which resulted 242-band image.

Remove bad band

The Hyperion visible and near-infrared (VNIR) spectrometer has only 50 calibrated bands, while the short-wave infrared (SWIR) spectrometer has only 148 calibrated bands. The non-calibrated bands of the Hyperion imagery (1–7, 58–76, and 225–242) were removed. The residual 198 bands cover the entire spectrum from 426 to 2395 nm therefore, the Hyperion bands sensitive to water absorption (i.e., bands 120–132, 165–182, 185–187, and 221–224) were removed in order to reduce the influence of atmospheric scattering and water vapour absorption caused by mixed gasses to the data.[11] Bands 77 and 78 were also eliminated as such bands had a low signal to noise value, and overlapped with bands 56 and band 57, respectively.[12]

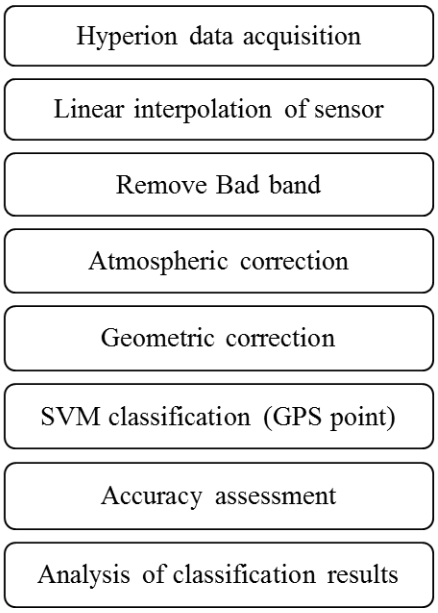

Figure 2. A flow diagram showing the processing scheme for the methodology.

Atmospheric correction

Remote sensing measurements of the Earth's surface are deeply influenced by atmosphere. Water vapour with smaller contributions from carbon dioxide, ozone and other gases dominates the absorption by atmospheric gases. To retrieve the surface reflection, the atmospheric components must be removed. In the study area, ENVI's Fast Line-of-sight Atmospheric Analysis of Spectral Hyper-cubes (FLAASH) module was applied on Hyperion data for atmospheric correction. The different parameters used in FLAASH atmospheric correction are contained in (Table 1).

Table1. Parameters used in FLAASH Atmospheric Correction

Parameter	Value
Latitude	30 29 23.28
Longitude	32 23 51.35
Sensor Altitude (km)	705
Ground Elevation (km)	0.050
Pixel Size (m)	30
Flight Date	1/9/2016
Flight Time	6:48:6
Atmospheric Model	Mid-Latitude Winter
Aerosol Model	Rural
Spectral Polishing	9 bands
CO2 Mixing Ration (ppm)	390
Zenith Angle	180
Azimuth Angle	132 57 21.96
Output Scale Factor	10000

FLAASH requires input image in BIL format and ASCII file of scale factors number. The scale factors for the VNIR and SWIR bands are 400 and 800 respectively in the case of nanometers (nm) while 40 and 80 for µm. The study area is rural and it located in winter climate. So, Mid-Latitude Winter atmospheric and rural aerosol model of FLAASH were selected and other parameters were defined based on metadata of the Hyperion image file. The change in the spectral reflectance curve of vegetation area before and after FLAASH correction can be seen in Figure 3.

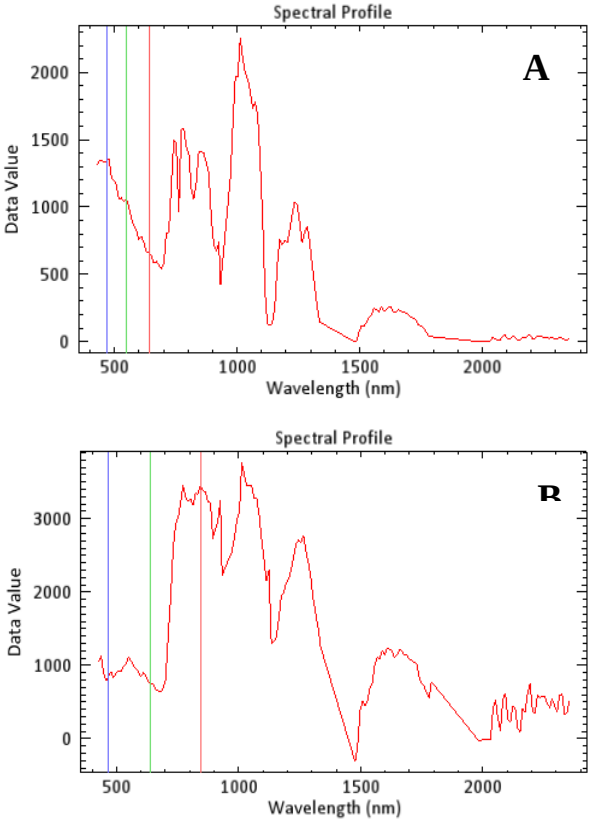

Figure 3.Spectral Curve of Vegetation (a) before FLAASH and (b) after FLAASH

Geometric correction

Geometric correction was carried out for the Hyperion image of 2016 using 40-ground control points (GCP's) obtained from a digital topographic map at a scale of 1:50,000 and Landsat ETM+ using the image- to-image technique. The geometric model used in the rectification process was three- order -polynomial and the resembling method is the nearest neighbor method.

The image was projected with Transverse Mercator projection in WGS-84 spheroid and datum. Finally, the root-mean-square error (RMSE) images were obtained as less than 0.4 pixels, which are acceptable.[13]

Support vector machine classification

Support vector machines classification (SVM) is a supervised machine learning method that performs classification based on the statistical learning.[14] Basically, SVM is based on fitting a separating hyperplane that provides the best separation between two classes in a multidimensional feature space. This hyperplane is the surface on which the optimal class separation takes place. The optimal hyperplane is the one that maximizes the distance between the hyperplane and the nearest positive and negative training. In order to represent more complex shapes than linear hyperplanes, a variety of kernels including the polynomial, the radial basis function (RBF), and the sigmoid can be used.[15] Also, a penalty parameter can be introduced to the SVM classifier to allow for misclassification during the training process.

Training data selection

It was necessary to use the training sites for the SVMs classification process which applied to the Hyperion image for land cover map. First, it has been defined for all land cover classes according to LCCS system shown in (Table 2). The classification system was based primarily on visual interpretation of the high resolution quick bird imagery acquired from Google Earth, furthermore topographic maps scale 1/ 50000. Chosen the date acquisition of the quick bird imagery was close to that of Hyperion imagery.

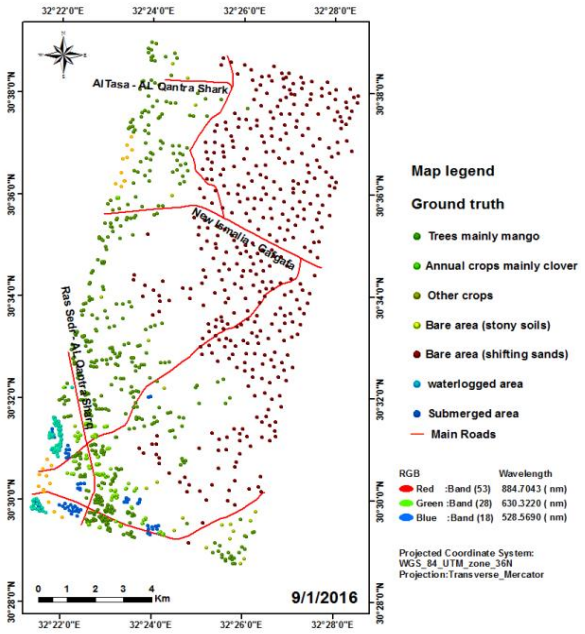

Figure 4. Training sites (GPS) for land cover classes.

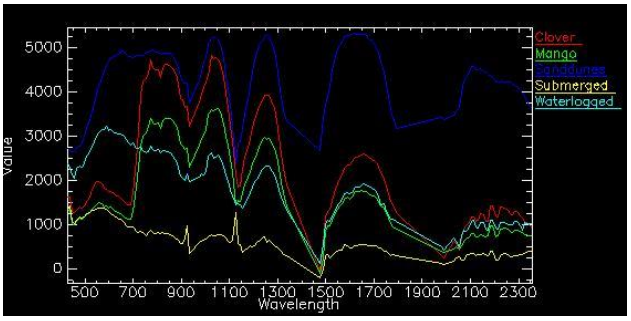

Figure 5. The component spectra for landcover classes used for SVM in test site.

Table 2. Classification key According to LCCS which was used in the land cover classes.

Class user name	LCCS code	Legend description
Permanently cropped area with surface irrigated tree crop(s) Dominant crop: fruits & mango (Mangiferaindica L.) Crop Cover: Orchard(s)	A1 = Trees A9 = Evergreen B1 = Large-Medium B5 = Continuous C1 = Single Crop D3 = Irrigated D4 = Surface D9 = Permanent S0606 = Citrus Fruits (Citrus spp.) W8 = Orchards or Other Type of Plantations	Continuous irrigated orchard of Citrus Fruits The field size varies from 2 to more than 5 fed. The class covers almost 80% of the polygon area
Permanently cropped area with small sized field(s) of surface irrigated Herbaceous crop(s) (one additional crop) (Herbaceous terrestrial crop sequentially). Dominant Crop: Clover	A3 = Herbaceous crop B2 = Small B5 = Continuous C2 = Multiple Crop C3 = 1 add. Crop C7 = Herbaceous Terrestrial C19 = Sequential D3 = Irrigated D4 = Surface D9 = Permanent Zs3 = Clover	Continuous clover crop The field size is less than 1fed. The class covers almost 80% of the polygon area
Shifting sands/Dune(s)	A2 = Unconsolidated A6 = Loose and shifting sands B1 = Dunes	Dunes
Non-perennial natural water bodies (surface aspect: sand) Scattered vegetation: scattered vegetation present	A1 = Natural Water bodies B2 = Non-Perennial B6 = Sand U1 = Scattered Vegetation	Non-perennial natural water bodies with scattered Vegetation Non-perennial natural water bodies with scattered Vegetation
Water seasonality	A23 = Waterlogged	The water table is very high and at or near the surface. These areas could be occasionally flooded but the main characteristic is the high level of the water table (e.g. bogs).

The training sites were accurately limited to include of all land cover classes. Second, training sites representative of the classes were collected from the Hyperion imagery following a stratified random sampling strategy. Third, it used for 1116 regions of interest which keeps of spectral signature from Hyperion image that has been used in SVM classification and compared it with the large spectral library of USGS are shown in (Figures 4 and 5), respectively.

Accuracy assessment

The final stage of the image classification process usually it include an accuracy assessment step.[16] Accuracy assessment is the quantification of mapping with the associate of remote sensing data, which is helpful in estimation of classification algorithms and also in limitation of the error level that might be associated with the image. The accuracy of classification is calculated in the form of an error matrix (also known as a confusion matrix).[17] Numerous methods for accuracy assessment have been explained in remote sensing. Accuracy assessment was based on the computation of the overall accuracy (OA), user's accuracy (UA), producer's accuracy (PA), and the Kappa (Kc) statistic.[18] The OA is the ratio of the number of validation pixels that have been correctly classified to the total number of validation pixels used for all classes and is expressed as a percentage (%).

RESULTS AND DISCUSSION

Figure 6 shows the land cover thematic map produced from the SVMs classification based on the Hyperion imagery acquired for our site selection. Land cover classes were extracted: cultivated land (including mango tree and clover), sand dunes, submerged area, waterlogged area, main roads and irrigation canal. According to the table 3, the total study area is 34165 feddans, mango 9445 feddans (28%), sand dunes 22940 feddans (67 %), had the highest level of the area. In contrast, the submerged area 311feddans (1 %), clover5 47 feddans (2 %) and waterlogged area 419 feddans (1 %) had the lowest level of the area while the

main roads had 435 feddans and the irrigation canal 68 feddans. moreover,[19] studied the comparison of 2 classification methods (MLC and SVM) to extract land use and land cover in Johor Malaysia. The results showed that the SVM classification based on kappa coefficient 0.86 was the most accurate method.[20] It further concluded that SVM is better than other traditional classifiers (i.e., the ML and the SAM classifier) in respect of classification accuracy and processing time.[21] The authors evaluated various algorithms for classification in land use mapping, and concluded that the SVM algorithm in comparison with the MLC algorithms and decision trees has a higher accuracy in the preparation of land use maps.

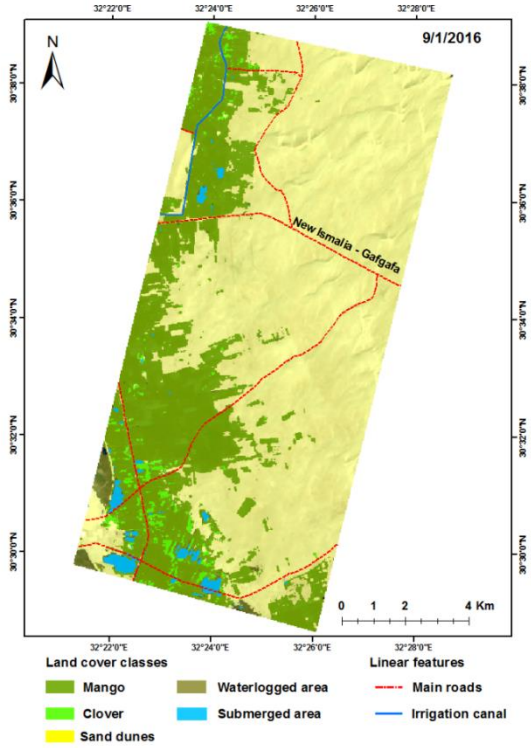

Figure 6. Land cover map for sit selection (2016) based on Hyperion image

Table 3. Land cover distribution for the study area(2016).

LC classes	Area / feddans	%
Clover	547	2%
Mango	9445	28%
Sand dunes	22940	67%
Submerged	311	1%
waterlogged	419	1%
Irrigation canal	68	0%
Main roads	435	1%
Total area	34165	100%

* Feddans = .42 hectare

Classification accuracy assessment

Classification accuracies of land cover classes using SVM classification are depicted in table 4 and table 5. A total number of 280 ground control points (GCPs) were used for accuracy assessment. 35 point of GCPs within clover, 109 point of GCPs to mango, 105 point of GCPs to sand dunes, 13 point of GCPs to submerged area and 17 point of GCPs to waterlogged area were taken. The overall classification accuracy obtained was 97.85 %. With producers and users accuracies from 92.86 % to 100 % for the individual classes, corroborating the standard accuracy of 85–90 % for land cover mapping studies as has been reported earlier.[22] The overall result showed that SVM classification process employed has got very promising potential to discriminate crops and tree classes, with high classification accuracies, when combined with high spectral resolution hyperspectral remote sensing data. The high accuracy produced by the SVM classifier may be due to the ability of the algorithm to identify the optimally separating hyperplanes for classes in comparison to other pixel-based techniques (e.g., artificial neural networks)[14] which may not be able to find such optimal hyperplanes.

Table 4. Confusion matrix for the land covers classification for the study area.

LC classes	Cl	Ma	Sa	Su	wa	Total
Unclassified	0	0	1	0	0	1
Clover	34	1	0	0	0	35
Mango	1	108	0	0	0	109
Sand dunes	0	2	103	0	0	105
Submerged	0	0	0	13	0	13
waterlogged	0	0	0	1	16	17
Total	35	111	104	14	16	280

CL=clover, Ma= Mango, Sa= sand dunes, Su= submerged, Wa= waterlogged

$$Accuracy_{Total} = \frac{38 + 108 + 1103 + 13 + 16}{280} * 100 = 97.85\ \%$$

Table 5. Accuracy totals for the classified images.

LC Classes	Producers Accuracy	Users Accuracy
Clover	97.14%	97.14%
Mango	97.30%	99.08%
Sand dunes	99.04%	98.10%
Submerged	92.86%	100.00%
waterlogged	100%	94.12%

Overall classification accuracy = 97.85 %

CONCLUSIONS

The aim of this research is to produce a land cover map for east Suez Canal area using hyperspectral data acquired by the EO-1 Hyperion instrument in conjunction with the support vector machines (SVM) classification techniques. SVM has a good generalization potentiality which stems from the selection of the hyperplane that maximizes the geometric margin between classes which helped to discriminate between the classes of land cover and various

special types of plants. The result showed that SVM classification process has got high classification accuracies, when combined with high spectral resolution hyperspectral remote sensing data. More research will be done to improving the classification accuracy and reducing the calculation time.

REFERENCES

[1]Kavzoglu, T., Colkesen, I., *Int. J. Appl. Earth Obs. Geoinf.*, **2009**, *11(5)*,352-359.https://doi.org/10.1016/j.jag.2009.06.002

[2]Castillejo-González, I. L., López-Granados, F., García-Ferrer, A., Peña-Barragán, J. M., Jurado-Expósito, M., de la Orden, M. S., González-Audicana, M., *Comput. Electron. Agr.* **2009,** *68, 2*, 207-215. Doi:10.1016/j.compag.2009.06.004

[3]Chintan, A. S., Arora, M. K., Pramod, K.V., *Int. J. Remote Sens.*, **2004**,*25*,481–487. doi.org/10.1080/01431160310001618040

[4]Mathur, A., Foody, G. M., *Int. J. Remote Sens.*, **2008**, *29(8)*,2227-2240. http://dx.doi.org/10.1080/01431160701395203

[5]Borak, J.S., Strahler, A.H., *Int. J. Remote Sens.*, **1999**, *20*, 919–938. http://dx.doi.org/10.1080/014311699212993

[6]Lu, D., Weng, Q., *Int. J. Remote Sens.*, **2007**,*28, 5*, 823–870. http://dx.doi.org/10.1080/01431160600746456

[7]Xu, D.Q., Ni, G.Q., Jiang, L.L., Shen, Y.T., Li, T., Ge, S.L., Shu, X.B., *ASR.*, **2008**, *411*,1800–1817. https://doi.org/10.1016/j.asr.2007.05.073

[8]Dalponte, M., Bruzzone, L., Vescovo, L., Gianelle, D., *Remote Sens Environ.*, **2009**, *113(11)*, 2345-2355. https://doi.org/10.1016/j.rse.2009.06.013

[9]USGS, **2006**, EO1 User's Guide. http://eo1.usgs.gov

[10]Galvão, L. S., Formaggio, A. R., Tisot, D. A., *Remote Sens Environ.*,**2005**, *94(4)*,523-534. https://doi.org/10.1016/j.rse.2004.11.012

[11]Beck, R., *EO-1 user guide, version 2.3*, Satellite Systems Branch, USGS Earth Resources Observation Systems Data Center (EDC).**2003**.

[12]Datt, B., McVicar, T. R., Van Niel, T. G., Jupp, D. L., Pearlman, *IEEE Trans. Geo. Sci. Rem. Sens.*, **2003**,*41(6)*,1246-1259. https://doi.org/10.1109/TGRS.2003.813206

[13]Tucker, C. J., Grant, D. M., Dykstra, J. D., *Eng. Rem. Sens.*, **2004**,*70(3)*,313-322. https://doi.org/10.14358/PERS.70.3.313

[14]Licciardi, G., Pacifici, F., Tuia, D., Prasad, S., West, T., Giacco, F., Gamba, P., *IEEE Trans. Geosci. Remote Sens.* **2009**. *47 (11)*, 3857-3865. https://doi.org/10.1109/TGRS.2009.2029340

[15]Fauvel, M., Benediktsson, A., Chanussot, J., Sveinsson, IEEE Trans. Geosci. Remote Sens. **2008**, *46(11)*, 3804–3814. https://doi.org/10.1109/ICASSP.2006.1660467

[16]Patil, M. B., Desai, C. G., Umrikar, B. N., *Int. J. Geol. Earth Environ. Sci.*, **2012**, *2*, 189-196. [17]Congalton, R. G., *RSE.*, **1991**, *37(1)*, 35-46.https://doi.org/10.1016/0034-4257(91)90048-B

[18]Congalton, R. G., Green, K., *Assessing the accuracy of remotely sensed data: Principles and applications*, *CRC* press.**1999**, 137.

[19]Deilmai, B. R., Ahmad, B. B., Zabihi, H., *IOP Conf. Ser. Earth Environ. Sci.*, **2014**.*20, (1),* 012052. Doi:10.1088/issn.1755-1315

[20]Moughal, T. A., *J. Phys. Conf. Ser.*, **2013**, *439*, 012042 Doi:10.1088/issn.1742-6596

[21]Otukei, J. R., Blaschke, T., *Int. J. Appl. Earth Observ. Geoinform.* **2010**, *12*,S27-S31. https://doi.org/10.1016/j.jag.2009.11.002

[22]Anderson, J. R., *A land use and land cover classification system for use with remote sensor data*, US Government Printing Office, **1976**, 964.

Received: 11.20.2017.
Accepted: 10.12.2017.