SEABUCKTHORN (HIPPOPHAE) MAPPING
IN LADAKH REGION OF JAMMU AND KASHMIR

Kiran Chauhan[a, *], Sangeeta Khare [a], S.C.Jain [a]

[a] Defence Electronics Applications Laboratory
Raipur Road, Dehradun, INDIA – 248001
Tel:091-135-2787083/84, Fax:091-135-2787265

KEY WORDS: Remote Sensing, Seabuckthorn, Image mosaic, training samples, Classification, Estimation.

ABSTRACT:

Remote sensing is a powerful tool for the regional mapping of natural resources. Satellite imagery has been demonstrated to be a cost effective method for classifying land use and land cover types throughout the world. This paper evaluates the potential of satellite remote sensing technology by presenting a case study for mapping ground coverage of Seabuckthorn plant, which grows in high altitudes in Ladakh region of Jammu & Kashmir. Mosaic of fourteen IRS-1D multispectral scenes covering the entire Ladakh district was prepared. Training samples were collected from Nubra-Shyok area based on the ground truths. This scene was then classified using the supervised classification technique to identify and extract information on spatial extent, distribution and estimation of Seabuckthorn plant with an overall accuracy of 96%. The study is helpful in exploitation of this highly beneficial natural resource.

1. INTRODUCTION

1.1 Thematic mapping with Remote Sensing data

Land is the most important natural resource, which embodies soil, water and associated flora & fauna involving the total ecosystem. Remote sensing has been used for environmental monitoring including forest cover and is helpful to obtain land use/ land cover for inaccessible dense forests on mountains & in the marshy low lands of the costal region. There exist many natural resources on the earth, which are highly useful to the mankind. Exploitation and utilization of such resources requires proper understanding of how to use the resources, its quantity and its geographical extent. In Himalayan region there grows a wild plant named as Seabuckthorn. Studies have shown that it is highly beneficial for health. It is rich in many vitamins and minerals. It is considered good for heart as well as known for its anti-stress quality.

This plant grows only in high altitudes. Assessment of geographical extent of coverage by conventional surveying technique is difficult because of undulating high altitude terrain. The present paper is based on a case study to estimate the ground coverage of Seabuckthorn in Ladakh region of Jammu & Kashmir so as to find out the potential for setting up food processing units for the use of this plant.

Photography, multispectral data, thermal images, airborne radar are some of the remote sensing data used for resource monitoring. An accurate land cover classification system is essential for providing information for assessment and effective management of natural resources. Satellite imagery has been demonstrated to be a cost effective method for classifying land use and land cover types throughout the world. Satellite remote sensing has an advantage of providing information rapidly and repetitively by means of sensors operating in different spectral bands. Remotely sensed data in Visible and NIR bands are used for crop growth and condition assessment.

1.2 Earlier studies

Earlier studies of resource mapping & monitoring using remote sensing multispectral data are of Snow cover monitoring using LISS and SWIR data (Kulkarni et al, 2002), Forest mapping using LISS and TM data (Nelson, 1983), ocean color chlorophyll assessment using P4 data (Chauhan et al, 2003), Detection of Diseased Rubber Plantations using IRS LISS III data (Ranganath et al, 2004), finding ground water prospect zones using LISS III data (Khan and Moharana, 2002).

This paper is organised as follows. In section 2, the Study area is described. Section 3 describes the methodology including data used, training site selection, techniques used for classification and classification of complete Ladakh area. Section 4 shows the results and section 5 draws the conclusions indicating the importance of remote sensing data in such type of studies. Few references are provided in the last section of the paper.

2. STUDY AREA

The study area is the complete Ladakh region in Jammu & Kashmir (figure 1), which is situated in Northern Region of India covering the Himalayan ranges. It is bounded by 74º45'E-80º26'E Longitude and 32º30'N-36º15'N Latitude. The complete region is a high altitude region where climate remains cold throughout the year and the region is covered with snow for around 5-6 months. This approximately 145 Lakh Hectare region is the major source of Seabuckthorn in India, especially the Nubra-Shyok valley.

---

* Corresponding author - kchauhanji@rediffmail.com

Figure 1: Area under study

3. METHODOLOGY

The job entailed collection and supervised classification (Jensen, 1996) of satellite imagery for the Ladakh region. Since the IRS imagery product size is only 140 km x 140 km, there is a requirement to mosaic many scenes and then perform the multi-spectral classification on this image-mosaic to classify the image pixels corresponding to the Seabuckthorn category. The task was carried out in the following sequential activities (Chauhan, 2003).

- I. Selection of IRS-1D LISS-III scenes pertaining to Ladakh region.
- II. Training Site selection.
- III. Spectral signature profiles generation & classification of training sites
- IV. Validation of training results
- V. Prepare mosaic of required LISS scenes.
- VI. Estimation of Seabuckthorn coverage by classifying scene-mosaic.

3.1 Data Used

The Indian Remote sensing Satellite IRS-1D data of LISS III sensor with spatial resolution of 23.5m in digital format for paths from 92-95 and rows from 45-47 is used for the study as shown in figure 2.

Taking into account the large area of the study region WIFS data can also be used. But the advantage of LISS III data over WIFS data is that LISS data has 3 spectral bands in visible range where as WIFS data has only 2 bands thus provides better spectral resolution. Moreover spatial resolution of WIFS data is 180m and this plant grows only in small patches in the region. Thus 23.5 m LISS III data is suitable for accurate training sample selection. Bands 2, 3 and 4 were considered and false colour composite of these bands was prepared.

For marking the boundary of the region on the image ancillary data like Jammu and Kashmir map is used. All the images were of months between July to September, which is the leaves and fruit-bearing season for the plant. In November and December the leaves shed off and fruits also start shedding off. After December the region is mainly covered with snow and is

dry up to May-June. So the best season for the study is from July to September.
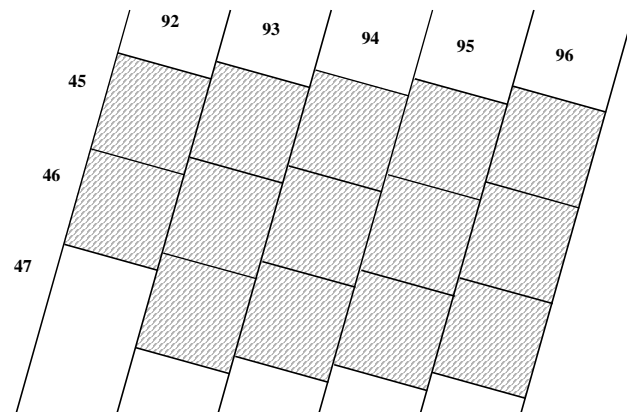

Figure 2 – IRS-1D LISS III scenes covering Ladakh region

3.2 Selection of Training Sites

Training sites are necessary to define classes. Training is a critical step in supervised classification of image. As the training samples should be representative of the land cover classes, they are collected from relatively homogeneous area on the ground. This plant is mainly present in the Nubra-Shyok valley of Jammu and Kashmir. There are many patches, which are purely of Seabuckthorn bushes. No other plant is present in those patches. These training sites were selected and the class was identified as *Pure Seabuckthorn class*. Seabuckthorn is also present in the region mixed with some local plant in Nubra-Shyok as well as in Leh area. From these areas sites of Seabuckthorn mixed with local plant – a *Mixed class* were selected. Training site selection was facilitated by the availability of ground truth information.

3.3 Classification & Validation

Supervised classification technique is applied to find out spatial extent and distribution of the plant. In-house developed multispectral classification software (Chauhan, 2004) is used to carry out the study. This software has the capability to process very large scenes. Conventional as well as fuzzy logic based supervised classification algorithms are provided. Using this software the spectral signature profiles are generated for each band of the training data and training set statistics were computed to understand the reflectance behaviour of Seabuckthorn plant covered land. The statistical parameters computed for each training class include univariate parameters - mean, minimum & maximum gray values of each band along with variance, standard deviation and multivariate parameters - variance-covariance matrix & correlation matrix. Signature files store the statistics of the samples gathered from the training site, that are later applied to the entire image during classification procedure. These parameters were analysed to ensure that the classes were well separated from each other. Statistical parameters of the two classes are shown in table3.

Class : Seabuck					
Band	mean	min	max	var	sd
4	134.65	110.00	150.00	56.00	7.48
3	65.44	57.00	81.00	28.35	5.32
2	93.38	78.00	111.00	36.70	6.06
Covariance matrix					
56.00	-11.60	-13.95			
-11.60	28.35	26.36			
-13.95	26.36	36.70			
Correlation matrix					
1.00	-0.29	-0.31			
-0.29	1.00	0.82			
-0.31	0.82	1.00			
Class : Mixed					
Band	mean	min	max	var	sd
4	178.45	156.00	225.00	172.55	13.14
3	60.28	51.00	74.00	21.68	4.66
2	100.03	82.00	127.00	66.71	8.17
Covariance matrix					
172.55	16.28	49.36			
16.28	21.68	29.13			
49.36	29.13	66.71			
Correlation matrix					
1.00	0.27	0.46			
0.27	1.00	0.77			
0.46	0.77	1.00			

Table 3 – Statistical Parameters for Classes considered

Training sets were plotted to see if there is any conflict between them or not. Fig. 4 shows the mean spectral reflectance pattern of the signatures chosen for the analysis.
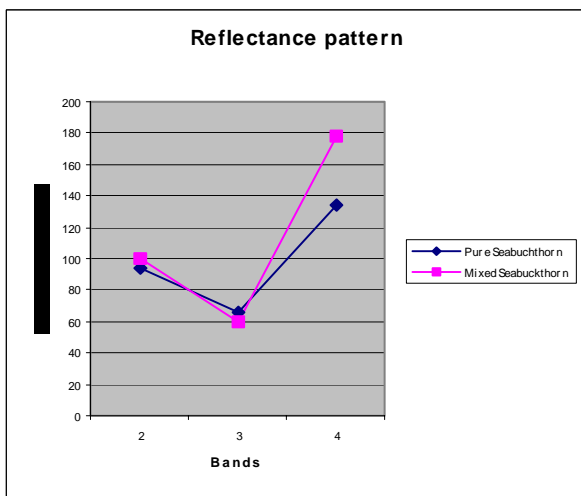

Figure 4 – Mean spectral reflectance pattern for all bands

Nubra-Shyok area and Leh area subscenes were classified using the selected training samples. There are many supervised classification techniques. But because of the large area constraint, computationally extensive and time-consuming techniques like fuzzy classification were not preferred. Comparatively simple technique with comparable accuracy is used. The accuracy is maintained by correct judgment and selection of training sites such that training classes do not overlap. Band four is the most discriminating band for the two classes as depicted by fig 4. The training samples for the two classes are well separated which can be seen in fig 5, thus thresholded maximum likelihood classification algorithm has produced accurate classification.
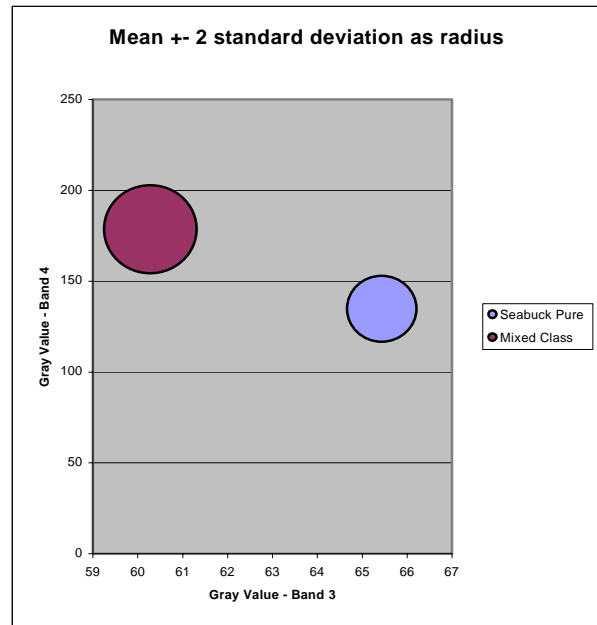

Figure 5 – Plot for reflectance values between bands 3 & 4

According to this algorithm (Jensen, 1996) the decision rule applied to the unknown measurement vector, X is : Decide X is in class c, iff

$$P_c \geq P_i \qquad (1)$$

where i = 1,2,3,……,m possible classes and

$$P_c = -0.5\log_e\left(\det(V_c)\right) - 0.5(X - m_c)^T(V_c)(X - m_c) \qquad (2)$$

and *det(V_c)* is the determinant of covariance V_c . Therefore, to classify the measurement vector X of an unknown pixel into a class, the maximum likelihood decision rule computes the value P_c for each class. Then it assigns the pixels to the class that has largest (or maximum) value only if its probability of belonging to that class is more than 80% else it is assigned to unknown class.

The coverage estimated by the satellite image in the training areas was in agreement with those present on the ground.

Accuracy assessment (Jensen, 1996; Lillisand and Kiefer, 1999) of classification is an essential component of classification process. Therefore, a complete accuracy assessment was performed on the training samples collected using the above mentioned classification algorithm. This generated an error matrix also called the contingency matrix between the reference data points and the classified data points. Descriptive statistics evaluated from the error matrix include user's accuracy (error of commission), producer's accuracy (error of omission), overall accuracy (total percent correct) and the Kappa index of agreement (KHAT accuracy). Contingency matrix is given in table 6. It was observed that only 3-4% pixels in each class were not classified i.e. they were assigned unknown class but none of the pixels were assigned a wrong class.

	Reference data samples		Row total	User's Accuracy
	Pure Seabuck	Mixed class		
Classified as				
Pure Seabuck class	*112*	*0*	112	1
Mixed class	*0*	*70*	70	1
Unknown	*5*	*2*	7	0
Total pixels	117	72	189	
Producer's accuracy	0.9573	0.9722		0.963

Table 6 – Error matrix for accuracy assessment

$$Overall\ Accuracy = \frac{Sum\ of\ diagonal\ elements}{Total\ no.\ of\ pixels} \quad (3)$$

Overall accuracy = 96%

$$K_{hat} = \frac{N\sum_{i-1}^{r} x_{ii} - \sum_{i=1}^{r}\left(x_{i+} \times x_{+i}\right)}{N^2 - \sum_{i=1}^{r}\left(x_{i+} \times x_{+i}\right)} \quad (4)$$

Where

x_{ii} is the diagonal elements
N is the total no. of data points
x_{i+} is the marginal total of rows
x_{+i} is the marginal total of cols
r is the number of classes

K_{hat} = .9247

The accuracy with Kappa coefficient is less because it includes the marginal totals also which is not considered in the case of overall accuracy.

Error of omission is 4.27% for Seabuck and 2.78% for mixed class. There is no error of commission as none of the pixels are assigned to wrong class. Only some of the pixels are not assigned any class. The classified results of these zones were found satisfactory and matching with the ground coverage.

3.4 Classification of complete study area

To find out ground coverage in whole Ladakh area (Chauhan, 2003), mosaic of 14 Raw scenes of IRS-1D LISS III data was prepared using ERDAS Imagine software (ERDAS Imagine 8.5) by selecting registration control points from the overlapping areas in the two adjacent scenes. In these overlapping areas, pixels of only one of the images were kept in the scene-mosaic so as to retain the original multispectral values. A subset was created by clipping the image mosaic coinciding with administrative boundaries of the study area from the political map. Full clipped image mosaic of 21702 * 21109 * 3 pixels ~ 1.3 GB of approximately 145 lakh hectare area was subjected to classification using the training samples selected for *Pure Seabuckthorn class* and the *mixed class* which transformed the multispectral data into a classified image. Zone wise analysis for the plant was also done interactively using the multispectral classification software.

4. OUTPUT AND RESULTS

In the present study spatial extent and distribution of Seabuckthorn is successfully estimated in the study area using the satellite data. Figure 7 shows the training sites of pure and mixed seabuckthorn classes in Nubra Shyok area. Figure 8 shows the classified Nubra Shyok area. Figure 9 shows the original image in Leh area. Figure 10 shows the classified Leh area. *Pure Seabuckthorn* is overlaid on the original image in red color and *Mixed Seabuckthorn* is shown in green color. The presence / absence of the plant is estimated in the whole Ladakh region. The area covered with *Pure Seabuckthorn* is estimated to be approx. 7184 Hectares. Area covered with *Mixed Seabuckthorn* class is approx. 2083 Hectares.
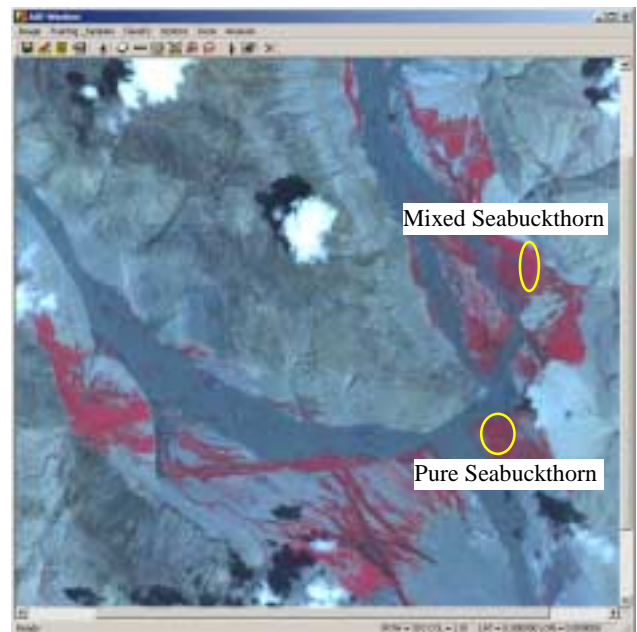

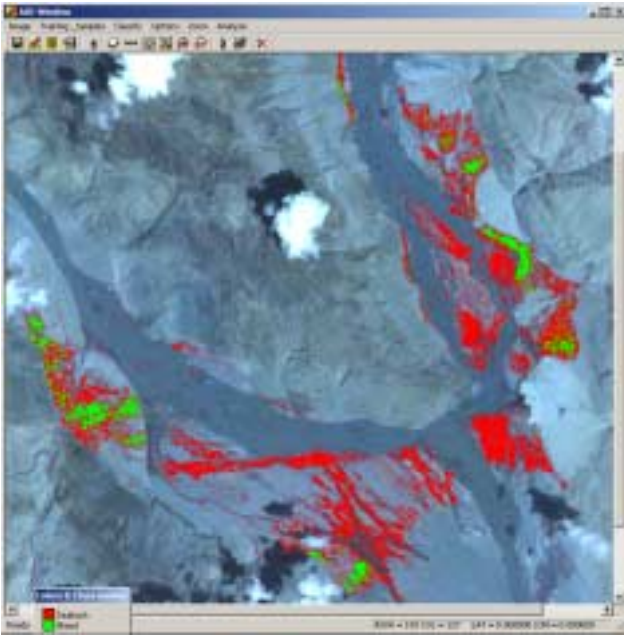Figure 7. Training Sites in Nubra-Shyok area
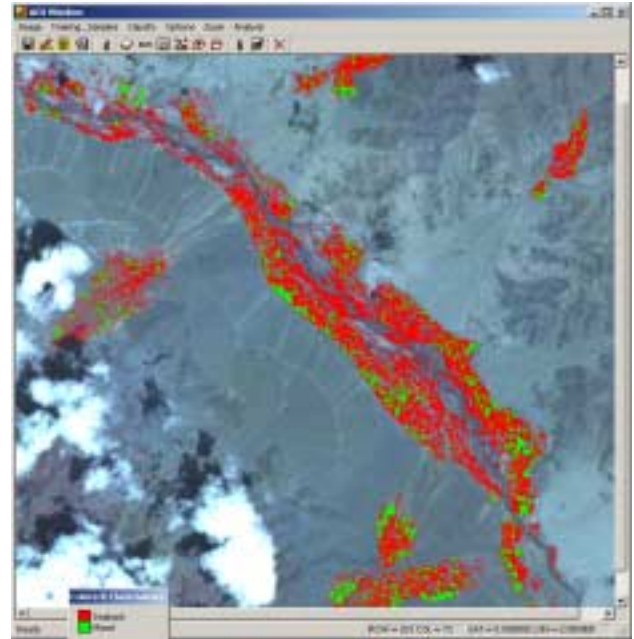
Figure 8. Classified Nubra Shyok area

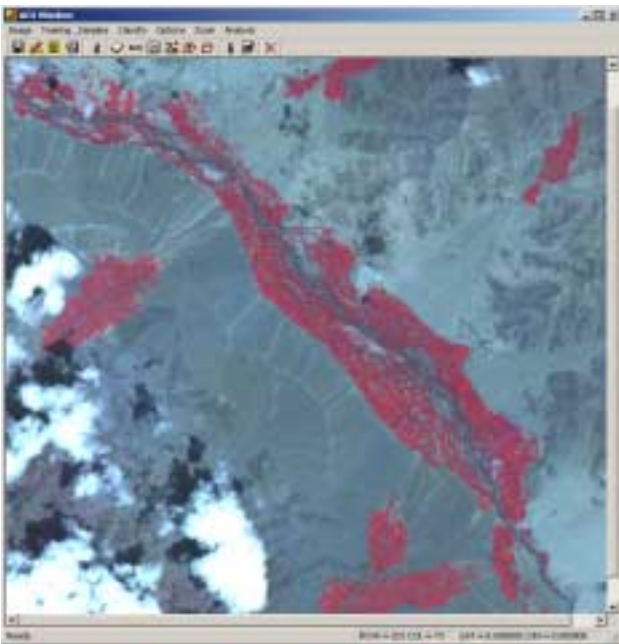Figure 10. Classified Leh area

Figure 9. False color composite of Leh area

5. CONCLUSION

The present study has demonstrated the potential of remote sensing technique for mapping, monitoring and assessment of Land resources. The study has focused on the utility of remote sensing data in identification of suitable zones for Seabuckthorn exploitation even in a complex terrain like the study area. Zonal analysis by the software helps in identifying areas with high Seabuckthorn yield.

REFERENCES

References from Journals:

Chauhan P., Mohan M., and Nayak S. 2003. Comparative analysis of ocean color measurements of IRS-P4 OCM and SeaWiFS in the Arabian Sea. *IEEE Transactions on Geoscience and Remote Sensing*, 41(4), pp. 922-926.

Khan M.A. and. Moharana P.C, 2002. Use of Remote Sensing and Geographical information system in Delination and Characterisation of Ground water prospect Zones. *The Indian Society of Remote Sensing*, 30(3), pp. 131-141.

Kulkarni A.V., Srinivasulu J., Manjul S.S. and Mathur P., 2002. Field based Spectral Reflectance studies to develop NDSI method for Snow Cover Monitoring. *The Indian Society of Remote Sensing*, 30(1&2), pp. 73-80.

Nelson R.F, 1983. Detecting Forest Canopy due to insect activity using Landsat MSS. *Photogramm. Engg. & Remote Sensing*, 49, pp. 1303-1314.

Ranganath B.K., Pradeep N. et al, 2004. Detection of Diseased Rubber Plantations Using Satellite Remote Sensing. *The Indian Society of Remote Sensing*, 32(1), pp. 49-58.

References from Books:

Jensen J.R, 1996. *Introductory Digital Image Processing- A Remote Sensing Perspective*, Second edition, Prentice Hall, New York, USA., pp. 199-251.

Lillisand T.M. and Kiefer R.W, 1999. *Remote Sensing and Image Interpretation*, Jhon Wiley and Sons, New York, USA., pp. 611-618.

References from Other Literature:

Chauhan Kiran, 2003. Technical document – *Estimation of Seabuckthorn (a wild plant) coverage in ladakh region of jammu & kashmir*, DEAL, Dehradun.

Chauhan Kiran, 2004. Technical document – *Multispectral Classification Software*, DEAL, Dehradun.

ERDAS Imagine Ver 8.5 - User's Guide