ISSN: 0974-2115
www.jchps.com Journal of Chemical and Pharmaceutical Sciences

Yield per hectare of rice crop from EOS hyperspectral data analysis

Kavitha P[1*], Shanmugha Sundaram G A[2]
[1]Centre for Excellence in Computational Engineering and Networking, ASE-Coimbatore,
Amrita Vishwa Vidyapeetham University, Coimbatore-641112, TN, India
[2]Department of Electronics and Communications Engineering, ASE-Coimbatore,
Amrita Vishwa Vidyapeetham University, Coimbatore- 641112, TN, India.
***Corresponding author: E-Mail: kavithapalattu09@gmail.com**

ABSTRACT

Rice is one of the routinely consumed diets of Kerala that is commonly cultivated in Palakkad and Malappuram districts. The study here describes the disparity in paddy yield based on different disease infestation levels on paddy crops by analyzing hyper spectral data that are obtained from the EO-1 Hyperion Sensor. In an optical wavelength range the hyper spectral data contains a large number of narrow spectral channels, and they can vary from several tens to a few hundred in number. In this paper, remote sensing technology is used to recognize disease infested plants based on the spectral reflectance of each band in the image; spectral reflectance may change according to the chlorophyll contents. Hyper spectral data is preprocessed using ERDAS IMAGINE 9.2 software. An end member spectrum is then developed by unsupervised classification (K-Mean algorithm) using ENVI 4.7, for automatically cluster pixels to each classes namely, high yield, low yield, moderate yield, and severe yield. The cultivation of rice crop in tonnes of grains per hectare area (6257.11ha) for different variety is determined as an outcome of this study.

KEY WORDS: EO-1 Hyperion Sensor, Hyper spectral data, K-Mean algorithm.

1. INTRODUCTION

Remote sensing is an emerging technology, which deals with different applications like, agriculture, urbanization & transportation, natural resource management, satellite image deforestation etc. In agriculture there are many applications like soil sensing, farm classification, farm condition assessment, agriculture estimation, mapping of farm and agricultural land characteristics etc. Rice remote sensing is focused on the mapping of field area and estimation of production (yield). The crop production can be increased by calculating the different crop parameters like leaf area index (LAI), spectral reflectance, water content, nitrogen content and also the production may decreased due to disease infestation.

Many of the disease infestations are affecting on rice crops, thereby it decreases the yield of the crops. Plant diseases can cause very large crisis through destroying the large areas of agricultural fields. Many insects like species of amphibians, reptiles, and insects are largely destroying the rice fields and some diseases are due to the climatic conditions.

Here, EO-1 Hyperion data set of rice is taken and which is downloaded from earthexplorer.usgs.gov. The Hyperion data can be used for pre-processing, unsupervised classification, yield estimation etc., because of higher resolution (spectral & spatial), data volume and redundant pixels of the hyper spectral image. So it can be used for analyzing the yield of rice crop.

This paper contains three stages like pre-processing, unsupervised classification (K-Mean algorithm), and yield estimation by using the end member collection spectrum. Pre-processing is done by using ERDAS IMAGINE 9.2; it includes different stages, Sub setting, de noising, and atmospheric correction (IARR) etc. EO1 Hyperion image is taken in the form of Geo TIFF file format. First stage of pre-processing is sub setting; it is done for cropping the unwanted portion of an image with correct proportions, after sub setting de noising the image, it removes the negative effects from the image and atmospheric correction is the last stage of preprocessing, it contains several types of methods like log residuals, IARR, flat field, here IARR is taken for preprocessing.

Unsupervised classification is done by using ENVI 4.7 software, clustering the image pixels automatically to each class. The classification automatically assigning the colors to each class. The algorithm used here is k -mean, it calculates the initial class means and distributed the each pixel in to the nearest class using minimum distance technique. It contains several iterations, and this iteration recalculates and reclassifies the pixels with respects to the new means, and some pixels may be unclassified. When a maximum number of iteration is reached, the process will continues until the number of pixels in each class changes.

Estimating the yield of rice crops by using an end member collection spectrum according to the spectral reflectance and band number, the each band number corresponds to different wavelength as per the Hyperion dataset. Each infestation levels contain different peaks, and the each peak denotes different infestation levels. When an insect is infested on the plants, their growth will be stressed due to the damage in chlorophyll pigments and structure of tissue and infestation deeply affect the photosynthesis and metabolism, due to this reason chlorophyll content will change. So health of the crop will vary. ROI file is made by using a classified image and plot the spectrum of paddy crops in ENVI software.

2. RESOURCES AND METHODOLOGY

2.1. Study area: As a case study, the Palakkad area, extending between 12°00'41.76''N to 10°11'36.92''N latitude and 76°40'38.95''E to 76°17'14.76''E longitude .The study area is taken on the basis of where paddy is cultivated, Fig. 1 shows the study area of Palakkad.

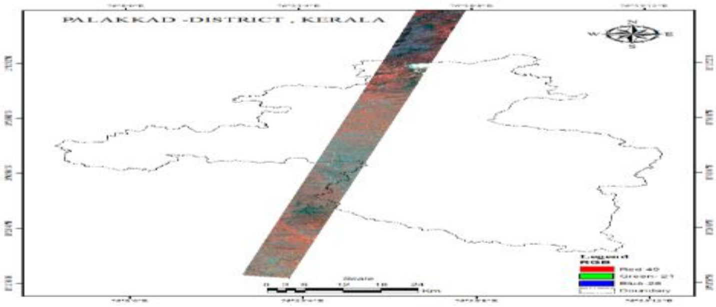


Fig. 1. Study area of Palakkad

2.2. Hyperion sensor and used dataset: The Hyperion imagery (EO-1) is used in the present study was acquired on January 30th, 2010, and obtained with free of cost from the United States geological survey (USGS) archive. Data contains 242 spectral bands, 0.35-2.5µm wavelength. The Hyperion sensor consists of 2 detectors, VNIR and SWIR, which covers spectral range of 400-1000nm and 900-2500nm wavelength. Hyperion sensor is the first developed space borne hyper spectral instrument; it has spectral resolution of 10nm and spatial resolution of 30nm. Hyperion data provides spectral coverage of 242 spectral bands, Out of 242 spectral bands, some of the bands are eliminated (1-7, 58-82, 120-134, 165-182, and 221-242) and remaining 132 bands are in use. The sensor has a swath width of 7.7km and the length of the scan line is extended from 42km to 185km.

The hyper spectral is downloaded is in the form of a full long strip and it has a length of 185 km and the processing level is level 1(L1GST) in Geo TIFF format, written as band-interleaved-by-line (BIL). The L1G product is already corrected radio metrically, and resampled geometrically. The parameters of hyper spectral data will vary depends on the study area that we have taken.

2.3. Image preprocessing: Preprocessing of hyper spectral image is done for getting a good quality image. It has three main steps, first one is de noising the image for removing the noise, second one is sub setting, for dimensionality reduction, and third one is atmospheric correction (IARR-Internal average Relative Reflectance Correction) [6] for reducing the effects of atmospheric conditions on the image values. The radiation is reflected from the surface, the radiation scattered and emitted by the atmosphere where the value recorded at a given pixel, this method is used for calculating the average radiance for each and every band of the image and divides it into actual radiance for each band of each pixel, in order to create an image of apparent reflectance. These are three main steps of preprocessing and it is done by using ERDAS IMAGINE 9.2 software. The below Fig. 2 shows the preprocessed image.

Table.1.Specifications of the used EO-1 hyperion dataset

Basic Parameter	Value	Basic Parameter	Value
entity-Identification(ID)	EO1H1440522010030110PL_PF2_01	Scene Start time	2010:030:04:57.030
Acquisition of data date/time(UTC)	2010/01/30	Scene Stop Time	2010:030:04:58:18.30
Orbit Path	144	Sun Azimuth	131.636783
Orbit Row	52	Sun Elevation	47.004766
Station	PF2	Sensor look angle	-5.0034
cloud coverage	0 to 9% Cloud Cover	Satellite Inclination	98.19
Preprocessing level	Level 1R	Geographical Extent	NE:76°44'44.57''E SW:76°17'14.76''E

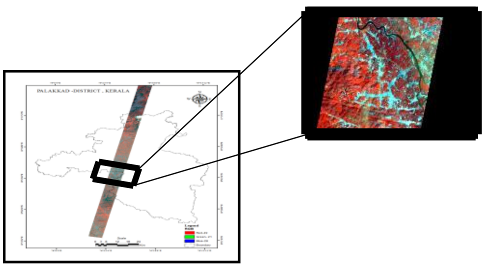


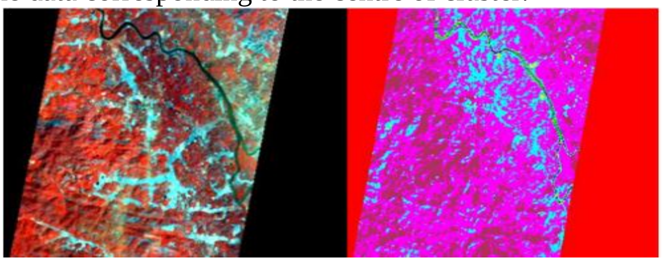
Fig. 2. Preprocessed image

2.4. Unsupervised Classification: The unsupervised classification analyzes and classifies a certain number of pixels (raster cells). It is an automatic classification method, no need of training data .The unsupervised classification methods are Simple one-pass clustering, K -Means, Fuzzy c Means, ISO data classification etc. Here the K- Mean algorithm is taken for the unsupervised classification.

2.5. K-Mean classification or clustering: Clustering is the classification of an objects in to different groups (class), more exactly we can say division of a data set into subsets or clusters, so the data in each subset maintaining a distance measure. K-Mean is a partitional clustering algorithm, it determine all cluster at once. The common distance measures are Euclidean distance, Manhattan distance, Mahalanobis distance, Hamming distance etc. K-mean is an algorithm which equally divides and classifies the n number of objects in to k partitions. K-Mean algorithm is somewhat relative to expectation –maximization algorithm. K-mean calculates the initial class means and distributed the each pixel in to the neighboring class using minimum distance technique. It contains several iterations, and this iteration recalculates and again classifies the pixels with respects to the new means, and some pixels may be unclassified. The classification is continuing until the number of pixels in each class changes. Separating N number of data points into K disjoint subsets, it contains data points in (1) so as to minimize the sum-of squares criterion.

$$J = \sum_{j=1}^{k} \sum_{n \in s_j} \left| x_n - \mu_j \right|^2 \qquad (1)$$

Where x_n a vector represents the n^{th} data point and u_j is the geometrical center of the hyper spectral data point in s_j. Using K-mean clustering the data can be grouped in K number of group; it can be any number of groups, based on the image features. Here K is a positive integer number. Grouping is based on the minimum value of square of distance between the data corresponding to the centre of cluster.



Original Image **Classified Image**

Fig.3.Classified image by using un-supervised classification (K-Mean) algorithm

The above Fig. 3 shows that a classified image, it contains different colors of classification Maroon, Magenta, Yellow, and Blue. The each color is indicating different disease infestation levels of paddy field. Infestation levels are categorized by using different spectral reflectance. Spectral reflectance means how much light is reflected from the paddy field, according to the spectral reflectance chlorophyll content will also change. If chlorophyll content is increased thereby decreasing the insect damage levels, similarly if chlorophyll content is decreased thereby increasing the insect damage levels. The below TABLE II shows four different levels of disease infestation, each level of infestation occupied in an hectare area according to the pixel size and 16.23, 39.33, 9.402, 0.568% are the percentage level of classification is done in the particular area. The percentage of classification means pixel wise classification, similar pixels classified in to the same classes. Class1 is healthy and class 2 is light healthy, the area occupied in the class1 is comparatively low with class 2, but the class1 has higher healthier than class 2, because of disease infected plants in the class 1 is less. For calculating the hectare area primarily the area of one pixel has to be calculated. The area of one pixel is 30m×30m (3om being the spatial resolution of the Hyperion data). Finally the total area of a class can be calculated by multiplying the number of pixels in the class and area of a single pixel and converting the square meter (unit of area) in to hectare by dividing 10,000(1 ha= 10,000m^2).

Table.2.Four levels of infestation and the area (hectares) occupied in the each level

Colors & type of infestation levels in the each classes	Points(pixels contained in the each polygon)	Area occupied for each region of paddy field (hectares)	% of classification	Yield per class (in tonnes)
Maroon-Healthy	17,238	1550.2937	16.239	3864.57
Magenta-Light	41,753	3755.0419	39.334	9360.56
Cyan-Moderate	9,980	897.5479	9.402	2237.40
Yellow-Severe	603	54.2306	0.568	135.186

3. RESULTS AND DISCUSSIONS

Here the pre-processing is done by using EO-1 Hyperion image and the image is obtained from EO-1 Hyperion sensor. Preprocessing has been accomplished by using steps like sub setting, denoising, atmospheric correction, for getting a clear image. After preprocessing classified the image separately in to separate classes, by using unsupervised classification, it is an automatic algorithm, and classes are categorized separately in to high yield, low yield, moderate, and severe, based on the chlorophyll content. Each class is indicating as disease infestation levels. The Fig. 4.b shows the end member collection spectra (Spectral Profile) of each infestation levels, where x axis is plotted as band number and Y axis plotted as Digital value (DN). The information contained in the each peak of the graph is 548.9200, 742.2500, 772.7800, 854.1800, 1027.1600, 1033.4900, 1346.2500nm wavelength range. The information contained in the each wavelength ranges are shown in the TABLE III. The highest peak of the graph is in between the 40-60 bands (854.1800nm), where the spectral reflectance is increasing (chlorophyll content increases) and decreasing the infestation level, it denotes at the crop is healthy. The second level has comparatively low chlorophyll content, so the chances of disease infested rice crop increasing, and in the third and fourth level the chlorophyll content remains low, when compared with the above two levels.

Table.3.Information contained in different wavelength range

Hyperion Band	Average Wavelength (nm)	FWHM, Full Width at Half Maximum(nm)	Spatial Resolution (m)
B20	548.9200	11.0245	30
B39	742.2500	10.6933	30
B42	772.7800	10.7907	30
B50	854.1800	11.2816	30
B67	1027.1600	11.2754	30
B89	1033.4900	11.0423	30
B12	1346.2500	10.7058	30

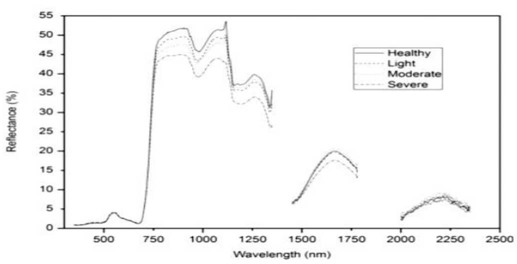


Fig. 4. a Spectral reflectance of various infestation levels are plotted using reflectance and wavelength.

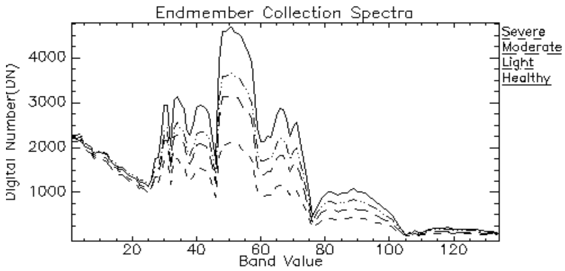


Fig. 4. b Spectral reflectance of various infestation levels are plotted using digital number and band value

When the Fig. 4.a is compared with Fig. 4.b, the reflectance is increasing at the peak of graph is same about 854.1800nm (Band 50), and 1027.1600nm (Band 67) wavelength. In this wavelength region the spectral reflectance is high and low disease infestation level.

Histogram Study: Histogram is a graphical representation of the distribution of numerical data shown in Fig. 5. Using the Histogram plot, it was able to analyze which of the rice disease infestation level is higher in the area. Three bands were (band 49, 21, 28) taken for finding histogram of four disease infestation level. According to the graph the Light healthier paddy crops were commonly seen in the area because of disease infestation level is higher compared to the other levels due to the stress factors, wider climate fluctuations, and shortage of soil parameters.

The most commonly found rice variety in the study area. By statistics the total area of rice cultivation in Palakkad in the year 2010 was 87511ha and the production was 218155tonnes. So the yield per hectare area is 2.4928tonnes/ha (218155ton÷87511ha), then the area of the study sites in Palakkad is 6257.11ha. So the yield of paddy cultivation for 6257.11ha will be 15597.7tonnes (2.4928×6257.11).The Table.2 shows the Yield in tonnes for each class which are 3864.57, 9360.56, 2237.40 & 135.186tonnes for Healthy, Light Moderate & Severe respectively.

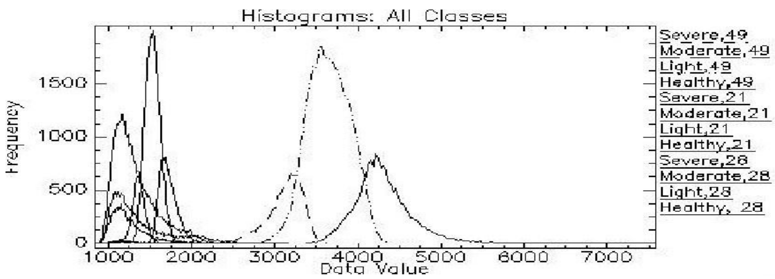


Fig.5.Histogram of all four classes, Healthy, Light, Moderate, and severe at different

4. CONCLUSION

The present work analyzed the use of hyper spctral data for monitoring rice crop yield in Palakkad district, Kerala. Steps like preprocessing, unsupervised classification and assessment of yield using the spectral reflectance has also been carried out. The enormous information contained in the Hyperion data helped to analyze the variations in rice crop identified in the study area. Using K-Mean clustering algorithm these variations have been tracked and four different classes of paddy crop based on its health have been produced. The difference in these classes have been plotted graphically in terms of its spectral signature and quantified in terms of the area in hectares covered by these classes and the corresponding yield per hectare.

5. ACKNOWLEDGMENT

The authors are very much thankful to the Hyperion data provider which is the USGS Earth Explorer site. Thanks to S.Anandraj (Project Associate) for all the technical support. And also thanks to the Department of CEN (Center for Computational Engineering and Networking), Amrita Vishwa Vidyapeetham for the entire lab assistants.

REFERENCES

Datt, Bisun, Preprocessing EO-1 Hyperion hyper spectral data to support the application of agricultural indexes, Geoscience and Remote Sensing, IEEE Transactions, 41 (6), 2003, 1246-1259.

Petropoulos, George P, Kostas Arvanitis, and Nick Sigrimis, Hyperion hyper spectral imagery analysis combined with machine learning classifiers for land use/cover mapping, Expert systems with Applications, 39 (3), 2012, 3800-3809.

Qin, Zhihao, and Minghua Zhang, Detection of rice sheath blight for in-season disease management using multispectral remote sensing, International Journal of Applied Earth Observation and Geo information, 7 (2), 2005, 115-128.

Shwetank S, Kamal Jain, and Karamjit Bhatia, Development of Digital Spectral Library and Supervised Classification of Rice Crop Varieties Using Hyper spectral Image Processing, Asian Journal of Geo informatics, 11 (3), 2012.

Telgarsky, Matus, and Andrea Vattani, Hartigan's Method: k-means clustering without Voronoi, International Conference on Artificial Intelligence and Statistics, 2010.

Wright, Dennis L, Canopy reflectance estimation of wheat nitrogen content for grain protein management, GI Science & Remote Sensing, 41 (4), 2004, 287-300.

Zhao, Jin-Ling, Identification and Characterization of Spectral Response Properties of Rice Canopy Infested by Leaf Folder, International Journal of Agriculture and Biology, 15 (4), 2013, 694-700.