Extraction of Water Body from LANDSAT7 ETM+ Thermal Band using Decision Tree Algorithm

1Ashu Sharma, 2Sharad Kumar Gupta, 3Dr. S. K. Katiyar
1,2Scholar, 3Professor
GIS Division, Civil Engineering Department,
Maulana Azad National Institute of Technology, Bhopal, India
e-mail: ashesh123@gmail.com, sharaqgupta27@gmail.com, skatiyar7@rediffmail.com

Abstract – This research paper investigates the approaches for extracting main water bodies and excluding shadow effect and other features which were reflecting falsely in water using DECISION TREE ALGORITHM. To extract water-body, firstly, all other noninterest earth features (like vegetation, building, soil) were removed by using different indices (e.g. NDVI, Combinations of Different Bands) and other methods based on spectral properties of different earth features. After removing all earth features other than water we discriminated shadow from main water bodies. We used the temperature difference of shadow on land and water bodies by calculating the temperature from thermal band of ETM+ of Landsat. Finally model was developed using ENVI software and water area was extracted and the resultant images were validated using Google Earth.

Keywords: water bodies, decision tree, NDVI.

I. INTRODUCTION
Satellite images acquired in visible & infrared bands are being used extensively for Environmental monitoring. Landsat ETM+ is a good resource for global change analysis and applications in agriculture, geology, forestry and regional Planning [5]. Water Body extraction from satellite imagery is an important task in natural resources management and monitoring like - River Dynamics, Flood Mapping, Coast Line Change Detection, Water Quality and this requires precise classification of water[3][4].Water has the spectral behaviour of absorbing most of visible and infrared wavelength bands [3]. There are many methods of water body extraction form multiband images. One of them is based on finding the difference between signatures of water and other ground objects based on analysis and using decision tree for delineating the water body [8].

In the literature various researchers have suggested the methodology for water body extraction from satellite images. Min Li et al, 2011 [2] proposed Water body extraction based on Oscillatory Network. Duong, 2012 [7] used spectral pattern analysis for extraction of water body. The developed model is based on 6 Visible Bands b1, b2, b3, b4, b5, b7 and b6 of Thermal Band of ETM+. NDVI [10] was calculated for subtracting vegetation form the image and then band combinations; was applied to subtract the effect of other earth features.[7].

In this research work investigations have been done for discriminating water bodies from other features on earth which was not classified by the separating the earth features like vegetation, buildings, barren land, plow land etc. from the image. As after eliminating these features still the classifier may get confused in discriminating the water by the very small quantity of water filled in a pit or some small low level area and shadows on the images. In this work, the above problem has been addressed by using the thermal band imagery.

II. STUDY AREA AND DATA RESOURCES
Image of Landsat 7 ETM+ acquired on October 29, 2001, of Path 144, Row 48 for Hyderabad City was used for analysis. The Hyderabad city is situated in Andhra Pradesh, India and having Latitude of 17° 23’ 06” N and Longitude of 078° 29’ 12” E.

III. METHODOLOGY
In this work investigations the Decision Tree Algorithm and logic operations for eliminating the earth features at every node of tree have been used [6]. At last after eliminating the earth features at every step one by one we have tried to discriminate the rest of the features on the image from water, which may be giving the false results due to the reflectance value. In the last step the water samples and other features were discriminated on the basis of the temperature. For this we have used the thermal band (band 6 high gain) of ETM + sensor in Landsat 7.

Following are the other steps of research work investigations:

Normalized Difference Vegetation Index (NDVI):
It was developed by Rouse et al. (1973) as the difference between near-infrared (NIR) and red reflectance values received by the sensors normalized over the sum of the two.

\[ NDVI = \frac{(b4 - b3)}{(b4 + b3)} \]
\[ b4= \text{Near Infrared Reflectance} \]
\[ b3 = \text{Red Reflectance} \]

For eliminating the various earth features we have used some inequalities which are as follows for different earth features. [4]

1. For vegetation used inequality is-
\[ NDVI > 0.15 \]  
(1)

2. For eliminating Buildings we have used
\[ b2 - b4 < 10 \]
\[ b4 < b5 \]  
(2)

3. For eliminating Plow land used equation is-
\[ b4- b3 > -1 \]
\[ b5/b4 > 1.06 \]
\[ b2+b3+b4+b5+b7 > 360 \]

Or
b4 - b3 > -1
b5/b4 > 1.12
b2+b4+b5+b7 > 288

(3)

4. For eliminating Cloud equation used is-
b1+b2+b3 + b4+b5 > 400

(4)

5. For eliminating Bare Land equation used is-
b5 > b4
b4 > b3

(5)

6. Now for distinguishing other remaining features from water we have used thermal band and at this node now the tree will classify on the basis of the temperature calculated from thermal band. For this we have extracted water by the below inequality as on the basis of various water samples the temperature of water varies in between the given temperature in Kelvin.

b6 >= 289.734680
b6 <= 298.612946

(6)

Here bi is the ith band of ETM+ sensor.

After applying the above inequality the classifier differentiates the water from the non-water feature which were previously wrongly classified by the decision tree up to using the equations 5.

The flow chart of the overall research work methodology is also shown in Fig-1. In this finally we will get the image I6 in which all other features which were classified as water would be eliminated.

IV. RESULTS AND ANALYSIS

For validation of the proposed work more than 30 points were selected and validated on Google Earth. We found that all those were accurately classified with our proposed methodology.

Fig.1: (Model of Decision Tree).

Fig.2: Location having co-ordinates as 17° 24’ 13.57” N, 78° 29’ 55.16” E shown with Box A.

Fig.3: (Google Earth Image of Box A).
One of such location is having co-ordinates as 17˚ 24’ 13.57” N, 78˚ 29’ 55.16” E and shown in the Box A of Fig. 2. As it was actually under non water area but wrongly classified as water because of reflectance similarity. With the use of proposed thermal band temperature input, the decision tree algorithm was able to distinguish the non-water body from water.

Another sample in support of proposed work is shown in Box B of Fig.4. The location having co-ordinates 17˚ 27 ’ 34.5”N, 78˚ 26’ 5.53”E was actually under non water area but the same was wrongly classified as water.

After validation of other points also, it was found that all those points were accurately classified except some points (e.g. some points at river) where the width of water body was very narrow (less than the resolution of the thermal band of ETM+ sensor).

The advantage of the proposed approach is that the main water bodies can be extracted clearly and the effect of the shadow can also be reduced by incorporating the temperature of land and water. If the shadow is on ground then the temperature at that point will differ from the temperature of water.

V. CONCLUSIONS AND LIMITATIONS OF WORK
The inclusion of ground temperature from thermal band has the capability to enhance the classification accuracy of decision tree algorithm investigated in this research work.

The limitation of proposed methodology is that the spatial resolution of thermal band is not same as the other bands (visible, NIR and SWIR) of ETM+. Because of this reason every pixel classified as water without temperature can be misclassified due to the lack of exact temperature at that particular co-ordinate value.

One more limitation of this approach is that, there is a need to train the samples according to the image which we are having, as the temperature will vary from season to season. Therefore as the image changes the threshold value in equation 6 will be also change. That’s why the equation (6) will not be same for each image. It will vary because of various reasons like- time of the day at the time of acquisition of image, weather condition, year of acquisition of image etc.

REFERENCES

Fig.4: Location having co-ordinates 17˚ 27’ 34.5”N, 78˚ 26’ 5.53”E shown with Box B.

Fig.5: (Google Earth Image of Box B).