International Journal of Engineering & Technology, 7 (1.3) (2018) 161-165

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Mapping the vegetation soil and water region analysis of Tuticorin district using Landsat images

Cynthia. J [1], Suguna. M [2], Senthil. S [3]

[1] *Professor, Department of Computer Science and Engineering, Kumaraguru College of Technology, Coimbatore, Tamilnadu*
[2] *Assistant Professor-II, Department of Computer Science and Engineering, Kumaraguru College of Technology, Coimbatore, Tamilnadu*
[3] *PG Scholar, Department of Computer Science and Engineering, Kumaraguru College of Technology, Coimbatore, Tamilnadu.*
**Email: senthil.tuticorin@yahoo.co.in*

Abstract

Mapping of water bodies, soil and vegetation region from satellite imagery has been widely explored in the recent past. Several approaches have been developed to detect water bodies and identify the soil types from different satellite imagery varying in spatial, spectral, and temporal characteristics. Due to the introduction of a New Operational Land Imager (OLI) sensor on Landsat 8 with a high spectral resolution and improved signal-to-noise ratio, the quality of imagery sensed is increased. Its imagery produces a better result in classifying the soil and water regions. The current study puts forward an approach to map water bodies, soil and vegetation region from a Landsat satellite imagery using the various processing models. In this study, to identify the water region and soil region, we go with water index, vegetation index and soil index measures. By using reflectance bands, it is easy to analyze the water, vegetation and soil regions. The proposed method accurately and quickly discriminated the water, vegetation and soil region from other land cover features.

Keywords: *Remote sensing; Landsat; Water; Index; Soil Index.*

1 Introduction

Remote Sensing plays an key role in providing the land coverage mappings and classification of land cover features which mainly includes vegetation, roads, water bodies etc. A chief use of remotely sensed data is to produce a classification map of the identifiable or meaningful features or classes of land cover types in a scene1. As a result, the chief product is a thematic map with themes such as land use, geology and vegetation types. The concept of image classification is a process by which the basic features of the image are assigned to the classes.
Water, Soil and Vegetation is an essential component of ecosystems for the sustainability of life on earth. It balances ecosystems and maintains climate variation, carbon cycling, etc. It is equally important to humans and other forms of life. Its presence causes increases in forest and grassland, or vice versa, whereas its excess or absence could lead to disasters and extreme land use change. Hence, identification of water bodies, soil and vegetation region is an essential process in science and engineering research. The identification can be useful in various ways, such as estimation of water areas, demarcation of flooded regions, wetland inventories, change detection, and so on. The availability of water helps in the estimation of agricultural land irrigation, productivity, hydropower energy, and many others. Soil surveys are the main information source for sustainable agriculture and land use management. Soil survey mapping units are defined by the soil properties that affect management practices, such as drainage, erosion control, tillage and nutrition, and they involve the whole soil profile. Vegetation survey defines vegetation types and helps to understand differences among them, which is essential for both basic ecological research and applications in biodiversity conservation and environmental monitoring.
n past decades, Landsat satellite sensors have been used for land use classification and water body identification. Therefore, it is necessary to explore the most appropriate and practical water identification methods that take advantage of the improved image quality and use the fewest inputs based on the original OLI bands. By using reflectance bands from Landsat 8 OLI imagery it is easy to analyze the water and soil regions from the Landsat images.

1.1 Remote Sensing Sensors

A remote sensing sensor is a key device that captures data about an object or scene remotely. Since objects (including vegetation, water and soil) have their unique spectral features (reflectance or emission regions), they can be identified from remote sensing imagery according to their unique spectral characteristics. A good case in vegetation and soil mapping by using remote sensing technology is the spectral radiances in the red and near-infrared regions, in addition to others. The radiances in these regions could be incorporated into the spectral vegetation indices, soil indices and water indices that are directly related to the intercepted fraction of photosynthetically active radiation. Over the past half century, remote sensing imagery has been acquired by a range of airborne and space-borne sensors from multispectral sensors to hyperspectral sensors with wavelengths ranging from visible to microwave, with spatial resolutions ranging from sub-meter to kilometers and with temporal frequencies ranging from 30 min to weeks or months. Since different sensors have different spatial, temporal, spectral and radiometric characteristics, the selection of appropriate sensors is very important for mapping vegetation cover. The selection of images acquired by adequate sensors is largely determined by four related factors: (i) the mapping objective, (ii) the cost of images, (iii) the climate conditions (especially atmospheric conditions) and (iv) the technical issues for image interpretation. In the field of vegetation and soil mapping, the most commonly applied sensors include Landsat (mainly TM and ETM+), SPOT, MODIS, NOAA–AVHRR, IKONOS and Quick Bird.

1.2 Literature Survey

In [1], Baker et al proposed a Decision-Tree-Based Model to map wetlands and riparian areas for Landsat Etml Imagery. Classification Tree Analysis (CTA) and Stochastic Gradient Boosting (SGB) decision-tree-based classification algorithms were used to distinguish wetlands and riparian areas from the rest of the landscape. CTA creates a single classification tree using a one-step-look-ahead procedure to reduce variance. SGB uses classification errors to refine tree development and incorporates multiple tree results into a single best classification. The SGB classification (86.0% overall accuracy) was more effective than CTA (73.1% overall accuracy) at detecting a variety of wetlands and riparian zones present on the landscape. In [9], S. S. Panda et al, proposed a novel method for estimation of lake water quality from satellite image. They includes an indirect method of determining the concentrations of chlorophyll-a (chl-a) and suspended matter(SM), two optically active parameters of lake water quality. Radial basis function neural (RBFN) network models are developed to predict the chl-a and SM concentrations in the lake. In [5], Leif G. Olmanson et al, discovered a method for examining the precision of repeated measurements on individual lakes within short time periods. They obtain a mean water clarity at the statewide level also remained stable with an average around 2.25 m from 1985 to 2005.Ustunera et al.[6] proposed a technique Rapid Eye Imagery to classify the type of crop. Radial Basis Function (RBF) kernel was used here for the Support Vector Machines (SVMs) classification. A GIS Framework is used to map the soil for Landsat satellite imagery and Digital Elevation Model (DEM) data [2].In hyperspectral remote sensing of vegetation and agriculture shows significant enhancement over conventional remote sensing, leading to improved and targeted modeling and mapping of specific agricultural characteristics such as: (a) biophysical and biochemical quantities (b) crop type\species (Thenkabail et al., 2013), (c) management and stress factors such as nitrogen deficiency, moisture deficiency, or drought conditions and (d) water use and water productivities (Thenkabail et al., 2013)[10].They carried out to perform supervised and unsupervised techniques on remote sensing data for land cover classification and to evaluate the accuracy result of both classification techniques. The sample point represented 25% of the total study area. The results showed that the overall accuracy for the supervised classification was 90.28% where Kappa statistics was 0.86, while the unsupervised classification result was 80.56% accurate with 0.73 Kappa statistics[4].Park et.al[8],Monitoring and understanding climate-induced changes in the boreal and arctic vegetation is critical to aid in prognosticating their future. In 2014 Feyisa, Gudina et.al, proposed a new Automated Water Extraction Index (AWEI) improving classification accuracy in areas that include shadow and dark surfaces that other classification methods often fail to classify correctly [3]. There main is to investigate the application of the probabilistic-based frequency ratio (FR) model in groundwater potential mapping at Langat basin in Malaysia using geographical information system[7].

2 Description of Dataset

2.1 Study Area

Landsat represents the world's longest continuously acquired collection of space-based moderate-resolution land remote sensing data. Four decades of imagery provides a unique resource for those who work in agriculture, geology, forestry, regional planning, education, mapping, and global change research. Landsat images are also invaluable for emergency response and disaster relief. The origin of the data is from “Image courtesy of the U.S. Geological Survey”. The request id of the data is “0701701038994_00002". On “2017-01-04T09:03:59Z” file is dated. At "05:06:07.3862760Z" the scene is captured. The latitude of upper left corner is 9.72713, longitude of upper left corner is 76.85499, latitude of upper right corner is 9.70972, longitude of upper right corner is78.91719, latitude of lower left corner is 7.63064, longitude of lower left corner is 76.84472, longitude of lower right corner is 7.61703 and the longitude of lower right corner is 78.89556.

2.2 Multispectral Image

A multispectral image acquired by a push-broom system consists of two spatial dimensions (along-track and across track) and one spectral dimension (wavelength). The image is registered by the instrument in a data-cube where the along track dimension y corresponds to the image lines; the across track dimension x is associated to the pixel line; and the spectral dimension λ represents the image bands. The size of the multi spectral data-cube can be written in the form l × p × b, where l is the number of image lines, p is the number of pixels per line, and b is the number of spectral channels. The raw information must be preprocessed in order to estimate TOA reflectance. This allows us to remove in practice the dependence on particular illumination conditions (day of the year and angular configuration) and illumination effects due to rough terrain (cosine correction), since the method is intended to work under many situations. TOA apparent reflectance is estimated according to

$$\rho(x,y,\lambda)=\frac{\pi L(x,y,\lambda)}{\cos(\theta(x,y))\,F_o(\lambda)} \tag{1}$$

where L(x, y, λ) is the provided at-sensor upward radiance at the image location (x, y), F0(λ) is the extraterrestrial instantaneous solar irradiance, and θ(x, y) is the angle between the illumination direction and the vector perpendicular to the surface.

3 Methodology

3.1 Region Segmentation

Image segmentation within recognition is based on a threshold fusion value that is a function of both spectral and shape heterogeneity. The threshold is calculated from user-defined parameters of scale, color, shape, compactness, and smoothness. Scale is a measure of the maximum change in heterogeneity allowable in the merging of two objects. Color and shape are weighted relative to one another such that the sum of their assigned values must be 10. A similar weighting is given to compactness and smoothness. We selected the plane neighborhood and only considered the four pixels that are connected along a plane border as neighbors. The imagery was segmented using a scale of 10 and a ratio of 8:2 for the relative importance of reflectance versus shape and 7:3 for compactness versus smoothness. This combination of parameters was established after several iterations in which the produced segments or objects appeared to capture the shape and size of the landslides within the study area. We selected an equal weighting of the six ETM+ bands and the NDVI channel. This ensured that the segments or objects created were consistent with the spectral and vegetative patterns and were not dependent on the DEM data or the two morphological data layers.

3.2 Water Index Method

The selection of these wavelengths maximizes the reflectance properties of water, that is maximize the typical reflectance of water features by using green wavelengths; minimize the low reflectance of NIR by water features; maximize the high reflectance of NIR by terrestrial vegetation and soil features. The NDWI is expressed as follows:

$$NDWI = \frac{(Green - NIR)}{(Green + NIR)} \qquad (2)$$

Where Green is a green band such as TM band 2, and NIR is a near infrared band such as TM band 4. This index is designed to (1) maximize reflectance of water by using green wavelengths; (2) minimize the low reflectance of NIR by water features; and (3) take advantage of the high reflectance of NIR by vegetation and soil features. As a result, water features have positive values and thus are enhanced, while vegetation and soil usually have zero or negative values and therefore are suppressed. However, the application of the NDWI in water regions with a built-up land background does not achieve its goal as expected. The extracted water information in those regions was often mixed with built-up land noise. This means that many built-up land features also have positive values in the NDWI image.

3.3 Vegetation Index Method

In general, if there is much more reflected radiation in near-infrared wavelengths than in visible wavelengths, then the vegetation in that pixel is likely to be dense and may contain some type of forest. If there is very little difference in the intensity of visible and near-infrared wavelengths reflected, then the vegetation is probably sparse and may consist of grassland, tundra, or desert. The result of this formula is called the Normalized Difference Vegetation Index (NDVI). The equation for NDVI is:

$$NDVI = \frac{(NIR - \mathrm{Re}\,d)}{(NIR + \mathrm{Re}\,d)} \qquad (3)$$

Where NIR is the reflectance or radiance in a near infrared channel and RED is the reflectance or radiance in a visible channel.Figure 1(a) shows the results of the vegetation map in the region. An advantage for using NDVI estimation is that NDVI is a routinely produced product, available globally typically every 10 days using satellite instruments such as the Advanced Very High Resolution Radiometer (AVHRR). Jackson et al. (2002) used this approach for grasslands during previous microwave soil moisture remote sensing investigations because use of NDVI increased the accuracy of soil moisture predictions. Even with these limitations in many situations, the only available option has been to use the NDVI because it was the only information available.

3.4 Soil Index Method

The threshold method is one of the most widely used algorithms for the extraction of water bodies from satellite imagery. The method is based on the fact that the reflected radiance of water in short-wave-infrared (SWIR) band is lower than that of other objects like vegetation, buildings, bare soil, and roads. The equation for MNDWI is

$$MNDWI = \frac{(Green - SWIR)}{(Green + SWIR)} \qquad (4)$$

Where SWIR is the reflectance or radiance in a short wave infrared wavelength channel. For Landsat TM/ETM+, NIR and SWIR correspond to bands 4 and 5, respectively. Classic operational instruments such as the AVHRR did do not include a SWIR band. However, new satellite sensors such as the Moderate Resolution Imaging Sensor (MODIS) on NASA's Terra and Aqua satellites now make such data routinely available. Gao (1996) recommended the use of a SWIR band centered at 1.24 Am, now available on MODIS, for NDWI because this band has similar atmospheric transmittance as the NIR band. Ceccato et al. (2002a,b) developed and tested an index similar to the NDWI using data from the SPOT-vegetation sensor. They and many others have also concluded that indices contrasting the SWIR channel with the NIR channel were sensitive to the mass or volume of water and not to the fractional percentage of water. We utilize Landsat TM and ETM+ imagery for the investigation presented here because of the high spatial resolution (30 m). The higher spatial resolution facilitates coordination of ground data collection with the imagery, thus assuring homogeneity of vegetation conditions within pixels (i.e., no mixtures).

3.5 Region Clustering Algorithm

K-means is one of the simplest unsupervised learning algorithms that classify a given data set into certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. These centroids should be placed in a cunning way, because different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed and an early grouping is done. Again re-calculate k new centroids as barycenter's of the clusters (resulting from the previous step). After having these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid. Repeat the process until centroids do not move any more. In the successive loops, the k centroids change their location step by step. The K-mean algorithm uses the following distance formula to compute the distance of the n data points from their respective jth cluster center

$$j = \sum_{j=1}^{k}\sum_{i=1}^{x}\left\|x_i^{(j)} - c_j\right\|^2 \tag{5}$$

Where $\left\|x_i^{(j)} - c_j\right\|^2$ is a distance measure between a data point $x_i^{(j)}$ and the cluster Centre, c_j. In our study we fixed the k values as 3 for the classes low, medium and high.

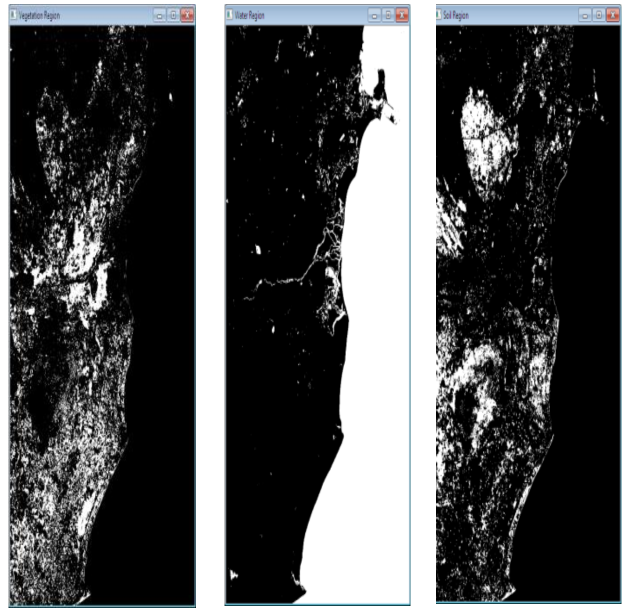
4 Experimental Results and Analysis

In this paper Tuticorin district Satellite image is gathered from the LANSAT. To identify the water region and soil region, we go with water index, vegetation index and soil index measures. By using reflectance bands, it is easy to analyze the water, vegetation and soil regions. The proposed method accurately and quickly discriminated the water, vegetation and soil region from other land cover features. In this model first input is given as RBG,NIR and SWIR. The image is segmented into vegetation, water and soil region by using Region of Intersect (ROI). Further the Clustering of soil and vegetation is carried out.

In general, due to reflection there is variation in vegetation region hence for different time the result may vary figure 1 (a) represent the result of vegetation map. The reflectance properties of water features by green wavelength by using water index method the water region is map as shown in figure 1(b). The reflected radiance of water in short-wave-infrared (SWIR) band is lower than that of other objects like vegetation, buildings, bare soil, and roads. Soil has high intensity then vegetation and water as shown in figure 1(c). Region Clustering methods are used to cluster the vegetation and soil region. Based on the centroid location cluster are created, if the region is dense then there is an possibility of vegetation. Figure 2 (a) shows the region that are cluster to vegetation region. A new cluster is created based on centroids and the regions are grouped together into soil region. Figure 2(b) illustrate the soil region that are gathered from the LANSAT. Finally figure3 shows the overall system model, which contains region mapping and clustering of the regions.

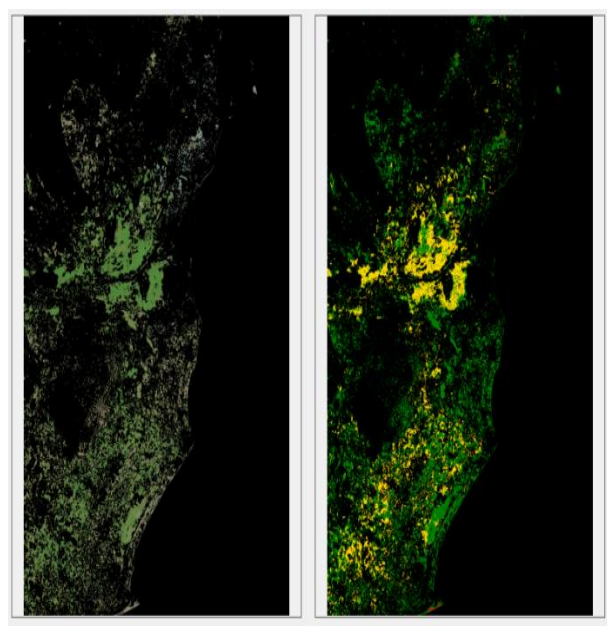
5 Conclusion

In this paper for the classification of water bodies, vegetation and soil region analysis with higher accuracy, we used the NDWI and MNDWI method. Both methods were capable of quickly extracting water, soil and vegetation information and obtained accurate information using an appropriate threshold. This study proved that the proposed method, using water, vegetation and soil index features of the Landsat 8 OLI sensor, obtained higher classification results compared to the TM and ETM+ sensors. This was because the Landsat 8 OLI sensor provides higher SNR imagery than the other sensors.

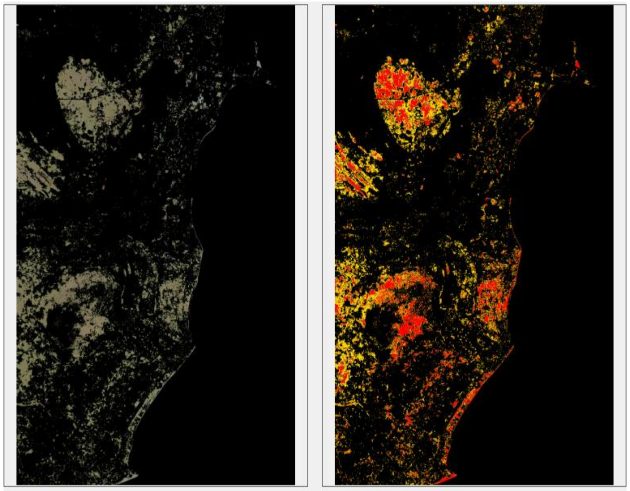


(a) (b) (c)

Figure 1. Experiment result Map of Tuticorin District (a) Vegetation. (b)Water. (c)soil.



(a)



(b)

Figure 2. Experiment result Clustered Region (a) Vegetation. (b)soil.

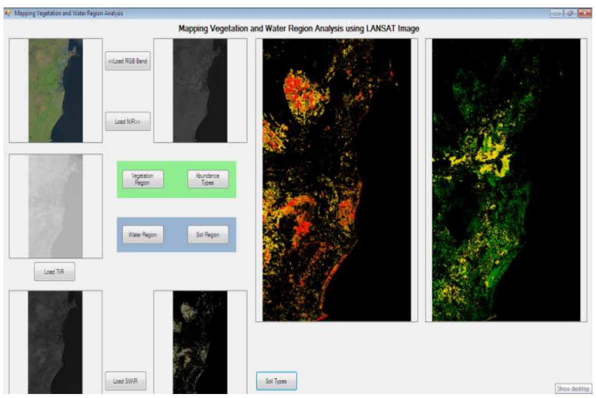


Figure 3. Overall System Model

6 References

[1]. Corey Baker, Rick Lawrence, Clifford Montagne, and Duncan Patten," Mapping Wetlands And Riparian Areas Using Landsat Etm1 Imagery And Decision-Tree-Based Models" WETLANDS, Vol. 26, No. 2, June 2006, pp. 465–474.

[2]. ErtugrulAksoy, GokhanOzsoy and M. SabriDirim," Soil mapping approach in GIS using Landsat satellite imagery and DEM data" African Journal of Agricultural Research Vol. 4 (11), pp. 1295-1302, November, 2009.

[3]. Feyisa, Gudina L., Henrik Meilby, Rasmus Fensholt, and Simon R. Proud. "Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery." Remote Sensing of Environment 140 (2014): 23-35.

[4]. Hasmadi, Mohd, H. Z. Pakhriazad, and M. F. Shahrin. "Evaluating supervised and unsupervised techniques for land cover mapping using remote sensing data." Geografia-Malaysian Journal of Society and Space 5, no. 1 (2017).

[5]. Leif G. Olmanson, Marvin E. Bauer, Patrick L. Brezonik (2008),"A 20-year Landsat water clarity census of Minnesota's 10,000 lakes" Remote Sensing of Environment 112 4086–4097.

[6]. M. Ustunera, F.B.Sanli , S.Abdikan, M.T.Esetlili , Y.Kurucu," Crop Type Classification Using Vegetation Indices Of Rapideye Imagery" The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-7, 2014.

[7]. Manap, Mohamad Abd, Haleh Nampak, Biswajeet Pradhan, Saro Lee, Wan Nor Azmin Sulaiman, and Mohammad Firuz Ramli. "Application of probabilistic-based frequency ratio model in groundwater potential mapping using remote sensing data and GIS." Arabian Journal of Geosciences 7, no. 2 (2014): 711-724.

[8]. Park, Taejin, Sangram Ganguly, Hans Tømmervik, Eugénie S. Euskirchen, Kjell-Arild Høgda, Stein Rune Karlsen, Victor Brovkin, Ramakrishna R. Nemani, and Ranga B. Myneni. "Changes in growing season duration and productivity of northern vegetation inferred from long-term remote sensing data." Environmental Research Letters 11, no. 8 (2016): 084001.

[9]. S. S. Panda, V. Garg and I. Chaubey (2004),"Artificial Neural Networks Application in Lake Water Quality Estimation Using Satellite Imagery"Journal of Environmental Informatics 4 (2) 65-74.

[10]. Thenkabail, Prasad. "Hyperspectral remote sensing of vegetation and agricultural crops." (2017).