COMPARISON OF LANDSAT 8 AND SENTINEL 2 DATA FOR ACCURATE MAPPING OF BUILT-UP AREA AND BARE SOIL

Kapil[1] and Mahesh Pal[2]
[1]Graduate Engineer Trainee, Afcons Infrastructure Limited, Mumbai 400 053
indiakapil1234@gmail.com
[2]Department of Civil Engineering, NIT Kurukshetra, Haryana, 136119
mahesh.pal@nitkkr.ac.in

KEYWORDS: Normalized Built-up Area Index, Band Ratio for Built-up Area, Bare Soil Index, support vector machines

Abstract: Landsat 8 and Sentinel 2 data for February 2016 were used to derive three indices: Normalized Built-up Area Index (NBAI), Band Ratio for Built-up Area (BRBA) and Bare Soil Index (BSI) and used in combination of actual images for accurate discrimination between bare soil and built-up area over Kurukshetra (Haryana). Bands 11 and 12 of Sentinel 2 data were resampled to 10 m resolution so as use them with Bands 2, 3, 4 and 8 for indices calculation and further classification. Three different dataset consisting of (1) three indices only (2) six actual bands and (3) combination of 3 indices and 6 bands for both Landsat 8 and Sentinel 2 data were used for classification using four classes i.e. bare soil, built-up area, water and vegetation. Support vector machine classifier was used to classify different combination of images used in this study. A post classification field visit was also carried using different classified image and a GPS set. Comparison of results in terms of area for both built-up area and bare soil using classified images and field visit suggest that Sentinel 2 data consisting of 6 wavebands and three indices was able to better discriminate both classes (12.99 and 27.72 Km^2) in comparison to 6 bands and three indices (3.5 and 38.98 Km^2) and other combinations using Landsat 8 Data.

1. INTRODUCTION

Impervious surfaces such as roads, parking areas, roof tops of buildings etc in urban areas are manmade features and emerged as an indicator of the degree of urbanization, which itself is a major indicator of environmental quality of an area (Weng, 2007). With the population growth rate of 2.3% and increasing urban population, effective urban land use planning and management are required. Up to 1970ŝs land survey records were mainly used for urban studies, but with the availability of satellite remote sensors, this trend has shifted to the use of data acquired by them. To determine the extent of urban expansion, land use mapping primarily employs the multispectral remote sensing data and auto mated image classification methods. Remote sensing has been successfully used for mapping urban built up area over last two decades which can be a major input for land use planning and resource management (Weng, 2007). However, urban built-up area extraction from moderate spatial resolution satellite data (e.g. Landsat, 30 m spatial resolution) is challenging because of significant intra-urban heterogeneity and spectral confusion between built up area and the bare soil. As the spectral characteristics of urban built up area and bare soil being similar, it may lead to confusion among these classes thus generating erroneous maps. Thus, mapping the built-up and bare soil and their proper differentiation during urban mapping is an important research area because proper differentiation between built-up area and bare soil can be used as an urban development and environmental quality.

Keeping in view the confusion in accurate classification of built up area and bare soil using moderate resolution satellite data, different indices enhancing the chances of better discrimination between built up area and bare soil are proposed in literature (Piyoosh and Ghosh, 2017). Indices are the most common form of spectral enhancement techniques used to improve the classification accuracies of built-up areas and bare soil. These are basically empirical relations which are formed by combining spectral band ratios of satellite image data used for the study under consideration. Normalized difference built-up index (NDBI; Zha et al. 2003), normalized difference soil index (NDSI; Takeuchi and Yasuoka 2004), normalized difference bareness index (NDBaI; Zhao and Chen, 2005), index-based built-up index (IBI; Xu, 2008), normalized difference impervious surface index (NDISI; Xu, 2010), enhanced built-up and bareness index (EBBI; As-Syakur et al., 2012), biophysical composition index (BCI; Deng and Wu 2012), bare soil index (BI; Li and Chen, 2014) and ratio normalized difference soil index (RNDSI; Deng et al. 2015) are some of the indices used by various researchers to improve the predictive accuracy of built up area and bare soil using medium resolution satellite datasets. In spite of availability of several indices it is not possible to justify the usefulness of one or other to correctly differentiate between built-up area and bare soil using Landsat data. Keeping in view this problem, this study aims in using a high resolution (10m) satellite data provided by Sentinel 2 to discriminate built up area from bare soil. To compare the results obtained by Sentinel 2 data, Landsat 8 data of the same area was also used. For both the dataset, 3 indices (Normalized Built up Area Index; NBAI, Band Ratio for Built-up Area; BRBA and Bare Soil Index; BSI) were derived and used in combination with actual datasets to judge their influence on discrimination between built up area and bare soil.

2. DATASET

The study area used for this work covers city of Kurukshetra (Haryana) lying to the north of New Delhi, Capital if India. Landsat-8 and Sentinel-2 satellite imagery obtained from USGS and ESA websites for the year 2016 were used (Table 1). In comparison to the availability of 11 wavebands with Landsat 8 data (bands 1-7, 9: 30m, Band 8:15m, bands 10 and 11:100m spatial resolution), sentinel 2 acquires data in 13 wavebands (bands 2-4,8:10m, bands 5-7,8a,11-12: 20m, bands 1,9, 10:60m spatial resolution). Several band combinations were used to classify images of the study area with both Landsat 8 and Sentinel 2 data. Different image combination obtained using both datasets (Table 1) were classified in four land cover categories: bare soil, vegetation, built up area and water.

Table 1. Dataset used for classification and acquisition dates

Dataset	Date of acquisition	Dataset used for final classification
Sentinel-2	05/02/2016	(1) Bands 2-4, 8 (2) Combination of (1) and BRBA,NBAI and BSI (3) Combination of (1), resampled Bands 11-12 and BRBA,NBAI and BSI
Landsat-8	22/02/2016	6 Band +BRBA+NBAI+BSI

The technique of ratioing bands involves, at its most basic form, dividing the spectral response value of a pixel in one band with the spectral value of the corresponding pixel in another band. This usually helps in suppressing, or normalizing, varying effects such as viewing angles, sun shading, atmospheric effects, soil difference, and so on. It is also applied to maximize sensitivity to the feature of interest, such as the relative health of vegetation. To achieve this, most indices go beyond simple band division to include differencing, weighting, and the introduction of other variables. Keeping in view the usefulness of different indices in differentiating the bare soil and built up areas with Landsat data, three indices as mentioned in Table 2 are used with both dataset in this study. Bands 11 and 12 of Sentinel 2 were under-sampled to a resolution of 10 m to use them for bare soil index calculation and further classification in combination with indices and bands 2-4 and 8.

Table2. Various Built-up area and Bare soil indices used in this study

Indices	Landsat-8 data	Sentinel-2 data
BRBA	B3/B5	B3/B8
NBAI	(B7-B5/B2)/(B7+B5/B2)	(B12-B8/B2)/(B12+B8/B2)
BSI	{(B6+B4)-(B5+B2)}/{(B6+B4)+(B5+B2)}	{(B11+B4)-(B8+B2)}/{(B11+B4)+(B8+B2)}

3. SUPPORT VECTOR MACHINE (SVM)

The SVM is based on statistical learning theory (Vapnik, 1998) and works on the principle of optimal separation of classes. In the case of a two-class pattern recognition problem in which the classes are linearly separable, the SVM selects from among the infinite number of linear decision boundaries the one that minimises the generalisation error. Thus, the selected decision boundary (represented by a hyperplane in feature space) will be one that leaves the greatest margin between the two classes, where margin is defined as the sum of the distances to the hyperplane from the closest cases of the two classes (Vapnik, 1998). For linearly non-separable classes, the restriction that all training data of a given class lie on the same side of the optimal hyperplane can be relaxed by introducing a slack variable. In this case, the SVM works by selecting a hyperplane that maximizes the margin and at the same time minimises a quantity proportional to the number of misclassification errors. For non-linear decision surfaces, the feature vector is mapped into a higher dimensional Euclidean space (feature space) and using the concept of a kernel function to reduce the computational cost.

Several user-defined parameters are required to achieve optimal performance by the SVM classifier for land cover classification. While dealing with multiclass land cover classification problems, choice of a suitable multi-class method, suitable value of regularization parameter (C), type of kernel and kernel specific parameters need to be selected for effective implementation of SVM. For the SVM based classification in this study the óone against oneô multiclass approach, a radial basis kernel function, which has been found to work well with remote sensing datasets, with kernel specific parameter (ɔ), was used. In order to find the optimal value for each of the user-defined parameters with SVM classification algorithm, a trial and error method was used.

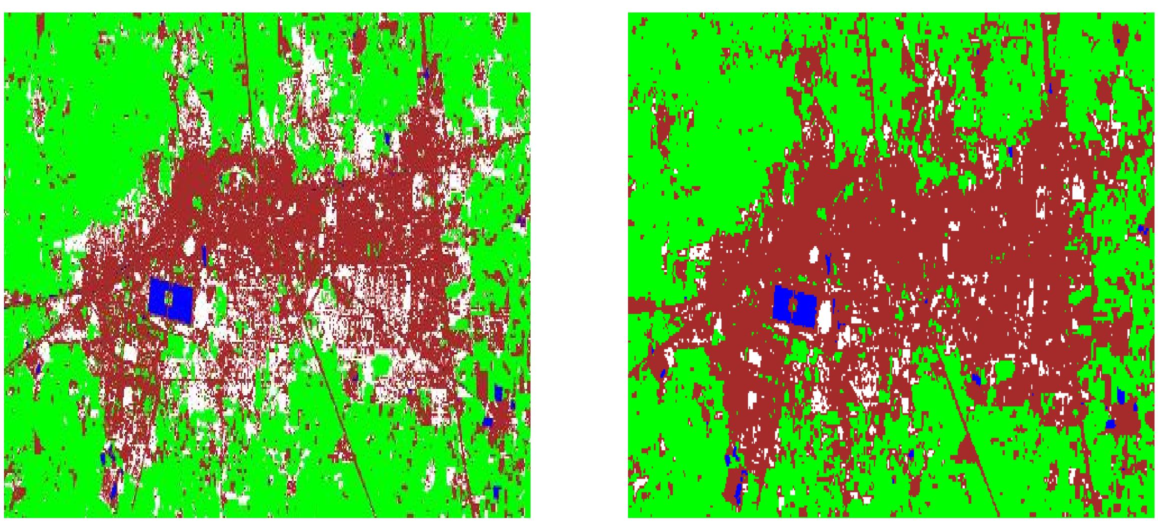
4. RESULTS

Different band combination (Table 1) were used to classify the images acquired over study area using support vector machine classifier in four classs. Area for different classes obtained from various image combinations were calculated and provided in Km2. Classified images from combination 3 of Sentinel 2 and and Landsat 8 data combination are provided in Figure 1.

Table 3. Area calculation from classified images of different datasets

Land cover class		Sentinel 2 (Km2)			Landsat 8 combination (Km2)
		1	2	3	
Bare soil		16.02	28.03	12.99	3.50
Vegetation		47.59	43.80	45.37	44.18
Built up Area		21.36	13.67	27.72	38.98
Water		1.79	1.26	0.69	0.69

Results from Table 3 and field visit to the study area using classified images suggests that area calculation for classes bare soil and built up area as provided by band combination (3) using Sentinel 2 data are found to be more close to the ground data than other band combinations of Sentinel 2 data. Thus suggesting that the use of resampled SWIR bands (Bands 11 and 12) were able to contribute towards improving the classification accuracy of both bare soil and built up area with Sentinel 2 data. Comparison of results with Landsat 8 data (Tables 1 and 3) suggests poor performance by this data combination in discriminating between bare soil and built up area. Figure 1 provides classified images of band combination (3) of Sentinel 2 and Landsat 8 data combination. Comparison of classified images (Figure 1) suggests that the combination of Landsat 8 and derived indices was not able to differentiate bare soil and built up area accurately. Major part of the study area, which is actually bare soil, being classified as built up area (Figure 1(b)) by SVM classifier using Landsat 8 dataset combination. A possible reason of poor performance of Landsat 8 data may be attributed to the spatial resolution of Landsat 8 data vis a vis the size of land parcels in the study area.



(a) (b)

Figure 1. Classified image (a) Sentinel 2 (combination 3) and (b) Landsat 8 with three indices

5. CONCLUSIONS

Satellite data from Sentinel 2 and Landsat 8 was used to discriminate built-up area and bare soil in combination with three indices derived from both datasets. A major conclusion from this study is that Sentinel 2 data performs better than Landsat 8 data in differentiating bare soil from built up area for urban mapping studies using different indices. As present study uses only three built up area and bare soil indices, it is proposed to carry out further studies using several other indices for better differentiation of bare soil from built up area. Further, future study will also include other study areas having more heterogeneity in built up area.

References

As-Syakur, A.R., Adnyana, I., Arthana, I.W. and Nuarsa, I.W., 2012. Enhanced built-up and bareness index (EBBI) for mapping built-up and bare land in an urban area. *Remote Sensing*, *4*(10), pp. 2957-2970.

Deng C, Wu C. 2012. BCI: A biophysical composition index for remote sensing of urban environments. *Remote Sensing of Environment*. 127 (12), pp. 247ï259.

Deng, Y., Wu, C., Li, M. and Chen, R., 2015. RNDSI: A ratio normalized difference soil index for remote sensing of urban/suburban environments. *International Journal of Applied Earth Observation and Geoinformation*, *39*(7), pp. 40-48.

Li, S. and Chen, X., 2014. A new bare-soil index for rapid mapping developing areas using LANDSAT 8 data. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *40*(4), pp. 139-144.

Piyoosh, A. K., & Ghosh, S. K. 2017. Development of a modified bare soil and urban index for Landsat 8 satellite data. *Geocarto International*, http://dx.doi.org/10.1080/10106049.2016.1273401.

Takeuchi W, Yasuoka, Y. 2004. Development of normalized vegetation, soil and water indices derived from satellite remote sensing data. C-9.4 *Remote Sensing Applications*, 25th ACRS 2004, Chiang Mai, Thailand.

Vapnik, V., 1998, Statistical Learning Theory, John Wiley and Sons, New York.

Weng, Q. (Ed.). 2007. *Remote sensing of impervious surfaces*. CRC Press, Boca Raton, Florida.

Xu, H., 2008. A New Index for Delineating Built-Up Land Features in Satellite Imagery. I*nternational Journal of Remote Sensing*. 29914), pp. 4269-4276.

Xu, H., 2010. Analysis of impervious surface and its impact on urban heat environment using the normalized difference impervious surface index (NDISI). *Photogrammetric Engineering & Remote Sensing*. *76*(5), pp. 557-565.

Zhao, H. and Chen, X., 2005. Use of normalized difference bareness index in quickly mapping bare areas from TM/ETM+. *Proceedings of IEEE Geoscience and Remote Sensing Symposium, 2005. IGARSS'05.* Vol. 3, pp. 1666-1668.

Zha, Y., Gao, J., and Ni, S., 2003. Use of Normalized Difference Built-Up Index in Automatically Mapping Urban Areas from TM imagery. *International Journal of Remote Sensing*. 24(3), pp. 583-594.