Vegetation Change Detection of Multispectral Satellite Images Using Remote Sensing

G. Sai Geethika, V. Sai Sreeja, T. Tharuni, and V. Radhesyam

Abstract Change Detection is depicted to compare the spatial representation of two points in time, with variations in the variables of interest causing changes. Remote sensed data of Landsat 8 satellite imagery can be used to detect changes in multispectral images. Vegetation change detection plays a prominent role in tracking the alteration of vegetation in selective areas. The vegetation change analysis of a specific area for a given time period involves a lot of information that can be used for predicting the impact of change over years. Multispectral images for vegetation change of Landsat 8 satellite for a specific location can be obtained by stacking selective bands together. Over the years, due to urbanization, the vegetative index of cities dropped drastically. To take necessary measures, the impact must be analyzed. To overcome this problem, we are using vegetation change detection analysis. The image stacking is performed using QGIS (Quantum Geographic Information System) software which is facilitated using various image enhancement options. The location Vijayawada is tracked for detecting change over a time period of eight years using the NDTS (Normalized Difference Between Time Series) algorithm followed by comparison with the PCA and K-means algorithm. This paper gives a detailed visualization of the results acquired in this project using metrics like RMSE and PSNR.

Keywords Landsat 8 · Remote sensed data · Multispectral images · Satellite imagery · Normalized difference between time series (NDTS) · PCA (principal component analysis) · K-means · Vegetative index · QGIS (quantum geographic information system)

G. S. Geethika (✉) · V. S. Sreeja · T. Tharuni · V. Radhesyam
Department of Information Technology, VR Siddhartha Engineering College, Vijayawada, India
e-mail: saigeethikagudapati@gmail.com

1   Introduction

Satellites gather an enormous amount of data every day. This data comes into use when the data is processed to gain useful insights. The essence of satellite imagery is that it gives the benefit of accessing remotely sensed data [1]. Remote sensed data is the relevant land cover data that is accessed anywhere irrespective of time and location. There is no requirement of visiting a particular location to explore the data. The data which is gathered in such a way can be used for analyzing the vegetation cover over a period of time. This information gives us the insights to predict cases like the increase or depletion of vegetation. The scenarios where depletion of vegetation cover is observed can be used for researching the drastic impacts on nature as a result.

Change detection techniques are not limited to land cover changes. The selective bands give us the detail of what is being detected. A few other observations that can be explored using change detection techniques are changes in land cover, and water bodies like rivers, lakes, and oceans. Predicting the differences using computer-based algorithms makes the process of finding the change faster and more efficient with improvements and enhancements that can be accommodated easily. The ecosystem stays balanced when there is a significant index of vegetation and forest cover.

Over 26% of the vegetation area was lost to habitation between the years 2005 and 2020. It is clear that the approaching era will be undesirable if these patterns of land use change persist. The majority of the land is covered by built-up areas (industries, commercial areas), as opposed to agricultural areas, however, agricultural areas shouldn't be thought for conversion to built-up areas [2]. Digital agricultural applications are crucial for tracking remote harvest and assessing the state of farmlands. As a result of their effectiveness in identifying land cover components, high-resolution pictures have drawn much more attention [3]. This estimate can be applied mainly to urban areas where the vegetation index decreases due to the increase in population. But it is not limited to any region.

The NDVI approach is commonly used to detect changes in the amount and distribution of vegetation, as well as in land use and land cover. Thanks to advancements in the spatial and spectral (from broadband range to small range) resolution of remote sensing data, it is now possible to work at the micro level [4].

1.1   Literature Study

According to Sumanta Bid, it is observed that NDVI technique with specific thresholds can detect changes in the vegetation of a place using remote sensing [4]. You Y, Cao J, Zhou W produced research article focused on different kinds of change detection techniques that can be used for urban change detection [5]. Song A, Choi J, Han Y, Kim Y research paper included change detection in hyperspectral imagery by using convolutional neural networks [6, 7]. Previous research on change detection

for vegetation on multispectral images is either primarily focused on a single algorithm like NDVI or focused on urban changes and hyperspectral images. But there is a need to detect vegetation change accurately using satellite images and compare it to other techniques to identify the best technique.

Our proposed study provides extended research by comparison study of K-means and NDVI for vegetation change detection on multispectral satellite images and determines the performance of both algorithms primarily for detecting accurate changes in vegetation over years.

2 Proposed Work

Change detection identifies the spatial changes induced by man-made or natural processes in multi-temporal satellite images. Remote sensing for change detection can help with urban planning and expansion by considerably improving land utilization. The objective of the project is to gather Landsat 8 satellite images and apply NDVI and compare results with PCA and K-means between 2 timestamps of the Vijayawada area.

2.1 Datasets

The dataset is taken from earth explorer [8] with USGS credentials. The images are located with a specified coordinate set enclosed in circular shape. The image radius considered for this project is 250 m near Vijayawada which can be observed in Fig. 1. The datasets that are available are displayed as search results. There are alternative variations for each dataset which include the year, cloud cover, area covered, and temperature constraints.

The selected dataset can be viewed with a footprint market on the map which can be observed in Fig. 2 and the specification can be viewed in Table 1. The image is displayed as a tile on the map. When the dataset is downloaded, the bands that are available for the Landsat 8 satellite are displayed [9]. To download the data, USGS account credentials are necessary. The data collected is Landsat Collection 2 with Level-2 data including Landsat 8–9 OLI (Operational Land Imager)/TIRS C2 L2 as specifications. The image data with feature class as vegetation features with minimum cloud cover is selected to acquire accurate results.

2.2 Data Extraction and Pre-processing

QGIS acts as GIS (Geographic Information System) software which allows different users to explore and edit spatial data. Beyond that, it allows composing and exporting

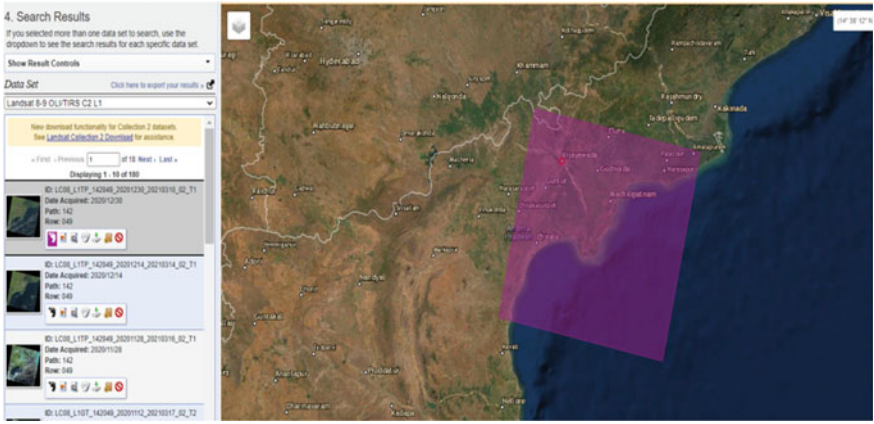**Fig. 1** Image location view in earth explorer

Fig. 2 Image location tile view in earth explorer

Table 1 Feature details of the dataset

Feature	Value
Geocoding method	Feature (GNIS)
Name	Vijayawada
Country	India
Center longitude	80.5991
Radius	250 m

various kinds of maps. The images are selected and pre-processed using QGIS Semi-automatic Classification Plugin (SCP) [10]. The images are cropped according to required coordinates to map a specific location.

2.3 Design Methodology

The multispectral images that are collected in the area of Vijayawada with timestamps 2013 and 2021 are collected as eleven different bands individually as Landsat Collection 2 Level 1 data which can be viewed in Fig. 3. The specification for level 1 data is considered as Landsat 8–9 OLI/TIRS C2L1. The bands that correspond to different change detection are listed in Table 2.

For vegetation analysis, bands 4, 5, 6 are considered. These image bands are stacked using QGIS software and pre-processed to remove noise. Our first change detection approach is calculating NDVI, and second approach is PCA and K-means.
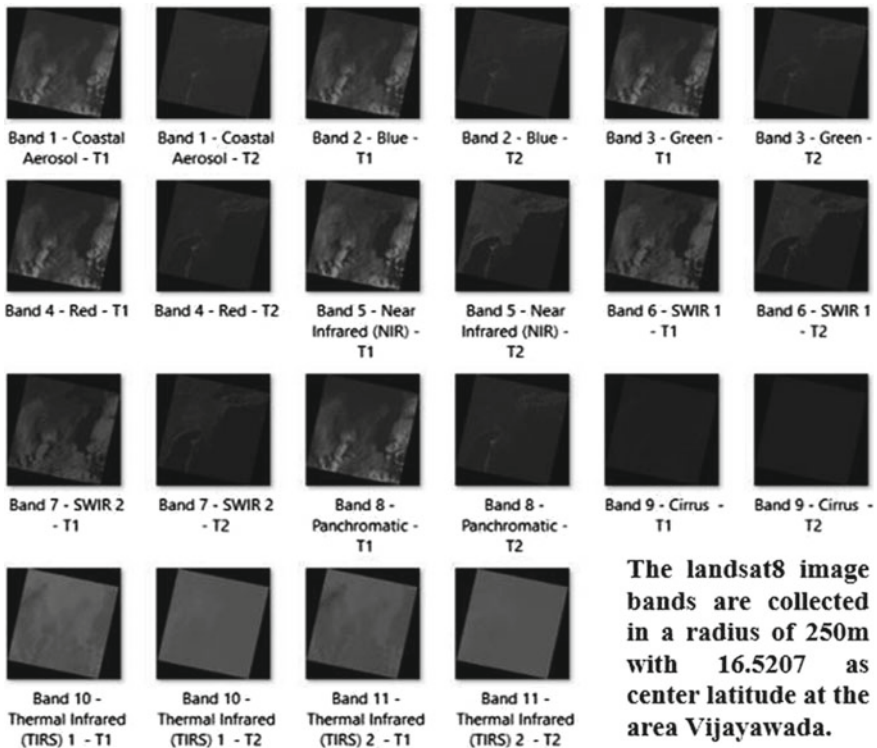

Fig. 3 Image bands of two timestamps

Table 2 Landsat 8 band designations

Sr. no.	Bands	Wavelength (μm)
1	Coastal aerosol	0.43–0.45
2	Blue	0.45–0.51
3	Green	0.53–0.59
4	Red	0.64–0.67
5	Near Infrared	0.85–0.88
6	SWIR1	1.57–1.65
7	SWIR2	2.11–2.29
8	Panchromatic	0.50–0.68
9	Cirrus	1.36–1.38
10	TIRS1	10.60–11.19
11	TIRS2	11.50–12.51

2.4 Procedure

NDTS

First step is importing libraries that are needed to open Geo TIFF format images. The libraries required are osgeo, gdalconst, NumPy, SciPy, IPython, matplotlib which are built-in libraries in python. Load the datasets after pre-processing using QGIS SCP plugin. The datasets used here are the stacked images. The images are loaded, and the files are named to apply the functions that are available in the libraries. Editing spatial data in QGIS can be enabled using digitized tools [11].

Plot the histograms for both bands. By taking multiple data points and organizing them into logical ranges or bins for the pixel values, the histogram condenses the given picture data series into an easily understood visual. NDTS can be computed in vegetation as NDVI with 0.1 as threshold value.

$$\text{NDTS} = \frac{(I_{t2} - I_{t1})}{(I_{t2} + I_{t1})} \tag{1}$$

Here, I_{t2} represents Image at timestamp t_1 (2013) with red reflectance and I_{t2} represents Image at timestamp t_2 (2021) with near infrared reflectance.

The NDVI approach is a straightforward arithmetical indicator that may be applied to remote sensing readings to determine whether or not the target or object being examined has significant vegetation. The array values of two images are used to find NDVI by using 0.1 as threshold value to identify the best satellite image differentiating using the algorithm which is specifically applied to calculate Normalized Difference Vegetation Index in the images. The connection between fluctuations in vegetation growth and spectral variations rate has been extensively studied using the NDVI (Normalized Difference in Vegetation Index). Determining the growth of green vegetation and spotting changes in the vegetation are both useful [12].

The image files are obtained as an output after calculation of NDTS. The square of values is calculated for comparison with the original image. The square of NDTS is calculated to increase scope of detecting change between marginal values that are estimated during NDTS calculation and remove noise using 3×3 mode filter. This improves the image accuracy and resolution to display changes.

Display the changes along with the detected output graph. The output is displayed in the form of a histogram and change map. The change map indicates the change using two colors. The black color signifies no change, and the white color indicates change observed. NDVI signifies vegetation change detection in two images.

Principal Component Analysis (PCA)

The primary purpose of PCA is to reduce dimensionality. A high number of variables are condensed into a smaller number in order to minimize the dimensionality of large data sets while retaining the majority of data.

PCA includes the following operations:

The first action is standardization. This stage involves normalizing the range of continuous beginning variables such that each one contributes equally to the analysis. Mathematically, the mean for each value of each variable may be subtracted, and it can be divided by standard deviation.

$$Z = \frac{\text{value} - \text{mean}}{\text{standard deviation}} \tag{2}$$

Computation of the covariance matrix. In this stage, we try to figure out how the variables in the given input dataset vary from mean about each other, or if there is any link between them.

$$\begin{pmatrix} \text{cov}(a, a) & \text{cov}(a, b) & \text{cov}(a, c) \\ \text{cov}(b, a) & \text{cov}(b, b) & \text{cov}(b, c) \\ \text{cov}(c, a) & \text{cov}(c, b) & \text{cov}(c, c) \end{pmatrix} \tag{3}$$

Calculate the eigenvectors and eigenvalues of the covariance matrix to find the primary components. By linearly integrating the initial variables, principal components are formed as completely new variables. n major components are given from n-dimensional data. The ten main components are thus provided by 10-dimensional data. The second component, however, has less information than the first component, since PCA concentrates the most information on to the first component.

The last step is to select the feature vector. The major components of this feature vector are quite important. Low-importance principal elements may be discarded. We can determine the primary components in terms of importance by computing the eigenvectors and sorting them by their eigenvalues in descending order.

K-Means Clustering

By using the K-Means clustering method, an informative index is divided into K unique, non-covering groups and it can be used for change detection [13]. Characterize the number of bunches you require (*K*) before using K-Means grouping. The K-implies calculation will then assign each perception to one of the K groups [14].

The K-Means uses a centroid that minimizes the idleness between the points. It can be represented by below equation

$$\sum_{i=0}^{n} \min(||x_i - \mu_j||)(||x_i - \mu_j||) \tag{4}$$

Determine the number of clusters (*K*) and then randomly assign *K* different centroid locations. Finding the Euclidean distance across each point and the centroid is the following step. Each point should be assigned to the closest cluster before the cluster mean is determined as the new centroid. After the new point is assigned, the new centroid's position (*X*, *Y*) is:

$$X = \frac{(x_1 + x_2 + x_3 + x_4 + \cdots + x_{n-1} + x_n)}{n} \tag{5}$$

$$Y = \frac{(y_1 + y_2 + y_3 + y_4 + \cdots + y_{n-1} + y_n)}{n} \tag{6}$$

2.5 System Architecture

The architecture of the entire change detection system can be seen in Fig. 4.

3 Results and Observations

Different band combinations can be observed in Table 3. The bands selected for vegetation change detection are Red, Near Infrared, and Short-wave Infrared 1. The wavelength required for the detection of change ranges from 0.64 to 1.65 μm. The image resolution is around 30 for every image band of Landsat 8 [15].

The image features are displayed in the form of a dense peak with values from negative value range to positive value range. A change map indicates the change between two images at different timestamps. From the analysis of the vegetation index in the area of Vijayawada, there is a significant decrease in the vegetation of the city. The reports suggest that this decrease in vegetation index is due to urbanization and the development of cities due to the increase in population over the past few years.
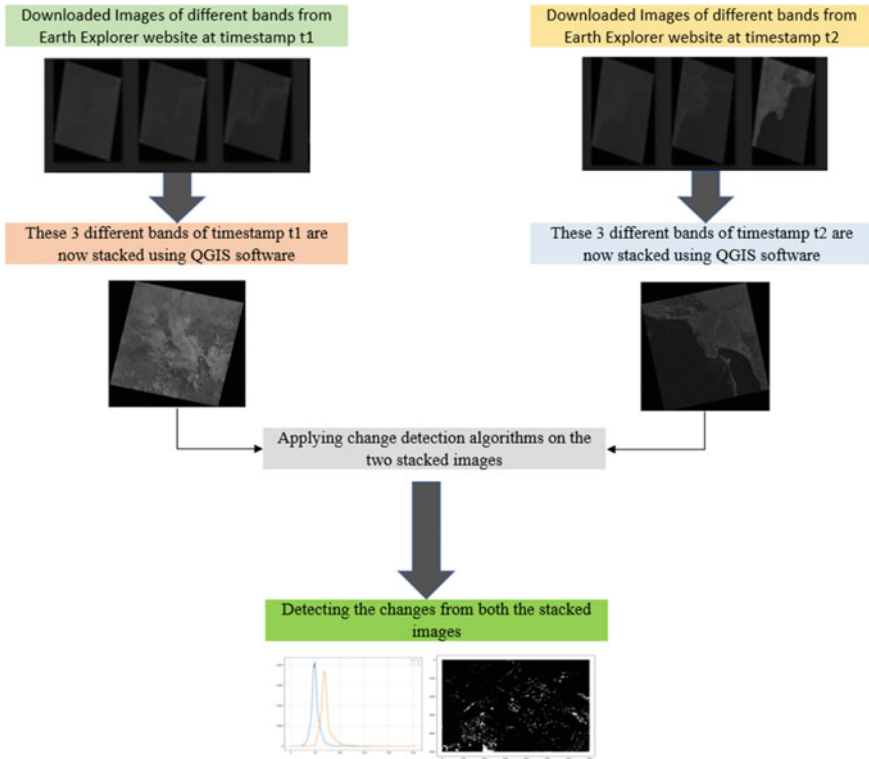
Fig. 4 System architecture diagram

Table 3 Bands associated with landscape

Property	Bands
Natural color band	4, 3, 2
False color (urban) band	7, 6, 4
Color infrared (vegetation) band	5, 4, 3
Agriculture band	6, 5, 2
Atmospheric penetration band	7, 6, 5
Healthy vegetation band	5, 6, 2
Land/water band	5, 6, 4
Natural with atmospheric removal band	7, 5, 3
Short-wave infrared band	7, 5, 4
Vegetation analysis band	6, 5, 4

This caused a hike in agricultural area removal and deforestation which decreased the vegetation index to a lower rate. The change map and associated graph results convey where the exact change is observed which gives the estimate of the vegetative index between the past few years.

RMSE is very helpful to measure a model's performance, during training, cross-validation, or even monitoring after deployment. The root mean square error is popularly used metrics for this. It is a reasonable rating scale that adheres to some of the most popular statistical hypotheses and is simple to understand. $y(i)$ is the ith measurement, $y^\wedge(i)$ is its corresponding forecast, and N is the total number of data points.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} \| y(i) - y \wedge (i) \|^2}{N}} \tag{7}$$

Image compression quality is also contrasted by the mean square error (MSE) and peak signal-to-noise ratio PSNR. PSNR indicates a measure of the peak error, whereas the MSE represents the cumulative squared error among original and compressed images. The error is inversely correlated with the value of MSE.

$$\text{PSNR} = 10 \log_{10}\left(\frac{R^2}{\text{MSE}}\right) \tag{8}$$

The histogram presented in Fig. 5 shows how the images are differentiated with each other. Blue curve indicates stacked image 1 and orange curve indicates stacked image 2. The change maps in Figs. 6 and 7 indicate particulary where the change is detected. White indicates change observed and black area indicates no change. The evaluation metrics obtained by calculating RMSE (Root Mean Square Error) is 0.0043071182 for NDTS and 0.0151108205 for K-means. It is used to calculate the variation between source image and obtained image. The metrics obtained by calculating PSNR (Peak Signal–Noise Ratio) can be seen in Table 4, which is 47.31626752546268 for NDTS and 36.41423705747562 for K-means. If PSNR is higher, then the image is of higher quality. If RMSE is higher, the produced image will be of lesser quality. Hence NDTS shows best change results compared to K-means.

4 Conclusion

The project mainly deals with the change of vegetation. The analysis is done using remote sensing techniques with the aid of satellite imagery which gives accurate results in less amount of time. The extended and depleted vegetation gives the idea of how the land cover is being used over the years. As the estimation of change using change detection algorithms overcomes the problem of identifying the location where the plantation is required, it helps various organizations to take steps toward

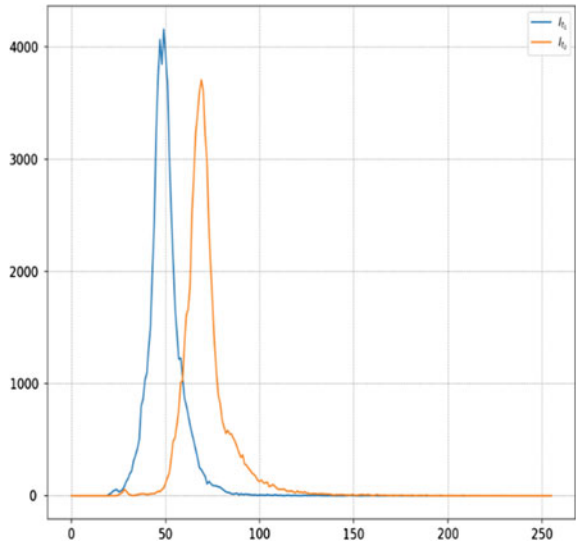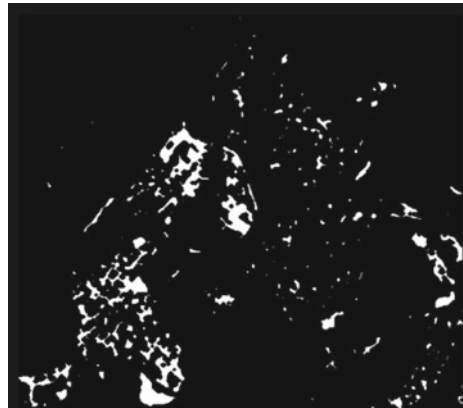**Fig. 5** Histogram depicting change between two images

Fig. 6 Change map obtained by NDTS

growth and control over vegetation. The project mainly dealt with vegetation change detection in multispectral images, and it can be concluded that NDTS performs better than PCA k-means for vegetation change detection.

The future study of this project can be extended for land cover change detection and water body change detection. Research using change detection algorithms can help to assess the situation of increase in water level over the years through global warming. But the process can be more perspicuous if the work is extended to hyperspectral images. The algorithms that are newly employed using complex deep neural networks can give a presumable spike in accuracy and decrease the time taken for the process.

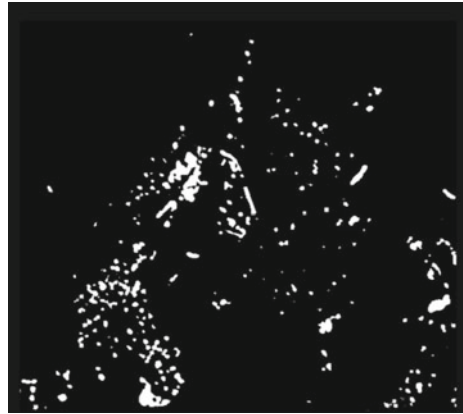**Fig. 7** Change map
obtained by PCA K-means

Table 4 Metrics obtained

Algorithm/metrics	RMSE	PSNR
NDTS	0.0043071182	47.31626752546268
PCA and K-means	0.0151108205	36.41423705747562

References

1. Navalgund RR, Jayaraman V, Roy PS (2007) Remote sensing applications: an overview. Curr Sci, pp 1747–1766
2. Prasad AS, Ramamurthy C, Krishna KL (2022) Land use and land cover change and sustainability assessment of Vijayawada city by RS&GIS. In: IOP conference series: earth and environmental science, vol 982, No 1. IOP Publishing
3. Joseph S (2021) Accurate segmentation for low resolution satellite images by discriminative generative adversarial network for identifying agriculture fields. J Innov Image Process 3(4):298–310
4. Bid S (2016) Change detection of vegetation cover by NDVI technique on catchment area of the Panchet Hill Dam, India. Int J Regul Gov 2(3):11–20
5. You Y, Cao J, Zhou W (2020) A survey of change detection methods based on remote sensing images for multi-source and multi-objective scenarios. Remote Sens 12(15):2460
6. Mateen M, Wen J, Akbar MA (2018) The role of hyperspectral imaging: a literature review. Int J Adv Comput Sci Appl 9(8)
7. Song A et al (2018) Change detection in hyperspectral images using recurrent 3D fully convolutional networks. Remote Sens 10(11):1827
8. Bensi P et al (2007) A new earth explorer-the third cycle of core earth explorers. ESA Bull 131:30–36
9. Ridwan MA et al (2018) Applications of landsat-8 data: a survey
10. Tobias MM, Mandel AI (2021) Literature mapper: a QGIS plugin for georeferencing citations in Zotero. Air Soil Water Res 14:11786221211009208
11. Graser A (2016) Learning Qgis. Packt Publishing Ltd
12. Gandhi GM et al (2015) Ndvi: vegetation change detection using remote sensing and gis–A case study of Vellore District. Procedia Comput Sci 57:1199–1210
13. Vignesh T, Thyagharajan KK, Ramya K (2019) Change detection using deep learning and machine learning techniques for multispectral satellite images. Int J Innov Technolo Expl Eng 9(1S):90–93

14. Lv Z et al (2019) Novel land cover change detection method based on K-means clustering and adaptive majority voting using bitemporal remote sensing images. Ieee Access 7:34425–34437
15. He G et al (2018) Generation of ready to use (RTU) products over China based on Landsat series data. Big Earth Data 2(1):56–64