ORIGINAL PAPER

A novel fuzzy Harris hawks optimization-based supervised vegetation and bare soil prediction system for Javadi Hills, India

Sam Navin MohanRajan[1] · Agilandeeswari Loganathan[2]

Abstract

For several decades, researchers throughout the world have been motivated and contributed to the research on land use/ land cover (LU/LC) change prediction analysis. This research work builds the ensemble learning model using unsupervised metaheuristic optimization and supervised machine learning algorithms for obtaining an LU/LC classification and prediction map in the area of Javadi Hills, India. The unsupervised fuzzy Harris hawks optimization (HHO) algorithm was used for finding the unknown features from the pre-processed satellite image. The feature-extracted map was used as an input for finding the LU/LC classes by using the supervised machine learning classifiers (support vector machine, random forest, and maximum likelihood). The principal component analysis (PCA) has been used for fusing the results of the supervised machine learning classifiers. The fused results were combined with the Adjusted Vegetation and Bareness Index (AVBaI) map and the final ensemble LU/LC (E-LU/LC) map was achieved for the years 2012 and 2015 with good classification accuracy, with an average of 95.275%. The impact of the LU/LC changes through the Land Surface Temperature (LST) map has been calculated and used as the input along with the E-LU/LC map during the process of LU/LC prediction. The ensemble-based prediction (EP)-LU/LC map for the years 2018, 2021, 2024, and 2027 has been forecasted by using the Markovian-Cellular Automata with the Multilayer Perceptron Neural Network model. The result of the EP-LU/LC map provides an average prediction accuracy of 95.763%. Our research on the LU/LC prediction will assist the concerned government officials in taking essential measures for protecting the LU/LC environment.

Keywords Land use/land cover · Harris hawks optimization · Principal component analysis · Adjusted vegetation and bareness Index · Land surface temperature · Markovian-cellular automata · Multilayer perceptron neural network

Introduction

In the developing technology of science, the area of remote sensing (RS) provides the advantage of acquiring information about the surface of the earth in the past, present, and also for future. The RS is the knowledge of finding information about the physical features of an area on the earth's surface through the measurement of reflected radiations from a particular distance by using a satellite or aircraft. The geospatial satellite RS technology plays a dynamic role for RS researchers in many aspects like geo-structural mapping, monitoring of the soil moisture, vegetation, forest fires, road monitoring, floods and droughts, deforestation, LU/LC, or urban planning (Yuan et al. 2020, Weiss et al. 2020, Wellmann et al. 2020, Bauer 2020, Cheng et al. 2020). The LU/LC prediction has been considered as needed and important research in the field of RS. Land cover refers to the earth's surface cover areas like urban infrastructure, water bodies, bare soil, vegetation, agricultural lands, mountain regions, and forest-covered area. Land use refers to the utilization of the land by people on earth for different socio-economic activities. The information about the LU/LC over several decades around the world has been captured and stored in the database of different satellites and the information has been sensed by scientists for performing research on LU/LC prediction. The importance of LU/LC prediction

Responsible Editor: Biswajeet Pradhan

✉ Agilandeeswari Loganathan
  agila.l@vit.ac.in

  Sam Navin MohanRajan
  samnavin@karunya.edu

[1] School of Computer Science and Technology, Karunya Institute of Technology and Sciences, Coimbatore, India

[2] School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India

research is to provide the LU/LC information for the particular area to the government officials, forest department, urban planners, and land resource management to take necessary actions for the protection of the LU/LC environment (Mohan et al. 2021, Vivekananda et al. 2021; Talukdar et al. 2020). From the observations of many researchers, the research on the LU/LC prediction is processed with the following steps that include data acquisition, pre-processing, LU/LC classification, accuracy assessment for validating the classified map, LU/LC change detection, modeling of environmental variables that affecting the LU/LC environment, LU/LC prediction, and the validation of the LU/LC prediction. Different algorithms and techniques have been used by researchers for predicting the LU/LC map for the specific area on the surface of the earth (Navin and Agilandeeswari 2020a, Ridding et al. 2020, Hasan et al. 2020, Li et al. 2020, Alharthi et al. 2020).

In the scientific field of RS, the data was collected through satellite systems or airborne sensors with different spectral, spatial, temporal, and radiometric resolutions. The use of multispectral data has an advantage in providing higher accuracy in the research of LU/LC prediction. A few of the high-resolution multispectral geospatial data used by researchers around the world include Sentinel, Landsat Series, IKONOS, SPOT, REIS (Rapid Eye Earth Imaging System), Linear Imaging Self-Scanning sensors (LISS-III and LISS-IV), Sentinel, Moderate Resolution Imaging Spectroradiometer (MODIS), ASTER Global DEM (Digital Elevation Model), and Cartosat DEM. The other data used for performing the research on LU/LC prediction include aerial photographs, ground survey data, Google Earth Engine (GEE) images, open street maps, and government-dependent data (Kodheli et al. 2020; Feng et al. 2020; Li et al. 2021). The noise, geometric, atmospheric, and other sensor errors present in the raw geospatial satellite images are corrected through the process of pre-processing. The most used techniques by researchers for the rectifying of geospatial satellite images are image de-striping, Quick Atmospheric Correction (QUAC) module, Dark Object Subtraction (DOS) module, Apparent Reflectance Model (ARM), ASCII Coordinate Conversion, georeferencing, FLAASH (Fast Line-of-sight Atmospheric Analysis of Hypercubes) module, orthorectification, wavelet dimensionality reduction, rescaling, histogram equalization resampling, Linear Discriminant Analysis (LDA), F mask method, and Minimum Noise Fraction (MNF) method (Tamiminia et al. 2020; Navin and Agilandeeswari 2020b; Kuchkorov et al. 2020). The features in the geospatial satellite images have been extracted using different unsupervised algorithms and the pixels in each feature help the researchers in performing further processing in obtaining information about the LU/LC classes. The unsupervised optimization technique has been widely used in the process of digital image processing for extracting features from geospatial satellite data. For extracting the features from the geospatial satellite image, optimization algorithms have been introduced for the clustering process. The researchers used the clustering objective as the minimum sum of the squared error for the optimization process. Based on the clustering objective, some of the optimization algorithms used for extracting the features from the geospatial satellite images are Genetic Algorithms (GA), Cuckoo Search Optimization (CSO), Ant Colony Optimization (ACO), Artificial Bee Colony (ABC), Shuffled Frog Leaping Optimization (SFLO), HHO, and Particle Swarm Optimization (PSO) (Sawant and Prabukumar 2020, Sheoran et al. 2020, Thyagharajan and Vignesh 2019, Sheoran et al. 2021, Rodríguez-Esparza et al. 2020, Singh et al. 2020).

The unsupervised classification algorithms are used for classifying the unknown features in an image into clusters or groups without any training process. Some of the most used unsupervised classification algorithms used by researchers in the field of RS are the Gaussian Mixture model, K-means, Fuzzy C-means (FCM), and ISODATA (Iterative Self-Organizing Data Analysis) clustering. The supervised classification algorithms are used for classifying the image by training the known features of an image with the reference image for obtaining good classification accuracy. Some of the most used supervised classification algorithms used by researchers in the field of RS are support vector machine (SVM), Mahalanobis distance, random forest classification (RFC), multivariate adaptive regression spline (MARS), parallelepiped classification, maximum likelihood classification (MLC), spectral angle mapper (SAM), k-nearest neighbor (kNN), and minimum distance to mean classification (Reddy and Kumar 2022, Singh et al. 2021, Alshari and Bharti 2021, Kulkarni et al. 2020, Yashin et al. 2020, Balha et al. 2021). The ground truth data is compared with the LU/LC classified data for calculating accuracy. Based on the accuracy results of the classification algorithms, the LU/LC classification map has been processed for further processing. The LU/LC change detection has been calculated for the classified LU/LC map between different periods (Chughtai et al. 2021; Mishra and Jabin 2020).

Some of the independent variables that provide the impact on LU/LC change include slope, aspect, elevation, precipitation data, hill shade, Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), Normalized Difference Soil Index (NDSI), census data, and distance variables (distance from edge of the forest, road, urban area, agricultural land, and water bodies) (Loganathan et al. 2022, Zeferino et al. 2020, Shafizadeh-Moghadam et al. 2020, Kayet et al. 2021). Based on the inputs of dependent and independent variables of the specific location, the LU/LC prediction was performed. The LU/LC change classification map for the specific study area was considered the dependent variable, and the factors that affect the LU/LC map of the study area were considered the independent variables. In the developing field of science and technology, the RS researchers

used different prediction algorithms for finding future LU/LC changes around the world. Some of the prediction methods used by the researchers are Artificial Neural Network (ANN), Markovian and Cellular Automata (MCA), Logistic Regression (LR), Autoregressive Integrated Moving Average Model (ARIMA), Long-Short-Term Memory Neural Network (LSTM), CNN (Convolutional Neural Network), and Back Propagation Neural Network (BPNN) (Mohanrajan and Loganathan 2022, Floreano and Moraes 2021, Said et al. 2021, Seydi et al. 2021, Gavrilovskaya et al. 2021, Sankarrao et al. 2021). The scope of the research is based on the performance of the different algorithms used for different application areas. For providing better predictive results, the ensemble system or multi-classifier system learning-based approach was used by the researchers. The ensemble learning system is the process of combining multiple classification models for improving the model's accuracy. The use of ensemble learning is to provide a less error rate on regression and high classification accuracy for statistical intelligence problems. The algorithms used in the ensemble learning system have been fused and the results are attained through the voting and averaging-based methods (Dou et al. 2021, Benbriqa et al. 2021; Phan et al. 2021; Kutlug and Colkesen 2021; Hu et al. 2021).

The environmental changes that have been observed in the past and present have been a concern for the future because of the worsening of nature and its impact on human health. The findings of the LU/LC changes for the past, present, and future were used for the appropriate planning and utilization of the LU/LC environment and its resources. So, the area of RS has been a significant tool for understanding the changes that happened in the past and present with the available satellite datasets. In the scientific field of the RS environment, the importance of the LU/LC prediction problem has been analyzed for the time series data by the community of scientists around the world. From the detailed observations, the selection of the satellite data, and the suitable algorithms for extracting the LU/LC classes for different periods is considered the first challenge in performing LU/LC change prediction research. The precise selection of spatial drivers that brings the impact of LU/LC change has been a challenge during the LU/LC prediction analysis. Providing the impact of multi-satellite information of LU/LC features through surface temperature maps by using LU/LC prediction models remains the challenge. The sustainable growth of the LU/LC environment for the time series data requires an accurate classification and prediction map, which was considered the strong motivation for this research. The main objective of our work is to provide the ensemble model of feature extraction-based LU/LC classification and prediction model with less misclassification rate. In our research work, we have proposed the ensemble-based LU/LC prediction model with less misclassification rate.
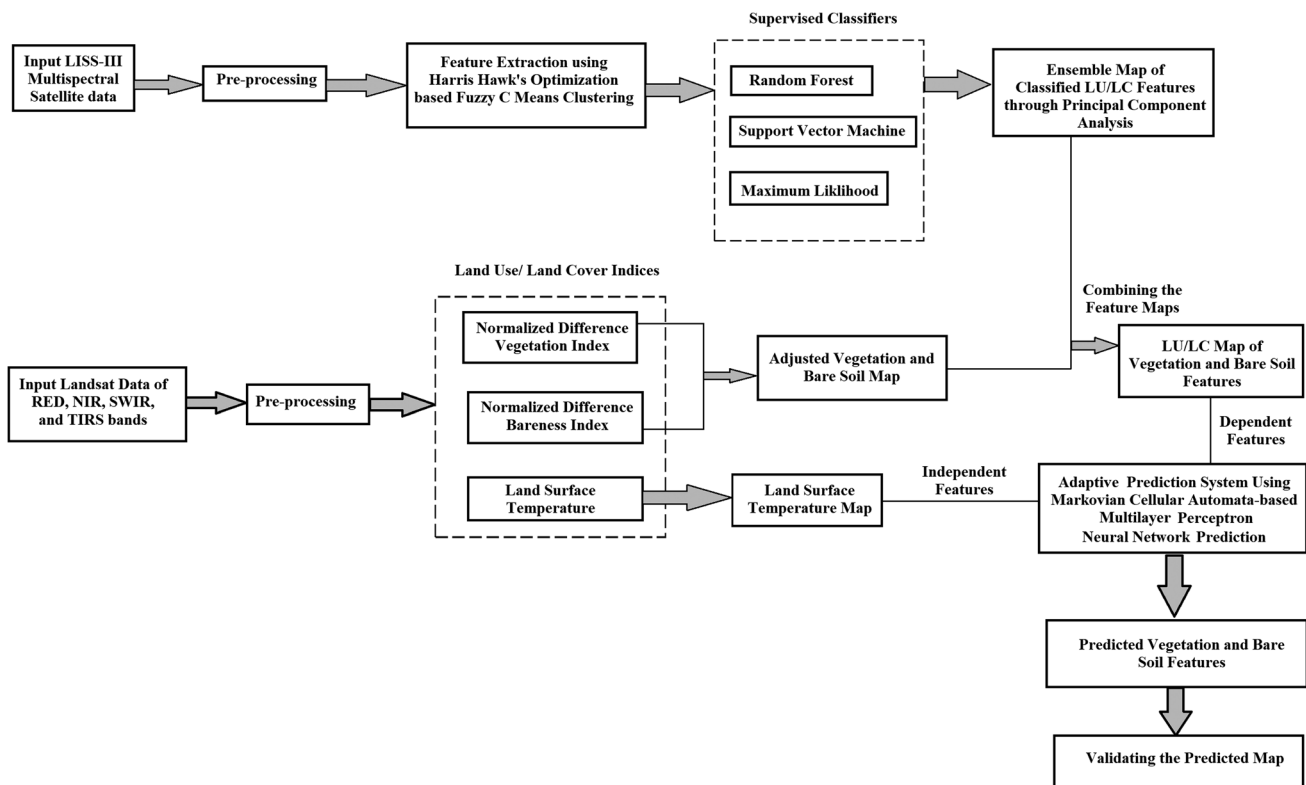

Fig. 1 Proposed ensemble-based LU/LC prediction flow chart

The main contribution of our research provides advanced research on the LU/LC prediction analysis through an ensemble model of unsupervised and supervised algorithms. The results will assist the government officials of land resource management in taking needed actions in protecting the LU/LC environment and its nature. In our research work, we have used the impact of multi-satellite information through the LISS-III and Landsat satellite images for the area of Javadi Hills, India. Our proposed flow of LU/LC prediction analysis is shown in Fig. 1. The main contributions of our work are as follows:

i. An efficient LU/LC classification system has been developed by using the ensemble model of fuzzy HHO and supervised machine learning classifiers through a PCA-based fusion analysis.

ii. The AVBaI has been calculated by using the LU/LC indices of the NDVI and Normalized Difference Bareness Index (NDBaI).

iii. The E-LU/LC vegetation and bare soil map have been obtained by combining the features of the AVBaI with the results of the LU/LC classification system.

iv. The LST map that affects the vegetation and bare soil features of Javadi Hills has been calculated and used as the spatial variable map during the process of LU/LC prediction.

v. With the inputs of the E-LU/LC and LST map of Javadi Hills, the EP-LU/LC map has been obtained by using the MC-MLP neural network model.

The rest of the paper is organized as follows: the "Area of the study and data acquisition" section presents the area of the study. The methods used for performing the LU/LC prediction are described briefly in the "Materials and methods" section.

The experiments and the results of the proposed ensemble models are given in the "Results" section. The detailed discussions of this research work are described in the "Discussion" section. The "Conclusion" section presents the conclusions of this work.

Area of the study and data acquisition

Our research aims to perform the LU/LC prediction for the forest and non-forest-covered region of Javadi Hills, Tamil Nadu, India. The Eastern Ghats extension of Javadi Hills separates the range of the Vellore and Tiruvannamalai districts. It lies between the geographic coordinates of 78.75E12.5N–79.0E12.75N. The GCS-UTM (Geographic Coordinate System-Universal Transverse Mercator) of the 44N projection system was used for extracting the satellite image of Javadi Hills.

The research in Javadi Hills provides information about how much the impact of humans and nature had influenced the LU/LC area for different periods IN the forest and non-forest-covered regions. The thematic view of Javadi Hills was prepared using the shapefiles of India, Tamil Nadu, and Javadi Hills map which has been extracted using DIVA-GIS system software (http://www.diva-gis.org/) and Google Earth Engine platform (https://www.google.com/earth/). The Javadi Hills location map is shown in Fig. 2. Table 1 shows the specification of the geospatial time-series satellite images. The multispectral LISS-III time-series satellite images of Javadi Hills, India, for the years 2012 and 2015 were collected from the Bhuvan RS data of the NRSC, the Indian Geo-platform of ISRO (www.bhuvan.com). The LISS-III time-series satellite images were used for extracting the LU/LC features and the process for generating the LU/LC classification and prediction map of Javadi Hills. The geospatial Landsat (8 and 7) time-series satellite images of Javadi
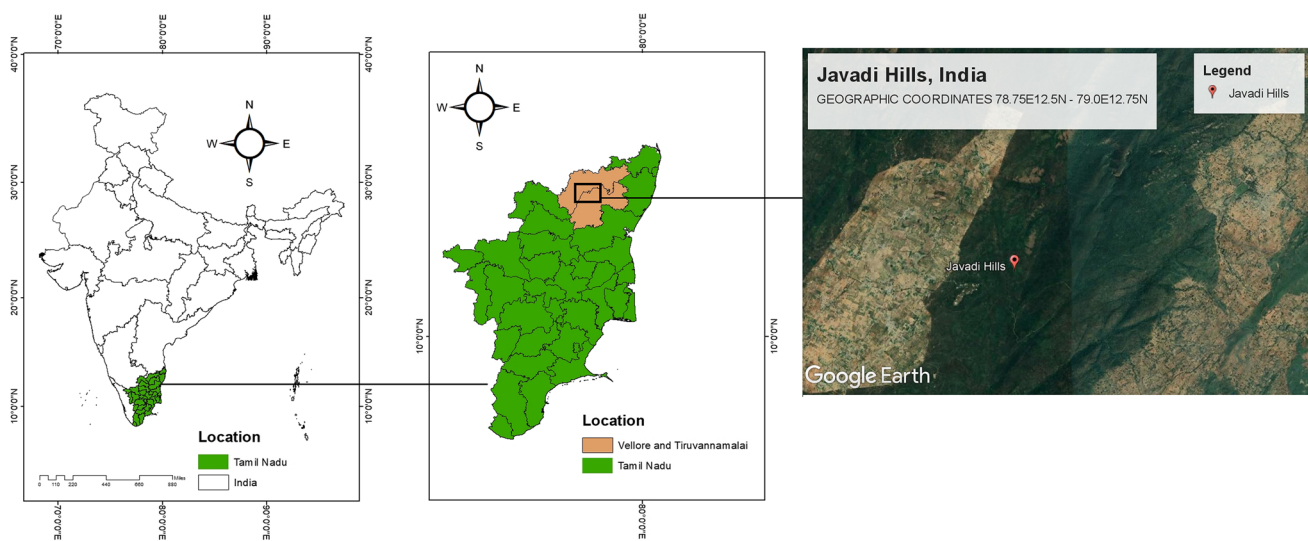

Fig. 2 Javadi Hills — location map

Table 1 Satellite image specification

Satellite/sensor (Tiff format)	Path/row	Spectral and spatial information			Date	Source
		Bands	WL (µm)	RL (m)		
LISS-III—Resourcesat-1/ Resourcesat-2	101/064	B2: Green	0.52–0.59	23.5	18 February 2012, 22 March 2015	Bhuvan RS data of the National Remote Sensing Centre (NRSC), Indian Geo-platform of Indian Space Research Organisation (ISRO) (www.bhuvan.com)
		B3: Red	0.62–0.68	23.5		
		B4: NIR	0.77–0.86	23.5		
		B5. SWIR	1.55–1.75	23.5		
Landsat 8—Operational Land Imager (OLI) and the Thermal Infrared (TI) Sensor	143/51	B1:Coastal/aerosol	0.43–0.45	30	27 March 2015, 19 March 2018, 27 March 2021	United States Geological Survey (USGS), USA (https://earthexplorer.usgs.gov)
		B2: Blue	0.45–0.51	30		
		B3: Green	0.53–0.59	30		
		B4: Red	0.64–0.67	30		
		B5: NIR	0.85–0.88	30		
		B6: SWIR 1	1.57–1.65	30		
		B7: SWIR 2	2.11–2.29	30		
		B8: Panchromatic	0.50–0.68	15		
		B9: Cirrus	1.36–1.38	30		
		B10: TIRS 1	10.6–11.19	100		
		B11: TIRS 2	11.50–12.51	100		
Landsat 7—Enhanced Thematic Mapper Plus (ETM +)	143/51	B1: Blue	0.45–0.52	30	26 March 2012	USGS, USA (https://earthexplorer.usgs.gov)
		B2: Green	0.52–0.60	30		
		B3: Red	0.63–0.69	30		
		B4: NIR	0.77–0.90	30		
		B5: SWIR 1	1.55–1.75	30		
		B6: TIRS	10.40–12.50	60		
		B7: SWIR-2	2.09–2.35	30		
		B8: Panchromatic	0.52–0.90	15		

WL, wavelength; *RL*, resolution; *NIR*, near infrared; *SWIR*, short wave infrared; *TIRS*, thermal infrared sensor

Hills, India, for the years 2012 to 2021 were collected from USGS, USA (https://earthexplorer.usgs.gov). The TIRS Landsat bands of Javadi Hills were used for estimating the LST. The TIRS bands were not present in the LISS-III satellite image. The Landsat NIR, SWIR, TIRS, and RED bands of Javadi Hills were used for estimating the vegetation and bare soil values. The Google Earth Engine data (https://www.google.com/earth/) for the years from 2012 to 2021 were used as the reference data during the process of accuracy assessment.

Materials and methods

In this section, we provided the detailed concepts of geospatial technology followed by pre-processing, feature extraction, supervised classification, fusion, prediction, and evaluation metrics of the proposed model. The geometric, atmospheric, and radiometric corrections were used in our research for correcting the extracted geospatial satellite image. The fuzzy HHO-based unsupervised algorithm was used for extracting the needed features in the pre-processed time-series geospatial satellite image. The supervised machine learning classifiers (SMC) used in our work includes SVM, RF, and ML algorithms. The PCA was used for fusing the results of SMC. The AVBaI has been calculated by using the LU/LC indices of the NDVI and Normalized Difference Bareness Index (NDBaI). The E-LU/LC vegetation and bare soil map have been obtained by combining the features of the AVBaI with the results of the LU/LC classification system. The LST map that affects the vegetation and bare soil features of Javadi Hills has been calculated and used as the spatial variable map during the process of LU/LC prediction. With the inputs of the E-LU/LC and LST map of Javadi Hills, the EP-LU/LC map has been obtained by using the MC-MLP neural network model.

Geospatial technology

In the developing field of geospatial technology, the RS, Geographic Information System (GIS), and Global Positioning System (GPS) play an important part in monitoring, identifying, locating, measuring, and modeling

geospatial data. The initial stage of modeling the geospatial satellite data is to collect information from the satellites. The RS is the science of acquiring the needed information about the surface of the earth through aerial photographs or satellites. Based on different applications in the field of RS satellite, image processing is the latest technology used for collecting high-resolution images. There are different satellites operated around the world by different countries with different spectral, spatial, and temporal resolutions. Some of the satellites are Earth Observing-1 (EO-1), Resourcesat-1, and Resourcesat-2, Landsat series, Sentinel 2 missions, Moderate Resolution Imaging Spectroradiometer (MODIS), Rapid Eye Earth Imaging System (REIS), Quick Bird, and Digital Elevation Model (DEM) (Lechner et al. 2020, Radočaj et al. 2020).

The LU/LC prediction has been found in emerging research for more than 50 decades, and currently, it had been considered an important research problem for scientists around the world. The information on the LU/LC prediction helps to insist the concerned government officials, urban planners, and forest departments take action in protecting the LU/LC environment (MohanRajan et al. 2020). In our research work, we have used the multispectral LISS-III images from Resourcesat-1, and Resourcesat-2, Landsat Images from Landsat 8 (Operational Land Imager (OLI) and the Thermal Infrared (TI) Sensor), and Landsat 7 (Enhanced Thematic Mapper Plus (ETM+)). The geospatial satellite data is extracted in the raster format, where each pixel in the data is associated with an exact geographical location, and the value represents the features of the region. For better understanding, we presented the raster format of LISS-III satellite data of Javadi Hills in Fig. 3. The histograms of our raster data provide information about the pixel value distribution as shown in Fig. 4.

Satellite image pre-processing

Pre-processing is considered to be the necessary progression in the field of RS. The geospatial image restoration or rectification is done for correcting the radiometric, geometric, atmospheric, and topographic distortions in the satellite data. The importance of pre-processing is to provide good visibility to the satellite images to attain better accuracy during the further process of LU/LC prediction (Kumar and Jain 2020, Ma et al. 2020). In our research work, we have used pre-processed methods like layer stacking, band rendering, haze reduction, georeferencing, and line drop-out error correction to provide good visibility to the satellite images.

Layer stacking and band rendering

The process of combining the multiple bands of a geospatial satellite image with the same rows and columns into a single targeted multispectral image is said as layer stacking. The images that needed to be stacked should have the same set of coordinates, whereas the spectral and temporal resolutions should be the same (Dez et al. 2021, Rahman et al. 2020). For better representation of the raster data, band rendering of contrast enhancement has been processed based on the presence of the multiband band, single band, and topographic bands of the satellite data (Gonzalez and Yamamoto 2020, Nádudvari et al. 2020).

$$\left(X_M, Y_M\right) = \frac{\sum_{i=1}^{N}\left(x_i, y_i\right)}{N} \tag{1}$$

Equation (1) represents the calculation of the layer stacking of the images. For each band i, the $1 \leq i \leq N$ represents the bands with the same set of pixel coordinate values x, y. When all the inputs have the same coordinates, the values in the images are added and the average value

Fig. 3 Raster format of the subset LISS-III satellite image of Javadi Hills

LISS − III data − Javadi Hills

Pixel Location (XY)

Rows

Columns
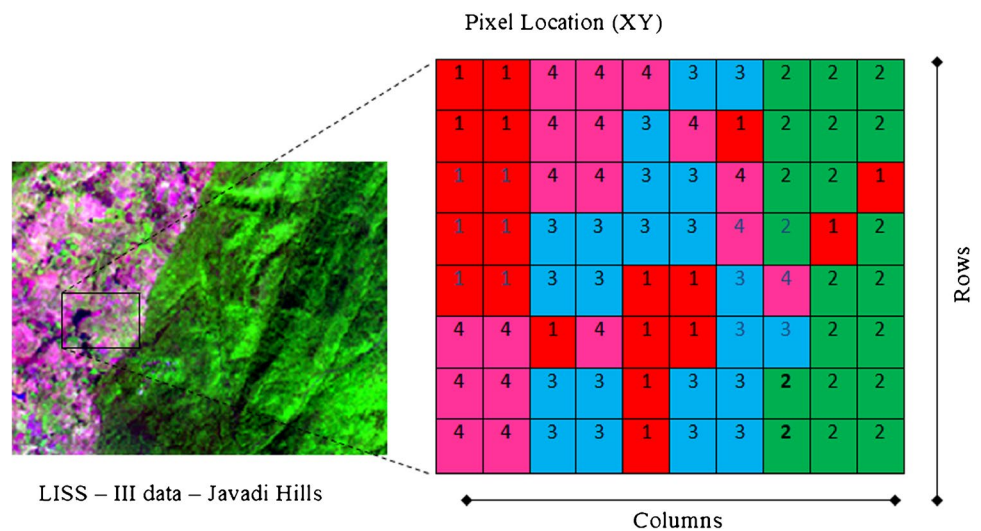
Fig. 4 Histogram representation of LISS-III satellite data of Javadi Hills
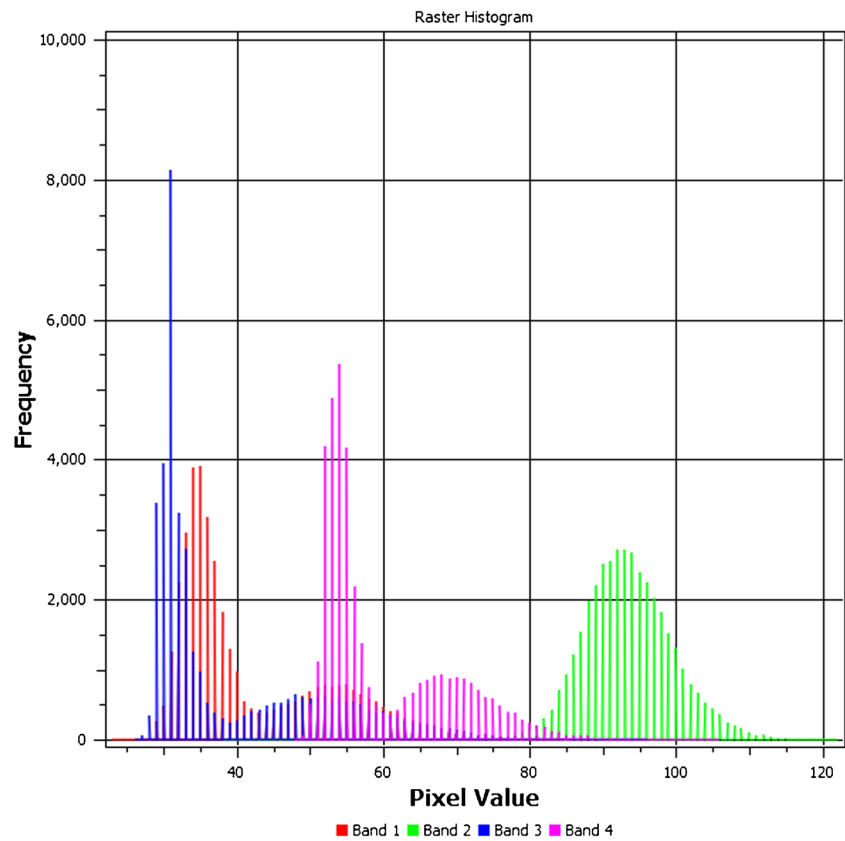
is calculated for representing the multispectral image *M* with the same pixel coordinates *X* and *Y*. All the bands (GREEN, RED, NIR, and SWIR) in the LISS-III sensors were stacked and presented as a multispectral satellite image. For better understanding, we presented the layer-stacked and band-rendered multispectral LISS-III satellite data of Javadi Hills in Fig. 5.

Haze reduction

The presence of atmospheric aerosols like cloud shadows, fog, and other darkened substances in the geospatial satellite data is considered to be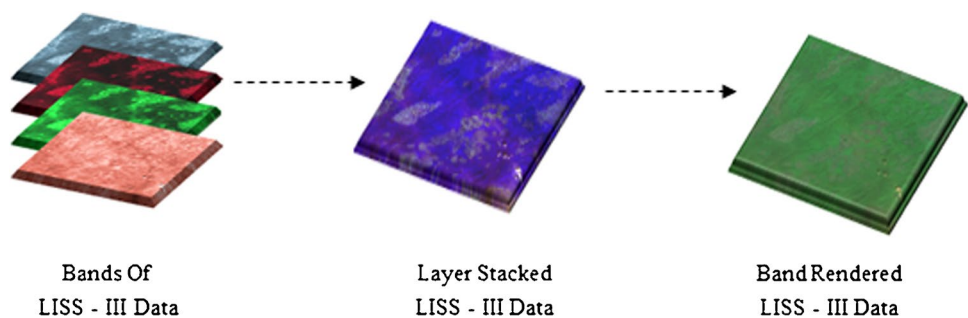 noisy data. The haze reduction process of atmospheric correction increases the accuracy of the geospatial satellite data visually and statistically (Nazeer et al. 2021, Ilori et al. 2019).

Equation (2) represents the calculation of the haze-removing function. For better understanding, we showed the haze-reduced LISS-III satellite data of Javadi Hills in Fig. 6.
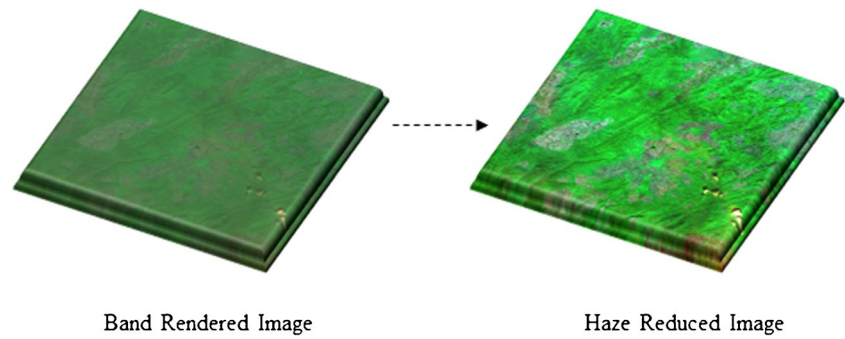
$$I(X) = J(X)t(X) + AL(1 - t(X)) \tag{2}$$

where *I* denotes the input multispectral image, *X* represents the pixel coordinates (*x*, *y*) of the input multispectral image, *J* denotes the scene radiance of the multispectral image, *AL* indicates the atmospheric light, and *t* denotes the multispectral image transparency.

Fig. 5 Layer stacking and band rendering of LISS-III satellite data of Javadi Hills

Bands Of LISS - III Data   Layer Stacked LISS - III Data   Band Rendered LISS - III Data

Fig. 6 Haze reduced LISS-III satellite data of Javadi Hills

Band Rendered Image

Haze Reduced Image

Georeferencing

Georeferencing is the process of correcting the geometrical coordinates (*X*, *Y*) of the geospatial satellite image in the field of RS. The main use of georeferencing is to extract the region of interest (RoI) and to correct the position through each pixel that represents the coordinate or location values in the geospatial satellite data. The shifting, scaling, and rotation transformation function of the first-order derivative helps in resampling the input geospatial data by providing the exact Ground Control Points (GCP) and Digital Number (DN) value of each pixel to the output geospatial satellite data. The Bilinear Interpolation Sampling (BIS) method uses the weighted average of the input DN values to assign the highest weighted value to the output pixel of the geospatial satellite data (Van Ha et al. 2018, Pham et al. 2020).

Equations (3) and (4) show the input and output geometric coordinate system of the geospatial satellite image. Equations (5) and (6) show the equation for the first-order polynomial transformation. Equation (7) shows the equation for the BIS method. For better understanding, we showed the georeferenced RoI LISS-III satellite data of Javadi Hills in Fig. 7. The georeferencing has been performed for extracting the RoI (region of interest) in the geospatial data of Javadi Hills and hence the RoI for our research work falls between 78.80E12.56N and 78.85E12.60N. The size of each pre-processed satellite image of Javadi Hills is 256*200 pixels.
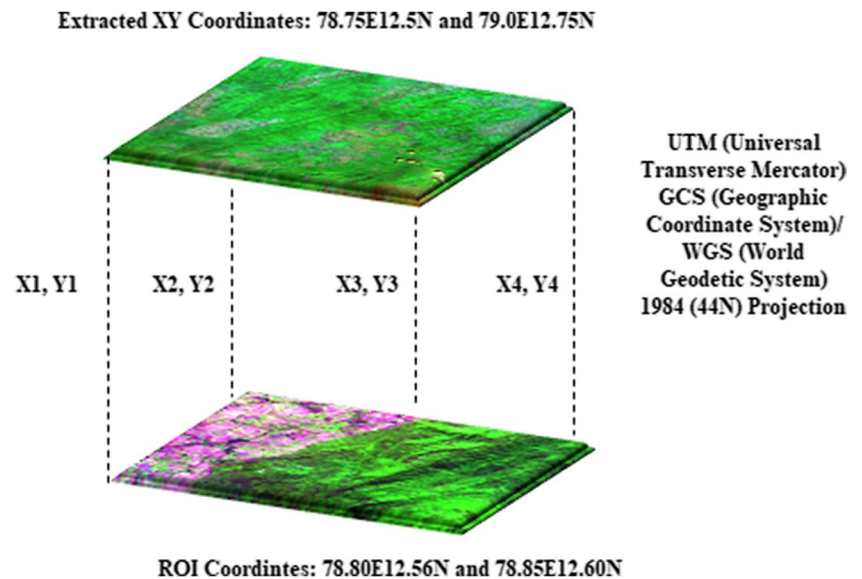
$$x = TF_1(X, Y) \tag{3}$$

$$y = TF_2(X, Y) \tag{4}$$

where *x*, *y* denote the geometric coordinates of the input geospatial satellite image, *X*, *Y* represent the coordinates of the RoI geospatial image, and TF_1, TF_2 denote the transformation functions.

$$x = A_0X + B_0Y + C_0 \tag{5}$$

Fig. 7 Region of interest (RoI) map of Javadi Hills

Extracted XY Coordinates: 78.75E12.5N and 79.0E12.75N

X1, Y1   X2, Y2   X3, Y3   X4, Y4

UTM (Universal Transverse Mercator) GCS (Geographic Coordinate System)/ WGS (World Geodetic System) 1984 (44N) Projection

ROI Coordintes: 78.80E12.56N and 78.85E12.60N

$$y = A_1 X + B_1 Y + C_1 \tag{6}$$

where A_0 denotes the pixel width of the geospatial satellite image, B_0 denotes the rotation term, C_0 denotes center most value x of the upper-left pixel of the geospatial satellite image, A_1 denotes the rotation term, B_1 denotes the negative pixel height of the geospatial satellite image, and C_1 denotes the centermost value y of the upper left pixel.

$$DN_{wt} = \frac{\sum_{n=1}^{4} \frac{R_n}{D_n^2}}{\sum_{n=1}^{4} \frac{1}{D_n^2}} \tag{7}$$

where DN_{wt} denotes the weighted values of DN, R_n denotes the neighboring DN values, and D_n^2 represents the squared distance of the input location in the geospatial satellite image.

Line drop-out error correction

Line drop-out error is the sensor error present in the geospatial satellite image. Radiometric correction is the process used for correcting the line drop-out error of the geospatial satellite image. The zero value information in the bands will lead to the line dropout error. The inverse

distance weighting (IDW) interpolation method was used method for filling the raster region with no data values. The average of the above and below the line of the no-data values was calculated using the IDW method. By using the gap mask data or by using the nearby known weighted location values, the average values were calculated for the missing value locations (Srivastava et al. 2019; Charrua et al. 2021).

The IDW expression for fixing the line drop-out error in the geospatial satellite image is shown in Eq. (8). For better understanding, we have shown the process of line drop-out error correction in Fig. 8 by using the gap-filled image of Landsat 7 data through the process of IDW.

$$Z_j = \frac{\sum_i Z_i / d^n_{ij}}{\sum_i 1 / d^n_{ij}} \tag{8}$$

where Z_i denotes the known data points in the geospatial satellite image, Z_j represents the unknown data points in the geospatial satellite image, d_{ij} mentions the distance measure to the known points, and n denotes the user-weighted parameter. The gaps that cause the scan line error in the RED, NIR, SWIR, and TIRS bands of the Landsat satellite were corrected by using radiometric correction. Figure 9 represents the overall pre-processed multispectral LISS-III Time Series
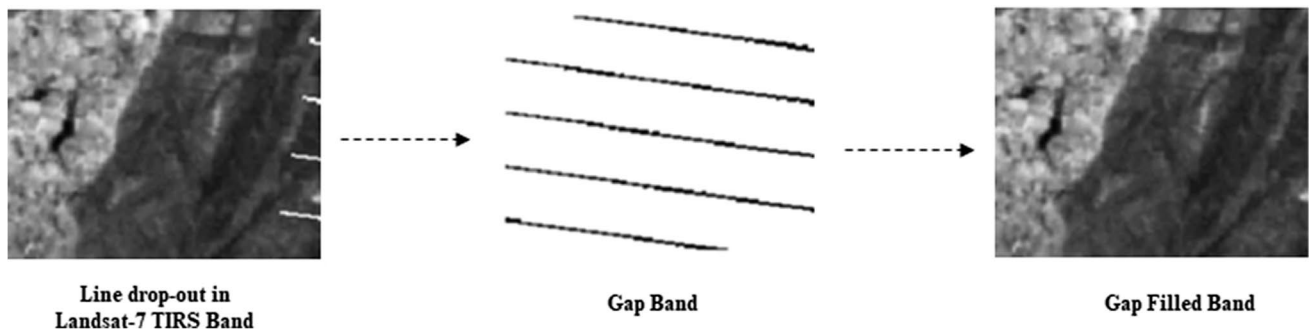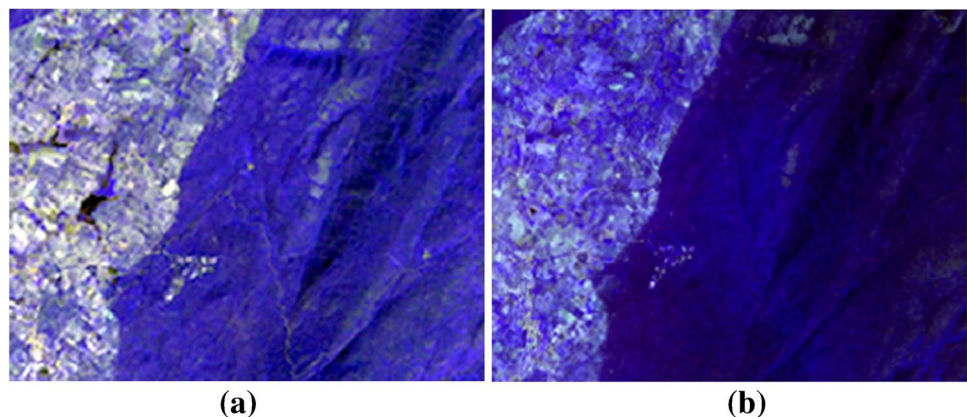

Line drop-out in Landsat-7 TIRS Band　　　　　**Gap Band**　　　　　**Gap Filled Band**

Fig. 8 Line drop-out error corrected Landsat TIRS image of Javadi Hills

Fig. 9 Pre-processed multispectral LISS-III bands of Javadi Hills. **a** 2012. **b** 2015

(a)　　　　　　　　　　　　　　　　　**(b)**

Satellite Images of Javadi Hills for the years 2012 and 2015. The results derived from the pre-processed Landsat bands are shown in the "Extracting the dependent features through the ensemble LU/LC map of Javadi Hills" section.

Fuzzy HHO-based supervised machine learning classifiers for extracting the LU/LC features

The unsupervised fuzzy HHO algorithm was used for finding the unknown features from the pre-processed satellite image. The feature-extracted map was used as an input for finding the LU/LC classes by using the supervised machine learning classifiers (SVM, RFC, and MLC). The PCA was used for fusing the results of the supervised machine learning classifiers.

Fuzzy Harris hawks optimization algorithm

Feature extraction is an important process that helps in increasing the accuracy of the classification models by extracting the needed features from the geospatial satellite data. The features are part of the information from the content of the geospatial satellite data. The new set of extracted feature values will be different from the original values. In the field of RS, object-based clustering and optimization methods were used consistently for extracting the features from the geospatial satellite data. The features are unknown and the process of an unsupervised-based optimization algorithm helps in extracting the unknown features in an image into clusters or groups without any training process (Matci and Avdan 2020; Al-Qaisi et al. 2021; Sihwail et al. 2020). In our research work, we have proposed the unsupervised fuzzy HHO technique for the optimal selection of cluster centers for extracting the needed features in the pre-processed LISS-III satellite images.

The swarm intelligence-based optimization algorithm helps the population members by the group information and through each other's position. Harris hawks are one of the smartest birds in the world. The uniformity and cooperation among these birds were recognized as the best solution to targeting the rabbits. The different views on the attacking and escaping behavior of Harris hawks were observed and mathematically modeled for handling optimization problems. In HHO, we have a group of Hawks as the candidate solutions, and the best-obtained solution is called the rabbit. The exploration and exploitation phase has been employed for performing the HHO (Heidari et al. 2019; Abualigah et al. 2021; Iswisi et al. 2021). In the exploration phase, Harris hawks search for the rabbit (Prey). The gripping eyes of the hawks help them to spot and track the rabbit. In a few cases, it finds difficult to see the rabbit. In this case, the hawks will

wait and monitor the place more visibly for tracking the rabbit. For each iteration, all the hawks are considered as the candidate solutions, and the fitness value has been calculated for each hawk based on the planned rabbit. Later, the hawks may wait in some positions for detecting the rabbit through Eq. (9). HHO uses a direct method for calculating the average position of the hawks using the following Eq. (10).

$$X(t+1) = \begin{cases} X_{\text{rand}}(t) - r_1 |X_{\text{rand}}(t) - 2r_2.X(t)| & rand \geq 0.5 \\ \left(X_{\text{rabbit}}(t) - X_M(t)\right) - r_3\left(LB + r_4(UB - LB)\right) & rand < 0.5 \end{cases}$$

(9)

$$X_M(t) = \frac{1}{N}\sum_{i=1}^{N} X_i(t)$$

(10)

$$E = 2E_0(1 - \frac{t}{T})$$

(11)

In the HHO, the energy coefficient changes the behavior of the hawks from searching to hunting. HHO assumes that the rabbit's energy is reducing while escaping from the hawks. This assumption has been calculated using Eq. (11). When the escaping energy of the rabbit is more than or equal to 1, the HHO sends the hawks to travel to different locations in search of the rabbit. When the rabbit's energy got reduced, then the hawks will search the neighborhood for the solution during the exploitation phase. The exploitation stage focuses on the vicinity of the acquired solutions through the hard and soft besieges. In this phase of the exploitation stage, the hawks target the rabbit through its previous position. The hawks monitor and follow the chasing strategy as soon as the rabbit challenges to escape. The two variables, r and $|E|$, have been used for the indication of events that occurred during the hawk's optimization. The variable $|E|$ denotes the escaping energy of the rabbit, and r denotes the probability of escaping. When $r < 0.5$, the rabbit has a higher chance to escape from the hawk, and $r \geq 0.5$, there is less chance for the rabbit to escape from the hawk.

$$X(t+1) = \Delta X(t) - E|J.X_{\text{rabbit}}(t) - X(t)|$$

(12)

$$\Delta X(t) = X_{\text{rabbit}}(t) - X(t)$$

(13)

$$J = 2(1 - r_5)$$

(14)

The HHO has been performed based on the four possible tactics, and they are soft besieges, hard besieges, soft besieges with progressive rapid dives, and hard besieges with progressive rapid dives. In soft besieges, the rabbit has some energy to escape from the hawk where $r \geq 0.5$ and $|E| \geq 0.5$. Hence, the hawks make the rabbit lose energy by surrounding it before its plans for the surprise jump. The

soft besieges are mathematically formulated in Eqs. (12), (13), and (14). In hard besieges, the rabbit is tired with a low chance of escaping from the hawk where $r \geq 0.5$ and $|E| < 0.5$. Hence, the hawks hardly surround the rabbit before the surprise jump. The hard besieges update the next location of the hawk through Eq. (15).

$$X(t+1) = X_{\text{rabbit}}(t) - E|\Delta X(t)| \tag{15}$$

In soft besiege with progressive rapid dives, the rabbit still has the energy to escape from hawk attack where $r < 0.5$ but $|E| \geq 0.5$. Hence, the hawk makes a clever zigzag movement with several dives around the rabbit before the surprise jump. In this stage of soft besiege, the hawks will update the position in two steps. In the first attempt, the hawks will target the rabbit by calculating the next move. The first step is shown in Eq. (16). Through the comparison of the last dive and the possible results, the decision has been made to either dive or not. If it is not, the hawks will start performing irregular dives while attacking the rabbit, based on the *LF* concept. Hence, the quick dive mechanism is formulated and shown in Eq. (17). The levy flight function is formulated in Eq. (18). The position of the Harris hawks in the soft besiege phase is updated and the mathematical expression is shown in Eq. (20).

$$Y = X_{\text{rabbit}}(t) - E|JX_{\text{rabbit}}(t) - X(t)| \tag{16}$$

$$Z = Y + S \times LF(\text{D}) \tag{17}$$

$$LF(x) = 0.01 \times \frac{\mu \times \sigma}{|v|^{\frac{1}{\beta}}} \tag{18}$$

$$\sigma = \left(\frac{\Gamma(1+\beta) \times \sin(\frac{\pi\beta}{2})}{\Gamma\left(1+\frac{\beta}{2}\right) \times \beta \times 2^{\frac{\beta-1}{2}}} \right) \tag{19}$$

$$X(t+1) = \begin{cases} Y \text{ if } F(Y) < F(X(t)) \\ Z \text{ if } F(Z) < F(X(t)) \end{cases} \tag{20}$$

In hard besiege with progressive rapid dives, the rabbit does not have sufficient energy to escape where $r < 0.5$ but $|E| < 0.5$. The Harris hawks perform some rapid dives to target and kill the rabbit. The movement of the Harris hawks in the phase of hard besiege is formulated through some levy flight functions if they are failed. The position of the hawks in iteration $t + 1$ has been controlled by Eq. (20). Hence, the position of X and Y is updated and shown in Eqs. (21) and (22). The symbolic representation of HHO is shown in Table 2.

$$Y = X_{\text{rabbit}}(t) - E|JX_{\text{rabbit}}(t) - X_M(t)| \tag{21}$$

Table 2 Meaning of Symbols in HHO

Symbol	Description
$X(t)$	Current position of a hawk
$X(t+1)$	Position of a hawk in the new iteration
$X_{\text{rabbit}}(t)$	Position of the most optimal solution
$X_{\text{rand}}(t)$	Random position in the problem space
$X_M(t)$	Center of gravity of the population of hawks
r_1, r_2, r_3, r_4, r_5	Random numbers ranging from 0 to 1
UB, LB	Lower and upper limits of the solutions
N	Swarm size
$X_i(t)$	Position of the hawks in iteration t
E	Escaping energy of the rabbit
E_0	The initial state of energy that randomly changes over the interval $(-1, 1)$
T	Maximum number of iterations
t	Current iteration
J	Random jump strength of the rabbit
$\Delta X(t)$	Difference between the Harris hawk and rabbit location in iteration t
LF	Levy flight function
D	Dimension of the optimization problem
Y	Position based on the soft siege
Z	Position change based on the quick dive
S	Randomly generated vector of size $1 * D$
μ and v	Random values between the intervals 0 and 1
β	Default fixed constant value of 1.5
F	Calculated fitness function for Y and Z

$$Z = Y + S \times LF(\text{D}) \tag{22}$$

The unsupervised FCM or fuzzy membership clustering algorithm is widely used in the field of digital image processing, for extracting information from larger image datasets. Fuzzy membership clustering is considered as a grouping of the data in different clusters where the location of the data in each cluster depends on the degree of fuzzy membership of the data. The membership degree value is in the range of 0 and 1. The similarity between the data in the clusters is identified through the higher or lower membership value. Fuzzy membership clustering is considered an iterative optimization method that minimizes the cost function to reduce the dissimilarities between the clusters (Loganathan and Kaliyaperumal 2016, Valdez et al. 2021, Arora and Meena 2020, Li and Endo 2020). Equation (23) represents the minimization of the cost function in fuzzy membership clustering.

$$J_{\text{FCM}} = \sum_{j=1}^{N} \sum_{i=1}^{c} (u_{ij})^m \|x_j - v_i\|^2 \tag{23}$$

where u_{ij} represents the fuzzy membership of pixel x_j in the ith cluster, v_i represents the ith cluster center, $\|.\|$ represents the inner product normalization metrics, and m represents the constant value. When the pixels are found to be nearer to the centroid of their clusters, the cost function is reduced and the clusters are allocated with a higher membership value. The lower membership values are allocated to the pixels that are far from the centroid of their clusters. The membership function determines the probability of the pixels fitting into the specific cluster. In this clustering method, the probability merely relies on the distance between the pixel and all the cluster centers in the feature domain. In the theory of fuzzy membership clustering, each cluster center is initialized with random values and then converges to the ideal solution for the v_i that represents the local minimum point of the cost function. The termination of the fuzzy membership clustering will be processed when there is no change in the cluster center during the iterative process. By using the outputs of the preprocessed image of Javadi Hills, the features in the LISS-III satellite images were extracted by using the proposed fuzzy HHO algorithm. During feature extraction, the unsupervised metaheuristic HHO algorithm was executed for selecting the optimum values from the membership matrix. The main aim of the feature extraction technique using fuzzy HHO is to use the optimum membership matrix for finding the cluster centers and to detect the needed features from the pre-processed satellite image. Each iteration has a set of membership matrices where each of them is considered a member of the HHO population. Each membership matrix represents the clusters evaluated by using the cost function of the fuzzy membership algorithm for determining the cluster centers. From the matrices, the optimal clusters were selected as the optimum membership matrix. The feature extraction-based fuzzy HHO algorithm consists of the following steps:

i. Cluster centers were randomly created from the pre-processed LISS-III image, and the membership matrix of each cluster is generated.

ii. A membership matrix is considered for each hawk in the HHO population.

iii. The Harris hawk and its movement functions are used to update the membership matrix, and they are employed for creating the new cluster centers in the following iterations.

iv. The new cluster centers are calculated by using the fuzzy evaluation function. The matrix with the minimum possible error is observed to be the optimal membership matrix.

v. The optimal matrix is the position of the rabbit in the HHO algorithm and it is more responsible for pointing the other membership matrices. The optimal matrix is used in the last iteration for finding the optimal cluster centers.

vi. The optimal cluster centers can be used for clustering the needed features from the pre-processed LISS-III satellite image.

While executing the fuzzy HHO algorithm, each pre-processed LISS-III image is used separately as an input. The parameters like population size, number of clusters, and the total number of iterations are initialized. A membership matrix has been used as the member of the Harris hawk population, and it is shown in Eq. (24).

$$\text{HHO}_i = \begin{bmatrix} u_{1,1}^i & \cdots & u_{1,c}^i \\ \cdots & \cdots & \cdots \\ u_{n,1}^i & \cdots & u_{n,c}^i \end{bmatrix}_{n \times c} \tag{24}$$

Every membership matrix represented in the above equation has n rows which indicate the total number of satellite image pixels, and c columns which indicate the fuzzy degree of an image pixel to cluster centers. The HHO_i represents the member of the HHO algorithm and finds the optimum solution through the membership matrix. According to Eq. (25), every hawk represents a membership matrix of fuzzy clustering for feature extraction. Each component of a fuzzy membership matrix is between 0 and 1, and it is represented in Eq. (26).

$$\text{HHO}_i = \left[U_{n.c} \right]^i \tag{25}$$

$$0 \leq u_{ij} \leq 1 \tag{26}$$

where i and j represent the pixel number in a pre-processed LISS-III image and a cluster center in the pre-processed LISS-III image, respectively. Every solution found by a hawk is a membership matrix. According to Eq. (27), the sum of the rows is 1, which signifies the sum of the fuzzy degrees of an image pixel to all the cluster centers is equal to 1.

$$\sum_{j=1}^{c} u_{ij} = 1 \tag{27}$$

where i is the number of pixels in a pre-processed LISS-III image, and j represents the number of cluster centers, which is denoted as c here. The fuzzy HHO algorithm generates an initial population of cluster centers where the membership matrices were defined to satisfy the conditions of Eq. (28).

$$\text{Pop} = \left\{ \left[U_{n.c} \right]^1, \left[U_{n.c} \right]^2, \ldots \ldots, \left[U_{n.c} \right]^n \right\} \tag{28}$$

where Pop represents the initial population of the membership matrices, which determines the cluster centers and n represents the number of members in the initialized population. The membership matrices are used for updating the cluster centers in the fuzzy membership clustering technique according to Eq. (29):

$$c_i = \frac{\sum_{k=1}^{n} u_{ik}^m I_k}{\sum_{k=1}^{n} u_{ik}^m} \tag{29}$$

In the membership matrix, each value of the satellite image pixel at any coordinates is shown by I_k, and the membership matrix or the members of the HHO algorithm has been updated through cluster centers and the position of the membership matrixes by using Eq. (30).

$$S(i) = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik})^m \text{sim}(I_k, c_i) \tag{30}$$

where $\text{sim}(I_k, c_i)$ represents the similarity of the satellite image pixel I_k to a cluster center c_i. The fuzzy HHO algorithm minimizes the target function by selecting the membership matrix optimally, where the cluster centers are optimized. In the HHO algorithm, the relevant equations are used in updating the membership matrix. For instance, Eqs. (31) and (32) show the behavior of the soft besieges and hard besieges implemented on the membership matrices to update them respectively.

$$\text{HHO}(t + 1) = \text{HHO}^* - \text{HHO}(t) - E|J.\text{HHO}^* - \text{HHO}(t)| \tag{31}$$

$$\text{HHO}(t + 1) = \text{HHO}^* - E|\text{HHO}^* - \text{HHO}(t)| \tag{32}$$

where $\text{HHO}(t)$ represents the membership matrix at each iteration t, and $\text{HHO}(t + 1)$ denotes the updated membership matrix. The HHO^* represents the optimal membership matrix, where the cluster centers are created for feature extraction. The feature extracted result of the unsupervised fuzzy HHO technique for the years 2012 and 2015 is shown in Fig. 10. The output obtained from the fuzzy HHO technique was used as input for finding the LU/LC classes.

Supervised machine learning classifiers

Supervised classification is the most often used technique in the field of digital image processing. Supervised learning algorithms are widely used for the classification of satellite data in the area of remote sensing. The supervised classification works on the known knowledge about the image where the user can select the sampled pixels from an image and train each of the pixels with the reference pixels of the training image to attain good classification accuracy (Manoharan et al. 2019; da Silva et al. 2020; Wang et al. 2020). In our research work, we have used RFC, SVM, and MLC-supervised classification algorithms for classifying the optimized geospatial satellite data. The RFC algorithm has been used for both regression and classification problems in the field of digital image processing. In the field of remote sensing, the RF classification method was mostly used in classifying the LU/LC classes from the geospatial satellite image. The RF builds multiple decision trees on the subset of the training data and then obtains the classified results from each decision tree and finally chooses the best solutions using averaging or voting the results. The disadvantage of overfitting is reduced by averaging the results of the decision trees. The selections of features in the training datasets were processed by calculating the Gini index, and this metric has been widely used during the decision tree classification process (Piao et al. 2021; Amoakoh et al. 2021). The Gini index calculates the sum of the square of the probability of each class and the expression is shown in Eq. (33).

$$\text{Gini index} = 1 - \sum_{i=1}^{C} (P_i)^2 \tag{33}$$

where P_i represents the probability of the classified features in the data, and C represents the total number of classes to be classified.

The SVM-supervised classification has been used for LU/LC classification problems in the field of remote sensing. The ideology of the SVM classification algorithm is to search the hyperplane and find it in n-dimensional space for classifying the dissimilar data points. The hyperplane helps in the extrication of the decision boundary between the data points. The features in the n-dimensional space

Fig. 10 Feature extracted map by using the fuzzy HHO technique. **a** 2012. **b** 2015

(a)　　　　　　　　　　　　　　　　**(b)**
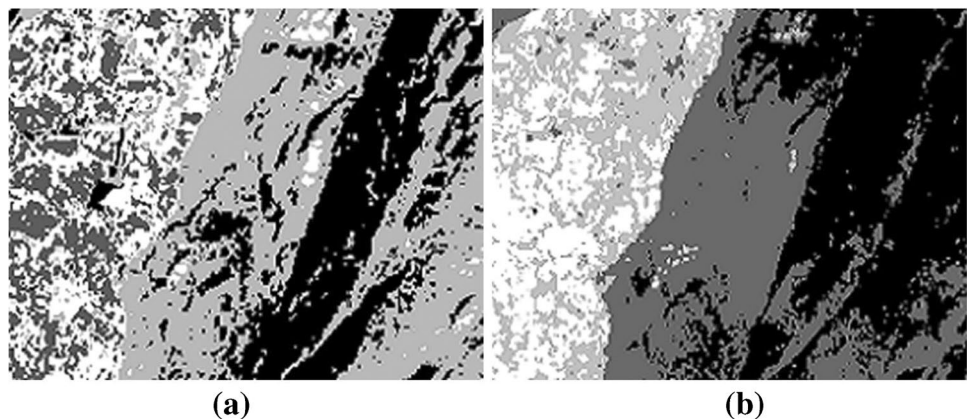
depend on the dimensions of the hyperplane. The data points that are nearer to the hyperplane were considered as the support vectors and hence help in maximizing the classifier's margin (Karimi et al. 2020; Hawash et al. 2021). Equations (34) to (41) show the expressions for describing the process of the SVMclassifier.

$$H : w^t x + b = 0 \tag{34}$$

$$H_1 : w^t x + b = -1 \tag{35}$$

$$H_2 : w^t x + b = 1 \tag{36}$$

where x represents the input column vector, w represents the weight of the column vector, b refers to the bias function by which the decision boundary of the hyperplane that separates the data points was represented as H, and the hyperplane H_1 and H_2 moves along the support vectors were represented as -1 and $+1$ respectively.

$$D_1 = \frac{(-1-b)}{|w|} \tag{37}$$

$$D_2 = \frac{(1-b)}{|w|} \tag{38}$$

$$M = \frac{D_1}{D_2} \tag{39}$$

$$M = \left(\frac{\frac{(-1-b)}{|w|}}{\frac{(1-b)}{|w|}} \right) \tag{40}$$

$$M = \frac{2}{|w|} \tag{41}$$

where the distance between the hyperplane H_1 and H_2 to its origin was represented as D_1 and D_2, and the formulation of M shows twice the margin. The maximization of the margin ($\max \frac{1}{|w|}$) through the adjusting parameters (w and b) was considered the key objective of the SVM classifier.

The MLC algorithm has been considered the well-known LU/LC classification system in the area of remote sensing. The ML classification system assigns the unknown pixels to the known classes in each band using the training data. The statistics of ML in classifying the LU/LC classes of each band in geospatial data are normally distributed by calculating the probability of LU/LC classes belonging to each pixel (Kumar et al. 2021, Bosquilia and Muller-Karger 2021). The class conditional density calculated during LU/LC classification is shown in Eq. (42). The standard Gaussian distribution function has

been used in this research work, where ω represents the set of parameters from each class j, x represents the image data of n bands (feature vector), and σ and μ represent the standard deviation and mean of each class j respectively.

$$p\left(x|\omega_j\right) = \frac{1}{\sigma_j \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_j}{\sigma_j}\right)^2} \tag{42}$$

Fusion using principal component analysis

Image fusion in the field of satellite image processing combines all spectral and spatial information of multiple satellite images of the same source with similar dimensions into a single image that preferably contains all the needed features from each of the original satellite images. PCA is an efficient way of fusing geospatial satellite images. The PCA method transforms the variables from correlated to uncorrelated for calculating the necessary weights to fuse the multiple images (Manoharan et al. 2018, Javan et al. 2021, Kaur et al. 2021, Batur and Derya 2018). The steps involved in fusing the image using PCA are as follows:

i.   The multiple images to fuse into a single image are stored in a matrix system by transforming each image into column vectors.
ii. Images that are stored in the column matrix are subtracted by their mean and the covariance matrix was calculated.
iii. The Eigenvectors and Eigenvalues of the covariance matrix were calculated and normalized in descending order.
iv. The Eigenvector that corresponds to the highest Eigenvalues has been used for calculating the weighted value for fusing the multiple images according to Eq. (43).
v.   The PCA-based fused image will be calculated through the summation of the scaled matrices according to Eq. (44).

$$Pc_1 = \frac{V(1)}{\sum V}, Pc_2 = \frac{V(2)}{\sum V}, Pc_n = \frac{V(n)}{\sum V} \tag{43}$$

$$Img\ Fuse = Pc_1 I_1 + Pc_2 I_2 + \cdots + Pc_n I_n \tag{44}$$

where I represents the input images to be fused, and $V(n)$ represents the Eigenvectors that correspond to the principal component Pc_n. The principal component analysis has been used for fusing the results of SMC. The SMC used in our work includes SVM, RF, and ML algorithms. The LU/LC classes (high vegetation, scarce vegetation, and bare soil) of Javadi Hills for the years 2012 and 2015 were obtained by

using the ensemble model of the fuzzy HHO technique with SMC through PCA.

The fused result for the years 2012 and 2015 is shown in Fig. 11. The algorithm of our ensemble model of fuzzy HHO-based SMC through PCA for extracting the LU/LC classes is presented in Algorithm 1. The overall flow chart of the ensemble model of fuzzy HHO-based SMC through PCA for extracting the LU/LC classes is shown in Fig. 12. We have used the front end of MATLAB 2021a for executing the fuzzy HHO-based feature extraction technique and Python software for executing the ensemble model of SMC algorithms.

Adjusted vegetation and bareness LU/LC indices

In the field of RS, the vegetation features of the geospatial data are measured using the vegetation Index, and the bare soil features are measured using the bare soil index. The NDVI has been used for extracting the vegetation features through NIR and RED bands of satellite data. The NIR band reflects the LU/LC features of vegetation strongly and the RED band absorbs the LU/LC features of vegetation for the particular area. The Normalized Difference Bareness Index (NDBaI) has been used for extracting the features of bare soil through the Short Wave Infrared (SWIR) and TIRS bands of satellite data (Useya et al. 2019; Mzid et al. 2021). Equation (45) shows the expression of NDBaI for Landsat 8 data, and Eq. (46) shows the expression of NDBaI for Landsat 7 data. Equation (47) shows the expression of the NDVI.

$$NDBaI_{Landsat\ 8} = \frac{SWIR - TIRS}{SWIR + TIRS} \tag{45}$$

$$NDBaI_{Landsat\ 7} = (RED + SWIR - NIR) \tag{46}$$

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)} \tag{47}$$

where SWIR, RED, NIR, and TIRS represent the reflectance values of the SWIR, RED, NIR, and SWIR bands of the Landsat satellite images. By using the outputs of pre-processed Landsat bands, the LU/LC indices have been calculated. The NDVI and NDBaI show the features of LU/LC vegetation and bare soil. In our research work, we have combined the NDVI and NDBaI indices for providing the effective features of LU/LC vegetation and a bare soil map. The equation used for extracting the features of vegetation and bare soil map for the bands of Landsat 8 and 7 are given in Eqs. 48 and 49.
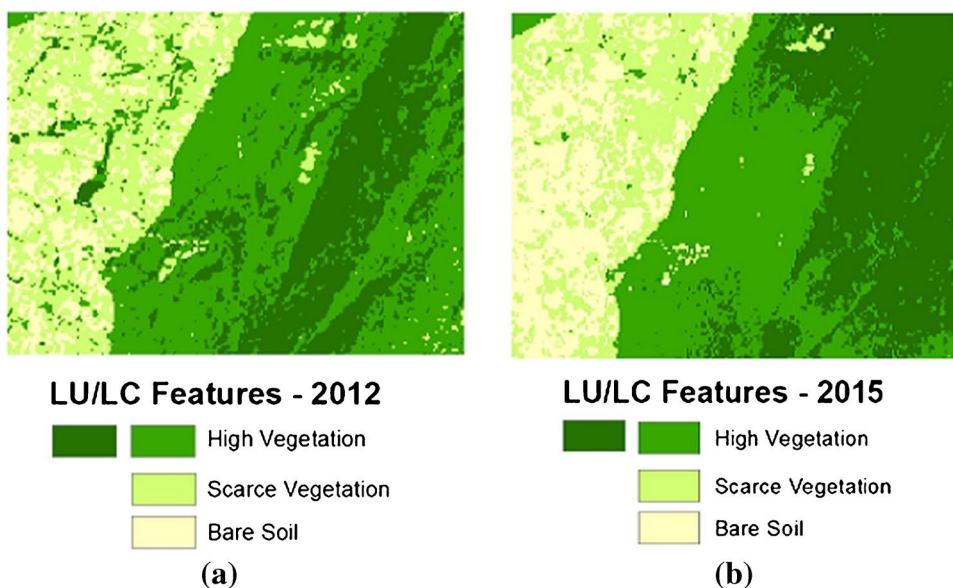
$$AVBaI_{Landsat\ 8} = (NDVI + NDBaL_{Landsat\ 8}) \tag{48}$$

$$AVBaI_{Landsat\ 7} = (NDVI + NDBaL_{Landsat\ 7}) \tag{49}$$

where AVBaI represents the Vegetation and Bareness Index, NDVI represents the Normalized Difference Vegetation Index, and NDBaI represents the Normalized Difference Bareness Index.

We have used the map algebra module of the python package in the back-end tool of ArcGIS geospatial software for evaluating the LU/LC indices of Javadi Hills. The AVBaI map for the years 2012 and 2015 is shown in Fig. 13. The combined map provides the clarity of vegetation and bare soil regions of satellite data for the area of Javadi Hills. The Vegetation and Bareness Index for the years 2018 and 2021

Fig. 11 LU/LC features of ensemble model of fuzzy HHO-based with SMC through PCA. **a** 2012. **b** 2015

LU/LC Features - 2012

High Vegetation
Scarce Vegetation
Bare Soil

(a)

LU/LC Features - 2015

High Vegetation
Scarce Vegetation
Bare Soil

(b)

Fig. 12 Flow chart of the ensemble model of fuzzy HHO-based SMC through PCA
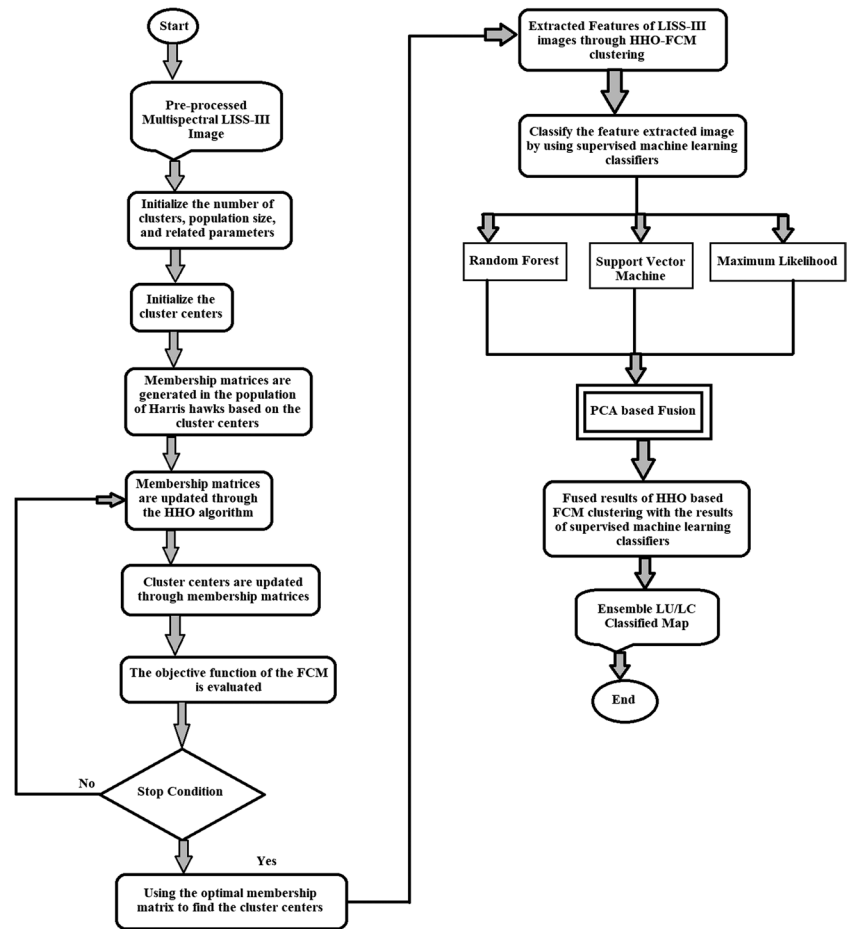
is also calculated and shown in Fig. 14. The AVBaI map for the years 2018 and 2021 has been reclassified using the Python package through ArcGIS software. The reclassified reference map for 2018 and 2021 is shown in Fig. 21. In our proposed research work, we have used the outputs of the AVBaI map of Javadi Hills for the years 2018 and 2021 as the reference map to validate the predicted LU/LC map of 2018 and 2021.

Extracting the dependent features through the ensemble LU/LC map of Javadi Hills

The dependent features of the E-LU/LC map define the changes that happened between the LU/LC features of Javadi Hills for the years 2012 and 2015. In our research work, we have analyzed the vegetation and bare soil features of Javadi Hills by using an efficient ensemble model. The dependent E-LU/LC map has been extracted by combining the outputs of the PCA-based fused LU/LC map (fuzzy HHO-based SMC) with the output of the AVBaI maps for the years 2012 and 2015. The ensemble features produce the combination of the pixel values with the associated attribute tables. Each coordinate of the ensemble LU/LC raster map contains the

dependent vegetation and bare soil feature values. We have shown the E-LU/LC map in Fig. 15. The E-LU/LC-dependent map for the years 2012 and 2015 has been used as an input for the prediction model for analyzing the future LU/LC changes in the area of Javadi Hills. We have used the front end of python and the back end of ArcGIS geospatial software for combining the ensemble features of LU/LC vegetation and bare soil map.

Extracting the independent LST features of Javadi Hills

In the field of RS, the skin temperature of the geospatial data was measured using the LST. The hot and cold on the surface of the earth were measured using the LST through the reflection of the radiant energy within the earth's surface. The TIRS data captured from the satellite were used in measuring the mixture of the hot and cold temperatures of the earth (Mukherjee and Deepika 2020, Sekertekin and Bonafoni 2020). In our research work, the LST has been used as one of the spatial data for analyzing the LU/LC prediction map. The TIRS bands of Landsat 7 and 8 satellite data have been used in our

Algorithm 1 Algorithm to construct the ensemble model of fuzzy Harris hawks optimization-based supervised machine learning classification algorithms through principal component analysis

Inputs (I_P): Multispectral LISS III Image

Output (O_P): Classified LU/LC Image

Begin

1   Initializing the Input Satellite Images

2   Pre-processing of the Input Images

3   Perform the Unsupervised Metaheuristic Optimization-based Fuzzy Membership Method for Extracting the Features from the Pre-processed Satellite Images

4   Implement the Fuzzy HHO-based Feature Extraction Algorithm

5   Initialize the number of clusters, population size, and related parameters

6   Initialize the cluster center

7   Membership matrices are generated in the population of Harris hawks based on cluster centers

8   Membership matrices are updated through the HHO algorithm

9   Cluster centers are updated through membership matrices or Harris hawk

10   Clustering or the objective function of the Fuzzy is evaluated

11   Using the optimal membership matrix to find the cluster centers

12   Stop when the conditions are satisfied and return the best solution

13   Extracted the needed features from the pre-processed Image

14   Return result of Fuzzy HHO-based Feature Extraction Technique

15   Execute the SMC for the output of the Fuzzy HHO-based Feature Extraction Technique

16   Train the SMC (SVM, MLC, and RFC) separately by using the training data

17   Return the results of each SMC

18   Implement the PCA-based Fusion Technique to ensemble the outputs of the SMC

19   Return the results of the ensemble model of Fuzzy HHO based SMC through PCA

End

Fig. 13 Adjusted Vegetation and Bareness Index map. **a** 2012. **b** 2015
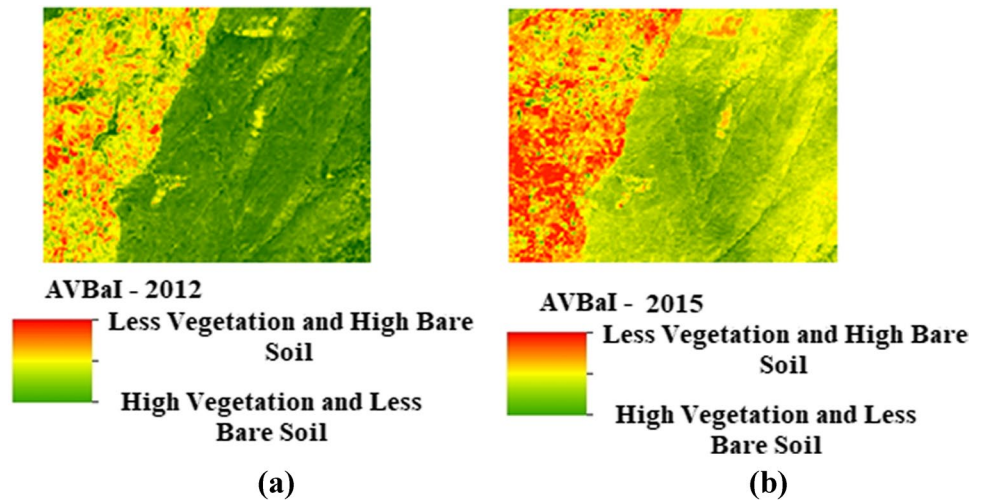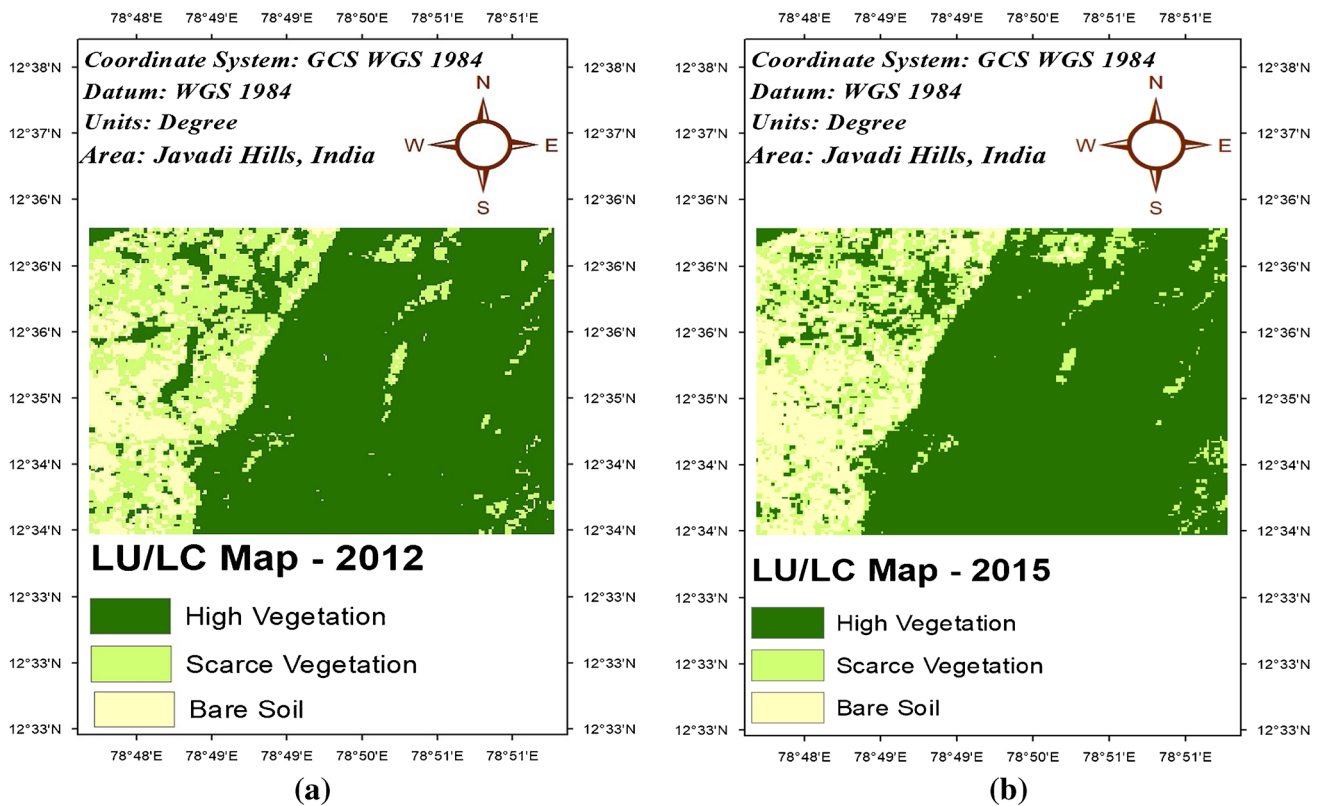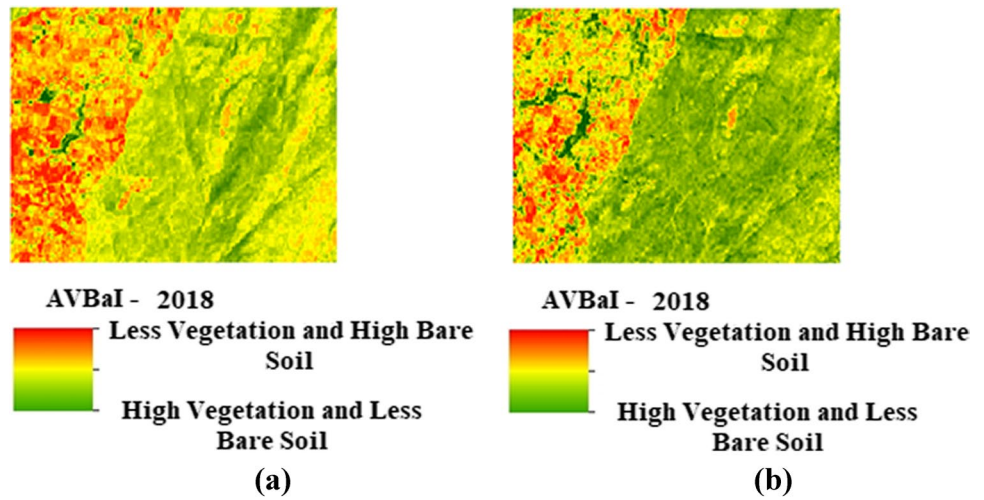
work. Equations (50), (51), and (52) show the calculation of LST for the TIRS bands of the Landsat 7 satellite. Equations (53) to (58) show the calculation of LST for the TIRS bands of the Landsat 8 satellite. Equations (50) and (51) show the conversion of the DN values of each pixel to the brightness temperature through the radiance of the TIRS band. Equation (52) shows the conversion of the values from Kelvin (K) to Celsius (C).

$$L_\lambda = \left(\frac{LMAX_\lambda - LMIN_\lambda}{QCALMAX - QCALMIN} \right) * (QCAL - QCALMIN) + LMIN_\lambda \tag{50}$$

where L_λ denotes the spectral radiance in Watts/(m^2 * sr^2 * μm), $QCAL$ denotes the quantized calibrated pixel value, and $QCALMAX$ and $QCALMIN$ denote the maximum and minimum quantized calibrated pixel value corresponding to the spectral radiance scale $LMAX_\lambda$ and $LMIN_\lambda$ respectively.

Fig. 14 Adjusted Vegetation and Bareness Index map. **a** 2018. **b** 2021

AVBaI - 2018

Less Vegetation and High Bare Soil

High Vegetation and Less Bare Soil

(a)

AVBaI - 2018

Less Vegetation and High Bare Soil

High Vegetation and Less Bare Soil

(b)

Coordinate System: GCS WGS 1984
Datum: WGS 1984
Units: Degree
Area: Javadi Hills, India

LU/LC Map - 2012

High Vegetation
Scarce Vegetation
Bare Soil

(a)

Coordinate System: GCS WGS 1984
Datum: WGS 1984
Units: Degree
Area: Javadi Hills, India

LU/LC Map - 2015

High Vegetation
Scarce Vegetation
Bare Soil

(b)

Fig. 15 Ensemble LU/LC Map of Javadi Hills. **a** 2012. **b** 2015

$$T_K = \frac{K2}{\ln\left(\frac{K1}{L_\lambda} + 1\right)} \quad (51)$$

$$C = T_K - 273.15 \quad (52)$$

where T_K signifies the satellite temperature effectiveness in Kelvin, and the calibration constants $K1, K2$ are represented

in Watts/(m^2 ∗ sr^2 ∗ μm). The calibration constant $K1, K2$ value in Landsat 7 satellite TIRS band is 666.09 and 1282.71.

$$TL_\lambda = ML * QCAL + AL - O_i \quad (53)$$

Equation (53) shows the conversion of TOA spectral radiance (TL_λ) of the TIRS band by using the radiance rescaling factor (ML) in Watts/(m^2 ∗ sr^2 ∗ μm), the quantized

calibrated pixel value is represented by $QCAL$, the radiance additive band rescaling factor AL, and the correction value O_i of TIRS band in Landsat 8 has been used respectively.

$$BT_P = \frac{K2}{\ln\left(\frac{K1}{TL_\lambda} + 1\right)} - 273.15 \tag{54}$$

Equation (54) shows the conversion of the TOA brightness temperature (BT_P) in Celsius (C) from spectral radiance data and the calibration constants $K1$, $K2$ are represented in Watts/($m^2 * sr^2 * \mu m$). The calibration constant K1 and K2 value in Landsat 8 satellite TIRS band is 774.8853 and 1321.0789. Equation (47) shows the expression for $NDVI$ of the Landsat satellite images.

$$PV = ((NDVI - NDVI_{min})/(NDVI_{max} - NDVI_{min}))^2 \tag{55}$$

$$E = 0.004 * PV + 0.986 \tag{56}$$

The land surface emissivity (E) is derived from the results of the NDVI values, where PV denotes the proportion of vegetation, and $NDVI_{max}$ and $NDVI_{min}$ represent the maximum and minimum reflectance values of the NDVI image of Landsat 8. The expressions for PV and E are shown in Eqs. (55) and (56). By using the results of Eqs. (54) to (56), the LST is calculated and expressed in Eq. (57).

$$LST = \frac{BT_P}{\left(1 + \left(\frac{\Delta P}{L}\right) * \ln(E)\right)} \tag{57}$$

$$c2 = \frac{\Delta P}{L} \tag{58}$$

where the wavelength of the emitted radiance is represented as λ, the Planck's constant (pk) value is $6.626 * 10^{-34}$Js, the velocity of light (vl) value is $2.998 * 108$m/s, and the Boltzmann constant (bc) value is $1.38 * 10^{-34}$JK. The use of independent features is to display the impact of dependent LU/LC changes in the Javadi Hills during the process of LU/LC prediction. The LST map of 2012 and 2015 has been extracted for analyzing the effect of LU/LC features of Javadi Hills. The LST provides the skin temperature of the earth, where the high temperature on LU/LC shows less vegetation and the less temperature on the LU/LC shows high vegetation. The key idea of our research is to provide the impact of LST on the E-LU/LC features of vegetation and bare soil. The LST independent map for the years 2012 and 2015 has been used as an input along with the dependent E-LU/LC features for implementing the prediction model to analyze the future LU/LC changes in the area of Javadi Hills. We have shown the LST-independent map of 2012 and 2015 in Fig. 16. The LST map for the years 2018 and 2021 is also calculated and used as the independent feature map during the LU/LC prediction for the years 2024 and 2027. The LST map for the years 2018 and 2021 is shown in Fig. 17.

Ensemble-based LU/LC prediction analysis

The LU/LC prediction has been considered an important study in the file of remote sensing. The importance of LU/LC prediction lies in providing future LU/LC information about the specific region for different periods using the prediction algorithms. Based on the accurate results of the prediction methods, the LU/LC data will be shared with the governmental officials or land resource planners for taking action in protecting the LU/LC environment from great loss and save nature (Anand and Bakimchandra 2020, Bakr et al. 2022). In our research work, we have used MC-MLP neural network model for predicting the LU/LC classes for Javadi Hills, India.

Markovian-Cellular Automata model

In the field of RS environment, the MC analysis is the most used model during the process of LU/LC prediction. The Markovian model was used for calculating the transition probability matrix (TPM) for the LU/LC classes between different periods. The Markovian model is discrete with both spatial and temporal data. The Cellular Automata model represents the nonlinearity and spatially distributed structure for providing the LU/LC patterns from a geospatial satellite image. The Cellular Automata is discrete with space, time, and state and it helps in simulating the spatio-temporal computations (MohanRajan and Loganathan 2021, Huang et al. 2020, Munthali et al. 2020, Armin et al. 2020).

$$Tm = P_{ij} = \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \cdots & \cdots & \cdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix} \tag{59}$$

$$\left(0 \leq P_{ij} < 1 \text{ and } \sum_{j=1}^{n} P_{ij} = 1\right) \tag{60}$$

$$S(t + 1) = P_{ij} * S(t) \tag{61}$$

In Eqs. (59), (60), and (61), the Tm indicates the TPM, P_{ij} indicates the probability of the LU/LC class i to class j, n represents the total number of LU/LC classes, i, j shows the LU/LC class of 1st and 2nd period, S represents the LU/LC information at time t, and $S(t + 1)$ refers to the time variants.

Multilayer Perceptron Neural Network

The MLP neural network is used as a prediction model for providing future LU/LC information for a different study area in the area of RS. The MLP follows the idea of the ANN. The idea of MLP investigates the simple representations of the human brain to solve difficult computational
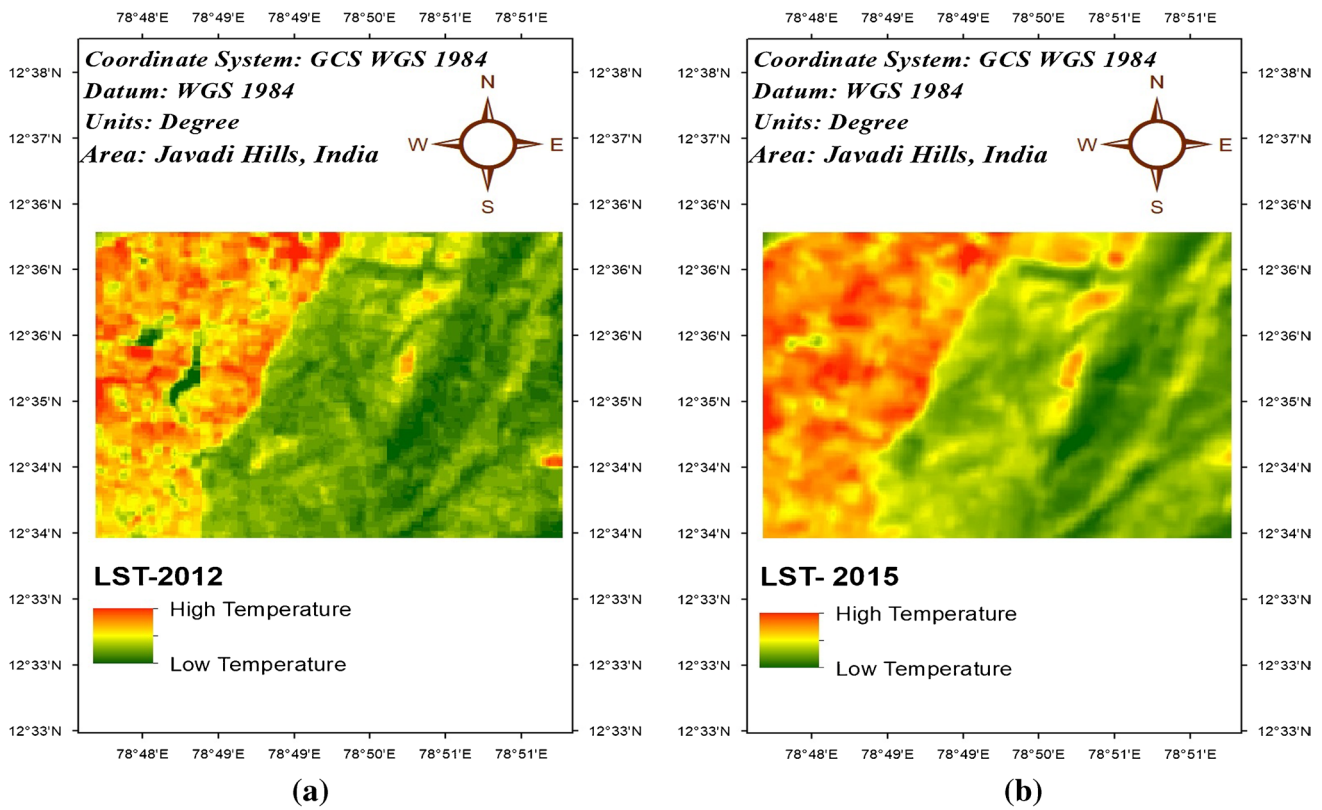
Fig. 16 LST map of Javadi Hills. **a** 2012. **b** 2015

problems. The MLP neural network model maps the non-linearity between the input and the corresponding output vector. In MLP, every neuron node transmits the needed information among them. The input and output processing units of the MLP neural network are interconnected with nodes. As the human brain reads up the rules and provides the results, the MLP also comprises learning rules for providing the results. The initial stage of the MLP neural network model recognizes the data patterns visually or textually. The MLP works with the supervised learning process by training and analyzing the data through the backpropagation method, where the data and its calculations have been performed in a single direction from the input data to the output data. The MLP model has been defined in three different layers and they are input, hidden, and output layers. The user-defined input (Neurons) has been passed in each layer and the inputs are represented as the dependent and independent variables. The computations of the neurons were performed by updating the weights in the hidden layer. Through the selection of the activation functions, the output layers are produced with the minimum error (Kafy et al. 2020, Shooshtari et al. 2020, Bose and Chowdhury 2020). Equations (62), (63), and (64) show the expressions for the MLP neural network.

$$Y = \varphi\left(\sum_{j}^{n} w_j x_j + b\right) \tag{62}$$

$$Y = \varphi\left(w^T x + b\right) \tag{63}$$

$$\varphi(z) = \frac{1}{1 + e^{-x}} \tag{64}$$

where the weights adjusted between the input and the hidden layer are represented as w_j, the neurons or user-defined inputs are represented as x_j, the output layer is represented as Y, φ, and b representing the activation function and bias value respectively. The activation or sigmoid function φ provides the resulting value (either 0 or 1) of the MLP neural network.

The dependent LU/LC map and independent LST map have been used as input for performing the LU/LC prediction analysis. The LU/LC prediction has been performed using the MC-MLP model. The EP-LU/LC map has been obtained by using the outputs of E-LU/LC and the LST map. The algorithm to implement the ensemble-based LU/LC prediction algorithm is presented in Algorithm 2. The overall flow chart of the ensemble-based LU/LC prediction algorithm is shown
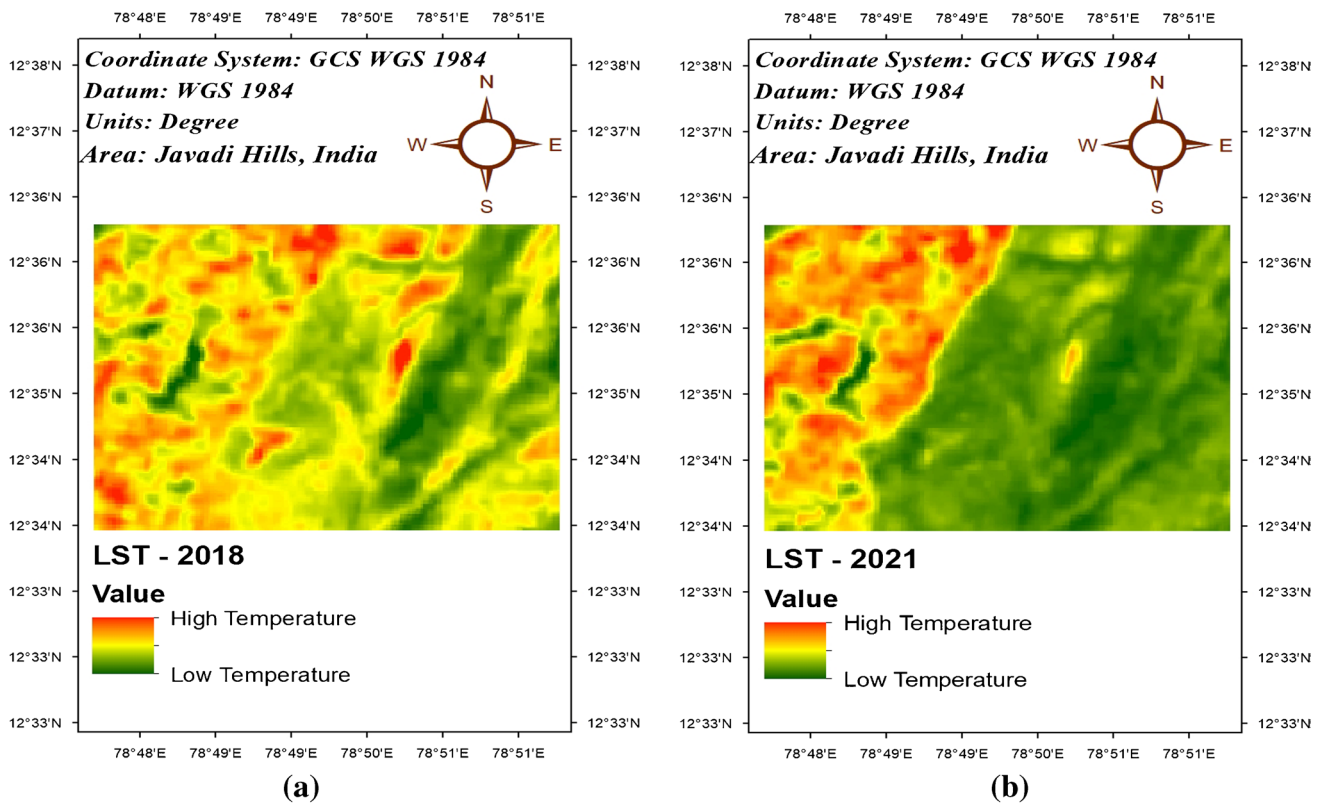
Fig. 17 LST map of Javadi Hills. **a** 2018. **b** 2021

in Fig. 20. We have used the QGIS simulation software for our research to show the predicted LU/LC map of Javadi Hills for the years 2018, 2021, 2024, and 2027. The MOLUSCE simulation tool provides the MC-MLP model for forecasting the LU/LC prediction results. The simulation analysis consists of different components like the input module, correlation analysis, change detection, prediction, and validation module. We have used the dependent and independent inputs of Javadi Hills for the years 2012 and 2015 for analyzing the LU/LC prediction for the years 2018, 2021, 2024, and 2027. The result of the EP-LU/LC map using the MC-MLP neural network model for the years 2018, 2021, 2024, and 2027 is shown in Figs. 18 and 19.

Validation of our proposed ensemble model

The LU/LC predicted results of Javadi Hills were obtained through the ensemble-based prediction model (Fig. 20). The inputs of dependent LU/LC and independent LST map of Javadi Hills for the years 2012 and 2015 have been used for predicting the LU/LC map of Javadi Hills for the years 2018. For acquiring better prediction accuracy, we have used the reclassified AVBaI map of 2018 as the reference map during the process

of validation. The same procedure has been executed for the year 2021. The reference map for 2018 and 2021 is shown in Fig. 21. With inputs from predicted 2018 and 2021 maps, we have predicted the LU/LC map for the years 2024 and 2027. We have acquired good validation accuracy for our proposed ensemble model.

Results

In this section, we explained the experimental setup of our research work and also we delivered the results of our proposed model. The experimentation analysis of the ensemble-based fuzzy HHO-based SMC models was presented briefly along with the growth patterns of the LU/LC analysis of the time series data for the Javadi Hills, India.

Experimental setup

In the field of RS, challenging research on LU/LC prediction has been performed in our research work. The geospatial satellite image of LISS-III and Landsat 7 and 8 has been extracted and used for the area of Javadi Hills, India. All our
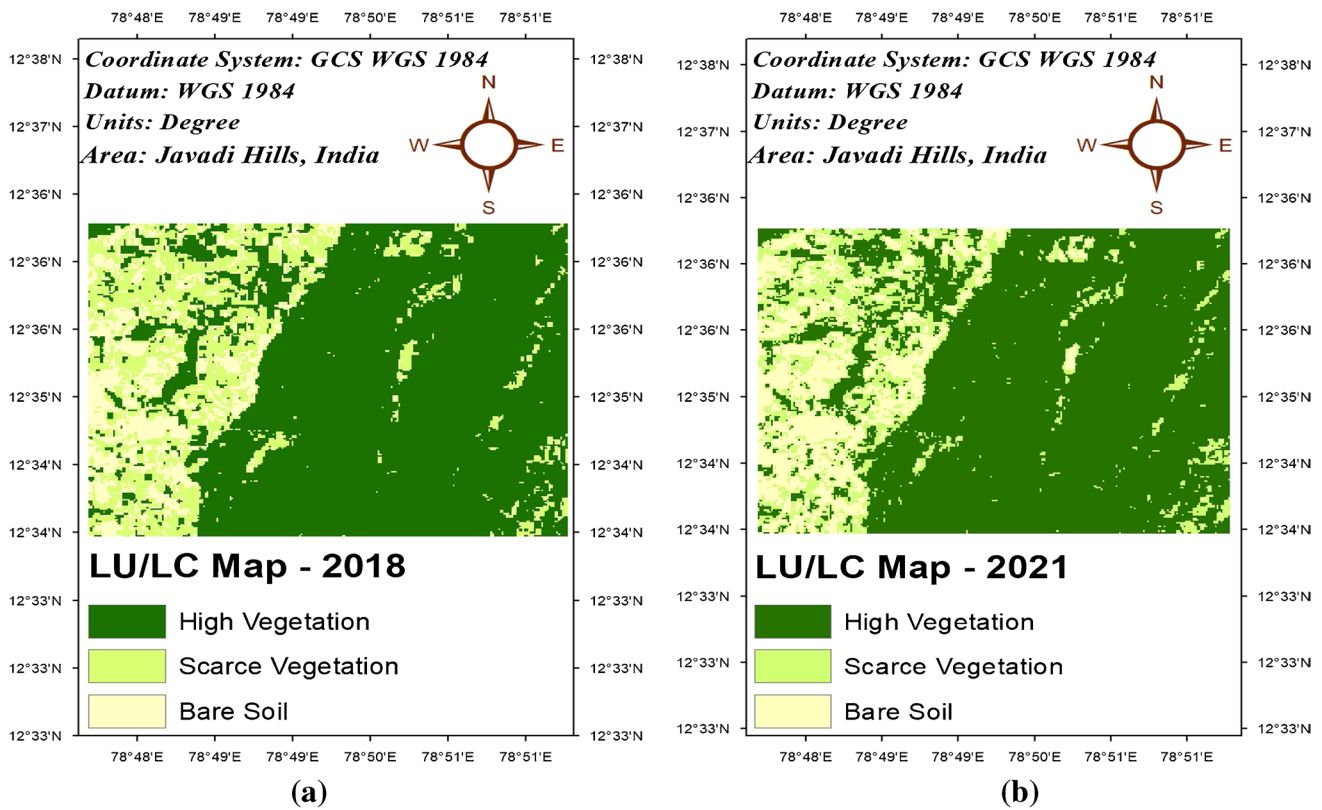
Fig. 18 Ensemble LU/LC predicted map of Javadi Hills. **a** 2018. **b** 2021

experiments were processed by using MATLAB R2021a, Python, QGIS, and ArcGIS geospatial software on the Intel Xeon processor 2.90 GHz CPU along with 128 GB RAM in Windows 10 (64 bit) environment.

Training data and parameter settings

The parameter values for the HHO were determined after several tests. The objective function for the HHO was defined according to the fuzzy membership clustering method, where the cluster centers within the input data are measured using distance metrics. The parameters defined for each HHO are shown in Table 3. The experiments were conducted on pre-processed LISS-III satellite images of Javadi Hills for the years 2012 and 2015. The size of the tested pre-processed image is 256×200 pixels.

The SMC (support vector machine, random forest, and maximum likelihood) were used for training the results of the fuzzy HHO-based feature extraction technique. The results of each SMC have been used separately and the results were obtained. The PCA has been used for fusing the results of all three SMC algorithms. The training parameters used for the SMC are shown in Tables 4, 5, and 6. The training parameters used for the MLP neural network model are shown in Table 7. The training data has been prepared by using the GEE time series data of Javadi Hills. We have randomly created the training samples with 300 random points (location) through the latitude and longitudinal coordinates of Javadi Hills. The training data has been prepared in the CSV (Comma Separated Values) file format and then converted into the vector shapefile. For a better understanding, the attribute table of a few training data is shown in Table 8. For better understanding, the validation of the Google Earth images associated with the training attribute values of the LU/LC map is shown in Fig. 22. The results of the ensemble model of the fuzzy HHO-based SMC algorithm through PCA have been combined with the results of the AVBaI map. The E-LU/LC map of 2012 and 2015 has been obtained and the accuracy assessment has been performed. The accuracy assessment has been performed for the E-LU/LC vegetation and bare soil feature map by using the test or reference data of GEE time-series images.

Based on the good accuracy results, we have used the E-LU/LC vegetation and bare soil map of 2012 and 2015 as the input for finding the EP-LU/LC map. The LST-independent spatial variable map provides the impact on the growth of LU/LC vegetation and bare soil in the area of Javadi Hills. We have used the LST map of 2012 and 2015 as the independent map during the process of LU/LC prediction. The less temperature represents the region with
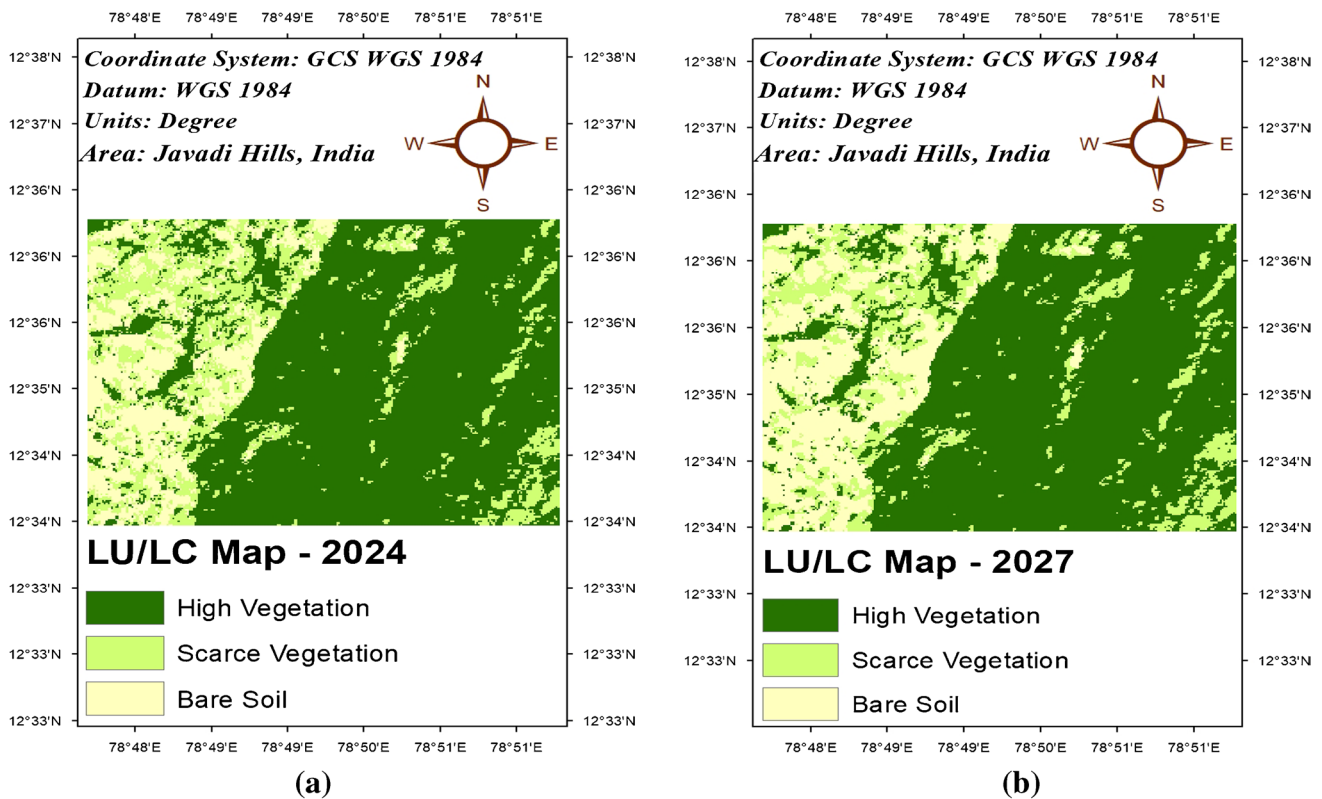
Fig. 19 Ensemble LU/LC predicted map of Javadi Hills. **a** 2024. **b** 2027

high vegetation, and the high temperature represents the region with less vegetation. The AVBaI values range from -1 to $+0.5$. The values obtained were separated into high vegetation, scarce vegetation, and bare soil. All the vegetation and bare soil values have been associated with the LST values. The LST value has been calculated in Celsius. For better understanding, we have shown the few attribute values associated with the dependent and independent features of Javadi Hills in Table 9. The input maps of dependent and independent features with the same geometric coordinates have been used during the process of LU/LC prediction. The MC model has been used for calculating the TPM for the years 2012 to 2021. The transition between each class has been computed and the LU/LC change map has been generated for further analysis of LU/LC prediction. The MLP neural network has been used for modeling the relationship between the dependent and independent maps through the adjustment of different parameters. Once the MLP neural network has been trained, the CA simulation has been performed for identifying the changes that happened during the next state (time-period). The validation of the results has been performed using the reclassified AVBaI map.

Accuracy assessment and computational complexity

The accuracy assessment provides information about the strength of the classification algorithms. The accuracy assessment is performed by comparing the reference or ground truth data with the classified results by building the confusion matrix. The confidence level of computing the LU/LC classification and prediction results generally depends on the statistical measure of the accuracy assessment (Li et al. 2020; Alharthi et al. 2020). In our research work, we have used the overall accuracy metrics for analyzing the correctness of the classification and prediction algorithms. The overall accuracy is considered the complete optimal function for the proposed ensemble model. The ultimate goal of our proposed ensemble model is to obtain better classification accuracy for LU/LC classification problems. Concerning the Google Earth Engine time series data, the accuracy assessment has been performed for the results of E-LU/LC vegetation and bare soil map. The random sampling points (pixel location) from the LU/LC images for the periods 2012 and 2015 were validated with the random sampling points of the Google Earth Engine through the confusion matrix. The results of the accuracy assessment are shown in Table 10.

The fused results were combined with the Adjusted Vegetation and Bareness Index (AVBaI) map and the final ensemble LU/LC (E-LU/LC) map was achieved for the years 2012 and 2015 with good classification accuracy, with an average of 95.275%. The results of the ensemble

Algorithm 2 Algorithm to implement the ensemble-based LU/LC prediction algorithm

Inputs (I_P): LU/LC Feature Map, Landsat Satellite Bands

Output (O_P): Predicted LU/LC map

Begin

1   Prepare the Adjusted Vegetation and Bare Soil Map (AVBaI)

2   Initialize the Landsat Bands (RED, SWIR, TIRS, NIR)

3   Calculate the NDVI and NDBaI for the Landsat Bands

4   Combine the NDVI and NDBaI features

5   Return AVBaI

6   Prepare the Land Surface Temperature Map (LST)

7   Initialize the Landsat Bands (RED, SWIR, TIRS, NIR)

8   Calculate the LST map

9   Return LST

10   Prepare Dependent LU/LC Feature Map

11   Combine the LU/LC Feature Map with the AVBaI Map

12   Analyze the dependent LU/LC vegetation and bare soil feature map

13   Calculate the accuracy for the dependent feature map using the Google Earth Engine Image

14   Analyze the changes that happened in LU/LC vegetation and bare soil feature map

15   Return Dependent LU/LC Feature Map

16   Perform LU/LC Prediction Model

17   Implement the LU/LC prediction by using the inputs of the dependent LU/LC feature map and the independent LST Map

18   Initialize the inputs for LU/LC Prediction

19   Compute the Transition Probabilities for LU/LC classification map using the Markovian model

20   Execute the prediction algorithm using Multilayer Perceptron (MLP) neural network

21   CA simulation has been performed for identifying the changes that happened during the next state (Time-Period).

22   Obtain the LU/LC predicted map

23   Return LU/LC Predicted Map

End

model of fuzzy HHO-based SMC through PCA and AVBaI techniques were used for predicting the LU/LC map of Javadi Hills. The E-LU/LC vegetation and bare soil map of 2012 and 2015 have been used as the dependent map, and the LST map has been used as an independent map during the process of LU/LC prediction. The LU/LC prediction has been performed using the ensemble-based MC-MLP neural network. The LU/LC map of 2012 and 2015 has been used as input along with the LST map for determining the LU/LC predicted map for 2018. The EP-LU/LC map of 2018 has been validated with the reference data of reclassified AVBaI map of 2018 to determine the correctness of our proposed model. The same procedure has been followed for predicting the 2021 map. With the available data, we have prepared the reference map for 2018 and 2021 through AVBaI. With inputs of EP-LU/LC map for 2018 and 2021, we have predicted the LU/LC map for the years 2024 and 2027. The validation has not been performed for 2024 and 2027. With the high percentage of correctness of 2018 and 2021, we have projected the results for 2024 and 2027. The validated results of the predicted LU/LC map for the years 2018 and 2021 using the MC-MLP model are presented in Table 11. The result for the ensemble-based MC-MLP model provides good accuracy with 95.616% for 2018 and 95.910% for 2021. The result of the EP-LU/LC map provides an average prediction accuracy of 95.763%.

The total running time taken for executing the algorithm has been defined through computational complexity. Each algorithm varies with computational complexity in terms of input data, training data, and tuning parameters. The computational complexity of our proposed algorithm is estimated and shown in Table 12. We have shown the elapsed time for each process of our proposed algorithm. The execution time taken for running our proposed

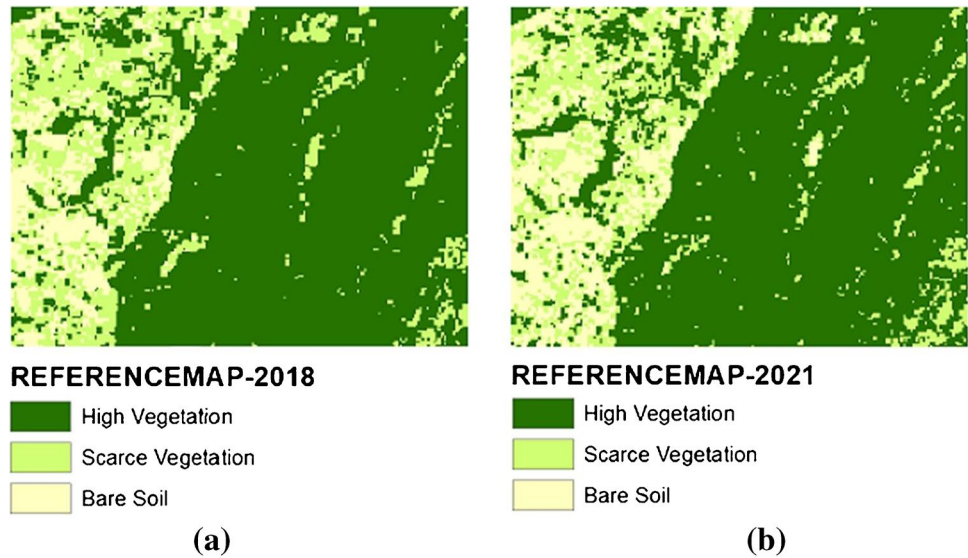**Fig. 20** Flow chart of the ensemble-based LU/LC prediction model
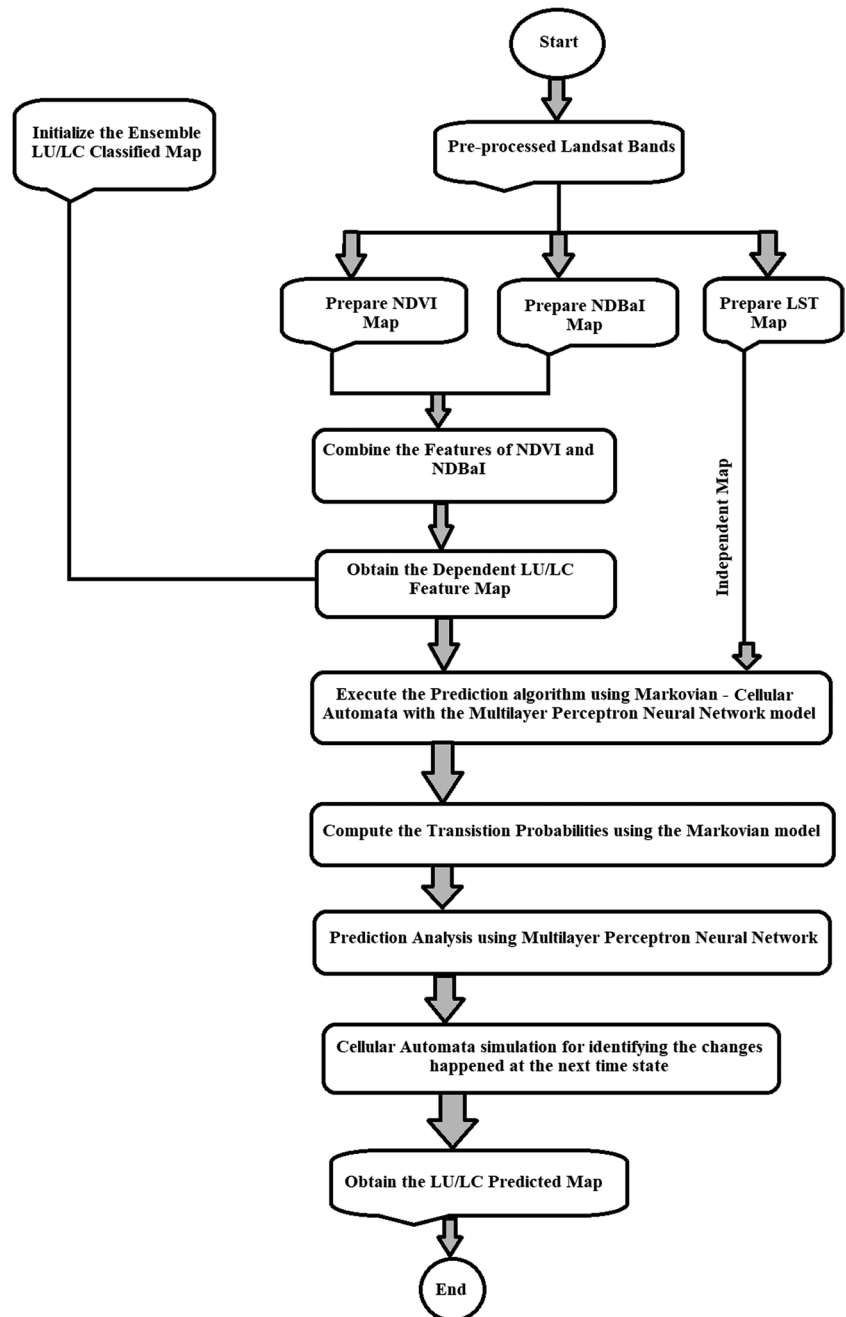
ensemble model is 1110.27 s. The results have been calculated in MATLAB and python by using the function tic-toc. The back-end tool of ArcGIS and QGIS also provides the elapsed time through the inbuilt python package. We have also compared the execution time of our fuzzy HHO algorithm with other optimization algorithms that include CSO, ACO, ABC, and PSO. The overall execution time taken by HHO shows good results when compared with other optimization algorithms, and the overall results along with the accuracies are shown in Table 13.

Correlation analysis

The correlation analysis is performed between the spatial variables that affect the LU/LC of the particular area. Pearson correlation analysis provides the statistical relationship between the spatial variables of the LU/LC analysis. The linear association varies from $+1$ to -1, which shows the positive and the negative correlation between the spatial variables. The correlation values help for expressing the strength of the association between the spatial variables (Riva and Nielsen

Fig. 21 Reference map of
Javadi Hills

REFERENCEMAP-2018

☐ High Vegetation
☐ Scarce Vegetation
☐ Bare Soil

(a)

REFERENCEMAP-2021

☐ High Vegetation
☐ Scarce Vegetation
☐ Bare Soil

(b)

Table 3 Parameter setting of each HHO technique

Input image (size)	Objective function	Optimization algorithm	Parameter	Value
Pre-processed satellite image (256*200)	Fuzzy membership clustering	HHO	Population size	10
			Maximum iteration	100
			Number of clusters	3
			Energy of rabbit	$E \in [0, 2]$
			Constant	$\beta = 1.5$

2021; Khamchiangta and Dhakal 2020). In our research work, we have used the Pearson correlation analysis for displaying the relationship between the LST and AVBaI maps of Javadi Hills, India. Equation (65) provides the expression for calculating the association between the spatial variables.

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \tag{65}$$

where r represents the correlation coefficient, x_i represents the values of the x variable, y_i represents the values of the y variable, $\bar{x}$ represents the mean values of the x variable, and $\bar{y}$ represents the mean values of the y variable. Pearson correlation analysis was constructed between the spatial LST maps that affect the impact of the LU/LC environment.

For signifying the impact of the LU//LC environment of Javadi Hills, we have constructed the correlation analysis

Table 4 Parameter setting for the SVM model

Parameters	Value
Input image type	Raster
Input image size	256*200 (pixels)
Training samples (random sampling)	300 points
Training data format	Shapefile (vector storage format of geographic location)
Kernel	Linear
C (cost parameter)	1

Table 5 Parameter setting for the random forest classification model

Parameters	Value
Input image type	Raster
Input image size	256*200 (pixels)
Training samples (random sampling)	300 points
Training data format	Shapefile (vector storage format of geographic location)
n_estimators (number of trees)	100
Criterion	Gini
max_depth (maximum depth of the tree)	5
Minimum number of samples in each node	10

Table 6 Parameter setting for the maximum likelihood classification model

Parameters	Value
Input image type	Raster
Input image size	256*200 (pixels)
Training samples (random sampling)	300 points
Training data format	Shapefile (vector storage format of geographic location)
Probability threshold (reject_fraction)	0
a_priori_probabilities	Equal

Table 7 Parameter setting for the Multilayer Perceptron Neural Network

Parameters	Value
Input image type	Raster
Input image size	256*200 (pixels)
Activation function	Sigmoid
Learning rate	0.010
Optimizer	Stochastic gradient descent
Momentum	0.9
Hidden layer	20
Number of iterations	100

Table 8 Attribute table of the training data values

Longitude	Latitude	Class value	Class label
78.854097	12.596853	1	High vegetation
78.810688	12.599685	3	Bare soil
78.81915	12.569452	1	High vegetation
78.809091	12.602053	3	Bare soil
78.807748	12.573319	2	Scarce vegetation
78.832464	12.590825	1	High vegetation
78.816307	12.589858	2	Scarce vegetation
78.805651	12.585313	3	Bare soil
78.808002	12.603136	2	Scarce vegetation
78.854301	12.599941	1	High vegetation
78.803583	12.595604	2	Scarce vegetation
78.808344	12.604199	2	Scarce vegetation
78.810493	12.584733	3	Bare soil
78.835309	12.583966	1	High vegetation
78.803979	12.583208	3	Bare soil

between the LST (2012, 2015, 2018, and 2021), and the AVBaI (2012, 2015, 2018, and 2021). From the different analyses for our study area, we found that the correlation that occurred between the LSTs of 2012 and 2021 provides a strong positive maximum value of 0.929343. The main significance of our research work provides no negative correlation values. Table 14 represents the Pearson correlation analysis of the spatial variables.

Growth pattern of LU/LC analysis

The LU/LC change information has been computed between the two different periods and also the information about the LU/LC change is provided to land resource managers or government officials to take appropriate actions for protecting the LU/LC environment (Chamling and Biswajit 2020, Chemura et al. 2020).

We have used the LU/LC change formula for calculating the area statistics of our study area. Equation (66) provides the expression for calculating the percentage of the LU/LC change (P_{change}) between the LU/LC map of two different periods (T_1 and T_2).

$$P_{\text{change}} = \left(\frac{T_2 - T_1}{T_1} \right) * 100 \tag{66}$$

Our research is based on ensemble-based prediction models for analyzing the LU/LC patterns of Javadi Hills from the period 2012 to 2027. From the deep analysis and validation, we found good classification accuracy for the E-LU/LC vegetation and bare soil map of Javadi Hills for 2012 and 2015. The E-LU/LC vegetation and bare soil map have been used as the dependent feature map during the process of LU/LC prediction analysis using MC-MLP Neural Network model. The LST map has been used as the independent spatial variable map during the LU/LC prediction analysis. Our prediction model has shown a higher percentage of correctness. The TPM has been calculated using the MC model for finding the impact of LU/LC changes on the dependent LU/LC map for the years 2012 to 2021 during the process of LU/LC prediction. With the results of the LU/LC change analysis from the MC model, the MLP neural network has been performed for the input maps, and the predicted results had been analyzed for the area of Javadi Hills. The TPM results for the dependent LU/LC map from 2012 to 2021 are shown in Table 15.

In our research, we have shown the growth patterns for the area of Javadi Hills from the period 2012 to 2027. The area statistics of the LU/LC map is shown in Table 16. The percentage of change that happened in the LU/LC classes is analyzed and shown in Table 17. At the end of our results, we analyzed that for every 3 years (2012 to 2027), the LU/LC change had happened. The continuous change of vegetation from high to low in the area of Javadi Hills from the period of 2012 to 2027 indicates that the necessary actions need to be taken by the land resource planners to take actions in protecting nature from huge loss. We had shown the graph in

Fig. 22 Validation of Google Earth Map of Javadi Hills
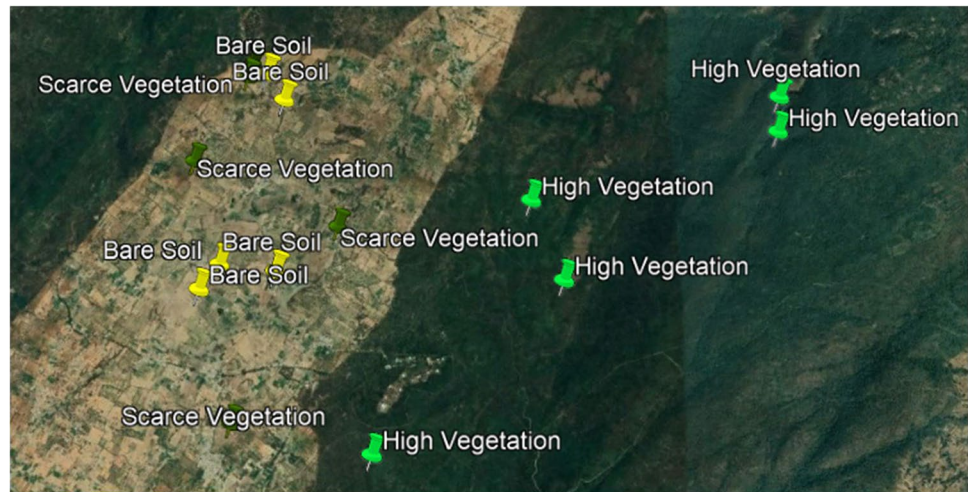
Table 9 Values associated with dependent and independent features of Javadi Hills

Year	LU/LC indices values range		LU/LC class
	Adjusted VBaI (-1 to $+0.5$)	LST (°C)	
2012	$-0.69 \leq AVBaI < -0.13$	$26.36 \leq LST < 35.45$	High vegetation
	$-0.13 \leq AVBaI < -0.008$	$35.45 \leq LST < 37.75$	Scarce vegetation
	$-0.008 \leq AVBaI \leq 0.2$	$37.75 \leq LST \leq 42.60$	Bare soil
2015	$-0.62 \leq AVBaI < -0.14$	$31.38 < = LST < 36.95$	High vegetation
	$-0.14 \leq AVBaI < -0.07$	$36.95 \leq LST < 38.91$	Scarce vegetation
	$-0.07 \leq AVBaI \leq 0.1$	$38.91 \leq LST \leq 42.23$	Bare soil
2018	$-0.68 \leq AVBaI < -0.021$	$22.57 \leq LST < 25.63$	High vegetation
	$-0.021 \leq AVBaI < -0.0008$	$25.63 \leq LST < 26.48$	Scarce vegetation
	$-0.0008 \leq AVBaI \leq 0.1$	$26.48 \leq LST \leq 29.91$	Bare soil
2021	$-0.61 \leq AVBaI < -0.05$	$25.80 \leq LST < 31.16$	High vegetation
	$-0.05 \leq AVBaI < -0.02$	$31.16 \leq LST < 32.83$	Scarce vegetation
	$-0.02 \leq AVBaI \leq 0.1$	$32.83 \leq LST \leq 36.59$	Bare soil

Table 10 LU/LC accuracy assessment for the feature extraction-based LU/LC map of Javadi Hills

LU/LC vegetation and bare soil map		2012	2015
Classified pixel count (51,200 pixels)	High vegetation	34,213	34,095
	Scarce vegetation	10,491	7719
	Bare soil	6496	9386
Overall accuracy		96.11%	94.44%
Overall average accuracy		95.275%	

Table 11 Validation for the EP-LU/LC map

S. No	Input		Output predicted LU/LC map	Reference AVBaI map	Percentage of correctness (accuracy)	Average prediction accuracy
	Dependent map	Independent map				
1	LU/LC map (2012, 2015)	LST map (2012, 2015)	2018	2018	95.616%	95.763%
2	LU/LC map 2015, predicted LU/LC map 2018	LST map (2015, 2018)	2021	2021	95.910%	

Table 12 Computational complexity of our proposed ensemble model

Techniques/algorithms		Computational time (s)
Pre-processing		58.91
Feature extraction (fuzzy HHO-based feature extraction)		142.04
Classification (SMC)	SVM	73.52
	RF	68.24
	ML	66.02
Extracting LU/LC map through PCA		12.62
Estimation of LU/LC indices	NDVI	16.81
	NDBaI	15.68
	LST	18.26
Assessment of AVBaI		4.64
E-LU/LC map extraction (combining AVBaI and LU/LC map)		5.52
EP-LU/LC map using MC-MLP neural network		628.01
Total elapsed time		1110.27

Table 13 Comparative analysis of the accuracy and execution time for the fuzzy HHO with other optimization algorithm-based ensemble prediction model

Algorithms	Overall average accuracy	Execution time
Proposed fuzzy HHO-based ensemble prediction model	**95.763**%	**1110.27**
Fuzzy CSO-based ensemble prediction model	95.138%	1102.34
Fuzzy PSO-based ensemble prediction model	94.841%	1097.072
Fuzzy ACO-based ensemble prediction model	93.948%	1076.28
Fuzzy ABC-based ensemble prediction model	94.047%	1092.86

Table 14 Pearson correlation analysis of the spatial variables for Javadi Hills, India

LU/LC feature maps	LST (2012)	LST (2015)	LST (2018)	LST (2021)	AVBaI (2012)	AVBaI (2015)	AVBaI (2018)	AVBaI (2021)
LST (2012)	1	0.920753	0.720835	0.929343	0.837789	0.755928	0.801087	0.732425
LST (2015)		1	0.710623	0.908381	0.800599	0.776331	0.759473	0.664009
LST (2018)			1	0.698723	0.646244	0.529183	0.652915	0.574943
LST (2021)				1	0.8007254	0.738074	0.777385	0.770513
AVBaI (2012)					1	0.825109	0.872812	0.800342
AVBaI (2015)						1	0.788863	0.747161
AVBaI (2018)							1	0.812991
AVBaI (2021)								1

Table 15 Transition probability matrix for LU/LC maps of Javadi Hills during the years 2012–2021

LU/LC class	2012–2015			2015–2018			2018–2021		
	High vegetation	Scarce vegetation	Bare soil	High vegetation	Scarce vegetation	Bare soil	High vegetation	Scarce vegetation	Bare soil
High vegetation	0.936282	0.049718	0.014001	0.940756	0.043095	0.016149	0.927042	0.054118	0.018840
Scarce vegetation	0.179392	0.489181	0.331427	0.297837	0.512631	0.189532	0.238531	0.519451	0.242018
Bare soil	0.012706	0.140724	0.846569	0.092074	0.410941	0.496985	0.149643	0.342844	0.507513

Table 16 Area statistics of the classified and predicted LU/LC map of Javadi Hills

LU/LC class	Year					
	2012	2015	2018	2021	2024	2027
	Area (ha)	Area (ha)	Area (ha)	Area (ha)	Area (ha)	Area (ha)
High vegetation	1598.14	1589.74	1644.84	1674.80	1538.58	1505.46
Scarce vegetation	493.92	357.79	433.68	407.55	474.32	413.85
Bare soil	295.83	440.36	309.37	305.54	374.99	470.58
Total	2387.89	2387.89	2387.89	2387.89	2387.89	2387.89

Table 17 LU/LC growth pattern of Javadi Hills from 2012 to 2027

LU/LC class	Year				
	2012–2015	2015–2018	2018–2021	2021–2024	2024–2027
	Area (%)	Area (%)	Area (%)	Area (%)	Area (%)
High vegetation	−0.52	3.46	1.82	−8.13	−2.15
Scarce vegetation	−27.56	21.21	−6.02	16.38	−12.74
Bare soil	48.85	−29.74	−1.23	22.73	25.49

Fig. 23 LU/LC growth pattern of Javadi Hills from the years 2012 to 2027

Growth Pattern of Javadi Hills

	Area (ha) 2012	Area (ha) 2015	Area (ha) 2018	Area (ha) 2021	Area (ha) 2024	Area (ha) 2027
High Vegetation	1598.14	1589.74	1644.84	1674.8	1538.58	1505.46
Scarce Vegetation	493.92	357.79	433.68	407.55	474.32	413.85
Bare Soil	295.83	440.36	309.37	305.54	374.99	470.58
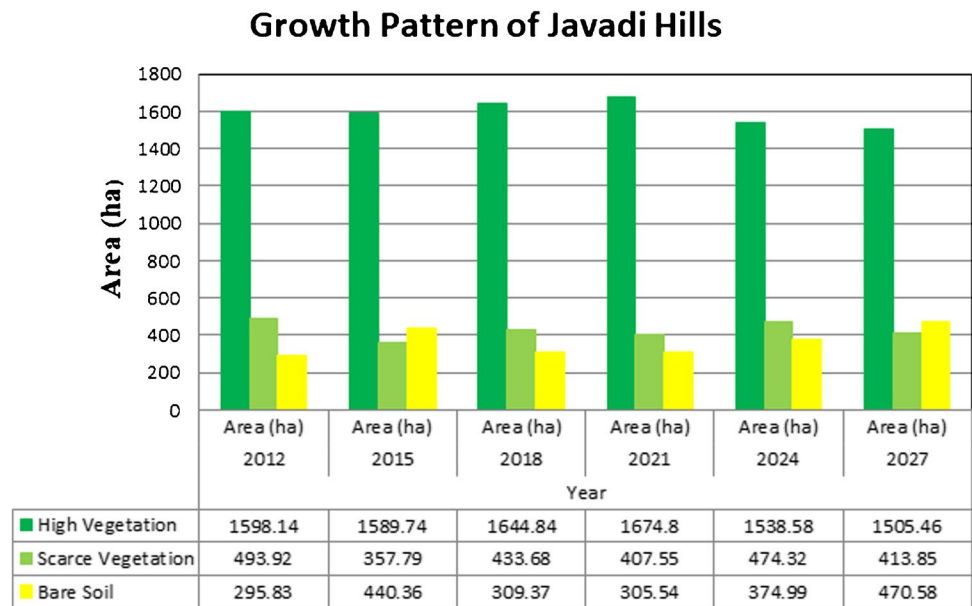
Fig. 23, which displays the LU/LC changes in the area of Javadi Hills from the years 2012 to 2027.

Discussion

The LU/LC change prediction analysis has been performed for many decades and the importance of this research is to assist the concerned land resource planners or government officials to take necessary actions in the protection of nature from huge loss. From the brief observation by the different researchers in the field of remote sensing, LU/LC change has been analyzed through different types of datasets that include satellite images, aerial images, and the images taken when carrying out the ground truth or field survey. The advantage of using satellite images is to find the LU/LC information for the larger areas, and also for the areas which are inaccessible through airborne or ground surveys. The multispectral images provide the advantage in performing the LU/LC change prediction analysis since it provides a less number of bands than hyperspectral. The execution time and choosing of bands for performing the LU/LC change prediction analysis look easier and more effective when choosing the multispectral bands. The

interesting tasks when performing the LU/LC change prediction analysis are as follows:

i. The selection of precise satellite systems for a particular location through geographical coordinates is an interesting task.
ii. The selection of the classification and prediction algorithms and their implementation through different training data, and the tuning of parameters is the most challenging and interesting part of LU/LC prediction research.
iii. Finding the reference data of each pixel in the satellite image to calculate the classification and prediction accuracy is an important part of the LU/LC prediction research.
iv. The use of more than one satellite information for one particular location in predicting the LU/LC change remains the challenge.

All the above interesting tasks were performed effectively in this research work. Instead of using one satellite information, this research work provides the information from two satellites that includes LISS-III and Landsat. The satellite images for the area of Javadi Hills have been collected from both satellite systems. LISS-III provides only 4 bands (Green, RED, NIR, and SWIR), and the information from LISS-III has been used for finding the LU/LC classes by using the ensemble model fuzzy HHO-based SMC through PCA analysis. The AVBaI map has been extracted by using the Landsat bands (NIR, SWIR, TIRS, and RED). The TIRS bands are not available in LISS-III and hence the multi-satellite system provides an advantage in using the TIRS bands of the Landsat satellite system. Also, the LST is the only independent spatial variable used for providing the impact of LU/LC changes in Javadi Hills by using the dependent E-LU/LC map. This work also provides the advantage of implementing both unsupervised metaheuristic optimization and supervised machine learning algorithms. We proposed the ensemble model of unsupervised fuzzy HHO-based SMC through PCA and AVBaI for avoiding misclassification errors and to provide good accuracy for the E-LU/LC classification model. Each optimization technique varies with accuracy and execution time. The importance of using the HHO algorithm over other optimization techniques HHO is an effective optimization method with many active and time-varying stages of exploration and exploitation. The searching strategies of hawks provide a flexible structure and high-quality results. The main idea behind the HHO technique is based on the support and chasing style of hawks with their surprise jump. The escaping energy parameters of the HHO have an active randomized time-varying nature for improving and harmonizing the exploratory and exploitive patterns. The importance of HHO displays a sequence

of searching tactics and then, it goes for the best movement for catching the rabbits which are said to be the optimum solution. The importance of HHO over other methods is that the Harris hawks can chase the detected rabbits to exhaustion, which helps them to increase their vulnerability. We have also compared the execution time of our fuzzy HHO algorithm with other optimization algorithm-based ensemble prediction model and the results are shown in Table 13. The fuzzy HHO-based ensemble prediction model provides good accuracy results with less execution time. This work also provides an advantage in fusing the results of three effective LU/LC classifiers (SVM, RFC, and MLC) through PCA for providing all the information regarding the LU/LC classes without any loss. The use of PCA provides the advantage of fusing the results without any information loss, and also it helps in reducing the noise level. The AVBaI has been calculated from the vegetation and bareness index map. The results of AVBaI were combined with the fused LU/LC map and it hence shows the importance of using multiple satellite data for one study area.

The use of more than one spatial variable like slope, aspect, hill shade, and distance variables that include distance from the road, forest edge, water bodies, and built-up areas has been overcome by using an LST map in predicting the vegetation and bare-soil map for any study area around the world. The LST map has been used as the only spatial variable since it provides the advantage of finding the impact of LU/LC changes happened in the vegetation and bare soil regions. The LST map has been extracted from Landsat TIRS bands and it displays the mixture of the hot and cold temperatures of the LU/LC of Javadi Hills. The results were measured using Celsius where the high temperature (hot) is considered as the bare soil regions and the less temperature (cold) regions are considered to be the vegetation regions. For better understanding, we have shown the few attribute values associated with the dependent and independent features of Javadi Hills in Table 9. The impact of the known past and present LST on the E-LU/LC map provides the advantage in finding the unknown LU/LC predicted map. The proposed ensemble model will help in finding the LU/LC prediction map through the processes of the MC-MLP neural network model. For indicating the impact of the LU//LC environment of Javadi Hills, we have calculated the Pearson correlation analysis between the LST (2012, 2015, 2018, and 2021) and the AVBaI (2012, 2015, 2018, and 2021) map of Javadi Hills, India. The main importance of our research work provides no negative correlation values. Table 14 represents the Pearson correlation analysis of the spatial variables. The use of the Markovian-Cellular Automata model provides an advantage in finding the transition probabilities between the LU/LC classes of time series data and analyzing the future LU/LC patterns of Javadi Hills, India. The MLP neural network model is

used in training the dependent (E-LU/LC map) and independent (LST map) during the process of LU/LC prediction. In our proposed work, we have validated the predicted maps with the reclassified AVBaI maps of 2018 and 2021. The validation results provide good accuracy for the years 2018 and 2021. Instead of validating the single LU/LC predicted map, we have checked and validated the two LU/LC predicted maps to manifest our model effectively. The Google Earth Engine images provide the advantage of finding reference maps during the process of accuracy assessment. Also, the information about the changes that happened in the vegetation and bare soil region of Javadi Hills will assist the concerned land resource planners or forest departments to take necessary actions in protecting the vegetation from a huge loss, especially in the forest-covered regions of the hill stations.

Conclusion

This research provides the growth patterns of LU/LC analysis of Javadi Hills, India. We have used the multispectral LISS-III geospatial satellite images for the periods of 2012 and 2015 as input for analyzing the LU/LC prediction for the years 2018, 2021, 2024, and 2027 for the area of Javadi Hills. The impact of LU/LC change has been analyzed through the LST for the years 2012, 2015, 2018, and 2021 of the Javadi Hills. The preprocessing of the geospatial satellite images has been performed for rectifying the radiometric, atmospheric, geometric, and topographic errors in the geospatial satellite images. The proposed ensemble model of fuzzy HHO-based SMC through PCA has been used for extracting the LU/LC classes from the geospatial LISS-III satellite images of Javadi Hills for the years 2012 and 2015. We have calculated the AVBaI map through the Landsat bands. The AVBaI map obtained from the Landsat satellite images was combined with the results of the fused LU/LC map of the ensemble model. Our proposed ensemble model provides good accuracy for the input E-LU/LC maps of 2012 and 2015 with an average of 95.27%. The time series Google Earth Engine platform has been used as reference data during the processes of accuracy assessment. The results of the dependent E-LU/LC maps (2012 and 2015) and the independent LST maps of Javadi Hills for the years 2012, 2015, 2018, and 2021 have been used during the predictive analysis of the MC-MLP neural network model. We had validated the EP-LU/LC map of 2018 and 2021 with the reclassified AVBaI reference data map. The result provides a good percentage of correctness with an average of 95.763%. The LU/LC change analysis has been calculated for each LU/LC map from 2012 to 2027, and we found that the high vegetation area of Javadi Hills has decreased and the less vegetated areas have increased simultaneously. The LU/LC

information about the vegetation changes in our study area will insist the government administrators of land resource management take action for the protection of the LU/LC environment. The future scope of our research is to provide the LU/LC information about the water bodies and urbanized areas by including additional spatial variables like Normalized Difference Water Index and Normalized Difference Built-up Index.

Declarations

References

Abualigah L et al (2021) Hybrid Harris hawks optimization with differential evolution for data clustering. In: Oliva, D., Houssein, E.H., Hinojosa, S. (eds) Metaheuristics in machine learning: theory and applications. Studies in computational intelligence. Springer, Cham, 967:267-299. https://doi.org/10.1007/978-3-030-70542-8_12

Alharthi A et al (2020) Remote sensing of 10 years changes in the vegetation cover of the northwestern coastal land of Red Sea, Saudi Arabia. Saudi J Biol Sci 27(11):3169–3179. https://doi.org/10.1016/j.sjbs.2020.07.021

Al-Qaisi L et al (2021) A review of evolutionary data clustering algorithms for image segmentation. In: Aljarah, I., Faris, H., Mirjalili, S. (eds) Evolutionary data clustering: algorithms and applications. Algorithms for intelligent systems. Springer, Singapore, pp 201–214. https://doi.org/10.1007/978-981-33-4191-3_9

Alshari EA, Bharti WG (2021) Development of classification system for LULC using remote sensing and GIS. Global Transitions Proc 2(1):8–17. https://doi.org/10.1016/j.gltp.2021.01.002

Amoakoh AO et al (2021) Tropical peatland classification using multisensor Sentinel imagery and random forest algorithm in Greater Amanzule, Ghana. IEEE IGARSS, pp 5910–5913. https://doi.org/10.1109/IGARSS47720.2021.9554615

Anand V, Bakimchandra O (2020) Future land use land cover prediction with special emphasis on urbanization and wetlands. Remote Sens Lett 11(3):225–234. https://doi.org/10.1080/2150704X.2019.1704304

Armin M, Majidian M, Kheybari VG (2020) Land use/land cover change detection and prediction in the Yasouj City suburbs in Kohgiluyeh Va Boyerahmad Province in Iran. Arid Ecosyst 10(3):203–210. https://doi.org/10.1134/S2079096120030129

Arora J, Meena T (2020) An enhanced spatial intuitionistic fuzzy c-means clustering for image segmentation. Proc Comput Sci 167:646–655. https://doi.org/10.1016/j.procs.2020.03.331

Bakr N, Morsy I, Yehia HA (2022) Spatio-temporal land use/cover detection and prediction in Mediterranean region: a case study in Idku ecosystem, Egypt. Remote Sens Appl: Soc Environ 25:100673. https://doi.org/10.1016/j.rsase.2021.100673

Balha A et al (2021) A comparative analysis of different pixel and object-based classification algorithms using multi-source high spatial resolution satellite data for LULC mapping. Earth Sci Inform 14(4):2231–2247. https://doi.org/10.1007/s12145-021-00685-4

Batur E, Derya M (2018) Assessment of surface water quality by using satellite images fusion based on PCA method in the Lake Gala, Turkey. IEEE Trans Geosci Remote Sens 57(5):2983–2989. https://doi.org/10.1109/TGRS.2018.2879024

Bauer ME (2020) Remote sensing of environment: history, philosophy, approach and contributions, 1969–2019. Remote Sens Environ 237:111522. https://doi.org/10.1016/j.rse.2019.111522

Benbriqa H et al (2021) Deep and ensemble learning based land use and land cover classification. ICCSA, Springer, Cham, pp 588–604. https://doi.org/10.1007/978-3-030-86970-0_41

Bose A, Chowdhury IR (2020) Monitoring and modeling of spatio-temporal urban expansion and land-use/land-cover change using markov chain model: a case study in Siliguri Metropolitan area, West Bengal, India. Model Earth Syst Environ 6(4):2235–2249. https://doi.org/10.1007/s40808-020-00842-6

Bosquilia RWD, Muller-Karger FE (2021) Analysis of the wetland classification using optical satellite imagery in the environmental protection area of Guaraqueçaba, PR Brazil. J S Am Earth Sci 112:103615. https://doi.org/10.1016/j.jsames.2021.103615

Chamling M, Biswajit B (2020) Spatio-temporal patterns of land use/land cover change in the Bhutan-Bengal foothill region between 1987 and 2019: study towards geospatial applications and policy making. Earth Syst Environ 4(1):117–130. https://doi.org/10.1007/s41748-020-00150-0

Charrua AB et al (2021) Impacts of the Tropical Cyclone Idai in Mozambique: a multi-temporal Landsat satellite imagery analysis. Remote Sens 13(2):201. https://doi.org/10.3390/rs13020201

Chemura A et al (2020) The impact of land-use/land cover changes on water balance of the heterogeneous Buzi sub-catchment, Zimbabwe. Remote Sens Appl: Soc Environ 18:100292. https://doi.org/10.1016/j.rsase.2020.100292

Cheng G et al (2020) Remote sensing image scene classification meets deep learning: challenges, methods, benchmarks, and opportunities. IEEE J Sel Top Appl Earth Obs Remote Sens 13:3735–3756. https://doi.org/10.1109/JSTARS.2020.3005403

Chughtai AH, Abbasi H, Ismail RK (2021) A review on change detection method and accuracy assessment for land use land cover. Remote Sens Appl: Soc Environ 22:100482. https://doi.org/10.1016/j.rsase.2021.100482

da Silva VS et al (2020) Methodological evaluation of vegetation indexes in land use and land cover (LULC) classification. Geol Ecol Landsc 4(2):159–169. https://doi.org/10.1080/24749508.2019.1608409

Dou P et al (2021) Time series remote sensing image classification framework using combination of deep learning and multiple classifiers system. Int J Appl Earth Obs Geoinf 103:102477. https://doi.org/10.1016/j.jag.2021.102477

Feng Y, Negrón-Juárez RI, Chambers JQ (2020) Remote sensing and statistical analysis of the effects of hurricane María on the forests of Puerto Rico. Remote Sens Environ 247:111940. https://doi.org/10.1016/j.rse.2020.111940

Floreano IX, de Moraes LAF (2021) Land use/land cover (LULC) analysis (2009–2019) with Google Earth Engine and 2030 prediction using Markov-CA in the Rondônia State, Brazil. Environ Monit Assess 193(4):1–17. https://doi.org/10.1007/s10661-021-09016-y

Gavrilovskaya N et al (2021) Advances in space-scale farming: UAV and satellite monitoring of wheat production in Krasnodar, Russian Federation. Int J Environ Res 1–14. https://doi.org/10.1080/00207233.2021.1997220

Gonzalez L, Yamamoto H (2020) The wall: the Earth in true natural color from real-time geostationary satellite imagery. Remote Sens 12(15):2375. https://doi.org/10.3390/rs12152375

Hasan SS et al (2020) Impact of land use change on ecosystem services: a review. Environ Dev 34:100527. https://doi.org/10.1016/j.envdev.2020.100527

Hawash E et al (2021) Change detection and urban expansion of Port Sudan, Red Sea, using remote sensing and GIS. Environ Monit Assess 193(11):1–22. https://doi.org/10.1007/s10661-021-09486-0

Heidari AA et al (2019) Harris hawks optimization: algorithm and applications. Future Gener Comput Syst 97:849–872. https://doi.org/10.1016/j.future.2019.02.028

Hu X et al (2021) Improving wetland cover classification using artificial neural networks with ensemble techniques. Gisci Remote Sens 58(4):603–623. https://doi.org/10.1080/15481603.2021.1932126

Huang Y et al (2020) Analysis of the future land cover change in Beijing using CA–Markov chain model. Environ Earth Sci 79(2):1–12. https://doi.org/10.1007/s12665-019-8785-z

Ilori CO, Pahlevan N, Knudby A (2019) Analyzing performances of different atmospheric correction techniques for Landsat 8: application for coastal remote sensing. Remote Sens 11(4):469. https://doi.org/10.3390/rs11040469

Iswisi AFA, Karan O, Rahebi J (2021) Diagnosis of multiple sclerosis disease in brain magnetic resonance imaging based on the Harris hawks optimization algorithm. Biomed Res Int 2021:1–12. https://doi.org/10.1155/2021/3248834

Javan FD et al (2021) A review of image fusion techniques for pan-sharpening of high-resolution satellite imagery. ISPRS J Photogramm Remote Sens 171:101–117. https://doi.org/10.1016/j.isprsjprs.2020.11.001

Kafy AA et al (2020) Modelling future land use land cover changes and their impacts on land surface temperatures in Rajshahi, Bangladesh. Remote Sens Appl: Soc. Environ 18:100314. https://doi.org/10.1016/j.rsase.2020.100314

Karimi H, Jafarnezhad J, Kakhani A (2020) Landsat time-series for land use change detection using support vector machine: case study of Javanrud District, Iran. IEEE CSASE, pp 128–131. https://doi.org/10.1109/CSASE48920.2020.9142087

Kaur H, Koundal D, Kadyan V (2021) Image fusion techniques: a survey. Arch Comput Methods Eng 28(7):4425–4447. https://doi.org/10.1007/s11831-021-09540-7

Kayet N et al (2021) Deforestation susceptibility assessment and prediction in hilltop mining-affected forest region. J Environ Manage 289:112504. https://doi.org/10.1016/j.jenvman.2021.112504

Khamchiangta D, Dhakal S (2020) Time series analysis of land use and land cover changes related to urban heat island intensity: case of Bangkok Metropolitan Area in Thailand. J Urban Manag 9(4):383–395. https://doi.org/10.1016/j.jum.2020.09.001

Kodheli O et al (2020) Satellite communications in the new space era: a survey and future challenges. IEEE Commun Surv Tutor 23(1):70–109. https://doi.org/10.1109/COMST.2020.3028247

Kuchkorov T et al (2020) Satellite image formation and preprocessing methods. IEEE ICISCT, pp 1–4. https://doi.org/10.1109/ICISCT50599.2020.9351456

Land use land cover change detection through GIS and unsupervised learning technique. Information and communication technology for sustainable development. Springer, Singapore 933: 239-247.https://doi.org/10.1007/978-981-13-7166-0_23

Kumar P et al (2021) Temporal dynamics change of land use/land cover in Jhansi district of Uttar Pradesh over past 20 years using LANDSAT TM, ETM+ and OLI sensors. Remote Sens Appl: Soc Environ 23:100579. https://doi.org/10.1016/j.rsase.2021.100579

Kumar S, Jain K (2020) A multi-temporal Landsat data analysis for land-use/land-cover change in Haridwar Region using remote sensing techniques. Proc Comput Sci 171:1184–1193. https://doi.org/10.1016/j.procs.2020.04.127

Kutlug SE, Colkesen I (2021) Performance analysis of advanced decision tree-based ensemble learning algorithms for landslide susceptibility mapping. Geocarto Int 36(11):1253–1275. https://doi.org/10.1080/10106049.2019.1641560

Le Dez M, Robin M, Launeau P (2021) Contribution of Sentinel-2 satellite images for habitat mapping of the Natura 2000 site 'Estuaire de la Loire'(France). Remote Sens Appl: Soc Environ 24:100637. https://doi.org/10.1016/j.rsase.2021.100637

Lechner AM, Foody GM, Boyd DS (2020) Applications in remote sensing to forest ecology and management. One Earth 2(5):405–412. https://doi.org/10.1016/j.oneear.2020.05.001

Li J, Endo Y (2020) Fuzzy c-means with improved particle swarm optimization. FUZZ-IEEE, pp 1–8. https://doi.org/10.1109/FUZZ48607.2020.9177673

Li Q et al (2020) Mapping the land cover of Africa at 10 m resolution from multi-source remote sensing data with Google Earth Engine. Remote Sens 12(4):602. https://doi.org/10.3390/rs12040602

Li Y, Ma J, Zhang Y (2021) Image retrieval from remote sensing big data: a survey. Inf Fusion 67:94–115. https://doi.org/10.1016/j.inffus.2020.10.008

Loganathan A, Kaliyaperumal G (2016) An adaptive HVS based video watermarking scheme for multiple watermarks using BAM neural networks and fuzzy inference system. Expert Syst Applicat 63:412–434. https://doi.org/10.1016/j.eswa.2016.05.019

Loganathan A et al (2022) Crop classification for agricultural applications in hyperspectral remote sensing images. Appl Sci 12(3):1670. https://doi.org/10.3390/app12031670

Ma Z et al (2020) Uncertainty analysis for topographic correction of hyperspectral remote sensing images. Remote Sens 12(4):705. https://doi.org/10.3390/rs12040705

Manoharan P et al (2018) Three-dimensional discrete cosine transform-based feature extraction for hyperspectral image classification. J Appl Remote Sens 12(4):046010. https://doi.org/10.1117/1.JRS.12.046010

Manoharan P, Loganathan A, Kaliyaperumal G (2019) An intelligent lung cancer diagnosis system using cuckoo search optimization and support vector machine classifier. J Ambient Intel Human Comput 10:267–293. https://doi.org/10.1007/s12652-017-0655-5

Matci DK, Avdan U (2020) Optimization-based automated unsupervised classification method: a novel approach. Expert Syst Appl 160:113735. https://doi.org/10.1016/j.eswa.2020.113735

Mishra S, Jabin S (2020) Land use land cover change detection using LANDSAT images: a case study. IEEE 5th ICCCA, pp 730–735. https://doi.org/10.1109/ICCCA49541.2020.9250801

Mohan A et al (2021) Review on remote sensing methods for landslide detection using machine and deep learning. Trans Emerg Telecommun Technol 32(7):e3998. https://doi.org/10.1002/ett.3998

Mohanrajan SN, Loganathan A (2021) Modelling Spatial Drivers for LU/LC Change Prediction Using Hybrid Machine Learning Methods in Javadi Hills, Tamil Nadu, India. J Indian Soc Remote Sens 49:913–934. https://doi.org/10.1007/s12524-020-01258-6

Mohanrajan SN, Loganathan A (2022) Novel Vision Transformer-Based Bi-LSTM model for LU/LC prediction—Javadi Hills. India Appl Sci 12(13):6387. https://doi.org/10.3390/app12136387

MohanRajan SN, Loganathan A, Manoharan P (2020) Survey on land use/land cover (LU/LC) change analysis in remote sensing and GIS environment: techniques and challenges. Environ Sci Pollut Res 27:29900–29926. https://doi.org/10.1007/s11356-020-09091-7

Mukherjee F, Deepika S (2020) Assessing land use–land cover change and its impact on land surface temperature using LANDSAT data: a comparison of two urban areas in India. Earth Syst Environ 4(2):385–407. https://doi.org/10.1007/s41748-020-00155-9

Munthali MG et al (2020) Modelling land use and land cover dynamics of Dedza district of Malawi using hybrid Cellular Automata and Markov model. Remote Sens Appl: Soc Environ 17:100276. https://doi.org/10.1016/j.rsase.2019.100276

Mzid N et al (2021) An analysis of bare soil occurrence in arable croplands for remote sensing topsoil applications. Remote Sens 13(3):474. https://doi.org/10.3390/rs13030474

Nadudvari A et al (2020) The estimation of lava flow temperatures using landsat night-time images: case studies from eruptions of Mt. Etna and Stromboli (Sicily, Italy), Kīlauea (Hawaii Island), and Eyjafjallajökull and Holuhraun (Iceland). Remote Sens 12(16):2537. https://doi.org/10.3390/rs12162537

Navin MS and Agilandeeswari L (2020a) Comprehensive review on land use/land cover change classification in remote sensing. J Spectr Imaging 9. https://doi.org/10.1255/jsi.2020.a8

Navin MS, Agilandeeswari L (2020b) Multispectral and hyperspectral images based land use/land cover change prediction analysis: an extensive review. Multimed Tools Appl 79(39):29751–29774. https://doi.org/10.1007/s11042-020-09531-z

Nazeer M et al (2021) Evaluation of atmospheric correction methods for low to high resolutions satellite remote sensing data. Atmos Res 249:105308. https://doi.org/10.1016/j.atmosres.2020.105308

Pham VH, Nguyen TNT, Laffly D (2020) Remote sensing case studies. TORUS 2–Toward an Open Resource Using Services: cloud computing for environmental data. 191–224. https://doi.org/10.1002/9781119720553.ch7

Phan DC et al (2021) Ensemble learning updating classifier for accurate land cover assessment in tropical cloudy areas. Geocarto Int 1–18.https://doi.org/10.1080/10106049.2021.1878292

Piao Y et al (2021) Analysis of land use and land cover change using time-series data and random forest in North Korea. Remote Sens 13(17):3501. https://doi.org/10.3390/rs13173501

Radočaj D et al (2020) Global open data remote sensing satellite missions for land monitoring and conservation: a review. Land 9(11):402. https://doi.org/10.3390/land9110402

Rahman A et al (2020) Performance of different machine learning algorithms on satellite image classification in rural and urban setup. Remote Sens Appl: Soc Environ 20:100410. https://doi.org/10.1016/j.rsase.2020.100410

Reddy GP, Kumar KC (2022) Machine learning algorithms for optical remote sensing data classification and analysis. Data Sci Agric Nat Resour Manag 195–220. https://doi.org/10.1007/978-981-16-5847-1_10

Ridding LE et al (2020) Modelling historical landscape changes. Landsc Ecol 35(12):2695–2712. https://doi.org/10.1007/s10980-020-01059-9

Riva F, Nielsen SE (2021) A functional perspective on the analysis of land use and land cover data in ecology. Ambio 50(5):1089–1100. https://doi.org/10.1007/s13280-020-01434-5

Rodríguez-Esparza E et al (2020) An efficient Harris hawks-inspired image segmentation method. Expert Syst Appl 155:113428. https://doi.org/10.1016/j.eswa.2020.113428

Said M et al (2021) Predicting land use/cover changes and its association to agricultural production on the slopes of Mount Kilimanjaro, Tanzania. Annals of GIS 27(2):189–209. https://doi.org/10.1080/19475683.2020.1871406

Sankarrao L, Ghose DK, Rathinsamy M (2021) Predicting land-use change: Intercomparison of different hybrid machine learning models. Environ Model Softw 145:105207. https://doi.org/10.1016/j.envsoft.2021.105207

Sawant S, Manoharan P (2020) Hyperspectral band selection based on metaheuristic optimization approach. Infrared Phys Technol 107:103295. https://doi.org/10.1016/j.infrared.2020.103295

Sekertekin A, Bonafoni S (2020) Land surface temperature retrieval from Landsat 5, 7, and 8 over rural areas: assessment of different retrieval algorithms and emissivity models and toolbox implementation. Remote Sens 12(2):294. https://doi.org/10.3390/rs12020294

Seydi ST et al (2021) Wildfire damage assessment over Australia using sentinel-2 imagery and MODIS land cover product within the google earth engine cloud platform. Remote Sens 13:220. https://doi.org/10.3390/rs13020220

Shafizadeh-Moghadam H et al (2020) Modeling the spatial variation of urban land surface temperature in relation to environmental and anthropogenic factors: a case study of Tehran, Iran. Gisci Remote Sens 57(4):483–496. https://doi.org/10.1080/15481603.2020.1736857

Sheoran S, Mittal N, Gelbukh A (2020) Analysis on application of swarm-based techniques in processing remote sensed data. Earth Sci Inform 13(1):97–113. https://doi.org/10.1007/s12145-019-00417-9

Sheoran S, Mittal N, Gelbukh A (2021) Improved change detection in remote sensed images by artificial intelligence techniques. J Indian Soc Remote Sens 49(9):2079–20921–14. https://doi.org/10.1007/s12524-021-01374-x

Shooshtari SJ et al (2020) Land use and cover change assessment and dynamic spatial modeling in the Ghara-su Basin, Northeastern Iran. J Indian Soc Remote Sens 48(1):81–95. https://doi.org/10.1007/s12524-019-01054-x

Sihwail R et al (2020) Improved Harris hawks optimization using elite opposition-based learning and novel search mechanism for feature selection. IEEE Access 121127–121145. https://doi.org/10.1109/ACCESS.2020.3006473

Singh M et al (2020) Detection of changes in Landsat images using hybrid PSO-FCM. Proc Comput Sci 167:423–430. https://doi.org/10.1016/j.procs.2020.03.251

Singh RK et al (2021) A machine learning-based classification of LANDSAT images to map land use and land cover of India. Remote Sens Appl: Soc Environ 24:100624. https://doi.org/10.1016/j.rsase.2021.100624

Srivastava PK et al (2019) GIS and remote sensing aided information for soil moisture estimation: a comparative study of interpolation techniques. Resources 8(2):70. https://doi.org/10.3390/resources8020070

Talukdar S et al (2020) Land-use land-cover classification by machine learning classifiers for satellite observations—a review. Remote Sens 12(7):1135. https://doi.org/10.3390/rs12071135

Tamiminia H et al (2020) Google Earth Engine for geo-big data applications: a meta-analysis and systematic review. ISPRS J Photogramm Remote Sens 164:152–170. https://doi.org/10.1016/j.isprsjprs.2020.04.001

Thyagharajan KK, Vignesh T (2019) Soft computing techniques for land use and land cover monitoring with multispectral remote sensing images: a review. Arch Comput Methods Eng 26(2):275–301. https://doi.org/10.1007/s11831-017-9239-y

Useya J, Chen S, Murefu M (2019) Cropland mapping and change detection: toward Zimbabwean cropland inventory. IEEE Access 7:53603–53620. https://doi.org/10.1109/ACCESS.2019.2912807

Valdez F, Castillo O, Melin P (2021) Bio-inspired algorithms and its applications for optimization in fuzzy clustering. Algorithms 14(4):122. https://doi.org/10.3390/a14040122

Van Ha P et al (2018) Assessment of georeferencing methods on MODIS Terra/Aqua and VIIRS NPP satellite images in Vietnam. IEEE KSE 282–287. https://doi.org/10.1109/KSE.2018.8573402

Vivekananda GN, Swathi R, Sujith AVLN (2021) Multi-temporal image analysis for LULC classification and change detection. Eur J Remote Sens 54(sup2):189–199. https://doi.org/10.1080/22797254.2020.1771215

Wang H et al (2020) Impacts of topography on the land cover classification in the Qilian Mountains, Northwest China. Can J Remote Sens 46(3):344–359. https://doi.org/10.1080/07038992.2020.1801401

Weiss M, Jacob F, Duveiller G (2020) Remote sensing for agricultural applications: a meta-review. Remote Sens Environ 236:111402. https://doi.org/10.1016/j.rse.2019.111402

Wellmann T et al (2020) Remote sensing in urban planning: contributions towards ecologically sound policies? Landsc Urban Plan 204:103921. https://doi.org/10.1016/j.landurbplan.2020.103921

Yashin JFM et al (2020) Comparative analysis of classification algorithms for landuse/landcover change over a part of the east coast region of Tamil Nadu and its environs. IEEE InGARSS 66–69. https://doi.org/10.1109/InGARSS48198.2020.9358945

Yuan Q et al (2020) Deep learning in environmental remote sensing: achievements and challenges. Remote Sens Environ 241:111716. https://doi.org/10.1016/j.rse.2020.111716

Zeferino LB et al (2020) Does environmental data increase the accuracy of land use and land cover classification? Int J Appl Earth Obs Geoinf 91:102128. https://doi.org/10.1016/j.jag.2020.102128