Earth Science Informatics (2021) 14:2231–2247
https://doi.org/10.1007/s12145-021-00685-4

RESEARCH ARTICLE

A comparative analysis of different pixel and object-based classification algorithms using multi-source high spatial resolution satellite data for LULC mapping

Akanksha Balha[1] · Javed Mallick[2] · Suneel Pandey[3] · Sandeep Gupta[4] · Chander Kumar Singh[1]

Received: 3 June 2021 / Accepted: 31 July 2021 / Published online: 3 September 2021

Abstract
The preparation of accurate LULC is of great importance as it is used in various studies ranging from change detection to geospatial modelling. Literature offers studies comparing different classification algorithms/approaches to prepare LULC maps. However, still there is a lack of studies that can provide a comprehensive analysis on widely used classification algorithms. Hence, in the present study, nine different pixel- and object-based classification algorithms have been used to compare their relative effectiveness in generating remotely sensed LULC maps. The algorithms include maximum likelihood, neural network, support vector machine (linear, polynomial, RBF (radial basis function), sigmoid kernels), random forest (RF) and Naive Bayes for pixel-based classification and maximum likelihood algorithm for object-based classification (OBC) approach. Additionally, the study has analysed the impact of different types of satellite datasets (i.e., high resolution image and resolution merged images of same resolution) on relative effectiveness of the algorithms in classifying the satellite imageries accurately. High resolution (5 m) satellite image LISS 4 MX70, resolution merged satellite images (5 m) LISS 3 merged with LISS 4 mono and LISS 3 merged with IRS-1D are employed for classification. 27 LULC maps (9 classification algorithms * 3 images) are evaluated for comparing classification algorithms. The accuracy assessment of the images is carried out using confusion matrix and Mc Nemar's test. It has been observed that (1) the performance of all classification algorithms differs from each other and (2) resolution merged data impacts classification accuracy differently when compared to other satellite image of same spatial resolution. RF and OBC are identified as potential classifiers with majority of datasets. The results suggest that due to heterogeneity in urban land-use, it is difficult to achieve higher overall accuracy in classifying a large urban area using 5 m resolution satellite dataset. Moreover, visual examination of LULC should be considered for choosing better classification approach as pixel-based approach produces salt-pepper effect in LULC, whereas OBC produces visually smoothened LULC and identifies even smaller objects in urban landscape. The comparative evaluation of different image types reveal that RF performs better with all images, however, the performance of OBC was found to be improved with original high-resolution data.

Keywords Radial basis function · Sigmoid · Kernel · Random forests · Naive Bayes · Mc Nemar's test

Communicated by H. Babaie.

✉ Chander Kumar Singh
chanderkumarsingh@gmail.com; chander.singh@terisas.ac.in

1 Department of Energy and Environment, TERI School of Advanced Studies, 10 Institutional Area, Vasant Kunj, New Delhi 110 070, India

2 Department of Civil Engineering, King Khalid University, Abha, Saudi Arabia

3 Environment and Waste Management Division, The Energy and Resource Institute (TERI), New Delhi, India

4 Indira Gandhi National Open University, Regional Centre, Jammu, Jammu and Kashmir, India

Introduction

Over the past few decades, the advancement of remote sensing and easier availability of satellite images have made land-use analysis using image classification a vital research topic (Robles Granda 2011; Tehrany et al. 2014). Automated image classification is one of the easiest and preferable techniques to prepare land use land cover (LULC) of an area (Rozenstein and Karnieli 2011). The studies carried out in the past have identified the best performing classification algorithm by comparing different classification algorithms. However, none of them provides a comprehensive comparative analysis of all the popular classification algorithms (Srivastava et al. 2012).

Literature offers various studies which address comparison between pixel-based and/or object-based classification approaches. The basic difference between the two approaches is that of the underlying spatial unit—pixel or object (Duro et al. 2012; Tehrany et al. 2014). Pixel-based classification uses the spectral information stored as digital numbers (DN) in each pixel, where each pixel represents different feature on the earth's surface. The object-based classification approach considers spatial features e.g. shape, size, tone/color, texture, pattern, association etc., and divides the image into homogeneous objects (Gao et al. 2011; Tehrany et al. 2014).

Srivastava et al. (2012) have evaluated three pixel-based classification algorithms—artificial neural network (ANN), support vector machine (SVM) and maximum likelihood (ML) using low resolution Landsat TM/ETM + images and found ANN as a better classifier than SVM and ML. Rozenstein and Karnieli (2011) have used low resolution Landsat TM image to compare different pixel-based classification algorithms: supervised (ML), unsupervised (ISODATA) and hybrid method (developed by combining spectral signatures from supervised and unsupervised classification). Their results revealed that hybrid method (73.5%) performed better than unsupervised (70.67%) and supervised (60.83%) algorithms. Similar order of performance by the three algorithms was observed after performing post-classification processing. Hybrid classification method was found to be statistically significant than supervised classification, but not in comparison to unsupervised method. Cleve et al. (2008) have compared pixel-based (ISODATA) and object-based classification (nearest neighbour) in wildlife-urban interface area using high resolution aerial photographs and found object-based approach (80.08%) to be better than pixel-based approach (62.11%). Using medium resolution (10 m) SPOT 5 data, Tehrany et al. (2014) has compared pixel-based (decision trees (DT)) and object-based (SVM and nearest neighbour (k-NN) classification approaches. The k-NN algorithm in object-based classification performed significantly better than SVM algorithm in object-based classification and pixel-based (DTs) classification. Duro et al. (2012) have used SPOT (10 m) data to compare pixel-based and object-based classification approaches using Random Forests (RF), DT, support vector machine (SVM) algorithms. Their statistical tests revealed that pixel-based and object-based classification are not statistically significant when same algorithm is applied, however, object-based (DT) classification algorithm was found to be statistically significant than object-based (RF) and SVM algorithms. Duro et al. (2012) stated that both pixel-based and object-based approaches produced similar overall accuracies; however, pixel-based classification was less time consuming to process. LULC map produced using object-based classification approach was found to be visually smoothened. Using pixel-based (ML) and object-based (k-NN) classification approaches to extract urban land cover from VHR quickbird data, it has been observed (Myint et al. 2011) that object-based classifier (90.40%) performs significantly better than pixel-based classifier (67.60%). Jozdani et al. (2019) evaluated comparative performance of different deep learning algorithms, common ensemble algorithms and SVM in classifying urban areas using object-based approach. The findings indicated multilayer perceptron (MLP) as the best classifier. Moreover, other classifiers such as SVM were observed as capable enough to map LULC in complex landscapes. While using Landsat 8 data, Qu et al. (2021) demonstrated that integration of auxiliary datasets improves pixel-based and object-based classification results. The performance of object-based approach was found to be higher than pixel-based approach. Object-based approach was observed to achieve higher accuracy while using only spectral datasets. Gudiyangada Nachappa et al. (2020) compared conventional pixel-based models (data-driven frequency ratio and expert based analytical hierarchical process) with geon-based object-based classification approaches. The results indicated that object-based approach provided higher accuracy than both pixel-based models and produced meaningful spatial units. Tassi et al. (2021) have compared pixel-based and object-based classification approaches and evaluated the impact of integrating textural details in classification process. The results revealed that the accuracy of pixel-based approach has not improved with the integration of textural details. The object-based approach was found to perform better than the pixel-based approach while employing 15 m panchromatic band of Landsat 8 data. Addition of panchromatic band did not improve the pixel-based classification results; however, it generated a detailed LULC with object-based classification approach.

Many LULC classification studies (Gao et al. 2009; Hu and Weng 2011; Duro et al. 2012; Tehrany et al. 2014) reveal that object-based classification approach is believed to provide more accurate results than pixel-based approach. Although more information is acquired from higher resolution images than coarser ones, high resolution images provide challenges for pixel-based classification (Cleve et al. 2008). Unlike natural landscapes, many features in a small space in an urban area can be captured precisely in higher spatial resolution image, but such higher level of detail may congest the details of urban features in spectral context (Myint et al. 2006, 2011). It occurs because pixel-based classification considers only spectral information and neglects spatial information, an attribute which is considered significant in object-based classification (Benz et al. 2004; Walter 2004; Myint et al. 2011; Duro et al. 2012). The similar spectra from different features in urban areas (e.g. buildings, rooftops, roads, sidewalks and other bright surface objects etc.) leads to "mixed pixel problem" or "salt and pepper effect" (Kelly et al. 2004; Cleve et al. 2008; Myint et al. 2011; Ouyang et al. 2011) in pixel-based classification. This causes higher intra-class spectral variability which lowers the statistical separability between classes, thereby leading to misclassification and low classification accuracy (Su et al. 2004) and thus, object-based classification approach is used to overcome these challenges (Cleve et al. 2008; Ouyang et al. 2011). Many researchers (Herold et al. 2003; Durieux et al. 2008; Hu and Weng 2011; Myint et al. 2011) have employed object-based classification for urban areas, considering urban areas to be too complex to be classified accurately by pixel-based methods. In comparison to pixel-based classification, besides higher accuracy, object-based classification approach has another advantage to classify objects with proper shape, which is rather hard to be achieved in pixel-based classification (Baatz et al. 2004; Ouyang et al. 2011). Although, object-based approach is believed to be more time consuming and labour intensive (Duro et al. 2012).

With the above background, the present study provides a comprehensive comparative view of most used classification algorithms while taking account of satellite datasets from different platforms. Nine different classification algorithms including maximum likelihood (ML), neural network (NN), support vector machine (SVM)—linear, polynomial, radial basis function (RBF) and sigmoid kernels, random forests (RF) and naive bayes (NB) in pixel-based and (maximum likelihood (ML) in object-based classification (OBC) are performed on different satellite datasets.

Study area

The study area is National Capital Territory (NCT) of Delhi (Fig. 1). While covering an area of 1,484 km^2, its latitudinal and longitudinal extent is 28.4084° N to 28.8845° N and 76.8328° E to 77.3377° E respectively. It has a completely urbanized landscape with some agricultural area at the outskirts and river Yamuna flowing through it.

Data and methodology

In this study, satellite images of different spatial resolution are used. Detailed specifications of images are mentioned in Table 1, which indicates the variation in datasets (in terms of data of different time periods and sensors) that has been considered in the present work. LISS and IRS-1D images were obtained from National Remote Sensing Centre (NRSC), Indian Space Research Organisation (ISRO) (https://www.nrsc.gov.in/). The years 2005, 2010 and 2016 have been chosen carefully considering two important things, i.e. (1) the availability of cloud-free satellite images and (2) optimal temporal variability so as to have changing land-use to be prominent. To avoid seasonal effects, all the images were acquired for the same season.

Image pre-processing

Before performing classification, all the satellite images for years 2005, 2010 and 2016 were pre-processed in ERDAS® IMAGINE 2016 (Hexagon Geospatial 2016). The

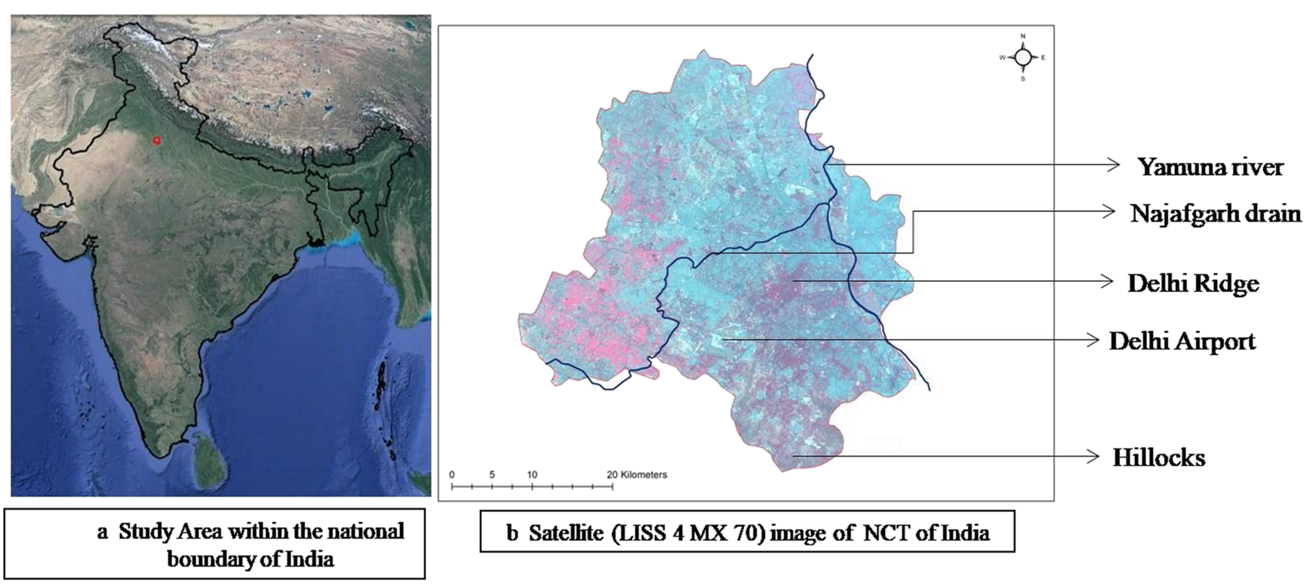


Fig. 1 Study area—NCT of Delhi

Table 1 Specifications of satellite images used for classification to prepare LULC of Delhi (2005–2016) (* in accordance with metadata of the image)

S. no.	Satellite/sensor	Scene no.	Date of image acquisition	Spectral wavelength/bands (μm)	Output spatial resolution* (m)
1	LISS 4 MX 70	96_51_A 96_51_C	Dec 04th, 2016	Band 2 (0.52–0.59) Band 3 (0.62–0.68) Band 4 (0.77–0.86)	5.0
2	LISS 4 MX 70	96_51_C	Oct 20th , 2011	Band 2 (0.52–0.59) Band 3 (0.62–0.68) Band 4 (0.77–0.86)	5.0
3	LISS 4 Mono	96_51_A	Nov 30th, 2010	Band 3 (0.62–0.68)	5.0
4	LISS 3	96_51	Nov 30th, 2010	Band 2 (0.52–0.59) Band 3 (0.62–0.68) Band 4 (0.77–0.86) Band 5 (1.55–1.70)	24.0
5	LISS 3	96_51	Dec 02nd, 2005	Band 2 (0.52–0.59) Band 3 (0.62–0.68) Band 4 (0.77–0.86) Band 5 (1.55–1.70)	24.0
6	IRS-1D (Panchromatic)	96_51_A 96_51_C	Oct 27th, 2005	0.5–0.75	5.0

methodology used for pre-processing the images is represented in Fig. 2. The images of 2005 and 2010 were resolution merged to generate higher spatial resolution images for the respective years. The highest spatial resolution among all the satellite images was 5 m, therefore the spatial resolution of resolution merged images was resampled to 5 m. For resolution merging, various techniques (including wavelet, high-pass filter (HPF), modified IHS (intensity, hue, saturation), principal component (PC) based resolution merging, projective resolution merging and hyperspherical color space (HCS) were performed, however, the images resolution merged using HCS technique for both 2005 and 2010 were chosen to be used further as they were more accurate among all.

Image classification

Different classification algorithms—pixel-based and object-based, were performed on the processed subset images. Nine LULC classes, namely water bodies, built-up, dense vegetation, sparse vegetation, cropland, fallow land, open land, scrubland/forest and sediment, were identified.

Pixel-based classification

For the training dataset, the spectral signatures from more than 3000 pixels from the satellite image of each year were selected to perform pixel-based classification. Thereafter, the training sets from spectrally similar pixels were merged. The same training dataset was used to perform all the algorithms used in the study in pixel-based classification.

All the pixel-based classification algorithms were performed in an open-source statistical computing software R version 3.3.2 (R Development Core Team 2016). Various add-on packages used in R to build and perform different classification algorithms include "rasclass" package for ML (Wiesmann and Quinn 2011), "nnet" package for NN (Venables and Ripley 2002), "kernlab" package (Karatzoglou et al. 2004) for SVM, "randomForest" package (Liaw and Wiener 2002) for RF and "naiveBayes" package for NB (Majka 2018).

Object-based classification (OBC)

The OBC was carried out in ArcGIS 10.5 (ESRI 2016). Firstly, image segmentation based on mean shift approach was performed to create segments or features of interest. There is no common scale (Myint et al. 2011) or fixed criterion to estimate the best parameters (Ouyang et al. 2011; Duro et al. 2012) for segmentation. The researchers (Chen et al. 2006; Ouyang et al. 2011) identify the scale that delineates the objects in the best visually corresponding manner to the real-world objects and consider it the appropriate scale level to be adopted for the classification. Initially, segmentation was tried with different values of parameters (Fig. 3). In parameters, the spectral detail value was kept constant and different spatial detail values were experimented to decide the parameters values for segmentation. Thereafter, based on visual attributes of the segmented image, images with criteria spectral detail = 20, spatial detail = 20 and minimum segment size = 5 pixels were found to be more appropriate and precise. The features of segmented image served as the underlying units

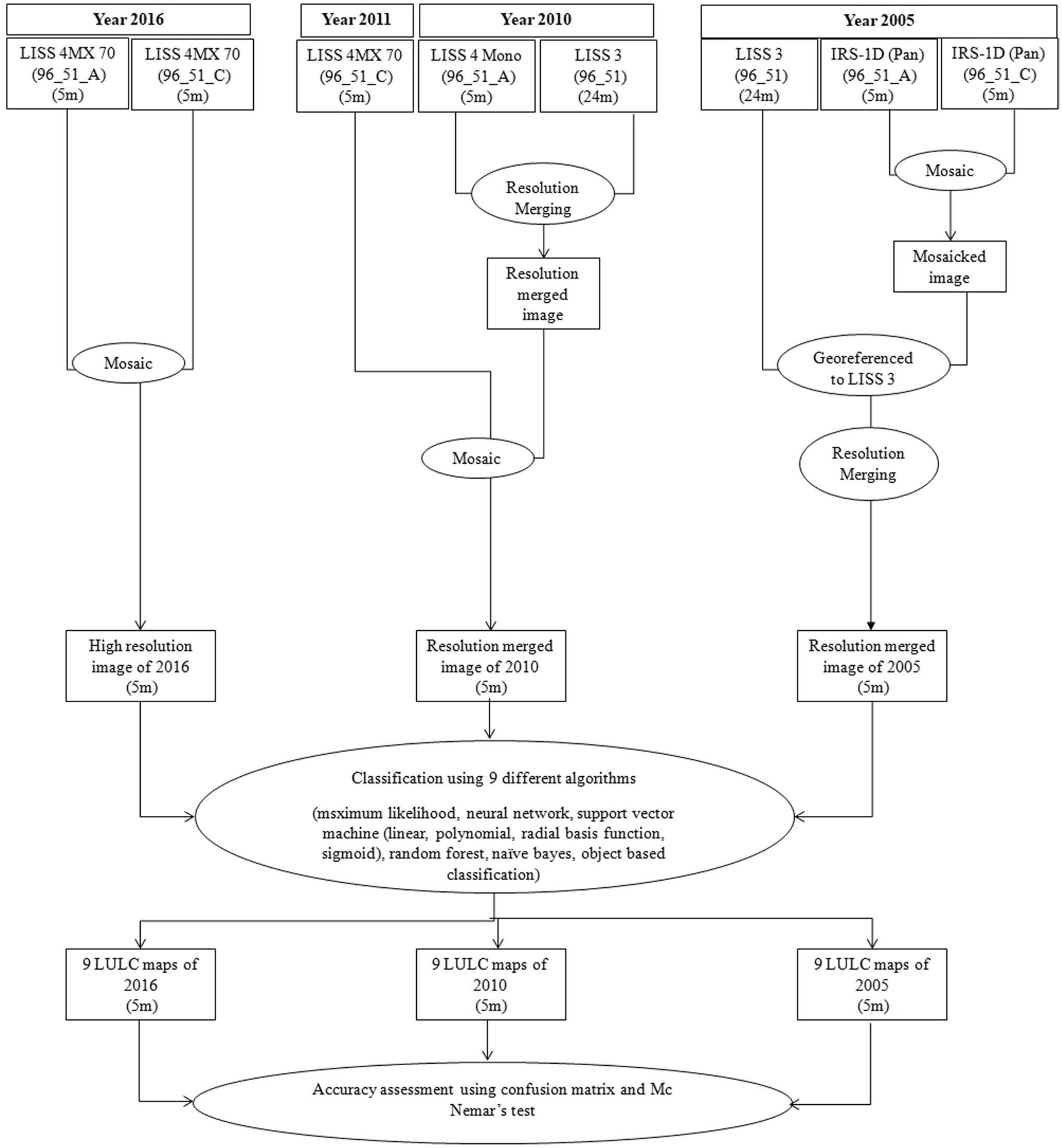


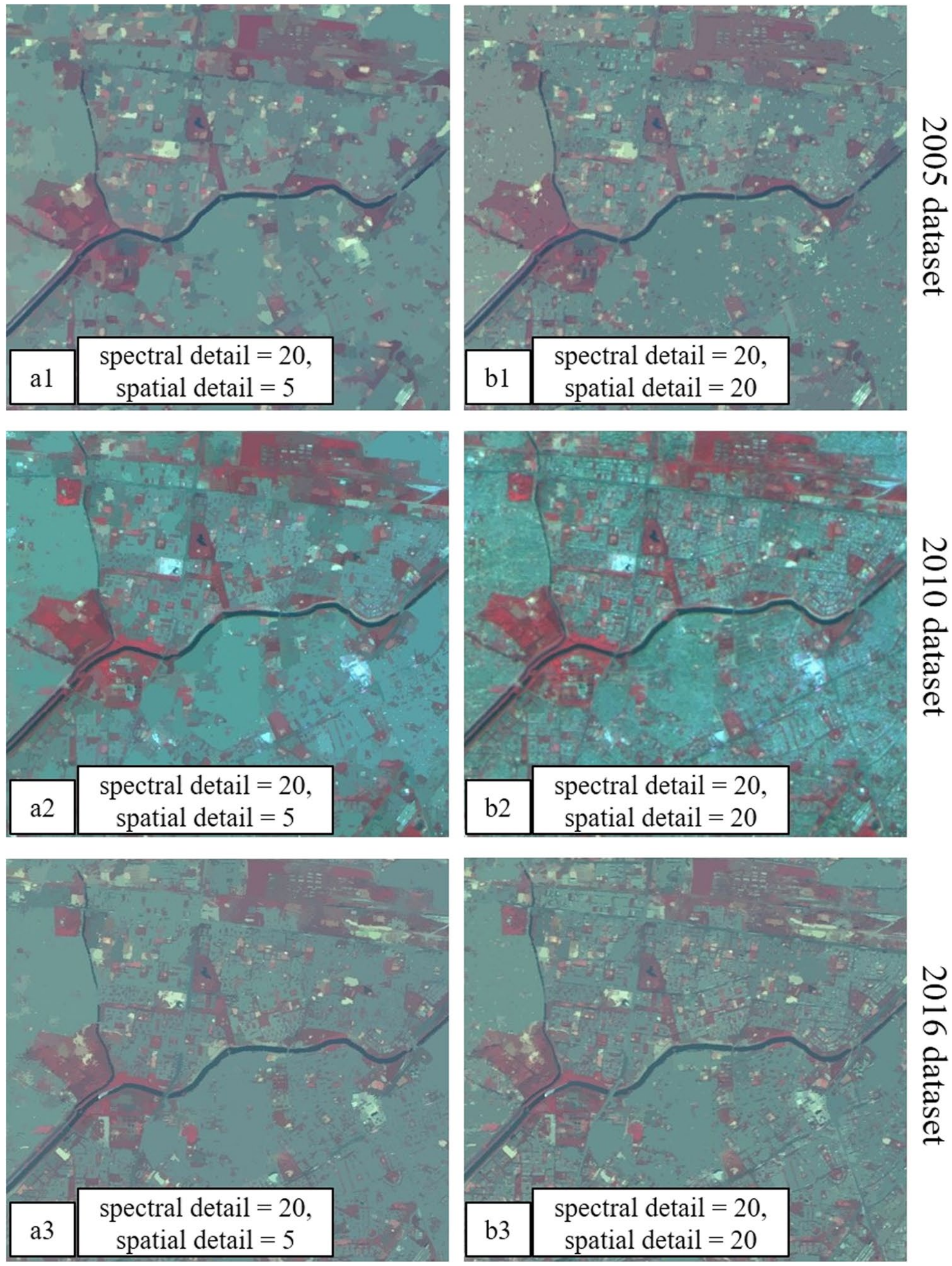
Fig. 2 Brief methodology

for OBC. On an average, each image was segmented into more than 200,000 image objects. Once image segmentation was done precisely, training samples were collected from the segmented raster. Using the training samples and ML classifier, a classifier file was generated. Subsequently, based on generated classifier file, OBC was executed.

Accuracy assessment

Accuracy assessment of thematic (LULC) maps is crucial since the reliability of remotely sensed LULC maps depends on their accuracy. In the present study, for accuracy assessment of the LULC maps, 542 points for 9 LULC classes in

Fig. 3 Image segmentation parameters in OBC approach used in the study



each year's dataset were selected based on equalized random sampling. The accuracy was determined using (1) Confusion (or Error) matrix; and (2) Mc Nemar's test (Kavzoglu 2017). Confusion matrix provides three accuracy measures, i.e., overall accuracy, producer accuracy, and user accuracy. The confusion matrix is based on the comparison between reference image and classified image (output). Columns of matrix refer to LULC classes of reference image whereas the rows of matrix LULC classes of classified image. The no. of pixels comprising a specific LULC class is show by the cells of a matrix, whereas the number of pixels accurately classified is show by the diagonal cells. The overall accuracy is determined by dividing accurately classified pixels by total number of pixels. The overall accuracy decides the classification accuracy of the entire image whereas producer's accuracy and user's accuracy decide the accuracy of individual LULC classes. The producer's accuracy is calculated as accurately classified pixels divided by the sum of total pixels in the reference image. The user's accuracy is calculated as accurately classified pixels divided by sum of total pixels in the classified image.

Mc Nemar's test is a statistical test used to evaluate statistical significance in the differences in the performance of different classifiers (Dietterich 1998). The test is applied to 2×2 contigency table where cells indicate number of samples incorrectly and correctly classified by two methods, the number of samples only correctly classified by one method (Kavzoglu 2017). The test statistic for Mc Nemar is give as Eq. (1)

$$\chi^2 = \frac{\left(\left|a_{ij} - a_{ji}\right| - 1\right)^2}{a_{ij} + a_{ji}} \tag{1}$$

where a_{ij} refers to pixels incorrectly classified by method *i* but classified correctly by method *j*, a_{ji} refers to pixels incorrectly classified by method *j* but not by method *i*. χ^2 follow chi-square distribution with degree of freedom 1. If estimated test values > χ value in the tale, two methods are said to perform differently, which means the difference in accuracy obtained by methods *i* and *j* are statistically significant.

Many researchers (Cohen 1960; Foody 2004; Rozenstein and Karnieli 2011; Duro et al. 2012) have pointed out that the cases wherein the same validation samples are used to assess different algorithms; the presumption that every algorithm is evaluated independently is infringed. In such instances, statistical comparison using kappa remains unjustifiable (Foody 2004; Duro et al. 2012; Rozenstein and Karnieli 2011). Hence, in such circumstances, Agresti (2002) and Zar (2009) recommends the use of Mc Nemar's test for comparing classification algorithms. It is a non-parametric statistical measure for assessing the accuracy of thematic maps (Yan et al. 2006; Dingle and King 2011; Rozenstein and Karnieli 2011; Whiteside et al. 2011; Duro et al. 2012).

Mc Nemar's test gives *p* value and chi-square value which determines the statistical significance of the difference between two algorithms (Foody 2004; De Leeuw et al. 2006; Rozenstein and Karnieli 2011). It is suggested to be performed as not every difference between two algorithms shall be significant. Assessing 27 LULC maps using Mc Nemar's test revealed the statistically significant difference between any of the pixel-based algorithms and OBC approach or among different pixel-based algorithms.

Temporal analysis of LULC change

LULC maps of 2005, 2010 and 2016 were compared to analyse the change in LULC over the specified period. Post classification comparison technique was adopted as it is widely used and considered to provide more accurate results than other techniques including PCA, image differencing etc. (Dingle Robertson and King 2011). LULC class-wise area statistics was tabulated to analyse the nature and trend of land-use change shown by different algorithms temporally.

Theory

A brief description of the algorithms used in pixel and object-based classification is mentioned here in Table 2.

Results

The LULC maps classified using all the studied algorithms are shown in Figs. 4, 5 and 6. The accuracy assessment of all the maps was performed using confusion matrix. The accuracy measures (overall accuracy (OA), producer's accuracy (PA), user's accuracy (UA) and kappa statistic) for all the years are given in Table 3 and the results of Mc Nemar's test for years 2005, 2010 and 2016 are given in Tables 4, 5 and 6 respectively.

On an average, the overall accuracy of all the LULC maps is approximately 50%. This is far below the established standard that states that the accuracy of the LULC maps should be at least 85% for the maps to be useful for planning and management of the areas (Anderson et al. 1976). However, in the present research work, the prepared LULC maps are not to be used for planning and management purposes but to compare the relative effectiveness of the different algorithms in classifying the remotely sensed satellite images accurately. Therefore, the output (LULC) of the algorithms as is produced have been taken into account for evaluation of algorithms and decided not to manipulate it with any post-classification processing i.e. filtering or recoding to increase the overall accuracy (Rozenstein and Karnieli 2011).

Accuracy assessment of LULC maps using confusion matrix

Overall accuracy (OA)

From Table 3, it is evident that among all studied algorithms, RF with OA (54.98% in 2005; 52.58% in 2010; and 56.83% in 2016) has performed as the best classification algorithm and Naive Bayes (39.11% in 2005; 41.14% in 2010; and 35.42% in 2016) the least. The performance of all the four kernels of SVM has been better than that of ML and NN in all the three years. However, no trend in the relative performances of the kernels across the three datasets is observed.

In comparison to all the pixel-based algorithms, object-based classification approach (44.46% OA in 2005 and 43.91% OA in 2010) has performed quite low; however, for year 2016, the performance of object-based classification (54.98% OA) has been very close to the best performed (pixel-based) classification algorithm i.e. RF (56.83% OA).

Table 2 A brief description of the algorithms used in pixel-based and object-based classification

	Algorithms	Attributes	Description	References
Pixel-based classification	ML	Principle and formula involved	Based on Bayes' theorem of decision making $D = \ln\left(a_c\right) - [0.5\ln\left(\lvert cov_c\rvert\right)] - [0.5\left(X - M_c\right)T\left(cov_c - 1\right)\left(X - M_c\right)]$ (1) where, D = likelihood, c = particular class, X = the measurement vector of the candidate pixel, Mc = the mean vector of the sample of class c, ac = percent probability that any candidate pixel is a member of class c (defaults to 1.0, or is entered from a priori knowledge), Cov_c = the covariance matrix of the pixels in the sample of class c, $\lvert Cov_c\rvert$ = determinant of Cov_c (matrix algebra), Cov_c^{-1} = inverse of Cov_c (matrix algebra), ln = natural logarithm function, T = transposition function (matrix algebra)	Otukei and Blaschke (2010), Srivastava et al. (2012)
		Merits	Performs better than other known parametric classification algorithms for normally distributed data	
		Demerits	For non-normally distributed data, it may not produce desirable results	
	NN	Principle, and formula involved	The logistic activation function used, is expressed as Eq. (2) $o_j = 1/\left(1 + e^{-\lambda net_j}\right)$ (2) where, o_j is the output of external input j, λ is a gain factor $net_j = \sum w_{ji} o_i$ (3) where, w_{ji} is the weight of interconnection channel to unit j from unit i and o_i is the output of external unit i	Tu (1996), Shao and Lunetta (2012), Srivastava et al. (2012), Zhang (2014)
		Merits	Doesn't consider assumptions about the statistical nature of data distribution	
		Demerits	black box' nature, greater computational burden, proneness to overfitting and empirical nature	
	SVM	Principle, and formula involved	Based on the statistical learning theory, it classifies the data by identifying a hyperplane (Eq. 4) which provides the function that forms relationship between input and output data while training the data in statistical learning $f(x) = \sum^{l} K\left(x, x_i\right)$ (4) where the x_i, i = 1,…, 1 are the input data, K is the kernel function, and c_i is a group of parameters which are estimated from the input data	Foody and Mathur (2004), Nemmour and Chibani (2006), Chi et al. (2008), Mountrakis et al. (2011), Tien Bui et al (2012), Ren (2012), Tehrany et al. (2014
		Merits	considered more accurate than widely used ML and NN; identify patterns and objects; Self-adaptability, robust to limited requirement of training data	
		Demerits	Selecting the appropriate kernel—a challenge	
	RF	Principle, and formula involved	An ensemble (forests) of classification trees; classifies via a voting system The time required for training Random Forests is expressed by Eq. (5) $cT\sqrt{MN}\log N$ (5) where, c is a constant dependant of the data complexity, T is the number of tree, M is the number of variables and N is the number of instances	Brieman (2001), Rodriguez-Galiano et al. (2012a), Long et al. (2013), Hayes et al. (2014), Puissant et al. (2014)
		Merits	Robus to large datasets, high dimensional and redundant data; noise and over fitting and does not require a particular statistical distributional data	
		Demerits	'Black box' nature	
	NB	Principle, and formula involved	Depends on Bayesian theorem and presumes all the variables participate in classification and are mutually correlated	Domingos and Pazzani (1997), Kotsiantis et al. (2007), Tien Bui et al. (2012), Jadhav and Channe (2016)
		Merits	Short computational time to train the data; robust to huge, missing, noisy, data; doesn't require many parameters	
		Demerits	Requires huge number of records and independent data attributes. Contradictory views on performance	
Object-based classification	OBC (ML)	Principle, and formula involved	Using spatial and spectral information, it divides the image into objects (Yan et al. 2006; Gao et al. 2011; Tehrany et al. 2014), which are a group of similar neighbouring pixels (Cleve et al. 2008) having similar DN values and inherent spatial features corresponding with real objects on the earth's surface (Hay et al. 2001). The process is termed as image segmentation (Petropoulos et al. 2012) (Fig. 3). The segmented features serve as the units for carrying out the OBC Once segmentation is done, classification is performed	Duro et al. (2012), Tehrany et al. (2014), Ouyang et al. (2011)
		Merits	Higher accuracy than pixel-based classification; smooth and contiguous LULC; identifies objects with proper shape	
		Demerits	labour intensive and time consuming	

Apparently, it indicates that object-based classification approach has performed better with original high-resolution dataset (i.e., LISS4 MX70 of 2016) than resolution-merged datasets (i.e., LISS 3 merged with LISS 4 and LISS 3 merged with IRS-1D of 2010 and 2005 respectively); although the resolution of resolution-merged datasets is similar that of the original dataset i.e. 5 m. This observation is not seen in any of the algorithms used in pixel-based classification approach.

Producer's accuracy (PA) and User's accuracy (UA)

Referring to PA and UA in Table 3, it is seen that no algorithm has highest PA and/or UA with respect to all the LULC classes in one or more years. Considering the notion of > 85% accuracy individually class-wise, it is observed that NN in 2003 and NB in 2016 have the highest PA for dense vegetation (87.1%) and water (90.32%) respectively. ML in 2005 (89.47%), ML in 2010 (95.45%), NN in 2005 (89.47%), RF in 2010 (95.83%), RF in 2016 (96.15%), NB in 2010 (88.46%) and OBC in 2005 (90.48%), in 2010 (91.67%) and in 2016 (100.00%) have the highest UA for water. RF in 2016 has the highest UA for cropland (85.17%). All the SVM kernels in 2010 have the highest UA for sediment (100.00%). It shows that the performance of classification algorithms is better with respect to UA in comparison to PA. Analysing the results class-wise, it is observed that the highest PA in 2005 is related to built-up (75.56%) classified using RF; in 2010 to water (68.43%) classified using SVM sigmoid and in 2016 to water classified using RF (80.65%). Similarly, the highest UA in 2005 is related to water (90.48%) classified using OBC; in 2010 to sediment (100.00%) classified using all the SVM kernels and;in 2016 to water (100.00%) classified using OBC.

Statistical significance assessment of LULC maps using Mc Nemar's test

From Table 4–6, the results of Mc Nemar's test with 5% significance level reveal that in 2005, OBC is statistically significant ($p < 0.05$) than many pixel-based algorithms (NN ($p = 0$), SVM linear ($p = 0.015$), SVM polynomial ($p = 0.023$), SVM RBF ($p = 0.038$), RF ($p = 0$) and NB ($p = 0.046$)). In year 2010, OBC is statistically significant than RF ($p = 0.001$) and in 2016, OBC is statistically significant than ML ($p = 0.033$), NN ($p = 0$) and NB ($p = 0$).

Analysis of Mc Nemar's test with 5% significance level for different pixel-based algorithms reveal that statistical significance ($p < 0.05$) exists between many pixel-based algorithms; however, no consistent pattern regarding statistical significance among the algorithms is observed. A statistically significant comparison of any two or three pixel-based classification algorithms can be done using p values mentioned in Tables 4, 5 and 6.

Temporal analysis of LULC change

Table 7 shows the temporal LULC change (in km^2) derived from all the classification algorithms for all the years. As RF and OBC, based on OA, are found out to be the best classification algorithms, only these two are discussed in detail here. The performance of rest of the algorithms can be studied from Table 7.

The area covered by water has decreased from year 2005 to 2016 for RF (46.41 km^2 to 20.38 km^2) as well as for OBC (29.08 km^2 to 19.79 km^2). These declining results in the are an obvious error as such sharp decline in the amount of water bodies in Delhi is not feasible. RF shows an overall decline in built-up (564.17 km^2 in 2005 to 450.592 km^2 in 2016) which is incorrect for a study area like Delhi which is constantly urbanising. The OBC shows a realistic trend of an increase in built-up (374.98 km^2 in 2005 to 564.93 km^2 in 2016), though the accuracy of the amount of area mapped as built-up cannot be relied upon as OBC has low P.A. (45.31% in 2005; 53.57% in 2010; and 69.63% in 2016) and U.A. for built-up for all the years. For dense vegetation, both RF (24.84 km^2 in 2005 to 67.62 km^2 in 2016) and OBC (20.59 km^2 in 2005 to 64.65 km^2 in 2016) shows an overall increase over the years, which is an incorrect detail considering the land-use of Delhi. To its credit, Delhi has only central ridge forest as dense vegetation and it has not increased with this huge magnitude over the given period. For sparse vegetation, both RF and OBC shows contrary overall trends. RF shows an increase in sparse vegetation (111.37 km^2 in 2005 to 247.80 km^2 in 2016) and OBC shows an overall decrease in sparse vegetation (258.232 km^2 in 2005 to 242.68 km^2 in 2016). Similar contrary trends are observed for cropland and fallow land by RF as well as OBC. The trend of LULC change for open land by RF as well as OBC is inaccurate as open land in Delhi, over the period, has a likelihood to get converted into either built-up or green spaces. Hence, decline in open land from 2005 to 2010 (49.30 km^2 in 2005 to 27.84 km^2 for RF and 81.11 km^2 to 51.79 km^2 for OBC) is understandable and justifiable, however, a sudden increase in open land area in 2016 is an error. Similar erroneous trend observations are seen for scrubland/forest class by RF and OBC. For sediment, RF and OBC show different trends. In RF, the area of sediment has increased in 2010 (42.33 km^2 from 12.10 km^2 in 2005) and then has declined to 10.43 km^2 in 2016. This sudden increase of sediment in 2010 is an error and unexplanatory. OBC shows a constant decline in sediment (30.38 km^2 in 2005 to 7.73 km^2 in 2016) though the amount of change appears huge considering sediment is located only along the banks of Yamuna River in Delhi.

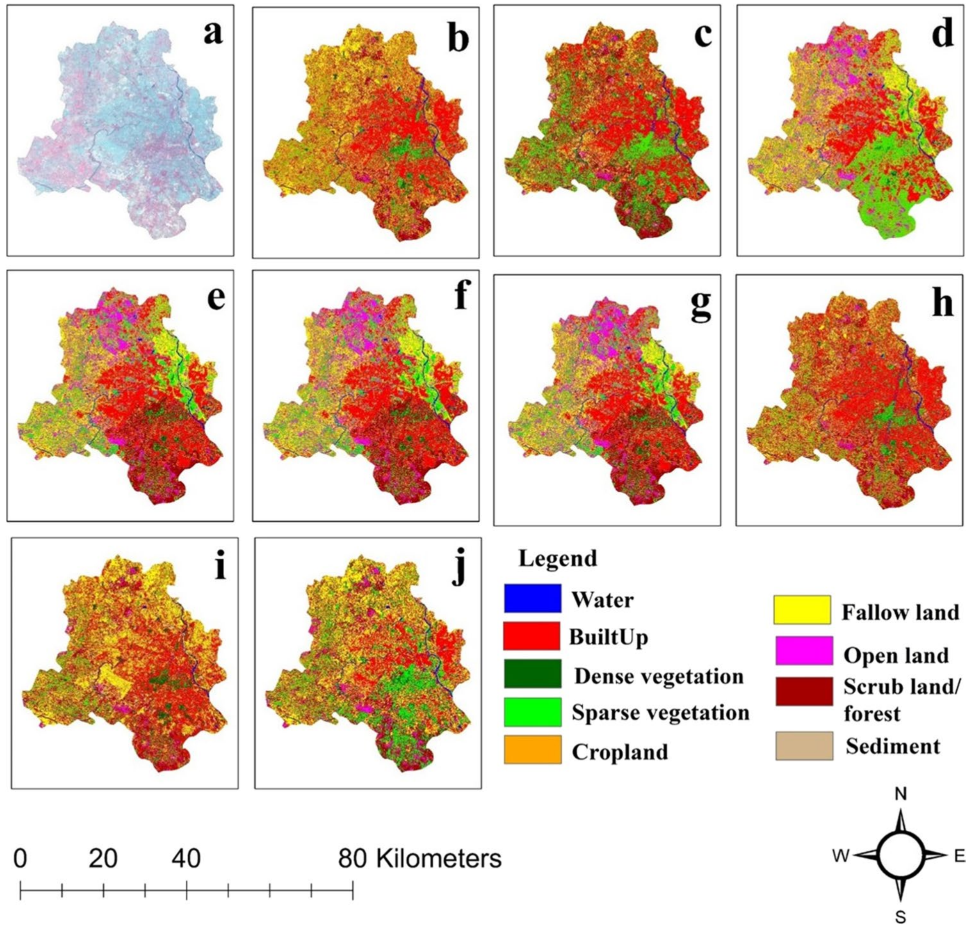


Fig. 4 Different classification algorithms performed on Delhi year 2005 dataset. **a** Standard False Colour Composite of Delhi satellite image, **b** ML, **c** NN, **d** SVM (linear), **e** SVM (polynomial), **f** SVM (RBF), **g** SVM (sigmoid), **h** RF, **i** NB, **j** OBC (ML)

Discussion

The results of the study infer that it is difficult to achieve higher overall accuracy in classifying large urban areas with detailed information using 5 m resolution satellite imageries. This is in consistence with the findings of Myint el al. (2011) stating that higher accuracy is difficult to be attained in detailed mapping of large urban areas. Also, the visual analysis of LULC maps reveal that the LULC maps prepared using pixel-based approach possess salt and pepper or mixed pixel effect and LULC prepared using object-based approach has provided a visually smoothened landscape in output LULC map that gives the appearance of earth-like landscape as found in Duro et al. (2012). This smoothening occurs because the heterogeneity in urban landscapes due to the presence of many different sized features in a small space in the area congest the spectral details of the urban features (Myint et al. 2006, 2011). This is the reason pixel-based classification leads to salt and pepper effect considering only spectral information. On the other hand, OBC considers spatial as well as spectral information of the features (Benz et al. 2004; Walter 2004; Myint et al. 2011; Duro et al. 2012) and it identifies the objects more precisely and leads to more accurate classification (Kelly et al. 2004; Cleve et al. 2008; Ouyang et al. 2011). Thus, the study demonstrates that OBC (ML) approach is preferable than pixel-based classification approach to prepare LULC for urban areas using satellite images having original high (5 m) spatial resolution. Among pixel-based classification, RF performs better compared to other algorithms. Despite having similar resolution (i.e. 5 m), original and resolution-merged dataset affects the performance of OBC. It illustrates that besides complex landscapes and classification algorithms, the type of remotely sensed data is another factor that affects the accuracy of the prepared LULC maps (Manandhar et al. 2009). In our study, it happens because the resolution merging technique used, i.e. hyperspherical color space (HCS) (Padwick et al. 2010) merges the edges of features with the shadow region in the image and thereby leads to the disappearance of smaller edges (Dahiya et al. 2013; Ghosh and Joshi 2013). Thus, it lacks spatial details (Ghosh and Joshi 2013) to some extent, which is a significant attribute in OBC. This is the reason, why resolution-merged datasets of 2005 and 2010 have shown lower accuracy for object-based approach. In this study, HCS resolution merging is used as it has generated resolution-merged datasets for years 2005

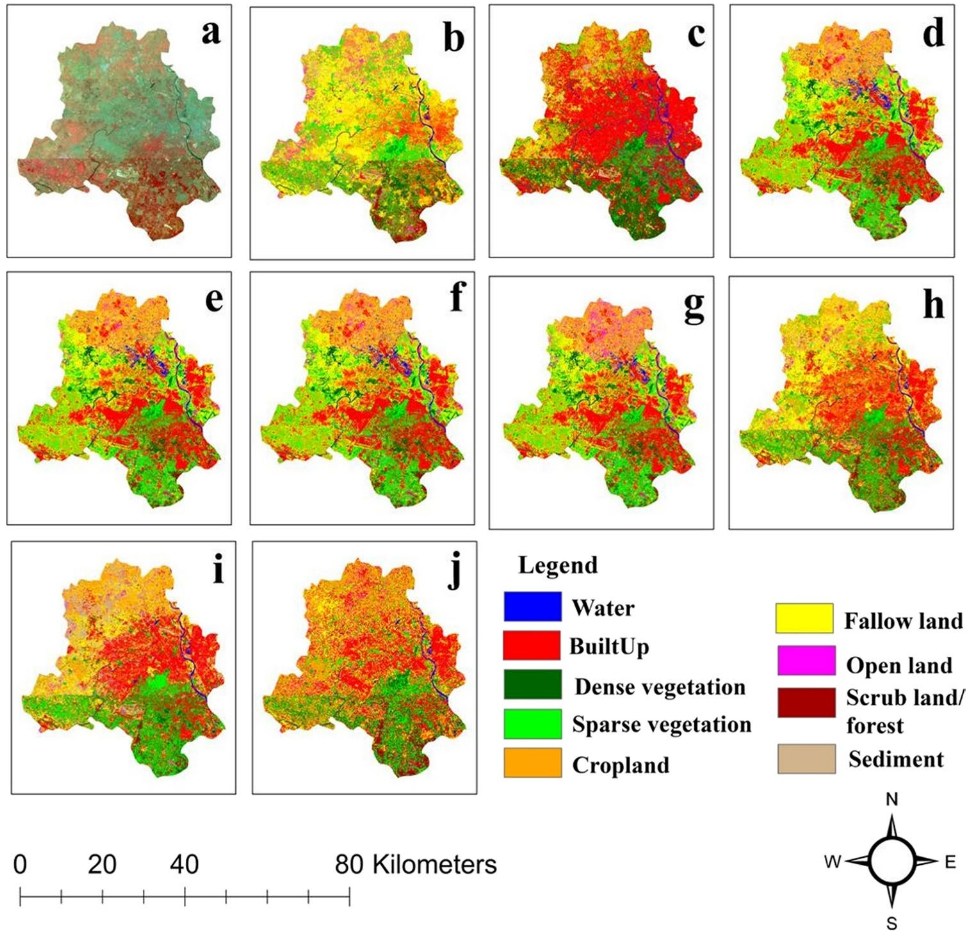


Fig. 5 Different classification algorithms performed on Delhi year 2010 dataset. **a** Standard False Colour Composite of Delhi satellite image, **b** ML, **c** NN, **d**SVM (linear), **e** SVM (polynomial), **f** SVM (RBF), **g** SVM (sigmoid), **h** RF, **i** NB, **j** OBC (ML)

and 2010 which appear visually more accurate (Agrafiotis and Georgopoulos 2015) than those generated using Ehler's fusion, wavelet, HPF, modified IHS and subtractive resolution merging methods. But considering the results and the fact that resolution-merging techniques affect the quality of the resolution merged products (Wang et al. 2005; Ghosh and Joshi 2013), it is suggested that before performing classification, the accuracy of merged datasets prepared using different techniques should be assessed by different measures and not only visually.

It has been observed that the time consumed in selecting the object features for OBC approach is almost equal to that of consumed in selecting the training data for pixel-based classification, provided the user has expertise in carrying out OBC. Otherwise, it can be very time consuming and labour intensive. With reference to the procedure of accuracy assessment of OBC, few researchers (e.g. Cleve et al. 2008) believe that a procedure that can assess the shape and topology of the features should be adopted because OBC takes into account the spatial topology, shape etc. of the classified features. In our study, to assess the performance of different pixel-based algorithms as well as OBC approach, pixel-based accuracy assessment method is used considering it to be the most suitable one as a pixel is the smallest unit of LULC map (Myint et al. 2011).The results reveal that unlike OA, the type of dataset (original or resolution merged) has no clear impact on the PA and UA of LULC classes with respect to different algorithms.

Higher PA of NN in 2005 for dense vegetation (87.1%) and that of NB in 2016 for water (90.32%) suggest NN and NB as the most powerful algorithms to classify the respective classes. Higher UA of OBC in 2005 for water (90.48%); that of RF and SVM in 2010 for water (95.83%) and sediment (100.00%) respectively and that of OBC and RF in 2016 for water (100.00%) and cropland (85.71%) suggest these algorithms as the most reliable ones in classifying the respective classes as accurate as their presence on the earth's surface. These results reveal that though based on OA, RF and OBC have performed as the best classifiers; class-wise, neither of them has higher (> 85%) PA for any of the LULC classes and UA for any other class except the ones mentioned a while ago. The PA and UA statistics (Table 3) show some shortcomings depicted by classification algorithms in few of the LULC classes. None of the algorithms has classified sediment class accurately in 2005 datasets, resulting into 0.00% PA and UA. The reason behind this could be

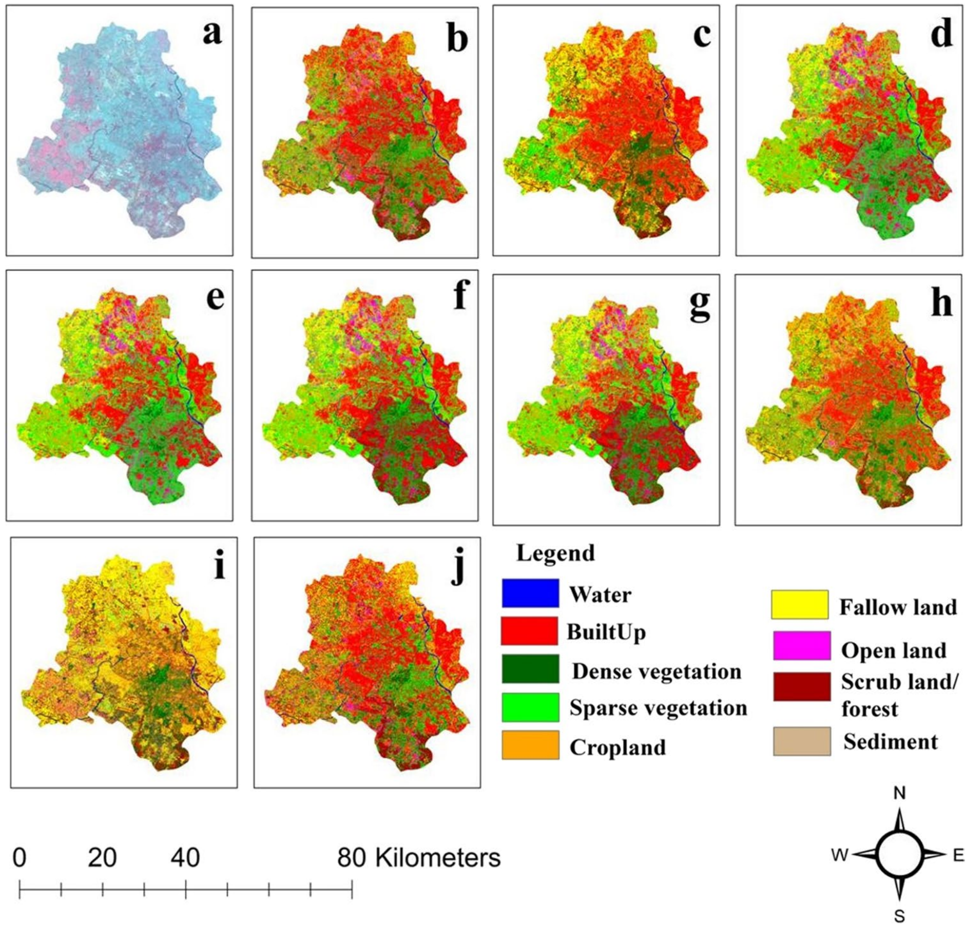


Fig. 6 Different classification algorithms performed on Delhi year 2016 dataset. **a** Standard False Colour Composite of Delhi satellite image, **b** ML, **c** NN, **d** SVM (linear), **e** SVM (polynomial), **f** SVM (RBF), **g** SVM (sigmoid), **h** RF, **i** B, **j** OBC (ML)

smaller percentage area of sediment in the study area. Similarly, NB in 2016 has not classified open land accurately. NN in 2005 and 2016 has not classified cropland and cropland and sediment respectively in the image. On visual basis, it was observed that the LULC maps in question do not contain enough number of pixels in the respective class that the accuracy of that class can be evaluated. Hence, it does not provide any PA or UA.

In addition to this, all the LULC maps were employed to analyse temporal LULC analysis. The only aim was to analyse the trend that how efficiently different algorithms mapped different LULC classes over the years. The nature and trend of LULC change was evaluated based on the knowledge of development occurred in the study area over the period. Comparison among different algorithms on LULC change or quantification of LULC change was not considered as the overall accuracy of all the LULC maps was quite low. The results revealed that neither of the two, RF and OBC, had shown satisfactory performance although OBC mapped LULC change trends correctly for built-up class.

Conclusion

In the present study, comparative evaluation of different classification algorithms and the impact of different types of satellite images on classification has been performed using confusion matrix and Mc Nemar's test. The results indicate that OBC is found to be statistically significant ($p<0.05$) than other algorithms in all the years (2005, 2010, 2016). Also, various pixel-based algorithms in the three years show statistical significance ($p<0.05$) although no consistent pattern has been observed. With an overall accuracy (54.98% in 2005; 52.58% in 2010; 56.83% in 2016), RF has performed as the best classification algorithm whereas Naive Bayes shows the least overall accuracy (39.11% in 2005; 41.14% in 2010; 35.42% in 2016). OBC exhibits lower overall accuracy (44.46% in 2005; 43.91% in 2010; 54.98% in 2016) in comparison to pixel-based algorithms. Moreover, the visual investigation of LULC reveals that despite lower accuracy, OBC derived LULC are visually smooth and contiguous in nature in comparison to pixel based derived LULC which possess salt and pepper effect. The assessment of different types of satellite data with respect to classification reveals that OBC has performed significantly better with original high-resolution dataset. The

Table 3 Accuracy assessment results of different classification algorithms for years 2005, 2010 and 2016. (PA: Producer accuracy; UA: User accuracy; Dense v.: dense vegetation; Sparse v.: sparse vegetation; Crop l.: cropland; Fallow l.: fallow land; Open l.: openland; Scrub l/f.: scrubland/forest)

LULC classes	Accuracy	ML 2005	NN 2005	SVM linear 2005	SVM Polynomial 2005	SVM RBF 2005	SVM Sigmoid 2005	RF 2005	NB 2005	OBC 2005	ML 2010	NN 2010	SVM linear 2010	SVM Polynomial 2010	SVM RBF 2010	SVM Sigmoid 2010	RF 2010	NB 2010	OBC 2010	ML 2016	NN 2016	SVM linear 2016	SVM Polynomial 2016	SVM RBF 2016	SVM Sigmoid 2016	RF 2016	NB 2016	OBC 2016
Water	PA	62.9	62.9	55.5	55.6	55.6	63.0	63.0	51.9	70.4	55.2	68.4	68.4	65.8	68.4	68.4	60.5	60.5	57.9	74.1	48.4	77.4	77.4	77.4	77.4	80.7	90.3	77.4
	UA	89.4	89.4	79.0	79.0	79.0	81.0	81.0	82.4	90.5	95.4	81.3	74.3	75.8	76.5	81.3	95.8	88.5	91.7	58.9	83.3	53.3	53.3	53.3	54.6	96.2	77.8	100.0
Built-up	PA	57.0	59.3	60.2	59.4	59.4	76.6	76.6	41.1	45.3	20.0	71.4	58.6	59.3	57.9	58.6	54.3	40.7	53.6	46.6	68.9	47.4	47.4	43.7	47.4	75.6	8.2	69.6
	UA	50.6	49.6	72.0	73.1	72.4	55.1	55.1	32.9	50.9	62.2	49.8	76.6	75.5	77.1	76.6	61.8	61.3	51.4	59.4	61.6	74.4	75.3	75.6	77.1	73.4	61.1	59.1
Dense v	PA	50.0	8.3	41.7	50.0	50.0	58.3	58.3	91.7	41.7	51.6	87.1	51.6	48.4	48.4	54.8	38.7	35.5	19.4	84.3	84.4	75.0	75.0	75.0	75.0	75.0	87.5	43.8
	UA	30.0	4.0	55.6	54.6	50.0	63.6	63.6	42.3	38.5	41.0	32.1	41.0	40.5	40.5	39.5	41.4	36.7	18.2	54.0	43.6	53.3	55.8	55.8	57.1	57.1	48.3	42.4
Sparse v	PA	13.1	50.0	22.6	19.1	16.7	43.4	43.4	6.0	48.8	52.3	28.6	34.5	36.9	36.9	36.9	58.3	44.1	40.5	48.3	9.9	46.2	48.4	48.4	49.5	42.9	15.4	50.6
	UA	30.5	32.0	29.2	26.2	24.1	63.2	63.2	18.5	41.0	43.1	42.9	26.6	28.4	27.9	28.2	50.0	39.0	33.0	41.9	13.0	33.3	33.6	33.1	33.1	45.4	30.4	54.1
Crop l	PA	64.4	–	35.6	35.6	35.6	51.1	51.1	28.9	33.3	53.2	40.3	59.7	59.7	59.7	53.2	46.8	62.9	64.5	51.6	–	59.7	59.7	59.7	59.7	48.4	51.6	77.4
	UA	26.1	–	69.6	69.6	66.7	40.4	40.4	36.1	83.3	53.2	67.6	59.7	56.9	58.7	60.0	58.0	51.3	54.1	74.4	–	78.7	78.7	78.7	80.4	85.7	78.1	70.6
Fallow l	PA	48.3	29.1	58.9	58.9	58.3	48.3	48.3	56.3	41.7	65.6	24.5	54.9	55.9	55.9	52.0	66.7	37.3	35.3	26.7	57.0	40.7	41.9	44.2	44.2	52.3	64.0	30.2
	UA	58.4	55.0	61.8	62.2	62.0	70.2	70.2	51.5	52.1	36.8	41.7	42.1	43.2	43.2	41.4	46.0	40.9	39.6	44.2	35.3	52.2	52.9	54.3	58.5	39.1	27.6	39.4
Open l	PA	13.2	24.5	56.6	56.6	56.6	35.9	35.9	22.6	28.3	17.6	2.0	9.8	9.8	11.8	13.7	17.7	5.9	9.8	44.4	3.2	57.1	57.1	57.1	60.3	30.2	0.0	30.2
	UA	70.0	76.4	35.3	33.3	33.3	79.2	79.2	66.7	55.6	24.3	50.0	19.2	19.2	20.0	18.4	64.3	21.4	22.7	52.8	50.0	64.3	64.3	61.0	64.4	70.4	0.0	65.5
Scrub l/f	PA	57.5	67.5	72.5	72.5	72.5	62.5	62.5	47.5	62.5	57.6	38.5	34.6	34.6	34.6	34.6	61.5	38.5	65.4	65.6	62.5	56.3	53.1	53.1	56.3	71.9	59.4	78.1
	UA	31.1	25.0	32.2	32.2	31.9	28.7	28.7	23.2	22.3	36.5	23.3	32.1	33.3	33.3	34.6	36.4	22.7	40.5	39.6	20.2	34.0	34.0	34.0	34.6	35.4	21.6	34.3
Sediment	PA	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	25.0	50.0	25.0	25.0	25.0	25.0	37.5	62.5	37.5	70.0	–	40.0	40.0	40.0	50.0	10.0	50.0	20.0
	UA	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	16.6	14.8	100.0	100.0	100.0	100.0	27.3	7.0	42.9	17.5	–	25.0	25.0	25.0	35.7	16.7	13.9	40.0
O.A	44.1	40.6	51.7	51.1	50.6	52.2	54.1	39.1	44.5	43.3	44.7	48.3	48.7	50.6	55.0	52.6	41.1	43.9	49.4	39.7	52.4	52.8	48.7	48.0	56.8	35.4	55.0	
Kappa	0.3	0.2	0.4	0.4	0.4	0.4	0.4	0.2	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4	0.3	0.3	0.4	0.2	0.4	0.4	0.3	0.3	0.4	0.2	0.4	

Table 4 Results of Mc Nemar's test (p value, chi square) for year 2005 dataset (*denotes 2 tailed p value s; statistically significant values ($p<0.05$) are in bold)

	ML	NN	SVM linear	SVM polynomial	SVM RBF	SVM sigmoid	RF	NB	OBC
ML		**0, 24.797**	**0.011, 6.534**	**0.018, 5.619**	**0.031, 4.676**	0.307, 1.042	**0, 17.137**	**0.027, 4.87**	
NN	**0, 24.797**		**0.046, 3.965**	**0.03, 4.694**	**0.018, 5.61**	**0.001, 12.108**	0.404, 0.696	0, 39.2	**0, 24.797**
SVM linear	**0.011, 6.534**	**0.046, 3.965**		0.549*	0.14*	**0.002, 9.49**	0.234, 1.417	**0, 20.405**	**0.011, 6.534**
SVM polynomial	**0.018, 5.619**	**0.03, 4.694**	0.549*		0.25*	**0.005, 7.848**	0.16, 1.97	**0, 19.051**	**0.018, 5.619**
SVM RBF	**0.031, 4.676**	**0.018, 5.61**	0.14*	0.25*		**0.015, 5.953**	0.106, 2.619	**0, 17.388**	**0.031, 4.676**
SVM sigmoid	0.307, 1.042	**0.001, 12.108**	**0.002, 9.49**	**0.005, 7.848**	**0.015, 5.953**		**0.006, 7.656**	**0.003, 8.529**	0.307, 1.042
RF	0, 17.137	0.404, 0.696	0.234, 1.417	0.16, 1.97	0.106, 2.619	**0.006, 7.656**		**0, 42.006**	0, 17.137
NB	**0.027, 4.87**	0, 39.2	**0, 20.405**	**0, 19.051**	**0, 17.388**	**0.003, 8.529**	**0, 42.006**		**0.027, 4.87**
OBC	1, 0	**0, 21.114**	**0.015, 5.952**	**0.023, 5.138**	**0.038, 4.29**	0.344, 0.895	**0, 14.405**	**0.046, 3.967**	1, 0

Table 5 Results of Mc Nemar's test (p value, chi square) for year 2010 dataset. (*denotes 2 tailed p value s; statistically significant values ($p<0.05$) are in bold)

	ML	NN	SVM linear	SVM polynomial	SVM RBF	SVM sigmoid	RF	NB	OBC
ML		0.696, 0.153	0.078, 3.115	0.057, 3.613	0.055, 3.681	0.11, 2.56	**0, 12.376**	0.422, 0.644	0.891, 0.019
NN	0.696, 0.153		0.192, 1.703	0.151, 2.061	0.157, 2.005	0.245, 1.35	**0.003, 9.14**	0.186, 1.751	0.835, 0.043
SVM linear	0.078, 3.115	0.192, 1.703		0.687*	0.754*	0.86, 0.031	0.106, 2.616	**0.006, 7.64**	0.093, 2.814
SVM polynomial	0.057, 3.613	0.151, 2.061	0.687*		1*	0.607, 0.265	0.139, 2.186	**0.004, 8.466**	0.067, 3.36
SVM RBF	0.055, 3.681	0.157, 2.005	0.754*	1*		0.571, 0.321	0.141, 2.162	**0.003, 8.556**	0.07, 3.289
SVM sigmoid	0.11, 2.56	0.245, 1.35	0.86, 0.031	0.607, 0.265	0.571, 0.321		0.076, 3.148	**0.01, 6.715**	0.126, 2.346
RF	**0, 12.376**	**0.003, 9.14**	0.106, 2.616	0.139, 2.186	0.141, 2.162	0.076, 3.148		**0, 26.579**	**0.001, 11.079**
NB	0.422, 0.644	0.186, 1.751	**0.006, 7.64**	**0.004, 8.466**	**0.003, 8.556**	**0.01, 6.715**	**0, 26.579**		0.293, 1.107
OBC	0.891, 0.019	0.835, 0.043	0.093, 2.814	0.067, 3.36	0.07, 3.289	0.126, 2.346	**0.001, 11.079**	0.293, 1.107	

Table 6 Results of Mc Nemar's test (p value, chi square) for year 2016 dataset (*denotes 2 tailed p value s; statistically significant values ($p<0.05$) are in bold)

	ML	NN	SVM linear	SVM polynomial	SVM RBF	SVM sigmoid	RF	NB	OBC
ML		**0, 15.63**	0.174, 1.844	0.124, 2.369	0.199, 1.647	**0.024, 5.097**	**0.005, 7.922**	**0, 27.574**	**0.033, 4.571**
NN	**0, 15.63**		**0, 23.959**	**0, 25.128**	**0, 22.672**	**0, 29.944**	**0, 46.762**	**0, 2.384**	**0, 29.884**
SVM linear	0.174, 1.844	**0, 23.959**		0.625*	1*	**0.049***	0.106, 2.619	**0, 38.696**	0.351, 0.871
SVM polynomial	0.124, 2.369	**0, 25.128**	0.625*		0.453*	**0, 0.143**	0.141, 2.162	**0, 40.042**	0.432, 0.617
SVM RBF	0.199, 1.647	**0, 22.672**	1*	0.453*		**0.031***	0.094, 2.81	**0, 38.756**	0.316, 1.005
SVM sigmoid	**0.024, 5.097**	**0, 29.944**	**0.049***	**0, 0.143**	**0.031***		0.326, 0.966	**0, 45.662**	0.776, 0.081
RF	**0.005, 7.922**	**0, 46.762**	0.106, 2.619	0.141, 2.162	0.094, 2.81	0.326, 0.966		**0, 67.474**	0.487, 0.482
NB	**0, 27.574**	**0, 2.384**	**0, 38.696**	**0, 40.042**	**0, 38.756**	**0, 45.662**	**0, 67.474**		**0, 47.935**
OBC	**0.033, 4.571**	**0, 29.884**	0.351, 0.871	0.432, 0.617	0.316, 1.005	0.776, 0.081	0.487, 0.482	**0, 47.935**	

Table 7 Temporal LULC analysis (2005–2016) by different classification algorithms

LULC classes	ML 2005	NN 2005	SVM linear 2005	SVM Polynomial 2005	SVM RBF 2005	SVM Sigmoid 2005	RF 2005	NB 2005	OBC 2005	ML 2010	NN 2010	SVM linear 2010	SVM Polynomial 2010	SVM RBF 2010	SVM Sigmoid 2010	RF 2010	NB 2010	OBC 2010	ML 2016	NN 2016	SVM linear 2016	SVM Polynomial 2016	SVM RBF 2016	SVM Sigmoid 2016	RF 2016	NB 2016	OBC 2016
Water	29.2	29.5	24.8	25.1	24.5	25.4	46.4	30.9	22.7	16.9	48.9	45.0	41.4	40.7	34.6	23.6	23.6	26.8	24.7	8.0	32.7	32.8	32.8	34.4	20.4	33.1	25.1
Built-up	445.8	499.8	340.8	339.0	339.4	335.3	545.1	440.5	484.0	139.3	690.8	362.7	361.3	363.9	352.9	414.0	375.2	577.8	533.5	444.7	406.7	402.0	397.5	388.0	431.5	93.1	667.5
Dense v	45.1	65.4	13.9	14.3	14.4	15.7	24.8	53.1	10.5	74.2	171.8	79.5	83.9	84.0	96.3	47.9	55.5	13.9	75.7	108.0	51.4	47.2	47.2	47.3	67.6	100.5	11.0
Sparse v	86.4	342.7	208.9	207.6	204.5	212.0	111.4	68.6	228.1	257.2	163.9	296.1	293.7	292.8	282.4	245.8	237.4	254.6	296.4	144.8	402.0	412.3	412.2	415.7	247.8	110.0	227.5
Crop l	287.6	1.6	67.9	68.2	69.8	63.3	179.7	102.5	136.7	156.1	82.1	148.4	149.8	148.4	133.0	120.0	185.0	188.5	62.7	5.5	71.6	71.4	72.1	69.8	57.4	73.5	200.4
Fallow l	350.4	203.4	384.1	378.3	373.6	354.9	260.7	472.1	436.8	635.3	169.3	442.2	444.2	438.9	441.2	466.1	272.7	351.4	220.8	479.2	288.6	290.8	293.5	301.1	417.4	687.6	282.6
Open l	25.7	42.2	242.5	250.0	255.3	275.8	49.3	41.3	102.5	91.8	2.3	48.5	49.4	55.4	84.3	27.8	36.0	32.4	113.1	13.1	96.8	96.4	97.6	95.1	47.5	57.2	20.1
Scrub l/f	203.5	276.3	197.6	195.8	196.2	199.0	254.4	238.8	59.5	82.7	97.1	60.0	58.8	58.4	57.6	96.5	103.1	34.6	131.3	280.4	131.7	128.7	128.7	128.9	183.9	266.9	47.4
Sediment	10.3	23.1	3.6	5.7	6.2	2.6	12.1	36.1	3.4	30.5	57.8	1.4	1.4	1.5	1.7	42.3	195.7	4.0	25.8	0.4	2.5	2.4	2.4	3.7	10.4	62.1	2.3

poorer performance of OBC with resolution-merged images could be attributed to the reason that HCS resolution merging algorithm that is used in this study degrades the sharpness and spatial details to some extent in the output, an entity that is significant in OBC algorithm. Hence, the study suggests that to prepare LULC map of an urban area using satellite images of original 5 m spatial resolution, OBC approach is recommended whereas with resolution merged 5 m spatial resolution, RF algorithm in pixel-based approach is recommended. The findings of the study may be useful for future studies mapping urban land-use using higher resolution or resolution merged images.

Acknowledgements The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work under grant number (RGP2/194/42). The authors are thankful to National Aeronautics and Space Administration (NASA) and U.S. Geological Survey (USGS) for providing freely available Landsat data, used in this manuscript. The authors are thankful to Geoinformatics laboratory, TERI School of Advanced Studies for providing access to software.

Authors' contributions The author AB and CKS prepared the concept of manuscript. Analysis was conducted by AB, JM and supervised by CKS. Draft writing was done by AB, SP, JM and edited by SP, SG, and CKS.

Funding This research received partial grant from United Nations University-Institute for the Advanced Study of Sustainability. The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work.

Availability of data and material There are no linked research data sets for this submission. The data that has been used is confidential.

Code availability ERDAS® IMAGINE 2016, ArcGIS 10.5 and R version 3.3.2 softwares are used in this manuscript.

Declaration

Conflict of interest The authors declare that there are no conflict of interest associated with this manuscript.

Ethical statement All ethical practices have been followed in relation to the development, writing, and publication of the article.

References

Agrafiotis P, Georgopoulos A (2015) Comparative assessment of very high resolution satellite and aerial orthoimagery. Int Arch Photogram Remote Sens Spatial Inf Sci 40(3):1

Agresti A (2002) Categorical data analysis. Wiley, New York

Anderson JR, Hardy EE, Roach JT, Witmer RE (1976) A land use and land cover classification system for use with remote sensor data. U.S. Geological Survey Professional Paper 964:28.

Baatz M, Benz U, Dehghani S, Heynen M, Höltje A, Hofmann P, Lingenfelder I, Mimler M, Sohlbach M, Weber M, Willhauck G (2004) eCognition Professional 4.0 User Guide. Definiens Imaging GmbH. Definiens, Munich

Benz UC, Hofmann P, Willhauck G, Lingenfelder I, Heynen M (2004) Multiresolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS J Photogram Remote Sens 58(3–4):239–258

Chen Z, Zhao Z, Gong P, Zeng B (2006) A new process for the segmentation of high resolution remote sensing imagery. Int J Remote Sens 27:4991–5001

Chi M, Feng R, Bruzzone L (2008) Classification of hyperspectral remote-sensing data with primal SVM for small-sized training dataset problem. Adv Space Res 41(11):1793–1799

Cohen J (1960) A coefficient of agreement for nominal scales. Educ Psychol Meas 20(1):37–46

Cleve C, Kelly M, Kearns FR, Moritz M (2008) Classification of the wildland–urban interface: a comparison of pixel-and object-based classifications using high-resolution aerial photography. Comput Environ Urb Syst 32(4):317–326

Dahiya S, Garg PK, Jat MK (2013) A comparative study of various pixel-based image fusion techniques as applied to an urban environment. Int J Image Data Fusion 4(3):197–213

De Leeuw J, Jia H, Yang L, Liu X, Schmidt K, Skidmore AK (2006) Comparing accuracy assessments to infer superiority of image classification methods. Int J Remote Sens 27:223–232

Dietterich TG (1998) Approximate statistical tests for comparing supervised classification learning algorithms. Neural Comput 10:1895–1923

Dingle Robertson L, King DJ (2011) Comparison of pixel- and object-based classification in land cover change mapping. Int J Remote Sens 32(6):1505–1529

Domingos P, Pazzani M (1997) On the optimality of the simple Bayesian classifier under zero-one loss. Mac Learn 29:103–130

Durieux L, Lagabrielle E, Nelson A (2008) A method for monitoring building construction in urban sprawl areas using object-based analysis of Spot 5 images and existing GIS data. ISPRS J Photogram Remote Sens 63(4):399–408

Duro DC, Franklin SE, Dubé MG (2012) A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sens Environ 118:259–272

ESRI AD (2016) Release 10.5. Environmental Systems Research Institute, Inc., Redlands, CA

Foody GM (2004) Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. Photogram Eng Remote Sens 70:627–634

Foody GM, Mathur A (2004) Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. Remote Sens Environ 93(1):107–117

Gao Y, Kerle N, Mas JF (2009) Object-based image analysis for coal fire-related land cover mapping in coal mining areas. Geocarto Int 24(1):25–36

Gao Y, Marpu P, Niemeyer I, Runfola DM, Giner NM, Hamill T, Pontius RG Jr (2011) Object-based classification with features extracted by a semi-automatic feature extraction algorithm–SEaTH. Geocarto Int 26(3):211–226

Ghosh A, Joshi PK (2013) Assessment of pan-sharpened very high-resolution WorldView-2 images. Int J Remote Sens 34(23):8336–8359

Gudiyangada Nachappa T, Kienberger S, Meena SR, Hölbling D, Blaschke T (2020) Comparison and validation of per-pixel and object-based approaches for landslide susceptibility mapping. Geomat Nat Hazard Risk 11(1):572–600

Hay GJ, Marceau DJ, Dube P, Bouchard A (2001) A multiscale framework for landscape analysis: object-specific analysis and upscaling. Lands Ecol 16(6):471–490

Hayes MM, Miller SN, Murphy MA (2014) High-resolution landcover classification using Random Forest. Remote Sens Lett 5(2):112–121

Herold M, Liu X, Clarke KC (2003) Spatial metrics and image texture for mapping urban land use. Photogramm Eng Remote Sensing 69(9):991–1001

Hexagon Geospatial (2016) ERDAS IMAGINE 2016. Intergraph Geospatial, Huntsville

Hu X, Weng Q (2011) Impervious surface area extraction from IKONOS imagery using an object-based fuzzy method. Geocarto Int 26(1):3–20

Jadhav SD, Channe HP (2016) Comparative study of K-NN, naive bayes and decision tree classification techniques. Int J Sci Res 5(1):1842–1845

Jozdani SE, Johnson BA, Chen D (2019) Comparing deep neural networks, ensemble classifiers, and support vector machine algorithms for object-based urban land use/land cover classification. Remote Sens 11(14):1713

Karatzoglou A, Smola A, Hornik K, Zeileis A (2004) kernlab-an S4 package for kernel methods in R. J Stat Softw 11(9):1–20

Kavzoglu T (2017) Object-oriented random forest for high resolution land cover mapping using Quickbird-2 imagery. In: Handbook of neural computation, Academic Press, pp 607-619

Kelly M, Shaari D, Guo Q, Liu D (2004) A comparison of standard and hybrid classifier methods for mapping hardwood mortality in areas affected by sudden oak death. Photogram Eng Remote Sens 70(11):1229–1239

Kotsiantis SB, Zaharakis I, Pintelas P (2007) Supervised machine learning: A review of classification techniques. Emerg Artif Intell Appl Comput Eng 160(1):3–24

Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2(3):18–22

Long JA, Lawrence RL, Greenwood MC, Marshall L, Miller PR (2013) Object oriented crop classification using multitemporal ETM+ SLC-off imagery and random forest. Gisci Remote Sens 50(4):418–436

Majka M (2018) naivebayes: High Performance Implementation of the Naive Bayes Algorithm. R package version 0.9.2. https://CRAN.R-project.org/package=naivebayes. Accessed April 2018

Manandhar R, Odeh IO, Ancev T (2009) Improving the accuracy of land use and land cover classification of Landsat data using post-classification enhancement. Remote Sens 1(3):330–344

Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: A review. ISPRS J Photogram Remote Sens 66(3):247–259

Myint SW, Gober P, Brazel A, Grossman-Clarke S, Weng Q (2011) Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sens Environ 115(5):1145–1161.

Myint SW, Mesev V, Lam NSN (2006) Texture analysis and classification through a modified lacunarity analysis based on differential box counting method. Geograph Anal 38:371–390

Nemmour H, Chibani Y (2006) Multiple support vector machines for land cover change detection: an application for mapping urban extensions. ISPRS J Photogram Remote Sens 61(2):125–133

Otukei JR, Blaschke T (2010) Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. Int J Appl Earth Observ Geoinform 12:S27–S31

Ouyang ZT, Zhang MQ, Xie X, Shen Q, Guo HQ, Zhao B (2011) A comparison of pixel-based and object-oriented approaches to VHR imagery for mapping saltmarsh plants. Ecol Inform 6(2):136–146

Padwick C, Deskevich M, Pacifici F, Smallwood S (2010) WorldView-2 pan-sharpening. In: Proceedings of the ASPRS 2010 annual conference, San Diego, CA, USA, vol 2630

Petropoulos GP, Kalaitzidis C, Vadrevu KP (2012) Support vector machines and object-based classification for obtaining land-use/

cover cartography from hyperion hyperspectral imagery. Comp Geosci 41:99–107

Puissant A, Rougier S, Stumpf A (2014) Object-oriented mapping of urban trees using Random Forest classifiers. Int J Appl Earth Obser Geoinform 26:235–245

Qu LA, Chen Z, Li M, Zhi J, Wang H (2021) Accuracy improvements to pixel-based and object-based LULC classification with auxiliary datasets from Google Earth engine. Remote Sens 13(3):453

R Development Core Team (2016) R: a language and environment for statistical computing, Vienna, Austria. http://www.R-project.org/. Accessed September 2017

Ren J (2012) ANN vs. SVM: Which one performs better in classification of MCCs in mammogram imaging. Knowledge-Based Syst 26:144–153

Robles Granda PD (2011) A new image classification algorithm based on additive groves. Unpublished MSc Thesis. Carbondale (IL): Southern Illinois University at Carbondale.

Rodriguez-Galiano VF, Ghimire B, Rogan J, Chica-Olmo M, Rigol-Sanchez JP (2012) An assessment of the effectiveness of a random forest classifier for landcover classification. ISPRS J Photogram Remote Sens 67:93–104

Rozenstein O, Karnieli A (2011) Comparison of methods for land-use classification incorporating remote sensing and GIS inputs. Appl Geograp 31(2):533–544

Shao Y, Lunetta RS (2012) Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. ISPRS J Photogram Remote Sens 70:78–87

Srivastava PK, Han D, Rico-Ramirez MA, Bray M, Islam T (2012) Selection of classification techniques for land use/land cover change investigation. Adv Space Res 50(9):1250–1265

Su Y, Huang PS, Lin CF, Tu TM (2004) Target-cluster fusion approach for classifying high resolution IKONOS imagery. IEEE Proc Vis Image Sig Process 151:241–249

Tassi A, Gigante D, Modica G, Di Martino L, Vizzari M (2021) Pixel-vs object-based Landsat 8 data classification in google earth engine using random forest: the case study of Maiella National Park. Remote Sens 13(12):2299.

Tehrany MS, Pradhan B, Jebuv MN (2014) A comparative assessment between object and pixel-based classification approaches for land use/land cover mapping using SPOT 5 imagery. Geocarto Int 29(4):351–369

Tien Bui D, Pradhan B, Lofman O, Revhaug I (2012) Landslide susceptibility assessment in Vietnam using support vector machines, decision tree, and Naïve Bayes models. Math Probl Eng 2012:26. https://doi.org/10.1155/2012/974638

Tu JV (1996) Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidem 49(11):1225–1231

Venables WN, Ripley BD (2002) Modern applied statistics with S, 4th edn. Springer, New York. ISBN 0-387-95457-0

Walter V (2004) Object-based classification of remote sensing data for change detection. ISPRS J Photogram Remote Sens 58:225–238

Wang Z, Ziou D, Armenakis C, Li D, Li Q (2005) A comparative analysis of image fusion methods. IEEE Trans Geosci Remote Sens 43(6):1391–1402

Whiteside TG, Boggs GS, Maier SW (2011) Comparing object-based and pixel based classifications for mapping savannas. Int J Appl Earth Observ Geoinform 13(6):884–893

Wiesmann D, Quinn D (2011) Rasclass: supervised raster image classification. R package version 0.2.1. http://cran.r-project.org/web/packages/rasclass/index.html. Accessed January 2018

Yan G, Mas JF, Maathuis BHP, Xiangmin Z, Van Dijk PM (2006) Comparison of pixel-based and object-oriented image classification approaches—a case study in a coal fire area, Wuda, Inner Mongolia, China. Int J Remote Sens 27(18):4039–4055

Zar JH (2009) Biostatistical analysis, 5th edn. Prentice Hall, Upper Saddle River

Zhang A (2014) Collaboration in the Australian and Chinese mobile telecommunication markets. Springer, New York