IDENTIFICATION OF BEST MACHINE LEARNING TECHNIQUE AND MONITORING AND DETECTING CHANGES OF LAND USE LAND COVER (LULC) FROM 2005 TO 2019: A CASE STUDY IN COLOMBO DISTRICT, SRI LANKA

M.W.I.L. Somawansha[1]
[1] Survey Department of Sri Lanka,
Center for Remote Sensing, SGO, Colombo 05, Sri Lanka,
Email: lakminisomawansha@gmail.com

KEY WORDS: SVM, RF, ANN, Change detection

ABSTRACT: Monitoring and mapping landscapes of rapidly developing city can significantly contribute to understand the complex growth of urbanization. However, generating accurate and temporal Land Use Land Cover (LULC) maps of such cities are a requirement as well as challenge due to spatial heterogeneity as well as fast dynamic land use practices. To monitor the changes of occurring in Earth surface features, it is required to perform change detection using temporal remote sensing data. Now a days many machine learning (ML) classifiers are used for generating to LULC classified maps. Here most major challenge is identification the best technique. This study not only focuses on the comparison of some ML classification algorithms but also monitoring and detecting changes using time series data.

In this study an attempt is made to observe the changes in LULC features of Colombo, the capital of Sri Lanka. The Landsat multispectral data was used in this study and was acquired from 2005-2019. Three different ML classifiers, Random forest (RF), support vector machine (SVM) and artificial neural network (ANN) was performed for LULC classification on Landsat 2005-2019 data. Among this ML classification the best classification algorithm observed for this study was identified. The multispectral data was classified according to both standard color and scheme of LULC maps in Sri Lanka. There are seven classes according to base map. Such as urban, plantation, forest, paddy, grass, scrub land, and water. By using all the data sets from 2005 – 2019 (2005, 2008, 2011, 2015, 2019), was identified both the best ML classifier and change detection. It was observed from the classified data of Landsat-2019. Here ML classifier's out performances are SVM-(OA-92.500% kappa-0.9124), RF-(OA- 85.2778%, kappa-0.8280), and ANN-(OA-71.9403% kappa-0.6703). So SVM was used for classify other data. The urban land use patterns were increased from 2005 to 2019 and paddy fields were reduced. There has been no significant change in water and forest land use classes.

1. INTRODUCTION

Land use and land cover (LULC) can be defined as: Land cover is the physical material at the surface of the earth which includes trees, grass, water, plantation, etc. Whereas land use is the human use of land which includes urbanization, agriculture, industrialization, etc. LULC change is one of the major influencing factors for the landscape change. Although urban areas currently cover only 3% of earth land surface, they have marked effects on environment condition and both local and global scale. LULC Change is a general term for the human modification of Earth's terrestrial surface. Though humans have been modifying land to obtain food and other essentials for thousands of years, current rates, extents and intensities of LULC change are far greater than even in history, driving unprecedented change in eco system and environmental processes at local, regional and global scales. Monitoring and mediating the negative consequences of LULC change while sustaining the production of essential resources has therefore become a major priority of researchers and policymakers around the world (Elis, 2011).

2. LITERATURE REVIEW

1.1 Land Use Land Cover Classifiers the Importance of Machine Learning Classifier

Land Use Land Cover Classifiers Extracting accurate LULC data from remotely sensed images require good image classification techniques. In general, these classifiers can be grouped as supervised and unsupervised, or parametric and nonparametric, or hard and soft classification, or per-pixel and sub-pixel based classifiers. Many classifiers exist whose performance are affected by various factors such as choice of training samples, heterogeneity of study area, sensors, number of classes to identify and so on. (Lu, 2007). Machine Learning Classifiers are reported to produce higher accuracy even with complex data and higher number input features. (Aksoy, 2005). Few of the popular classifiers are RF, SVM, ANN etc. RF uses random subset of training data to construct multiple decision trees. Other classifiers such as ANN follow a neural network pattern and build multiple layer of nodes to passes input observations back and forth during the learning process (Multi-Layer Perceptron) until it reaches a termination condition (Mas, 2008).

On the other hand classifiers such as SVM find a subset of training data as support vectors by fitting a hyper plane that separates two classes in the best possible way (Justice, 2002). Among all these classifiers, most literatures suggest that RF and SVM have an upper hand in most classification scenarios as they outperform other machine classifiers (Belgiu, 2016)

Urban sprawl refers to the extent of urbanization, which is a global phenomenon mainly driven by population growth and large scale migration. Urban planners require information related to the rate of growth, pattern and extent of sprawl to provide basic amenities such as water, sanitation, electricity etc. GIS and remote sensing data along with collateral data help in analysing the growth, pattern and extent of sprawl. With spatial and temporal analyses along with modelling it was possible to identify pattern of sprawl. Apart from the extent of sprawl attempts were also made to describe some of landscape metrics required for quantifying sprawl. For understanding and modelling this dynamic phenomenon, prominent causative factors were considered (H S, 2004).

1.2 Support Vector Machine

SVM is one of the widely used classifiers in remote sensing field. SVM gained its importance due to highly accurate classification results with lesser training samples, which is usually a limitation in land use land cover classification scenarios.

SVM is a linear binary classifier which is based on the concept that training samples which are at closer proximity to the boundaries of a class will discriminate a class better than other training samples. Hence SVM focuses on finding an optimal hyper plane which separates the input training samples of various classes. The samples present close to the boundaries of a class and at minimum distance to the hyper plane are taken as support vectors, which are used for the actual training (Cortes, 1995)

Another technique adapted to deal with non-linear input data(x) is the transformation of an input space to another higher dimensional feature space where, the training samples can be linearly separated. This transformation is achieved through a kernel trick where a mapping function Φ transforms x into Φ(x). (Boser, 1996) Training problem appears in the form of dot product of two vectors (Φ(xi), Φ(xj)). The computational cost of higher dimensional space (Φ(xi), Φ(xj)) is less because the following kernel transformation k is applied as shown in equation 1.

$$\Phi(xi), \Phi(xj)) = k(xi\ , xj)$$

Equation 1 kernel transformation

Additionally, this has added advantage that the knowledge of the mapping function is not needed (HUANG, 2002). Only the user has to choose a kernel which follows Mercer’s Theorem. Various kernel functions exist such as polynomial kernel, linear kernel and radial basis kernel (RBF). The choices of kernels also affect the results of the classification. A kernel such as RBF has a user-defined γ parameter which controls the influence of a training sample on the decision boundary. Higher the values of γ, more tightly fit are the decision boundaries around the samples. But this can lead to over fitting. Hence it is necessary to strike a right balance (Foody, 2004)

The influence of user-defined parameters is also discussed by (Mountrakis, 2011)in their review of support vector machines where they conclude the choice of kernels being a major setback of SVM. This is evidenced by the different results obtained from different kernels. SVM, a non-parametric classifier, is still among the popular classifiers as it gives highly accurate results with limited training samples while generalizing well on new input data. It also works well with higher dimensional data which is a good advantage in remote sensing field as more and higher resolution; multi-spectral data are made available (Prashant K Srivastava, 2012) .

SVM is also widely used to solve multi-class classification problems using one-against-all and one-again stone techniques. While one-against-all compares one class with all other classes taken together, generating n(number of classes) classifiers, one-against-one forms $(n(n-1))/2$ classifiers by forming all two-class classifier pairs from the given input classes (Liangpei Zhang, 2010). A SVM optimally separates the different classes of data by a hyper plane (Kavzoglu, 2009).

1.3 Artificial Neural Network (ANN)

In recent years Artificial Neural Network (ANNs) are being increasingly used as modelling tool in a wide range of application, including among other system identification and control, pattern recognition and data processing. Due to their ability to identify patterns and to detect complex trends by taking in to account nonlinear relationship between complex input and output data, ANNs have also been employed in recent studies for improving our understanding of the ways in which land use and change and evolve (Vafeidis, 2008)

1.4 Random Forest (RF)

This is the basis for the ensemble classifier RF, which combines output from multiple decision trees to decide the label for a new input data based on maximum vote. Random Forest randomly selects a subset of training sample through replacement to build a single tree (Tumer, 1996). Provide that combining output from multiple classifiers for predicting an outcome gives very high classification accuracies.

One of the user defined parameters for RF is the number of trees (Breiman, 2001). That the generalization error always converges as the number of trees increase. Hence there is no issue of over fitting which can also be attributed to Strong Law of Large Numbers (Feller, 1971). Thus for RF, number of trees can be as large as possible but beyond a certain point, additional trees will not help in improving the performance of the classifier. (Belgiu M. D., 2016)Suggest in their review that most papers use 500 numbers of trees for RF classification while there are few other studies which use 5000, 1000 or 100 trees for RF. And among these, 500 are considered as the acceptable optimal value for number of trees. Number of variables required to decide the best split is another user-defined parameter which highly affects the performance of RF. And this is usually set to square root of the number of input variables.

RF has even proven to give good results when used in various applications such as urban landscape classification. Land cover classification on multi-temporal and multi-frequency SAR data and so on (Waske, 2009).

1.5 Change Detection

For detecting and analysing the change on the earth's surface, various techniques are employed. Before studying about various change detection techniques, it is necessary to know about the procedure of change detection. To detect the changes of the surface of the earth, six main steps are important as mentioned by Jensen which is as follows:

- Nature of change detection problems.
- Selection of remotely sensed data.
- Image pre-processing.
- Image processing or classification.
- Selection of change detection algorithm.
- Evaluation of change detection results.

The goal of change detection is to discern those areas on digital images that depict change in the feature of interest between two or more image dates. The reliability of the change detection process may be strongly influenced by various environmental factors that might change between image dates (Jensen R.R., 2007).

3. STUDY AREA

The study area, Colombo district, extends between 79° 48′ and 80° 15′ east and approximately between 6° 45′ and 6° 59′ south, covering an area of 699 km2. The average altitude is approximately 1 m above sea level with a mean annual rainfall of about 2348 mm. Land use within the study area is divided roughly into urban, forest, Grass land, Paddy, Scrub and Plantation. Much of the forests, however, have been removed as a result of both the agricultural and urban expansions.

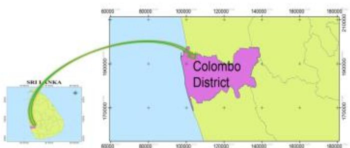

Figure 1 Study area

4. DATA USED

For monitoring LULC change, it is necessary to have at least data of two time periods for comparison. Remote sensing approach usually involves the usage of satellite images of two or multiple dates for quantifying the land use and land cover changes in any area. In this study, the selection of the imageries was made in light of their compatible spatial resolution (30 m). Landsat data archive having images sufficiently consistent with data from the earlier missions allows assessing long-term regional and global LULC change (Irons, 2012).

Table 1 Characteristics of the satellite data used in the present study

Landsat	Sensor	Resolution (M)	Acquisition Date	Bands
5	TM	30	2005-02-13	1, 2, 3, 4
5	TM	30	2008-11-04	1, 2, 3, 4
5	TM	30	2011-11-13	1, 2, 3, 4
8	OLI	30	2015-01-08	2, 3, 4, 5
8	OLI	30	2019-01-03	2, 3, 4, 5

5. METHODOLOGY

Phase I | Phase II

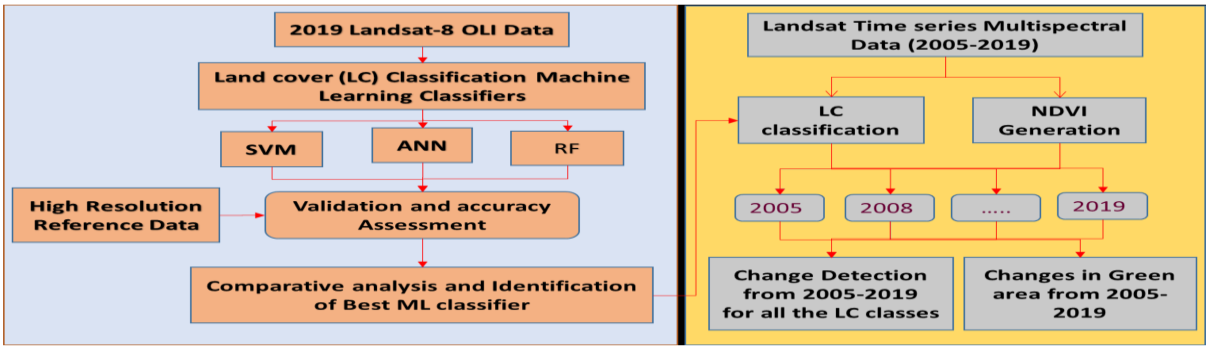

Figure 2 Describing the methodology adopted for the study is given below

1.6 Land Cover Classification Using Machine Learning Classifier

There are total seven classes used for the classification and following three machine learning classifiers are used for identify the best classifier. Land cover classification was performed using machine learning classifier following method are used.

1.7 Support Vector Machine (SVM)

In SVM Kernel type we used radial Basis Function for classification. Following default value we used for better classification result.

Table 2 SVM classification values

Gamma in kernel function	0.143
Penalty parameter	100
Pyramid levels	0
Classification probability threshold	0

1.8 Artificial Neural Network (ANN)

For better result Activation function used as Logistic and these parameters were use as default value for better results.

Table 3 ANN classification values

Training Threshold Contribution	0.9
Training Rate	0.2
Training Momentum	0.9
Training RMS Exit Criteria	0.1
Number of Hidden Layers	1
Number of Training Iterations	1000

1.9 Random Forest (RF)

According to Number of the training samples and Number of trees classification accuracy was change. In we used

Table 4 RF classification values

Number of training samples	500
Number of trees	100

All the classified data of Landsat 2019 accuracy was evaluated for identify the best machine classifier. The best technique which one observed was used for classifying the time series data.

1.10 Accuracy Assessment

Accuracy assessment was performed for identification the best machine learning clarifier. This helps evaluate the performance of various classifiers and also the effect of the underlying training sampling designs. In the study, test samples are randomly chosen from each strata defined over the dataset using Google Earth such that they don't overlap with the existing training data which are created using various sampling designs. Among the various available metrics, Overall Accuracy (OA) is used for assessing all classifiers performance as well as the effect of sampling designs through the performance of one of the classifiers.

1.11 NDVI

NDVI was calculated for 2005, 2008, 2011 Landsat 5 images and 2015 and 2019 Landsat 8 images. Band math was used to calculate NDVI. Output NDVI of each image was used to process change detection. Threshold values were given with respect to the output statistic file of the NDVI.

$$\mathrm{NDVI} = (\mathrm{NDVI} - \mathrm{R})/(\mathrm{NIR} + \mathrm{R})$$

Equation 2 NDVI

1.12 Change Detection

Using the best machine learning classifier all data from 2005 to 2019 was classified and changes were observed from 2005 to 2008, 2008 to 2011, 2011 to 2015, 2015 to 2019 and 2005 to 2019 to see which area are having more changes which class increase or decries to find change detection.

6. RESULTS AND DISCUSSIONS

This chapter include detail discussion of the result obtain as far the discussed of the methodology.

1.13 Land Cover Classification of 2019 LS-8 OLI Data Using Machine Learning Classifier

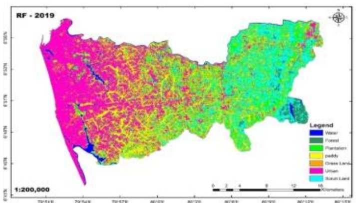

Figure 3 SVM classification 2019

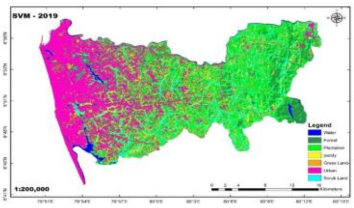

Figure 4 ANN classification 2019

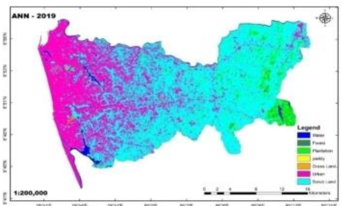

Figure 5 RF classification 2019

Table 5 Accuracy Result of Landsat 8 OLI Image Classification

	SVM	RF	ANN
Overall Accuracy (%)	92.5000	85.2778	71.9403
kappa coefficient	0.9124	0.8280	0.8280

Table 6 User & Producer Accuracy of Different ML Classifier

Class	User accuracy			Producer accuracy		
	SVM	RF	ANN	SVM	RF	ANN
Water	100.00	100.00	100.00	100.00	100.00	100.00
Forest	93.88	100.00	68.97	100.00	100.00	100.00
Plantation	100.00	60.87	100.00	76.42	53.85	4.76
Paddy	90.91	100.00	0	76.92	97.83	0
Grass	95.56	78.13	100.00	93.48	49.02	91.89
Urban	100.00	91.67	100.00	100.00	95.65	100.00
Scrub	74.29	68.42	40.63	100.00	100.00	100.00

SVM Classification technique was observed as best Machine Learning Classifier I n this study as it was providing higher overall accuracy (OA) and kappa coefficient compared to RF and ANN.

1.14 LC Classification of Landsat Multispectral Data from 2005 to 2019

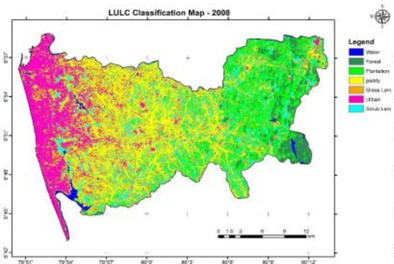

Figure 6 LC clalification2005

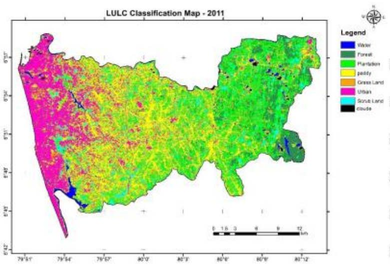

Figure 7 LC classification2008

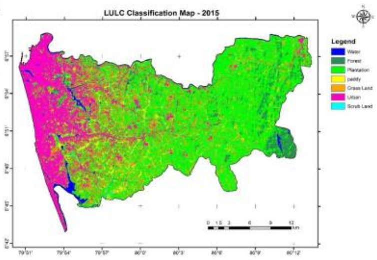

Figure 8 LC classification 2011

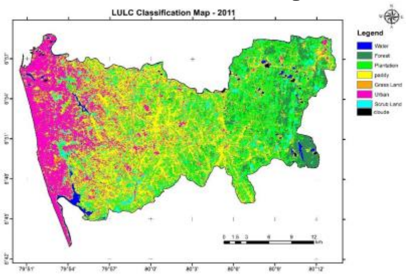

Figure 9 LC classificatio 2015

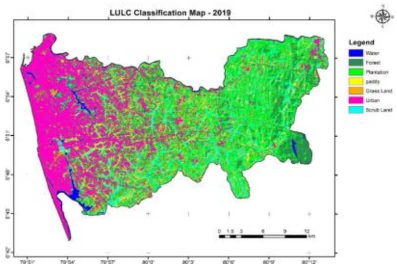

Figure 10 LC classification 2019

Table 7 Accuracy Result of Using SVM Classification Year 2005 - 2019

Year	2005	2008	2011	2015	2019
Overall Accuracy (%)	90.0463	86.4943	81.2698	82.6590	92.5000
kappa coefficient	0.8798	0.8421	0.7810	0.7982	0.9124

Table 8 Land Use Changes in Year 2005 - 2019

AREA(sq km)	2005	2008	2011	2015	2019
Water	12.5478	11.2428	12.8889	11.6874	11.5974
Forest	89.046	44.451	69.3927	50.1579	81.8577
Plantation	100.00	60.87	100.00	76.47	53.85
Paddy	90.91	100.00	0	76.92	97.83
Grass	95.56	78.13	100.00	93.48	49.02
Urban	100.00	91.67	100.00	100.00	95.65
Scrub	74.29	68.42	40.63	100.00	100.00

Area of paddy land cover class was decreases from 2005 to 2019. Urban land cover class was increases at that period. Forest, Grass Land also decreases due to Deforestation and Urbanization.

1.15 NDVI Change Map

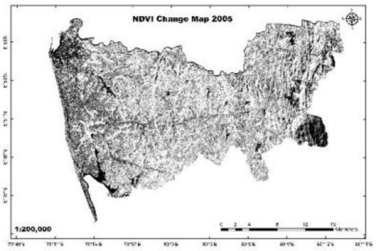

Figure 11 NDVI 2005

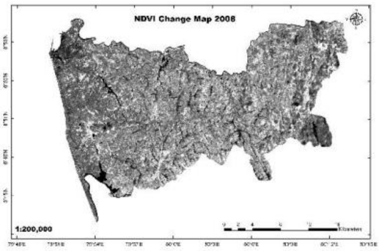

Figure 12 NDVI 2008

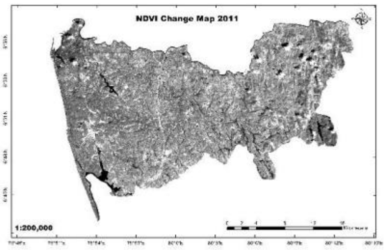

Figure 13 NDVI 2011

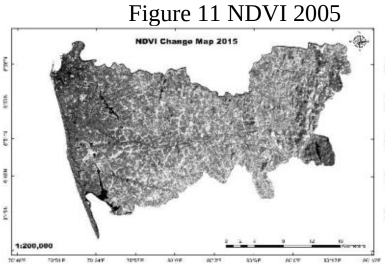

Figure 14 NDVI 2015

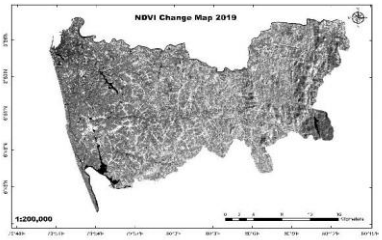

Figure 15 NDVI 2019

1.16 Change Detection Result

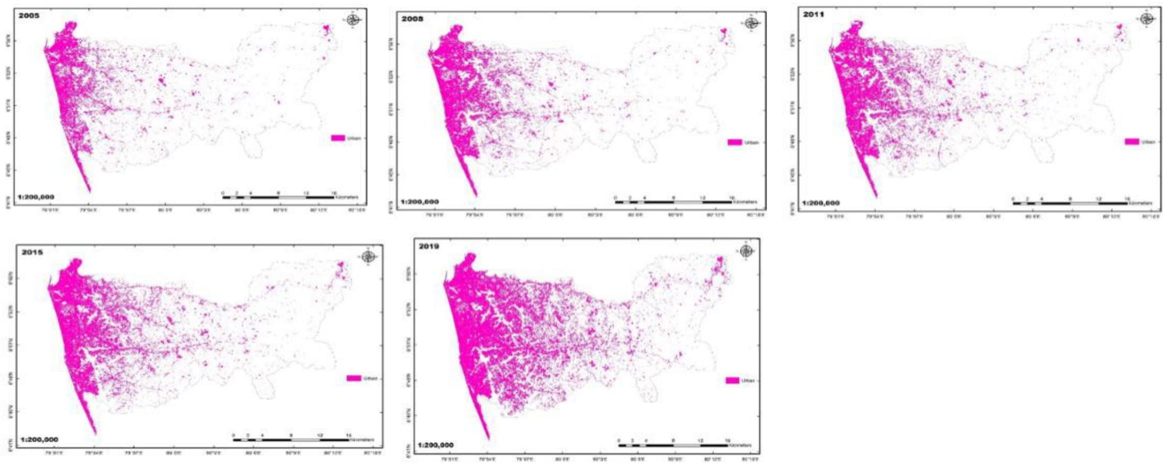

Figure 16 Urban land use pattern changes at different locations in the study area (2005–2008–2011-2015-2019)

1.17 Change Detection Output in Different Stages

Table 9 Change detection percentage of year 2008-2011

	Initial stage - 2008									
	Land Use classes	Water	Forest	Plantation	paddy	Grass Land	Urban	Scrub Land	Row Total	Class Total
Final stage - 2011	Water	82.373	1.737	0.556	0.565	0.175	0.356	0.374	100	100
	Forest	2.233	71.251	21.886	4.037	0.28	0.008	1.339	100	100
	Plantation	1.193	7.769	59.52	13.266	8.135	0.196	47.132	100	100
	paddy	12.152	14.912	11.188	71.058	23.047	8.579	17.075	100	100
	Grass	0.112	1.213	1.56	2.133	32.279	2.356	3.882	100	100
	Urban	0.664	0.476	0.238	5.091	16.298	87.806	0.78	100	100
	Scrub Land	0.752	1.348	4.482	3.654	19.626	0.302	29.193	100	100
	cloude	0.52	1.294	0.569	0.196	0.161	0.398	0.224	100	100
	Class Total	100	100	100	100	100	100	100		
	Class Changes	17.627	28.749	40.48	28.942	67.721	12.194	70.807		
	Image Difference	14.641	56.111	29.38	-2.167	-45.989	8.745	-42.24		

Figure 17 Change Detection 2008 – 2011

Table 10 Change detection percentage of year 2011-2015

	Initial stage - 2011									
	Land Use Classes	Water	Forest	Plantation	paddy	Grass Land	Urban	Scrub Land	Row Total	Class Total
Final stage - 2015	Water	74.757	0.053	0.087	0.584	0.092	0.154	0.494	99.49	100
	Forest	7.737	39.057	2.21	7.355	0.823	0.199	1.29	98.75	100
	Plantation	10.314	57.204	81.629	51.792	14.065	1.701	46.668	99.65	100
	paddy	1.383	0.991	5.846	8.21	2.92	0.123	24.761	99.77	100
	Grass Land	1.341	1.873	6.293	19.699	67.93	11.54	21.186	99.88	100
	Urban	4.092	0.386	0.663	9.671	13.754	86.253	2.263	99.65	100
	Scrub Land	0.377	0.436	3.272	2.691	0.416	0.03	3.337	99.86	100
	Class Total	100	100	100	100	100	100	100		
	Class Changes	25.243	60.943	18.371	91.79	32.07	13.747	96.663		
	Image Difference	-9.322	-27.719	104.006	-80.775	207.802	11.072	-75.338		

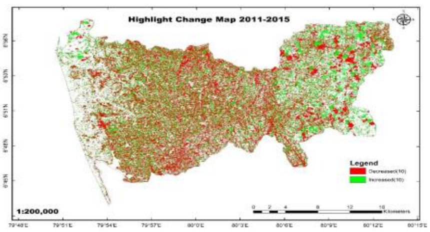

Figure 18 Change Detection 2011- 2015

Table 11 Change detection percentage of year 2015-2019

	Initial stage - 2015									
Final stage - 2019	Land Use Classes	Water	Forest	Plantation	Paddy	Grass Land	Urban	Scrub Land	Row Total	Class Total
	Water	90.205	1.123	0.021	0.128	0.009	0.255	0.092	100	100
	Forest	5.352	76.381	11.992	1.325	1.255	1.494	3.637	100	100
	Plantation	0.508	13.759	49.649	9.35	16.81	1.002	30.037	100	100
	Paddy	0.431	1.809	10.32	36.868	13.107	0.34	25.273	100	100
	Grass Land	0.1	0.486	2.339	2.884	14.489	0.828	1.996	100	100
	Urban	2.672	3.411	7.553	2.101	49.659	95.924	2.668	100	100
	Scrub Land	0.732	3.031	18.126	47.344	4.67	0.158	36.296	100	100
	Class Total	100	100	100	100	100	100	100		
	Class Changes	9.795	23.619	50.351	63.132	85.511	4.076	63.704		
	Image Difference	-0.77	63.2	-39.695	55.446	-75.62	54.896	563.052		

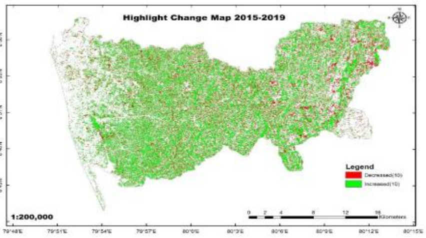

Figure 19 change detection 2015-2019

Table 12 Change detection percentage of year 2005-2019

	Initial stage - 2005									
Final stage - 2019	**Land Use Classes**	Water	Forest	Plantation	Paddy	Grass land	Urban	Scrub Land	Row Total	Class Total
	Forest	12.079	46.158	9.681	7.591	3.2	1.128	4.116	100	100
	Plantation	0.932	39.205	45.74	28.103	17.363	1.188	15.325	100	100
	paddy	0.825	2.337	11.622	11.084	25.433	0.639	32.302	100	100
	Grass Land	0.158	0.705	3.013	3.889	17.524	1.577	6.208	100	100
	Urban	3.622	4.901	4.662	36.582	21.515	94.75	8.789	100	100
	Scrub Land	1.757	6.222	25.19	12.657	14.685	0.336	32.023	100	100
	Class Total	100	100	100	100	100	100	100		
	Class Changes	19.373	53.842	54.26	88.916	82.476	5.25	67.977		
	Image Difference	-7.574	-8.073	40.452	-77.582	-32.386	155.172	391.949		

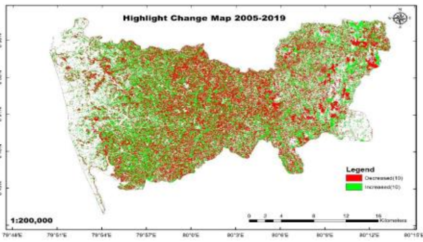

Figure 20 change detection 2005-2019

1.18 Producer Accuracy, User Accuracy Class Wise in Different Years.

Table 13 Producer & User Accuracy in Year 2005

Class	Prod. Acc.%	User Acc.%
Paddy	92.86	88.64
Forest	100	82.61
Plantation	69.23	61.36
Grass Land	45.45	95.24
Water	100	100
Scrub Land	57.14	46.15
Urban	96.49	100

Table 14 Producer & User Accuracy in Year 2008

Class	Prod. Acc.%	User Acc.%
Water	100	100
Grass Land	93.18	97.62
Plantation	43.48	90.91
Paddy	83.33	84.91
Scrub Land	86	59.72
Urban	98.25	100
Forest	97.73	86

Table 15 Producer & User Accuracy in Year 2011

Class	Prod. Acc.%	User Acc.%
Plantation	66.67	57.14
Water	100	100
Scrub Land	65	63.41
Forest	90.24	100
Grass land	97.37	88.1
Paddy	85.71	93.75
Urban	98.75	100

Table 16 Producer & User Accuracy in Year 2015

Class	Prod. Acc.%	User Acc.%
Plantation	82.61	52.78
Forest	100	97.87
paddy	91.84	68.18
Scrub Land	16.36	81.82
Grass Land	95.65	95.65
Urban	100	100
Water	100	100

Table 17 Producer & User Accuracy in Year 2019

Class	Prod. Acc.%	User Acc.%
Water	100	100
Grass Land	93.48	95.56
Plantation	76.47	100
Scrub Land	100	74.29
Urban	100	100
Forest	100	93.88
paddy	76.92	90.91

7. CONCLUSION

This research mainly aims to analyse different machine learning classifiers such as RF, SVM and ANN under the influence of certain factors such as sample size, data dimension and provide a comparative analysis of their sensitivity to other factors. The research also focuses particularly on sampling techniques for training data which requires more attention according to literatures. The fixed number of heterogeneous pixels produced good class-level accuracies. With all three methods favoured at different times, choice of sampling technique to obtain training samples depends on the requirement of a particular study. The effect of sampling can also be associated with the accuracy of reference maps to certain extent. With the absence of ground truth data in most of the studies, dependencies on reference maps are higher on large areas. Hence there is a need for more freely accessible high resolution LULC maps which are regularly updated. Accuracies of LULC maps also depend on the classifiers.

A multi-temporal data is best classified if the dataset includes the variation during the study period. The reference map used in the study contains certain errors and the presence of misclassified training pixels helped observe the sensitivity of classifiers. Tree- based RF classifier were less sensitive to such samples while the kernel-based SVM showed high sensitive to the quality of training samples.

This study concludes that RF, SVM and ANN are all powerful classifiers for LULC classifications. These results also indicate that the choice of classifier depends on the study area, thematic accuracy and quality of training samples and requirement of the map.

8. REFERENCE

Aksoy, S. K. (2005). Learning Bayesian classifiers for scene classification with a visual grammar. *Geoscience and Remote Sensing*.

Belgiu, M. a. (2016). Random Forest in Remote Sensing: A Review of Applications and Future Directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 24-31.

Belgiu, M. D. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 24-31.

Boser, B. G. (1996). A Training Algorithm for Optimal Margin Classifier. *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory*, (p. 5).

Breiman, L. (2001). Random Forests. *Machine Learning*, 5-32.

Cortes, C. a. (1995). *Support-Vector Networks. Machine Learning.* Boston: Kluwer Academic Publishers.

Elis, E. (2011). Anthropegenic transformation of the terrestial biosphere. *Philoshopical transaction. Series A, Mathematical, physical and Engineering Sciences*, 369.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications.* New York: John Wiley and Sons.

Foody, G. M. (2004). Toward intelligent training of supervised image classifications: Directing training data acquisition for SVM classification. *Remote Sensing of Environment*, 107-117.

H S, S. R. (2004). Urban sprawl: Metrics, dynamics and modelling using GIS. *International Journal of Applied Earth Observation and Geoinformation*, 29-39.

HUANG, C. D. (2002). An assessment of support vector machines for land cover classié cation. *International Journal of Remote Sensing*.

Irons, J. D. (2012). The next Landsat satellite: The Landsat Data Continuity Mission. *Remote Sensing of Environment*, 11-21.

Jensen R.R., G. J. (2007). *Remote Sensing Change Detection in Urban Environments.* Berlin: Springer.

Justice, C. T. (2002). Special issue on the Moderate Resolution Imaging Spectroradiometer (MODIS): A new generation of land surface monitoring. *Remote Sensing of Environment*, 1-2.

Kavzoglu, T. C. (2009). A kernel functions analysis for support vector machines for land cover classification. *Int. J. Applied Earth Observation and Geoinformation*, 352-359.

Liangpei Zhang, X. H. (2010). Object-oriented subspace analysis for airborne hyperspectral remote sensing imagery. *ISSN 0925-2312*, 927-936.

Lu, D. a. (2007). A Survey of Image Classification Methods and Techniques for Improving Classification Performance. *International journal of Remote Sensing*, 823-870.

Mas, J. F. (2008). The application of artificial neural networks to the analysis of remotely sensed data. *International Journal of Remote Sensing*, 617-663.

Mountrakis, G. I. (2011). Support vector machines in remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 247-259.

Prashant K Srivastava, D. H.-R. (2012). Selection of classification techniques for land use - land cover change investigation. *Advances in Space Research*, 1250-1265.

Tumer, K. G. (1996). Error Correlation And Error Reduction In Ensemble Classifiers. *Connection Science*.

Vafeidis, A. N. (2008). A New Global Coastal Database for Impact and Vulnerability Analysis to Sea-Level Rise. *Journal of Coastal Research*, 917-924.

Waske, B. B. (2009). Classifier ensembles for land cover mapping using multitemporal SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 450-457.