forests

MDPI

Article

Mapping of *Shorea robusta* Forest Using Time Series MODIS Data

Bhoj Raj Ghimire [1,*], Masahiko Nagai [1,2], Nitin Kumar Tripathi [1], Apichon Witayangkurn [3], Bhogendra Mishara [4] and Nophea Sasaki [5]

1 School of Engineering and Technology, Department of Information and Communication Technology, Asian Institute of Technology, Post Box No 4, Pathumthani 12120, Thailand; nagaim@yamaguchi-u.ac.jp (M.N.); nitinkt@ait.asia (N.K.T.)
2 Department of Sustainable Environmental Engineering, Yamaguchi University, Yamaguchi 755-8611, Japan
3 Center for Spatial Information Science, Tokyo University, Chiba 277-8568, Japan; apichon@ait.asia
4 Faculty of Geo-Information and Earth Observation (ITC), University of Twente, Enschede 7500 AE, The Netherlands; bhogendra@gmail.com
5 School of Environment, Recourses, and Development, Asian Institute of Technology, Post Box No 4, Pathumthani 12120, Thailand; Nopheas@ait.asia
* Correspondence: dr.bghimire@gmail.com; Tel.: +66-94-647-6365 or +977-9841-197-233

Received: 15 August 2017; Accepted: 30 September 2017; Published: 7 October 2017

Abstract: Mapping forest types in a natural heterogeneous forest environment using remote sensing data is a long-standing challenge due to similar spectral reflectance from different tree species and significant time and resources are required for acquiring and processing the remote sensing data. The purpose of this research was to determine the optimum number of remote sensing images and map the Sal forest through the analysis of Vegetation Index (VI) signatures. We analyzed the eight days' composite moderate resolution imaging spectroradiometer (MODIS) time series normalized differential vegetation index (NDVI), and enhanced vegetation index (EVI) for the whole year of 2015. Jeffries-Matusita (J-M) distance was used for the separability index. Performance of EVI and NDVI was tested using random forest (RF) and support vector machine (SVM) classifiers. Boruta algorithm and statistical analysis were performed to identify the optimum set of imageries. We also performed data level five-fold cross validation of the model and field level accuracy assessment of the classification map. The finding confirmed that EVI with SVM (F-score of Sal 0.88) performed better than NDVI with either SVM or RF. The optimum 12 images during growing and post monsoon season significantly decreased processing time (to one-fourth) without much deteriorating accuracy. Accordingly, we were able to map the Sal forest whose area is accounted for about 36% of the 82% forest cover in the study area. The proposed methodology can be extended to produce a temporal forest type classification map in any other location.

Keywords: phenology; Boruta; R; image reduction; random forest; Sal; EVI; NDVI; SVM

1. Introduction

Covering a total area of 6.0 million ha (or 40.4%) of the total land area in Nepal [1], forest has a significant role in social and economic development of Nepal. The contribution of forestry to national GDP is underestimated at 4.4% during 1990–2000. Another study, however suggests that the contribution can be as high as 15.0% [2]. Sal (*Shorea robusta*) is the sole dominant species in Sal forest and it is still one of the most important species in Nepali forests. Sal forest is distributed in varying altitudes from low land to up-hills being named Terai Sal in southern plains and Hill Sal in higher altitudes. Sal is usually the dominant tree in the forests where it occurs. Nationally, the proportion of *Shorea robusta* is highest (15.3%) followed by *Pinus roxburghii* (8.5%). Of the remaining, 60.0% are dominated by other

species found in the mix forest [1]. Above ground carbon stock of Sal were estimated at 80.0 t/ha [3] indicating that Sal forest is important for climate change mitigation and the potential benefits from the REDD+ scheme (Reducing emission from Deforestation and forest Degradation) of the United Nations Convention on Climate Change if the Sal forest is properly managed [4,5]. Currently, Nepal uses a definition of forest as having an area of land at least 0.5 ha, minimum width/length of 20 m, with tree crawn cover of more than 10% and tree heights of 5 m at maturity. Moreover, the definition of "forest type", for the purpose of this study, is adapted from forest resource assessment (FRA) of Nepal [1] and is defined as the forest where at least 60% of the area is occupied by that particular species; and if no species constitutes more than 60% of the basal area, such forests are defined as mix forests. Sal forest provides varied products that include timber, wood for tools and furniture; carvings for historical, religious and architectural structures; utensils, firewood, plates and bowls, gum, green manure, medicines and resin, and livestock browse [6–8]. Moreover, Sal forest is one of the favorite food sources for Asian elephants [9] and a suitable habitat for endangered Bengal tigers [10]. Sal is also important in terms of cultural beliefs. National forest inventories are less often updated working multiple years and spending huge budget. However, our study has potential to find Sal forest at the national, sub national, and local level and produce high resolution Sal maps in an economic and timely manner.

Forest mapping is needed for proper forest management because there are different types of forest and each forest would require different management regime [11]. Mapping Sal forest, is important for better forest policy formulation and management in Nepal. It can also be utilized for carbon monitoring under REDD+ scheme. Information about distribution of Sal forest can be very useful to forest managers for sustainable use of forest resources and environment. Up until now, there are no *Sal* forest maps generated for Nepal by the government nor are any found in the literature. Developing such maps solely by field assessment over a large area is extremely difficult and expensive. So, enhanced methods and the latest technologies are needed to obtain explicit species information from forest.

Analysis of remotely sensed data is a modern approach which has been quite popular in mapping and monitoring forests over the past few decades. With rapid advancement in the remote sensing technology, today we have many satellite systems which acquire earth observation data in high temporal, spectral, and spatial resolution compared to the past. Hyper spectral and Lidar [12–14], laser scanner [15] as well as fine resolution multi spectral data [16] have been used for tree detection and species classification. However, the high cost of acquisition and processing, the limited availability and coverage, and their applicability in operational levels are limited to a handful of organizations with sufficient resources. Hence, for larger area mapping and classification, MODIS and Landsat have been popular because of their free cost, global coverage, and high temporal availability [17–19]. One of the major drawbacks of using single instant imaging for forest type classification, though higher resolution, is that multiple species might have similar spectral signatures. Moreover, successive high-resolution data is not available. Even with Landsat, cloud free data for desired locations throughout the year is not always available. For example, out of 23 images a year (every 16 days), only 8 clearer images were available (with clouds <10%) in our study area. No useful data was acquired for four months, i.e., July, August, September, and November. At times, multiple scenes in the same study area acquired at different dates further complicate the issue as the vegetation structure of the same species changes over time. Hence, MODIS 16 days composite products developed from daily observation would be a reasonable choice for observing vegetation change.

Phenological study of different species in the past have shown differences in the time and duration of phenophases of individual species. Bajpai et al. [20] performed an in-situ experiment to investigate phenophases of two dominant species, *Shorea robusta* and *Ficus hispida*, in deciduous natural forests along the Indo-Nepal border. Data collected from 160 twigs showed clear differences between the species. The variation in vegetation phenological properties opens the possibility of differentiating the specie types using spectral signatures over time. Time series data analysis with vegetation indices is

popular in Land Use Land Cover (LULC) classification as well as agricultural and forestry applications. Yan et al. [21] utilized normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI) time series data to classify vegetation cover types in China. MODIS time series images have been used for particular crop identification [22–25], land cover classification [26] etc.

In the recent past, some studies have been conducted over *Sal* in Nepal, India, and the region. Most of them are related to carbon stock and above ground biomass calculation. A study by Pandey et al. [27] in Chitwan, Nepal shows that carbon density is highest in *Sal* followed by mix forest. Patel et al. [28] estimated biomass of *Shorea robusta* using principal components of vegetation indices from Landsat images. Chitale et al. [29] characterized different Sal species in Northern India based on visual interpretation of size, tone, shape, and texture generated from optical images. However, applicability of remote sensing data for forest type mapping has been the least explored in Nepal [30]. Forest type mapping in Nepal have usually been multi-year projects updated every decade. Reliable methods and economic technology are critically needed.

Maximum likelihood and minimum distance classifiers are quite popular conventional methods being used for classification with multispectral data. Recently, notable attention has been drawn by machine learning algorithms like support vector machines (SVM), random forest (RF) and neural networks, among others [14,31,32]. For supervised techniques, class separability can be improved by calculating metrics like Jeffries-Matusita (J-M) and Transformed Divergence (TD). Based on the score between each pair of classes, they can be merged, further partitioned or deleted. Some researchers prefer using unsupervised techniques like K neighborhood for clustering [26]. Remote sensing images have bigger size compared to ordinary images. Many such images are needed to cover larger area, which demands high processing costs and time. At times, multiple images might carry redundant information which is surely a processing overhead. So, for similar performance, fewer images are preferred. Identification of a minimum number of images that collectively carry maximum information and the least redundancy is not a simple task. However, there are reduction techniques that select the most important images and reduces the dimensionality of the input, thereby optimizing the processing time. Boruta algorithm [33], principal component analysis, and minimum noise fraction (MNF) [12] are very popular in the literature. Reduced images surely achieve gain in processing time; additionally, it sometimes yields even better accuracy.

This study aims to produce *Sal* map in Southern Nepal using the optimum number of vegetation indices (VI) products of MODIS time series. We focus on following two research questions: (1) Can *Sal* forest be mapped in coexisting heterogeneous forest environment using MODIS time series data? And (2) which period of the year provides optimum performance in terms of accuracy and processing cost?

2. Materials and Methods

2.1. Study Area

The study was performed in two districts, Chitwan and Nawalparasi, in Southern Nepal (Figure 1). It is extended between 83°29′10″ and 84°50′30″ East longitude and 27°10′10″ and 27°52′30″ North latitude with a total area of approximately 4400 square kilometers having a population of 1,223,492 (2011). It is one of the resource rich regions of the country, having immense natural and commercial forests and agricultural commodities. The region is mainly divided into three physiographic ranges; the Mahabharata range (middle mountains) in the north, the Siwalik range (first foot hills of massive Himalayas) in the south, and the inner Terai valley (flat plains) between these two ranges ranging from 83 to 2082 m from the mean sea level.

The average annual temperature is 24 °C ranging from 5° to 40° centigrade. The annual precipitation varies from 1584 to 2287 mm, with 1830 mm as the average. The climate is diverse due to the variation in the geography; however, tropical and subtropical are the dominating ones. The major vegetations are the Evergreen and semi-deciduous tropical forests (Dobremez, 1976). There are mainly

three types of forest dominances. Sal (*Shorea robusta*) forest is the first type dominated by Shorea robusta, a high commercial value tree species. It is also one of the protected species of natural forest. We can find larger plots of *Sal* dominant forest in the region. The second type is a hardwood mix forest that is a mix of *Sal* and other species like Asna (*Terminalia tomentosa*), Karma (Adina *cordifolia*), Khair (*Acacia catechu*), Simal (*Bombax ceiba*), etc. The third type is riverine (RVN) forest, which is not only dominated by Sissoo (*Dalbergia sissoo*) but also comprised of other species like Khair (*Acacia catechu*). Some small bushes, shrubs, grassland and softwood mix are also available in certain areas. Agricultural land is the major component of non-forest regions. Rice, wheat, maize and mustard are the main agricultural crops. The study area is home to endangered species, single-horned rhinoceros (*Rhinoceros unicornis*); more than 600 types of flowering plants; above 500 species of birds; 50 species of mammals; and reptiles, each and more than a hundred species of fish [34]. There is also an abundance of tigers and elephants in the region. Despite the great significance of this forest, Chitwan district had undergone massive deforestation up to 23% during the last two decades of 20th century [34].

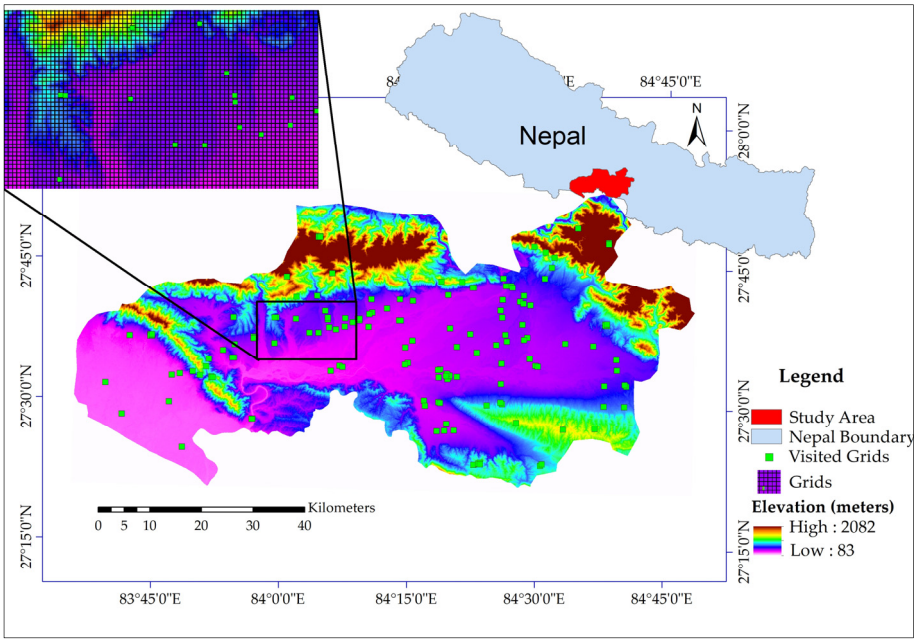


Figure 1. Study area map with altitude information.

2.2. Data

For NDVI and EVI (VI) products, MODIS MOD13Q1 (Terra) and MYD13Q1 (Aqua) were retrieved from the online Data Pool, courtesy of the NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://e4ftl01.cr.usgs.gov. They are 250 m by 250 m products calculated with best available pixel values within a 16-day acquisition period following the strategy of a low cloud, a low view angle and the highest reflectance value. Taking advantage of the phased production between the Terra starting day of year (DOY) 001 and the Aqua starting DOY 009, a total of 46 products of the study area of tile h25v06 were downloaded. This resulted in having 8 days of temporal frequency images which were projected to WGS84 UTM 45 N for the study area. The base map of the study area (1:125,000) was obtained from the Department of Survey, Nepal.

2.3. Overall Methodology

First, VI time-series values from all 46 images in the year are plotted in a graph for each forest type and agricultural area (cover types) to study multi temporal VI signature of different types.

The curves are statistically interpreted to find the similarity and differences of phenophases among them. J-M distance was calculated to verify the goodness of samples. Then, NDVI and EVI time series data is each classified using two supervised machine learning algorithms SVM and RF to know which datasets provide better results and which algorithm is superior to next for *Sal* mapping. Furthermore, Boruta algorithm is applied to reduce the number of images. Additionally, reduced images are also obtained by statistical analysis and visual interpretation of the VI graphs. Performance of different subsets of reduced images is again compared to identify the optimal images. The methodology is represented in Figure 2 and further explained in the sections below.

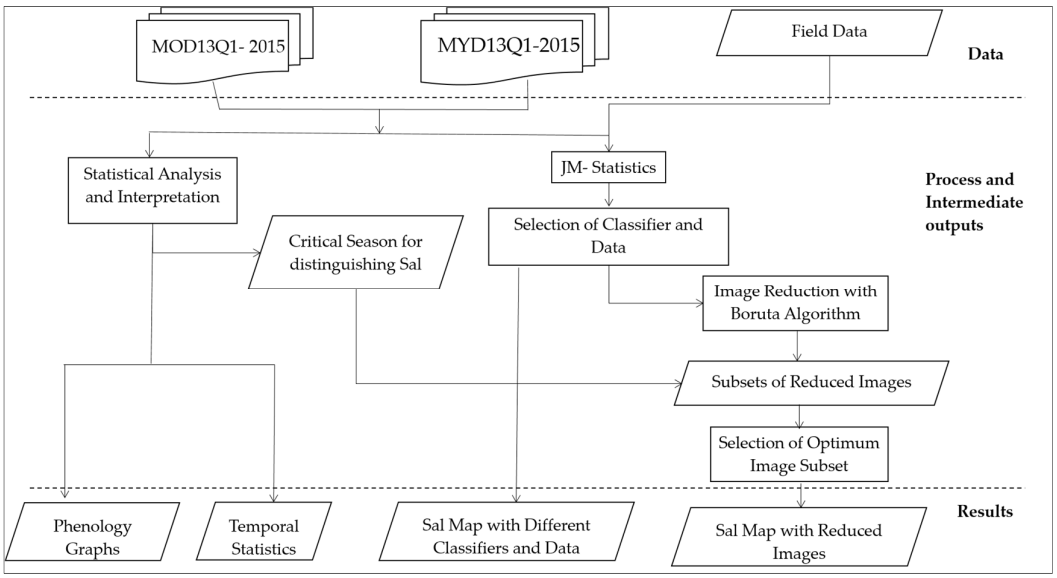


Figure 2. Overall methodology.

2.3.1. Graphical and Statistical Time Series Analysis

Generalized phenology of forest types is helpful; however, species-specific knowledge is preferable because reflectance signature varies over species and time. Hence, it is advantageous to synchronize time series image with the species' phenological cycle. In this study, VI values of four cover types with an 8-day lag, over the period of a year, was plotted in a graph. They were statistically analyzed to gain knowledge on similarity and difference of *Sal* VI signature compared with other types. Two major seasons (time windows) of the year, during which clear distinction was observed, were identified. Moreover, annual, as well as temporal, descriptive statistics namely mean, minimum, maximum, standard deviation were calculated. These are the common metrics calculated for phenological studies of tree species [21].

2.3.2. Region Of Interest (ROI) Separability

Signature separability is a measure of distance between two classes in multiple dimensions. It is calculated for 46 images and sample collection thereby ruling out bad performers. The J-M distance (Equations (1) and (2)) statistics were frequently used by the past researchers to investigate the separability of the classes [35]. This is an extended form of the Bhattacharya distance [36], which calculates the degree of scattering of the two classes. JM distance ranges from 0 to 2; value 2 between a pair of classes implies that these two can be completely distinguished within the datasets used,

whereas the value 0 indicates the classes cannot be separated from each other. Other values in between explain the corresponding degree of distributional distinction between the classes involved.

$$\alpha = \frac{1}{8}(\mu_i - \mu_j)^T \left(\frac{c_i + c_j}{2}\right)^{-1} (\mu_i - \mu_j) + \frac{1}{2} \ln \left(\frac{(c_i + c_j)/2}{\sqrt{|c_i| \times |c_j|}} \right) \quad (1)$$

$$JM_{ij} = \sqrt{2(1 - e^{-\alpha})} \quad (2)$$

where, α = Bhattacharya Distance; i, j = Two signatures (classes) being compared; c_i = covariance matrix of signature i; μ_i = mean vector of signature i; T = transposition function; ln = natural log.

2.3.3. Classification Algorithms

SVM classifier performs supervised classification on a set of images to predict the class of unknown pixel(s). SVM works with the principle where a hyperplane linearly separates the original data, maximizing the margin between the classes. Thus, the goal of SVM is to identify the optimal separating hyperplane using training vectors and kernel functions. It is intended to give the best balance to get the best promotion, by using the optimization methods to solve machine learning problem(s), and SVM is widely used in the regression analysis and statistical classification [10,30]. This method has many advantages in solving nonlinear characteristics, small samples, high dimensional pattern recognition, etc. It is also known as the maximum marginal zone classifier because of the support vector machine classifier's feature ability to minimize the experience error and maximize the geometrical marginal zone. Initially, this algorithm was developed for a binary classifier. However, most of the problems in remote sensing need multiple classifications. Hence, the algorithm is modified to work for multiple classifications of either one-against-all (OAA) or one-against-one (OAO) approach. OAA requires having training data prepared for each class; however, a single set of training data is sufficient for OAO.

Random forest algorithm uses the bagging technique, which combines multiple learning models to yield better classification accuracy. The algorithm starts with creating many random subsets of training samples and forming a decision tree for classification for each of these subsets so that a forest of decision tree is formed. For the prediction, the candidate data is fed into each decision tree and voting is performed among all the decisions to determine the final class. Random decision forests are supposed to correct for the decision trees' habit of over fitting to their training set. A package of random forest in R was used for this study. The functionality is an implementation of the concept proposed in [37].

To exhibit the potency of EVI and NDVI for mapping *Sal* forest and coexisting types, classification of the four classes was performed with SVM and RF, separately; relative performance was compared. The data that yielded better results was chosen. The implementation was performed using free and open source scripting language R. R studio is an easy-to-use development environment for R which can run on any platform. Packages like *gdata*, *rgdal*, and *raster* have made it possible to take advantage of many features of R with remote sensing images. *e1071* and *RandomForest* were the major packages applied for SVM and RF implementation, respectively.

2.3.4. Accuracy Assessment

The study area was divided into a 250 m-by-250 m grid with reference to the MODIS product. 91 grid points ground data of the study area collected during Forest Resource Assessment (FRA) 2011–2016 was obtained from the Department of Forest, Nepal [1]. Another 62 grids were randomly selected to increase the density of plots and to incorporate agriculture and riverine classes which were not obtained from department of forest. Plots in the selected grids were then visited for field verification. Crown coverage of tree species in each grid was visually observed and noted. If a grid had more than 60% *Sal* trees, for example, then it was marked as *Sal* grid and so forth. Garmin GPSMap 64 and Garmin eTrex30 series handheld devices were used to record the geographical coordinates.

For the inaccessible places, expert opinion of the local ranger was considered. Moreover, comparison of texture and structure of the places were made with that of already visited locations using high resolution images from google earth. Altogether, 153 grids were selected, of which 62 were *Sal*, 43 mix, 22 RVN and 26 agriculture. On average, a grid represents 28.7 square kilometer

Overall accuracy (OA) and kappa (k) coefficient were assessed as global measures over the confusion matrix; whereas users' accuracy (precision), producers' accuracy (recall) and the *F-score* measured performance at the individual class. *F-score* is a harmonic balance of both precision and recall. Equations (3)–(5) provide definition of the said measures [38].

$$Precision = \frac{\#True\ Positive}{\#True\ Positive + \#False\ Positive} \quad (3)$$

$$Recall = \frac{\#True\ Positive}{\#True\ Positive + \#False\ Negative} \quad (4)$$

$$F - Score = \frac{2\ Precision * Recall}{(Precision + Recall)} \quad (5)$$

where, *#True positive* is the number of true positives: the number of pixels correctly classified to their true class. *#False positive* is the number of false positives: number of pixels incorrectly classified as belonging to some class C. *#False negative* is the number of false negatives: the number of pixels belonging to some class C but not correctly classified.

2.3.5. Image Reduction

Boruta algorithm is a wrapper based on an all-important feature selection method, which is based on random forest. The algorithm determines if the image is relevant, not relevant or undecided for the classification of objects depending upon the z-score. First, copies of all the variables are added to the original data as shadow attributes and then shuffled for the removal of any correlations. Then, the random forest classifier is computed to get the z score for each of them. The maximum z score among shadow attribute is referenced as the key. Any image whose z-score is higher than the key is selected as important, and any less than the key is non-relevant. Some of them might be undecided if the values are very close to the key. The detail implementation has been explained in [39]. Rasanen et al. [40] has applied the Boruta algorithm to reduce a total of 328 features to only 100 important features for boreal forest habitat type classification and obtained even better results than using full features. In this study, Boruta algorithm was used to determine the subset of important images out of 46 images. Starting with one fourth of the total, i.e., 12 images, then adding the four next important images each time, a total of four subsets with 12 images, 16 images, 20 images and 24 images were selected with respect to the importance score of the individual images. The same process was repeated for NDVI as well. Moreover, one additional subset of images was also determined on the basis of statistical and visual interpretation of the plotted graphs.

3. Results

3.1. Time Series Statistics of VIs

As per the result of mapping EVI values for agricultural area, a clear inverted U shape curve was observed during mid-June to early November, with late August being the highest peak point (Figure 3a). A small bump was observed during mid-March and June, too. In case of RVN, the lowest value was observed in early March increasing slowly and gradually to reach its peak in early September. The growing curve was long and slow. The lowest value of the mix forest was observed during mid-March and highest during July end. For *Sal*, the lowest and highest values of EVI were observed during mid-March and mid-July respectively.

Comparing EVI of Sal forest with that of mixeforest it was worth noting Sal EVI remained higher during the growing season (June–September) while lower post monsoon until the end of the year. Overall EVI signature for all types showed that the values were at their minimum around March, after which these values began increasing until they reached their peak and started declining. The peak value and time-reaching peak were different for different types. For agriculture, there were two peaks; one smaller peak in April end and the next during early September. Riverine reached its peak during early September, whereas mix and Sal had the highest values a bit earlier in July end and mid-July, respectively. All of them were declining post monsoon, and the rate of declination was different for each type. Not much difference was seen in the VI curve for all types during the end of December until February. For the rest of the year, Sal EVI was higher than all the others with an exception of mix forest.

Contrary to EVI values, *Sal* NDVI values (Figure 3b) were always higher than mixforest throughout the year. The difference of NDVI values between *Sal* and mix forest was less in comparison to the difference observed in the case of EVI values during the same time and location. The NDVI values of agricultural area were the least followed by RVN. It was more difficult to delineate the signature between mix and *Sal* forest from NDVI signature in comparison to that of EVI.

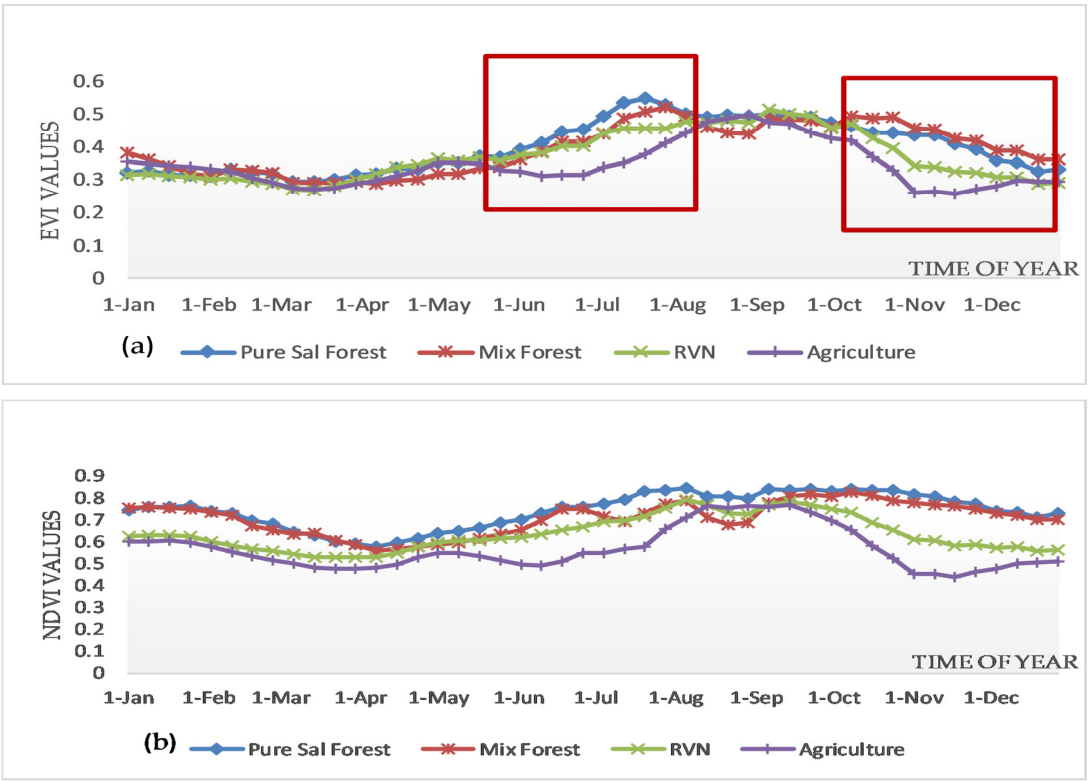


Figure 3. Comparison of (**a**) Enhanced Vegetation Index of Sal with all other types; (**b**) Normalized Differential Vegetation Index of Sal with all other types.

Furthermore, it was observed that *Sal* had the maximum annual values for average, min, max and standard deviation, followed by mix, RVN and agriculture, respectively (Figure 4). Average of Sal (0.397) and mix (0.392) were very close to each other. Standard deviation of mix (0.074) and RVN (0.074) was equal to each other. This provided a general understanding that Sal forest has higher annual values of EVI as compared to the rest of the forest types, but the difference was not wide enough to find a threshold value for the separation.

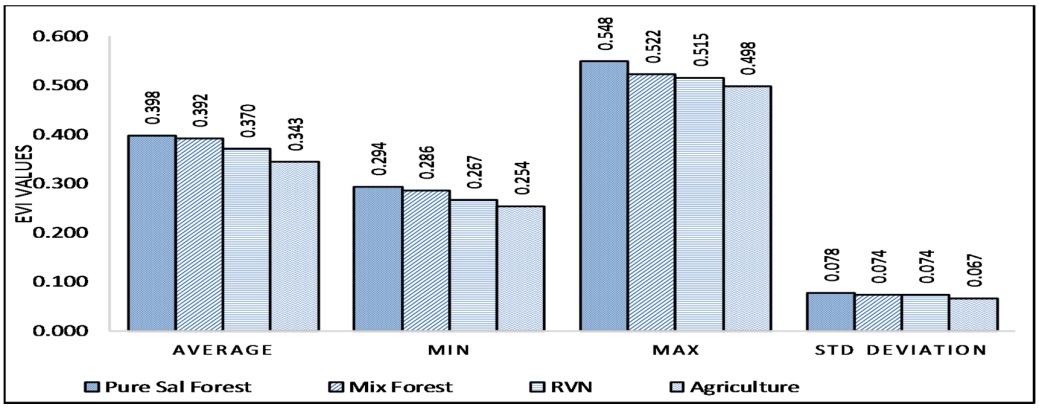


Figure 4. Annual EVI characteristics of different vegetation types.

As shown in plotted graph, it was observed that the EVI of all the types were distinct in two periods shown by the red boxes (Figure 3a); (i) first window (w1) during the growing season between mid-May and mid-August; and (ii) second window (w2) during post monsoon season between early October and mid-December. So, seasonal statistical measures for these periods were computed (Table 1). Sal forest had the highest cumulative values (sum) and mean in the w1, but the mix forest took its place during w2. Slopes during w1 remained all positive while they were all negative during w2. Sal and mix forest had higher variance during w1, while RVN and agriculture had higher values in w2. Similarly, maximum and minimum EVI value of *Sal* was higher than the mix forest during w1, while same measures were lower than that of the mix forest during w2 (Figure 5).

Table 1. Sum, mean, variance and slope comparison of EVI signatures during w1 and w2.

Statistical Features	Sum		Mean		Variance		Slope	
Type	**W1**	**W2**	**W1**	**W2**	**W1**	**W2**	**W1**	**W2**
Pure Sal Forest	5.167	4.046	0.470	0.405	0.004	0.002	0.015	−0.015
Mix Forest	4.821	4.359	0.438	0.436	0.004	0.002	0.017	−0.015
RVN	4.674	3.513	0.425	0.351	0.002	0.003	0.012	−0.018
Agriculture	3.984	3.023	0.362	0.302	0.003	0.003	0.016	−0.011

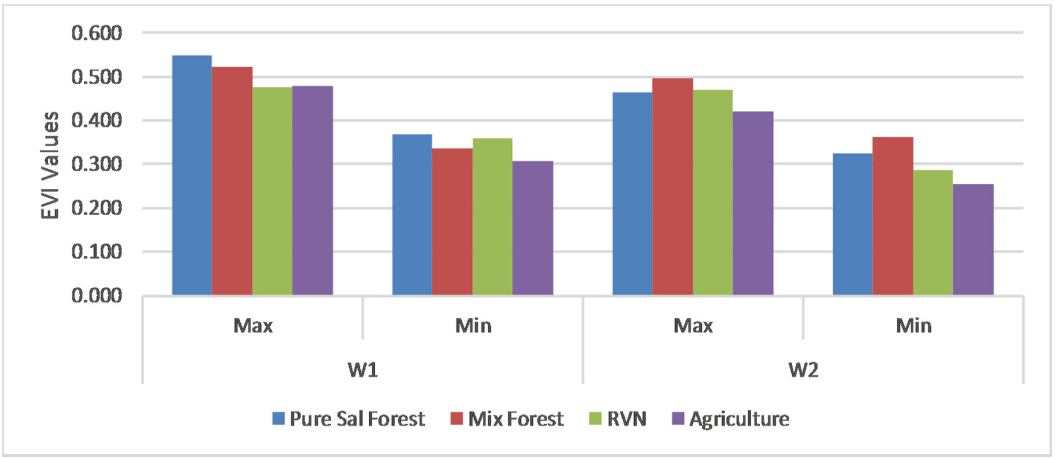


Figure 5. Seasonal minimum maximum.

3.2. Separability Indices

J-M distance was calculated for the selected samples over all 46 NDVI and EVI time series images. The values for JM ranged from 1.82 to 2.0 for different pairs of cases (Table 2), suggesting that the selected sample classes had distinct signatures and were good candidates for using them in a classification. Of the six pairs, the separability between mix forest and *Sal* forest was the poorest followed by RVN and agriculture. The best separation was found between *Sal* forest and agriculture. J-M distance calculated with NDVI produced results little different than that by EVI but the NDVI results were not significantly different.

Table 2. Region Of Interest separability comparison of Enhanced Vegetation Index and Normalized Differential Vegetation Index.

Pair Wise ROI Separability	Jeffries-Matusita Value	
	EVI	NDVI
Mix Forest with *Sal* Forest	1.95	1.82
RVN with Agriculture	1.95	1.96
Mix Forest with RVN	1.96	1.89
Sal Forest with RVN	1.99	1.8
Mix Forest with Agriculture	2	1.99
Sal Forest with Agriculture	2	2

3.3. Selection of Data and Classifier

All available data (46 imageries each for NDVI and EVI) were input to the classifier models for the classification, and the comparison of the results indicated EVI yielded better results classified with either of SVM or RF classifier (Table 3) in comparison to NDVI data sets. It was observed that overall accuracy of the Kappa coefficient and the F-score for *Sal* of NDVI with RF and SVM was (65.4%, 0.53, 0.74) and (68.6%, 0.57, 0.77), respectively (Table 3 (a,b)). However, these indices for the EVI data set, (Table 3 (c,d)) with RF (70.6%, 0.59, 0.82) and SVM (78.4%, 0.69, 0.88), were better. Irrespective of the datasets used, SVM was better for *Sal* mapping in comparison to RF since the above-mentioned parameters were always higher in the case of SVM. F-score of all the types was maximum for SVM-EVI followed by RF-EVI and SVM-NDVI; they were minimum for RF-NDVI. Overall accuracy and Kappa coefficient also followed the same trend. Analysis of the confusion matrix of SVM-EVI showed that most misclassification occurred in the mix forest falsely classified as pure Sal forest leaving users' accuracy (UA) to 82.9% but maintaining producers' accuracy of 93.5%. Forest types classification map using SVM with EVI dataset is presented in Figure 6. According to this study, more than two thirds of the study area is covered by forest and 45% of the forest is covered by *Sal* species. Area of each type as calculated using different datasets and classification algorithm is presented in Table 3 (e).

Table 3. Performance comparison between EVI and NDVI.

(a) RF NDVI							
Classes		**Observed**					**UA (%)**
		Sal	**MIX**	**RVN**	**Agri**	**Total**	
Classified	*Sal*	41	7	1	0	49	83.7
	MIX	10	29	0	4	43	67.4
	RVN	11	7	18	10	46	39.1
	Agri	0	0	3	12	15	80
	Total	62	43	22	26	153	
PA (%)		66.1	67.4	81.8	46.2	**OA**	65.4
F Score		0.74	0.67	0.53	0.59	**K**	0.53

Table 3. *Cont.*

(b) SVM NDVI

Classes		Observed					UA (%)
		Sal	**MIX**	**RVN**	**Agri**	**Total**	
Classified	*Sal*	45	9	1	0	55	81.8
	MIX	8	29	0	4	40	72.5
	RVN	9	3	18	10	40	45
	Agri	0	2	3	13	18	72.2
	Total	62	43	22	26	153	
PA (%)		72.6	67.4	81.8	50	**OA**	68.6
F Score		0.77	0.7	0.58	0.59	**K**	0.57

(c) RF EVI

Classes		Observed					UA (%)
		Sal	**MIX**	**RVN**	**Agri**	**Total**	
Classified	*Sal*	51	10	1	0	62	82.3
	MIX	5	25	0	5	35	71.4
	RVN	6	5	18	7	36	50
	Agri	0	3	3	14	20	70
	Total	62	43	22	26	153	
PA (%)		82.3	58.1	81.8	53.8	**OA**	70.6
F Score		0.82	0.64	0.62	0.61	**K**	0.59

(d) SVM EVI

Classes		Observed					UA (%)
		Sal	**MIX**	**RVN**	**Agri**	**Total**	
Classified	*Sal*	58	10	2	0	70	82.9
	MIX	3	29	0	6	38	76.3
	RVN	1	2	19	6	28	67.9
	Agri	0	2	1	14	17	82.4
	Total	62	43	22	26	153	
PA (%)		93.5	67.4	86.4	53.8	**OA**	78.4
F Score		0.88	0.72	0.76	0.65	**K**	0.69

(e) Comparison of forest types coverage area under different algorithm and datasets

Area	**Area in %**				**Area in Sq Km**			
Method and Data	**Sal**	**MIX**	**RVN**	**Non-Forest**	**Sal**	**MIX**	**RVN**	**Non-Forest**
RF NDVI	23%	29%	27%	21%	1006	1239	1157	888
SVM NDVI	24%	31%	27%	19%	1010	1320	1152	808
RF EVI	30%	24%	25%	21%	1277	1050	1070	892
SVM EVI	36%	29%	17%	18%	1526	1246	729	789

UA = Users' Accuracy (Precision), **PA** = Producers' Accuracy(Recall), **OA** = Overall Accuracy, **K** = Kappa coefficient.

Intel Core i5-6400, 2.7 GHz processor with 8 GB RAM computer was used for data processing in this study. The scripts were written and run in free and open source scripting language R, utilizing different packages written for raster data processing and machine learning. It was noticed that the processing time for all 46 images was always between 96 and 98 h.

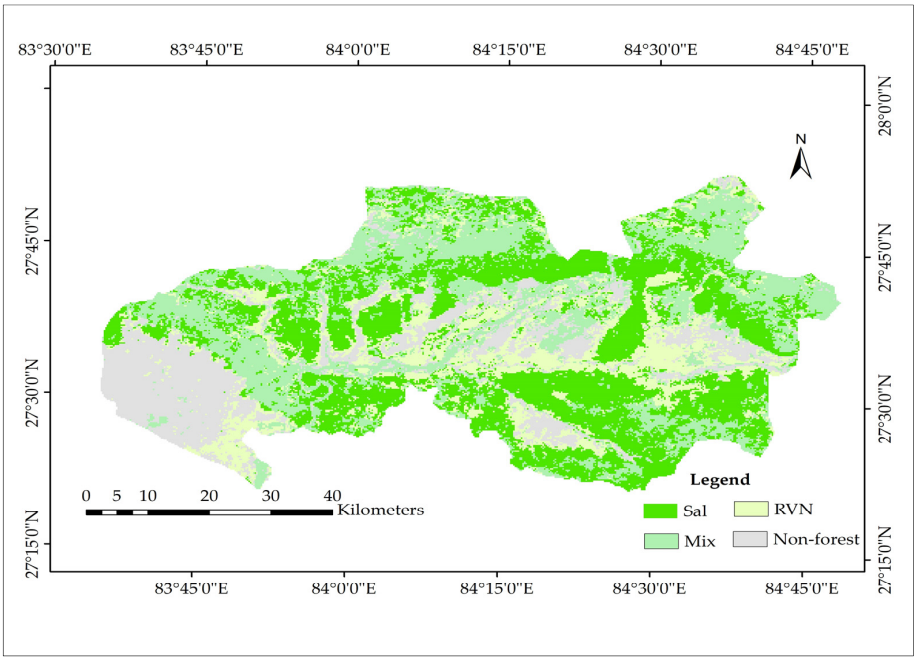


Figure 6. Map showing distribution of Sal and other forest types in the study area using SVM-EVI.

3.4. Reduced Feature Sets

The result obtained from image reduction using random forest based Boruta algorithm evinced that all of the images were important for the classification, with image band X36 being the most important and X3 the least important of all 46 EVI images (Figure 7). X1, X2, ... X46 represent the images of DOY 1, 8, 16, ... 357 of 2015. The most important 24 images for EVI using Boruta, 21 images determined by visual interpretation of the graph plot (Figure 3a) and the most important 24 images for NDVI using Boruta have been presented in Table 4. The image sequence is in order of importance. Interestingly, out of the 12 most important images selected by Boruta (first row of Table 4), 10 images were also present in the subset of 21 images derived from the graphical analysis presented in the third and fourth row of Table 4 and highlighted with bold character. The NDVI had a different hierarchy of importance with image X24 being the most important one.

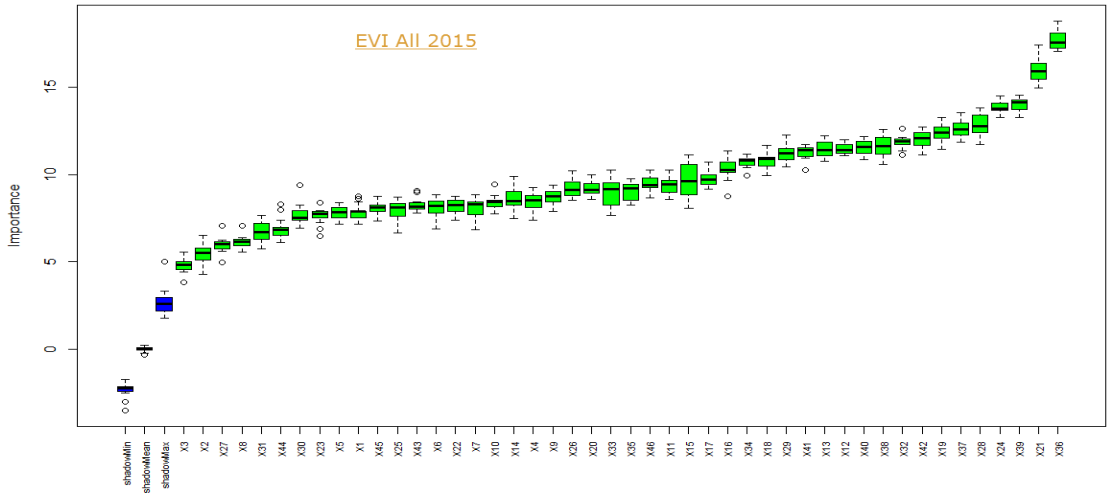


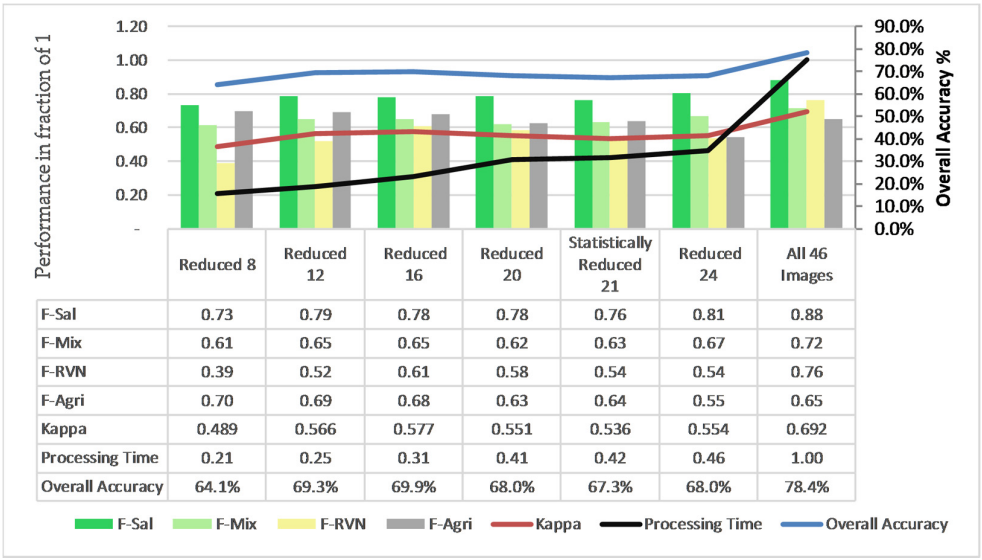
Figure 7. Feature Selection Importance Graph using Boruta Algorithm for EVI Data.

Table 4. Reduced image subsets from Boruta algorithm and statistical analysis.

Subset of 24 images—EVI	**X36**	**X21**	**X39**	**X24**	**X28**	**X37**	**X19**	**X42**	**X32**	**X38**	**X40**	**X12**
(Boruta)	X13	X41	X29	X18	X34	X16	X17	X15	X11	X46	X35	X33
Subset of 21 images—EVI	**X19**	X20	**X21**	X22	X23	**X24**	X25	X26	X27	**X28**	X29	**X36**
(Statistical)	**X37**	**X38**	**X39**	**X40**	X41	**X42**	X43	X44	X45			
Subset of 24 images—NDVI	X24	X29	X19	X36	X37	X42	X1	X41	X39	X28	X38	X30
(Boruta)	X35	X16	X15	X22	X44	X25	X7	X21	X9	X34	X13	X18

3.5. Comparison for Reduced Images

21 images selected during statistical analysis of the time series EVI and reduced most important 8, 12, 16, 20, 24 images (generated by Boruta algorithm) were classified using SVM which was considered superior in this study earlier (Table 3 (d)); where performance indicators were calculated. Overall accuracy of 64% and kappa value 0.49 were obtained with the most important eight images (Figure 8). It was further observed that the most important 12 images provided superior result with overall accuracy (69.3%); Kappa (0.57) and F score of *Sal* forest (0.79),. Six of them (X36–X40 and X42) were from post monsoon season, five (X19, X21, X24, X28, and X32) during monsoon and one (X12) from beginning days of warmer season after cold winter (December to March) in Nepal. After adding the next four important images, overall accuracy (69.9%) and kappa (0.58) were slightly increased; however, F-Sal (0.79) was decreased by 0.01. Obviously, processing time increased by a fraction of 0.06. Even with the most important 20 and 24 images, both the overall accuracy and kappa coefficient decreased. F score of *Sal* provided by the most important 12 images was 90% of the same score provided by all 46 images, but the processing time was, using the same configuration machine, only one fourth. Considering the investment in time and resources to collect, prepare and process the data, these 12 images provided optimum performance among all the reduced images for mapping of Sal forest. This result was further compared with the performance indicators of 21 images (statistically selected), and the results showed that there was only a fine marginal difference in the overall accuracy (2%): Kappa (0.03) and F-Sal (0.022). The classification map with Boruta reduced 12 images provided in Figure 9.



	Reduced 8	Reduced 12	Reduced 16	Reduced 20	Statistically Reduced 21	Reduced 24	All 46 Images
F-Sal	0.73	0.79	0.78	0.78	0.76	0.81	0.88
F-Mix	0.61	0.65	0.65	0.62	0.63	0.67	0.72
F-RVN	0.39	0.52	0.61	0.58	0.54	0.54	0.76
F-Agri	0.70	0.69	0.68	0.63	0.64	0.55	0.65
Kappa	0.489	0.566	0.577	0.551	0.536	0.554	0.692
Processing Time	0.21	0.25	0.31	0.41	0.42	0.46	1.00
Overall Accuracy	64.1%	69.3%	69.9%	68.0%	67.3%	68.0%	78.4%

Figure 8. Performance comparison of reduced images.

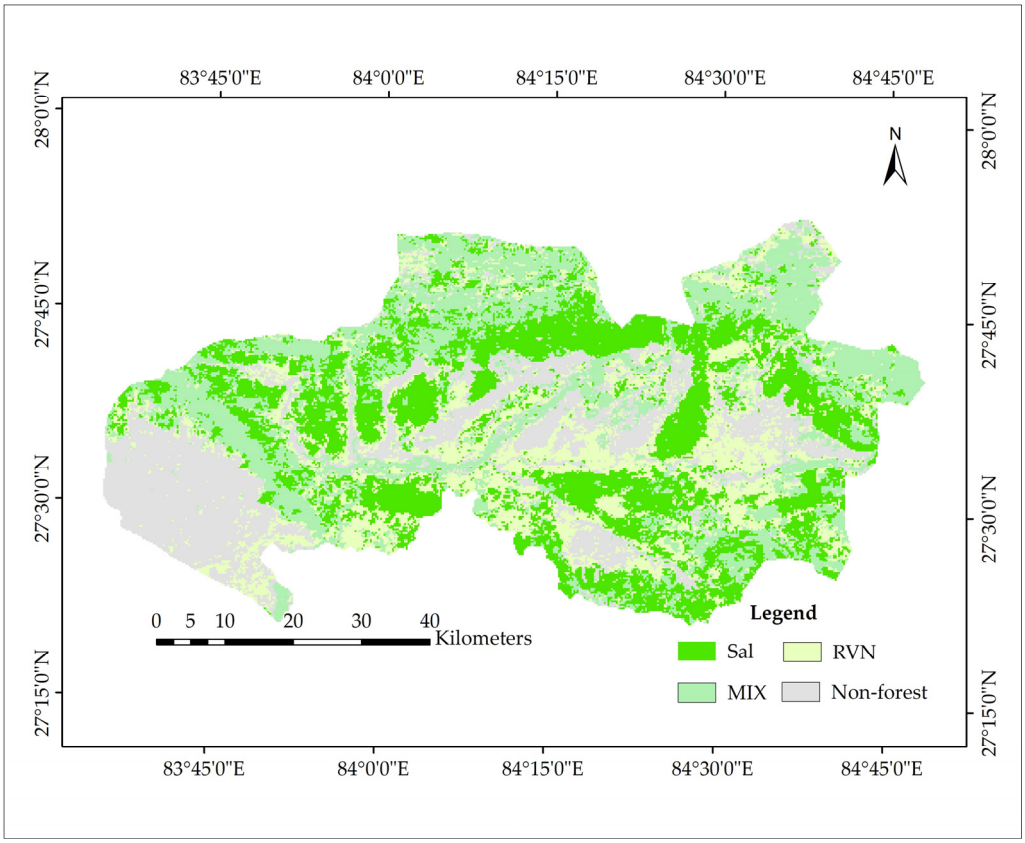


Figure 9. Forest types map with reduced 12 images.

4. Discussion

The agricultural area had shorter and multiple phenological cycles; one during April–June and next one between June–November. The area in the earlier period was for vegetables and short cash crops, like mustard, while the area in the later period had a paddy crop cycle. Pure Sal forest had a cycle starting in March and reached the peak in July. In an earlier in-situ research on a phenological study [20] , it was found that the flowering of *Shorea robusta* started in March, shortly after the dry season, and took 3–4 months to become fully matured in June. The trend is similar; however, there exists some time difference. This gap might have been contributed by other factors—such as temperature, rainfall, soil type, aspect, etc.—which are not considered in this study. Moreover, MODIS products used were of 16 days composite. The EVI values of all the types that started advancing in March are justifiable since the day light starts increasing around this time shortly after the chilling winter in Nepal that runs from December to February. Not much of a difference was seen in the EVI curve for all the types during the end of December until February. This is typically dry season when all deciduous trees either lose their leaves or undergo very inactive photosynthesis. The statistical analysis also revealed that the most significant months for Sal detection were June, July (X19–X29), October and November (X36–X42) for the year 2015 in this studied area. This could differ slightly for some other year and location due to effect of climatic and geological factors on phenology.

Statistical study, as well as machine learning classification algorithms, have suggested EVI to be better than NDVI data. The result of EVI was superior to that of NDVI when tested with SVM as well as RF. For EVI datasets, an overall accuracy and kappa value of 78.4% and 0.69 was observed with SVM and 70.6% and 0.59 with RF. But, OA (68.6%), kappa (0.57) with SVM as well as OA (65.4%), and kappa (0.53) with RF were less for the NDVI data set of the same time series. Even the visual interpretation of the time series graph aligned with this result. The reason for this might be that EVI was

introduced to minimize the effect of aerosol and overcome the saturating tendency of NDVI over high biomass and leaf area index (LAI) which is true for this studied area. According to this study, support vector machine (SVM) was preferred over the random forest (RF) for *Sal* mapping. Nevertheless, both classifiers are widely used in machine learning algorithms, and they have performed superior to each other in different past situations. The accuracy of the results obtained was comparable to that of the past researchers Chitle et al. [29] , Peter Burai et al. [41] , and Mathus Pinheiro et al. [42]. Field observations of many forests with varying species and interaction with local people, as well as forest officers and experts, was helpful in adding implicit knowledge of species distribution as well.

The Boruta algorithm suggested that all the features were relevant for classification. So, the best result can be expected using all the 46 imageries of the year. However, the use of minimum resources without significantly compromising the performance is always justifiable as it always costlier when using a bigger set of input data. Statistical properties of VIs showed that the growing (June–July) and post monsoon (October–November) season was the most critical period for *Sal* mapping. During June and July, the early phase of monsoon season, trees gain new leaves and hence VI values increase. It was found that Sal has higher values than other trees during this season. During post monsoon phase, shrubs and under tree vegetation cover is generally high. However, Sal dominated forest has less undergrowth vegetation. This might be the reason that mix forest has higher VI values compared to that of Sal. This selection of crucial season also coincides with Boruta results as 10 out of the 12 most important images suggested by Boruta were present within the 21 images that were statistically selected.

While the number of input images was reduced, the processing time also decreased. Most importantly, images 8, 12, 16, 20 and 24, suggested by Boruta algorithm, yielded results in 20, 24, 30, 40 and 45 h, respectively. The 12 most important images selected by Boruta algorithm outperformed all the remaining reduced subsets. The statistically reduced 21 images also provided almost equal performance. The F-score for *Sal* with 12 images was 0.79, and that with 21 images was 0.77, which was 90% and 87% of the best result obtained when full images was used. It was observed that even by increasing the number of images to 16, 20 and 24, there was no better result; although processing time increased in proportion to the increased images. There might be duplicate information carried by different images, which would have resulted in giving either a similar performance even by increasing the total number of input images. Utilizing the importance score of the Boruta algorithm, the researcher can enjoy the liberty of choosing different subsets of a varying number of images and determine optimum images, considering the tradeoff between accuracy and time.

Major misclassification occurred between Sal and mix classes, as well as in between Agriculture and RVN. This was constantly indicated by all the results in graphical representation, JM score, and the confusion matrix. In the field, it was observed that the mix forest also had *Sal* species mixed up in many places. Moreover, most species in the mix forest were broad-leaved like Sal. Though *Sal* was found mostly in its homogeneity, covering quite a bigger spatial extent, the case was different for mix and RVN. In certain areas, they were found amid an agricultural field or a river bank or in the middle of a settlement area covering comparatively less area. So, coarser resolution of the MODIS image with the length of 250 m was certainly a limitation. Obtaining an accurate measurement of the location in the deep forest by a handheld GPS device was another challenge and might be one possible source of error. The varying age of trees and the varying canopy density might have contributed to some errors. Geological, geographical and climatic factors, such as soil type, aspect, elevation, rainfall and surface temperature, were not investigated in this research. In further studies, these factors could be considered for the use of enhancing the understanding of species phenology and increasing the accuracy of *Sal* mapping. Since this study has identified major seasons of the year for better image acquisition, the spatial resolution of classification can be improved by using higher resolution satellite images from other sensors like Landsat or Sentinel if cloud free images are available.

5. Conclusions

This study investigated the use of MODIS NDVI and EVI time-series, remote-sensing data for mapping *Sal* in a natural forest of Southern Nepal. The result showed that MODIS time series data has been beneficial to understand the distribution of *Sal* forest found in the heterogeneous forest. This understanding is important to introduce appropriate management interventions. Moreover, EVI data sets were preferable to NDVI, and the support vector machine method of classification was more accurate than that used in the random forest. The 12 most important images, suggested by the Boruta algorithm, yielded good results, gaining 75% in processing time alone. Those images were from growing and post monsoon seasons (June, July, October, and November in this study). It was further verified that merely increasing the number of input images would not guarantee a gain in performance, but doing so would certainly add time and resource overhead. Simple implementation with a straightforward replicable methodology in R scripting language applied in this study can help future researchers to extend the concept in other dominant forest types and locations taking a few samples of each type. Both related organizations and the government could benefit from the study for producing species specific temporal forest type maps. These maps can serve as a good tool for sustainable forest management, change monitoring, and for monitoring carbon stock in the forest.

Acknowledgments: We would like to acknowledge the data and resource support from The Department of Forest, the Nepal government, the staff members and the nature scientists from Chitwan National park. We are thankful for Kitamoto Asanobu from National Institute of Informatics, Japan for his guidance and support. Sincere thanks to Bikash Pathak and Nabaraj Poudel who assisted during entire data collection, despite the vast forest and challenging weather. We would like to provide special thanks to Amrit Kadel and Raul A. Ramirez for their contribution in running through English style and grammar.

Author Contributions: Bhoj Raj Ghimire and Bhogendra Mishara conceived and designed the data and experiments. Bhoj Raj Ghimire, Masahiko Nagai, and Apichon Witayangkurn developed the methodology and performed experiments. Bhoj Raj Ghimire, Nitin Kumar Tripathi, and Nophea Sasaki contributed to the analysis and interpretation of the results. Bhoj Raj Ghimire wrote the original draft and all the authors contributed to review and editing of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ministry of Forests and Soil Conservation. *State of Nepal'S Forests*; Nepal Government: Kathmandu, Nepal, 2015.
2. Ministry of Forests and Soil Conservation (MOFSC). *Nepal Forestry Outlook Study*; Ministry of Forests and Soil Conservation: Kathmandu, Nepal, 2009; p. 83.
3. Ulvdal, P. Stand Dynamics and Carbon Stock in a Sal Dominated Forest in Southern Nepal. Master's Thesis, Swedish University of Agricultural Sciences, Alnarp, Switzerland, 2016.
4. Mbaabu, P.R.; Hussin, Y.A.; Weir, M.; Gilani, H. Quantification of carbon stock to understand two different forest management regimes in Kayar Khola watershed, Chitwan, Nepal. *J. Indian Soc. Remote Sens.* **2014**, *42*, 745–754. [CrossRef]
5. Gilani, H.; Krishna, S.; Murthy, M.S.R.; Koju, U.A.; Uddin, K.; Karky, B. Monitoring the performance of community forestry to achieve redd + goals through geospatial methods. *Int. Arch. Photogr. Remote Sens. Spat. Inf. Sci.* **2014**, *40*, 9–12. [CrossRef]
6. Vedaraman, N.; Puhan, S.; Nagarajan, G.; Ramabrahmam, B.V.; Velappan, K.C. Methyl ester of Sal oil (*Shorea robusta*) as a substitute to diesel fuel-A study on its preparation, performance and emissions in direct injection diesel engine. *Ind. Crops Prod.* **2012**, *36*, 282–288. [CrossRef]
7. Khan, M.Y.; Ali, S.A.; Pundarikakshudu, K. Wound healing activity of extracts derived from *Shorea robusta* resin. *Pharm. Biol.* **2016**, *54*, 542–548. [CrossRef] [PubMed]
8. Islam, M.A.; Quli, S.M.S.; Rai, R.; Singh, P.K. Livelihood promotion through value addition to household traditional sal (*Shorea robusta* Gaertn.) leaf plate making in Jharkhand, India. *Indian J. Nat. Prod. Resour.* **2015**, *6*, 320–325.
9. Koirala, R.K.; Raubenheimer, D.; Aryal, A.; Pathak, M.L.; Ji, W. Feeding preferences of the Asian elephant (*Elephas maximus*) in Nepal. *BMC Ecol.* **2016**, *16*, 54. [CrossRef] [PubMed]

10. Singh, G.; Velmurugan, A.; Dakhate, M.P. Geospatial approach for tiger habitat evaluation and distribution in Corbett Tiger reserve, India. *J. Indian Soc. Remote Sens.* **2009**, *37*, 573–585. [CrossRef]
11. Sasaki, N.; Asner, G.P.; Pan, Y.; Knorr, W.; Durst, P.B.; Ma, H.O.; Abe, I.; Lowe, A.J.; Koh, L.P.; Putz, F.E. Sustainable Management of Tropical Forests Can Reduce Carbon Emissions and Stabilize Timber Production. *Front. Environ. Sci.* **2016**, *4*. [CrossRef]
12. Dian, Y.; Pang, Y.; Dong, Y.; Li, Z. Urban Tree Species Mapping Using Airborne LiDAR and Hyperspectral Data. *J. Indian Soc. Remote Sens.* **2016**, *44*, 595–603. [CrossRef]
13. Alonzo, M.; Bookhagen, B.; Roberts, D.A. Urban tree species mapping using hyperspectral and lidar data fusion. *Remote Sens. Environ.* **2014**, *148*, 70–83. [CrossRef]
14. Zhang, C.; Qiu, F. Mapping individual tree species in an urban forest using airborne lidar data and hyperspectral imagery. *Photogramm. Eng. Remote Sens.* **2012**, *78*, 1079–1087. [CrossRef]
15. Lin, Y.; Herold, M. Tree species classification based on explicit tree structure feature parameters derived from static terrestrial laser scanning data. *Agric. For. Meteorol.* **2016**, *216*, 105–114. [CrossRef]
16. Engler, R.; Waser, L.T.; Zimmermann, N.E.; Schaub, M.; Berdos, S.; Ginzler, C.; Psomas, A. Combining ensemble modeling and remote sensing for mapping individual tree species at high spatial resolution. *For. Ecol. Manag.* **2013**, *310*, 64–73. [CrossRef]
17. Leckie, D.G.; Gougeon, F.; McQueen, R.; Oddleifson, K.; Hughes, N.; Walsworth, N.; Gray, S. Production of a Large-Area Individual Tree Species Map for Forest Inventory in a Complex Forest Setting and Lessons Learned. *Can. J. Remote Sens.* **2017**, *43*, 140–167. [CrossRef]
18. Thompson, S.D.; Nelson, T.A.; White, J.C.; Wulder, M.A. Mapping Dominant Tree Species over Large Forested Areas Using Landsat Best-Available-Pixel Image Composites. *Can. J. Remote Sens.* **2015**, *41*, 203–218. [CrossRef]
19. Fan, H.; Fu, X.; Zhang, Z.; Wu, Q. Phenology-based vegetation index differencing for mapping of rubber plantations using landsat OLI data. *Remote Sens.* **2015**, *7*, 6041–6058. [CrossRef]
20. Bajpai, O.; Kumar, A.; Mishra, A.K.; Sahu, N.; Behera, S.K.; Chaudhary, B.L. Phenological Study of Two Dominant Tree Species in Troppical Moist Deciduous Forest from the Northern India. *Int. J. Bot.* **2012**, *8*, 66–72.
21. Yan, E.; Wang, G.; Lin, H.; Xia, C.; Sun, H. Phenology-based classification of vegetation cover types in Northeast China using MODIS NDVI and EVI time series. *Int. J. Remote Sens.* **2015**, *36*, 489–512. [CrossRef]
22. Zeng, L.; Wardlow, B.D.; Wang, R.; Shan, J.; Tadesse, T.; Hayes, M.J.; Li, D. A hybrid approach for detecting corn and soybean phenology with time-series MODIS data. *Remote Sens. Environ.* **2016**, *181*, 237–250. [CrossRef]
23. Grzegozewski, D.M.; Johann, J.A.; Uribe-opazo, M.A.; Mercante, E.; Coutinho, A.C. Mapping soya bean and corn crops in the State of Paraná, Brazil, using EVI images from the MODIS sensor. *Int. J. Remote Sens.* **2016**, *37*, 1257–1275. [CrossRef]
24. Sun, H.; Xu, A.; Lin, H.; Zhang, L.; Mei, Y. Winter wheat mapping using temporal signatures of MODIS vegetation index data. *Int. J. Remote Sens.* **2012**, *33*, 5026–5042. [CrossRef]
25. Wardlow, B.D.; Egbert, S.L.; Kastens, J.H. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains. *Remote Sens. Environ.* **2007**, *108*, 290–310. [CrossRef]
26. Xue, Z.; Du, P.; Member, S.; Feng, L. Phenology-Driven Land Cover Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 1142–1156. [CrossRef]
27. Pandey, S.S.; Maraseni, T.N.; Cockfield, G. Carbon stock dynamics in different vegetation dominated community forests under REDD+: A case from Nepal. *For. Ecol. Manag.* **2014**, *327*, 40–47. [CrossRef]
28. Patel, N.; Majumdar, A. Biomass estimation of *Shorea robusta* with principal component analysis of satellite data. *J. For. Res.* **2010**, *21*, 469–474. [CrossRef]
29. Chitale, V.S.; Behera, M.D.; Matin, S.; Roy, P.S.; Sinha, V.K. Characterizing *Shorea robusta* communities in the part of Indian Terai landscape. *J. For. Res.* **2014**, *25*, 121–128. [CrossRef]
30. Fassnacht, F.E.; Latifi, H.; Stereńczak, K.; Modzelewska, A.; Lefsky, M.; Waser, L.T.; Straub, C.; Ghosh, A. Review of studies on tree species classification from remotely sensed data. *Remote Sens. Environ.* **2016**, *186*, 64–87. [CrossRef]
31. Adelabu, S.; Mutanga, O.; Adam, E.; Cho, M.A. Exploiting machine learning algorithms for tree species classification in a semiarid woodland using RapidEye image. *J. Appl. Remote Sens.* **2013**, *7*, 073480. [CrossRef]

32. Shang, X.; Chisholm, L.A. Classification of Australian native forest species using hyperspectral remote sensing and machine-learning classification algorithms. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 2481–2489. [CrossRef]
33. Li, X.; Chen, W.; Cheng, X.; Liao, Y.; Chen, G. Comparison and integration of feature reduction methods for land cover classification with RapidEye imagery. *Multimed. Tools Appl.* **2017**, 1–17. [CrossRef]
34. Panta, M.; Kim, K.; Joshi, C. Temporal mapping of deforestation and forest degradation in Nepal: Applications to forest conservation. *For. Ecol. Manag.* **2009**, *256*, 1587–1595. [CrossRef]
35. Liu, Y.; Wu, C.; Peng, D.; Xu, S.; Gonsamo, A.; Jassal, R.S.; Arain, M.A.; Lu, L.; Fang, B.; Chen, J.M. Improved modeling of land surface phenology using MODIS land surface reflectance and temperature at evergreen needleleaf forests of central North America. *Remote Sens. Environ.* **2016**, *176*, 152–162. [CrossRef]
36. Simin, C.; Rongqun, Z.; Wenling, C.; Hui, Y. Band selection of hyperspectral images based on Bhattacharyya distance. *WSEAS Trans. Inf. Sci. Appl.* **2009**, *6*, 1165–1175.
37. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
38. Multi-label, G.; Karalas, K.; Tsagkatakis, G.; Zervakis, M.; Tsakalides, P. Land Classification Using Remotely Sensed Data: Going Multilabel. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 3548–3563.
39. Kursa, M.B.; Rudnicki, W.R. Feature Selection with the Boruta Package. *J. Stat. Softw.* **2010**, *36*, 1–13. [CrossRef]
40. Räsänen, A.; Kuitunen, M.; Tomppo, E.; Lensu, A. Coupling high-resolution satellite imagery with ALS-based canopy height model and digital elevation model in object-based boreal forest habitat type classification. *ISPRS J. Photogramm. Remote Sens.* **2014**, *94*, 169–182. [CrossRef]
41. Burai, P.; Deák, B.; Valkó, O.; Tomor, T. Classification of herbaceous vegetation using airborne hyperspectral imagery. *Remote Sens.* **2015**, *7*, 2046–2066. [CrossRef]
42. Ferreira, M.P.; Zortea, M.; Zanotta, D.C.; Shimabukuro, Y.E.; de Souza Filho, C.R. Mapping tree species in tropical seasonal semi-deciduous forests with hyperspectral and multispectral data. *Remote Sens. Environ.* **2016**, *179*, 66–78. [CrossRef]