The Republic of the Union Myanmar
Ministry of Natural Resources and Environmental Conservation
Forest Department

Analysis of Land Cover Changes and Deforestation Drivers: Illustrated to Explore Potential Restoration Areas in Kawlin District, Sagaing Region

U Chan Myae Aung-2
Staff Officer
Forest Research Institute

February, 2024

Contents

Analysis of Land Cover Changes and Deforestation Drivers: Illustrated to Explore Potential Restoration Areas in Kawlin District, Sagaing Region

U Chan Myae Aung-2 (Staff Officer, Forest Research Institute, Yezin)

Abstract

Forest cover in Myanmar has gradually declined from 58.00% in 1990 to 42.19% in 2020. This study was undertaken in Kawlin District, which exhibited high loss of intact forest during the past decades. Aiming to support forest conservation and restoration strategies, this research focuses on the dynamics of land cover changes and the identification of deforestation drivers in Kawlin District from 2000 to 2021. Land cover changes were examined using Landsat 5-TM and Landsat 8-OLI satellite imagery. The findings indicate a substantial shift in land cover patterns. Densed forest cover declined from 49.33% to 40.50%, while degraded forest areas increased from 25.13% to 30.60%. Concurrently, mining areas expanded from 0.02% to 0.20%, water bodies reduced from 1.08% to 1.02%, and bare land decreased from 2.65% to 1.59%. Agricultural land use rose from 21.32% to 24.94%, and built-up areas increased from 0.49% to 0.68%. To understand the factors influencing deforestation, a binary logistic regression model (BLR) and a random forest (RF) model were employed. The BLR identified accessibility to settlements, roads, lower altitude, and slope as key factors driving deforestation. The RF model, showing higher accuracy (AUC=0.9072) than BLR (AUC=0.881), highlighted altitude, slope, distance from roads, and settlements as significant. Deforestation probability maps generated by these models indicated that 2.62% (12,494.16 ha) and 8.68% (41,449.77 ha) of forest lands have a very high probability of deforestation according to BLR and RF models, respectively, necessitating cautious monitoring and protection. Additionally, land suitability analysis was conducted to assess the potential for teak plantation, a key component for forest restoration in Kawlin District. Results showed that 10.25% (66,566.61 ha) of the area is highly suitable for teak plantations, 5.51% (35,742.69 ha) is moderately suitable, 11.32% (73,494.81 ha) is marginally suitable, and 72.92% (473,474.65 ha) is currently unsuitable. This assessment is crucial for planning effective forest restoration activities in the district. The study provides valuable insights into land cover changes and deforestation drivers, offering a foundation for informed decision-making in forest conservation and restoration initiatives.

Keywords: forest cover changes, Landsat satellite images, binary logistic regression, random forest, teak plantations

စစ်ကိုင်းတိုင်းဒေသကြီး၊ ကောလင်းခရိုင်၏ မြေယာဖုံးလွှမ်းပြောင်းလဲမှုနှင့် သစ်တောပြုန်းတီးမှုကို ဖြစ်ပေါ်စေသော အကြောင်းရင်းများကို ဆန်းစစ်လေ့လာခြင်း (ယင်းပြောင်းလဲမှုများအပေါ် အခြေခံ၍ သစ်တောစိုက်ခင်းများ ထူထောင်ရန် အလားအလာရှိသော မြေနေရာများကို လေ့လာဖော်ထုတ်ခြင်း)

ဦးချမ်းမြေ့အောင်–၂၊ ဦးစီးအရာရှိ၊ သစ်တောသုတေသနဌာန၊ သစ်တောဦးစီးဌာန

စာတမ်းအကျဉ်းချုပ်

မြန်မာနိုင်ငံ၏သစ်တောဖုံးလွှမ်းမှုသည် ၁၉၉၀ ခုနှစ်တွင် ၅ % မှ ၂၀၂၀ ပြည့်နှစ်တွင် ၄၂.၁၉% သို့တဖြည်းဖြည်းကျဆင်းလာခဲ့သည်။ ယခုသုတေသနကို ပြီးခဲ့သော ဆယ်စုနှစ်များအတွင်း သစ်တောပြုန်းတီးမှုမြင့်မားခဲ့သည့် ကောလင်းခရိုင်တွင်ဆောင်ရွက်ခဲ့ပါသည်။ သစ်တောထိန်းသိမ်းရေးနှင့် သစ်တောများပြန်လည်ထူထောင်ရေးကို ပံ့ပိုးကူညီရန် ရည်ရွယ်၍ ဤသုတေသနကို ၂၀၀၀ပြည့်နှစ်မှ ၂၀၂၁ခုနှစ်အထိ ကောလင်းခရိုင်၏ မြေယာဖုံးလွှမ်းပြောင်းလဲမှုနှင့် သစ်တောပြုန်းတီးမှုကိုဖြစ်ပေါ်စေသော အကြောင်းအရင်းများကို ဆန်းစစ်လေ့လာခဲ့ပါသည်။ Landsat 5–TM နှင့် Landsat 8–OLI ဂြိုဟ်တုဓာတ်ပုံများကိုအသုံးပြု၍ မြေယာဖုံးလွှမ်းပြောင်းလဲမှုများကို ဆန်းစစ်ခဲ့ပါသည်။ ယခုသုတေသန၏တွေ့ရှိချက်များအရ ရွက်အုပ်ပိတ်တောဖုံးလွှမ်းမှုမှာ ၄၉.၃၃ % မှ ၄၀.၅ % အထိကျဆင်းခဲ့ပြီး ရွက်အုပ်ပွင့်တောဖုံးလွှမ်းမှုသည် ၂၅.၁၃ % မှ ၃၀.၆ % အထိ တိုးလာခဲ့သည်။ တစ်ဆက်တည်းမှာပင် သတ္တုတူးဖော်သည့်ဧရိယာများမှာ ၀.၀၂ % မှ ၀.၂ % အထိတိုးလာပြီး ရေဧရိယာများသည် ၁.၀၈% မှ ၁.၀၂% အထိလျော့နည်းခဲ့ပါသည်။ သစ်တောသစ်ပင်ပေါက်ရောက်ခြင်းမရှိသည့်မြေ(Bare Land)ဧရိယာများမှာ ၂.၆၅%မှ ၁.၅၉% အထိကျဆင်းခဲ့သည်။ စိုက်ပျိုးမြေဧရိယာများမှာ ၂၁.၃၂ % မှ ၂၄.၉၄ % အထိ တိုးလာပြီး လူနေဧရိယာများမှာ ၀.၄၉ ရာခိုင်နှုန်းမှ ၀.၆၈ ရာခိုင်နှုန်းအထိ တိုးလာခဲ့သည်။ သစ်တောပြုန်းတီးမှုအပေါ်သက်ရောက်မှုရှိသော အကြောင်းရင်းများကိုနားလည်ရန် binary logistic regression model (BLR) နှင့် random forest (RF) model များကို အသုံးပြုခဲ့ပါသည်။ BLR modelအရ လူများအခြေချနေထိုင်မှုနှင့် နီးကပ်သည့်နေရာများ၊ လမ်းနှင့်နီးကပ်သောအကွာအဝေး၊ လျှောစောက်(Slope)နှင့် မြေမျက်နှာသွင်ပြင်အနိမ့်အမြင့်(elevation) နိမ့်သောနေရာများသည် ကောလင်းခရိုင်ရှိ သစ်တောပြုန်းတီးမှုကိုဖြစ်ပေါ်စေရန် အလားအလာများကြောင်း တွေ့ရှိရပါသည်။ ဆန်းစစ်ချက်များအရ random forest (RF) model ၏ တိကျမှန်ကန်မှုသည် ၉၀.၇၂ % နှင့် binary logistic regression model (BLR) ၏ တိကျမှန်ကန်မှုသည် ၈၈.၁ % ရှိသည့် အတွက် random forest (RF) model သည် တိကျမှန်ကန်မှုမြင့်မားကြောင်း လေ့လာတွေ့ရှိရပါသည်။ RF model အရ လူနေဧရိယာနှင့် နီးကပ်သည့်နေရာများ၊ လမ်းနှင့်နီးကပ်သောနေရာ၊ လျှောစောက်(Slope)၊ မြေမျက်နှာသွင်ပြင်အနိမ့်အမြင့် (elevation)နှင့် ချောင်းနှင့်နီးကပ်သည့်နေရာများသည်

သစ်တောပြုန်းတီးမှုကိုဖြစ်ပေါ်စေရန် အလားအလာများကြောင်း တွေ့ရှိရပါသည်။ BLR modelကိုအသုံးပြု၍ ထုတ်လုပ်ခဲ့သည့် သစ်တောပြုန်းတီးမှုဖြစ်နိုင်ခြေမြေပုံတွင် ကောလင်းခရိုင်အတွင်းရှိ သစ်တောမြေများ၏ သစ်တောပြုန်းတီးမှုဖြစ်နိုင်ခြေဧရိယာသည် ၁၂,၄၉၄.၁၆ ဟက်တာ (၂.၆၂%) ကို ပြသခဲ့ပြီး RF model မှ မြေပုံအရ သစ်တောမြေများ၏ ၄၁,၄၄၉.၇၇ ဟက်တာ(၈.၆၈%)သည် သစ်တောပြုန်းတီးမှုဖြစ်နိုင်ခြေ အလွန်မြင့်မားသည်ကိုတွေ့ရှိရပြီး ယင်းဧရိယာများကို အလေးထားစောင့်ကြည့်ထိန်းသိမ်းသင့်ပါသည်။ ထို့အပြင် ကောလင်းခရိုင်၏ သစ်တောပြန်လည်ထူထောင်ရေးတွင် အဓိကကျသည့် အခန်းကဏ္ဍတွင်ပါဝင်သော ကျွန်းစိုက်ခင်းတည်ထောင်ရန် အလားအလာရှိသည့် ဧရိယာများ ဖော်ထုတ်ခြင်းကိုလည်း ဆောင်ရွက်ခဲ့ပါသည်။ ရလဒ်များအရ ၁၀.၂၅% (၆၆,၅၆၆.၆၁ ဟက်တာ)သည် ကျွန်းစိုက်ခင်းများ တည်ထောင်ရန်အတွက် အလွန်သင့်လျော်ကြောင်း (Highly suitable areas)၊ ၅.၅၁ % (၃၅,၇၄၂.၆၉ ဟက်တာ)သည် အတန်အသင့် သင့်လျော်ကြောင်း (Moderately suitable areas)၊ ၁၁.၃၂ % (၇၃,၄၉၄.၈၁ ဟက်တာ)သည် အနည်းငယ်သင့်လျော်ကြောင်း (Less suitable areas)နှင့် ကောလင်း ခရိုင်၏ ၇၂.၆၅ % (၄၇၃,၄၇၄.၆၅ ဟက်တာ)သည် သစ်တောစိုက်ခင်းတည်ထောင်ရန် မသင့်လျော်ကြောင်း (Currently not suitable areas)ဟူ၍ အဆင့်၄ဆင့်သတ်မှတ်၍ မြေပုံထုတ်လုပ်၍ ဖော်ထုတ်ခဲ့ပါသည်။ မြေယာဖုံးလွှမ်းပြောင်းလဲမှုနှင့် သစ်တောပြုန်းတီးမှုကို ဖြစ်ပေါ်စေသည့် အကြောင်းအရင်းများကိုဆန်းစစ်ခဲ့သည့် ယခုသုတေသနသည် သစ်တောများထိန်းသိမ်းရေးနှင့် သစ်တောများပြန်လည်ထူထောင်ရေးလုပ်ငန်းများအတွက် အခြေခံကျသည့် ကိန်းဂဏန်းအချက်အလက်များကို ပံ့ပိုးပေးနိုင်ပါသည်။

Analysis of Land Cover Changes and Deforestation Drivers: Illustrated to Explore Potential Restoration Areas in Kawlin District, Sagaing Region

1. Introduction

Myanmar is one of the Southeast Asian nations that have high forest coverage, with a total geographical area of 676,600 km^2 and an estimated population of 53.37 million people. Myanmar is renowned for its abundant forest resources, which include tropical rain forests and mangroves in the southern plain, dry and deciduous forests in the central dry habitats, and subalpine forests in the northern mountainous region (Bryant, 1997). However, although it still has a relatively high forest cover, Myanmar's forests have been struggling to meet the expanding resource demands associated with population growth and international demand for wood products (Brunner et al., 1998; Laurance, 2007). According to the Global Forest Resources Assessment, Myanmar's forest area declined from 57.9% of the country's area in 1990 to 42.19% of the country's area in 2020 (FAO, 2015). Myanmar has lost up to 407,100 ha of forest lands annually, mainly due to agricultural expansion and illegal logging (FAO, 2015). Between 1990 and 2020, the net yearly change was 1.05% each year, and Myanmar ranked seventh in the world for annual net loss of forest area from 2010 to 2020 (FAO, 2020). Although many forested regions have undergone significant deforestation rates, land-use planning for the forest lands has not taken place over time to align with the changes in forest resources.

Deforestation is a crucial aspect of land use/land cover changes and a factor that sends feedback to the land-use change drivers. The Forest Resources Assessment defines deforestation as 'forest clearance associated with a change in landscape cover from open or closed forest to cover by trees and/or ferns less than 10% and ground coverage by, palm and/or bamboo less than 20%' (FAO, 2020). Myanmar's accessible forests are decreasing spatially and qualitatively at an alarming pace, while forest loss is greatest in low-latitude regions with rapid economic growth, high population, and concentrated arable land (Yang et al., 2019). In addition to several environmental problems and natural disasters, logging, agriculture, infrastructure development, and closeness to human settlements/cities have been major causes of deforestation and degradation(Aung et al., 2020; Lim et al., 2017; Mon et al., 2010). The extent of deforestation and forest degradation caused by logging and the exploitation of forest products has a substantial influence on biodiversity and community livelihoods. Due to the difficulties in law enforcement and governance structure, the Myanmar Forest Department, which is practically responsible for almost all of the country's forest land, has had limited effective authority over logging activities in border regions and conflict zones such as Kachin State (Naing Tun et al., 2021). Identifying the patterns and causes of deforestation is crucial for providing the additional data required to enhance the management methods for efficiently conserving forests in Myanmar.

Assessing land cover changes and identifying the driving factors responsible for these changes enables the development of natural resource management and the prediction of future land cover trajectories (Giri et al., 2003). Spatial and statistical analysis of deforestation based on biophysical factors and their association with land use/land cover changes will improve the knowledge of deforestation and aid in the development of policy responses to deforestation (Lambin, 1994; Mas, 2000). The use of remote sensing and GIS technology is widespread for several natural resources management challenges, like land use/land cover changes and deforestation (Gautam et al., 2003).

The implementation of a set of forest plantation procedures has contributed significantly to the growth of forest cover in Myanmar from 2017 to 2026 (FD, 2017). Severe deforestation previously occurred in Myanmar, and the temporal and spatial characteristics of deforestation, the drivers of deforestation, and the conservation policies have been reported in several past studies. However, there are a limited number of studies examining suitable lands for forest plantation establishment or reforestation based on a long time series of land-use data. Land suitability analysis (LSA) is a technique for identifying a site's inherent and future capabilities and appropriateness for various purposes (Bandyopadhyay et al., 2009). In order to analyze land suitability with more accuracy and adaptability, GIS-based modeling is a useful tool. Previous studies have demonstrated that GIS capacity can be effectively applied in forest restoration activities. Recent studies have stated that GIS-based suitability models are developed for forest restoration sites in Myanmar (Htoo et al., 2020; Kyaw, 2018).

Kawlin District in the Sagaing Region in the upper part of Myanmar is one important forested area. However, logging, mining, and agricultural expansions pose a threat to the region's forests, leading to deforestation and forest degradation. From 2001 to 2021, Katha District (formerly Kawlin district is included as a part of Katha) lost 106 kha of tree cover, equivalent to a 9.4 % decrease in tree cover (Watch, 2020). Myanmar Reforestation and Rehabilitation Program (MRRP) is a national program to implement forest restoration activities all over Myanmar for ten years, from 2017 to 2026. According to this program, the reforestation activities like establishing teak plantations are also implemented in Kawlin district. However, just bringing attention to the situation does not help with the restoration of these important natural resources in the region. Based on the changes in land-cover and the likely causes of these changes, we need to take comprehensive measures to restore the degraded forest areas.

The geospatial analysis combined with satellite remote sensing is a powerful tool that may bridge the knowledge gap in a timely and cost-effective manner, providing a path ahead for forest restoration. Analysis of geographic information systems (GIS) and remote sensing assist in identifying potential locations for forest restoration and provide suggestions for investigating suitable sites for restoration actions. According to the standard operating procedures (SOP) in the forest management in Myanmar, proper site selection is the foundation for establishing prosperous forest plantations. In addition, the knowledge of land-cover changes and the deforestation process of the targeted area is also vital to implement suitable reforestation initiatives as today forests are being increasingly affected by various anthropogenic activities together with the increase of human population. The present study focuses on detecting land cover changes during the previous 21 years (2000-2021) and explores the distinct biophysical and locational factors affecting the possibility of deforestation in the Kawlin district using remote sensing and GIS techniques. While considering the current land-use planning, such as elevation, slope, settlements, distance from roads, and distance from streams, we identify the suitable sites for potential teak plantations in the Kawlin district. We hope this study may fulfil the research gap on the ongoing land use/land cover changes in the district and provide practical recommendations for successful rehabilitation and reforestation initiatives to support the MRRP's efficiency in the Kawlin district.

2. Objectives

The main objective of this study is to explore deforestation drivers and identify suitable forest restoration sites for teak based on the past land cover changes in Kawlin district, Sagaing Region of upper Myanmar. The specific objectives are as follows.

(1) to detect land cover changes between 2000 and 2021 in the study area,

(2) to investigate the biophysical and locational factors affecting deforestation in the study area, and

(3) to generate suitable lands for forest restoration activities especially teak plantations using remote sensing and GIS

3. Materials and Methods

3.1 Study Area

Kawlin District, Sagaing Region is located between the latitudes of 23˚ 30' N and 24 ˚ 21' N and longitudes of 95˚ 01' E and 96˚ 59' E, in the upper part of Myanmar as shown below in Figure 1. The area of the district is about 649,278.76 ha and consists of Kawlin, Pinlebu, and Wun Tho Townships. There are 26 reserved forests with an area of 164, 588 ha, 29 protected public forests with 126,010 ha, and one protected area with 17,025 ha. The total forest area is about 62 % of the total land area of the district, whereas the reserved forests area is about 33% and protected public forests area is about 25% (FD, 2020). Outside of the forest land, it encompasses forest formations around the villages, including secondary forests on fallows and bare/shrub land. The total population is about 330,660, and the local livelihoods depend mainly on agriculture, such as rice and vegetables, mining and exploitation of forest products (FD, 2020).

Mixed deciduous forest dominates in Kawlin District, and the vegetation of the region is dominated by commercially valuable forest tree species such as Teak (Tectona grandis) as major species, along with other hardwood species such as Pyin-Ka-Toe (Xylia dolabriformis), Tamanalan (Dalbergia oliveri), Padauk (Pterocarpus macrocarpus), Kanyin (Dipterocarpus grandiflorus), and In (Dipterocarpus tuberculatus) (FD, 2020). Kyathaung-Wa (Bambusa polymorpha) and Tin-Wa (Cephalostachyum pergracile) are characteristic bamboo species.

The topography of the study area consists of flat terrain in the center part and hilly areas in the west, north, and east with a slope up to 66.5° and an elevation ranging from 109 to 1,689 m above the sea level. Mu River, one tributary of Ayeyarwady River which is the most important River in the country, originates from Banmauk and Pinlebu townships and flows through Kawlin township into Thaphanseik dam (FD, 2020). Along the banks of the Mu River, the soil is composed of alluvial and sandy silt, which is also very fertile, and local people mostly engage in agricultural activities along the riverine areas (FD, 2020). The highest mean annual temperature is 42°C, and the lowest mean annual temperature is 6°C. The average annual rainfall is 57.08 inches, and the average moisture content is 78% all over the year. Kawlin District mainly includes two types of soil, namely red brown forest soils (Rhodic Ferralsol) and yellow brown forest soils (Xanthic Ferralsol) (FD, 2020).

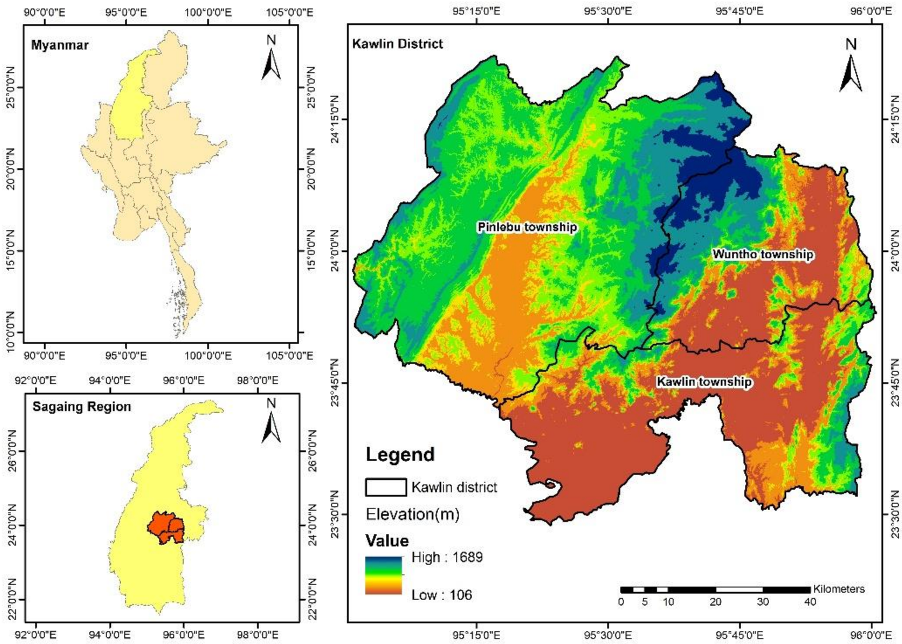


Figure 1: Location of Study Area

The framework of this study is shown in figure 2.

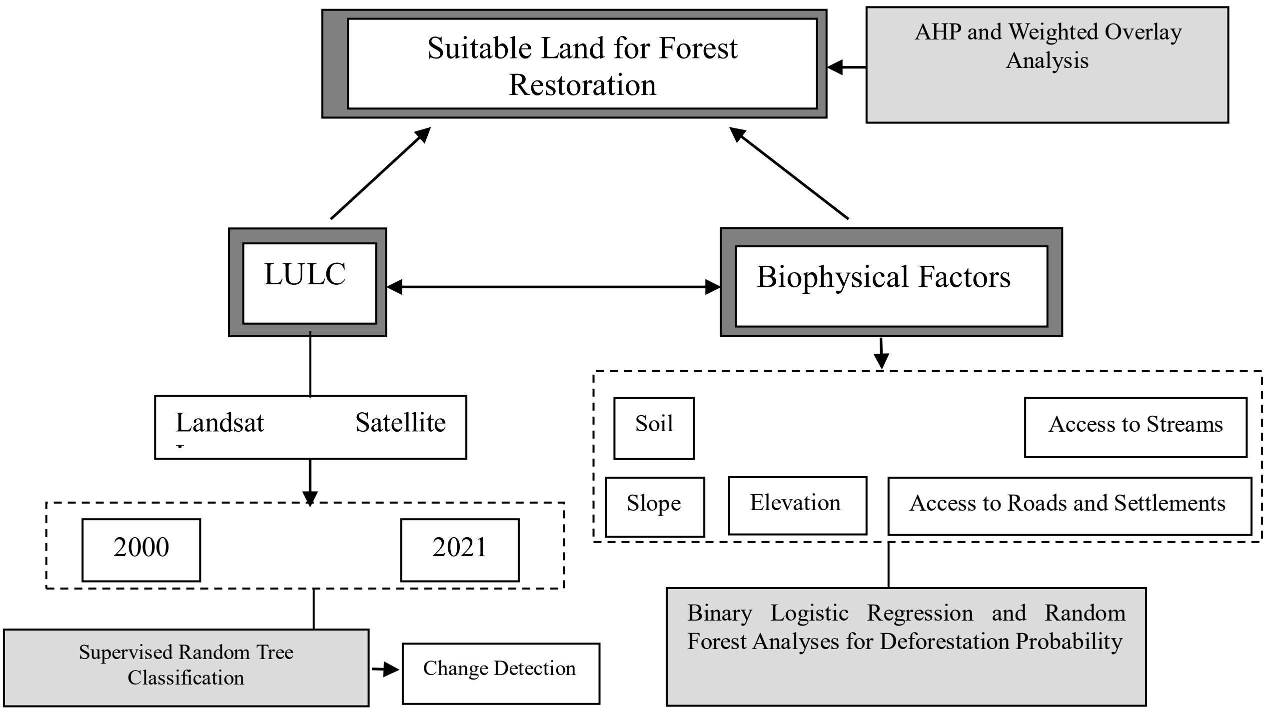


Figure 2: Framework of the Study

In this study, Landsat 5 and Landsat 8 satellite images with less than 10% cloud cover were downloaded from the United States Geological Survey (USGS) website (https://earthexplorer.usgs.gov) to make land cover maps. Landsat 5-TM images for the years of 2000 and Landsat 8 OLI image for 2021 were used for land cover classification and change detection analysis. Since 1972, the Landsat program has provided a continuous data set that is regularly updated. The Landsat images are freely accessible as cloud-free, georeferenced, and orthorectified genuine data with medium resolution. This is very beneficial for researchers in developing countries like Myanmar. The detail information of Landsat images is shown in table 1. For the pre-processing of the satellite images, layer stacking and study area extraction were carried out in Arc GIS (ESRI Inc). The geometric corrections (WGS 1984/UTM Zone 46 N) and standard false color composite were generated before mapping. ArcGIS software version 10.5.1 was used to produce the false color composite layer by stacking of bands 1 to 5 and 7 of Landsat 5-TM images and band 1 to 7 of Landsat 8-OLI image.

Tab.1 Descriptions of Landsat Satellite Images

No.	Year	Satellite Images	Bands	Path & Row	Spatial Resolution	Date of Acquisition
1.	2000	Landsat 5 TM Surface reflectance	Multispectral bands 1-5 and 7	134/43, 134/44	30m	January
2.	2021	Landsat 8 OLI Surface reflectance	Multispectral bands 1-7	134/43, 134/44	30m	January

Seven land cover classes (i.e., forest, degraded forest, mining, agriculture, water, scrub/bare land, and built-up areas) were mapped using supervised classification by implementing random tree classification algorithm in ArcGIS software (Table 2). The random forests algorithm is a machine learning technique that is increasingly being used for image classification. Performance of random trees classification algorithm is much easier to use, more forgiving to overfitting and outliers, and it is suitable for remotely sensed data classification and land use/land cover change mapping (Horning, 2010). Some common applications of random forests classification include land cover and land cover change mapping and cloud and shadow detection. An advantage of random forests is that it is not very sensitive to the parameters used to run it and it is easy to determine which parameters to use (L. Breiman, 2001).

Tab.2 Descriptions of Land Cover Classes

No.	Class Name	Description
1	Forest (Densed Forest)	Areas where trees dominate with dense canopy cover which can be detected in high resolution image
2	Degraded Forest	Areas predominant with small trees and bamboos together with a few large trees
3	Mining	Areas covered with the mining operations
4	Agriculture	Areas covered with paddy fields, crop lands and upland cultivation
5	Water	Streams, ponds, lakes, water bodies
6	Scrub/ Bare Land	It is not recognized as a forest. It includes scrubs, bushes, shrubs, bare lands and young plantations
7	Built-up	Settlement areas, man-made structures, roads and infrastructure

3.2 Accuracy Assessment and Change Detection

Accuracy assessment was carried out by employing stratified random sampling method using 'create accuracy assessment points' tool in ArcGIS. High resolution Google Earth images and field photos, ground truth points obtained from the field survey of Forest Research Institute during 2020, completed by 30 m resolution false color composite images of Landsat 5 and 8 satellite images were used as reference dataset in accuracy assessment of the classifications. The confusion matric and kappa statistics were produced to evaluate the accuracy of 2000 and 2021 classified maps. The confusion matrix compares the relationship between reference data and corresponding classification results.

Change detection analysis was carried out to understand the trends of changes in the past, current, and projections for the future. Analysis of the changes in forest cover at different spatial and temporal scales support appropriate planning and sustainable management of forests(Tripathi et al., 2020). After producing land cover classification maps for 2000 and 2021, change detection analysis was employed to detect changes between 2000 and 2021. The change matrix tables including "from-to" changes on a pixel-by-pixel basis for two periods were produced.

3.3 Factors Affecting Deforestation using Binary Logistic Regression and Random Forest

For preparing deforestation probability map of the study area, binary logistic regression and random forest (RF) have been applied. The sources of data are mentioned in Table 3.

Tab.3 Descriptions of Variables

Variables	Data Format	Source of Data
(1) Altitude	Raster grid	Digital Elevation Model from USGS https://earthexplorer.usgs.gov/
(2) Slope	Raster grid	Digital Elevation Model from USGS https://earthexplorer.usgs.gov/
(3) Soil	Raster grid	Soil Map of Myanmar (Ministry of Agriculture, Agriculture Department, Myanmar)
(4) Distance from roads (Euclidean distance)	Polyline	Digitized from topographic maps of compiled from 1:50,000 aerial photographs (Survey Department, Myanmar) and downloaded from https://www.openstreetmap.org/ (https://download.geofabrik.de/asia/myanmar.html)
(5) Distance from streams (Euclidean distance)	Polyline	Digitized from topographic maps of compiled from 1:50,000 aerial photographs (Survey Department, Myanmar)
(6) Distance from villages (Euclidean distance)	Point	downloaded from MIMU (Myanmar Information Management Unit) (http://www.themimu.info)

(7) 2000 and 2021 land cover maps	Raster grid	Landsat 5 TM and Landsat 8 OLI

3.4 Application of Binary Logistic Regression in Deforestation Probability

Binary logistic regression is a method for analyzing problems in which one or more independent variables determine an outcome, and the outcome is measured with a dichotomous variable (Menard, 2002; Pampel, 2020). The binary logistic regression model (BLR) has been widely used in different fields like determining probable deforestation areas, landslide susceptibility mapping, ground water potential mapping and hazard zonation etc (Ozdemir, 2011). For deforestation probability mapping, statistical analysis, spatial modeling, machine learning, and a common parametric model (Binary logistic regression model) are widely recognized (Aguilar-Amuchastegui et al., 2014; Dlamini, 2016). Deforestation probability mapping is essential for forest management as it measures the sensitivity zones of forest loss, and helps policymakers in managing deforestation problems effectively (Nandy et al., 2007; Srivastava et al., 2002). In this study, binary logistic regression model was conducted for analyzing which biophysical factors determine the probability of deforestation. In the binary logistic regression model, the spatial dataset of forest cover changes during the study period was used as the dependent variable and whereas biophysical characteristics were employed as independent variables to predict the likelihood of deforestation. The binary logistic regression model is appropriate for the application when the dependent variable is a binary value(Atkinson & Massari, 1998).

In each land cover classification map of 2000 and 2021, ‘forest’ and ‘degraded forest’ were combined into a single category ‘forest zones’, whereas ‘mining’, ‘agriculture’, ‘water’, ‘bare/ scrub land’ and ‘settlement’ were categorized together as ‘non-forest zones’. After comparing two maps, a raster dataset with a pixel size of 30m×30m was created in which four land cover classes, including changes from ‘forest to non-forest’, ‘forest to forest’, ‘non-forest to non-forest’, and ‘non-forest to forest’, were conducted. In this study, we focus on the process of deforestation, and thus ‘non-forest to non-forest’ class and ‘non-forest to forest’ (forest increase) were discarded. The final deforestation map consisted of just two classes such as ‘forest to forest’ (no deforestation) and ‘forest to non-forest’ (deforestation) classes. As a dependent variable, forest to non-forest (deforestation) was coded as “1” and forest to forest (no deforestation) was coded as “0”.

As independent variables in the logistic regression model, six biophysical deforestation determining factors, including distance to roads, distance to streams, proximity to settlements, soil types, elevation, and slope, were identified. From a topographic map produced by the Survey Department of Myanmar using 1:50,000 aerial photographs, the roads and streams were manually digitized. Euclidean distance analyses were performed based on digitized road and stream layers to generate the GIS layers for ‘distance from roads and streams’. The villages and town points were downloaded from MIMU (Myanmar Information Management Unit) website (http://www.themimu.info). Proximity to settlements map was generated by using ‘Euclidean distance tool’ in Arc GIS. Soil map of the study area was extracted from the Myanmar soil map which was produced by the Agriculture Department of Myanmar in accordance with the FAO soil classification system. Elevation and slope information were derived from a Digital Elevation Model (DEM) obtained from the USGS website. Distance to streams, distance to roads and proximity to settlements maps were generated using the Euclidean distance algorithm. The Euclidean distance provides the distance between each raster cell and the nearest source, thus it is common to construct raster layers from vector layers. Furthermore, it can be used when data representing the distance from a certain object is needed. The data format and sources of data for each independent variables are described in Table 3 and the raster dataset for each independent variable is shown in figure 3.

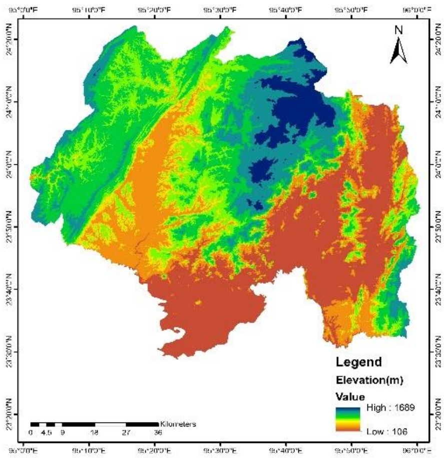
The above-mentioned six independent variables of GIS layers were used as raster datasets with the pixel size of 30m×30m. The deforestation and non-deforestation sampling sites are kept equal, and 1500 for each. Random points were created with the 'create random points' tool in Arc GIS throughout the deforestation and no-deforestation classes. The attribute values of dependent and independent variables were extracted by using the 'extract values to points' tool of Arc map. The collinearity problems of deforestation determining factors were tested through VIF and Tolerance. Then, randomly 80% of the points have been taken for the training dataset and the remaining 20% have been kept for validating the model. The binary logistic regression analysis was done by using R software version 2022.02.0. Binary logistic regression model is as follows.

$$y = a + b_1 x_1 + b_2 x_2 + b_3 x_3 + \cdots b_m x_m \tag{1}$$

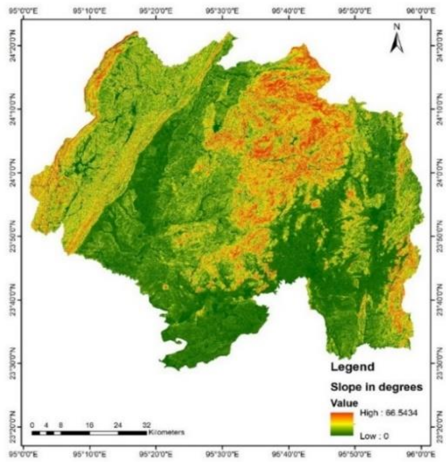
$$y = log_e \left[\frac{P}{1 - P} \right] = logit(P) \tag{2}$$

$$P = \frac{e^y}{1+e^y} \tag{3}$$

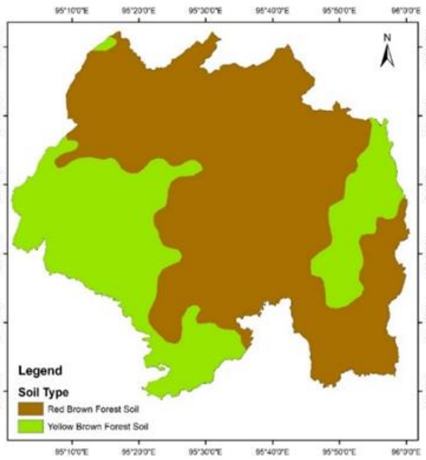
x_1,x_2 and x_3 are independent variables, b_1, $b_2$and b_3 are coefficients, and a stand for constant value, and P means Probability. Y describes $logit(P)$. The resulted deforestation model have been validated by receiver operating characteristics (ROC) curve, efficiency, true skill statistics (TSS) and Kappa.



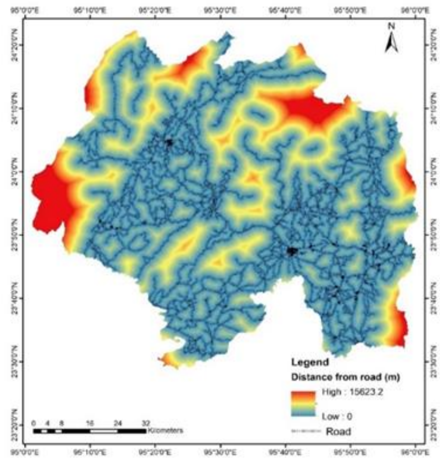
Altitude Map



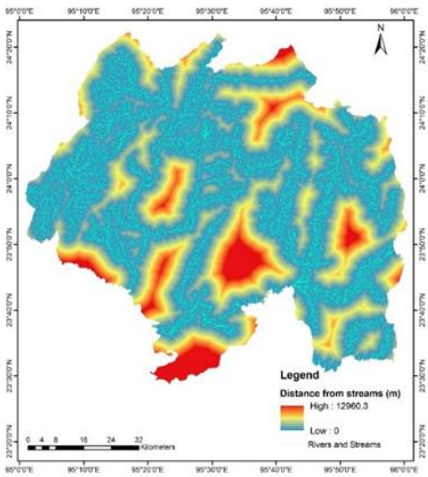
Slope Map



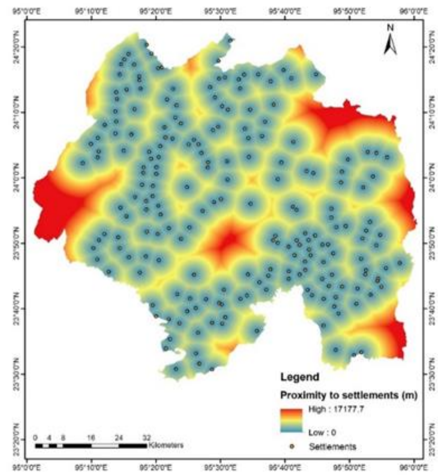
Soil Map



Distance from Road Map



Distance from Streams Map



Distance from Settlements Map

Figure 3: Raster Datasets of Independent Variables

3.5 Application of Random Forest in Predicting Deforestation

Random forest (RF) is a machine learning ensemble method widely used in classification and regression problems (L. J. M. l. Breiman, 2001). Random forest can successfully manage not only numerical and categorical variables but also high data dimensionality and multicollinearity, and is insensitive to overfitting (Belgiu & Drăguţ, 2016). It is an ensemble learning method that needs two parameters for implementation based on the number of trees and variables. In order to build a random forest, the number of trees (n) and predictive variables' split nodes (m) are required. Random forest is a method that constructs multiple decision trees on different samples during training phase and takes their majority voting for classification and average in case of regression as the final decision. In essence, RF is an extension of the baggings technique (Breiman, 1996).

In this study, random forest has been carried out using the "Randomforest" package in R software to generate a deforestation probability map. In the deforestation probability assessment, the dependent factor has been expressed as the binary variable as deforested and non-deforested. 1500 random points were allocated separately for each deforestation and non-deforestation classes (3000 total points that were used in BLR model). Factors affecting deforestation include both the numerical and categorical variables namely, distance to roads, distance to streams, distance to settlements, soil types, elevation and slope. 80% of total random points were randomly selected for training dataset and the remaining 20% were used for model validation. Based on the default for RF, RF has been set to 500 numbers of trees, and the number of variables at each split denoted by the best mtry value has been set to be 3 with the minimum out of bag (OOB) error. The mean decrease in Gini error have been computed, and the results are validated with ROC curve, efficiency, true skill statistics (TSS) and Kappa.

3.6 Validation Methods for Deforestation Probability

The receiver operating characteristics (ROC) is a well-accepted method of validation and it is the graphical representation of the true positive rate (TPR) or sensitivity values against the false positive rate (FPR). The TPR and FPR have been calculated using following equations:

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{4}$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} \tag{5}$$

where, TN stands True negative, TP presents True positive, FN means False negative and FP stands False positive. The area under a receiver operating characteristics (ROC) abbreviated as AUC, is a single scalar value that measures the overall performance of a binary classifier (Hanley & McNeil, 1982). The value of AUC ranges from 0 to 1, where 0.5-0.6 (weak), 0.6-0.7 (moderate), 0.7-0.8 (good), 0.8-0.9 (substantial or very good), and 0.9-1 (excellent) (Yesilnacar & Topal, 2005).

Moreover, in order to identify the robustness of the model, efficiency (E) and true skill statistics (TSS) were used. Efficiency is the proportion of true results (both true positives and true negatives) among the total number of cases examined. Efficiency is calculated as following equations:

$$\mathrm{E} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{6}$$

True skill statistics (TSS) is largely immune to prevalence (Allouche et al., 2006) and TSS is calculated as follows:

$$\mathrm{TSS} = \mathrm{TPR} - \mathrm{FPR} \tag{7}$$

Cohen's kappa index is the most popular statistical measure of accuracy, consistency and efficiency of the predictive models. The kappa statistic ranges from −1 to +1, where +1 indicates perfect agreement and values of zero or less indicate a performance no better than random (Cohen, 1960). Kappa index is calculated as follows:

$$\mathrm{k} = \frac{P_{obs} - P_{exp}}{1 - P_{exp}} \tag{8}$$

P_{obs} and P_{exp} can be calculated using following equations –

$$P_{obs} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{N}} \tag{9}$$

$$P_{exp} = \frac{(\mathrm{TP} + \mathrm{FN})(\mathrm{TP} + \mathrm{FP}) + (\mathrm{FP} + \mathrm{TN})(\mathrm{FN} + \mathrm{TN})}{N^2} \tag{10}$$

Where, "N" is the total number of pixels. The kappa index can be classified as 0-0.2 (low), 0.21-0.4 (moderate), 0.41-0.6 (good), 0.61-0.8 (substantial or very good) and >0.81 (excellent) accuracy (Landis & Koch, 1977).

3.7 Suitability Analysis for Potential Teak Plantation Areas

Teak (Tectona grandis) is one of the most important tropical timber species found in semi-evergreen forests, mixed deciduous forests, and dipterocarp forests (J. Blaser, 2011). Natural teak forests can be found only to four countries of Southeast Asia: India, Myanmar, Thailand and Laos (Keiding et al., 1986). The estimated natural teak forests were about 29 million ha, almost half of it grow in Myanmar (W. Kollert, 2012). Natural teak forests have been declined in all countries due to unsustainable logging. According to FAO, natural teak forests substantially declined in Myanmar (1.1 million ha), India (2.1 million ha) and Lao PDR (68.5 thousand ha) (W. Kollert, 2012). In Myanmar, teak and other commercially valuable hardwoods have been selectively logged for economic contribution of the country since historical times, leading to deforestation and forest degradation with increasing human growth (Naing Tun et al., 2021). Legal but unmonitored logging of teak and other hardwoods for decades and the supply of illegal timber market has contributed significantly to Myanmar's deforestation and forest degradation (Johnson, 2010). As a result, timber production in Myanmar has gradually decreased in both quantity and quality, while natural teak forests are rapidly exploited (Springate-Baginski et al., 2014). Therefore, selecting suitable sites is necessary to grow prosperous teak plantations, and it helps to achieve one of the restoration targets of the Myanmar Forest Department to extend its plantation areas nationwide.

Suitable site selection is the primary factor for any valuable or commercial tree plantation and significantly affects the sustainable timber production and management system. In order to assess suitable sites for teak plantations, Myanmar Forest Department's criteria for plantation establishment were used. To better understand environmental requirements of teak, the literature on teak was reviewed. The process of land suitability analysis needs to consider current and previous vegetation coverage and the change conditions within the corresponding timeframe. The suitability analysis includes five main steps, (1) GIS data conversion and projection, (2) reclassification of selected variables, (3) assigning ranking and weighting scores of variables, (4) spatial analyses, and (5) site suitability mapping. The suitable areas for potential teak plantations were computed using the weighted overlay procedure in ArcGIS. Six parameters such as slope, land cover, distance from streams and rivers, distance from roads, elevation, and distance from settlements were selected for identifying site suitability.

3.8 Standardization of Selected Parameters Maps

All of the selected slope, land cover, distance from streams and rivers, distance from roads, elevation, and distance from settlements raster maps were converted into uniform units (30 m× 30 m) cell size maps because Weighted Overlay Method require to be converted into same units.

3.9 Reclassification of Criteria

For land suitability identification, all parameters were reclassified based on the importance of each of the criteria. All parameters were reclassified in four different categories in this study. The sub-criterion will be classified into 1- 10 different ranking scales where 10 is the highest significance and 1 is the least significant. When giving different ranking scales, it is applied from the combination of expert opinions and considerable experiences of local forest officers from three townships of Kawlin district who are actually working on teak plantations establishment.

Criteria for Suitability Analysis of Teak Plantations in Kawlin District

(a) Slope

Slope value ranges from 0 to 66.54 degrees in the study area. It was observed that more than half of the study area (4468.6638 km2 or 68.81% of the study area) has slope degree of less than 10 degrees which is easy to cultivate teak plantations typically. One of the significant limiting factors for teak plantations is slope because the lower the slope, the higher the moisture content of the slope is (Greulich, 1999). Therefore, lower slope areas are more suitable for plantations. The Standard Operating Procedures (SOP) from Myanmar Forest Department recommended that the areas with slopes more than 30 degrees should not be selected for site selection of teak plantations (FD, 2016).

(b) Land Cover

Land cover data were produced using satellite remote sensing from Landsat 5 TM and Landsat 8 OLI. Landsat 5 TM was used to classify the land cover of study area in 2000 and Landsat 8 OLI was used for 2021. All landcover maps were created using a Random tree classifier in Arc GIS. Identifying land cover change areas is the key component in determining suitable plantation sites. Change analysis of land cover maps of 2000 and 2021 are implemented using the "Combine" tool in Arc GIS. This tool produces a table for each combination of input values indicating the changes. The initial land cover map consists of 7 land classes. Change from Densed Forest/ Degraded Forest to other land classes are recorded as deforestation. Change to Scrub Land and Degraded Forest (bamboo forest) are suitable for prioritized plantations areas declared by Standard Operating Procedures (SOP) from Myanmar Forest Department. All changes to these other land classes (Mining, Agriculture, Water, Built-up) are ignored for suitable sites because they are challenging to restore. Then, the change map is converted using the 'Reclassify' tool to generate the four categories of land cover (Forest, Degraded Forest, Scrub Land, and Other Land Class). The reclassified land cover map is used for suitability analysis.

(c) Distance from Streams and Rivers

The streams and rivers were digitized from topographic maps compiled from 1:50,000 aerial photographs (Survey Department, Myanmar). Distance from streams and rivers was generated using "Euclidean distance" tool in Arc GIS. Typically, the Myanmar Forest Department selects the plantation sites near streams because nurseries are established near plantation areas, and water resources are essential services for young seedlings and the nursery for plantations (FD, 2016). The nursery should be near the plantation sites to reduce transportation and labor costs for seedlings to be planted.

(d) Distance from Roads

Road networks are digitized from topographic maps compiled from 1:50,000 aerial photographs (Survey Department, Myanmar) and downloaded from Open Street Map (https://download.geofabrik.de/asia/myanmar.html). The accessibility of roads was considered for growing the forest plantations because road accessibility affects plantation management expenses and outcomes over the long run. From a practical aspect, local forest staffs conduct plantation establishment during the rainy season; the distance they could travel from roadways was also a key component. When the plantation sites are too far from the road, the potential extension, maintenance and management for plantations might be difficult. Based on the experiences of local forest officials, 0-700 m was the best choice for frequent access to the planting site.

(e)Elevation

The elevation of study area ranges from 106 to 1689 m. Elevation is also extracted from Digital Elevation Model from USGS website. Elevation is also one of the limiting factors for teak because teak cannot grow naturally in all elevation gradients. In Myanmar, teak is not found naturally in areas of more than 914 m above sea level and the most suitable range is between 400 and 700 m (Htoo et al., 2020; Myint, 2012). Therefore, the elevation limit for teak is approximately 900m.

(f) Distance from Settlements

Settlements (village) units are downloaded from MIMU (Myanmar Information Management Unit) website (http://www.themimu.info) and verified with Google Earth Pro. Distance from settlements map was produced using Euclidean distance tool in Arc GIS. In Myanmar, teak plantations were established using the Taungya method. The taungya is a system that gives the right to villagers and sometimes forest plantation workers to cultivate agricultural crops during the early phases of forest plantation establishment (Chamshama et al., 1992). With the Taungya method, local indigenous people mainly contributed to establishing Forest Department's teak plantations. Therefore, plantation sites should not be very far from settlements because enough labor from near villages is needed for the entire plantation establishment process.

(g) AHP

Analytical Hierarchy Process (AHP) is one of the most effective multiple decision-making methods (Thomas & Doherty, 1980). It is designed to use several criteria to determine complex decisions. In the AHP process, pairwise comparison judgments have to be implemented by the decision-maker in order to develop overall priorities for ranking the alternatives. The AHP process sets criteria or sub-criteria to make a hierarchical structure by giving the weight of each criterion an incomplete decision-making process (Kiker et al., 2005). The weight value analyses the relative significance of individual criteria and is chosen intentionally. The following steps were used in order to assign weights of the factors using the AHP technique: (1) defining the goal (potential teak plantations) (2) deciding the factors for teak plantations and scaled weights of each factor of Saaty scale from 1 to 9 (Table 4); (3) pairwise comparison matrix establishment based on the relative scale weights of the selected factors (4) obtaining overall weights to the factors and calculating consistency ratio. For calculating the pairwise comparison matrix, a scale of 1-9 is used where 9 indicates extreme importance and 1 indicates equal importance, which is shown in Table 4.

Tab.4 Fundamental Scale for Pairwise Comparison Matrix (Saaty, 1980)

Intensity of importance rank	Definition	Explanation
1	Equal importance	Two criteria contribute equally to the objective
3	Moderate importance	Experience and judgement moderately favor one criterion over another
5	Strong importance	Experience and judgement strongly favor one criterion over another
7	Very strong importance	A criterion is favored very strongly

		over another; its dominance is demonstrated in practice
9	Extreme importance	The evidence favoring one criterion over another is of the highest possible order of affirmation
2,4,6,8	Intermediate values between the two adjacent importance or judgements	When adjustment is needed
Reciprocals	If criteria i has one of the above numbers designated to it when compared with criteria j, then j has the reciprocal value when compared with i.	

Pairwise Comparison Matrix and Consistency Ratio (CR)

According to AHP techniques, the pairwise comparison matrix was formed to determine the weights of the parameters. Pairwise matrix was calculated by the following expression.

$$\begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \tag{11}$$

The sum of each column in pairwise matrix was described as follows.

$$C_{ij} = \sum_{i=1}^{n} C_{ij} \tag{12}$$

Each element in the matrix was divided by its column total in order to generate a normalized pairwise matrix using the following equation.

$$X_{ij} = \frac{C_{ij}}{\sum_{i=1}^{n} C_{ij}} = \begin{bmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{bmatrix} \tag{13}$$

Then, the following expression denotes that the sum of the normalized column of matrix was divided by the number of criteria used (n) to generate weighted matrix.

$$W_{ij} = \frac{\sum_{j=1}^{n} X_{ij}}{n} = \begin{bmatrix} W_{11} \\ W_{21} \\ W_{31} \end{bmatrix} \tag{14}$$

After calculating the weighted matrix, the eigenvector (λ_{max}) was calculated using Saaty's method (Saaty, 1980). Firstly, the pairwise matrix was multiplied by the weights vector in order to get the consistency vector (Cv).

$$\begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} * \begin{bmatrix} W_{11} \\ W_{21} \\ W_{31} \end{bmatrix} = \begin{bmatrix} Cv_{11} \\ Cv_{21} \\ Cv_{31} \end{bmatrix} \tag{15}$$

Then, we divide the weighted sum vector with criterion weight to accomplish consistency vector (Cv) using the following expressions.

$$\begin{aligned} Cv_{11} &= \frac{1}{W_{11}}[C_{11}W_{11} + C_{12}W_{21} + C_{13}W_{31}] \\ Cv_{21} &= \frac{1}{W_{21}}[C_{21}W_{11} + C_{22}W_{21} + C_{23}W_{31}] \end{aligned} \tag{16}$$

$$Cv_{31} = \frac{1}{W_{31}}[C_{31}W_{11} + C_{32}W_{21} + C_{33}W_{31}]$$

The eigenvector (λ_{max}) was calculated by averaging the value of consistency vector (Cv).

$$\lambda_{max} = \sum_{i=1}^{n} Cv_{ij} \tag{17}$$

Consistency index (*CI*) was calculated in order to measure the deviation $\lambda_{max}$from n i.e., a measure of inconsistency.

$$CI = \frac{\lambda_{max} - n}{n} \tag{18}$$

Consistency Ratio (*CR*) was calculated to detect logical inconsistency of the judgments and possible errors.

$$CR = \frac{CI}{RI} \tag{19}$$

Saaty (1980) suggested that if the value of CR > 0.1, the weight values of the pairwise comparison matrix are inconsistent, and AHP may not give meaningful results. CR value up to 0.1 is acceptable for results (Table 5).

Tab.5 Random inconsistency indices (*RI*) for n=10

n	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49

Weighted Overlay Analysis

The weighted overlay is one method used for modeling suitability. In the ArcGIS environment, each raster layer must be assigned a weight value to individual criteria. The cell values of the selected raster are converted into a common scale. Raster layers are overlaid, multiplying each raster cell's suitability value by its layer weight and totaling the values to derive a suitability value.

Land Suitability Analysis

Suitability analysis for potential forest plantation areas is calculated as follows.

$$LS = \sum_{i=1}^{n} W_n C_i \tag{20}$$

LS = total land suitability score
W_n = weight of land suitability criteria
C_i = the criteria (i) that was reclassified

The land suitability for plantation sites categorization was based on the FAO framework of four suitability classes: (1) highly suitable, (2) moderately suitable, (3) marginally suitable, and (4) currently not suitable.

4. Results

4.1 Accuracy Assessment Results of Land Cover Classification

The confusion matrices and kappa statistics of 2000, 2011 and 2021 land cover classification maps were shown in Table 6, 7 and 8. High resolution Google Earth images and field photos, ground truth points obtained from the field survey of Forest Research Institute during 2020, completed by 30 m resolution false color composite images of Landsat 5 and 8 satellite images were used as reference dataset in accuracy assessment of the classifications (figure 4). The overall accuracy for 2000 and 2021 classification maps were 89.81%, and 91.79% and kappa coefficient were 84.64% and 87.91% respectively.

Tab.6 Confusion Matrix for 2000 Classification Map

Class value	Forest	Degraded forest	Mining	Agriculture	Water	Scrub/ bare land	Built-up	Total	U_ accuracy (%)	Kappa (%)
Forest	689	39	0	0	0	7	0	735	93.74	0
Degraded forest	29	315	0	14	0	23	0	381	82.68	0
Mining	0	0	9	0	0	1	0	10	90.00	0
Agriculture	0	15	0	287	0	16	0	318	90.25	0
Water	0	0	0	0	13	4	0	17	76.47	0
Scrub/ bare land	0	2	0	2	0	37	0	41	90.24	0
Built-up	0	0	0	1	0	1	8	10	80	0
Total	718	371	9	304	13	89	8	1512	0.00	0
P_accuracy (%)	95.96	84.91	1	94.41	1	41.57	1	0	89.81	0
Kappa	0	0	0	0	0	0	0	0	0	84.64

Tab.7 Confusion Matrix for 2021 Classification Map

Class value	Forest	Degraded forest	Mining	Agriculture	Water	Scrub/ bare land	Built-up	Total	U_ accuracy (%)	Kappa (%)
Forest	581	21	0	0	0	2	0	604	96.19	0
Degraded forest	43	411	0	5	0	4	0	463	88.77	0
Mining	0	1	8	1	0	0	0	10	80	0
Agriculture	1	9	1	342	0	29	0	382	89.53	0
Water	1	0	0	1	14	0	0	16	87.5	0
Scrub/ bare land	0	3	0	1	0	21	0	25	84	0
Built-up	0	1	0	0	0	2	14	17	82.35	0
Total	626	446	9	350	14	58	14	1517	0	0
P_accuracy (%)	92.81	92.15	88.89	97.71	1	36.21	1	0	91.69	0
Kappa	0	0	0	0	0	0	0	0	0	87.91

Class Name Field Photo Google Earth Image Satellite Image

Forest

Degraded Forest

Mining

Agriculture

Water

Scrub/ Bare Land

Built-up

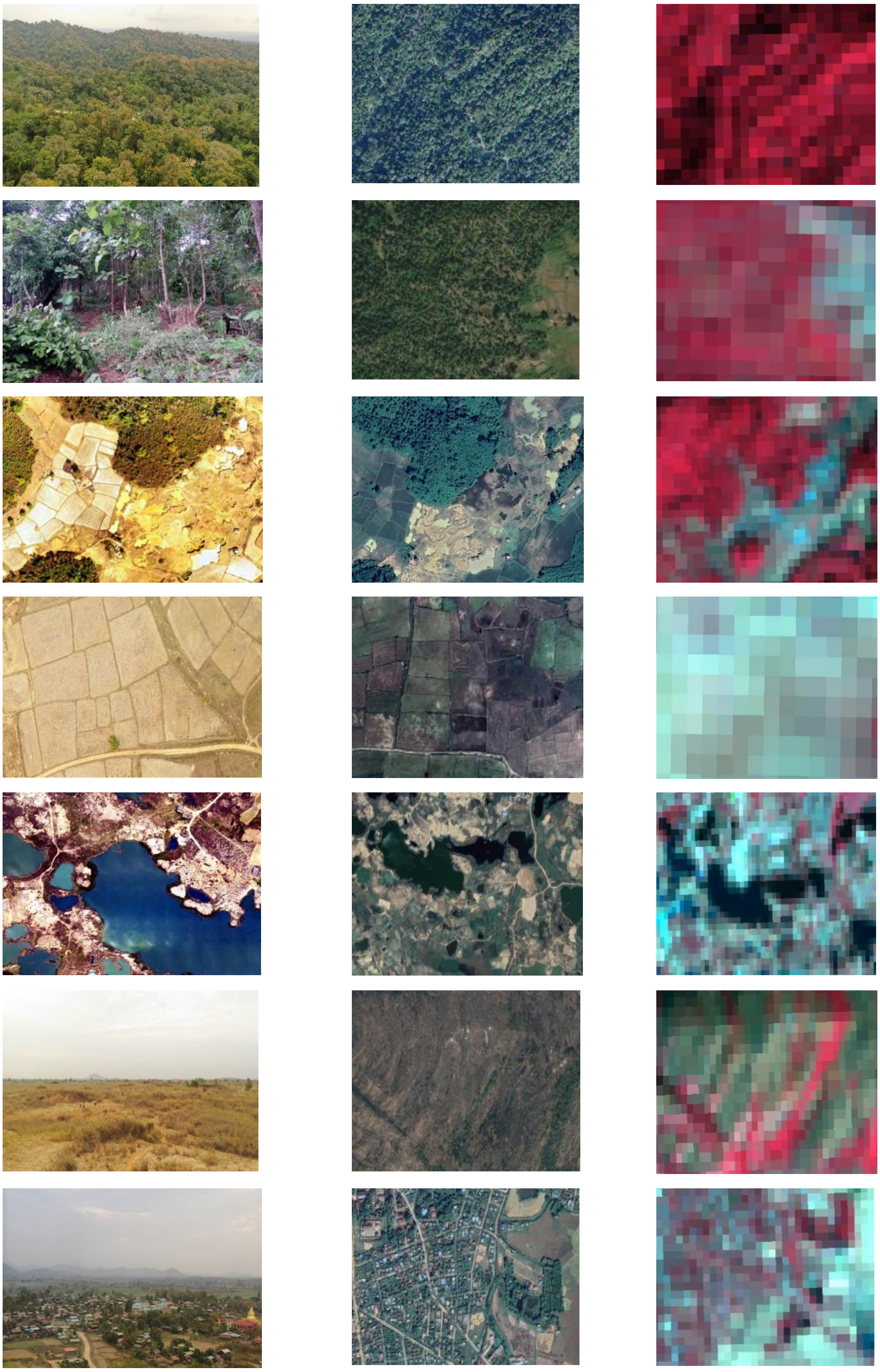


Figure4: Description of land cover classes with field photos, google earth and satellite images

4.2 Land Cover Changes between 2000 and 2021

Three land cover maps for 2000 and 2021 are shown in Figure 5 and 6.

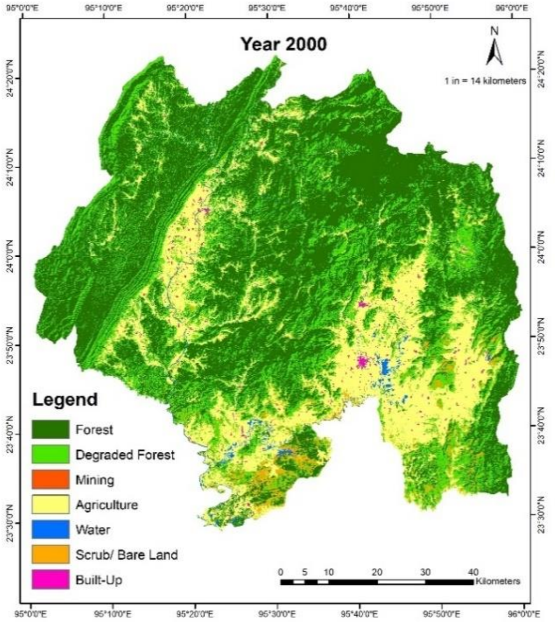


Figure 5: Land Cover Classification Map for 2000

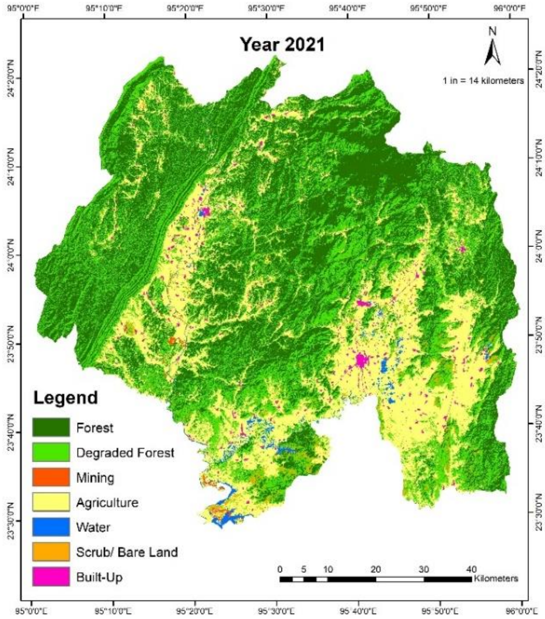


Figure 6: Land Cover Classification Map for 2021

Net area coverage of each land cover class in each year of Kawlin district obtained from the supervised classification by random tree classifier are shown Table 8. Figure 7 shows the net area changes during the study period. The area of forests in the study area was becoming decreased with time. In 2000, densed forest area was 320,275.34 hectares, which constituted 49.33% of the entire land cover. In 2021, it decreased to 262,957.08 hectares. This represented a decline of 8.83% within the periods of 21 years. From the analysis, a total densed forest area of 57318.26 hectares is converted to other land classes during 2000 to 2021.

The degraded forests occupied 163,144.56 hectares, which represented 25.13% of the entire land cover in 2000. Between 2000 and 2021, degraded forests increased from 159,055.01 to 198,702.77 hectares, an increase of 5.47% during the past 21 years. This represented an increase of 39,647.76 hectares of degraded forests within the period of 21 years.

The pixel statistics in Table 8 shows that mining area was 105.77 hectares in 2000, which was 0.02% of total land area. By 2021, mining areas were expanded from 105.77 to 1311.06 hectares, an increase of 1205.29 hectares (0.18%) within the periods of 21 years.

In 2000, the area occupied by agriculture was 138,398.81 hectares in 2000, which was 21.32% of total land cover. Between 2000 and 2021, agriculture areas increased from 138,398.81 to 161,932.89 hectares, an increase of 3.62% during 21 years, which put an increase at 23,534.08 hectares.

In 2000, the area of water body status was 6,999.84 hectares, which was 1.08% of total land area. Between 2000 and 2021, the status of water body areas decreased from 6,999.84 to 6,635.71 hectares, a decrease of 0.06% within the periods of 21 years.

During the period of 2000 to 2021, the scrub land areas decreased from 17,197.92 to

10,316.48 hectares, which puts a decrease of 1.06%. This represented a decrease of 6,881.44 hectares of scrub lands during the past 21 years.

The development of built-up areas in the study area has been gradually increased. This accounted for the proportion of 3,156.53 hectares in 2000. During 2000 to 2021, the built-up areas increased from 3,156.53 hectares to 7,422.77 hectares, an increase of 4266.24 hectares (0.66 %) within the period of 21 years.

Tab.8 Net Area (ha) for 2000 and 2021 Classification Maps

Class	2000		2021	
	Area (ha)	%	Area (ha)	%
Densed Forest	320,275.34	49.33	262,957.08	40.50
Degraded Forest	163,144.56	25.13	198,702.77	30.60
Mining	105.77	0.02	1,311.06	0.20
Agriculture	138,398.81	21.32	161,932.89	24.94
Water	6,999.84	1.08	6,635.71	1.02
Barrens/Scrub	17,197.92	2.65	10,316.48	1.59
Built Up	3,156.53	0.49	7,422.77	1.14
Total	649,278.76	100.00	649,278.76	100.00

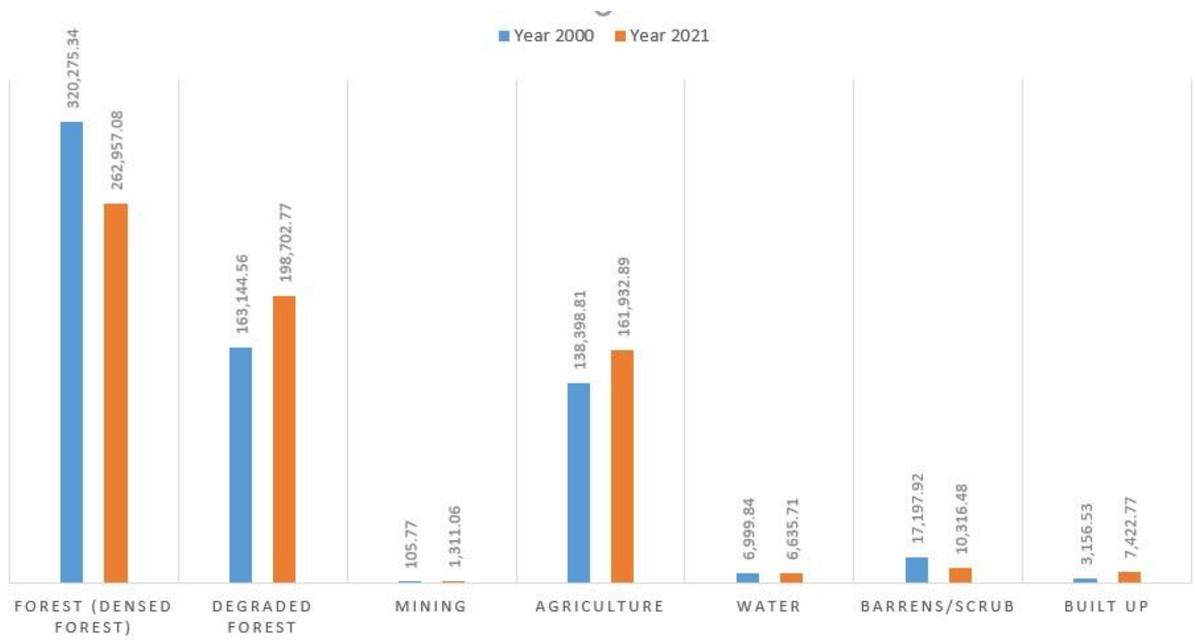


Figure 7: Contribution of Net Area (ha) in 2000 and 2021

4.3 Change Detection Analysis Results for 2000 and 2021

The land cover change matrix for the periods 2000-2021 showed the main transitions among seven land use/cover categories in this study (Table 9). Indeed, in the twenty-one (21) year period of study (i.e., from 2000 to 2021), a total of 83978.34 ha of densed forest cover was lost at the expense of other land use/cover types mainly converting to degraded forest (72,109.10 ha), agriculture (9,365.69 ha), scrub/bare land (1,539.93 ha), water (579.58 ha), mining (249.31 ha) and built up (134.73 ha) (Figure 8, Table 9). However, 26660.08 ha of forest cover was added to the existing forest cover due to conversion from other land use/cover types in which 24,140.30 ha of densed forest was restored from degraded forest (Table 12). In the case of the period between 2000 and 2021, most of the forest and degraded forest area was converted to mainly agriculture and scrub/bare land. The land cover class with the least conversion rate was the mining site. Overall, there was a net decrease in forest cover whereas degraded forest, agriculture, scrub/bare land, built-up, mining and water experienced gains. In terms of total forest cover which include both forest (densed forest) and degraded forest, an estimated 39,996.5 ha were lost, and an estimated 18,236.45 ha of forest were gained between 2000 and 2021. Figure 11 shows the map of land cover changes from 2000 to 2021, showing together with each sample area of deforestation and forest gain area.

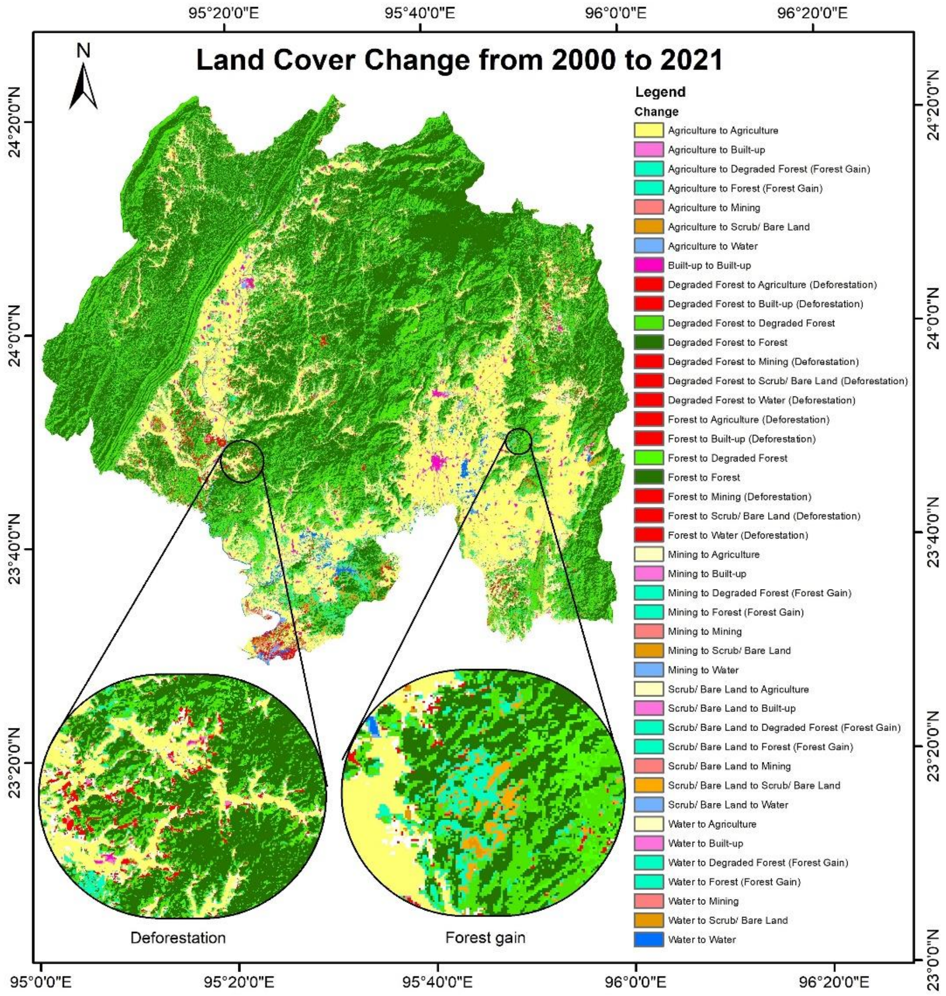


Figure 8: Map showing land cover changes from 2000 to 2021

Tab.9 Change Matrix from 2000 to 2021

Land Cover		2021							
		Forest	Degraded Forest	Mining	Agriculture	Water	Scrub/ Bare Land	Built up	Total
2000	Forest (Densed Forest)	**236,297.00**	72,109.10	249.31	9,365.69	579.58	1,539.93	134.73	320,275.34
	Degraded Forest	24,140.30	**110,877.00**	323.06	24,416.80	386.43	2,435.39	565.58	163,144.56
	Mining	0.17	0.26	**51.60**	34.04	17.61	1.90	0.20	105.77

Agriculture	979.56	8,478.10	514.58	**119,787.00**	2,520.82	2,823.02	3,295.73	138,398.81
Water	309.06	450.11	113.84	2,852.62	**2,889.69**	293.21	91.31	6,999.84
Scrub/ Bare Land	1,227.72	6,774.40	58.68	5,439.88	240.58	**3,222.30**	234.36	17,197.92
Built up	3.27	13.80	-	36.87	0.99	0.74	**3,100.86**	3,156.53
Total	262,957.08	198,702.77	1,311.06	161,932.89	6,635.71	10,316.48	7,422.77	649,278.76

Note: Bold values indicate the areas (ha) remained unchanged during 2000 and 2021.

4.4 Deforestation Probability by Binary Logistic Regression Results

Collinearity statistical test is conducted by linear regression in order to check variance inflation factor (VIF) and tolerance for the validation of independent variables to fit in logistic regression model. If tolerance value < 0.1 and VIF value >5 also indicate multi-collinearity problem in the dataset (Johnston et al., 2018). There is no multicollinearity among the independent variables with the tolerance value > 0.1 and VIF <5 (Table 10).

Tab.10 Collinearity Statistics of Different Independent Variables

Variables	Collinearity Tolerance	Statistics VIF
Soil type	0.827	1.209
Distance from settlements	0.535	1.868
Distance from roads	0.441	2.266
Distance from streams	0.938	1.066
Altitude	0.571	1.749
Slope	0.645	1.550

Correlation matrix shows the correlation coefficients among the independent variables to identify the presence of multicollinearity (Table 11, Figure 9). If simple correlation coefficient is greater than 0.8 or 0.9 multicollinearity is a serious problem between datasets. There is no serious concern about multicollinearity between variables.

Tab.11 Correlation matrix showing the relationship among independent variables using Pearson's correlation coefficient method

Variables	Soil type	Distance from villages	Distance from roads	Distance from streams	Altitude	Slope
Soil type	1	0.041	0.027	0.229	-0.310	-0.242
Distance from villages	0.141	1	0.676	-0.010	0.229	0.249
Distance from roads	0.027	0.676	1	0.043	0.432	0.380
Distance from streams	0.229	-0.010	0.043	1	-0.102	-0.077
Altitude	-0.310	0.229	0.432	-0.102	1	0.567
Slope	-0.242	0.249	0.380	-0.077	0.567	1

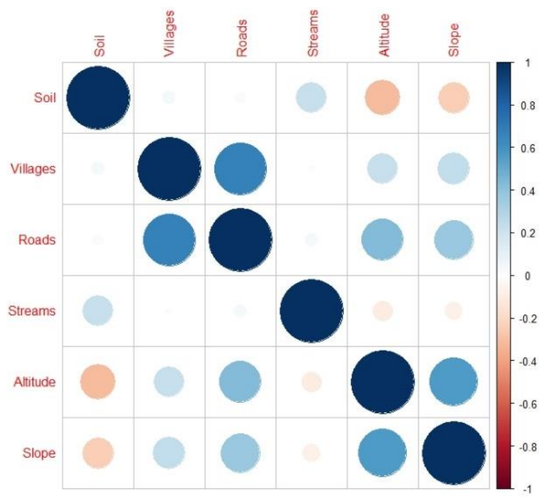


Figure 9: Correlation among the independent variables which reflect by a color matrix scale

The logistic regression model revealed that distance from settlements, distance from roads, altitude, and slope were significant factors in indicating the probability of deforestation (Table 12). For the case of distance from settlements, that factor is negatively correlated with deforestation at the significant high level of 0.001, and it means that areas located near settlements are a high possibility of deforestation. In the case of distance from roads, that factor is negatively correlated with deforestation at the significant level of 0.001. It means that areas located near roads are a high chance of deforestation. Other variables, altitude and slope are negatively related to deforestation at high significant levels of 0.001. According to logistic regression results, if the altitude and slope become higher, the probability of deforestation becomes lower.

Tab.12 Binary logistic regression results

Independent Variables	B	S.E	Z value	Pr(>|z|)	Exp (B)	95% C.I for Exp (B)	
						Lower	Upper
Intercept	4.770	0.295	16.171	2e-16 ***	27.465	4.192	5.348
Soil	0.030	0.116	0.258	0.796	0.712	-0.197	0.257
Distance from Settlements	-0.223	0.033	-6.819	9.17e-12 ***	0.743	-0.287	-0.159
Distance from Roads	-0.405	0.061	-6.635	3.24e-11***	0.784	-0.524	-0.285
Distance from Streams	0.044	0.028	1.563	0.118	1.106	-0.011	0.100
Altitude	-0.010	0.001	-10.984	2e-16***	0.988	-0.012	-0.009
Slope	-0.119	0.012	-9.611	2e-16***	0.870	-0.143	-0.095

Note: Signif. codes: '***' 0.001 (Dispersion parameter for binomial family taken to be 1) Null deviance: 3324.3 on 2397 degrees of freedom, Residual deviance: 2207.9 on 2391 degrees of freedom, AIC: 2221.9, Number of Fisher Scoring iterations: 6, Chi-squared test: X2 = 514.9, df = 7, P(> X2) = 0.0, -2log likelihood= -1103.946, Nagelkerke R square= 0.496

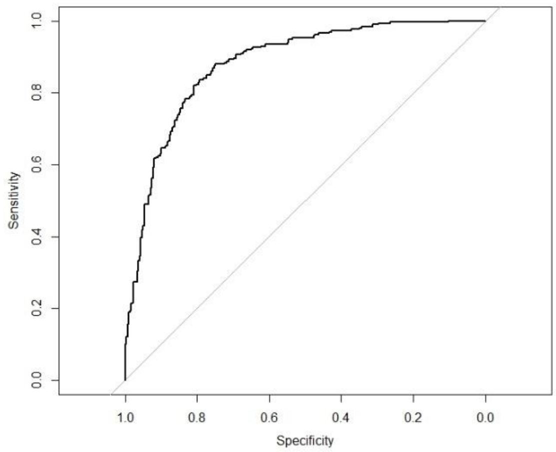


Figure 10: ROC curve for binary logistic regression model

Area under the curve: 0.881

The accuracy and reliability of binary logistic regression and random forest models have been evaluated based on 20% validation dataset. ROC curve (receiver operating characteristic curve) is represented graphically when True Positive Rate (TPR) or sensitivity values is in Y-axis, and the specificity values is in X-axis, which exhibits the performance of the classification model at all classification thresholds (Figure 10). Table 13 shows values of true positive rate (TPR), false positive rate (FPR), efficiency (E), true skill statistics (TSS) and Kappa co-efficient of binary logistic regression and random forest. The sensitivity (TPR) values of binary logistic regression model is 0.753 which means that 75.3% of the zones with deforestation were correctly identified by binary logistic regression model. The false positive rate (FPR) value is 0.247 that also shows the ability of the model. True skill statistics (TSS) takes into account both omission and commission errors, and success as a result of random guessing, and ranges from −1 to +1, where +1 indicates perfect agreement with values below 0.4 indicate poor model discrimination (Allouche et al., 2006) and values of zero or less indicate a performance no better than random. TSS values of binary logistic regression model is 0.63, and efficiency is 0.815. The kappa co-efficient of binary logistic regression is 0.63, which is also substantial or very good accuracy (Landis & Koch, 1977). Area under the curve (AUC) can explain model validation. According to the generated AUC (Areas Under Curve) of the binary logistic regression model from this study, AUC value is 0.881 which indicates this model has 88.1% accuracy (Figure 10). Deforestation probability map was generated based on the computed value of the binary logistic regression model using the following equation.

$$P = \frac{e^y}{1 + e^y} \tag{3}$$

The deforestation probability map also shows that the eastern part and lower part of the study area have high chance of deforestation (Figure 141. The binary logistic regression value ranges from 0 to 1, where nearer to 0 means deforestation probability is very low and nearer to 1 is very high. Deforestation probability result has been subdivided with equal interval into five categories as very low, low, moderate, high and very high occupying 53.85%, 19.43%, 14.84%, 9.26% and 2.62% of the forested areas respectively (Table 14).

Tab.13 Values of true positive rate(TPR), false positive rate(FPR), efficiency(E), true skill statistics (TSS) and Kappa co-efficient

Metrics	Binary Logistic Regression	Random Forest
TPR	0.753	0.866
FPR	0.247	0.187
Efficiency	0.815	0.838
TSS	0.63	0.679
Kappa	0.63	0.677

Tab.14 Deforestation probability of forested areas in Kawlin District

Deforestation probability zones	Binary Logistic Regression	
	Areas in ha	%
Very low	257,206.9	53.85
Low	92,814.74	19.43
Moderate	70,887.05	14.84
High	44,244.88	9.26
Very high	12,494.16	2.62

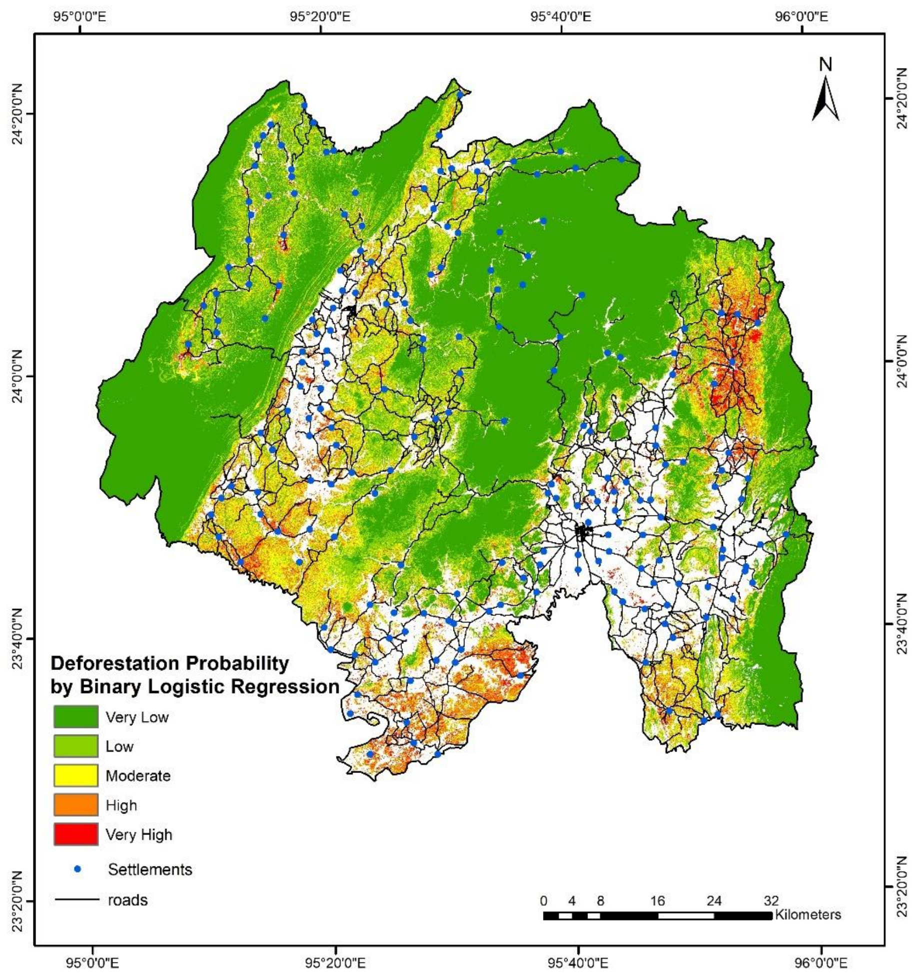


Figure 11: Map showing deforestation probability using binary logistic regression model

4.5 Deforestation Probability by Random Forest Results

The mean decrease in Gini coefficient quantifies how each variable contributes to the homogeneity of the resultant random forest nodes and leaves. The higher the value of mean decrease accuracy or mean decrease Gini score, the higher the importance of the variable in the model. The outcomes of random forest showed that the highest weight has been achieved by the altitude with a mean decrease Gini of 370.592 and lowest by soil (24.488). Based on Random Forest, altitude, distance from settlements, slope, distance from roads and distance from streams are key controlling factors of deforestation in this region and soil has no role to play here (Table 15).

Tab.15 Relative importance of the factors calculated by Random Forest

Variables	Mean Decrease Gini
Altitude	370.592
Slope	257.529
Distance from roads	209.694
Distance from settlements	182.655
Distance from Streams	151.873
Soil	24.488

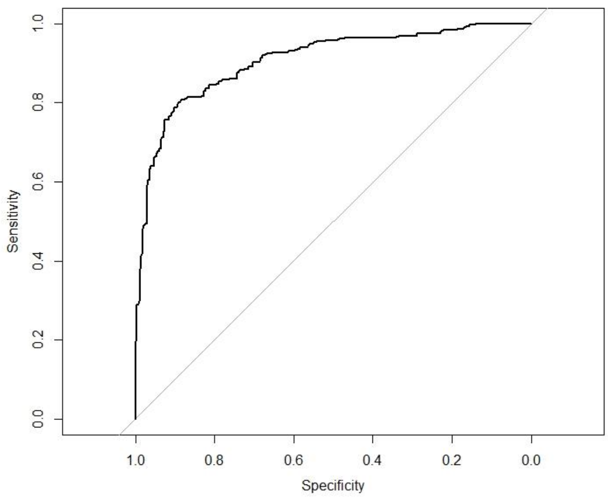


Figure 12: ROC curve of random forest

Area under the curve: 0.9072

The sensitivity (TPR) values of random forest is 0.866 which means that 86.6% of the zones with deforestation were correctly identified by random forest. The false positive rate (FPR) value is 0.187 that also shows the ability of the model. True skill statistics (TSS) values of random forest is 0.679, and efficiency is 0.838. The kappa co-efficient of random forest is 0.677, which have better accuracy than binary logistic regression model. According to the generated AUC (Areas Under Curve) of the Random Forest from this study, AUC value is 0.9072. The following deforestation probability map generated from random forest shows that 52.22%, 14.77%, 12.00%, 12.33% and 8.68% of forested areas under very low, low, moderate, high, and very high respectively (Table 16, Figure 13).

Tab.16 Deforestation probability of forested areas in Kawlin District

Deforestation probability zones	Random Forest	
	Areas in ha	%
Very low	249,417.36	52.22
Low	70,537.41	14.77
Moderate	57,350.07	12.00
High	58,893.12	12.33
Very high	41,449.77	8.68

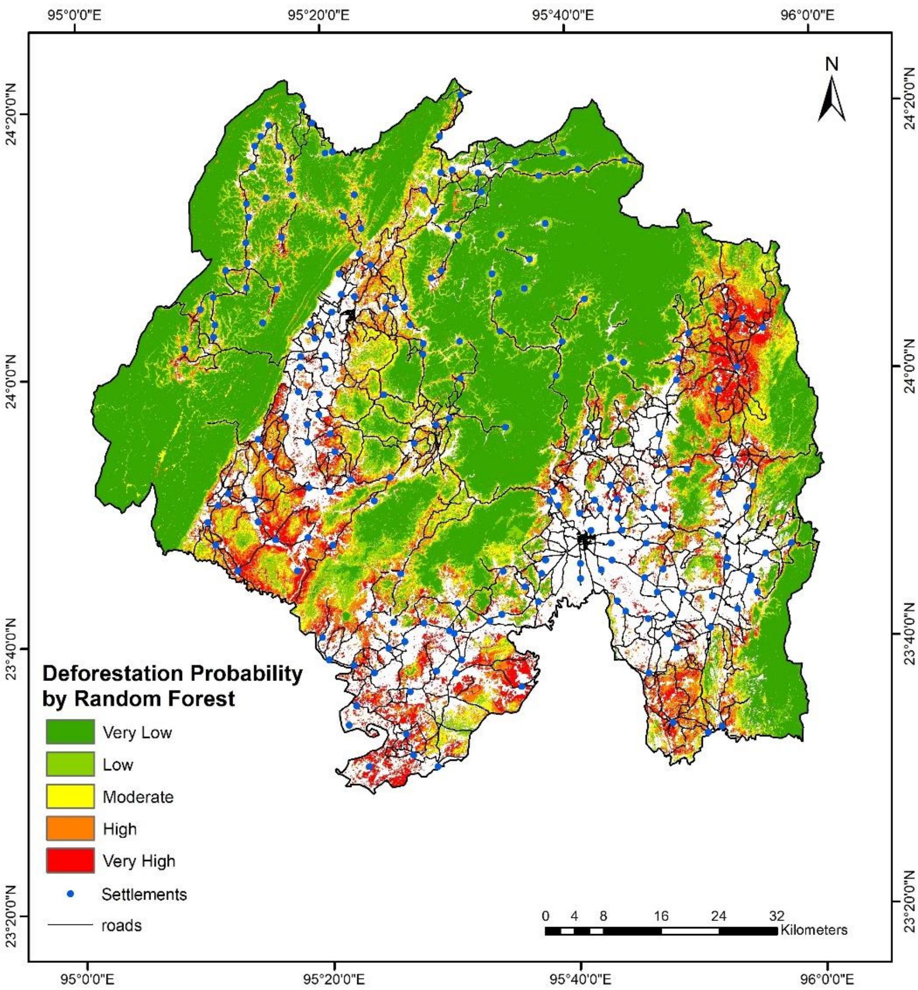


Figure 13: Map showing deforestation probability using random forest

4.6 Land Suitability Analysis Results for Potential Teak Plantation Areas

Pairwise comparison matrix used a scale of 1-9 where 9 indicates extreme importance and 1 indicates equal importance. According to AHP techniques, the pairwise comparison matrix was developed to determine the weights of the parameters (Table 20). After calculating the pairwise matrix, the eigenvector (λ_{max}) was calculated using Saaty's method (Thomas & Doherty, 1980) and the resulted maximum eigenvalue (λ_{max}) is 6.2664 with the consistency index (CI) of 0.05329 (Table 18). And then, consistency Ratio (CR) was calculated to detect logical inconsistency of the judgments and possible errors. Saaty suggest that if the value of CR > 0.1, the weight values of the pairwise comparison matrix are inconsistent, and AHP may not give meaningful results, but CR value up to 0.1 is acceptable for results. In the present study, the calculated Consistency ratio (CR) value was 0.0429 which is under acceptable limits and computed weight values are valid (Table 18).

Tab.17 Pairwise comparison matrix

Criteria	Slope	Land Cover	Stream	Road	Elevation	Settlement
Slope	1	2	4	5	8	9
Land Cover	1/2	1	3	4	7	9
Stream	1/4	1/3	1	3	6	7

Road	1/5	1/4	1/3	1	3	3
Elevation	1/8	1/7	1/6	1/3	1	2
Settlement	1/9	1/9	1/7	1/3	1/2	1

Tab.18 Synthesized matrix for multi-criteria decision making

Criteria	Slope	Land Cover	Stream	Road	Elevation	Settlement	wi
Slope	0.457	0.521	0.463	0.366	0.314	0.290	0.402
Land Cover	0.229	0.261	0.347	0.293	0.275	0.290	0.282
Stream	0.114	0.087	0.116	0.220	0.235	0.226	0.166
Road	0.091	0.065	0.039	0.073	0.118	0.097	0.080
Elevation	0.057	0.037	0.019	0.024	0.039	0.065	0.040
Settlement	0.051	0.029	0.017	0.024	0.020	0.032	0.029

When n=6 and maximum eigenvalue $\lambda_{\max} = 6.2664$,

$\lambda_{max} - n = 0.2664$, Consistency index (CI) = $(\lambda_{max} - n) / (n-1) = 0.2664/5 = 0.05329$

Random inconsistency indices (RI) = 1.24, Consistency ratio (CR) = (CI / RI) = 0.05329/1.24 = 0.0429

For suitable site identification, six parameters, namely slope, land cover, stream, road, elevation and settlements, were considered during potential teak plantation sites analysis. The sub-criteria were classified into 1-10 different ranking scales where 1 had the least importance, and 10 had the highest importance. The weight values of selected six parameters calculated in the Analytic Hierarchy Process and assigned scores of sub-criteria are used in Weighted Overlay Analysis to generate a land suitability map for teak plantations in Kawlin district (Table 19).

Tab.19 Weights of the criteria and scores of the sub-criteria

Main criteria	Weight	Influence (%)	Sub-criteria	Score
Slope	0.402	40.2	0°-10°	10
			10°-20°	8
			20°-30°	6
			30°-66.5°	1
Land Cover	0.282	28.2	Forest	1
			Degraded Forest	10
			Other Land Class	1
			Scrub Land	8
Stream	0.167	16.7	0-400 m	10
			400-800 m	7
			800-1200m	6

			1200-12960.3m	1
Road	0.080	8	0-700 m	10
			700-1400 m	7
			1400-1700 m	3
			1700-15623.2 m	1
Elevation	0.040	4	106-400 m	4
			400-700 m	10
			700-900 m	6
			900 -1689 m	1
Settlements	0.029	2.9	0-4000 m	10
			4000-6000 m	7
			6000-8000 m	4
			8000-17177.7m	1

The final land suitability map for potential teak plantation sites was classified up to four classes (1) highly suitable sites, (2) moderately suitable sites, (3) marginally suitable sites, and (4) currently not suitable sites based on the FAO framework of suitability classes (Figure 14). Total area coverage and percentage distribution of suitable teak plantation areas within the range of four suitability level are described in Table 20.

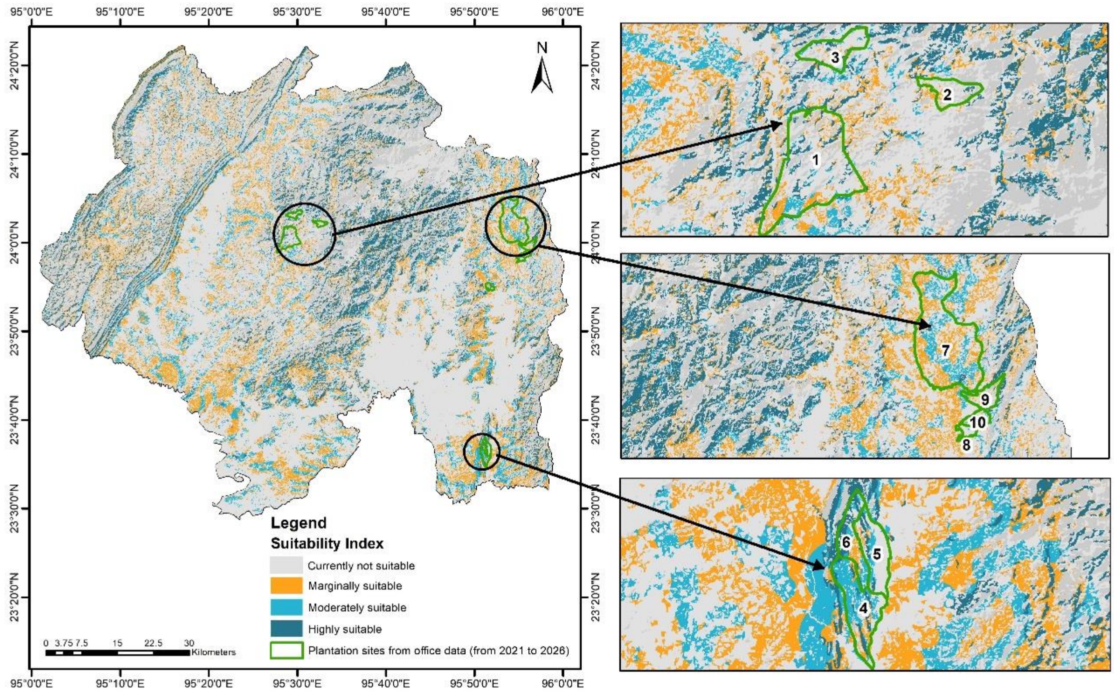


Figure 14: Teak plantation sites selected by office data from 2021 to 2026 and suitability areas for teak recommended by this study

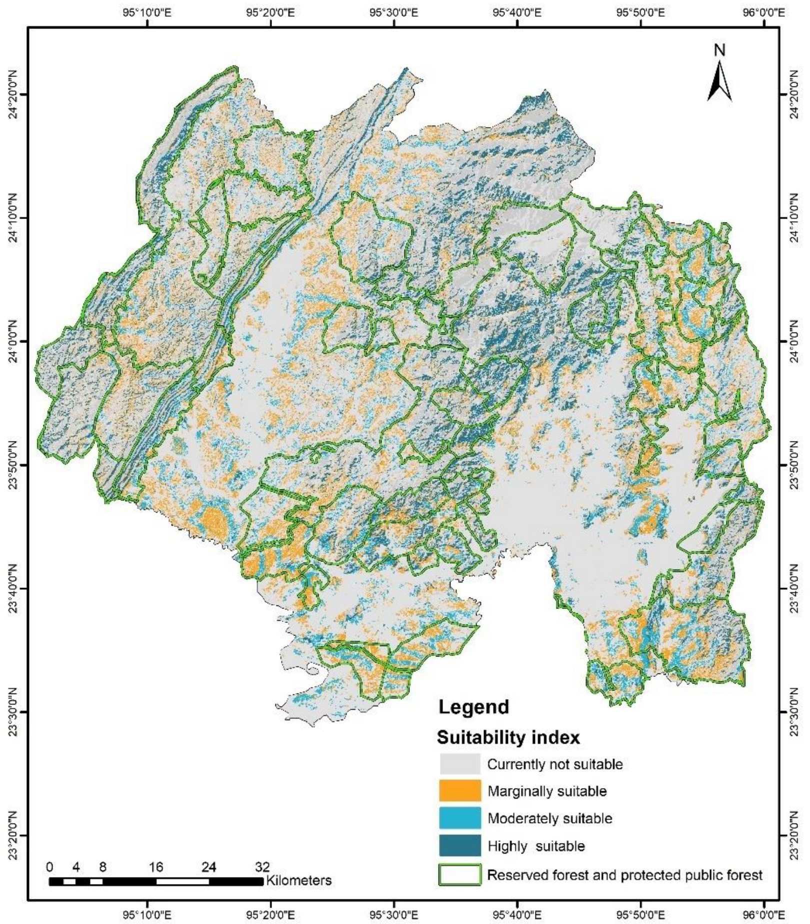


Figure 15: Site suitability map of teak plantations in Kawlin district

Tab.20 Area and percentage distribution of suitable teak plantation areas in Kawlin District

Suitability level	Suitable areas for teak plantation areas		Suitable areas within 54 reserved forests and protected public forests	
	ha	%	ha	%
Highly suitable	66,566.61	10.25	38,029.59	13.52
Moderately suitable	35,742.69	5.51	16,731.81	5.95
Marginally suitable	73,494.81	11.32	36,983.88	13.15
Currently not suitable	473,474.65	72.92	189,568.12	67.39
Total	649,278.76	100	281,313.4	100

According to the results, it was concluded that 10.25 % (66,566.61 ha) of the study area be highly suitable for teak plantations areas, 5.51% (35,742.69 ha) is moderately suitable, 11.32% (73,494.81 ha) is marginally suitable land, 72.92% (473,474.65 ha) is currently not suitable for teak plantations (Table 20). In Myanmar, reforestation and forestry operations are mainly conducted in reserved forests and protected public forest areas that are manageable and authorized by the Forest Department. Therefore, it was determined within reserved forests and protected public forests in which 13.52% (38,029.59 ha) be highly suitable, 5.95% (16,731.81 ha) is moderately suitable, 13.15% (36,983.88 ha) is marginally suitable land, 67.39% (189,568.12 ha) is currently not suitable for teak plantations. Figure 15 and 16 show the area (ha) of potential suitable teak plantation areas within 54 reserved forests (RFs) and protected public forests (PPFs).

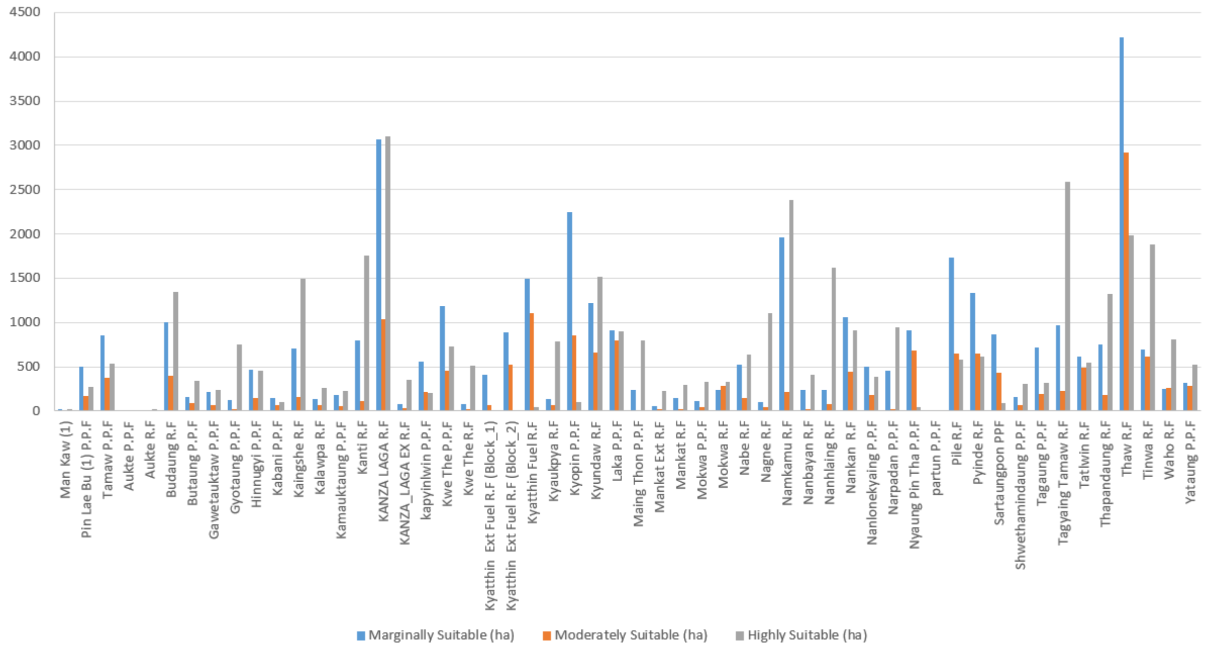


Figure 16: Potential suitable teak plantation areas within (54) reserved forests and protected public forests

Myanmar Forest Department (FD) developed Myanmar Reforestation and Rehabilitation (MRRP) from 2017 to 2026 (FD, 2017). Establishing forest plantations is one important activity in this ten-year program for the purpose of restoration of valuable forest tree resources in the country. For selection of suitable sites for proposed plantation areas, Forest Department use the standard operating procedure (SOP) and guidelines issued for the easy establishment of forestry work. GIS-aided site suitability analysis has not been used for the selection of sites for plantations. This study made the comparison between teak plantation sites to be established during 2021 and 2026 selected by Forest Department and the resulted suitable areas of teak plantations through weighted overlay analysis by this study. The proposed teak plantation sites within each compartment of reserved and protected public forests of Kawlin District which are preliminarily selected for MRRP by Forest Department were extracted and overlaid to the resulted suitability map of this study. Figure 14 is the comparison map for plantation sites between this study and MRRP. We can visually see that the plantation sites proposed by MRRP are nearly overlap or are close to each other by the sites suggested by this study. Table 21 showed the comparison of areas (ha) of each plantation site proposed by MRRP (from 2021 to 2026) and the sites recommended by this study.

Tab.21 Statistics of Forest Department teak plantation sites from 2021 to 2026 and recommended teak suitability zones from this study

Site Id.	areas of site(ha)	Proposed teak plantation areas(ha) by FD	Suitability areas from this study				
			Currently not suitable (ha)	Marginally suitable (ha)	Moderately suitable (ha)	Highly Suitable (ha)	Toal sum of suitable areas (ha)
1	1,154.71	161.87	931.15	95.67	28.8	99.09	223.56
2	202.42	161.87	176.14	19.62	0	6.66	26.28
3	268.47	80.94	217.08	16.02	13.5	21.87	51.39
4	348.71	121.4	123.62	38.7	142.11	44.28	225.09

5	250.56	40.47	172.26	11.61	39.69	27	78.3
6	199.52	40.47	62.9	24.66	41.04	70.92	136.62
7	3,462.63	161.87	1823.28	914.76	680.4	44.19	1639.35
8	125.33	80.94	106.25	15.75	3.33	0	19.08
9	360.53	80.94	284.48	53.46	15.48	7.11	76.05
10	203.08	80.94	168.34	20.88	12.51	1.35	34.74
total	6,575.95	1011.71	4065.5	1211.13	976.86	322.47	2510.46

From 2021 to 2026, a total of 1011.71 ha of teak plantations are planning to be established by Forest Department within 10 reserved forests (RF) and protected public forests (PPF) in Kawlin district. According to table 21, the total teak plantation suitability area from this study for site Id.2 is only 26.28 ha, which is much smaller than the proposed teak plantation areas (161.87 ha) by Forest Department when compared. For site Id. 1,4, 5, 6, and 7, the total suitability area from this study is larger than the proposed areas by Forest Department, which is more than enough areas for plantations to be established. The suitability area from this study for site no.9 is not much different from the Forest Department's proposed areas. For sites Id.3, 8, and 10, the total suitability area from this study is also lesser than the proposed teak plantation areas by Forest Department. Therefore, sites Id.2, 3, 8, and 10 have major differences from the proposed sites by Forest Department. Before teak plantations establishment begin, site Id.2,3,8 and 10 should be investigated on the ground whether those areas are suitable or not.

5. Discussion

5.1 Land Cover Changes between 2000 and 2021

The matrices of changes generated by intersecting of land cover maps of 2000 and 2021 of Kawlin District show changes at different units of occupation of seven land cover classes. From 2000 to 2021, 476,225.45 ha out of 649,278.76 ha of total area have not changed, accounting for 73.35% of Kawlin District. Meanwhile, there is an area of 173,053.31 ha or 26.62% of conversions from 2000 to 2021. From 2000 to 2021, approximately 57318.26 ha of densed forest had been converted to other land cover classes such as degraded forests, mining, agriculture and built-up areas. It is consistent with other studies in Myanmar which reported that intact forest loss increases with the expansion of other land cover classes(Kyaw, 2018; Mon et al., 2009).

From 2000 to 2021, the majority of the forest (72,109 ha) was directly converted to degraded forest; nevertheless, certain forest areas were restored as a result of the conversion of degraded forests into dense forests, indicating the region's forest succession. The conversion from forest to degraded forest may have been affected by illegal logging and overexploitation of teak and other commercial tree species beyond the annual allowable cut (AAC) over the past two decades (DEV, 2016). In Myanmar, selective logging of commercially valuable tree species is still practiced in almost 35% of natural forest areas, which leads to forest degradation(Mon et al., 2010; Mon et al., 2012). The silvicultural operations such as assisted natural regeneration and enrichment planting should be carried out in the degraded forest land to improve the natural regeneration and quality of the degraded forests. Establishing forest plantations with native tree species could be a proven way to prevent deforestation in the severely degraded land in the area. The Myanmar Reforestation and Rehabilitation Programme (MRRP) (2017-2026) provides financial support for establishing forest plantations in the Kawlin district. Thus, the forest managers may conduct proper forestry operations and establish forest plantations in the heavily depleted forest lands.

During the 21-year study period, the forest lands (both forest and degraded forest) were mainly transformed into agricultural land due to the expansion of agriculture into the forested areas. Cultivation of crops such as rice, groundnut, sesame and maize are a significant source of income for the district's residents. Thus, we can see the continuous agricultural lands across the entire area while visualizing the satellite images and the land cover map in 2021. However, agricultural encroachment into the forest lands may reduce forest cover, deplete forest resources and threaten the region's biodiversity. Other studies also reported that the expansion of crop land at the expense of the forest caused the significant land use changes leading to deforestation in India and Ethiopia (Behera et al., 2018; Tadese et al., 2020). The Kawlin district has around 330,660 residents, a 1.65% annual growth rate, and the majority of them work in the agricultural industries, according to statistics from the 2014 national census (DOP, 2014). The government and other relevant organizations should collaborate to invest in long-term agricultural development in the region to ease the farmers' financial and technical difficulties and obtain sustainable agrarian methods to prevent the loss of precious forest lands for agricultural purposes.

In addition, 6,881.44 ha of bare land had significantly gone from bare areas to vegetated areas and crop land. This may be due to climatic factors like increase of rainfall or other social factors like increase population over the district which contributed to the densification of vegetation and crop lands in certain areas. This result is in line with the study of Kamga(2020) who reported the conversion of bare land to vegetated areas in certain regions of east Cameroon while evaluating the land cover changes from 1987 to 2017. There is an increase appearance of mining activities over the study period. This may be due to the interest of people in the study area in the search of gold for economic benefits. Oo et al., (2022) stated that gold mining and agricultural land expansions caused severe land use/ land cover changes in the Kyaukpahto gold mining area in the Kawlin district. The increase in agriculture and the increase in the built-up area revealed the population growth in Kawlin District for 21 years. There was a slight decrease in water cover (364.13 ha) from 2000 to 2021, which may be due to the expense of agriculture and mining activities which unfavored the flow of natural streams.

5.2 Driving Factors of Deforestation

Finding of the results show that random forest has been found as better in predicting deforestation probability comparing the performance of binary logistic regression model. The random forest model provides numerous advantages over other classification methods. Random forest is not needed for assumptions about the independent variables and allows the categorical and numerical variables. Another advantage of the random forest was determining the importance of the input variables. Based on relative importance of the factors calculated by random forest, altitude, slope, distance from roads, distance from settlements and distance from streams are key controlling factors of deforestation in this region and soil has no role to play here. Deforestation probability maps generated by random forest and binary logistic regression showed that 41,449.77 ha (8.68%) and 12,494.16 ha (2.62%) of forested areas have very highly probability for deforestation, which is a big reason to worry if proper care is not taken. According to deforestation probability maps, eastern and southern parts of the study area were identified as having higher deforestation risks. In these high deforestation risk areas along roads, low elevation areas, gentle slopes with low land and around settlement areas should be focused by cautious monitoring in order to avoid potential risks. Deforestation probability maps prepared could be used by forest managers for designating and locating potential deforestation areas in order to prepare restoration planning and forest management related works.

Regarding six biophysical factors, most of them showed negative significant impact on the probability of deforestation in binary logistic regression model (BLR). Based on p values, distances from settlements, distances from roads, lower altitude and slope were the strongly significant factors of deforestation in BLR. The probability of deforestation increases with the decrease of distance from the settlements and roads. The spatial binary logistic regression model results confirm that the increase of deforestation is related with demographic factors such as the increase of population growth and road network, as has also been suggested by Lambin et al., (2001) and Plata-Rocha et al., (2021). However, more scientific research is required to predict the drivers of deforestation based on local communities with increased population and their socioeconomic livelihoods, i.e., how much that community relies on surrounding forests in order to make sustainable forest management strategies for the conservation of forest resources and development. In the present study, the biophysical factors such as altitude and slope affected deforestation. It is suggested that lowland or gentle slope areas with the low production of forest products have higher probability for deforestation. This result is in line with the studies of Mon et al., (2009) and Plata-Rocha et al., (2021) who reported that lower altitude and slope are significant factors for deforestation. Moreover, Wang et al., (2021) also reported that slope and elevation exerted the great influence on vegetation change in the Poyang lake basin in China. Although the likelihood of deforestation is low in high altitude (elevation) sites, it would be cautious to regularly monitor the characteristics contributing to deforestation. In the present study, the trigger events such as socio-economic conditions, population, climate and cultural factors were not spatially distributed although they were considered important in the process of deforestation. Wang et al., (2022) pointed out that the contribution of anthropogenic activities and climate variations affect on both forest restoration and degradation in the Yangtze River basin in China.

5.3 Suitable Site Selection for Teak Plantations

With decreasing availability of logs from natural forests, plantations have become an important source of timber. Before forest plantations are established, assessing suitable sites is important for accelerating plantation forestry because suitable site selection is the primary factor for any valuable tree species and significantly influences the cost and success of long-term plantation management.

Weighted overlay analysis was used in this study because it is widely used for land suitability analyses in different regions (Ahmad et al., 2017; Ambarwulan et al., 2016; Kar & Hodgson, 2008; Li & Nigh, 2011; Muñoz-Flores et al., 2017). However, researchers' experience and perspectives about the area is important for assigning the suitable ranks and weights. Thus, cross-referencing with the local Forest Department foresters and comparing the perspectives and experience of the author confirmed the ranks and weights of the weighted overlay analysis.

In this study, degraded forests and other bare/scrub land were considered as potential areas for teak plantations in weighted overlay analysis. In the degraded forests, teak plantation establishment is the most practical option because of difficult natural regeneration status (Khai et al., 2016), as has been suggested by Kyaw (2018) that degraded forests/ bamboo forests and other wooded land (such as scrubland, bushes and completed shifting cultivation areas) were considered as potential areas for teak plantations in Bago Yoma in Myanmar. Myanmar Forest Policy 1995 also stated that establishing plantations and providing institutional finance of man-made forests on degraded forest lands. However, if plantation forests were established on degraded forests (bamboo forests) or scrubland areas covered with small trees and bushes, it could provide wood production and economic benefits. Other land classes such as mining, agriculture, water and built-up were not considered as potential sites for teak plantations. Areas covered with paddy fields, croplands, and upland cultivation are recognized as traditional farming for the livelihood of indigenous people. Change of agriculture areas to teak plantation sites would conflict with local people's social, environmental, and economic impacts (Barlow & Cocklin, 2003). In addition, it is impossible for the option of converting mining, water, and built-up areas into teak plantations. For the success of teak plantations, it is most important to identify suitable sites which should be determined based on geology, topography, soil, rainfall, temperature and light (Tanaka et al., 1998). In this study, soil type is not considered for suitability analysis because of data availability limitations, and the soil types in this district are also teak forest trees preferred soil types such as red-brown forest soil and yellow-brown forest soil (MOAI, 2004). Climatic data (temperature and rainfall) are not considered because this district has only one weather station point. Global datasets (e.g., climate and soil) might not be convenient because of large spatial scale differences to the study area. Therefore, Kiran et al., (2015) described that GIS could provide a tremendous advantage for result data on accuracy and reliability, relying on the available spatial data.

6. Conclusion

Land cover maps for 2000 and 2021 were developed to detect recent land cover changes. Through change detection analysis, deforestation and forest degradation were quantified and the drivers of deforestation were estimated through spatial binary logistic regression model and random forest. The mapping of deforestation probability is also crucial because it can become an essential tool to prevent future deforestation. The final step was to provide spatial map for suitable teak plantation areas in Kawlin District.

The change detection results from 2000 to 2021 implied that 57,318.26 ha of forest was declined and a corresponding 35,558.21 ha of degraded forest was increased. The expansion of degraded forests reveals the decline of forests of both quantity and quality. However, deforestation can be prevented through effective management. Regarding to this point, the divergent effects of biophysical factors on deforestation were estimated using binary logistic regression and random forest. The results revealed that distance from settlements, distance from roads, elevation and slope significantly influenced on deforestation. Based on the results, effective measures can be conducted for the improvement of the study area. In the future, socio-economic variables and environmental factors of the study area should be added to improve the quality of this research. The weighted overlay analysis using AHP through implication of GIS and RS identified 649,278.76 ha of suitable sites for teak plantations with four level of suitability classes.

This research provides practically actionable information of site selection for the restoration and reforestation activities with the help of cost-effective technology. The results generated from this study can be used as a reference for reviewing the existing plantation establishment of Kawlin District. The procedures of this study can be integrated with the traditional site selection method to achieve more cost and time-effective and reliable spatial information.

7. References

Aguilar-Amuchastegui, N., Riveros, J. C., & Forrest, J. L. (2014). Identifying areas of deforestation risk for REDD+ using a species modeling tool. *Carbon balance and management*, *9*, 1-10.

Ahmad, F., Goparaju, L., Qayum, A. J. A. A., & Food. (2017). Agroforestry suitability analysis based upon nutrient availability mapping: a GIS based suitability mapping. *2*(2), 201-220.

Allouche, O., Tsoar, A., & Kadmon, R. J. J. o. a. e. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *43*(6), 1223-1232.

Ambarwulan, W., Santoso, P. B., Sabiham, S., & Hikmat, M. J. P. E. S. (2016). Remote sensing and land suitability analysis to establish local specific inputs for paddy fields in Subang, West Java. *33*, 94-107.

Atkinson, P. M., & Massari, R. (1998). Generalised linear modelling of susceptibility to landsliding in the central Apennines, Italy. *Computers & Geosciences*, *24*(4), 373-385.

Aung, T. S., Fischer, T. B., & Buchanan, J. (2020). Land use and land cover changes along the China-Myanmar Oil and Gas pipelines–Monitoring infrastructure development in remote conflict-prone regions. *PloS one*, *15*(8), e0237806.

Bandyopadhyay, S., Jaiswal, R., Hegde, V., & Jayaraman, V. (2009). Assessment of land suitability potentials for agriculture using a remote sensing and GIS based approach. *International journal of remote Sensing*, *30*(4), 879-895.

Barlow, K., & Cocklin, C. J. J. o. r. s. (2003). Reconstructing rurality and community: plantation forestry in Victoria, Australia. *19*(4), 503-519.

Behera, M., Tripathi, P., Das, P., Srivastava, S., Roy, P., Joshi, C., Behera, P., Deka, J., Kumar, P., & Khan, M. J. J. o. E. M. (2018). Remote sensing based deforestation analysis in Mahanadi and Brahmaputra river basin in India since 1985. *206*, 1192-1203.

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS journal of photogrammetry and remote sensing*, *114*, 24-31.

Breiman, L. (2001). Random forests. statistics department. *University of California, Berkeley, CA*, *4720*.

Breiman, L. J. M. l. (1996). Bagging predictors. *24*(2), 123-140.

Breiman, L. J. M. l. (2001). Random forests. *45*(1), 5-32.

Brunner, J., Talbott, K., & Elkin, C. (1998). *Logging Burma's frontier forests: resources and the regime*. World Resources Institute.

Bryant, R. L. (1997). *The political ecology of forestry in Burma, 1824-1994*. University of Hawaii Press.

Chamshama, S., Monela, G., Sekiete, K., & Persson, A. (1992). Suitability of the taungya system at North Kilimanjaro forest plantation, Tanzania. *Agroforestry Systems*, *17*, 1-11.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*, 37 - 46.

Department, F. (2016). *Standard Operating Procedure*.

Department, F. (2017). *Myanmar Reforestation and Rehabilitation Program (MRRP)*.

Department, F. (2020). *Kawlin District Forest Management Plan*.

DEV, O. S.-B. (2016). Legally and illegally logged out-The status of Myanmar's timber sector and options for reform-Oliver Springate-Baginski, Thorsten Treue, Kyaw Htun.

Dlamini, W. M. (2016). Analysis of deforestation patterns and drivers in Swaziland using efficient Bayesian multivariate classifiers. *Modeling earth systems and environment*, *2*, 1-14.

DOP. (2014). *The 2014 Myanmar Population and Housing Census (Sagaing Region)*.

FAO. (2015). Global Forest Resources Assessment 2015 (Desk Reference).

FAO. (2020). Global Forest Resources Assessment 2020: Main report. Rome.

Gautam, A. P., Webb, E. L., Shivakoti, G. P., & Zoebisch, M. A. (2003). Land use dynamics and landscape change pattern in a mountain watershed in Nepal. *Agriculture, ecosystems & environment*, *99*(1-3), 83-96.

Giri, C., Defourny, P., & Shrestha, S. (2003). Land cover characterization and mapping of continental Southeast Asia using multi-resolution satellite sensor data. *International journal of remote Sensing*, *24*(21), 4181-4196.

Greulich, F. R. (1999). A primer for timber harvesting.

Hanley, J. A., & McNeil, B. J. J. R. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *143*(1), 29-36.

Horning, N. (2010). Random Forests: An algorithm for image classification and generation of continuous fields data sets.

Htoo, M. T., Mainmit, N., & Trisurat, Y. (2020). Suitable Site Selection for Teak Plantations using GIS Technique at Phyu Township, Taungoo District, Bago Yoma, Myanmar. Remote Sens. *GIS Assoc. Thail. J*, *21*, 1-19.

J. Blaser, A. S., D. Poore, and S. Johnson. (2011). "Status of Tropical Forest Management 2011. ITTO Technical Series No 38., International Tropical Timber Organization, Yokohama, Japan, p 420," ed.

Johnson, F. X. (2010). What woodfuels can do to mitigate climate change.

Johnston, R., Jones, K., Manley, D. J. Q., & quantity. (2018). Confounding and collinearity in regression analysis: a cautionary tale and an alternative procedure, illustrated by studies of British voting behaviour. *52*(4), 1957-1976.

Kamga, M. A., Nguemhe Fils, S. C., Ayodele, M. O., Olatubara, C. O., Nzali, S., Adenikinju, A., & Khalifa, M. J. G. (2020). Evaluation of land use/land cover changes due to gold mining activities from 1987 to 2017 using landsat imagery, East Cameroon. *85*(4), 1097-1114.

Kar, B., & Hodgson, M. E. J. T. i. G. (2008). A GIS‐based model to determine site suitability of emergency evacuation shelters. *12*(2), 227-248.

Keiding, H., Wellendorf, H., & Lauridsen, E. B. (1986). *Evaluation of an international series of teak provenance trials*. Danida Forest Seed Centre.

Khai, T. C., Mizoue, N., Kajisa, T., Ota, T., Yoshida, S. J. G. e., & conservation. (2016). Stand structure, composition and illegal logging in selectively logged production forests of Myanmar: Comparison of two compartments subject to different cutting frequency. *7*, 132-140.

Kiker, G. A., Bridges, T. S., Varghese, A., Seager, T. P., & Linkov, I. (2005). Application of multicriteria decision analysis in environmental decision making. *Integrated environmental assessment and management: An international journal*, *1*(2), 95-108.

Kiran, G. S., Baria, P., Malhi, R. K. M., Revdandekar, A. V., Mudaliar, A. N., Joshi, U. B., & Shah, K. A. J. I. J. A. R. S. G. (2015). Site Suitability Analysis for JFM Plantation Sites using Geo-Spatial Techniques. *4*, 920-930.

Kyaw, T. Y. (2018). The Use of Landsat Imagery and Spatial Analysis to Detect Forest Cover Change and Degradation, and Determine Forest Management Suitability Indices in the Bago Mountain Region of Myanmar.

Lambin, E. F. (1994). Modelling Deforestation processes, a review. *Research Report, Joint Research Center, Institute for Remote Sensing Applications; European Space Agency, Luxembourg*.

Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., Coomes, O. T., Dirzo, R., Fischer, G., & Folke, C. J. G. e. c. (2001). The causes of land-use and land-cover change: moving beyond the myths. *11*(4), 261-269.

Landis, J. R., & Koch, G. G. J. b. (1977). The measurement of observer agreement for categorical data. 159-174.

Laurance, W. F. (2007). Forest destruction in tropical Asia. *Current Science*, 1544-1550.

Li, Y., & Nigh, T. J. A. G. (2011). GIS-based prioritization of private land parcels for biodiversity conservation: A case study from the Current and Eleven Point Conservation Opportunity Areas, Missouri. *31*(1), 98-107.

Lim, C. L., Prescott, G. W., De Alban, J. D. T., Ziegler, A. D., & Webb, E. L. (2017). Untangling the proximate causes and underlying drivers of deforestation and forest degradation in Myanmar. *Conservation Biology*, *31*(6), 1362-1372.

Mas, J.-F. (2000). Modelling deforestation in the region of the Lagoon of Términos, South East Mexico. Proceedings of the ASPRS (American Society of Photogrammetry and Remote Sensing) annual convention, Washington, 2000,

Menard, S. (2002). *Applied logistic regression analysis*. Sage.

MOAI. (2004). *Soil types and characteristics of Myanmar*.

Mon, M. S., Kajisa, T., Mizoue, N., & Yoshida, S. J. J. o. F. P. (2009). Factors affecting deforestation in Paunglaung Watershed, Myanmar using remote sensing and GIS. *14*(1), 7-16.

Mon, M. S., Kajisa, T., Mizoue, N., & Yoshida, S. J. J. o. f. p. (2010). Monitoring deforestation and forest degradation in the Bago Mountain Area, Myanmar using FCD Mapper. *15*(2), 63-72.

Mon, M. S., Mizoue, N., Htun, N. Z., Kajisa, T., Yoshida, S. J. F. E., & Management. (2012). Factors affecting deforestation and forest degradation in selectively logged production forest: A case study in Myanmar. *267*, 190-198.

Muñoz-Flores, H. J., Castillo-Quiroz, D., Castillo-Reyes, F., Sáenz-Reyes, J. T., Avila-Flores, D., & Rueda-Sánchez, A. J. O. J. o. F. (2017). Potential Areas for Commercial Timber

Plantations of Tabebuia rosea (Bertol.) DC. in Michoacan, Mexico. *7*(01), 48.

Myint, T. T. (2012). *A study on the sustainable management of natural teak forests in Myanmar* 東京大学].

Naing Tun, Z., Dargusch, P., McMoran, D., McAlpine, C., & Hill, G. (2021). Patterns and drivers of deforestation and forest degradation in Myanmar. *Sustainability*, *13*(14), 7539.

Nandy, S., Kushwaha, S., & Mukhopadhyay, S. (2007). Monitoring the Chilla–Motichur wildlife corridor using geospatial tools. *Journal for Nature Conservation*, *15*(4), 237-244.

Oo, T. K., Arunrat, N., Sereenonchai, S., Ussawarujikulchai, A., Chareonwong, U., & Nutmagul, W. J. S. (2022). Comparing Four Machine Learning Algorithms for Land Cover Classification in Gold Mining: A Case Study of Kyaukpahto Gold Mine, Northern Myanmar. *14*(17), 10754.

Ozdemir, A. (2011). Using a binary logistic regression method and GIS for evaluating and mapping the groundwater spring potential in the Sultan Mountains (Aksehir, Turkey). *Journal of hydrology*, *405*(1-2), 123-136.

Pampel, F. C. (2020). *Logistic regression: A primer*. Sage publications.

Plata-Rocha, W., Monjardin-Armenta, S. A., Pacheco-Angulo, C. E., Rangel-Peraza, J. G., Franco-Ochoa, C., & Mora-Felix, Z. D. J. L. (2021). Proximate and underlying deforestation causes in a tropical basin through specialized consultation and spatial logistic regression modeling. *10*(2), 186.

Saaty, T. L. (1980). *The Analytical Hierarchy Process; McGraw-Hill: New York, NY, USA*.

Springate-Baginski, O., Thein, A. K., Neil, A., Thu, W. M., & Doherty, F. (2014). Democratising timber: an assessment of Myanmar's emerging 'Forest law enforcement, governance and trade'(FLEGT) process. *Forest Policy and Economics*, *48*, 33-45.

Srivastava, S., Singh, T., Singh, H., Kushwaha, S., & Roy, P. (2002). Assessment of large-scale deforestation in Sonitpur district of Assam. *Current Science*, 1479-1484.

Tadese, M., Kumar, L., Koech, R., Kogo, B. K. J. R. S. A. S., & Environment. (2020). Mapping of land-use/land-cover changes and its dynamics in Awash River Basin using remote sensing and GIS. *19*, 100352.

Tanaka, N., Hamazaki, T., & Vacharangkura, T. J. J.-j. A. R. Q. (1998). Distribution, Growth and Site Requirements of Teak. *32*, 65-77.

Thomas, P. G., & Doherty, P. C. (1980). The Analytic Hierarchy. Process: Planning, Priority Setting, Resource Allocation, McGraw-Hill,

Tripathi, S., Subedi, R., & Adhikari, H. J. R. S. (2020). Forest cover change pattern after the intervention of community forestry management system in the mid-hill of Nepal: A case study. *12*(17), 2756.

W. Kollert, L. J. F. P. F. C., and R. (2012). Teak resources and market assessment 2010. *Trees Working Paper FP/47/E*, *206*.

Wang, Y., Zhang, Z., & Chen, X. J. F. (2022). The Dominant Driving Force of Forest Change in the Yangtze River Basin, China: Climate Variation or Anthropogenic Activities? , *13*(1), 82.

Wang, Y., Zhang, Z., & Chen, X. J. R. S. (2021). Quantifying influences of natural and anthropogenic factors on vegetation changes based on geodetector: A case study in the Poyang Lake Basin, China. *13*(24), 5081.

Watch, G. F. (2020). "Katha, Sagaing, Myanmar deforestation rate and statistics.".

Yang, R., Luo, Y., Yang, K., Hong, L., & Zhou, X. (2019). Analysis of forest deforestation and its driving factors in Myanmar from 1988 to 2017. *Sustainability*, *11*(11), 3047.

Yesilnacar, E., & Topal, T. J. E. G. (2005). Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). *79*(3-4), 251-266.