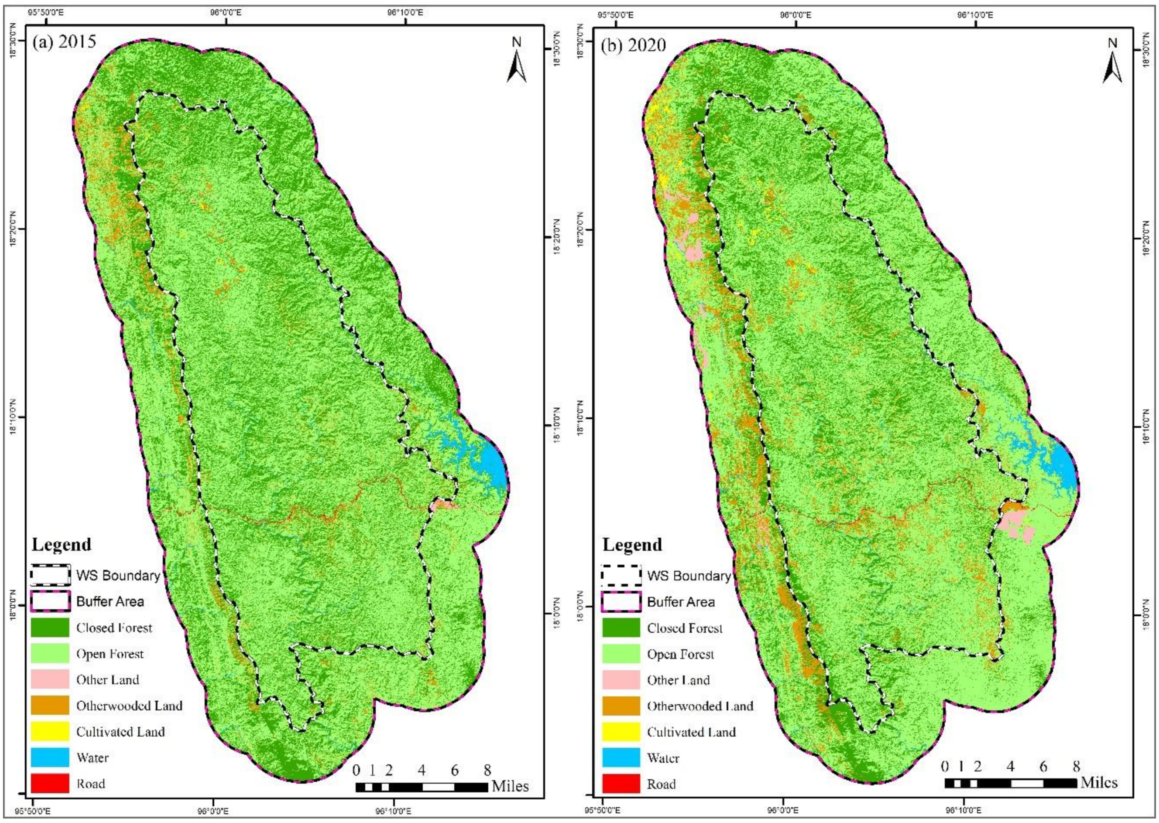
The Republic of the Union Myanmar

Ministry of Natural Resources and Environmental Conservation

Forest Department

Factors Affecting Forest Cover Change in North Zamari Wildlife Sanctuary using RS and GIS Approach



U Chan Myae Aung, Staff Officer, Forest Research Institute, Yezin

U Nay Aung, Staff Officer, Planning and Statistics Division

Dr. Thin Thin, Assistant Director, Planning and Statistics Division

U Khine Zaw Wynn, Assistant Director, Planning and Statistics Division

Dr. Yuya Aye, Assistant Director, Forest Research Institute, Yezin

February, 2024

Table of Contents

စာတမ်းအကျဉ်း

မြောက်ဇာမရီတောရိုင်းတိရစ္ဆာန်ဘေးမဲ့တော၏ သစ်တောဖုံးလွှမ်းမှု ပြောင်းလဲခြင်းအပေါ် သက်ရောက်နိုင်သည့် အကြောင်းအရင်းများအား ပထဝီဝင်သတင်းအချက်အလက် နည်းပညာအသုံးပြု၍ ဆန်းစစ်လေ့လာခြင်း

ဦးချမ်းမြေ့အောင် (ဦးစီးအရာရှိ၊ သစ်တောသုတေသနဌာန)
ဦးနေအောင် (ဦးစီးအရာရှိ၊ RS & GIS ဌာနစိတ်၊ စီမံကိန်းနှင့်စာရင်းအင်းဌာန)
ဒေါက်တာသင်းသင်း(လက်ထောက်ညွှန်ကြားရေးမှူး၊ RS&GISဌာနစိတ်၊ စီမံကိန်းနှင့်စာရင်းအင်းဌာန)
ဦးခိုင်ဇော်ဝင်း (လက်ထောက်ညွှန်ကြားရေးမှူး၊ RS & GIS ဌာနစိတ်၊ စီမံကိန်းနှင့်စာရင်းအင်းဌာန)
ဒေါက်တာယုယအေး (လက်ထောက်ညွှန်ကြားရေးမှူး၊ သစ်တောသုတေသနဌာန)

သစ်တောဂေဟစနစ်သည် ဇီဝမျိုးစုံမျိုးကွဲထိန်းသိမ်းခြင်း၊ ရာသီဥတုညီညွတ်မျှတစေခြင်းနှင့် အခြားသော ဂေဟစနစ်ဝန်ဆောင်မှုများကို အထောက်အပံ့ပေးခြင်းတို့တွင် အလွန်အရေးကြီးသော အခန်းကဏ္ဍမှ ပါဝင်လျက် ရှိပါသည်။ သစ်တောဖုံးလွှမ်းမှု ပြောင်းလဲစေသည့် အကြောင်းအရင်းများကို သိရှိနားလည်ခြင်းဖြင့် သစ်တောထိန်းသိမ်းရေးလုပ်ငန်းများဆောင်ရွက်ရာတွင် များစွာအထောက်အကူ ဖြစ်စေပါသည်။ ထို့အပြင် သစ်တောများသည် ဇီဝမျိုးစုံမျိုးကွဲများအတွက် အရေးကြီးသော ဂေဟစနစ် တစ်ခုဖြစ်သည့်အပြင် ကာဗွန်သိုလှောင်ရုံရာနေရာများလည်း ဖြစ်ပါသည်။ ဤသုတေသန၏ အဓိက ရည်ရွယ်ချက်မှာ ၂၀၁၅-၂၀၂၀ ခုနှစ်အတွင်း မြောက်ဇာမရီတောရိုင်းတိရစ္ဆာန်ဘေးမဲ့တောနှင့် ၎င်း၏ ကြားခံနယ်မြေတို့တွင် သစ်တောဖုံးလွှမ်းမှု ပြောင်းလဲခြင်းအား ဆန်းစစ်လေ့လာနိုင်ရန်နှင့် သစ်တော ဖုံးလွှမ်းမှုပြောင်းလဲစေသည့် အကြောင်းအရင်းများနှင့် ပြောင်းလဲရန် အလားအလာရှိသည့် ဧရိယာများ အား ခန့်မှန်းဖော်ထုတ်နိုင်ရန်တို့ ဖြစ်ပါသည်။ Binary Logistic Regression (BLR) Model နှင့် ပထဝီဝင်သတင်းအချက်အလက်နည်းပညာ (RS&GIS) တို့ကို ပေါင်းစပ်အသုံးပြု၍ ဆန်းစစ်ခဲ့ပါသည်။ သုတေသနတွေ့ရှိချက်အနေဖြင့် မြောက်ဇာမရီတောရိုင်းတိရစ္ဆာန်ဘေးမဲ့တောအတွင်း၌ ၂၀၁၅-၂၀၂၀ ခုနှစ်အတွင်း ရွက်အုပ်ပိတ်တောသည် ၃၄.၂၃% မှ ၂၈.၁၆% အထိ လျော့နည်းသွားပြီး ရွက်အုပ်ပွင့် တောရေိယာသည် ၆၁.၇၇% မှ ၆၃.၈၁% အထိ တိုးပွားလာသည်ကို တွေ့ရှိရပါသည်။ ထို့အပြင် တောနိမ့်ဧရိယာနှင့် စိုက်ပျိုးမြေဧရိယာလည်း တိုးလာပါသည်။ Binary Logistic Regression (BLR) Model ၏ တိကျမှန်ကန်မှုသည် ၈၅.၉၆ % ရှိပြီး ၎င်းမှ ထွက်ထုတ်ပေးသော သစ်တောပြုန်းတီးရန် အလားအလာရှိသည့် တည်နေရာပြ မြေပုံအရ မြောက်ဇာမရီတောရိုင်းတိရစ္ဆာန်ဘေးမဲ့တောဧရိယာ အတွင်း သစ်တောဧရိယာ ၁၀၅၃.၅၆ ဟက်တာ (၁.၀၈ %) နှင့် ကြားခံနယ်မြေအတွင်း ၃၉၆၅.၆၁ ဟက်တာ(၄.၄၂%)တို့သည် ပြုန်းတီးနိုင်ခြေရှိကြောင်း တွေ့ရှိရပါသည်။ မတ်စောက်သော လျှောစောက် နေရာများတွင် သစ်တောပြုန်းတီးနိုင်ခြေနည်းသော်လည်း ရေရရှိရန်လွယ်ကူသည့် ချောင်းနှင့်နီးသော နေရာများတွင် စိုက်ပျိုးမြေချဲ့ထွင်ခြင်းနှင့် ရွှေ့ပြောင်းတောင်ယာများကြောင့် သစ်တောပြုန်းတီးနိုင်ခြေ ပိုများကြောင်း တွေ့ရှိရပါသည်။ ဤသုတေသနတွေ့ရှိချက်များသည် REDD+ လုပ်ငန်း၏ အဓိက ရည်မှန်းချက်ဖြစ်သည့် သစ်တောပြုန်းတီးမှုလျော့ချခြင်းမှတဆင့် သစ်တော ကာဗွန်သိုလှောင်မှုပမာဏ တိုးပွားလာစေရန်အတွက် သစ်တောထိန်းသိမ်းရေးလုပ်ငန်းများ ဦးစားပေး ဆောင်ရွက်သင့်သည့် နေရာများကို ဖော်ထုတ်ပေးခြင်း ဖြစ်ပါသည်။

ABSTRACT

Factors Affecting Forest Cover Change in North Zamari Wildlife Sanctuary using RS and GIS Approach

U Chan Myae Aung (Staff Officer, Forest Research Institute, Yezin)
U Nay Aung (Staff Officer, RS & GIS Section, Planning and Statistics Division)
Dr. Thin Thin (Assistant Director, RS & GIS Section, Planning and Statistics Division)
U Khine Zaw Wynn (Assistant Director, RS & GIS Section, Planning and Statistics Division)
Dr. Yuya Aye (Assistant Director, Forest Research Institute, Yezin)

Forest ecosystems play a vital role in biodiversity conservation, climate regulation, and the provision of critical ecosystem services. Understanding the factors influencing forest cover change is crucial for effective conservation, considering the region's significance as a crucial ecosystem supporting diverse species and its role as a substantial carbon sink. The primary objectives of this study were to assess the patterns of forest cover changes in the North Zamari Wildlife Sanctuary and its adjacent 5 km buffer zone between 2015 and 2020, utilizing a combination of Remote Sensing (RS) and Geographic Information System (GIS) approach with Sentinel-2 satellite images. Additionally, the study aimed to investigate the factors influencing deforestation and the probable deforestation areas in this region through the application of the Binary Logistic Regression (BLR) model. The results revealed a notable reduction in closed forest cover within the protected area, decreasing from 34.23% to 28.16%, accompanied by a slight increase in open forest area from 61.77% to 63.81%. Furthermore, an increase in other wooded land and cropland was observed in the study area. The BLR model demonstrated strong performance, with an AUC of 0.8596 and Kappa of 0.6216, highlighting that factors such as proximity to roads, elevation, slope and streams significantly influence deforestation. The areas located at farther distances from roads and higher elevations were observed to be more susceptible, potentially contributed to limited monitoring and enforcement in remote areas, along with specific land-use practices such as crop cultivation. Steeper slope areas were found to be less likely to experience deforestation, possibly due to the challenges associated with accessing and utilizing these terrains for anthropogenic activities. In contrast, areas near streams were more susceptible to deforestation, particularly influenced by shifting cultivation practices near watercourses. The deforestation probability map generated by the BLR model revealed that 1053.56 hectares (1.08%) of forested areas within the protected area and 3965.61 hectares (4.42%) in the buffer zone are at high risk of deforestation. The findings of the study highlight the urgent need for conservation and monitoring strategies, directly aligning with the core objectives of REDD+ initiatives in combating deforestation and preserving crucial carbon sinks.

Keywords: Forest cover change, Deforestation, Binary Logistic Regression (BLR), North Zamari Wildlife Sanctuary

1. Introduction

Forest is the most terrestrial ecosystem on the earth (Xie et al., 2010), and it plays a vital role in biodiversity conservation, climate regulation, and the provision of critical ecosystem services (Krieger, 2001). According to Global Forest Resource Assessment, forests cover nearly one-third of the total land area of the world, which occupied 4.06 billion ha with 31% of the world's total land area. Among these areas, the most significant proportion of the world's forests (45%) is occurred in the tropical forests, followed by the boreal, temperate and subtropical forests, respectively (GFRA, 2020). However, tropical forests depleted at alarming rates with significant impacts on ecosystems, climate and livelihoods (Pendrill et al., 2022). It was estimated that the annual rate of deforestation ranged from 10 to 12 million ha in 2010-2015 (FAO, 2020).

Myanmar is one of the countries in Southeast Asia, which has a wide range of forest cover (Naing Tun et al., 2021). Like other countries, Myanmar has been experiencing a significant deforestation and forest degradation rate. The country's forest cover in 2020 was about 42.19% of the total land area, decreasing from 31 million ha in 2010 and 28 million ha in 2020 (FAO, 2020). As a result, Myanmar had the third-highest rate of deforestation in the world during 2010-2015 (FRA, 2015), and it ranked seventh in the world for deforestation in 2020. Although there are many proximate and underlying causes of deforestation and forest degradation in Myanmar, agricultural expansion, legal and illegal logging, fuelwood and charcoal consumption, road construction, mining, and dam construction are significant drivers of forest loss (Naing Tun et al., 2021). In other words, forest cover change has significantly occurred in the country, conversion of forests into agricultural land, degraded forest land, built-up area, bare land, and road.

Information on changes in land use and forest cover is crucial in the management of natural forests and evaluation of their impacts on the environment (Win et al., 2009). Understanding the factors influencing forest cover change play a critical role for effective conservation, considering the region's significance as a crucial ecosystem supporting diverse species and its role as substantial carbon sink (Kafy et al., 2023). However, the patterns and processes of deforestation have been conducted in many studies, but information about factors influencing forest degradation is still limited (Mon et al., 2012). Therefore, studying factors affecting forest cover change and the probable deforestation area support decision making processes in the sustainable forest management and conservation.

As science and technology progress, remote sensing (RS) is an essential tool and widely used for the studies of LULC changes (Alqurashi & Kumar, 2013; Baynard, 2013; Rwanga & Ndambuki, 2017) because of the increasing availability of data, time and cost-effectiveness, and good quality of the resulting data (Baynard, 2013). Remote sensing satellite data is the main source for assessing, quantifying and mapping LULC changes and patterns due to its repetitive data acquisition (Abd El-Kawy et al., 2011). Moreover, remote sensing (RS) and geographic information system (GIS) are applied to conduct monitoring, mapping, and management of forest resources (Franklin, 2001). Moreover, the remote sensing method becomes capable of identifying forest health condition and deforestation rate (Guhathakurta & Roy, 2000). Many previous studies have used RS and GIS approach to conduct land use and land cover change (LULC), and forest cover change (Belayneh et al., 2020; Negassa et al., 2020; Spruce et al., 2020; Yaung et al., 2021).

Binary Logistic Regression (BLR) model is frequently applied in the analysis of land cover change (Bera et al., 2020; Buya et al., 2020; Chang et al., 2021). The BLR model is a regression method applied to determine the correlation between the binary dependent variable (y) and the independent variable (x). It is particularly suitable for phenomena with binary properties, such as land cover change (Liao et al., 2010; Muzdalifah et al., 2020). The BLR model has been used in mapping landslide susceptibility, hazard zonation, predicting probable deforestation area, and assessing ground water potential (Ozdemir, 2011). In the recent years, numerous studies have been conducted to estimate deforestation probability areas, applying a combination of RS and GIS techniques along with the BLR model. The main objective of the present study is to evaluate changes in forest cover, and predict deforestation probability area within North Zamari Wildlife Sanctuary (WS) and the surrounding 5 km buffer area during the period of 2015-2020.

2. Objective of the Study

The overall objective of the study is to assess forest cover changes in the North Zamari Wildlife Sanctuary and its adjacent 5 km buffer area from 2015 to 2020, with the use of RS and GIS approach with Sentinel-2 satellite images. To achieve the primary objective of the research, the specific objectives are as follows:

(1) To estimate forest cover changes of the study area between 2015 and 2020
(2) To investigate the factors influencing deforestation and the probable deforestation areas of the study area, using Binary Logistic Regression (BLR) model.

3. Materials and Methods

3.1.Study Area

The North Zamari Wildlife Sanctuary (WS) is located in central Myanmar, spanning the western part of the Bago Yoma mountain range from north to south. The sanctuary was officially designated on May 22, 2014. It covers an area of 98,314 hectares and is a vital conservation site located between approximately 18°00'N to 18°30'N latitude and 96°00'E to 96°20'E longitude. It spans across the Tharawady and Bago districts. The sanctuary spans multiple townships: Gyobingauk, Okpho, Minhla, Letpadan, Kyauktaga, and Bago.

The North Zamari WS exhibits a complex geography characterized by significant variations in elevation. The northern part is characterized by its dominant steep landscape, with elevations reaching up to 815 meters. Conversely, the southern part of the sanctuary has altitudes of a more moderate kind. The vegetation of the sanctuary consists mainly of a combination of mixed deciduous and dry deciduous forests. The North Zamari Wildlife Sanctuary is surrounded by several Reserved Forests (R.F.) that are crucial for conserving the Sanctuary's diverse ecosystems. Pyu Kum R.F. surrounds the sanctuary to the north, Bawbin R.F. to the northwest, Yenwe R.F. and Yenwe Extension R.F. to the northeast, Min Hla R.F. to the east, Baing Dar R.F. to the southeast, Mokka R.F. to the south, Kadin Bilin R.F. to the southwest, and Thanonze R.F. to the west. These reserved forests serve as vital areas that act as buffers for conserving Wildlife Sanctuary's biodiversity.

3.2.Data Acquisition

Sentinel-2 satellite images have been used to detect forest cover change in the North Zamari Wildlife Sanctuary between 2015 and 2020. Level-1C Sentinel-2 satellite images were acquired from the United States Geological Survey (USGS) website. The Level-1C products provides Top-Of-Atmosphere (TOA) multispectral reflectance which has been corrected for sensor distortions, atmospheric effects, and standardized for land monitoring studies. Only the images from December to February were downloaded to ensure tree canopy of deciduous forest and avoid cloud cover. The same common spectral bands with a spatial resolution of 10-20 m for the two different periods were downloaded (Table 1). Other auxiliary data such as DEM data from the USGS, topographic maps developed by Myanmar Survey Department, road and stream shapefiles from Myanmar Forest Department, were collected to improve image interpretation and classification process.

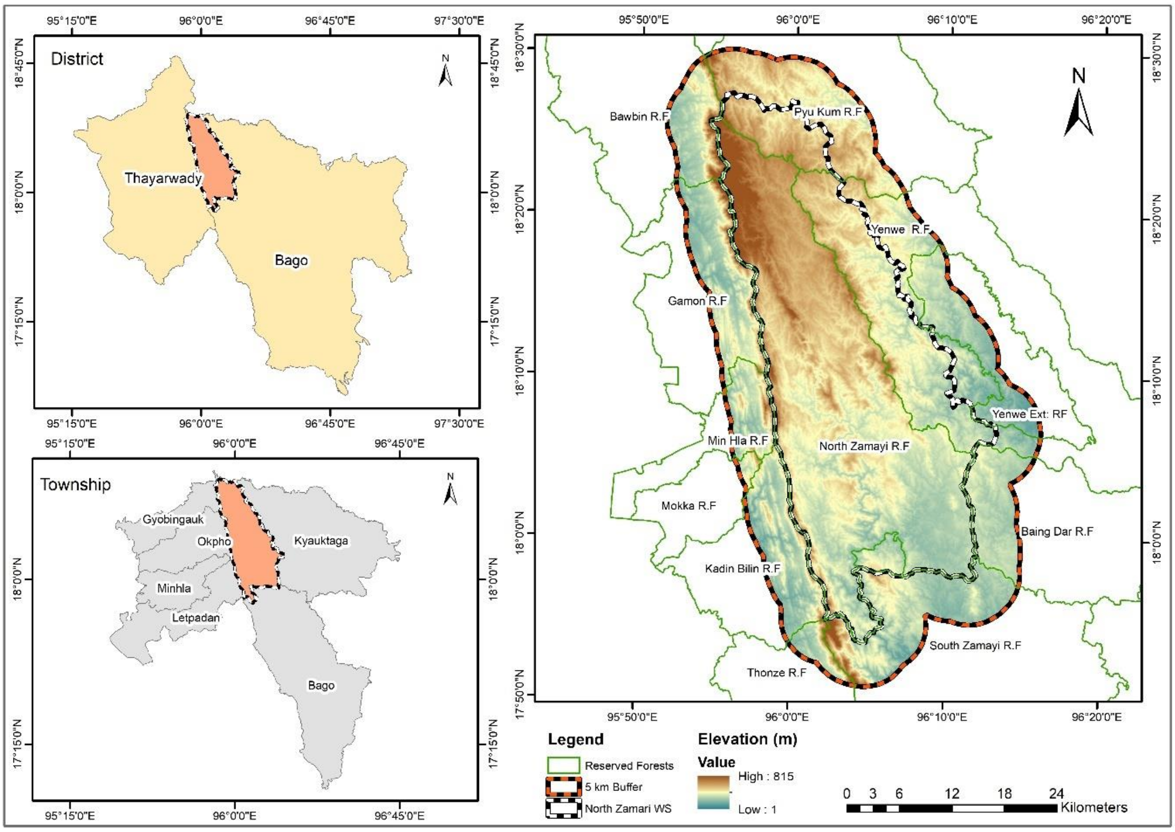


Figure 1: Study Area Map (within North Zamari WS and its adjacent 5 km Buffer Area)

Table 1: Spectral bands of Sentinel-2 satellite images used in the study

Band Name	Spatial Resolution (m)	Wavelength (nm)	Acquisition Date
Blue	10 m	496.6nm (S2A) / 492.1nm (S2B)	1^{st} December to 28^{th} February
Green	10 m	560nm (S2A) / 559nm (S2B	1st December to 28^{th} February
Red	10 m	664.5nm (S2A) / 665nm (S2B)	1^{st} December to 28^{th} February
NIR	10 m	835.1nm (S2A) / 833nm (S2B)	1st December to 28^{th} February
SWIR 1	20 m	1613.7nm (S2A) / 1610.4nm (S2B)	1^{st} December to 28^{th} February
SWIR 2	20 m	2202.4nm (S2A) / 2185.7nm (S2B)	1st December to 28^{th} February

3.3.Image Classification and Change Detection

Image interpretation was conducted to classify land cover classes using a supervised classification method with a maximum likelihood classification (MLC) algorithm in ENVI 5.3 software. Seven different land use and land cover (LULC) types have been classified in the study, namely closed forest, open forest, otherwooded land, other land, road, and water (Table 2). Initially, visual interpretation of the study area was conducted using satellite imagery acquired from the Google Earth Pro. Then, Regions of Interest (ROIs) were selected manually as training samples for the supervised classification. The ROIs were chosen to represent designated LULC classes, with an emphasis on grouping homogenous pixels. A minimum of

30 training datasets for each land class were selected for the supervised classification. The same selection process was replicated to ensure the classification map accurately represented the conditions depicted in the satellite imagery. Manual editing was performed in the post-classification process using ArcMap 10.5 software to minimize misclassification errors. This process was applied to produce the forest cover maps for 2015 and 2020. The overlay procedure in ArcMap Software was applied to compute and quantify the changed area from one land class to another during the five-year study.

Table 2: Description of classified Land Use and Land Cover (LULC) classes

LULC Classes	Description
Closed Forest	Land spanning more than 0.5 hectares with trees higher than 5 meters and a canopy cover of more than 40 percent, or trees able to reach these thresholds in situ.
Open Forest	Land spanning more than 0.5 hectares with trees higher than 5 meters and a canopy cover between 10 and 40 percent, or trees able to reach these thresholds in situ. It includes pure bamboo forest and bamboo forest with trees.
Otherwooded Land	Land spanning more than 0.5 hectares with trees higher than 5 meters and a canopy cover of 5-10 percent, or trees able to reach these thresholds in situ, or with a combined cover of shrubs, bushes and trees above 10 percent. It does not include land that is predominantly under agricultural use.
Other Land	Land that is not classified as forest or otherwooded land (rocks, bareland, sandbanks, clear-cutting land for plantation)
Road	Land that is used for transportation (car road)
Water	Creeks, streams, ponds, lakes, dams and reservoirs

3.4.Accuracy Assessment

Accuracy assessment is the evaluation of the accuracy or correctness of the classified image by comparing with reference data that are assumed to be true (Story & Congalton, 1986). Google Earth Pro and Sentinel-2 satellite image were used to conduct accuracy assessment as reference data. The accuracy assessment of the classified maps was performed using the stratified random sampling method. Each land class was considered as a separate stratum. The adequate number of sampling points for the accuracy assessment is essential to reflect the actual ground condition of the study area. For stratified random sampling, the sample size was determined using the equation provided by Cochran (1977) to cover the research area (Olofsson et al., 2014):

$$N = \left(\sum_{i=1} (W_i * S_i)/S_o\right)^2 \tag{1}$$

where W_i is mapped area proportion of class i; S_i is the standard deviation of stratum i; and S_o is the expected standard deviation of overall accuracy.

3.5.Exploring the Determinants of Deforestation through Binary Logistic Regression

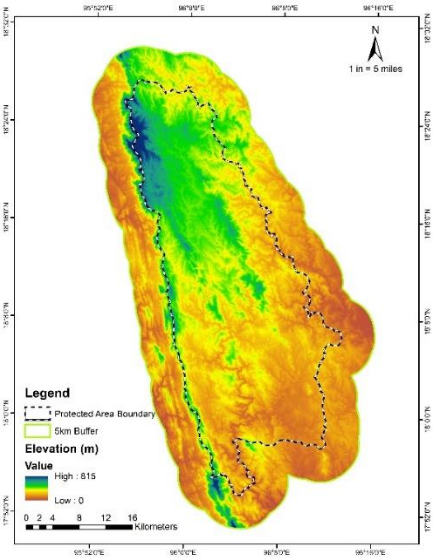
Binary logistic regression is a statistical method used to analyze problems with one or more independent variables determining a binary outcome (Menard, 2002; Pampel, 2020). The binary logistic regression model is a widely recognized parametric model in the fields of deforestation probability mapping, statistical analysis, spatial modeling, and machine learning. The binary logistic regression model suits situations where the dependent variable takes on binary values (Atkinson & Massari, 1998). This study utilized a spatial dataset of forest cover changes as the dependent variable, while proximity and biophysical factors were used as independent variables to identify factors that affect deforestation and to predict the probability of deforestation.

The spatial dataset of changes in land cover over the study period was utilized as the dependent variable in the binary logistic regression model. The land cover classification maps from 2015 and 2020 were modified to conduct binary logistic regression analysis. The categories 'Closed forest', 'Open forest' and "Otherwooded land" were merged into a single category called 'forest zones'. Similarly, the categories 'Other Land', 'Cultivated land', 'Water' and 'Road' were grouped together as 'non-forest zones'. After conducting change detection on two maps, a raster dataset (pixel size of 10m×10m) with four land cover classes, showing changes from forest to non-forest, forest to forest, non-forest to non-forest, and non-forest to forest was created. In this study, only deforestation was considered, and thus, 'non-forest to non-forest' class and 'non-forest to forest' (forest increase) were excluded. The final deforestation map comprised only two categories: 'forest to forest' (indicating no deforestation) and 'forest to non-forest' (indicating deforestation). Forest to non-forest (deforestation) was coded as “1”, and forest to forest (no deforestation) was coded as “0” as the dependent variable. In contrast, elevation, slope, distance from roads, distance from streams and distance from 5km buffer were used as independent variables to predict the likelihood of deforestation. The five independent variables from GIS layers were used as raster datasets with a pixel size of 10m x 10m.

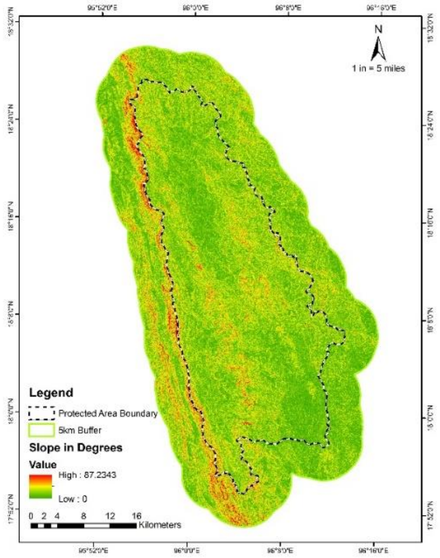
The sources of data and data format for each independent and dependent variable are described in Table 3. The raster dataset for each independent variable is shown in Figure 2.

Table 3: Descriptions of Variables

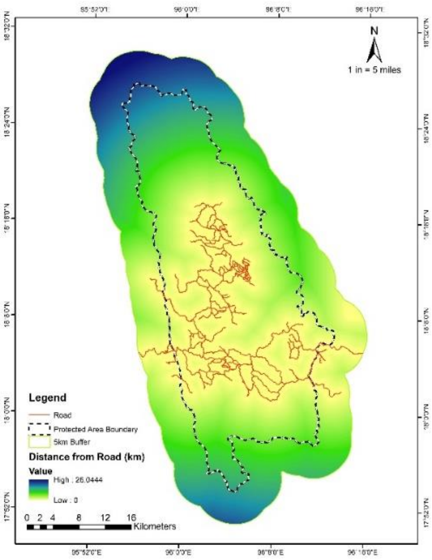
Variables	Data Format	Source of Data
(1) Elevation	Raster grid	Digital Elevation Model from USGS https://earthexplorer.usgs.gov/
(2) Slope	Raster grid	Digital Elevation Model from USGS https://earthexplorer.usgs.gov/
(3) Distance from roads (Euclidean distance)	Polyline	Digitized from topographic maps 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)
(4) Distance from streams (Euclidean distance)	Polyline	Digitized from topographic maps of compiled from 1:50,000 aerial photographs (Survey Department, Myanmar)
(5) Distance from 5km buffer (Euclidean distance)	Polygon	Generated from 5 km buffer around Wildlife Sanctuary
(6) 2015 and 2020 land cover maps	Raster grid	Sentinel 2A



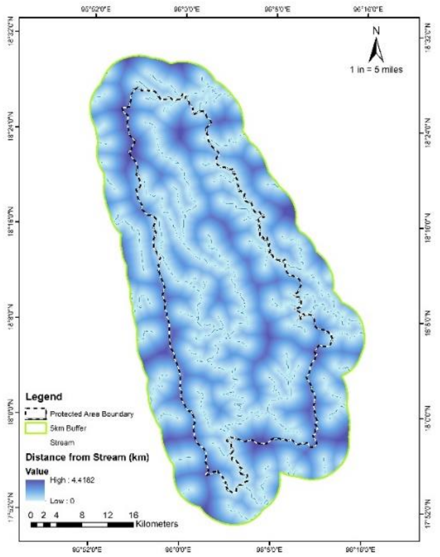
Elevation



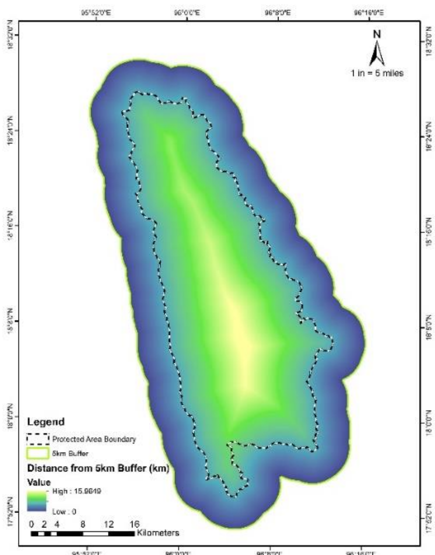
Slope



Distance from road



Distance from stream



Distance from 5km Buffer

Figure 2: Raster Datasets of Independent Variables

A total of 1600 sampling points were produced using Arc GIS across the deforestation and no-deforestation classes. Collinearity issues of deforestation determining factors were evaluated through VIF and Tolerance. Subsequently, 70% of the points were randomly selected for the training dataset, while the remaining 30% were kept for validating the model. Using R software version 2022.02.0, binary logistic regression analysis was performed. Binary logistic regression model is as follows.

$$y = a + b_1x_1 + b_2x_2 + b_3x_3 + \cdots b_mx_m \tag{2}$$

$$y = log_e\left[\frac{P}{1-P}\right] = logit(P) \tag{3}$$

$$P = \frac{e^y}{1+e^y} \tag{4}$$

The model that employs independent variables denoted by x_1,x_2 and x_3, paired with coefficients b_1, $b_2$and b_3 , a, stands for constant value. The probability $'P'$ is expressed by $'y'$, which represents the $logit(P)$. The model was validated using the TSS, Kappa, efficiency, and ROC curve.

3.6.Validation Methods for Deforestation Probability

The Receiver Operating Characteristics (ROC) is a widely used method to evaluate the performance of a binary classifier (Hanley & McNeil, 1982). It is a graphical representation that displays the relationship between the True Positive Rate (TPR) or sensitivity values and the False Positive Rate (FPR) of the classifier.

The TPR represents the proportion of actual positive cases that are correctly identified by the classifier, while the FPR represents the proportion of negative cases that are incorrectly classified as positive. The following equations were used to compute the TPR and FPR:

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{6}$$

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).

Furthermore, to assess the strength of the model, efficiency (E) and true skill statistics (TSS) were employed. 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:

$$\text{E} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{7}$$

True Skill Statistics (TSS) is a performance metric used to evaluate the accuracy of a model's predictions (Allouche et al., 2006). TSS is calculated as the difference between the hit rate (sensitivity) and the false alarm rate.

$$\text{TSS} = \text{TPR} - \text{FPR} \tag{8}$$

Cohen's kappa index is a widely used statistical measure for evaluating the accuracy, consistency and efficiency of predictive models. The kappa value ranges from -1 to +1, where a value of +1 indicates perfect agreement and a value of zero or less indicates a performance no better than random chance (Cohen, 1960). Kappa index is calculated as follows:

$$\mathrm{k} = \frac{P_{obs} - P_{exp}}{1 - P_{exp}} \tag{9}$$

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

$$P_{obs} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{N}} \tag{10}$$

$$P_{exp} = \frac{(\mathrm{TP} + \mathrm{FN})(\mathrm{TP} + \mathrm{FP}) + (\mathrm{FP} + \mathrm{TN})(\mathrm{FN} + \mathrm{TN})}{N^2} \tag{11}$$

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).

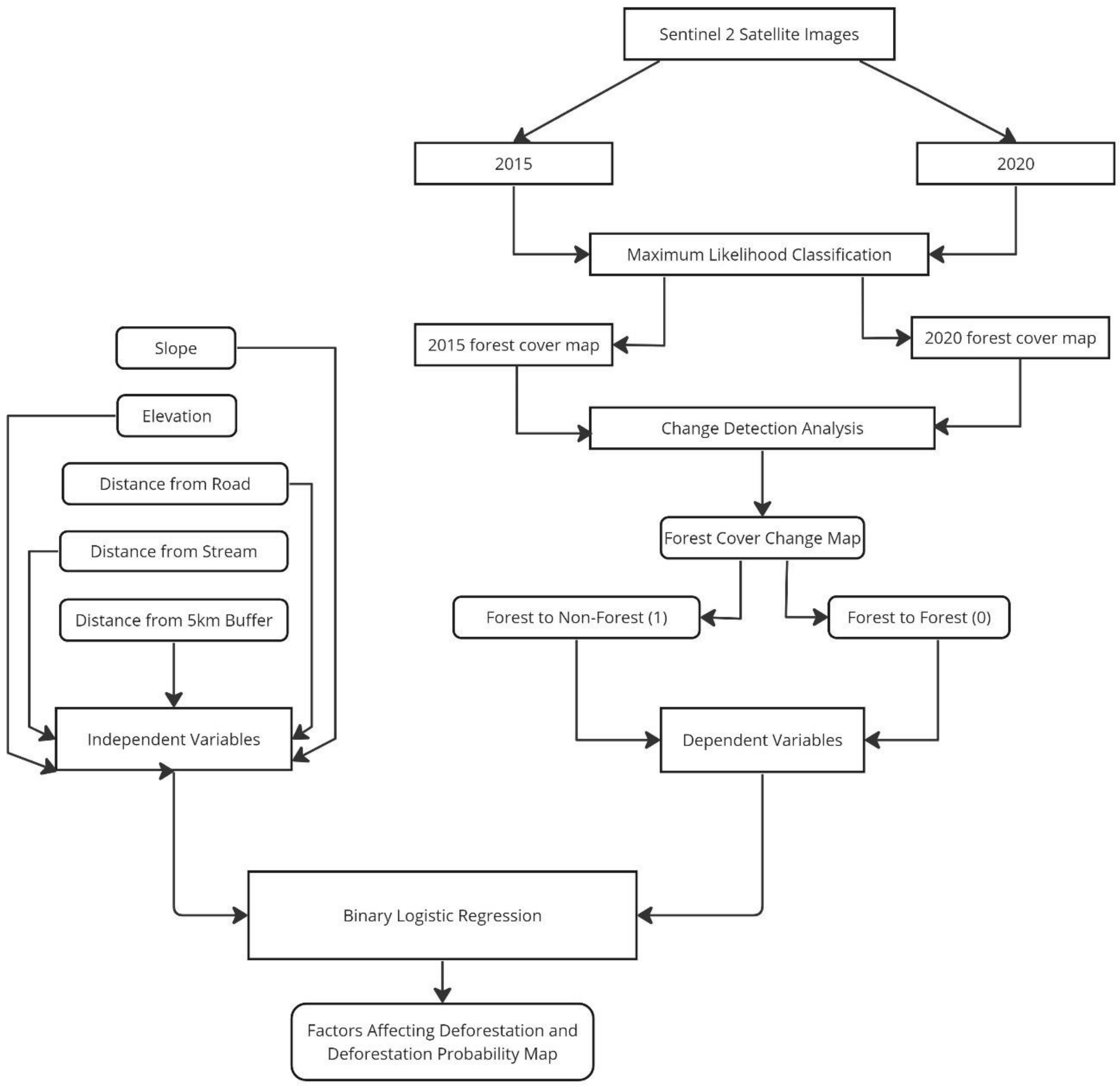


Figure 3: Overall Methodological Framework of the study

4. Results

4.1. Accuracy Assessment of the Study Area

The overall accuracy, Kappa coefficient, user accuracy (commission error) and producer accuracy (omission error) of the classification for the reference periods were described in Table 4 and 5. The overall accuracies of the research area were observed 88.65% in 2015 and 89.61% in 2020, indicating a fulfillment of the minimum 85% accuracy set by Anderson's classification scheme (Anderson, 1976). Moreover, the corresponding Kappa Coefficients for two different study periods were all greater than 0.81 for different study years, accounting for 0.81 (81.04%) in 2005, and 0.82 (82.08%) in 2020. This indicated that there was a strong agreement between the classified data and the reference data for all periods.

Table 4: Accuracy assessment of 2015 Land Cover in the study area

LAND CLASS	Closed Forest	Open Forest	Other Land	Otherwooded Land	Cultivated Land	Water	Road	Row Total
Closed Forest	142	7	0	0	1	0	0	150
Open Forest	33	222	0	3	0	0	0	258
Other Land	0	0	9	1	0	0	0	10
Otherwood-ed Land	1	4	0	14	0	0	0	19
Cultivated Land	0	0	0	1	9	0	0	10
Water	0	0	0	1	0	9	0	10
Road	0	1	0	0	0	0	9	10
Column Total	176	234	9	20	10	9	9	467
Overall Accuracy = 88.65%			**Kappa Coefficient = 81.04%**					

Producer's Accuracy (omission error)				User's Accuracy (commission error)			
Closed Forest	= 80.68%	19.32%	omission error	Closed Forest	= 94.67%	5.33%	commission error
Open Forest	= 94.87%	5.13%	omission error	Open Forest	= 86.05%	13.95%	commission error
Other Land	= 100%	0.00%	omission error	Other Land	= 90%	10%	commission error
Otherwood-ed Land	= 70%	30.00%	omission error	Otherwooded Land	= 73.68%	26.32%	commission error
Cultivated Land	= 90%	10.00%	omission error	Cultivated Land	= 90%	10%	commission error
Water	= 100%	0.00%	omission error	Water	= 90%	10%	commission error
Road	= 100%	0.00%	omission error	Road	= 90%	10%	commission error

Table 5: Accuracy assessment of 2020 Land Cover in the study area

LAND CLASS	Closed Forest	Open Forest	Other Land	Otherwooded Land	Cultivated Land	Water	Road	Row Total
Closed Forest	96	14	0	1	0	0	0	111
Open Forest	17	253	0	1	1	1	0	273
Other Land	0	1	8	1	0	0	0	10
Otherwood-ed Land	0	8	0	30	0	0	0	38
Cultivated Land	0	0	0	1	9	0	0	10
Water	1	0	0	0	0	9	0	10
Road	0	0	0	1	0	0	9	10
Column Total	114	276	8	35	10	10	9	462
Overall Accuracy = 89.61%			**Kappa Coefficient = 82.08%**					

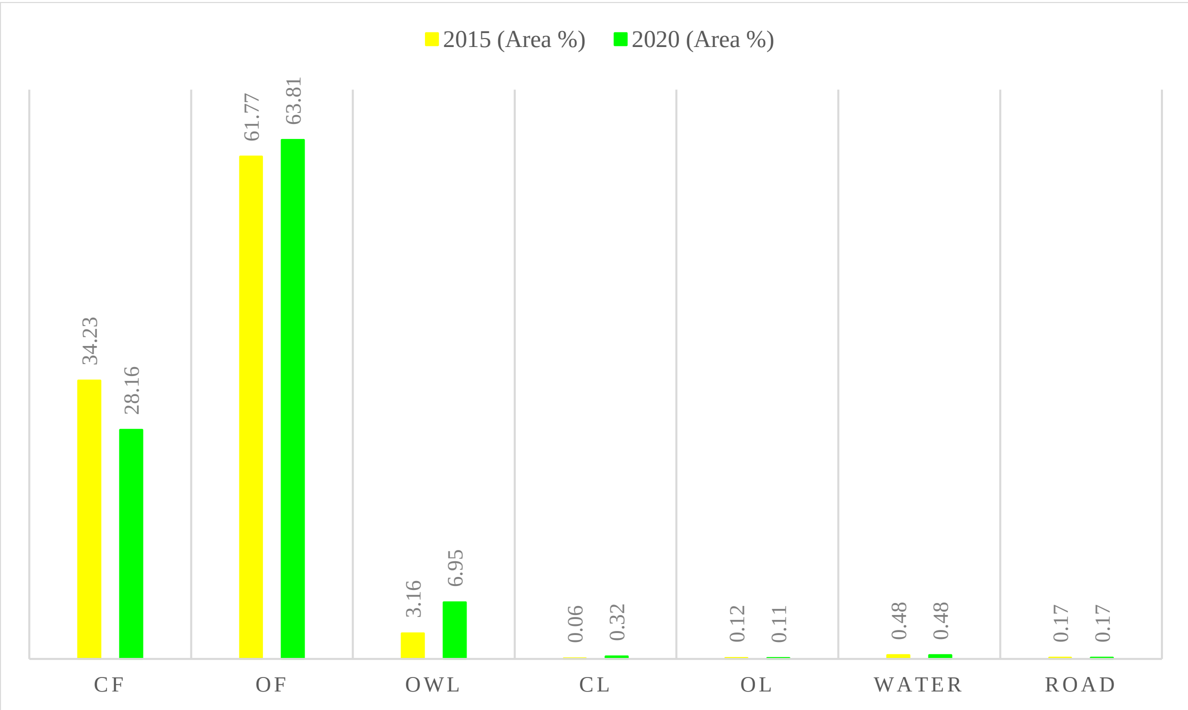
Producer's Accuracy (omission error)				User's Accuracy (commission error)			
Closed Forest	= 84.21%	15.79%	omission error	Closed Forest	= 86.49%	13.51%	commission error
Open Forest	= 91.67%	8.33%	omission error	Open Forest	= 92.67%	7.33%	commission error
Other Land	= 100%	0.00%	omission error	Other Land	= 80%	20%	commission error
Otherwood-ed Land	= 85.71%	14.29%	omission error	Otherwooded Land	= 78.95%	21.05%	commission error
Cultivated Land	= 90%	10.00%	omission error	Cultivated Land	= 90%	10%	commission error
Water	= 90%	10.00%	omission error	Water	= 90%	10%	commission error
Road	= 100%	0.00%	omission error	Road	= 90%	10%	commission error

4.2.Forest Cover Status within North Zamari Wildlife Sanctuary

Forest cover assessment was conducted into two parts: North Zamari Wildlife Sanctuary (within the protected area) and its adjacent 5 km buffer area (outside the protected area). Table 6 and Figure 4 summarize the condition of forest cover in North Zamari WS from 2015 to 2020. The spatial distribution of forest cover in the wildlife sanctuary for the two different periods is described in Figure 5. The dominant land cover of the study area is forest cover because it is a protected area for the habitat of wild animals, particularly elephants. In 2015, open forest area occupied the highest percentage in the protected area, covering 61.77% of the total land area of North Zamari WS. The second dominated land cover was closed forest (34.23%), followed by otherwooded land (3.16%), water (0.48%), road (0.48%), other land (0.12%), and cultivated land (0.06%). In 2020, the largest land cover was dominated by open forest (63.81%), followed by closed forest (28.16%), otherwooded land (6.95%), water (0.48%), cultivated land (0.32%), road (0.17), and other land (0.11%). In the study period (2015-2020), the increasing trend was observed in open forest (2.04%), otherwooded land (3.79%), and cultivated land (0.26%), while the decreasing trend was experienced in closed forest (6.07%) and other land (0.01%). Water and road showed stable percentage from 2015 to 2020. The forest area, including closed forest, open forest, and otherwooded land, decreased from 99.16% (97496.36 ha) of forest area to 98.92% (97256.22 ha) of the protected area between 2015 and 2020 (Table 6) (Figure 4). In other words, 0.24% (240.14 ha) of the total forest area was converted into other types of land cover in the North Zamari WS and the net annual rate of deforestation in the protected area was 0.05% (48.03 ha) during the five years.

Table 6: Forest Cover Status and Change within North Zamari WS in 2015 and 2020

Land Cover	**2015**		**2020**		**Change (2015-2020)**	
	Area (ha)	%	Area (ha)	%	Area (ha)	%
Closed Forest	33654.95	34.23	27686.92	28.16	-5968.03	-6.07
Open Forest	60731.08	61.77	62735.07	63.81	2003.99	2.04
Otherwooded Land	3110.33	3.16	6834.23	6.95	3723.9	3.79
Cultivated Land	60.28	0.06	314.55	0.32	254.27	0.26
Other Land	118.74	0.12	104.6	0.11	-14.14	-0.01
Water	468.42	0.48	468.44	0.48	0.02	0
Road	170.2	0.17	170.2	0.17	0	0
Total	98314	100	98314	100		



(CF, OF, OWL, CL, OL refers to Closed Forest, Open Forest, Otherwooded Land, Cultivated Land, and Other Land.)

Figure 4: Forest Cover Status within North Zamari WS in 2015 and 2020

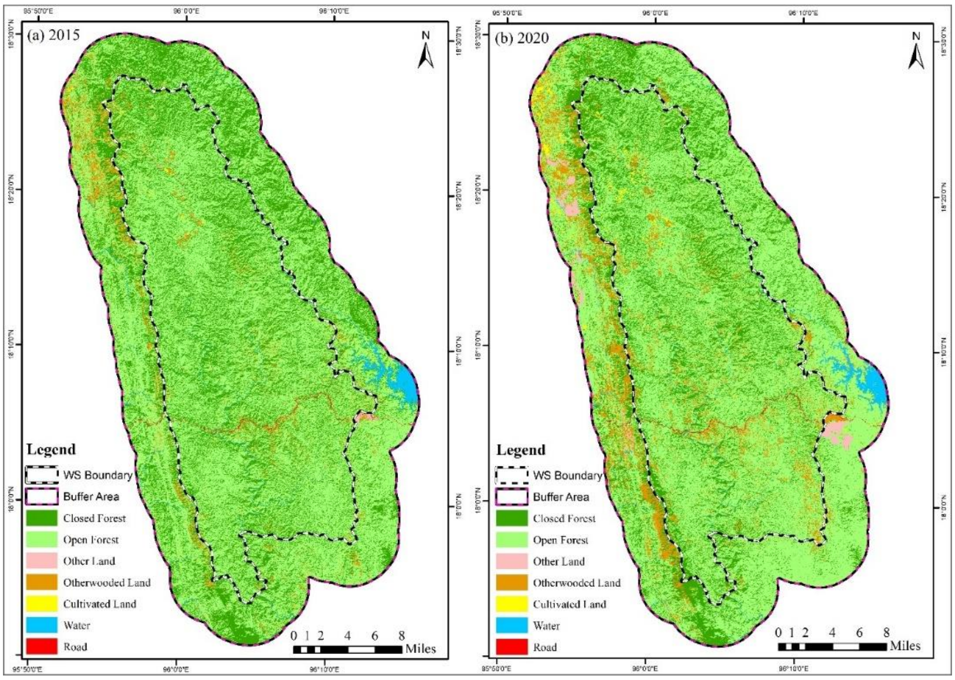


Figure 5: Forest Cover Map within North Zamari WS and 5 km Buffer Area

4.3.Forest Cover Change within North Zamari Wildlife Sanctuary

Figure 6 and Table 7 show the changes in forest cover within the North Zamari WS and the 5 km buffer area. The vast majority of the area within the North Zamari WS remained as forest area (Forest to Forest), accounting for 98.84% (97,177.07 ha) of the sanctuary area. There was a very small fraction of deforestation area (Forest to Non-Forest) in the WS, which was 0.33% (322.8 ha). On the other hand, reforestation area (Non-Forest to Forest) was observed, representing 0.08% (82.68 ha) of the area. The area that remained as non-forest area was 0.74% (731.46 ha) over the study period.

The change matrix cross-tabulation (Table 8) shows the comparable table of the total area of land classes changed with the corresponding percentages in North Zamari WS from 2015-2020. During the study period, no changes of water and road areas were observed, as 100% remained intact, followed by open forest at 80.71%, other land at 71.21%, otherwooded land at 67.91%, closed forest at 59.61%, and cultivated land at 19.56%. Most of the closed forest was converted to open forest (38.90%), otherwooded land (1.33%), cultivated land (0.15%), and other land (0.01%). Open forest was converted to closed forest (11.91%) of the total area, otherwooded (6.98%), cultivated land (0.38%), and other land (0.02%). Most of the otherwooded land was changed into open forest (31.16%), cultivated land (0.81%), and other land (0.12%). Cultivated land (48.49 ha) and other land (34.19 ha) were converted to otherwooded land between 2015 and 2020.

Table 7: Forest Cover Change within North Zamari WS and 5 km Buffer Area in 2015 and 2020

Change	North Zamari WS		5 km Buffer Area	
	Area (ha)	Area (%)	Area (ha)	Area (%)
Forest to Forest (F - F)	97177.07	98.84	88248.6	95.23
Forest to Non-Forest (F - N)	322.8	0.33	2149.3	2.32
Non-Forest to Forest (N - F)	82.68	0.08	140.73	0.15
Non-Forest to Non-Forest (N - N)	731.46	0.74	2129.37	2.30

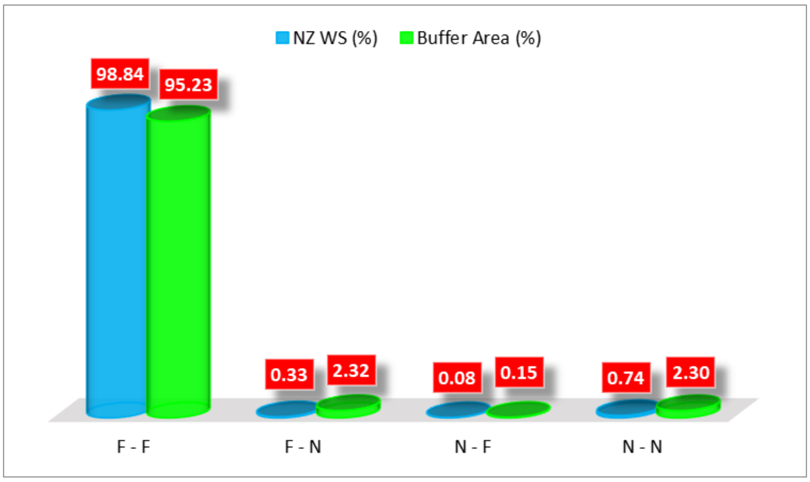


Figure 6: Forest Cover Change within North Zamari WS and 5 km Buffer Area in 2015 and 2020

Table 8: Transition Matrix showing Land Cover Change in North Zamari WS in 2015 and 2020

LULC Classes	Unit	CF	OF	OWL	CL	OL	Water	Road
CF	ha	**20063.26**	13091.65	447.89	49.60	3.79	0.00	0.00
	%	**59.61**	38.90	1.33	0.15	0.01	0	0
OF	ha	7232.62	**49019.08**	4241.02	227.92	12.60	0.00	0.00
	%	11.91	**80.71**	6.98	0.38	0.02	0	0
OWL	ha	0	969.12	**2112.43**	25.24	3.65	0.00	0.00
	%	0	31.16	**67.91**	0.81	0.12	0	0
CL	ha	0	0	48.49	**11.79**	0	0	0
	%	0	0	80.44	**19.56**	0	0	0
OL	ha	0	0	34.19	0	**84.56**	0	0
	%	0	0	28.79	0	**71.21**	0	0
Water	ha	0	0	0	0	0	**464.91**	0
	%	0	0	0	0	0	**100**	0
Road	ha	0	0	0	0	0	0	**170.20**
	%	0	0	0	0	0	0	**100**

(CF, OF, OWL, CL, OL refers to Closed Forest, Open Forest, Otherwooded Land, Cultivated Land, and Other Land. The bold numbers on the diagonal are unchanged land area with the respective percentages, and the others are the changed area from one land class to another.)

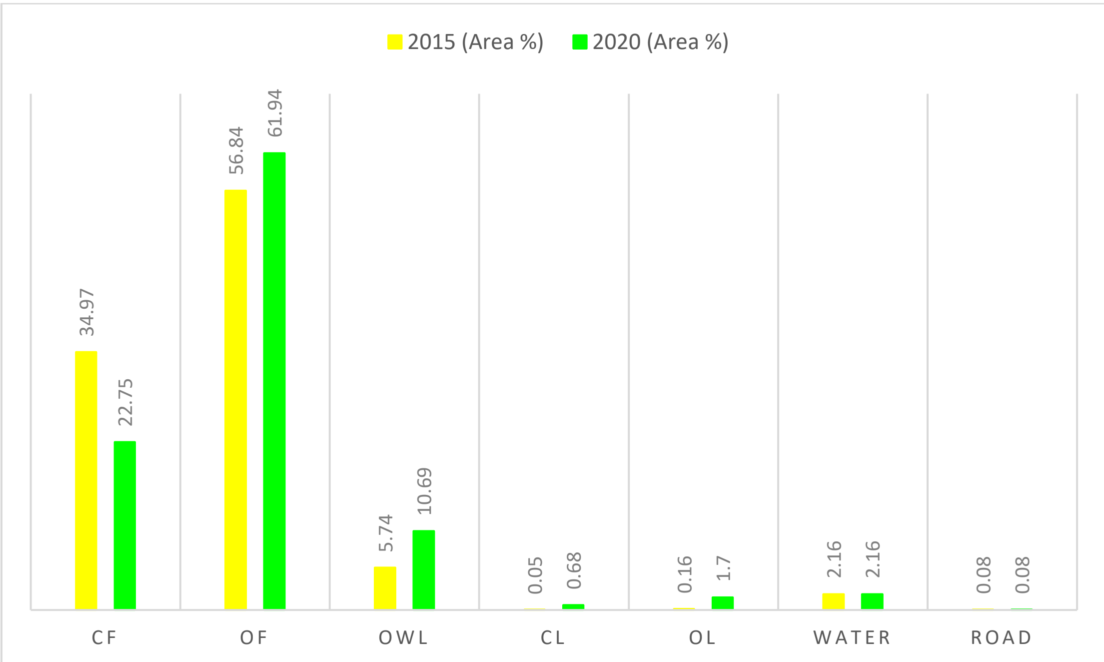
4.4.Forest Cover Status in 5 km Buffer Area

Assessment of forest cover was conducted outside the protected area (5 km buffer area) in order to assess forest condition of the surrounding area of the North Zamari Wildlife Sanctuary. Table 9 and Figure 7 summarized forest and land cover status within 5 km buffer area in 2015 and 2020. Seven types of land classes, namely closed forest, open forest, otherwooded land, cultivated land, other land, water, and raod, were classified based on the actual condition on the ground. Forest is the main land cover type within the 5 km buffer area. In 2015, the largest dominant land cover was open forest (56.84%) of the total 5 km buffer area, followed by closed forest (34.97%), otherwooded land (5.74%), water (2.16%), other land

(0.16%), road (0.08%), and cultivated land (0.05%). In 2020, open forest (61.94%), closed forest (22.75%), and otherwooded land (10.69%) occupied the three largest land cover types in the same buffer area, followed by water (2.16%), other land (1.7%), cultivated land (0.68%), and road (0.08%). The trend changes for all land classes were experienced in the buffer area, except water and road. Areas of open forest, otherwooded land, cultivated land, and other land increased during the study period. On the other hand, a significant decrease in closed forest was observed in the buffer zone, ranging from 34.97% in 2015 to 22.75% in 2020 (Table 9) (Figure 7). No changed area of water (2.16%) and road (0.08%) were occurred in the same period. During the reference period, 2009.37 ha the total forest area, comprising closed forest, open forest and otherwooded land, were changed into other classes. The annual deforestation rate in the buffer area was 0.44% (401.87 ha) of the forest area in the study area between 2015 and 2020.

Table 9: Forest Cover Status and Change within 5 km Buffer Area in 2015 and 2020

Land Class	**2015**		**2020**		**Change (2015-2020)**	
	Area (ha)	%	Area (ha)	%	Area (ha)	%
Closed Forest	32405.62	34.97	21079.08	22.75	-11326.54	-12.22
Open Forest	52675.9	56.84	57400.16	61.94	4724.26	5.1
Otherwooded Land	5316.39	5.74	9909.3	10.69	4592.91	4.95
Cultivated Land	43.18	0.05	630.1	0.68	586.92	0.63
Other Land	149.32	0.16	1570.97	1.7	1421.65	1.54
Water	2000.46	2.16	2001.25	2.16	0.79	0
Road	77.13	0.08	77.14	0.08	0.01	0
Total	92668	100	92668	100		



(CF, OF, OWL, CL, OL refers to Closed Forest, Open Forest, Otherwooded Land, Cultivated Land, and Other Land.)

Figure 7: Forest Cover Status within 5 km Buffer Area in 2015 and 2020

4.5.Forest Cover Change in 5 km Buffer Area

The area that remained forest (Forest to Forest) was 95.23% (88248.6 ha) of the 5 km buffer area, while non-forest area (Non-Forest to Non-Forest) was stable at 2.30% (2129.37 ha) during the period of 2015-2020 (Table 7) (Figure 6). The deforestation area (Forest to Non-Forest) was occurred in the buffer zone, accounting for 2.32% (2149.3 ha). Only a very small fraction of the area experienced reforestation (Non-Forest to Forest), which was 0.15% (140.73 ha).

Table 10 describes the cross-tabulation change matrix for the changed areas and their corresponding percentages from one LULC class to another in comparison with the total area of each LULC class from 2015-2020. During the study period, 100% of water and road remained unchanged, followed by cultivated land (85.80%), open forest (83.03%), closed forest (59.90%), otherwooded land (54.56%), and other land (9.86%). The majority of closed forest were converted to open forest (36.75%) (11909.97 ha), and the rest to otherwooded land (2.56%) (830.96 ha), cultivated land (0.14%) (46.61 ha), and other land (0.64%) (207.37 ha). Most of the open forest were converted to otherwooded land (11.47%) (6044.27 ha), followed by closed forest (2.79%) (1468.88 ha), other land (2.05%) (1079.53 ha), and cultivated land (0.66%) (346.89 ha). Otherwooded land was converted to open forest (36.62%) (1946.74 ha), other land (5.07%) (269.35 ha) and cultivated land (3.75%) (199.55 ha). Cultivated land (6.13 ha) and other land (134.6 ha) were converted to otherwooded land.

Table 10: Transition Matrix showing Land Cover Change in 5 km Buffer Area in 2015-2020

LULC Classes	Unit	CF	OF	OWL	CL	OL	Water	Road
CF	ha	**19410.7**	11909.97	830.96	46.61	207.37	0	0
	%	**59.90**	36.75	2.56	0.14	0.64	0	0
OF	ha	1468.88	**43736.3**	6044.27	346.89	1079.53	0	0
	%	2.79	**83.03**	11.47	0.66	2.05	0	0
OWL	ha	0	1946.74	**2900.76**	199.55	269.35	0	0
	%	0	36.62	**54.56**	3.75	5.07	0	0
CL	ha	0	0	6.13	**37.05**	0	0	0
	%	0	0	14.20	**85.80**	0	0	0
OL	ha	0	0	134.6	0	**14.72**	0	0
	%	0	0	90.14	0	**9.86**	0	0
Water	ha	0	0	0	0	0	**2000.47**	0
	%	0	0	0	0	0	**100**	0
Road	ha	0	0	0	0	0	0	**77.13**
	%	0	0	0	0	0	0	**100**

(CF, OF, OWL, CL, OL refers to Closed Forest, Open Forest, Otherwooded Land, Cultivated Land, and Other Land. The bold numbers on the diagonal are unchanged land area with the respective percentages, and the others are the changed area from one land class to another.)

4.6.Assessing Deforestation Risks using Binary Logistic Regression Analysis

The collinearity statistical test is performed to assess the variance inflation factor (VIF) and tolerance for validating the reliability of independent variables for inclusion in a logistic regression model. If tolerance value < 0.1 and VIF value >5, it indicates the presence of a multicollinearity problem in the dataset (Johnston et al., 2018). There is no presence of multicollinearity among the independent variables, as indicated by the tolerance value > 0.1 and VIF <5 (Table 11).

Table 11: Collinearity Statistics of Different Independent Variables

Variables	Collinearity Tolerance	Statistics VIF
Slope	0.964	1.039
Distance from Stream	0.843	1.187
Distance from Road	0.529	1.887
Distance from 5km Buffer	0.342	2.919
Elevation	0.504	1.985

The correlation matrix displays the correlation coefficients between independent variables to identify multicollinearity. If the value of the simple correlation coefficient exceeds 0.8 or 0.9, it indicates a significant issue of multicollinearity between the datasets. There are no serious concerns regarding the presence of multicollinearity among the variables.

Table 12: Correlation matrix showing the relationship among independent variables using Pearson's correlation coefficient method

Variables	Slope	Distance from Road	Distance from Stream	Distance from 5km Buffer	Elevation
Slope	1	-0.128	0.102	0.045	0.107
Distance from Road	-0.128	1	-0.095	-0.645	-0.284
Distance from Stream	0.102	-0.095	1	0.007	0.277
Distance from 5km Buffer	0.045	-0.645	0.007	1	0.619
Elevation	0.107	-0.284	0.277	0.619	1

Correlation Matrix of Independent Variables

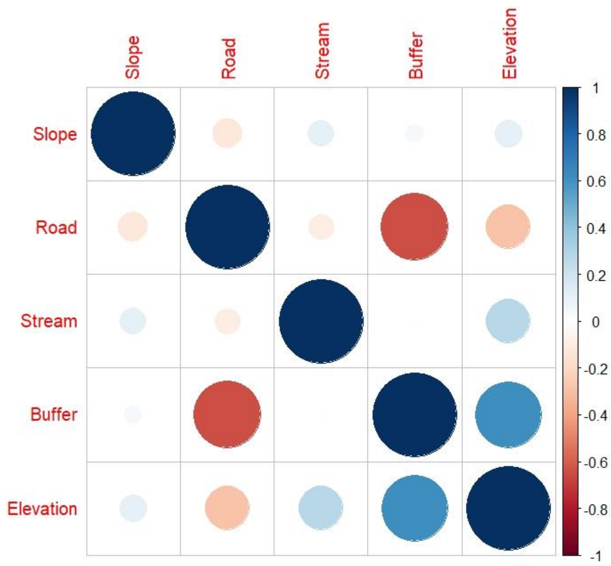


Figure 8: Correlation among the independent variables which reflect by a color matrix scale

The variable "Distance from Streams" showed a significant negative coefficient (B = -0.554) and a highly significant p-value (p < 0.001), showing a significant negative relationship between the probability of deforestation and the proximity to streams. The interpretation of the odds ratio (Exp(B) = 0.575) indicates that as distance from streams increases by one kilometer, the probability of deforestation decreases by 42.5%. Therefore, it can be inferred that areas closer to streams are more susceptible to deforestation than those located further away.

The variable "Distance from 5km Buffer" revealed that there was no significant relationship (B = -0.027, p = .416) with an odds ratio of 0.974. Hence, it can be inferred that the distance to the 5km buffer did not have a substantial impact on the likelihood of deforestation within the range of data examined (Exp(B) = 0.974). The negative sign of the estimate suggests a negative relationship, indicating that the chance of deforestation very slightly decreases with increasing distance from the buffer zone. The high p-value, however, suggests that this result is not statistically significant and may just be the result of chance rather than a real influence.

The variable 'Elevation' shows a significant positive coefficient (B = 0.05, p < 0.001) with an odds ratio of Exp(B) = 1.05. This finding suggests that as elevation increases, so does the likelihood of deforestation. Specifically, the probability of deforestation increases by 5% for each unit increase in elevation.

Table 13: Results of Binary Logistic Regression Analysis

Independent Variables	B	S.E	Z value	Pr(>|z|)	Exp (B)	95% C.I for Exp (B) Lower	Upper
Intercept	-0.808	0.296	-2.733	0.006	0.446	0.248	0.793
Slope	-0.176	0.017	-10.139	2e-16 ***	0.839	0.809	0.867
Distance from Road	0.174	0.015	11.777	2e-16 ***	1.189	1.157	1.226
Distance from Streams	-0.554	0.114	-4.843	1.28e-06***	0.575	0.458	0.717
Distance from 5km Buffer	-0.027	0.033	-0.813	0.41612	0.974	0.913	1.038
Elevation	0.005	0.001	5.289	1.23e-07***	1.005	1.003	1.006

Note: Signif. codes: '***' 0.001, AIC: 1040.8. Area under the curve AUC: 0.8596, Efficiency:0.8106

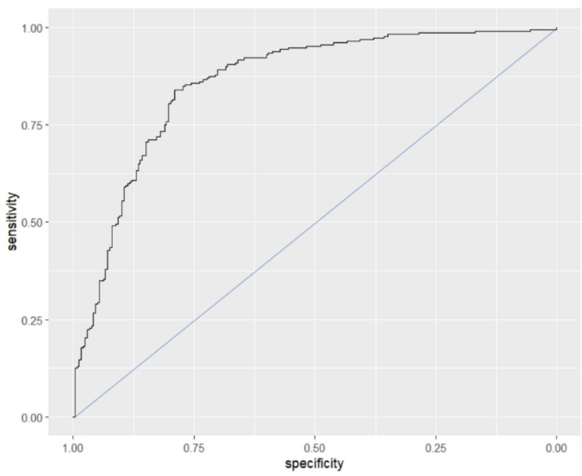


Figure 9: ROC curve for binary logistic regression model

Area under the curve: 0.8596

The assessment of the binary logistic regression model's accuracy and reliability was conducted using a 30% testing dataset. The graphical representation of the receiver operating characteristic curve (ROC) involves the Y-axis representing the True Positive Rate (TPR) or sensitivity values and the X-axis representing the specificity values. This curve illustrates the classification model's performance across all thresholds of classification. The values of efficiency (E), true skill statistics (TSS), false positive rate (FPR), and Kappa coefficient of binary logistic regression are presented in Table 14.

With a True Positive Rate (TPR) of 0.8491, or sensitivity, the algorithm effectively detects 84.91% of deforestation events. In contrast, the model's ability to incorrectly categorize locations as having deforestation (false positives) around 22.69% of the time when they are free of deforestation is shown by the FPR (false positive rate) value of 0.2269. The True Skill Statistics (TSS) measure considers both omission and commission errors, as well as success due to random guessing. It is scaled from -1 to +1, with +1 indicating perfect agreement and values below 0.4 indicating poor model discrimination. Values of zero or less indicate a performance that is no better than random. A TSS value of 0.6222 indicates that the model shows a moderate to good capability in distinguishing between positive cases (deforestation) and negative cases (no deforestation).

The Kappa coefficient, with a value of 0.6216, measures the agreement between the model's predictions and the observed data, considering the agreement that could happen by chance. The kappa co-efficient is 0.6216, which is also substantial or very good accuracy. In addition, the model's predictive ability is measured by the Efficiency, which is reported as 0.8106. This represents the proportion of true results among the total number of cases examined, including both true positives and true negatives. Based on the generated AUC curve, the AUC value is 0.8596, suggesting that the model can distinguish between deforestation and non-deforestation classes. A higher AUC value indicates a stronger level of predictive accuracy.

Table 14: Values of true positive rate (TPR), false positive rate (FPR), efficiency (E), true skill statistics (TSS) and Kappa co-efficient

Metrics	Binary Logistic Regression
TPR	0.8491
FPR	0.2269
Efficiency	0.8106
TSS	0.6222
Kappa	0.6216

Using the following formula, a deforestation probability map was generated based on the binary logistic regression model's computed value.

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

The probability value ranges from 0 to 1, with values closer to 0 indicating a very low likelihood of deforestation and values closer to 1 indicating a very high probability. The probability of deforestation inside the Wildlife Sanctuary was categorized into five categories: very low, low, moderate, high, and very high. These categories occupied 75.47%, 14.66%, 5.24%, 3.55%, and 1.08% of the forested areas respectively (Table 15). Based on the results, a significant proportion of the Wildlife Sanctuary, specifically 75.47% (73873.32 ha), falls within the 'Very Low' probability category. Conversely, the 'Low' probability category encompasses 14.66% (14349.56 ha) of the area and signifies a risk of deforestation that is also negligible. The identification of 'Very High' probability zones for deforestation risks within Wildlife Sanctuary, covering only 1.08% (1,053.56 ha) and requires cautious monitoring.

Table 15: Deforestation probability of forested areas within Wildlife Sanctuary

Deforestation probability zones	Binary Logistic Regression	
	Areas in ha	%
Very low	73873.32	75.47
Low	14349.56	14.66
Moderate	5133.67	5.24
High	3476.99	3.55
Very high	1053.56	1.08

The 5 km area surrounding the Wildlife Sanctuary was analyzed to identify deforestation risks. This area was also divided into 5 categories: very low, low, moderate, high, and very high. These categories occupied 60.09%, 19.35%, 8.55%,7.59% and 4.42% of the forested areas respectively (Table 16). Based on the results, specifically 60.09% (53907.86 ha), falls within the 'Very Low' probability category. Conversely, the 'Low' probability category encompasses 19.35% (17363.82 ha) of the area and signifies a risk of deforestation that is also negligible. In this buffer zone, 'Very High' probability areas are slightly more extensive than within the PA, covering 3,965.61 ha, which constitutes 4.42% of the buffer zone area.

Table 16: Deforestation probability of forested areas within 5km buffer area

Deforestation probability zones	Binary Logistic Regression	
	Areas in ha	%
Very low	53907.86	60.09
Low	17363.82	19.35
Moderate	7669.3	8.55
High	6807.28	7.59
Very high	3965.61	4.42

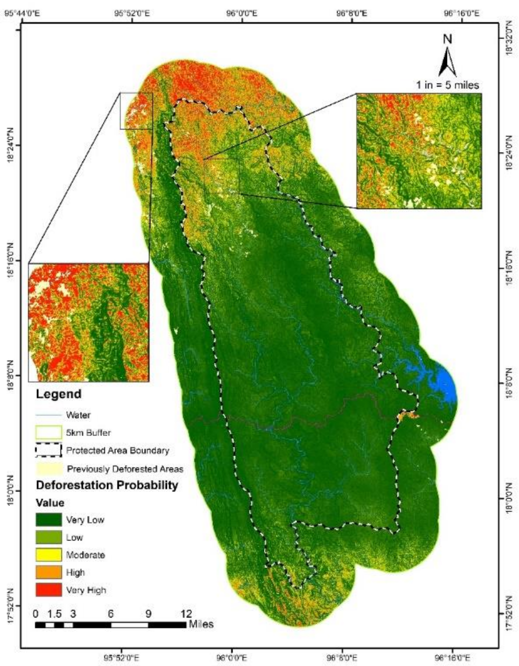


Figure 10: Map showing deforestation probability using binary logistic regression model

5. Discussion

5.1.Forest Cover Changes between 2015 and 2020

In this study, the main finding was that most of the closed forest was intensively converted to open forest and otherwooded land both in North Zamari WS and the surrounding 5 km buffer area between 2015 and 2016. In the same way, open forest was mostly changed into otherwooded land. This indicates that there was a serious degradation in the forest area of the research area. On the other hand, areas of closed forest, open forest and otherwooded land were significantly changed into cultivated land and other land not only in the North Zamari WS and but also in the 5 km buffer area. Thus, it was clearly found that North Zamari WS and its adjacent area experienced significant forest degradation and remarkable deforestation from 2015 to 2020. It was found in the study that forest cover and other type of land cover was dynamic and linear, meaning that the conversions follow similar patterns.

Anthropogenic activities such as illegal logging, charcoal making, hunting activities, forest fires and shifting cultivation, are the major courses of deforestation and forest degradation in the study area. The local people living near the study area traditionally and highly depend on natural resources, mainly forest resources, for their livelihoods. Before the North Zamari WS was gazetted in 2014, there was a two-lane road that passed through the sanctuary from east to west. The illegal timber log and charcoal were extracted through the road. Moreover, evidence of human activities such as illegal hunting, illegal logging and a widespread forest fire across the sanctuary, were detected in camera-traps (Hein et al., 2020), leading to forest degradation and deforestation in the North Zamari WS and its surrounding area. Thus, the results showed consistency with the previous findings.

On the other hand, agricultural expansion was a major driver of deforestation in Myanmar (Hasan et al., 2020; Naing Tun et al., 2021). In the past, upland rice was grown using shifting cultivation systems. Now, this traditional cultivation had already changed to short-fallow shifting cultivation (Aung et al., 2018), which contributed to rapid LULC dynamics. In the hilly region of Myanmar, shifting cultivation was one of the significant causes of deforestation and forest degradation (Aung et al., 2018). Shifting cultivation, which caused to change from forest to cultivated land, was a traditional framing system in Myanmar (Mon et al., 2009; Zin, 2005). In Myanmar, ethnic minority groups mainly live in the mountainous areas. According to the socioeconomic and cultural system of the groups, shifting cultivation is recognized as a major source of the rural economy in East and West of Bago Mountain Range (Billy et al., 2015). Moreover, a lot of local ethnic groups like Karen group residing in the reserved forests of the Bago Mountains usually conduct shifting cultivation for their livelihoods (Rosy & Billy, 2015). Thus, the conversion of forest into cultivated land and scrubland was observed in this study during the reference periods.

The North Zamari WS was established to preserve flora and fauna, especially to protect the endangered Asian elephant (MONREC, 2015). However, it was significantly experienced that there was a remarkable increase in cultivated land in the North Zamari WS and the surrounding 5 km buffer area. Cultivated land increased from 60.28 ha in 2015 to 314.55 ha in 2020 within the sanctuary, and 43.18 ha to 630.1 ha in buffer area. Cultivated land was mainly occurred in the northern part of the study area. Hein et al., (2020) reported that farming was observed in the northern peripheries of the North Zamari WS, and it may be a severe threat to habitat loss for the wildlife. Thus, this finding is consistent with the previous studies.

There are some limitations in the study to assess forest cover change using RS and GIS approach. Collection of ground-truthing information could not conducted for the classification accuracy due to the security reasons and the current political situations in Myanmar. Limited time period to assess forest cover change and write research paper is one of the limitations in this study. In the assessment of forest cover changes, good-quality satellite imageries are not enough available even in the growing season. It is challenging to differentiate closed forest, open forest and otherwooded land in some areas, especially mountainous areas, through the application of open assessed satellite imageries. On the other hand, forest canopy of mixed deciduous forest is naturally fallen the leaves in the opening season. Thus, some minor misclassifications may still have in the assessment of forest cover. In the validation process, Google Earth Pro cannot provide clear images for some areas, and no updated information could assess from it. Thus, it is recommended in further research that it is necessary to conduct image classification with the integration of field validation and high-resolution satellite imageries.

5.2.Driving factors of deforestation

The variables that had the most significant effect on deforestation in BLR, as determined by p values, were elevation, slope, and distances from roads and streams. The study's findings highlight that deforestation is less likely to occur in areas with steeper slopes because the difficulties and greater costs associated with clearing land on steep slopes for logging and agriculture. Linkie et al., (2004) found that forests on gentle slopes are more likely to be threatened because they are easy to access and can be used for cultivation and other land uses. Gayen and Saha (2018) conducted a study that offers further evidence affirming that the proper use of lower slopes for agricultural activities is an important factor of deforestation in certain areas. This finding supports the fundamental importance of terrain in influencing deforestation patterns, since lower slope areas are more prone to being transformed for agricultural purposes, resulting in a higher degree of forest cover decline.

The 'Distance from Roads' variable in the BLR model quantitatively represents the relation, suggesting that as the distance from roads increases, the probability of deforestation also increases. The study's results reveal an unexpected trend where locations that are further away from road infrastructure have a higher probability of deforestation. This pattern may be related to the limitations in accessibility and the limited regulatory monitoring in remote areas, increasing vulnerability to illegal deforestation cases. The road data included in this study were digitized using topographic maps derived from 1:50,000 scale aerial photographs obtained from the Survey Department of Myanmar. Subsequently, visual interpretations obtained from Google Earth and Open Street Map were integrated into these maps. However, not all existing paths may be captured by Google Earth satellite imagery, particularly those that are narrow rarely passed or covered by canopy cover. Furthermore, it should be noted that the Open Street Map depends on volunteer contributions and thus may not consistently provide the most current or complete collection of road data, especially in remote regions. Some insignificant pathways or less visible routes may have unintentionally been omitted throughout the digitization process, especially if they were unclear or indistinguishable in the employed images. These constraints highlight the potential for underestimating the existence of roads in less accessible regions, which might affect the interpretation of the relationship between the 'Distance from Roads' variable and deforestation.

The variable "distance from streams" indicates that areas near streams are more susceptible to deforestation. The availability of water is essential for both human settlement and agricultural activities. Consequently, human activities leading to deforestation, such as expanding human settlements and agricultural areas, often focus on watercourses. Shifting cultivation, a traditional agricultural method, is often used near water sources. While these approaches are generally acceptable in small quantities, excessive usage or insufficient fallow periods may lead to significant deforestation. In a study conducted by Bera et al., (2020), it was found that there is a negative correlation between deforestation and the distance to rivers. The research showed that areas located closer to water sources are more susceptible to deforestation.

The variable "distance from 5km buffer" had no significant effect on the probability of deforestation within the analyzed dataset. There is a possibility that the deforestation distribution within and around the buffer zone is relatively consistent, which could result in a lack of statistically significant differentiation. This may indicate that deforestation pressures are widespread, impacting regions near the buffer zone and those located further away. Furthermore, enforcement and monitoring efforts might be ineffective in the vicinity of the Wildlife Sanctuary, including the buffer zone.

The variable "Elevation" demonstrates a statistically significant positive relationship between 'Elevation' and deforestation. This suggests that as elevation increases, the likelihood of deforestation also rises. This scenario is also supported by Gayen and Saha (2018), who studied the factors affecting deforestation in the Pathro river basin. They used binary logistic regression to analyze various variables, including elevation, and their results indicated a positive correlation between elevation and deforestation. The results of our analysis also showed that this positive correlation represents a shift in land use patterns. As lowland regions get cultivated and decline in quality, communities may start to look towards higher elevations with gently sloped lands as the next opportunity for agricultural activities. This scenario may lead to unintentional opportunities for implementing shifting cultivation for people to engage in such activities in distant places that are difficult to supervise, in contrast to lowland regions subject to more monitoring. This scenario emphasizes the need for improved monitoring and implementation strategies in remote areas to ensure the efficacy of protected zones.

The deforestation probability map indicates that the western and northern regions of the North Zamari Wildlife Sanctuary were recognized as having higher deforestation risks. Based on binary logistic regression, the deforestation probability map indicates that 1.08% of the forested areas inside the Wildlife Sanctuary and 4.42% of the forested areas within the 5km buffer are very susceptible to deforestation. This is a concerning matter that requires attention in order to minimize additional threats to the forest cover. These high deforestation risk areas along gentle slope areas with high elevation and near stream areas should be focused by cautious monitoring to avoid potential risks. Forest managers could use a deforestation probability map prepared for designating and locating potential deforestation areas to prepare restoration planning and forest management-related works.

6. Conclusions

The analysis of the changes in forest cover within the North Zamari Wildlife Sanctuary from 2015 to 2020 indicated a significant reduction of 240.14 hectares in the forested area. It was also observed the same trend in the 5 km buffer area, amounting for 2009.37 ha of the forest area. In the study area, most of the closed forest was converted to open forest and otherwooded land, and only a very small fraction to cultivated land and other land. In other words, there was a significant forest degradation and remarkable deforestation both within the North Zamari WS and its adjacent 5 km buffer area. It was also observed that the rate of deforestation and forest degradation of the 5 km buffer area was higher than that of the North Zamari WS. This indicated that degraded forest and deforested area, and expansion of cultivated land were the results of illegal encroachment and high dependency on the forest of the local livelihoods, especially in the buffer area. Therefore, proper land-use planning, effective control and conservation of forest, and developing alternative local livelihoods with the support of REDD+ activities and other efficient ways in order to avoid forest degradation and conversion of forest to other land uses not only in the North Zamari WS but also in the 5 km buffer area.

A binary logistic regression analysis was conducted to identify the factors contributing to deforestation. The results revealed that deforestation was significantly impacted by proximity to roads, altitude, and slope. The findings can assist the planning and implementation of targeted measures to reduce deforestation in the study area. A True Positive Rate of 0.8491 signifies that the model can accurately detect approximately 85% of actual deforestation occurrences. The Efficiency score of 0.8106 signifies a substantial overall accuracy, considering both true positive and true negative predictions. Moreover, the AUC value of

0.8596 indicates that the model possesses a strong capability to differentiate between deforested and non-deforested categories.

Deforestation is a significant environmental concern within the North Zamari WS and its adjacent 5km buffer area. Aligning with the core objectives of the REDD+ initiatives, deforestation and forest degradation can be reduced through creating alternative and efficient livelihoods for expansion of cultivated land, raising awareness on importance of forest for combating illegal logging, distributing efficient cooking stoves for charcoal and fuel extraction, establishing plantations on degraded forest area for rehabilitating natural forest, conducting necessary silvicultural operations in degraded forest for regrowth of forests, and strengthening law enforcement for illegal encroachment.

To enhance the quality of future research on the factors affecting deforestation, it is essential to integrate socio-economic and environmental variables into the model. This will enhance the understanding of the issue and contribute to conservation planning. Considering these broader socio-economic factors allows for the development of management strategies that are better suited to address the driving factors of deforestation. This, in turn, will bring about in effective results for the forest ecosystem within the sanctuary.

7. Acknowledgements

Firstly, we would like to gratitude with our deepest regards to my supervisors, Dr. Thin Thin (Assistant Director, RS & GIS Section, Planning and Statistics Division) and U Khine Zaw Wynn (Assistant Director, RS & GIS Section, Planning and Statistics Division), for your valuable guidance. Our special thanks to “FD-KFS REDD+ Project” for financial support. We appreciate the contributions of Dr. Yuya Aye (Assistant Director, Forest Research Institute), and U Phyo Wai (Staff Officer, RS & GIS Section, Planning and Statistics Division) for providing precious comments and suggestions to improve the quality of our research paper. We acknowledge Forest Department and Forest Research Institute (Yezin) for providing a conducive research environment. Finally, thanks to our colleagues from RS and GIS Section, Planning and Statistics Division, for your constructive feedbacks. This research would not have been possible without your collective efforts.

Chan Myae Aung
Staff Officer
Forest Research Institute
14-02-2024

Nay Aung
Staff Officer
Planning and Statistics Division
14-02-2024

8. References

Abd El-Kawy, O. R., Rød, J. K., Ismail, H. A., & Suliman, A. S. (2011). Land use and land cover change detection in the western Nile delta of Egypt using remote sensing data. *Applied Geography*, *31*(2), 483–494. https://doi.org/10.1016/j.apgeog.2010.10.012

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, *43*(6), 1223–1232.

Alqurashi, A., & Kumar, L. (2013). Investigating the use of remote sensing and GIS techniques to detect land use and land cover change: A review. *Advances in Remote Sensing*.

Anderson, J. R. (1976). *A land use and land cover classification system for use with remote sensor data* (Vol. 964). US Government Printing Office.

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, M. M., Nwe, W. W., Saung, T., & Oo, T. N. (2018). Assessment of plant species diversity and wildlife in Pindaya Townsip, Taunggyi District, Shan State, Myanmar. *Forest*

Research Institute, Ministry of Natural Resources And.

Baynard, C. W. (2013). *Remote sensing applications: Beyond land-use and land-cover change*.

Belayneh, Y., Ru, G., Guadie, A., Teffera, Z. L., & Tsega, M. (2020). Forest cover change and its driving forces in Fagita Lekoma District, Ethiopia. *Journal of Forestry Research*, *31*, 1567–1582.

Bera, B., Saha, S., & Bhattacharjee, S. (2020). Forest cover dynamics (1998 to 2019) and prediction of deforestation probability using binary logistic regression (BLR) model of Silabati watershed, India. *Trees, Forests and People*, *2*, 100034.

Billy, N. W., Swe, S. T., & Phyu, P. H. (2015). Study on Soil Erosion in Abandon Shifting Cultivations in Bago Yoma Watershed Areas. *Forest Research Institute, Forest Department, Myanmar*, *Paper*(22/2015), 31 pages.

Buya, S., Tongkumchum, P., & Owusu, B. E. (2020). Modelling of land-use change in Thailand using binary logistic regression and multinomial logistic regression. *Arabian Journal of Geosciences*, *13*, 1–12.

Chang, X., Zhang, F., Cong, K., & Liu, X. (2021). Scenario simulation of land use and land cover change in mining area. *Scientific Reports*, *11*(1), 12910.

Cochran, W. G. (1977). *Sampling techniques*. john wiley & sons.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

FAO. (2020). *Global forest resources assessment 2020–key findings, Policy research paper*. FAO Rome.

FRA, F. A. O. (2015). Global forest resources assessment 2015 Desk reference. *Food and Agriculture Organization of the United Nations, Rome*.

Franklin, S. E. (2001). *Remote sensing for sustainable forest management*. CRC press.

Gayen, A., & Saha, S. (2018). Deforestation probable area predicted by logistic regression in Pathro river basin: a tributary of Ajay River. *Spatial Information Research*, *26*(1), 1–9.

GFRA. (2020). *Global forest resources assessment 2020: Main report*. Food & Agriculture Organization of the UN.

Guhathakurta, P., & Roy, S. (2000). *Joint forest management in West Bengal: a critique*. World Wide Fund for Nature.

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

Hasan, S., Shi, W., & Zhu, X. (2020). Impact of land use land cover changes on ecosystem service value–A case study of Guangdong, Hong Kong, and Macao in South China. *PloS One*, *15*(4), e0231259.

Hein, Z. M., William, A. C., Soe, P., Cox, N. J., Htun, N. Z., Oo, T. N., Aye, Y. Y., Htun, Y. L., & Yoganand, K. (2020). *Status of two species of threatened wild cattle (Bos gaurus and Bos javanicus birmanicus) in North Zamari Wildlife Sanctuary, Bago Region, Myanmar*.

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

Kafy, A.-A., Saha, M., Fattah, M. A., Rahman, M. T., Duti, B. M., Rahaman, Z. A., Bakshi, A., Kalaivani, S., Rahaman, S. N., & Sattar, G. S. (2023). Integrating forest cover change and carbon storage dynamics: Leveraging Google Earth Engine and InVEST model to inform conservation in hilly regions. *Ecological Indicators*, *152*, 110374.

Krieger, D. J. (2001). *Economic value of forest ecosystem services: a review*.

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

Liao, Q., Li, M., Chen, Z., Shao, Y., & Yang, K. (2010). Spatial simulation of regional land use patterns based on GWR and CLUE-S model. *2010 18th International Conference on Geoinformatics*, 1–6.

Linkie, M., Smith, R. J., & Leader-Williams, N. (2004). Mapping and predicting deforestation patterns in the lowlands of Sumatra. *Biodiversity & Conservation*, *13*, 1809–1818.

Menard, S. (2002). *Applied logistic regression analysis* (Issue 106). Sage.

Mon, M. S., Kijisa, T., Mizoue, N., & Shigejiro, Y. (2009). "Factors Affecting Deforestation in Paunglaung Watershed, Myanmar using Remote Sensing and GIS." *Japan Society of Forest Planningapan Society of Forest Planning*, *14*(1), 7–16.

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

MONREC. (2015). National Biodiversity Strategy and Action Plan 2015-2020. *Ministry of Natural Resources and Environmental Conservation of Myanmar Forest Department*.

Muzdalifah, Q. R., Deliar, A., Virtriana, R., Naufal, A., & Ajie, I. S. (2020). Using geographically weighted–binary logistic regression to analyze land cover change phenomenon (case study: northern west java development region). *IOP Conference Series: Earth and Environmental Science*, *448*(1), 12121.

Naing Tun, Z., Dargusch, P., McMoran, D. J., McAlpine, C., & Hill, G. (2021). Patterns and Drivers of Deforestation and Forest Degradation in Myanmar. *Sustainability*, *13*(14), 7539.

Negassa, M. D., Mallie, D. T., & Gemeda, D. O. (2020). Forest cover change detection using Geographic Information Systems and remote sensing techniques: a spatio-temporal study on Komto Protected forest priority area, East Wollega Zone, Ethiopia. *Environmental Systems Research*, *9*, 1–14.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V, Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, *148*, 42–57.

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* (Issue 132). Sage publications.

Pendrill, F., Gardner, T. A., Meyfroidt, P., Persson, U. M., Adams, J., Azevedo, T., Bastos Lima, M. G., Baumann, M., Curtis, P. G., & De Sy, V. (2022). Disentangling the numbers behind agriculture-driven tropical deforestation. *Science*, *377*(6611), eabm9267.

Rosy, N. W., & Billy, N. W. (2015). Analysis of shifting cultivation practices and teak regeneration and soil nutrient on fallow lands in Bago Mountains, Myanmar. *Forest Research Institute, Forest Department, Myanmar*, *Paper*(21/2015), 15 pages.

Rwanga, S. S., & Ndambuki, J. M. (2017). Accuracy assessment of land use/land cover classification using remote sensing and GIS. *International Journal of Geosciences*, *8*(04), 611.

Spruce, J., Bolten, J., Mohammed, I. N., Srinivasan, R., & Lakshmi, V. (2020). Mapping land use land cover change in the Lower Mekong Basin from 1997 to 2010. *Frontiers in Environmental Science*, *8*, 21.

Story, M., & Congalton, R. G. (1986). Accuracy assessment: a user's perspective. *Photogrammetric Engineering and Remote Sensing*, *52*(3), 397–399.

Win, R. N., Reiji, S., & Shinya, T. (2009). Forest cover changes under selective logging in the Kabaung Reserved Forest, Bago Mountains, Myanmar. *Mountain Research and Development*, *29*(4), 328–338.

Xie, G., Li, W., Xiao, Y., Zhang, B., Lu, C., An, K., Wang, J., Xu, K., & Wang, J. (2010). Forest ecosystem services and their values in Beijing. *Chinese Geographical Science*, *20*(1), 51–58.

Yaung, K. La, Chidthaisong, A., Limsakul, A., Varnakovida, P., & Nguyen, C. T. (2021). Land Use Land Cover Changes and Their Effects on Surface Air Temperature in Myanmar and Thailand. *Sustainability*, *13*(19), 10942.

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

Zin, M. T. (2005). *“Developing a Scientific Basis for Sustainable Management of Tropical Forest Watersheds: Case Studies from Myanmar.”* Universitätsverlag Göttingen.