HPE In-semester Project Report on

Forest Cover Changes Detection Using Landsat 8 Time Series Data in Western Ghats Indian Forests

Submitted by

Varsha P (171IT148)

Rakshatha Vasudev (171IT131)

Department of Information Technology

National Institute of Technology Karnataka, Surathkal

July 2020

Abstract

Timely and accurate detection of forest changes or disturbances can help in an effective forest management.The observations from time series data if combined with the radar satellite ones gives exceptional results which provides an excellent way to monitor forests.Forest monitoring is very much necessary to detect forest changes and forest degradation.Machine learning models are developed to detect forest changes for forest disturbances.Our project involves detecting forest changes with the help of landsat 8 imagery of Western Ghats particularly the Agumbe region using a machine learning model.Disturbance probability of the observations from the landsat 8 images are classified using Random forest classifier.We have predicted deforestation,forest degradation and stability(no change in forest) using random forest model with Landsat data and with the covariates obtained from landsat spectral and temporal metrics.

Keywords: *LTS-Landsat Time Series, spatial-temporal , Random forest method , forest monitoring , time series ,fmask algorithm-function of mask*

Contents

List of Figures

1 Introduction

Forests are the most widely covered natural vegetation and plays an important role in ecosystem balance.Forests are the most vital sources of many of our day to day needs like fuel,food,furniture etc and houses a huge variety of flora and fauna. As such, forests play a very vital role in the everyday life of humans.Since plants are responsible for providing the life saving oxygen and consumption of carbon dioxide, the sustenance of forests is always of prime importance.This considered it is also important that the forests are monitored timely and efficiently.

Landsat, being one of the most reliable sources for the earth data has been a part of earth observation programs for about four decades. The data is collected using a global acquisition strategy in a very systematic way.The increase in the level of processing and analysis of Landsat has led to many developments in technology and science.In particular, forest monitoring analysis from LTS can be attributed to both abrupt and gradual changes.

Forest monitoring is done using landsat to classify forest change types.This is accomplished by based on the temporal trajectory of the spectral variables over time.The range of spectral bands varies from single-band-reflectance to range of indices.The original spectral bands and their derivatives are manipulated to compute these ranges.The variables which are commonly used are known for their sensitivity toward chlorophyll pigments or structure of the forests.

LTS based forest monitoring methods allows for rapid forest disturbance detection and descriptions of forest change trajectories beyond reach of conventional methods.

2 Literature Survey

2.1 *Related Work*

Both current state and the temporal dynamics are considered while characterizing the forest change types.In their journal[1] Nilam et al suggest that Landsat can be used for change detection which is broadly categorized into two types: trajectory-based analysis and image classification.

Combined with the local experts' data landsat based trajectories can be analysed to detect forest changes which is investigated by Ben DeVries et al. in their work[2] .From each spectral band and index,series of temporal trajectory metrics were derived using a version of BFAST[7] algorithm.These metrics describe changes in trend and seasonal amplitudes between time series segments and also an overall time series trend and intercepts.

Land use/land cover change (LUCC) is one of the underlying cause of global environmental change.In their paper [3] Vijayasekaran et al have presented a study that Landsat satellite can be used to examine and estimate finer scale LUCC.The study area is in the Mula-Pravara River basin(a semi-arid region of India).They have identified hotspots from the LUCC and have investigated the driving forces that ha caused these changes.

Katsuto et al in their work[8] have investigated the importance of using time series data(Landsat 8 and Sentinel-1) in detecting forest disturbances in tropical seasonal forests.

Huang et al have created a Vegetation Change Tracker which uses integrated forest z-score value to detect any changes which is presented in article[5]. The algorithm is based on the temporal and spectral properties of the land cover and the forest changes.It requires little or no fine tuning for most forests.

Synthetic Aperture Radar (SAR) is used to observe land surfaces under cloudy conditions.It is detected as the radar signals penetrate though cloud.Studies have mapped forest degradation and deforestation using L-band in the SAR data as per[6] by Neha et al.

Tropical and sub-tropical sites were tested to apply synthetic and optical aperture radar data for classification of the land cover to support both REDD+ and MRV by Laura et al[9].

2.2 *Outcome of Literature Survey*

Existing methods of forest cover detection tend to focus on gradual changes over a time period and exclude abrupt changes. Inclusion of low level drivers of forest cover change and correlation factor between different covariates is found to have given better results in previous works.

2.3 *Problem Statement*

Forest change detection using Landsat 8 Time Series Data in Western Ghats(Agumbe) Indian Forests.

2.4 *Objectives*

1. Differentiation between deforestation,forest degradation and stability(no change) using Landsat time series data.
2. Description of key change processes by mapping of forest change types.
3. Deriving Temporal Metrics
4. Developing random forest model to detect forest change types as a function of landsat spectral-temporal metrics.

3 Methodology

3.1 *Dataset*

The dataset used for the study is Landsat 8 in which the area coverage is the **Western Ghats forests of India**-Agumbe region in particular.Landsat data consists of nine spectral bands.

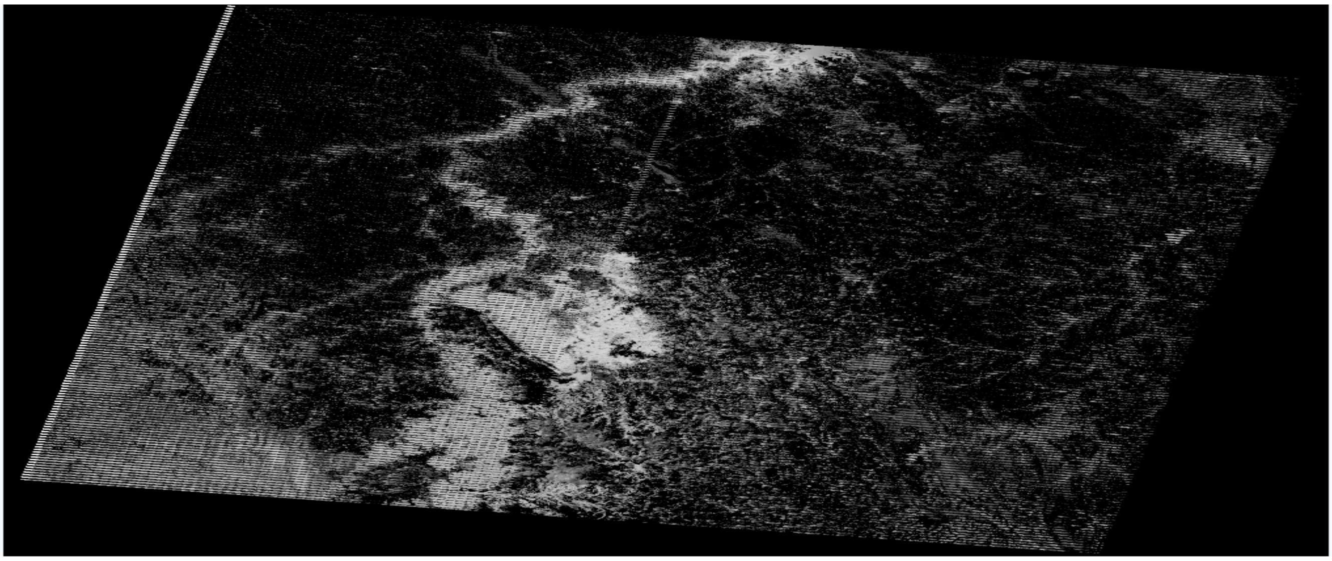

Figure 3.1: Imagery before clouds removal

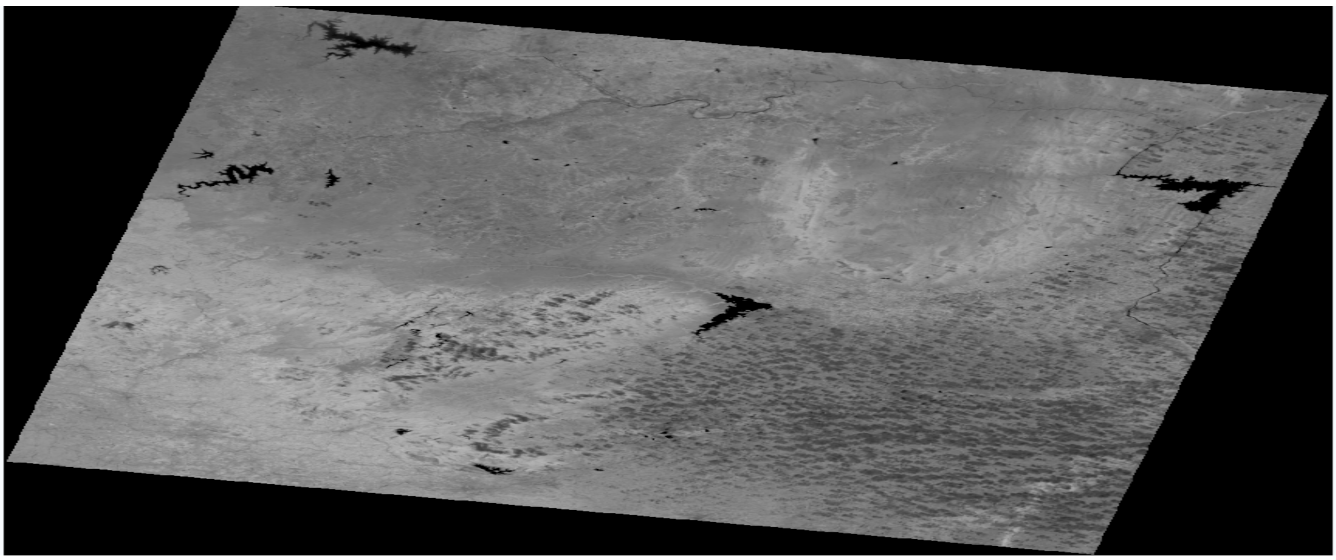

Figure 3.2: Clouds removal

3.2 *Methods and Evaluation Techniques*

3.2.1 *LTS pre-processing*

All the available imagery pertaining to our area of study was downloaded from Landsat-OLI sensor.The images with a cloud cover below 90% for each scene and with a processing level L1 were selected.Cloud mask which is derived from **FMASK** algorithm(Function of Mask) was applied to each scene, which effectively masked out the clouds,gaps and cloud shadows.

In order to reduce the number of cloud pixels which might be undetected by the cloud masks derived from FMASK algorithm, all images were applied with **Five pixel sieve**.The clusters of pixel that are around masked values ¡=5 pixels will be removed.

Name	Abbreviation	Equation	Remarks
Normalized Difference Vegetation Index	NDVI	$\frac{NIR-R}{NIR+R}$	sensitive to photosynethetic activity
Normalized Difference Moisture Index	NDMI	$\frac{NIR-SWIR1}{NIR+SWIR1}$	sensitive to canopy moisture content
Normalized Burn Ratio	NBR	$\frac{NIR-SWIR2}{NIR+SWIR2}$	sensitive to disturbances and fire
Normalized Burn Ratio 2	NBR2	$\frac{SWIR1-SWIR2}{SWIR1+SWIR2}$	sensitive to disturbances and fire
Tasseled Cap Brightness	TCB	$b_1B+b_2R+b_3G+b_4NIR+b_5SWIR1 + b_6SWIR2$	sensitive to surface brightness
Tasseled Cap Greenness	TCG	$g_1B+g_2R+g_3G+g_4NIR+g_5SWIR1 + g_6SWIR2$	sensitive to vegetation greenness
Tasseled Cap Wetness	TCW*	$w_1B + w_2R + w_3G + w_4NIR + w_5SWIR1 + w_6SWIR2$	sensitive to vegetation moisture content
Tasseled Cap Angle	TCA	$tan^{-1}\left(\frac{TCB}{TCG}\right)$	sensitive to above-ground biomass

*Spectral indices used to make final change probability maps.

Figure 3.3: Spectral Indices

The spectral indices were selected using pre-processed surface-reflectance layers as in Fig 3.3.These indices are considered to be more sensitive to changes and vegetation characteristics over time.

3.2.2 *Deriving Temporal Metrics*

We have considered three forest change types to be the classes of our model which are:

- Deforestation
- Forest Degradation

- Stable(no change)

A time series data usually consists of systematic components which are trend and seasonality, and a non-systematic component noise.

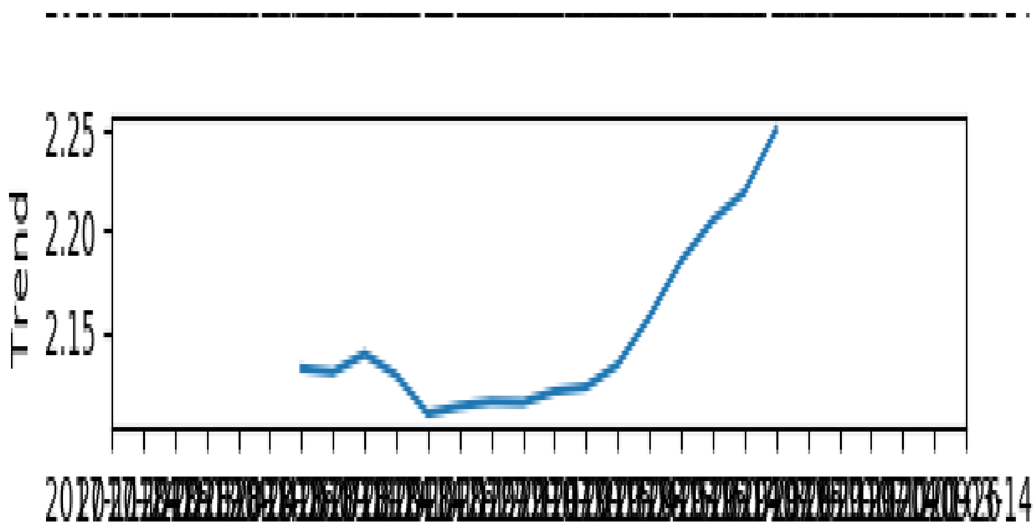

Figure 3.4: NDVI Trend

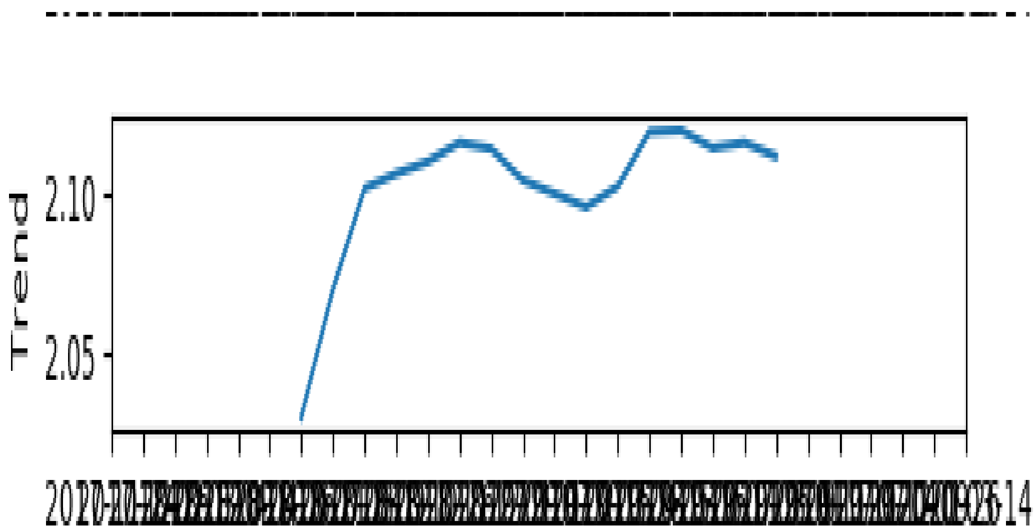

Figure 3.5: NDMI Trend

We have contrived a quadratic time series as a square of time step from 1-99 and have decomposed the data assuming that it is a multiplicative model.The extracted trend wholly characterizes a time series data. A linear function is used to fit each spectral band to the time series.Instead of commonly used linear regression models,**Robust linear**

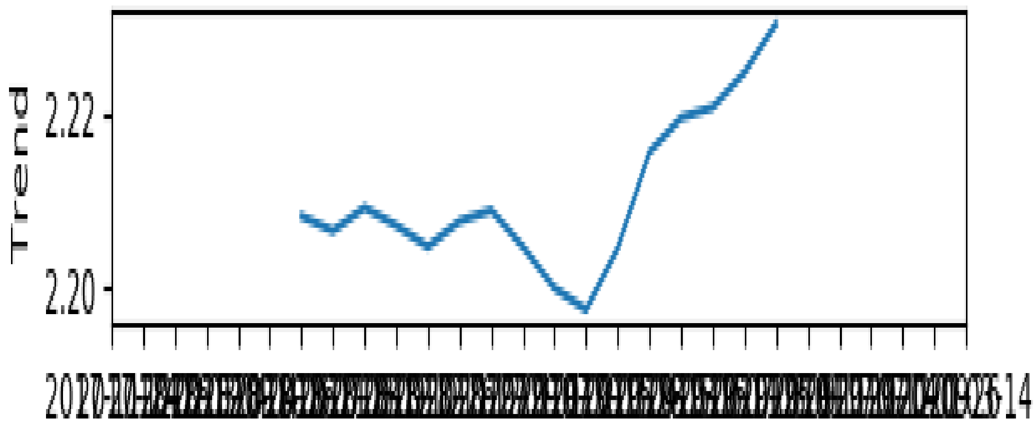

Figure 3.6: NBR2 Trend

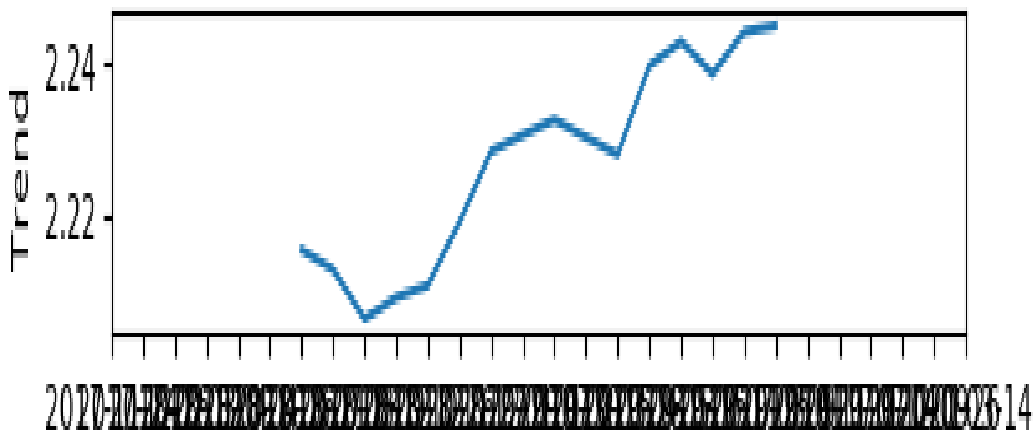

Figure 3.7: NBR Trend

regression [4] is used.This is because the Landsat data may usually contain noise because of some unmasked clouds.While fitting a least squares regression, there is a possibility of finding some outliers. The output of the time series that were segmented was applied to each spectral band.It will consist of necessary geographical, spectral temporal metrics.

An overall Robust Linear Regression trend cannot capture the abrupt or onset of gradual changes.To describe these changes,**Break point** method is used.It is implemented by testing the presence or absence of breaks for every spectral-band in every time series.An assumption is made that the land use or forest change will occur once throughout the

[h!]

NDVI	0.4146230029069114
NDMI	0.42374108594431656
NBR	0.2923316761955195
NBR2	0.1564983709526736

Table 1: Standard error of each index in overall trend fitting.

length of the time series (of the duration 2018-2020) to identify the important break.So the number of breaks(maximum) is set to 1.Seasonal trend models were fitted for each segment after the break point computation.

The formula used for fitting is given in the figure,

$$y_t = \alpha_j + \beta_j t + \gamma_j sin(\frac{2\pi t}{f} + \delta_j)$$

Figure 3.8: Sine formula for seasonal trend fitting

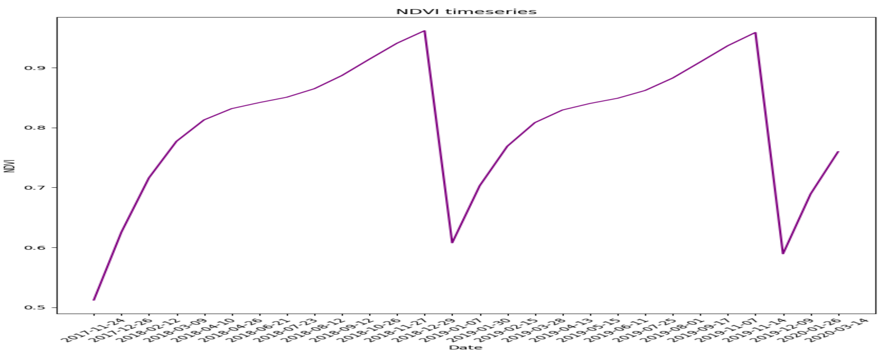

Figure 3.9: NDVI Time series - Seasonality

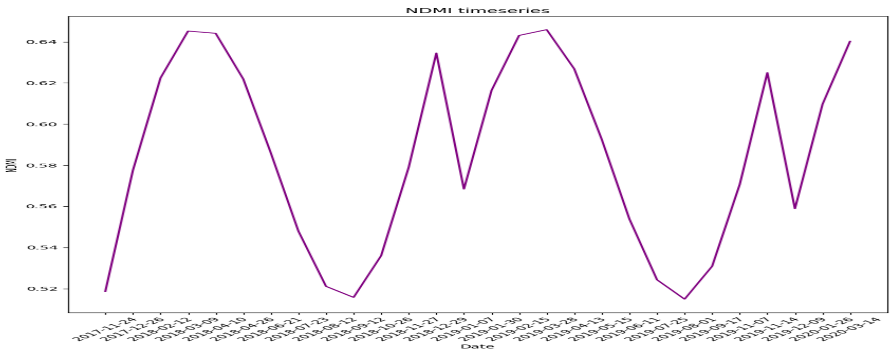

Figure 3.10: NDMI Time series - Seasonality

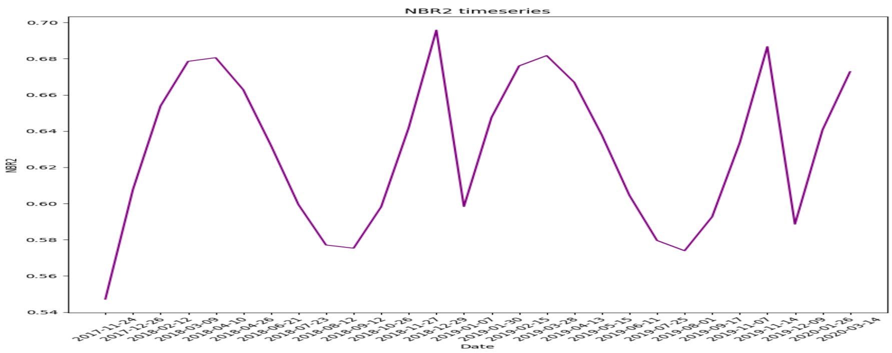

Figure 3.11: NBR2 Time series - Seasonality

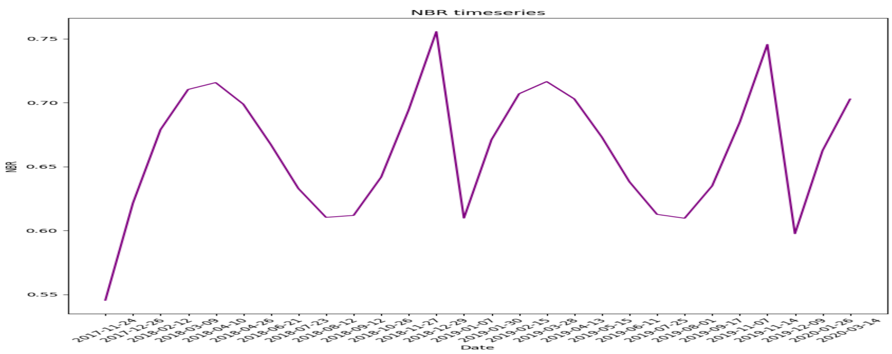

Figure 3.12: NBR Time series - Seasonality

3.2.3 ***Random Forest Model***

Random forests was used to model the forest change types as a function of landsat spectral and temporal metrics.A Random forest classifier consists of many individual decision trees working as an ensemble.Each individual tree in the classifier provides a prediction of the class and the class with the highest votes is considered as the model's prediction. **Feature Randomness** — In decision tree classifiers, every possible feature is considered.The observations in the left node v/s observations in the right node is considered.The feature which is producing highest variation is picked when splitting a node.But in random forest classifier random subset of features is chosen by each tree.This results in low correlations among the trees and increases diversification.A low correlation results in accurate detection of important features i.e. spectrum variables.

NDVI	0.34735356
NDMI	0.36413809
NBR	0.15687626
NBR2	0.13163209

Table 2: Feature importance.

3.2.4 *Selecting Important Variables*

The random forest model produced maps with a subset of spectral and temporal covariates which are important.The interpretation of important variables poses a problem if the covariates are tightly correlated. A conditional measure proposed by Strobl et al. (2008) [10] is implemented to overcome this.A **Scoring Algorithm** is used on the permuted samples of the covariates to check for importance using the equation,

$$S_{j,\Delta} = \frac{1}{N} \sum_{i}^{N} \frac{x_{i,j} - 1}{n - 1}$$

Figure 3.13: Score

4 Results and Analysis

4.1 Model Accuracy and Covariates Importance

The random forest classifier gave an overall Out Of Bag error estimate of **21.428%**.Following table provides the class errors.

Deforestation	29.963%
Forest Degradation	22.628%
No change	26.147%

Table 3: Various class errors.

Even though multiple iterations showed inconsistencies in important metrics, overall robust regression trends from various spectral bands helped in consistently ranking important predictors. Feature importance were calculated for all the indices and **NDMI** was the prominent deciding factor followed by NDVI, NBR, and NBR2.

5 Conclusion & Future Work

From the model, we determined that NDVI and NDMI are the spectral bands which are important for detecting the changes between deforestation and forest degradation.The spectral and temporal covariates based on these bands were used to give forest change predictions.

The future work can be followed on this approach by expanding the data sources for example including Sentinel-2, airborne or terrestrial LiDAR or any other airborne remote sensing datasets using data fusion methods. Furthermore,the approach is extensible and can include other forest change types such as reforestation and afforestation.

References

[1] Asim Banskota, Nilam Kayastha, Michael Falkowski, Michael Wulder, Robert Froese, and Joanne White. Forest monitoring using landsat time series data: A review. *Canadian Journal of Remote Sensing*, 40:362–384, 12 2014.

[2] Ben DeVries, Arun Pratihast, Jan Verbesselt, Lammert Kooistra, and Martin Herold. Characterizing forest change using community-based monitoring data and landsat time series. *PLoS ONE*, 11:e0147121, 03 2016.

[3] Vijayasekaran Duraisamy, Ramkumar Bendapudi, and Ajit Jadhav. Identifying hotspots in land use land cover change and the drivers in a semi-arid region of india. *Environmental Monitoring and Assessment*, 190, 08 2018.

[4] UCLA: Statistical Consulting Group. Robust regression r data analysis examples, https://stats.idre.ucla.edu/r/dae/robust-regression (March 11, 2020).

[5] Chengquan Huang, Samuel Goward, Jeffrey Masek, Nancy Thomas, Zhiliang Zhu, and Jim Vogelmann. An automated approach for reconstructing recent forest disturbance history using dense landsat time series stacks. *Remote Sensing of Environment*, 114:183–198, 01 2010.

[6] Neha Joshi, Edward Mitchard, Natalia Woo, Jorge Torres, Julian Moll-Rocek, Andrea Ehammer, Murray Collins, Martin Jepsen, and Rasmus Fensholt. Mapping dynamics of deforestation and forest degradation in tropical forests using radar satellite data. *Environmental Research Letters*, 10:034014, 03 2015.

[7] R-forge. Bfast, breaks for additive season and trend, http://bfast.r-forge.r-project.org (March 11, 2020).

[8] Katsuto Shimizu, Tetsuji Ota, and Nobuya Mizoue. Detecting forest changes using dense landsat 8 and sentinel-1 time series data in tropical seasonal forests. *Remote Sensing*, 11:1899, 08 2019.

[9] Laura Sirro, Tuomas Häme, Yrjö Rauste, Jorma Kilpi, Jarno Hämäläinen, Katja Gunia, Bernardus Jong, and Fernando Paz Pellat. Potential of different optical and sar data in forest and land cover classification to support redd+ mrv. *Remote Sensing*, 10:942, 06 2018.

[10] Carolin Strobl, Anne-Laure Boulesteix, Thomas Kneib, Thomas Augustin, and Achim Zeileis. Conditional variable importance for random forests. *BMC Bioinformatics*, 9(1):307, 2008.