TECHNICAL REPORT ON

Sal (*Shorea robusta*)

Method for rapid estimation of distribution and productivity of Sal in Jharkhand

ISB | Bharti Institute of Public Policy

Highlights

- Village-wise total number of Sal trees
- Village-wise average production of sal seeds
- First model of sal tree classification using satellite images and machine learning
- Overall accuracy of > 89% for district and state level estimates

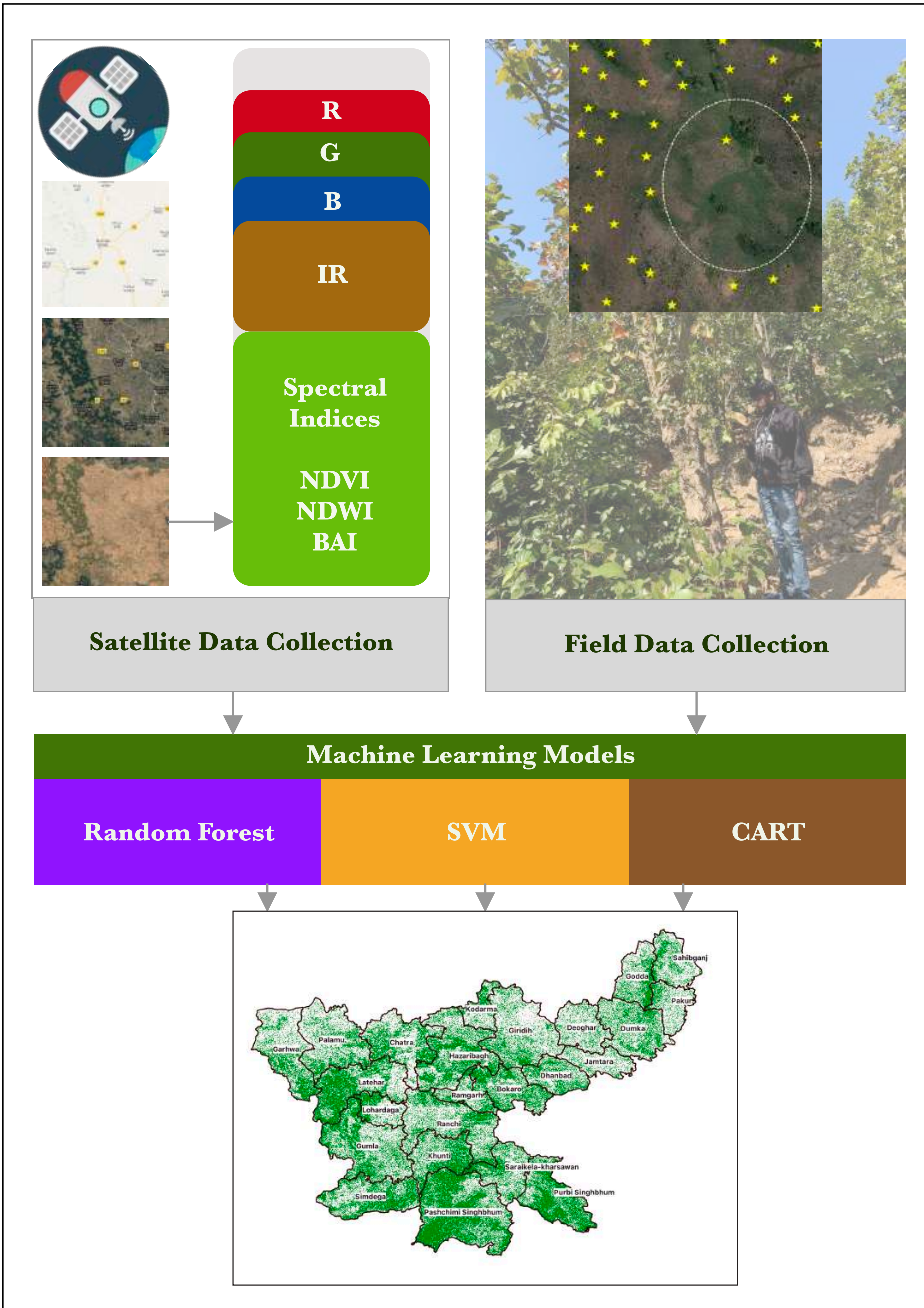
Information on distribution and abundance of tree species with reasonable accuracy is crucial for developing strategies for Seasonal Forest Products (SFPs) extraction, sustainable management, and ecological monitoring. Recent advances in remote sensing technologies have made tree species classification possible. We present here preliminary findings from machine learning based model for estimation of number of Sal trees in Jharkhand using satellite imagery.

For more information:

Initiative on the Forest Economy
Indian School of Business, Hyderabad

https://forestgovernance.in
forest_governance@isb.edu

Visual Summary



Technical Summary

- Data from Sentinel-2 satellite system have been used for analysis.

- The spatial resolution of Sentinel-2 data is 10 meters with a revisit time of 5 days at the equator. The satellite images consist of 13 bands, of which three are in red-edge spectrum and two in shortwave infrared spectrum.

- Stratified Random Sampling have been used for collection of field data.

- Three different machine learning models - Random Forest (RF), Support Vector Machine (SVM), Classification and Regression Tree (CART) - were used to produce the classification map. An overall accuracy of 89%, 87% and 82% have been achieved using RF, SVM and CART respectively.

- Future improvements will focus on sensor fusion with LiDAR and deep learning models.

The Science

The physical interaction between sunlight and tree leaves forms the scientific basis of the use of remote sensing satellite in tree species classification. Different wavelengths in the electromagnetic spectrum have different interaction pattern with the leaves. Leaves mainly absorb light in the visible spectrum due to the presence of chlorophyll and carotenoids. Leaf morphology affects how photons are scattered within the chlorophyll pigments and results in high reflectance in the near-infrared spectrum. Presence of water in the leaves is responsible for chemical absorption at 970 nm and 1200 nm and results in a drop in reflectance in these wavelengths. Leaves contain cellulose and lignin molecules that reflect strongly in the short-wave infrared spectrum, but the spectral absorption of water overshadows this in green leaves; cellulose and lignin molecules in dry leaves have a different reflectance, since water is absent. **Flowering, leaf-onset, and senescence change the biophysical and structural properties of leaves, which is used to differentiate between tree species in multi-temporal satellite datasets.**

Recent scholarship using Sentinel-2 imagery has shown that the red-edge bands and the SWIR bands are useful for discriminating between tree species. We combined multi-temporal Sentinel-2 imagery with machine learning models to detect Sal trees in the landscape with reasonable accuracy.

Materials and Method

Materials

The success of machine learning models depends on the quality and quantity of input training data. The following approach was adopted for collection of the input data:

A. Sentinel-2 data

Cloud free imagery at different phenological stages of Sal tree was acquired for the year 2021. For Sal, different phenological stages consist of leaf-onset, bloom and senescence.

B. Field data

Field data collection was carried out in December 2021. Coordinates of Sal trees were recorded using handheld GPS device. The recorded locations were taken from pure Sal stands to ensure that the distinct spectral reflectance signature of Sal was captured. A total of 1736 points were recorded.

Method

Spectral profile plots were constructed for Sal, water and other landcover type to examine the difference and to select of bands for inputs into the model. Apart from band values, major spectral indices like NDVI, BAI and NDWI were also calculated and used as input parameters in the model. Three different machine learning models were used to generate the classification. Random Forest produced consistently best results as compared to SVM and CART.

Accuracy assessment

A 10-fold cross-validation was used to evaluate the RF classification models. The input data set was randomly split into 10 samples of equal size. Nine samples (90%) were used to train the classifier and one sample (10%) was used for validation.

Production estimate

Data on production estimate was collected from key informants in local communities during the field work period. Another source of production data was found in the 2015-2018 Survey report on NTFP from Department of Forest, Environment, and Climate Change, Jharkhand. The production of Sal seeds was estimated by multiplying the number of Sal trees with average production per tree.

Results

Map showing the distribution of Sal trees in different district in Jharkhand

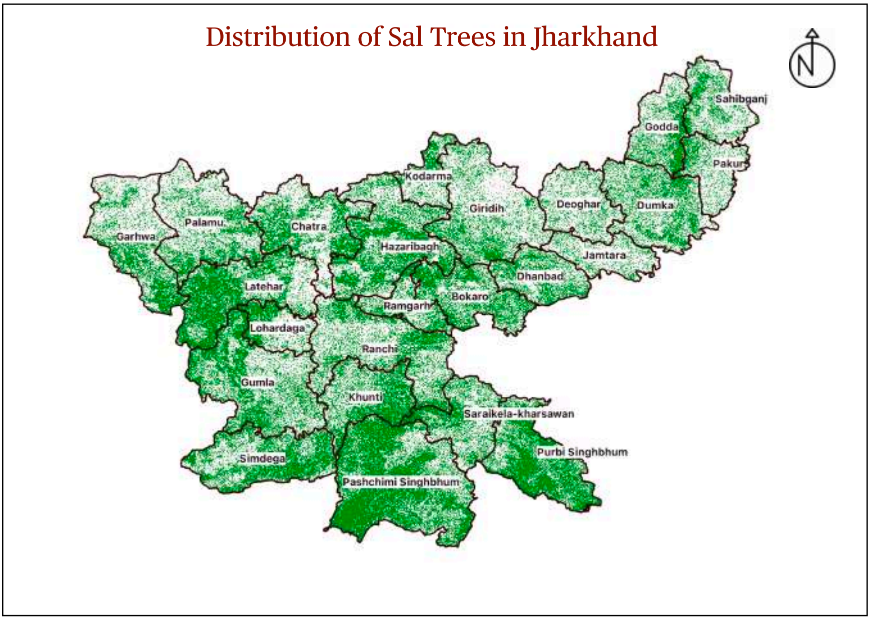


Table showing the number of Sal trees in different district in Jharkhand

District Name	Number of Sal Trees	Average Production	Lower limit	Upper limit
Jamtara	52,35,933	8,90,10,861	4,71,23,397	13,08,98,325
Kodarma	65,44,413	11,12,55,021	5,88,99,717	16,36,10,325
Pakur	65,64,687	11,15,99,679	5,90,82,183	16,41,17,175
Ramgarh	73,63,954	12,51,87,218	6,62,75,586	18,40,98,850
Lohardaga	76,14,070	12,94,39,190	6,85,26,630	19,03,51,750
Deoghar	77,51,370	13,17,73,290	6,97,62,330	19,37,84,250
Sahibganj	90,95,447	15,46,22,599	8,18,59,023	22,73,86,175
Godda	1,09,46,979	18,60,98,643	9,85,22,811	27,36,74,475
Dhanbad	1,13,06,474	19,22,10,058	10,17,58,266	28,26,61,850
Saraikela-kharsawan	1,41,11,898	23,99,02,266	12,70,07,082	35,27,97,450
Garhwa	1,73,31,457	29,46,34,769	15,59,83,113	43,32,86,425
Khunti	1,75,62,545	29,85,63,265	15,80,62,905	43,90,63,625
Bokaro	1,80,35,047	30,65,95,799	16,23,15,423	45,08,76,175
Chatra	1,80,65,041	30,71,05,697	16,25,85,369	45,16,26,025
Palamu	1,81,12,264	30,79,08,488	16,30,10,376	45,28,06,600
Dumka	1,82,51,638	31,02,77,846	16,42,64,742	45,62,90,950
Giridih	1,85,89,871	31,60,27,807	16,73,08,839	46,47,46,775
Ranchi	2,33,44,176	39,68,50,992	21,00,97,584	58,36,04,400
Simdega	2,41,11,855	40,99,01,535	21,70,06,695	60,27,96,375
Purbi Singhbhum	2,76,20,623	46,95,50,591	24,85,85,607	69,05,15,575
Hazaribagh	2,82,16,859	47,96,86,603	25,39,51,731	70,54,21,475
Latehar	2,85,73,790	48,57,54,430	25,71,64,110	71,43,44,750
Gumla	3,00,85,278	51,14,49,726	27,07,67,502	75,21,31,950
Pashchimi Singhbhum	5,55,39,581	94,41,72,877	49,98,56,229	1,38,84,89,525
Important considerations:				
1. Production estimates are in Tonnes				
2. Lower limit is calculated @ 9 kg sal seed / tree without wings				
3. Upper limit is calculated @ 25 kg sal seed / tree without wings				