USING MACHINE LEARNING CLASSIFIERS AND REGRESSION MODELS FOR ESTIMATING THE STAND AGES OF FALCATA PLANTATIONS FROM SENTINEL DATA

J. C. L. Escasio [1], J. R. Santillan [1, 2 *], M. Makinano-Santillan [1, 2]

[1] Department of Geodetic Engineering, College of Engineering and Geosciences, Caraga State University, Ampayon, Butuan City, 8600, Philippines – (johncarl.escasio, jrsantillan, mmsantillan)@carsu.edu.ph

[2] Caraga Center for Geo-Informatics, Caraga State University, Ampayon, Butuan City, 8600, Philippines

Commission IV, WG 7

KEY WORDS: Multivariate Regression, Machine Learning Classification, Sentinel, Falcata, Stand Age

ABSTRACT:

Falcata is a widely planted Industrial Tree Plantation (ITP) species in the Caraga Region, Mindanao, Philippines, significantly contributing to more than 80% of the country's Falcata log production in recent years. Currently, Falcata plantations face several challenges, especially regarding plantation monitoring and management. The information on stand age is essential for the efficient monitoring and sustainable management of ITPs. With advances in technology and freely available satellite data, primarily from the Sentinel satellites, remote sensing has provided an alternative approach to determining stand age information. In this context, this study used multivariate regression models and machine learning classifiers, namely Support Vector Machine (SVM) and Random Forest (RF) to estimate Falcata stand ages from single-date and multitemporal Sentinel-1 and Sentinel-2 images and vegetation indices (VIs). The eleven multivariate regression models have R^2 values ranging from 0.23 and 0.81 and with estimation errors of 1.72 to 3.58 years. The best multivariate regression model developed is an exponential model that relates Falcata stand age with the year 2021 Sentinel-2 surface reflectance bands and year 2021 Sentinel-1 VV and VH polarization bands. The model has training and validation data R^2 of 0.61 and 0.55, which are the most consistent among the 11 regression models developed. When used for stand age estimation, this model may underestimate, or overestimate stand age by 1.72 years. When a machine learning approach is to be employed, the RF classifier performed better than SVM, particularly when estimating stand ages using multitemporal Sentinel-1 and Sentinel-2 data and VIs. The classifier has an overall accuracy of 84.69%, the highest among the eight classification results generated by the study.

1. INTRODUCTION

Falcata (*Paraserianthes falcataria L. Nielsen*) is a fast-growing tree species that may reach 40 meters high and have a diameter of 20 to 100 centimetres. Logs from Falcata are one of the wood industry's most used raw wood products. Over the years, Falcata plantations have played a significant role in the Philippine timber industry. In Mindanao, Philippines, particularly in Caraga Region, Falcata is the most planted among the several Industrial Tree Plantation (ITP) species because of their rapid growth and being harvestable in four to six years after planting (Doloriel, 2017). From the Philippines' total Falcata log production of 560,970 and 632,574 cubic meters in 2018 and 2019, 483,768 and 555,966 cubic meters, or 86% of the country's Falcata log production, were from Caraga Region (Forest Management Bureau, 2018, 2019). To improve monitoring and sustainable management of forest and other industrial tree plantations (ITPs), it is essential to have efficient and accurate mapping (Santillan and Gesta, 2021). Determination and mapping of plantation stand ages are also helpful to conservationists, landowners, and land use planning by allowing practical accounting of resource availability (Spracklen and Spracklen, 2021). Moreover, information on the stand age of Falcata plantations is essential for proper planning, marketing, harvesting, logistics, and transport capacity for the forthcoming wood availability. Plantation-level carbon cycle studies (e.g., estimation of biomass, carbon pools, and fluxes) can also benefit from the availability of stand age information (Chen et al., 2018).

The traditional method of determining the stand age information of Falcata and other trees is carried out through land surveys. However, this approach is costly, labour-intensive, and impractical on a large-scale basis. Nowadays, advances in Remote Sensing technologies and techniques provide an alternative approach in determining tree parameters, such as the stand age. Satellite remote sensing data availability, frequency, and coverage have grown significantly in recent years, mainly due to the European Union's Copernicus program, which includes Sentinel-1 and Sentinel-2 satellites. Images provided by the Sentinel and Landsat satellites are utilized in studies regarding land cover and forest monitoring, as they are free to access by everyone.

Stand age can be estimated by correlating it with factors that can be determined remotely, such as the spectral reflectance of optical satellite imagery (Cohen et al. 1995), like Sentinel-2, and radar backscattering coefficients from Synthetic Aperture Radar (SAR) images, like Sentinel-1 (Attarchi and Gloaguen, 2014; Akbari et al., 2021). Recent studies used optical images to estimate plantation stand ages, while others combined optical and SAR images (Wang et al., 2015; Akbari et al., 2021).

The application of machine learning-based classifiers to predict forest parameters, such as stand ages, is a relatively new area of research in recent years. To our knowledge, the use of remotely sensed data, particularly Sentinel-1 and Sentinel-2 images, for estimating the stand ages of Falcata plantations is not yet fully explored. Earlier attempts in this research area focused on the application of Support Vector Machine (SVM), Random Forest (RF), Maximum Likelihood, and Artificial Neural Network (ANN) classifiers in mapping Falcata plantations and estimating stand ages. In the study of Santillan et al. (2021) using Sentinel-

* Corresponding author

https://doi.org/10.5194/isprs-archives-XLVIII-4-W6-2022-465-2023

2 imagery, the ML classifier was found to be the best classifier in mapping Falcata plantations compared to three variants of SVM, NeuralNet (an ANN), and RF classifiers. For stand age estimation, a study by Pastor and Gono (2021) utilized all the bands of a single-date Sentinel-2 imagery for SVM and RF classification. Results showed SVM to have the highest overall accuracy, although RF was more consistent in estimating Falcata stand ages. However, the accuracies of the estimated stand ages are low, with Producer’s and User’s Accuracies well below 50%. Also, the approach implemented failed to detect all stand age classes. The use of a single-date image was considered to be the main reason for the unsatisfactory results.

The objective of this study is to evaluate the performance of machine learning classifiers, particularly SVM and RF, as well as multivariate regression models in estimating the stand ages of Falcata plantations in Caraga using multitemporal Sentinel-2 and Sentinel-1 images, including topographic variables such as elevation and slope.

2. METHODOLOGY

2.1 Overview

Figures 1 and 2 present the process flows for the estimation of Falcata stand ages using Sentinel-1 and Sentinel-2 images and derivatives, including topographic variables (elevation and slope).

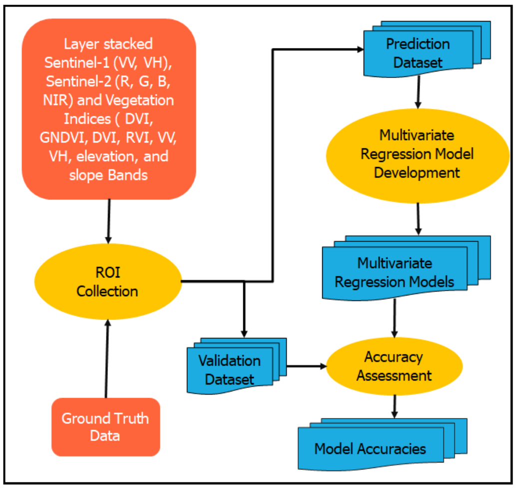

Figure 1. Steps employed in the development and accuracy assessment of multivariate regression models for Falcata stand age estimation.

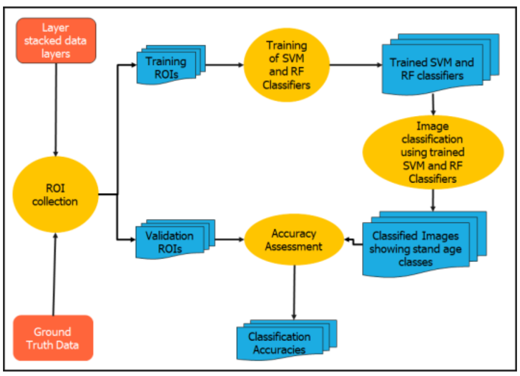

Figure 2. Steps employed in the application and accuracy assessment of Support Vector Machine (SVM) and Random Forest (RF) classifiers for Falcata stand age estimation.

2.2 Study Area

We selected Butuan City, in Caraga Region, Mindanao, Philippines (Figure 3) as the study area because of the large tracts of Falcata plantations established in various parts of the city. An ITP mapping conducted in 2021 estimated that 5,811.23 hectares of Falcata plantations exist in the city (Caraga State University, 2021).

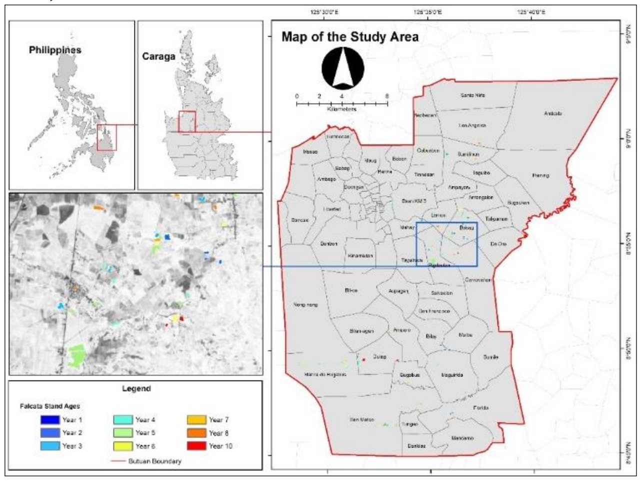

Figure 3. The study area: Butuan City, Caraga Region, Mindanao, Philippines. Also shown are the locations of plantations with different stand ages used in the regression model development and machine learning classifications.

2.3 Sentinel and Topographic Datasets Used

This study utilized four (4) Sentinel-1 images acquired in the years 2018, 2019, 2020 and 2021; and six (6) Sentinel-2 images acquired in the years 2015, 2016, 2018, 2019, 2020, and 2021. These images were all downloaded from the Copernicus Open Access Hub at https://scihub.copernicus.eu/dhus/#/home. The Sentinel-1 images were chosen to have image acquisition dates as close as possible to those of the Sentinel-2 images.

The multitemporal Sentinel-1 images are all Level-1 Ground Range Detected (GRD) Interferometric Wide Swath (IW) products, containing VV and VH polarization bands. Each product underwent pre-processing steps in the Sentinel Application Platform (SNAP) that includes orbital correction, thermal border noise removal, radiometric calibration, speckle filtering, terrain flattening, terrain correction, and conversion of backscatter coefficient values to decibels (dB). The output VV and VH polarization bands were exported as TIFF files with 10-m spatial resolution.

The multitemporal Sentinel-2 images are all Level 2A products, which means that they already passed through radiometric calibration, geometric correction, and atmospheric correction. To be consistent with the Sentinel-1 dataset, we selected the 10-m resolution images containing four (4) spectral bands, namely Band 2 (Blue), Band 3 (Green), Band 4 (Red), and Band 8 (Near-Infrared (NIR)). For each image date, four vegetation indices (VIs), namely Normalized Vegetation Index (NDVI), Difference Vegetation Index (DVI), Green Normalized Vegetation Index (GNDVI), and Ratio Vegetation Index (RVI) were also generated as additional bands.

The multitemporal Sentinel-1 and Sentinel-2 bands and image derivatives (i.e., Blue, Green, Red, NIR, NDVI, DVI, GNDVI, RVI, VV, and VH) were then layer-stacked together with elevation and slope layers derived from the Shuttle Radar Terrain Mission (SRTM) Digital Elevation Model (DEM). Then, it was subsetted to the portion of the study area for multivariate regression modelling and machine learning classification of Falcata stand ages.

2.4 Falcata Stand Age Data

The field data for the location and stand age information of Falcata plantations in the study area was requested from the Project 1. Development of a Geodatabase of Industrial Tree Plantations in Caraga Region Using Remote Sensing and GIS, Caraga Center for Geo-Informatics (CCGeo), Caraga State University. It consists of polygon shapefiles containing attribute data on the stand ages of Falcata plantations, ranging from 1-year-old to 10-year-old stand ages (Figure 3). This data was collected by the project in 2019 and 2020. Additional ground truth data collection was conducted in 2021 to increase the number of stand age data. Overall, the stand age data comprised of 117 polygons with area ranging from 80 m^2 to 92,000 m^2.

2.5 Training and Validation Data

Random points (pixels) within the Falcata stand age polygons were generated in Envi 5.3 software as regions of interests (ROIs). Then, it was partitioned into training (60%) and validation (40%). For multivariate regression modelling, these datasets are already sufficient. However, for machine learning classification, it was necessary to collect additional training and validation ROIs representing non-Falcata classes such as barren, built-up, cropland, forest, grassland, palm, and water bodies. For this purpose, the ground truth datasets gathered by Project 1 was also utilized and supplemented by corresponding high resolution satellite images provided by the Google Earth Pro application.

2.6 Multivariate Regression Modelling of Falcata Stand Ages

Values of year 2021 Sentinel-1 VV and VH polarization bands, year 2021 Sentinel -2 surface reflectance and vegetation indices, and SRTM DEM-derived elevation and slope, were extracted for each training and validation pixel and compiled into an MS Excel file for multivariate regression modelling. Prior to the regression model development, outliers were removed through the one-sigma rule approach resulting in the final training and validation dataset presented in Table 1.

Stand Age	No. of Training Pixels	No. of Validation Pixels
1	10	8
2	39	17
3	28	16
4	223	104
5	32	24
6	49	68
7	18	11
8	25	16
10	92	57
Total	516	321

Table 1. The sets of training and validation pixels for multivariate regression modelling.

Multivariate regression modelling was done using JASP 0.16 software (https://jasp-stats.org/). Both linear and nonlinear multivariate regression models were developed using different input explanatory variables with a backward data entry method using the prediction (training) dataset. The following are the sets of input explanatory variables considered in this study for the regression model development: i.) four spectral bands of Sentinel-2 2021 10-m resolution image; ii.) four spectral bands of Sentinel-2 10-m resolution 2021 image and derived NDVI, GNDVI, RVI, and DVI; iii.) four spectral bands of Sentinel-2 2021 image, derived NDVI, GNDVI, RVI, and DVI, and the VV and VH polarization bands of Sentinel-1 2021 image; and iv.) four spectral bands of Sentinel-2 2021 image, derived NDVI, GNDVI, RVI, and DVI, VV and VH polarization bands of Sentinel-1 2021 image, and the elevation and slope data layers.

With the backward data entry method, all variables ("predictors") that have been considered are initially entered into the model, and their statistical significance is calculated. Then, predictors with less than a given level of contribution ($p < 0.1$) are removed (Goss-Sampson, 2018). This process is repeated until all the remaining predictors are statistically significant to the outcome variable (stand age). This data entry method has been useful for fine-tuning regression models to select the best predictors available. For each model developed, multicollinearity between variables were examined using the tolerance statistic and Variance Inflation Factor (VIF) values. Variables with VIF > 10 and tolerance < 0.1 indicate the presence of multicollinearity (Goss-Sampson, 2018) which could affect the precision and reliability of the model prediction and can also be a cause of model overfitting. In such a case, these variables were dropped (because they are highly correlated with one or more of the variables), and the model is re-developed using the remaining variables. Each linear and nonlinear multivariate regression model developed were applied to the validation dataset to determine its coefficient of determination (R^2_{val}) and Root Mean Square Error (RMSE). The lower the RMSE value of a regression model, the more accurate it is in estimating the stand ages of Falcata plantations.

2.7 Machine Learning Classification-based Stand Age Estimation

SVM and RF classification of stand ages, as well as distinguishing the Falcata areas from the non-Falcata areas, was performed in Arc GIS Pro 2.7 utilizing the layer-stacked datasets and training ROIs. Each classifier was trained using eight (8) different classification input datasets (Table 2). The summary of classification training and validation pixels are presented in Table 3.

Classifier training default parameters in ArcGIS Pro were used. For the SVM classifier, the following parameters were used: maximum number of samples per class = 500; and RBF kernel type. For RF, the following parameters were used: maximum number of trees = 50; maximum tree depth = 30; and maximum number of samples per class = 1000. Several studies have indicated that default values for the SVM and RF classifier parameters are often a good choice (Wang et al., 2015; Zhao et al., 2018; Trisasongko, 2017; Xu et al., 2018). Then, the trained SVM and RF classifiers were applied to their corresponding input layer stacked images to generate classification results showing classes of Falcata stand ages, ranging from one (1) to 10 years old, and the other land cover classes, which are merged into one and labelled as "Non-Falcata". Each classification output was then subjected to accuracy assessment using the confusion matrix approach. Relevant accuracy measures such as Producer's and

User's Accuracy and Overall Accuracy were then calculated and used for comparison.

Classification Set	Input Data
1	Multitemporal Sentinel-1 and Sentinel-2 surface reflectance bands and vegetation indices (all years)
2	Multitemporal Sentinel-2 surface reflectance bands and vegetation indices (all years)
3	Year 2021 Sentinel-1 and Sentinel-2 images and vegetation indices
4	Year 2021 Sentinel-2 images and vegetation indices
5	Multitemporal Sentinel-1 and Sentinel-2 surface reflectance bands and vegetation indices (all years) + Elevation and Slope
6	Multitemporal Sentinel-2 surface reflectance bands and vegetation indices (all years) + Elevation and Slope
7	Year 2021 Sentinel-1 and Sentinel-2 images and vegetation indices + Elevation and Slope
8	Year 2021 Sentinel-2 images and vegetation indices + Elevation and Slope

Table 2. The different sets of input data for SVM and RF classifications.

Class	Number of Pixels for Training	Number of Pixels for Validation
Barren	2320	885
Built-up	2701	2597
Grassland	1795	1487
Cropland (Bare)	2122	1208
Cropland (Planted)	4478	1608
Palm	2180	584
Forest	2833	985
Water bodies	4459	1967
Falcata Year 1	141	19
Falcata Year 2	363	92
Falcata Year 3	239	231
Falcata Year 4	1510	633
Falcata Year 5	846	554
Falcata Year 6	898	480
Falcata Year 7	199	47
Falcata Year 8	276	208
Falcata Year 10	755	378

Table 3. Summary of training and validation data for pixel-based machine learning classification of Falcata stand age.

3. RESULTS AND DISCUSSION

3.1 Multivariate Regression Models for Stand Age Estimation

Eleven multivariate regression models were developed with R^2 values ranging from 0.23 to 0.81 by employing the surface reflectance bands of Sentinel-2 and derived VIs, polarization bands of Sentinel-1, and the elevation and slope data layers (Table 4). The modelling results suggest that the VV and VH polarization bands of Sentinel-1 imagery were useful predictors for the stand ages of Falcata plantations. Its inclusion in models that only utilized the spectral bands of Sentinel-2 images and derived VIs contributed to an improvement in the R^2 and RMSE values of the model.

When applied to the validation dataset, the models obtained estimation errors ranging from 1.72 to 3.58 years. The accuracy assessment showed that the eleven developed multivariate regression-based models for stand age estimation of Falcata plantations severely overestimated the actual stand ages of Falcata years 1 and 2, underestimated those of Falcata years 7, 8, and 10, and gave inaccurate results on the other Falcata stand ages classes. The tendencies of the models to overestimate and underestimate the actual stand ages of Falcata plantations must be considered in choosing this approach over the machine learning classifiers.

Model 4 and model 8 gained the highest model development R^2 of 0.81. However, these models performed poorly during validation. Model 4's validation R^2 was reduced to 0.36, and its RMSE is 2.92 years. On the other hand, model 8 validation R^2 was reduced to 0.14, and its RMSE of 3.58 years is the highest. These results indicate the unsuitability of the models for stand age estimation. Among the eleven models, it is model 7 that has the most consistent performance. It achieved an R^2 and R^2_{val} value of 0.61 and 0.55, respectively, and has the lowest estimation error (RMSE) value of 1.72 years, among other models The model includes the blue (B2), red (B4), and NIR (B8) bands of Sentinel-2, and the VV and VH polarization bands of Sentinel-1. A map showing the stand ages of Falcata plantations was then generated using the equation of model 7 and is presented in Figure 4.

Model	Equation	R^2	R^2_{val}	*RMSE (years)*
1	Age = 21.048 + 0.027(B4)-0.007(B8)	0.23	0.33	2.06
2	Age = 558.173-0.231(B3)-0.388(B4)+0.087(B8)-577.332(GNDVI)-17.556(RVI)	0.27	0.36	2.03
3	Age = 104.795 - 0.024(B2) - 0.021(B4) - 107.399(NDVI) - 1.941(VV)	0.54	0.47	1.84
4	Age = -14.631 - 0.037(B2) + 0.059(B3) - 0.125(B4) + 142.124(GNDVI) - 5.399(RVI) - 0.860(VV) - 0.196(VH) + 0.030(Elevation)	0.81	0.36	2.93
5	Age = exp (3.814 + 0.006(B4) - 0.001(B8)	0.23	0.37	2.11
6	Age = exp (47.352 - 0.057(B3) + 0.011(B8) + 83.986(NDVI) - 143.906(GNDVI) - 1.884(RVI))	0.26	0.36	2.08
7	**Age = exp (2.630 - 0.005(B2) + 0.006(B4) - 0.001(B8) - 0.262(VV) - 0.156(VH))**	**0.61**	**0.55**	**1.72**
8	Age = exp (-22.959 - 0.006(B2) + 0.018(B3) - 0.004(B8) + 45.341(GNDVI) - 0.138(VV) + 0.050(VH) + 0.007(Elevation))	0.81	0.14	3.58
9	Age = 592.771 - 55209.667(1/B4) + 554221.084(1/B8) - 482.509(1/NDVI)	0.25	0.35	2.03
10	Age = -6.322 + 13140.82(1/B2) - 4189.731(1/B4) + 69397.814(1/B8) + 96.841(1/VV)	0.50	0.46	1.85
11	Age = 452.399 + 8799.988(1/B2) - 40089.991(1/B4) + 3935656.600(1/B8) - 359.258(1/NDVI) + 35.136(1/VV) + 155.073(1/VH) - 42.348(1/Elevation)	0.75	0.48	2.14

Table 4. Multivariate regression models for Falcata stand age estimation. The R^2_{val} and *RMSE* values were computed using the validation dataset.

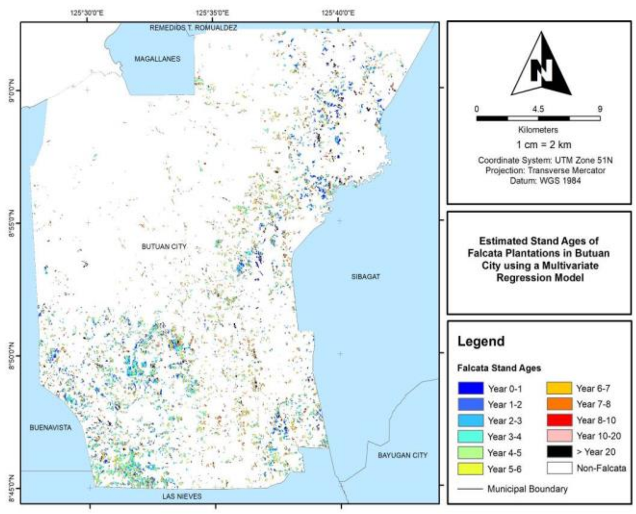

Figure 4. Falcata stand age map generated using Model 7.

3.2 Machine Learning-based Falcata Plantation Stand Age Estimation

A total of eight classification results (Figures 5 and 6) were generated using Support Vector Machine and Random Forest classifier when utilizing the spectral bands, derived VIs, and polarization bands of a single year and multitemporal Sentinel-2 and Sentinel-1 images in estimating the stand ages of Falcata plantations. An additional eight additional classification results were also generated, incorporating topographic variable-elevation and slope data layers in addition to the spectral, derived VIs, and polarization bands of a single year and of multitemporal Sentinel-2 and Sentinel-1 images. The results of accuracy assessment are summarized in Tables 6 and 7. The highest overall classification accuracy of 84.69% was achieved in classifying between the nine stand age classes of Falcata plantations and the non-Falcata class when using RF classifier and using multitemporal Sentinel-1 and Sentinel-2 data and VIs (Figure 7). On the other hand, the highest overall classification accuracy achieved by the SVM classifier is 82.50% in classifying between the nine stand age classes of Falcata plantations and the non-falcata class using the same dataset.

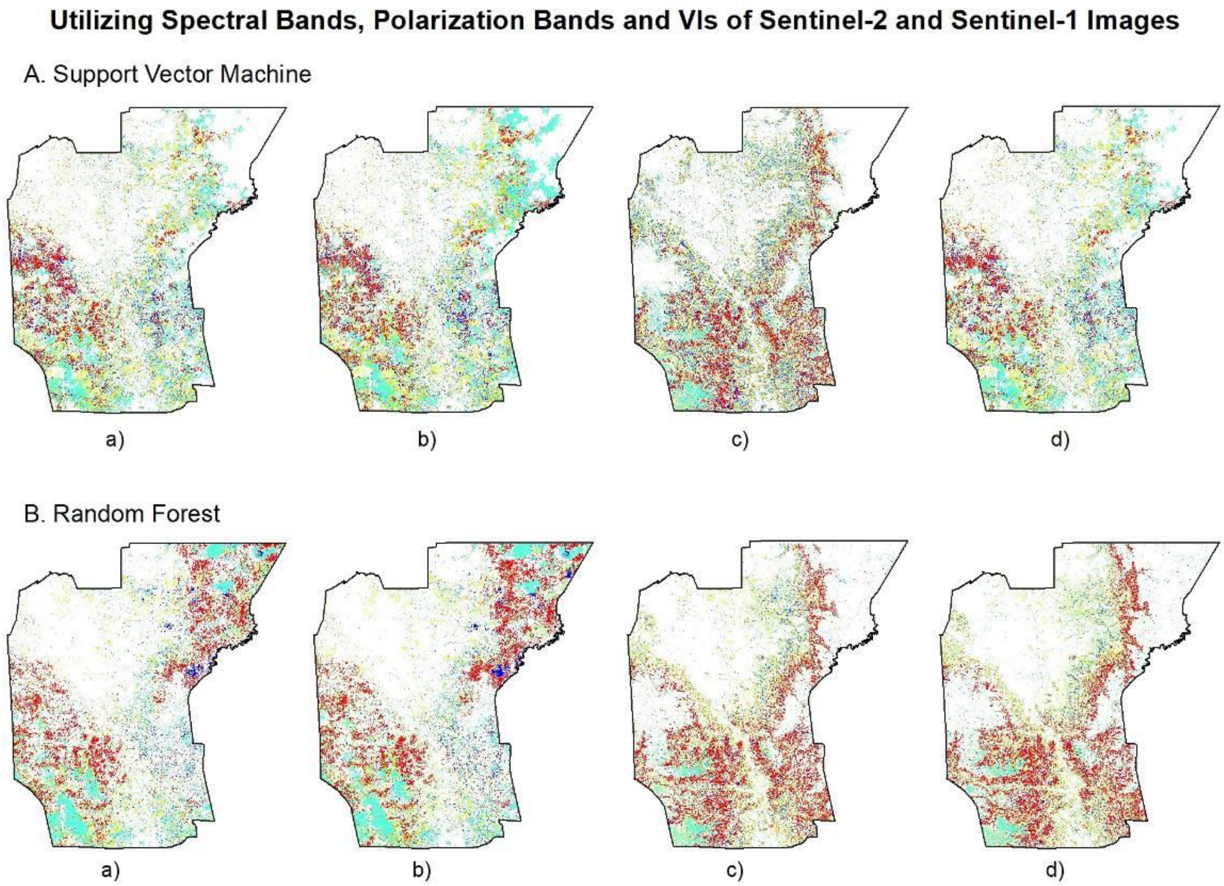

Figure 5. Classification results for Falcata stand ages using: a) Multitemporal Sentinel-1 and Sentinel-2 data with VIs, b) Multitemporal Sentinel-2 data with VIs c) Single Year Sentinel-1 and Sentinel-2 data with VIs, d) Single Year Sentinel-2 data with VIs.

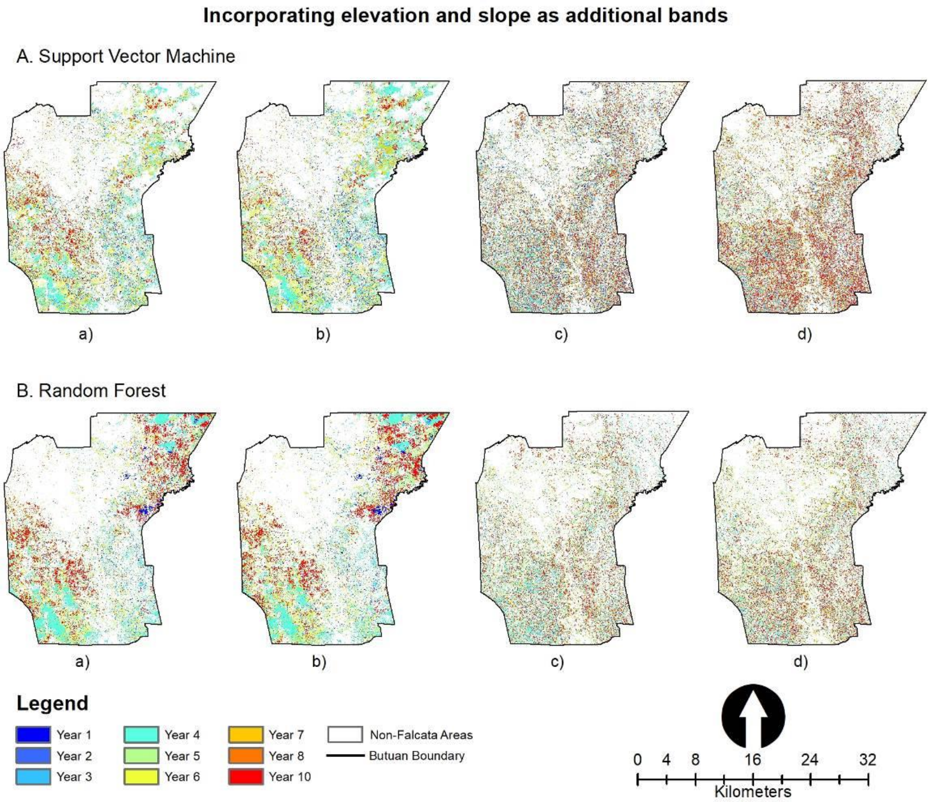

Figure 6. Classification results for Falcata stand ages incorporating elevation and slope as additional bands to the following data layers: a) Multitemporal Sentinel-1 and Sentinel-2 data with VIs, b) Multitemporal Sentinel-2 data with VIs c) Single Year Sentinel-1 and Sentinel-2 data with VIs, d) Single Year Sentinel-2 data with VIs.

			Year1 (%)	Year2 (%)	Year3 (%)	Year4 (%)	Year5 (%)	Year6 (%)	Year7 (%)	Year8 (%)	Year10 (%)	Non-Falcata (%)	OA
Single Year S2 + VIs	SVM	PA	15.79	0.00	18.61	7.27	17.15	19.38	0	0.48	33.33	92.09	77.58%
		UA	4.88	3.53	0	8.90	13.33	32.76	16.09	0	0.24	96.47	
	RF	PA	15.79	8.70	9.52	16.43	22.02	23.33	0	1.92	19.31	94.05	79.46%
		UA	5.08	7.02	13.10	17.90	25.36	19.34	0	2.90	11.20	95.62	
Single Year S2 and S1 + VIs	SVM	PA	26.32	18.48	6.06	20.06	14.26	15.63	4.26	2.40	14.02	91.62	76.98%
		UA	2.84	4.12	5.69	19.16	23.24	18.47	2.35	1.85	7.89	97.01	
	RF	PA	21.05	18.48	5.63	25.43	18.95	19.58	2.13	0.48	20.63	94.63	80.12%
		UA	7.02	8.72	15.85	20.13	23.44	20.61	2.27	0.91	12.72	96.01	
Multitemporal S2 + VIs	SVM	PA	36.84	13.04	16.45	31.60	40.07	27.92	10.64	3.85	4.76	93.94	81.11%
		UA	10.45	14.77	16.18	27.65	36.43	18.34	5.00	3.83	58.33	96.27	
	RF	PA	42.11	35.87	13.85	42.50	47.11	25.00	2.13	4.33	33.60	96.35	84.28%
		UA	27.59	30.84	25.00	52.33	37.45	22.14	1.30	10.34	31.83	95.83	
Multitemporal S1 and S2 + VIs	SVM	PA	36.84	20.65	22.08	33.49	41.16	27.08	2.13	8.65	52.91	94.10	82.50%
		UA	14.58	18.27	21.43	31.55	37.44	19.15	1.33	12.08	66.23	96.09	
	RF	PA	47.37	23.91	10.82	43.44	46.03	27.29	0	12.50	52.91	96.12	84.69%
		UA	26.47	22.92	21.74	52.18	35.66	23.14	0	35.62	47.62	95.91	

Table 5. Accuracies of the machine learning classification-based Falcata stand age estimations.

			Year1 (%)	Year2 (%)	Year3 (%)	Year4 (%)	Year5 (%)	Year6 (%)	Year7 (%)	Year8 (%)	Year10 (%)	Non-Falcata (%)	OA
Single Year S2 + VIs	SVM	PA	31.58	42.39	19.48	18.48	18.41	21.04	6.38	0	15.61	97.21	77.26%
		UA	4.88	7.56	14.11	16.53	31.10	18.13	2.61	0	13.75	97.36	
	RF	PA	26.32	30.43	21.21	16.75	17.69	21.46	4.26	0.48	20.90	95.02	80.41%
		UA	29.41	7.91	16.78	25.30	25.93	17.02	1.74	0.71	20.26	95.60	
Single Year S2 and S1 + VIs	SVM	PA	36.84	34.78	23.81	24.96	22.38	24.17	0	0.96	32.80	91.64	78.73%
		UA	7.53	6.03	17.24	28.83	36.36	17.90	0	0.83	27.49	97.37	
	RF	PA	31.58	38.04	15.58	20.54	16.97	24.17	0	0.96	21.96	95.11	80.71%
		UA	15.00	10.48	13.69	23.42	28.92	18.04	0	1.69	22.25	96.13	
Multitemporal S2 + VIs	SVM	PA	36.84	13.04	16.45	31.60	40.07	27.92	10.64	3.85	4.76	92.09	79.28%
		UA	16.28	12.50	15.45	22.86	39.02	17.52	4.42	2.86	12.24	96.27	
	RF	PA	36.84	28.26	15.58	39.65	46.21	28.13	4.26	9.13	23.81	95.35	83.20%
		UA	28.00	30.95	24.66	47.81	41.49	15.92	2.47	15.45	27.69	96.48	
Multitemporal S1 and S2 + VIs	SVM	PA	42.11	19.57	22.51	33.33	41.70	28.33	2.13	8.65	59.52	93.04	81.88%
		UA	14.81	16.67	20.31	26.64	41.25	19.05	1.32	12.59	68.39	96.36	
	RF	PA	47.37	22.83	19.48	42.50	45.31	31.88	0	11.06	24.87	95.67	83.77%
		UA	39.13	23.60	23.94	45.90	40.81	19.47	0	38.33	31.33	96.34	

Table 6. Accuracies of the machine learning classification based Falcata stand age estimations, with elevation and slope as additional bands.

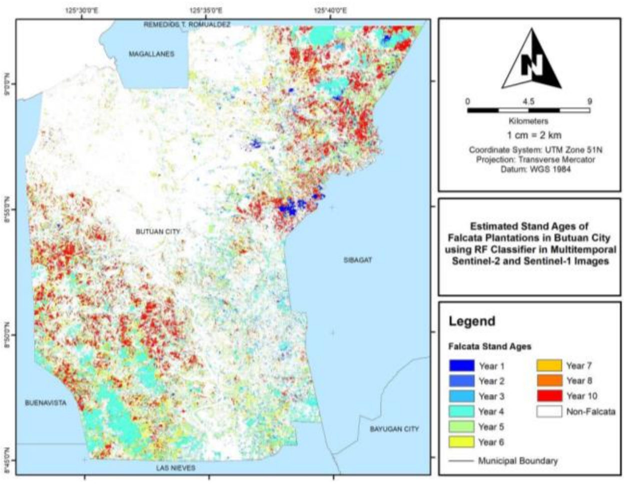

Figure 7. Map of the stand ages of Falcata plantations generated through RF classification of multitemporal Sentinel-1 and Sentinel-2 data and VIs.

Both RF and SVM classifiers performed best when using multitemporal Sentinel-2 and Sentinel-1 data. SVM and RF classification of Falcata plantation stand ages using multitemporal data obtained higher individual class accuracies for the Falcata stand age classes compared to when using only single-year datasets. It is also worth noting that the overall accuracy of the classification increases with the use of more data layers in the training and application of machine learning classifiers. Single-year datasets on their own were generally ineffective in classifying Falcata plantation stand ages as they produced low individual class accuracies for years 7, 8, and 10. On the other hand, the integration of Sentinel-1 to Sentinel-2 imagery made little improvement on the overall accuracy of RF and SVM classifiers as corresponding classification results were similar when only using Sentinel-2 data. This might be due to the nature of the Sentinel-1 C-Band SAR not being able to penetrate tree canopies to retrieve relevant tree structure data. Moreover, the incorporation of elevation and slope as additional bands did not improve the overall classification accuracies of the RF classifier and the SVM classifier, as the overall classification accuracies of the datasets that did not include elevation and slope data layers were not surpassed. One of the possible reasons for this result is that Falcata plantations are planted in the study area in both low and high elevations, with either flat or sloping terrain, Consequently, the plantation ages also vary regardless of the elevation or slope such that these variables are not useful in the stand age classification.

4. SUMMARY AND CONCLUSIONS

This study developed multivariate regression models and evaluated the performance of machine learning classifiers, particularly SVM and RF in estimating the stand ages of Falcata plantations in Caraga using multitemporal Sentinel-1 and Sentinel-2 images, vegetation indices, and topographic variables such as elevation and slope.

From the results, the following major conclusions are drawn:

- Stand ages of falcata plantations can estimated using an exponential, multivariate regression model employing the VV and VH bands of Sentinel-1 and the Blue, Red, and NIR bands of Sentinel-2 with an accuracy of ±1.72 years.
- The other bands of Sentinel-2, including vegetation indices, are not good predictors of Falcata stand ages

https://doi.org/10.5194/isprs-archives-XLVIII-4-W6-2022-465-2023 |

due to the low R^2 values and higher RMSE of the regression models incorporating these variables when applied to validation data.
- A pixel-based classification approach using RF can be used to estimate Falcata stand age, but input data layers shall include multitemporal Sentinel-1 and Sentinel-2 data and VIs for better accuracy.

ACKNOWLEDGEMENTS

This work is an output of "Project 1. Development of a Geodatabase of Industrial Tree Plantations (ITP) in Caraga Region Using Remote Sensing and GIS" under the Niche Centers in the Regions for R&D (NICER) Program: Industrial Tree Plantations Research and Innovation Center (ITP Center) for Upgrading the Wood-Based Industry, funded and supported by the Philippines’ Department of Science and Technology (DOST) and the Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development (PCAARRD).

REFERENCES

Akbari, V., Solberg, S., Puliti, S., 2021. Multitemporal Sentinel-1 and Sentinel-2 Images for Characterization and Discrimination of Young Forest Stands under Regeneration in Norway, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 14, pp. 5049–5063, doi: 10.1109/JSTARS.2021.3073101.

Attarchi, S., Gloaguen, R., 2014. Classifying complex mountainous forests with L-Band SAR and landsat data integration: A comparison among different machine learning methods in the Hyrcanian forest,” Remote Sens., vol. 6, no. 5, pp. 3624–3647, doi: 10.3390/rs6053624.

Caraga State University, 2021. Terminal Report - Project 1. Development of a Geodatabase of Industrial Tree Plantations in Caraga Region Using Remote Sensing and GIS. Retrieved from http://mindset.ccgeo.info:82/dataset/itp-center-project-1-terminal-report (July 8, 2022).

Chen, B., Xiao, X., Wu, Z., Yun, T., Kou, W., Ye, H., ... & Cao, J. (2018). Identifying establishment year and pre-conversion land cover of rubber plantations on Hainan Island, China using landsat data during 1987–2015. Remote Sensing, 10(8), 1240.

Cohen, W. B., Spies, T. A., Fiorella, M., 1995. Estimating the age and structure of forests in a multi-ownership landscape of western oregon, U.S.A, Int. J. Remote Sens., vol. 16, no. 4, pp. 721–746, doi: 10.1080/01431169508954436.

Doloriel, N. S., 2017. Marketing Practices of Falcata Growers in Tagbina, Surigao Del Sur, Philippines, Int. J. Contemp. Appl. Res., vol. 4, no. 6, pp. 54–61.

Forest Management Bureau, 2018. “Philippine Forest Statistics 2018,” vol. 52, [Online]. Available: https://forestry.denr.gov.ph/index.php/statistics/philippines-forestry-statistic.

Forest Management Bureau, 2019. “Philippine Forestry Statistics,” vol. 53, no. 9, pp. 1689–1699.

Gono, Y. E., Pastor, R., 2021. Evaluation of Machine Learning Classifiers in Estimating Stand Ages of Falcata Plantations,” 2021. Thesis. Retrieved from http://mindset.ccgeo.info:82/dataset/evaluation-of-machine-learning-classifiers-in-estimating-stand-ages-of-falcata-plantations (July 8, 2022).

Goss-Sampson, M. A., 2018. Statistical Analysis in JASP: A Guide for Students, vol. 1, no. 1, pp. 6–8, [Online]. Available: https://static.jasp-stats.org/Statistical Analysis in JASP - A Students Guide v1.0.pdf.

Santillan, J. R., Gesta, J. L. E., 2021. Evaluation of machine learning classifiers for mapping falcata plantations in sentinel-2 image, Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch., vol. 43, no. B3-2021, pp. 103–108, doi: 10.5194/isprs-archives-XLIII-B3-2021-103-2021.

Spracklen, B., Spracklen, D. V., 2021. Synergistic use of Sentinel-1 and Sentinel-2 to map natural forest and acacia plantation and stand ages in north-central Vietnam, Remote Sens., vol. 13, no. 2, pp. 1–19, doi: 10.3390/rs13020185.

Trisasongko, B. H., 2017. Mapping stand age of rubber plantation using ALOS-2 polarimetric SAR data, Eur. J. Remote Sens., vol. 50, no. 1, pp. 64–76, doi: 10.1080/22797254.2017.1274569.

Wang, H., Zhao, Y., Pu, R., Zhang, Z., 2015. Mapping Robinia pseudoacacia forest health conditions by using combined spectral, spatial, and textural information extracted from IKONOS imagery and random forest classifier, Remote Sens., vol. 7, no. 7, pp. 9020–9044, doi: 10.3390/rs70709020.

Xu, C., Manley, B., Morgenroth, J., 2018. Evaluation of modelling approaches in predicting forest volume and stand age for small-scale plantation forests in New Zealand with RapidEye and LiDAR, Int. J. Appl. Earth Obs. Geoinf., vol. 73, no. June, pp. 386–396, doi: 10.1016/j.jag.2018.06.021.

Zhao, Q., Yu, S., Zhao, F., Tian, L., Zhao, Z., 2018. Comparison of machine learning algorithms for forest parameter estimations and application for forest quality assessments, For. Ecol. Manage., vol. 434, pp. 224–234, 2019, doi: 10.1016/j.foreco.2018.12.019.