USING MAXENT IN FINDING SUITABLE LOCATIONS FOR ESTABLISHING FALCATA TREE PLANTATIONS IN CARAGA REGION, MINDANAO, PHILIPPINES

Jojene R. Santillan[1,2,3], Arnaldo C. Gagula[1,2,3], and Meriam Makinano-Santillan[1,2,3]

[1]Caraga Center for Geo-informatics, Caraga State University, Butuan City, Philippines
[2]Department of Geodetic Engineering, College of Engineering and Geosciences, Caraga State University, Butuan City, Philippines
[3]Industrial Tree Plantation Research and Innovation Center, Caraga State University, Butuan City, Philippines
Email: jrsantillan@carsu.edu.ph, acgagula@carsu.edu.ph, mmsantillan@carsu.edu.ph

KEY WORDS: Falcata, Suitability Mapping, MaxEnt, Maximum Entropy, Caraga Region, Mindanao, Philippines

ABSTRACT: The Caraga Region in Mindanao, Philippines, is considered a significant contributor in log production, specifically due to Falcata (*Paraserianthes falcataria*) plantations. Over 80% of the country's Falcata log production came from Caraga Region in 2019. Among the challenges faced by the tree-growers is finding a suitable location for the establishment of new plantations. We used MaxEnt, a machine learning Species Distribution Modeling (SDM) based on Maximum Entropy principles, for this study's Falcata plantation suitability modeling and mapping. This approach used 2,125 Falcata location points distributed in the region, biophysical factors (i.e., Elevation, Slope, Aspect, and the like), and bioclimatic factors (i.e., Annual Mean Temperature, Isothermality, and Annual Precipitation, among others). The model was found to have acceptable model performance based on the average training and test Area Under the Curve (AUC) values of 0.76 and 0.73. A 1 km x 1 km Falcata suitability map was generated using the model. The map shows that 12% of the region has high suitability, while 23% and 30% have moderate and low suitabilities. On the other hand, 35% of the region was not suitable for Falcata plantation establishment.

1. INTRODUCTION

Falcata *(Paraserianthes falcataria (L.) Nielsen)* (Fig. 1.) is a preferred tree species in industrial tree plantations because of its fast growth, grows in various soil textures, and is acceptable for wood production (Krisnawati H. et al., 2011). The 2019 Philippine Forest Statistics of Forest Management Bureau (FMB) of the Department of Environment and Natural Resources (DENR-FMB, 2020) showed that the Caraga Region contributed 64.12% of the country's total log production and 87.89% of the Falcata total log production. It is a large tree that grows up to 40 meters in height and 20 to 100 dbh (Alipon, M. et al., 2017). Woods of Falcata is an excellent source to produce pulpwood and plywood (Krisnawati H. et al., 2011) and the most preferred raw materials for wood (Alipon, M. et al., 2017).

Figure 1. Falcata Plantation in Caraga Region, Mindanao, Philippines.

An earlier study was conducted to assess the suitability of Falcata plantations in the Caraga Region (Santillan et. al., 2021). This study utilized MaxEnt, a machine learning Species Distribution Model (SDM), with bioclimatic datasets (i.e., BIO 2, BIO 4, BIO 6, BIO 7, BIO 14, BIO 15, BIO 17, BIO 18), biophysical datasets (i.e., elevation, aspect, slope, land cover, soil type, and land cover), solar radiation, and wind speed. In selecting the environmental variables, this study assessed the collinearity of each variable based on the result of Principal Components Analysis (PCA) of the raster files corresponding to the variables. From this study, 925,700 hectares in the region were found to be "suitable" for establishing Falcata plantations (Santillan et al., 2021). According to (Feng, X. et al., 2019), there were two types of collinearity influence in regression type models. First is the effect on model training caused by the degree of predictor collinearity. The other is the effect on model transfer caused by differences in the correlation structure of predictor variables between training and testing regions. Therefore, deciding what collinearity methods are essential.

In this work, we conducted Falcata suitability modeling and mapping using MaxEnt. Specifically, we utilized the extracted pixel values of bioclimatic, biophysical, solar radiation, and wind speed corresponding to the spatial locations of Falcata species in the collinearity analysis to find the final set of variables for inclusion in the model.

2. MATERIALS AND METHODS

2.1 Study Area

The Caraga Region (Fig. 2) has long been considered the biggest producer of significant forest products in the Philippines. It has been dubbed as "the timber corridor of the Philippines." It has a total land area of 1,884,697 ha, of which 71% are classified as forestlands, and 29% are certified alienable and disposable. The region is characterized by mountainous areas and flat and rolling lands, with its most productive agricultural area along the Agusan River Basin. It has a Type II climate, with no pronounced wet and dry season; heavy rains are usually experienced from November to February (Wikipedia, 2021).

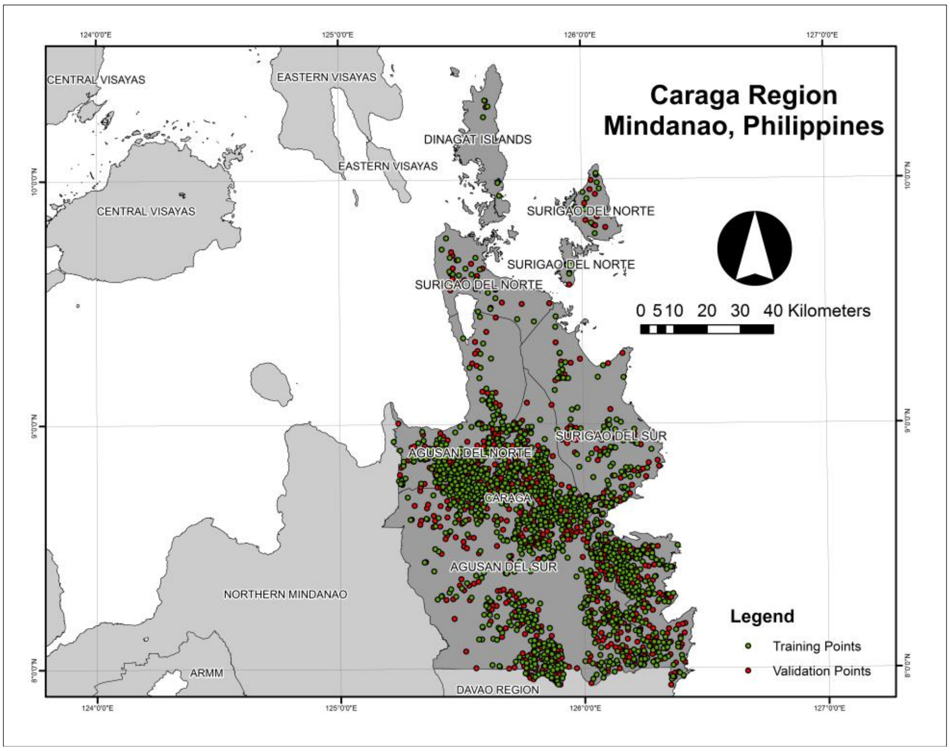

Figure 2. Map of Caraga Region, Mindanao, Philippines. Presented in the map are locations of Falcata Plantations used in MaxEnt modeling.

2.2 Falcata Location Data

The location data were gathered from the map of confirmed Falcata species derived using Sentinel-2 image using Maximum Likelihood Classification with greater than 90% accuracy and refined through high resolution google earth images. Only the Falcata with at least 1 hectare was considered with 1,224 m distance to each other. The "Random Points Generator" extension in ArcView software was used to select points from Falcata plantations' centroids randomly. This resulted in 1,065 points which were saved as "training" points, and another 1,060 points were saved as "validation" points (Table 1). A comma space value (CSV) file was generated containing the Universal Transverse Mercator (UTM 51) WGS 1984 grid coordinates of each point used in MaxEnt.

Table 1. The total number of Falcata location points in the Caraga Region.

Province Name	Training Points	Validation Points	Total Number of Points
Agusan del Norte	208	206	414
Agusan del Sur	557	555	1,112
Surigao del Norte	33	33	66
Surigao del Sur	261	262	523
Dinagat Islands	6	4	10
Total	1,065	1,060	2,125

2.3 Environmental Variables

There are 26 environmental variables identified, including 19 bioclimatic variables, solar radiation, wind speed, elevation, slope, aspect, soil type, and land cover (Figure 3). The 19 bioclimatic variables (BIO 01-BIO 19), including the solar radiation and wind speed, were downloaded from https://www.worldclim.org/data/worldclim21.html. It has a spatial resolution of approximately 1 km and is in GeoTIFF format, derivatives of the WorldClim version 2.1 climate data for 1970-2000. The biophysical variables include the Digital Elevation Model that was derived from SAR interferometry with a spatial resolution of 5 meters, soil data that was obtained from the Department of Agriculture - Bureau of Soils and Water Management (DA-BSWM), and landcover data from the Philippines' National Mapping and Resource Information Authority (NAMRIA). The DEM was resampled into 1 km spatial resolution to be consistent with the 19 bioclimatic variables, solar radiation, and wind speed. The resampled DEM becomes the base layer to generate aspect and slope data using ArcGIS 10.8 software. The soil and land cover data are also converted to raster data with 1 km spatial resolution for consistency. The data layers were layer-stacked using ArcGIS 10.8 software to ensure the exact spatial resolution (1 km) and coverage and raster dimensions. Each layer was then exported to ASCII (*.asc) format in preparation for MaxEnt modeling.

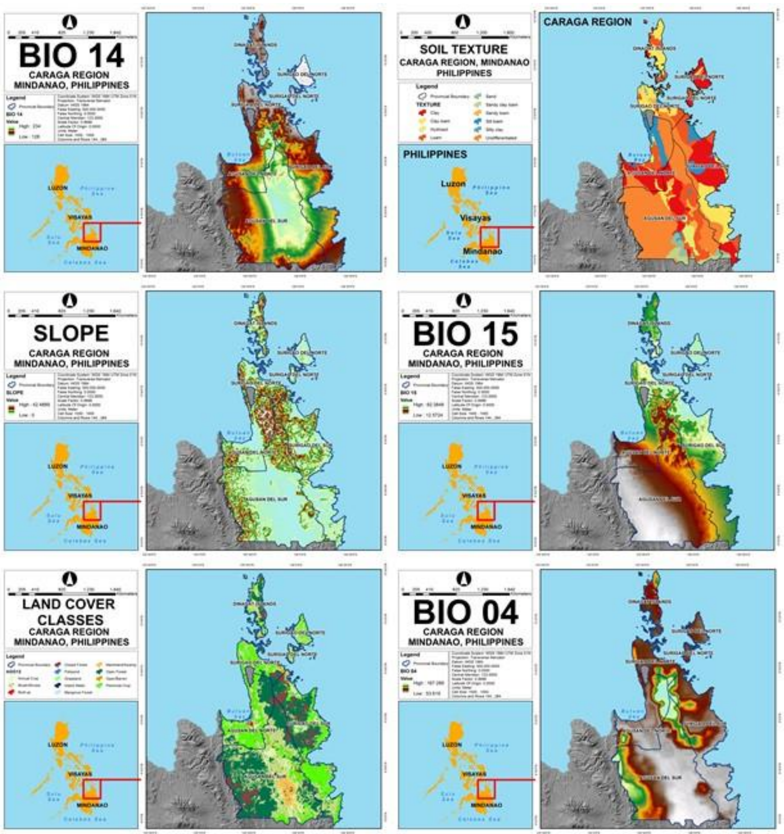

Figure 3. Example of environmental variables used in MaxEnt Modeling. These 6 variables have the highest percent contribution on determining suitable sites for Falcata plantations.

2.4 Environmental Variable Selection

Collinearity of the environmental variables needs to be quantified so that only variables that are not collinear with other variables must be considered in the Species Distribution Modeling (SDM). Pixel values were extracted from the 26 environmental variables using the locations of confirmed Falcata plantations. Afterward, the values underwent statistical analysis (correlation analysis). The Pearson's correlation coefficients (r) between the variables were calculated to facilitate collinearity analysis. Variables with a value of $r \geq 0.7$ are considered correlated, and only one of them should be selected and included in the MaxEnt model. The variable selection reported on (Garcia, K. et al., 2013) was used to minimize the finalizing of the environmental variables. This method underwent initial runs of all the environmental variables. Based on its outputs, only one variable (variable with highest percent contribution among others) was chosen from the variables with high correlation values in the final MaxEnt Model (Garcia, K. et al., 2013). As a result, only 14 out of 26 environmental variables were considered (Table 2).

Table 2. List of environmental variables. Highlighted variables are those included in the final MaxEnt model.

BIO1	Annual Mean Temperature
BIO2	Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO3	Isothermality [(BIO 02/BIO 07) x100)]
BIO4	Temperature Seasonality (standard deviation x 100)
BIO5	Maximum Temperature of Warmest Month
BIO6	Minimum Temperature of Coldest Month
BIO7	Temperature Annual Range (BIO 05 – BIO 06)
BIO8	Mean Temperature of Wettest Quarter
BIO9	Mean Temperature of Driest Quarter
BIO10	Mean Temperature of Warmest Quarter
BIO11	Mean Temperature of Coldest Quarter
BIO12	Annual Precipitation
BIO13	Precipitation of Wettest Month
BIO14	Precipitation of Driest Month
BIO15	Precipitation Seasonality (Coefficient of Variable)
BIO16	Precipitation of Wettest Quarter
BIO17	Precipitation of Driest Quarter
BIO18	Precipitation of Warmest Quarter
BIO19	Precipitation of Coldest Quarter
SRAD	Solar radiation
WIND	Wind speed
LCOV	Land-cover type
SOIL	Soil type
ELEV	Elevation
ASPECT	Aspect
SLOPE	Slope

2.5 MaxEnt Modeling

MaxEnt Version 3.4.1, downloaded from https://biodiversityinformatics.amnh.org/open_source/maxent/, was used in this study. In this study, the model runs in 15 replicates with 5000 iterations to have adequate time for convergence. Each iteration uses a random partition of the presence data through subsample, in which 75% was used for modeling, and 25% was used for testing. The convergence threshold used was 0.00001 and 10,000 for maximum background points. Random seeding, as well as jack-knife test for variable importance, were also enabled. The model was evaluated using the area Under the Curve (AUC) statistic, calculated from the Receiver Operating Characteristic (ROC) based on training and testing data. The following differentiation of performance levels was used: excellent (>0.9), good (0.8–0.9), accepted (0.7–0.8), poor (0.6–0.7), and unsatisfactory (<0.6) (Rojas-Briceño,

N. B., et al., 2020). This is the range of the value of the final potential species distribution map, which was reclassified into four branches of potential habitat: "high" (>0.6), "moderate" (0.4–0.6), and "low" (0.2–0.4), and 'not suitable' (<0.2) (Zhang, K. et al., 2019).

3. RESULTS AND DISCUSSIONS

3.1. MaxEnt Model Results

Based on the result, area Under the Curve (AUC) values range from 0.75-0.76 and 0.71-0.74 for training and testing presence data. The average training and test AUC values are 0.76 and 0.73, respectively, indicating acceptable MaxEnt model performance.

Figure 4 shows the percent contribution of each environmental variable in the MaxEnt Model. It reveals that BIO 04 (Temperature Isothermality) gains the highest percentage contribution with 32.9%. On the other hand, ASPECT is the least contributing variable with 0.2%.

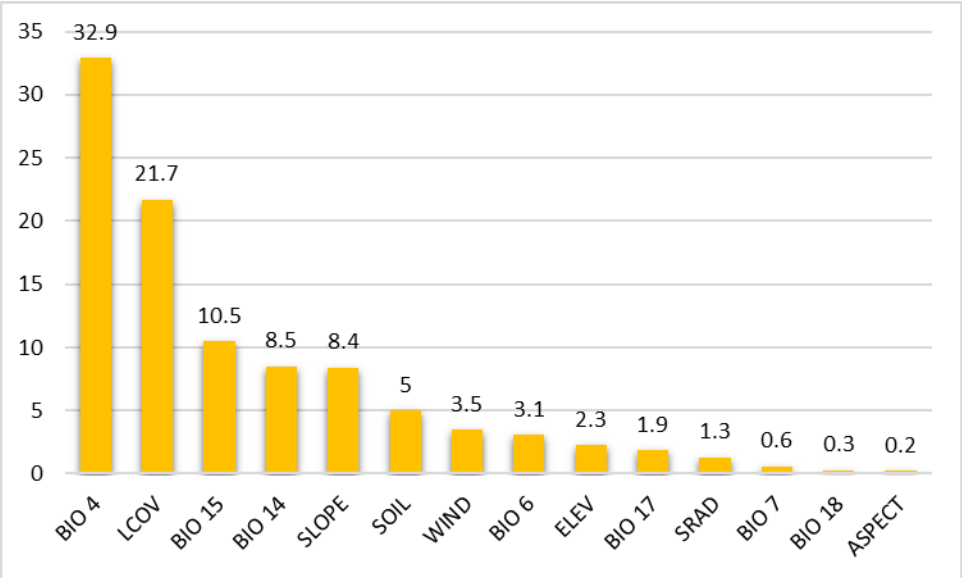

Figure 4. Percent Contributions of environmental variables to the Falcata MaxEnt model.

Marginal curves (Figure 5 and Figure 6) were generated to show how each environmental variable affects the MaxEnt prediction. The curves show how the predicted probability of presence changes as each environmental variable varies, keeping all other environmental variables at their average sample value. It shows that for BIO 4which has the highest percent contribution, higher logistic output was obtained during the lower temperature. It means that existing Falcata plantations in the region can be found in areas where temperature seasonality has lower values.

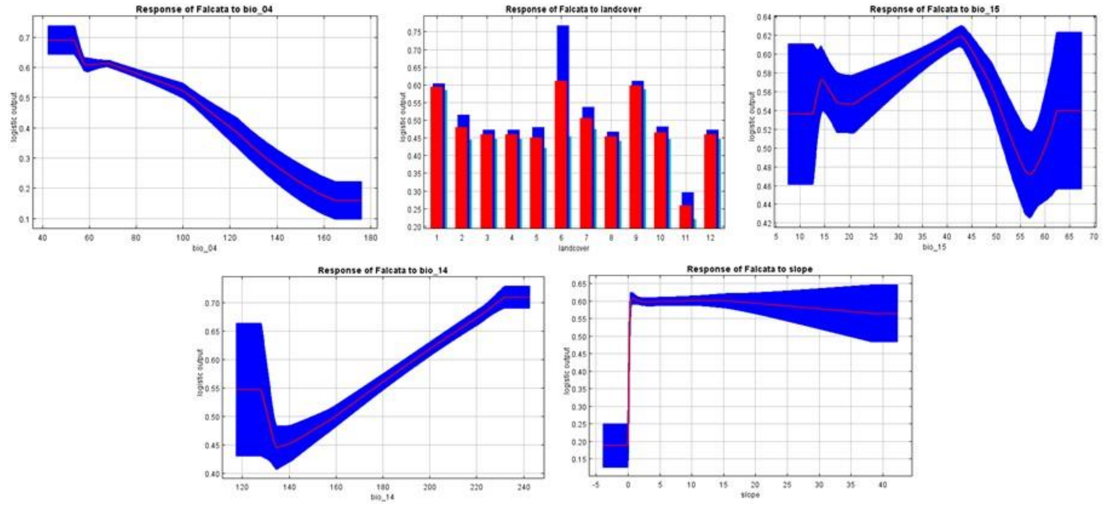

Figure 5. The marginal response curves of BIO 4, land cover, BIO 15, BIO 14, and Slope.

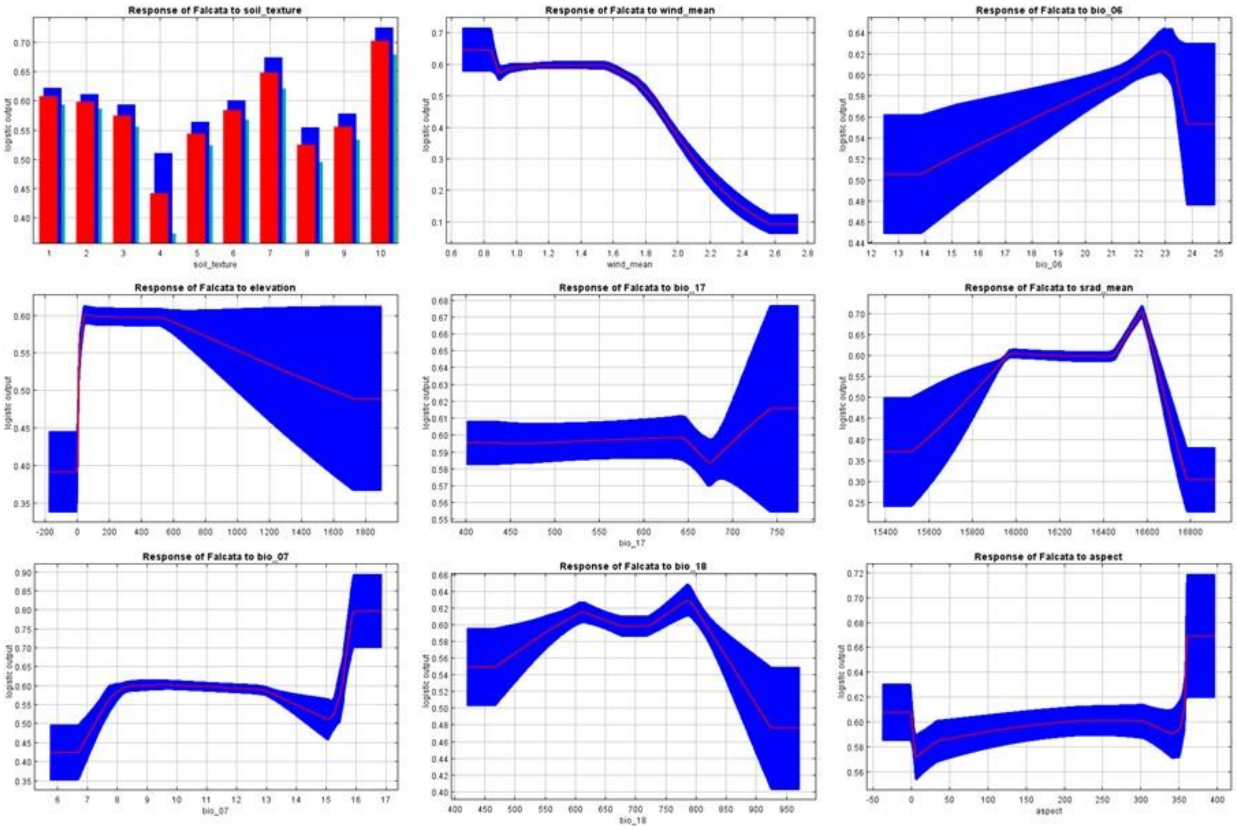

Figure 6. The marginal response curves of soil texture, wind speed, BIO 6, elevation, BIO 17, solar radiation, BIO 7, BIO 18, and aspect.

3.2 Falcata Suitability Map

Figure 7 shows the Falcata Suitability Map generated using MaxEnt. The result shows 212,985 hectares (12%) of the Caraga Region was classified as high suitable, 421,747 hectares (23%) were medium suitable, 538,951 hectares (30%) were low suitable, and 619,380 (35%) were classified as unsuitable. In terms of area, the total suitable areas in the region are 1,173,683 hectares.

Figure 8 shows the percentage comparison of Caraga Region provinces for Falcata suitability based on its land area in terms of statistics per province. It shows that Agusan del Sur has more "High Suitable" areas, followed by Agusan del Norte and Surigao del Sur with 129,864 hectares, 47,839 hectares, and 35,282 hectares, respectively. In contrary, Agusan del Sur has more “Not Suitable” areas followed by Surigao del Sur, Surigao del Norte, Agusan del Norte, and Dinagat Islands containing 238,305 hectares, 148,125 hectares, 92,522 hectares, 90,522 hectares, and 49,906 hectares, respectively.

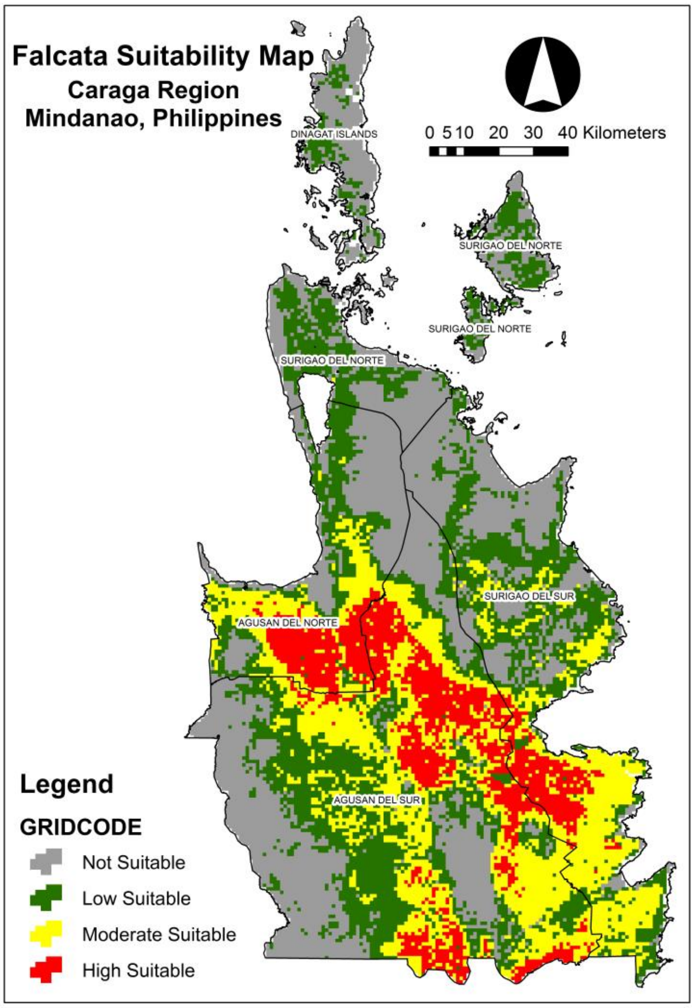

Figure 7. The Falcata Suitability Map for Caraga Region, Mindanao, Philippines.

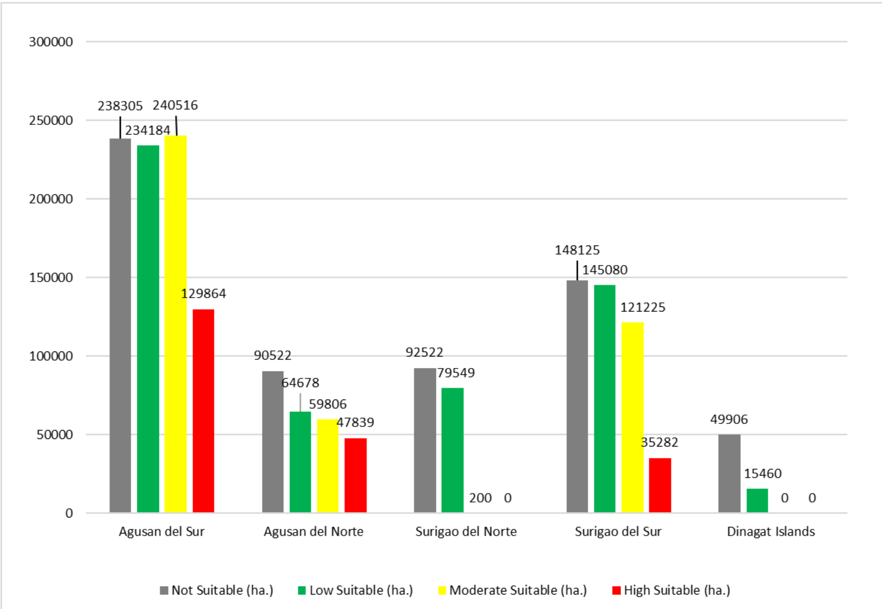

Figure 8. The comparison of Caraga Region provinces in terms of Falcata suitability. Values are in hectares.

Figure 9 shows the comparison of the total suitable areas in each province to the actual total areas of Falcata plantations. It shows that Agusan del Sur has gains the highest suitable areas, followed by Surigao del Sur, Agusan del Norte, Surigao del Norte, and Dinagat Islands. In addition, Caraga Region provinces have only less than 1% of the total suitable areas planted with Falcata. This implies that the region still has more areas that are potential in growing Falcata plantations.

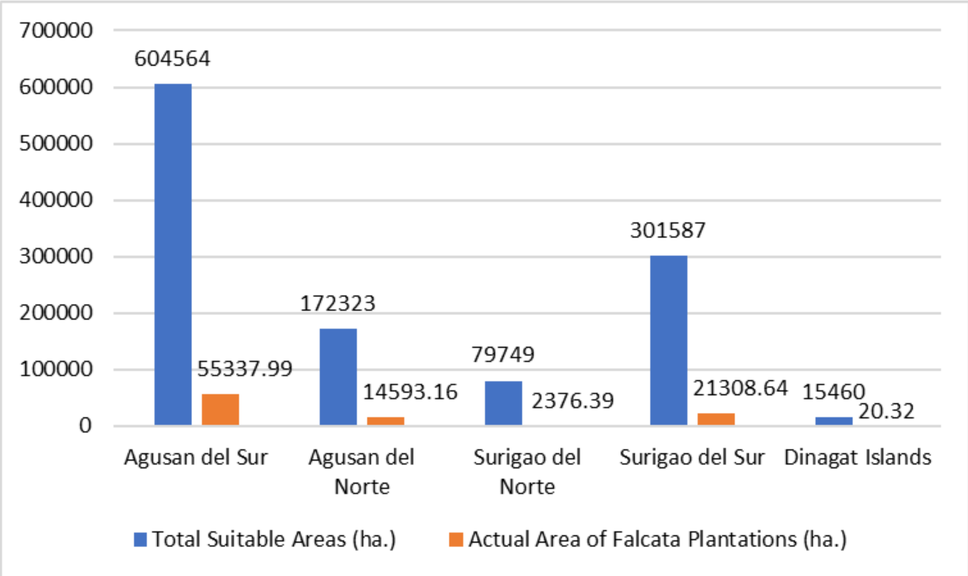

Figure 9. The comparison of the total suitable areas with the total actual area of Falcata plantations in the Caraga Region. Values are in hectares.

The results differ when applying another collinearity analysis approach in the previously conducted study (Santillan, J. et al., 2021). Using the extracted values of Falcata points instead of data layers in collinearity analysis, BIO 02 (Mean Diurnal Range) has been removed from the final environmental variables for MaxEnt modeling and the total suitable areas in the region increased by 247, 983 hectares.

4. CONCLUSION

In this study, we conducted suitability analysis for Falcata in Caraga Region using MaxEnt. The study has successfully identified areas for Falcata plantations with training and test AUC values of 0.76 and 0.73, respectively. BIO 4 (Temperature Seasonality) gains the highest percentage contribution with 32.9%, and ASPECT has the lowest percent contributions among others with 0.2%. The suitability map revealed 1,173,683 hectares of land where Falcata plantations can be potentially established.

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 (ITPS) 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

Alipon, M., Alcachupas, P., Bondad, Elvina, Cortiguerra, E., 2016. Assessing the Utilization of Falcata [Falcataria moluccana (Miq.) Barneby & J. W. Grimes] for Lumber Production. Philippine Journal of Science 145 (3): 225-235.

Feng, X., Park, D., Liang, Y., Pandey, R., Papes, M., 2019. Collinearity in ecological niche modeling: Confusions and challenges. Ecology and Evolution. Volume9, Issue1, Pages 10365-10376. https://doi.org/10.1002/ece3.5555

Garcia, K., Lasco, R., Ines, A., Lyon, B., Pulhin, F. 2013. Predicting geographic distribution and habitat suitability due to climate change of selected threatened forest tree species in the Philippines. Applied Geography, 44, 12-22.

Krisnawati, H., Varis, E., Kallio, M., Kanninen, M., 2011. Paraserianthes falcataria (L.) Nielsen; Ecology, silviculture and productivity. Center for International Forestry Research, Bogor, Indonesia.

Rojas-Briceño, N. B., Cotrina-Sánchez, D. A., Barboza-Castillo, E., Barrena-Gurbillón, M. Á., Sarmiento, F. O., Sotomayor, D. A., Salas-López, R., 2020. Current and Future Distribution of Five Timber Forest Species in Amazonas, Northeast Peru: Contributions towards a Restoration Strategy. Diversity, 12(8), 305.

Santillan, J., Gagula, A., Santillan, M., 2021. Machine Learning Approach for Tree Plantation Suitability Mapping. 2021 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 6865-6868.

Zhang, K., Zhang, Y., Tao, J., 2019. Predicting the Potential Distribution of Paeonia veitchii (Paeoniaceae) in China by Incorporating Climate Change into a Maxent Model. Forests, 10, 190; doi:10.3390/f10020190