SPECIES DISTRIBUTION MODELING IN FOREST PLANNING OF ANNUAL PRODUCTION UNITS IN THE SOUTHWEST AMAZONIA

Alexandra Bezerra de Menezes[2*] and Symone Maria de Melo Figueiredo[3]

[1] Received on 02.10.2020 accepted for publication on 04.05.2021.
[2] Universidade Federal do Acre, Graduada em Engenharia Florestal, Rio Branco, AC - Brasil. E-mail: <bezerrademenezesa@gmail.com>.
[3] Universidade Federal do Acre, Centro de Ciências Biológicas e da Natureza, Rio Branco, AC - Brasil. E-mail: <symone.figueiredo@ufac.br>.
*Corresponding author.

ABSTRACT – The generally limited resources for forest management and the growing need of forest production regulation requires the optimization of planning approaches for the spatialization of annual production units (APU). An APU planning methodology for forest species of high wood value (*Amburana acreana* (Ducke) ACSm., *Apuleia leiocarpa* (Vogel) JF Macbr. and *Castilla ulei* Warb.) in management area was proposed, using prediction of potential distribution of these species with data from the occurrence of a census forest inventory. It was used sample inventory data simulated in three sampling systems (random, conglomerate systematic, and systematic) and sample intensities (0.5% and 0.8%). As predictive variables, it was used the altitude, vertical distance to the nearest drain, individual bands of the TM sensor on board the Landsat 5, and vegetation index by normalized difference. Eighteen models were obtained, six per species. The test area under the curve (AUC) of the models ranged from 0.517 to 0.804. For all species, the best predictive model was considered the conglomerate system with a sample intensity of 0.8%. Altitude was the predictor variable that most contributed to the models. The AUC values for the *Amburana acreana* models were significantly different from *Apuleia leiocarpa* and *Castilla ulei* (p = 0.0138). For species of lower density, it is recommended greater sampling intensity and sampling systems that provide better spatialization of occurrence records. The use of data from sampling forest inventories in different sampling systems is capable of predicting environmental suitability for forest species and helps to define APUs. Thus, it is possible to strenghten the exploration strategies and management planning of management areas and to contribute to the perpetuation of the activity in the unequal forests of the Amazon region.

Keywords: Geotechnology; Maxent; Forest planning.

MODELAGEM DE DISTRIBUIÇÃO DE ESPÉCIES NO PLANEJAMENTO FLORESTAL DAS UNIDADES DE PRODUÇÃO ANUAL NO SUDOESTE DA AMAZÔNIA

*RESUMO – Considerando os recursos geralmente limitados para a gestão de florestas e a crescente necessidade de regulação da produção florestal, faz-se necessário otimizar abordagens de planejamento da espacialização das unidades de produção anual (UPA). Foi proposta uma metodologia de planejamento de UPA para espécies florestais de alto valor madeireiro (**Amburana acreana** (Ducke) A.C.Sm., **Apuleia leiocarpa** (Vogel) J.F Macbr. e **Castilla ulei** Warb.) em área de manejo, utilizando predição de distribuição potencial destas espécies com dados de ocorrência de inventário florestal censitário. Foram utilizados dados de inventário amostral simulados em três sistemas de amostragem (aleatório, conglomerado sistemático e sistemático) e intensidades amostrais (0,5% e 0,8%). Como variáveis preditoras, utilizamos a altitude, distância vertical à drenagem mais próxima, bandas individuais do sensor TM a bordo do Landsat 5 e índice de vegetação por diferença normalizada. Foram obtidos 18 modelos, seis por espécie. A área sob a curva (AUC) de teste dos modelos variou de 0,517 a 0,804. Para todas as espécies, o melhor modelo preditivo foi considerado o sistema conglomerado com intensidade amostral de 0,8%. A altitude foi a variável preditora que mais contribuiu nos modelos. Os valores de AUC para os modelos de Amburana acreana foram significativamente diferentes de **Apuleia leiocarpa** e **Castilla ulei** (p=0,0138). Para espécies de menor densidade, é recomendado utilizar maior intensidade amostral e sistemas de amostragem que proporcionem melhor espacialização dos registros de ocorrência. A utilização de dados*

provenientes de inventários florestais amostrais em diferentes sistemas de amostragem é capaz de predizer a adequabilidade ambiental para espécies florestais e auxilia a definição de UPAs. Assim, pode-se fortalecer as estratégias de planejamento da exploração e gestão das áreas de manejo e contribuir para a perpetuação da atividade nas florestas inequiâneas da região amazônica.

Palavras-Chave: Geotecnologia; Maxent; Planejamento florestal.

1. INTRODUCTION

New challenges have been triggered in planning the harvesting of native forests, in addition to environmental, social, and economic aspects, the regulation of forest production is one of the greatest challenges (Carvalho et al., 2015). The sustainable management of tropical forests must guarantee a constant flow of resources for their viability, however, this is not always possible, due to the lack of harvesting planning (Fernandes et al., 2013).

Bearing in mind that the planning of logging in unequal forests is determined by dividing Forest Management Areas (FMA) into Production Units (PUs), of similar sizes and shapes, and that this procedure does not necessarily correspond to a volume of regular harvest and income in each PU (Silva et al., 2018a), the spatial distribution of annual production units (APUs) is a problem that needs optimized solution.

In addition, due to environmental, economic factors, and production targets, there is an increasing need for detailed spatial planning. Managers have been gradually using computer systems in order to regulate forest production in an optimized way. Whether in equine forests (Binoti et al., 2014), which has presented a beneficial effect on management, or in areas of unequal forest management in the Amazon (Silva et al., 2018a), promoting the formation of PUs efficiently, grouping the trees in an optimized way.

The integration of computational systems also provided robust results in a study to verify the optimal selective cutting regime in models for allocation wooden storage yards, using models of entire linear programming, the tested scenarios were efficient and concise, ratifying the potential to increase the efficiency of logging and forest management plans in the Amazon (Silva et al., 2018b).

In general, the use of these geotechnologies and methods of optimization of exploitation constitute potential tools to solve the spatialization of forest resources. Its use proved to be more efficient than the intuitive or empirical methods normally used in the management of unequal forests (Martinhago, 2012).

Other potential tools for forest planning are the species distribution models, both in the planning of natural forests (Mateo et al., 2018), and to assist in the development of new approaches in the planning of PUs in forest management.

The species distribution models are based on the distribution of the species' occurrence points (geographic location) in the sub-space of conditions of its ecological niche and produce functions to predict where in the geographic space it is likely to occur (De Marco and Siqueira, 2009). In addition, the use of environmental predictors with higher spatial resolution (of the order of a few meters) can be applied to carry out the geographic distribution modeling of species of interest (Figueiredo et al., 2015).

In addition to the potential of species distribution models to solve problems related to conservation and to map the potential distribution of species of timber interest in the Amazon (Guisan et al., 2013; Pérez Chaves et al., 2018), they are a potential tool in the spatialization of forest resources, to formulate development strategies and plans, important for forest planning and regulation (Silva et al., 2018a).

Considering the challenges of planning the forest exploitation aiming at the regulation of wood production, the objective of this study was to propose a new approach to planning annual production units in the management area, through the modeling of the distribution of forest species, using data from sampling inventory and predictive environmental variables with spatial resolution of 30 meters.

2. MATERIAL AND METHODS

2.1 Study area

The study was carried out in the Antimary State Forest (ASF), located in the municipality of Bujari,

whose upper left and lower right vertices have the respective coordinates (SIRGAS 2000 datum): S 09° 29' 78.36"and W 68 ° 26 '70.02"and S 09° 34' 49.89"and W 68° 16' 86.34". It is a Conservation Unit for Sustainable Use, held by the State Secretariat for the Environment - SSE. The Sustainable Forest Management Plan area corresponds to 37,687.66 ha, accounting for 82.2% of the total area of the ASF, which is 45,686.5 ha (Figure 1A).

In ASF there is a predominance of Dense Forest vegetation associated with Open Forest with Bamboo and Dense Forest associated with Open Forest with Palm trees. It has rainfall above 2,000 mm, with a dry period with monthly precipitation below 60 mm of rain. The average annual temperature ranges from 24.6 °C to 25.0 °C, the soil comprises the Argisol Red Dystrophic Latossolic class associated with Oxisol Yellow Dystrophic Argisolic (SEMA, 2019).

2.2 Forest species data

We used the forest census of five APUs (APU-02, APU-03, APU-04, APU-07 and APU-08), the censuses took place among the years 2005 to 2015, data in digital format were provided by the Environment Institute of Acre (IMAC) and authorized by SSE.

In the inventory, all individuals from 40 cm DBH (diameter at breast height) were inventoried, a diametric class immediately below the harvesting limit (50 cm). About the mapped trees, the following information is available: geographic location, species, DBH, commercial height and basal area, among others. The individual volume of the trees was estimated by the volume equation based on the Schumacher-Hall model, which considers double entry (diameter and

Source: author
Fonte: autor.

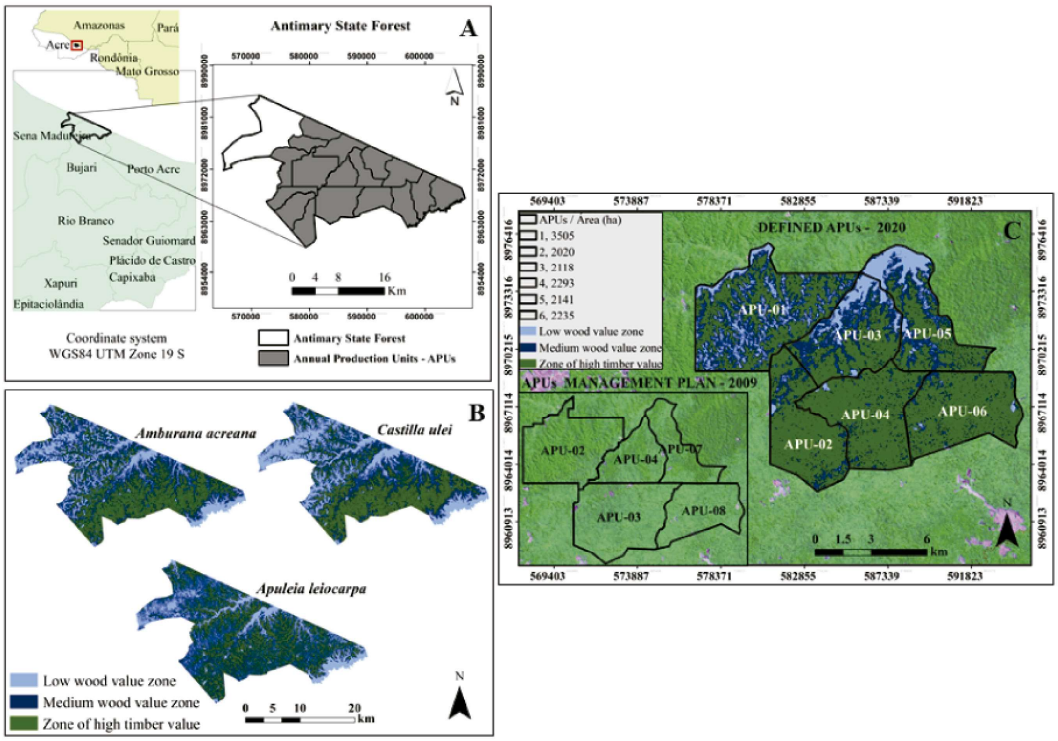

Figure 1 – Location map of the Antimary State Forest, with the current division of annual production units. Bujari, Acre, 2020 (A), map with the wood value zones for the species: *Amburana acreana*, *Apuleia leiocarpa* and *Castilla ulei* (*B*) and map with the division of annual production units for the area with the highest probability of occurrence in the Antimary State Forest (C).

***Figura 1** – Mapa de localização da Floresta Estadual do Antimary, com a atual divisão de unidades de produção anual. Bujari, Acre, 2020 (A), mapa com as zonas de valor madeireiro para as espécies: **Amburana acreana**, **Apuleia leiocarpa** e **Castilla ulei** (B) e mapa com a divisão das unidades de produção anual para a área de maior probabilidade de ocorrência na Floresta Estadual Antimary (C).*

height). The model uses the commercial height of the trees, estimated by botanical identifier (SEMA, 2019).

The area of the APUs is equivalent to 13,703.7ha, however, for the calculation of the number of plots of the simulated sampling inventory in a Geographic Information System (GIS) it was considered environment the entire value (13,000 ha). We rounded off the value of the area, as the ASF is cut by roads, access routes and other open areas, including facilities such as the headquarters, Integrated Environmental Management Unit (IEMU), located in an APU, we discount these areas, as they could not be used to distribute plots.

Three forest species of wood interest with high market demand in the state of Acre were selected. They are: *Amburana acreana* (Ducke) A.C.Sm., *Apuleia leiocarpa* (Vogel) J.F.Macbr. and *Castilla ulei* Warb. (Table 1).

We used the geographic position of the tree to define the presence. For each species, only one occurrence data per pixel was selected. The geographical coordinates of the trees, by the UTM cylindrical cartographic projection, were obtained through GPS, which replaced the false coordinates (X, Y) and the right and left position of the line used in traditional inventories. We consider the abundance sufficient for robust modeling, exceeding the minimum number of 30 observations in the study area (Giannini et al., 2012), and the reliable taxonomic identification at the species level.

The records of occurrence of forest species, derived from the census forest inventories provided by the Environment Institute of Acre (EIA), in digital format, were audited by the institution's technical staff at different stages of the harvesting activity, in addition to being an experimental area of Embrapa Acre, with a high frequency of "field truths" (field verification) made by experienced parabotanists, which guarantees the reliability of the worked information. The exsiccates of the species of interest are deposited in the herbarium of the Zoobotanic Park of the Federal University of Acre (UFAC-ZP) (Figueiredo et al., 2015).

2.3 Predictor variables

We used the following predictor variables: altitude, vertical distance to the nearest drain (HAND), spectral bands of the TM sensor (Landsat 5 satellite) and the vegetation index of the normalized difference of the acronym in English (NDVI), all with spatial resolution 30 meters. The individual bands of the Landsat 5 sensor referent to the year 2008 were used in order to provide images with better quality.

The altitude data were generated from data from the SRTM (Shuttle Radar Topographic Mission). The HAND (Height Above the Nearest Drainage) refers to the vertical distance of each point in relation to the nearest drainage, the values of the vertical distances were extracted from the HAND and processed according to the methodology of Rennó et al. (2008).

We used the mosaic of spectral bands derived from the Landsat TM satellite (Thematic Mapper), bands corresponding to the orbit/point 002/66 and 002/67, all of August 18, 2008. The bands were obtained from the image bank of the National Institute for Space Research (NISR), without clouds, covering the study area one year before the forest census was carried out, due to the better quality of the images. They were georeferenced based on the Landsat8 image, Sensor-Operational Land Imager (OLI) for the year 2013 available on Libra, the Landsat image browser.

Of the seven bands that the TM sensor has, four were used as predictor variables in the species distribution models. The red band 3 (0.63 - 069 μm) of the visible light spectrum and bands 4 (0.76 - 0.90 μm), 5 (1.55 - 1.75 μm) and 7 (2.08 - 2.35 μm) of the near and medium infrared spectrum. These bands

Table 1 – List of forest species selected for modeling and total number of occurrence records in the census forest inventory in the Antimary State Forest, considering the selected APUs.
Tabela 1 *– Relação das espécies florestais selecionadas para a modelagem e número total dos registros de ocorrência no inventário florestal censitário na Floresta Estadual do Antimary, considerando as UPAs selecionadas.*

Scientific name	Family	Common name	Total number of occurrence data
Amburana acreana (Ducke) A.C.Sm.	Fabaceae	Cerejeira, cumaru-de-cheiro.	1883
Apuleia leiocarpa (Vogel) J.F.Macbr.	Fabaceae	Cumaru-cetim, garapeira	4869
Castilla ulei Warb.	Moraceae	Caucho	7433

were selected, as there is high reflectance in the near infrared region. Meanwhile, the medium infrared medium presents sensitivity to the moisture content of the plants, serving to observe stress in the vegetation.

The index of vegetation by normalized difference is an index of significant importance in the characterization of vegetation (Neta et al., 2018). NDVI values indicate the amount of green vegetation present in the pixel (Prates Clark et al., 2008). For the calculation of the NDVI, we used the mosaic of the spectral bands (band 3 and band 4) corresponding to the orbit/point 002/66 and 002/67.

2.4 Models construction

The occurrence data was partitioned into two data sets: a. data for model calibration and b. data for model validation. The selection of data for the calibration of the models was obtained according to the sampling intensity and the simulated sampling system, thus, the occurrences for the calibration were those contained in the sampling units and the validation occurrences there were all other occurrences remaining, according to the sampling system and intensity.

For each selected species, six predictive models of species distribution were generated in function of the sampling systems and sampling intensity. Three sampling systems were simulated: random, systematic conglomerate and systematic in two sampling intensities: 0.5 and 0.8%, the intensities were used considering that sampling inventories in the Amazon are generally carried out with low sampling intensity, less than 1%, in function of depending on costs (Cavalcanti et al., 2011).

In the random and systematic sampling system, the sample units had the dimensions (50 x 200 m) in width and length, respectively. Corresponding to sampling units of 1 ha (10,000 m^2). With the difference that in the systematic sampling after the selection of the first sample unit, the plots followed a predetermined pattern at constant intervals of (1 x 1 km), considering the area available for the distribution of plots.

The sampling units per systematic conglomerate were made up of four rectangular subunits of fixed area, perpendicular to their central point, oriented in the direction of the cardinal points and numbered from 1 to 4. The subunits with a size of 20 m in width by 200 m of length (4,000 m^2) and 50 m among them, in total one sample unit corresponded to 1.6 ha (16,000 m2).

In all, there were 65 plots with an intensity of 0.5% and 104 plots with an intensity of 0.8% in the random and systematic system. In the systematic conglomerate system, there were 41 plots at 0.5% intensity and 65 plots at 0.8% intensity. The distribution of plots in the area, according to the simulated sampling system, and the selection of occurrence data for the three species were carried out in a GIS environment.

In modeling, we used the maximum entropy algorithm (Maxent), a method to perform prediction or inferences based only on presence data (Phillips et al., 2006), executed with the Maximum Entropy Species Distribution Modeling v. 3.3.3k (<www.cs.princeton.edu/_schapire/maxent/>).

Maxent estimates the probability of suitability of the habitat where the occurrence of the species is possible, finding the maximum entropy probability distribution (closest to the uniform) subject to a set of restrictions, in which the expected values in each environmental variable must correspond to the averages observed in the samples (Phillips et al., 2006; Elith et al., 2011).

At Maxent, according to Elith et al. (2011), the landscape of interest is denoted by L, y = 1 (presence), y = 0 (absence), z = vector of covariates representing environmental conditions, f (z) = probability density of covariates in L (a random sample of L) and f1 (z) = probability density of covariates in locations in L where the species is present.

The probability you want to find is the probability that the species is present given that the z environment is observed. Using presence data and random landscape data, it is possible to model f1 (z) and f (z) and the f1 (z) / f (z) ratio estimate, known as Maxent's "raw" output, giving an insight into which characteristics are important and estimates the relative suitability of one place in relation to another.

According to Elith et al. (2011), in order to estimate f1 (z), several of these distributions are possible, but is selected one that is closer to f (z), minimizing the distance from f (z) which is a null model for f1 (z), because without data on the occurrence, we would

have no reason to expect the species prefers all specific environmental conditions over any other. Restrictions are imposed so that the solution is one that reflects the information from the presence records, that is, the estimate of f1 (z) must be close to the average of the observed presence locations.

As the necessary information on prevalence is not available for calculating conditional probability of occurrence, an alternative solution, called Maxent's logistic output, was implemented, where prevalence is represented by the parameter τ, defined from "typical conditions" of a group of species, whose default value of $\tau = 0.5$. This logistical transformation is monotonically related to Maxent's "raw" output. Maxent's logistics output transforms the model from an exponential model belonging to the Gibbs distribution to a logistic model (Elith et al., 2011):

$$fk = \left(2.\frac{f1 + f2 + f3 + \cdots + f\left(\frac{n}{2}\right) - 1}{\left(\frac{n}{2}\right) - 1} - f\frac{n}{2} \right).1.10 \quad \text{Eq.1}$$

r: the relative entropy of $f1(z)$ that is obtained from the presence data in relation to $f(z)$ that is estimated from random points in the landscape.

n (z): it is the result of maximizing entropy subject to certain restrictions.

For each species, sampling system and sampling intensity, modeling was performed with one repetition (number of model runs. The data was resampled using the bootstrap technique, this technique involves randomly splitting the data, with replacement, into various training and test sets (Fielding and Bell, 1997).

We used the Jackknife test to estimate which are the most important variables in the models, consequently the contribution of the variables to the models (Yang et al., 2013). We analyzed the Pearson coefficient of variation in order to exclude from the variable models with a correlation greater than 0.7.

The samples of random points in the landscape (background) were extracted from a region that covers the entire study area containing the sampling distribution of the occurrence data of the forest inventory. This region was defined considering all occurrences, regardless of the species.

2.5 Evaluation of models

The performance of the models was evaluated through the area under the curve (AUC) of the ROC (Receiver Operating Characteristic) curve, graph of the receiver - operator, in which the AUC is the measure of performance, and where is possible to compare the estimated area of the models with those observed at the same point by the validation samples. AUC is a widely used method because it is a global measure of performance independent of cut limits (Fielding and Bell, 1997).

In the selection of the models, the omission rate was also verified, as an auxiliary measure to the AUC, considered important for the selection (Figueiredo et al., 2015). We used the Student's t test to assess the significance ($\alpha = 0.05$) of the variation in the mean AUC values between the different species distribution models generated as a function of the system and sampling intensity.

In this study, we used the Maximum test sensitivity plus specificity as the cut-off limit, provided by Maxent himself, above which is predicted the variable of interest, and it is important that the choice of the threshold ponders the intended use of the map (Freeman and Moisen, 2008).

2.6 Planning of annual production units

For the planning of the APU, the best model by species was selected according to the analysis of the AUC, the omission rate and we also considered the sampling effort. The probability of occurrence of species on the maps was categorized into four classes (1. High; 2. Medium; 3. Low and 4. Absence). The categorization was done using the Natural Breaks method, where the division of the probability values by class was done by the program itself. The areas with the highest probability values were categorized as high probability of species occurrence, and the areas with the lowest probability values were categorized as absence predicted by the models.

The maps of probability classes were vectorized by species. After that, we proceeded with the sum of the vectorized maps, thus creating zones to subsidize and define as APUs by group of species. Three zones were defined: Zone 1: High wood value, Zone 2: Medium wood value and Zone 3: Low wood value in the area of the five APUs where the occurrences were used for modeling.

Greater weight was given to the probability maps of *Amburana acreana* and *Apuleia leiocarpa*, in the

sum of the maps of these species was multiplied by two (2), the importance given to the distribution of these species is due to the demand by the timber consumer market (De Araujo, 2014).

As a result of the sum of the probability maps, a new probability map was obtained for the three species. The new map was divided into 15 classes of probability of occurrence, grouped every five according to the variation of probability, thus defining three zones of wood value for the species. In which, the probabilities of occurrence of the maps were distributed as follows: Zone 1 (2.31 - 3.5) with higher probability values, Zone 2 (1.15 - 2.30) with average values of probability and Zone 3 (0 - 1.14) with the lowest probability values.

We verified through descriptive statistics, coefficient of variation (CV), if the models provided more equiproductive APUs at the species level, for this, we compared the distribution of the volume of wood by species and APU considering the division of the units before and after the models of species distribution.

3. RESULTS

Eighteen models were obtained using the maximum entropy algorithm, six per species, one model for each combination of species, system and sample intensity. The test AUC value of the models ranged from 0.517 to 0.804 among species.

There was greater variation in the AUC values for the species *Amburana acreana*, than for *Apuleia leiocarpa* and *Castilla ulei*. The models generated following the three sampling systems (random, systematic conglomerate and systematic), as well as the two sample intensities (0.5% and 0.8%) performed well with an average AUC of 0.7925 (*Castilla ulei*), 0.658 (*Amburana acreana*) and 0.749 (*Apuleia leiocarpa*), above a randomly generated model (AUC = 0.5).

For *Amburana acreana*, there was a greater variation in the AUC values between the models. The data for the species show that for the same sampling system there can be a large difference of omission, as in the systematic sampling system. The AUC values for *Apuleia leiocarpa* had little variation (0.737 to 0.765), and omission values (14.3 to 22.4%), in most models the omission was greater using a smaller number of occurrence records.

As for *Castilla ulei*, all models showed good AUC values ranging from (0.783 to 0.804) and default values (16.1 to 20.5%). This species had the highest number of occurrences to generate the models (40 to 80 occurrences). The increase in the number of occurrence points increased the AUC and decreased the omission, for most models.

There was a significant difference when comparing the AUC values among species. The models of Amburana acreana were significantly different from *Apuleia leiocarpa* and *Castilla ulei* ($p = 0.0138$). There was no significant difference between the models of the latter.

Considering all the models evaluated, the omission was ≤ 30%. For all species, there was no significant difference in relation to the omission of the models, although the values of AUC and omission for the models of *Apuleia leiocarpa* and *Castilla ulei* were contained in less variation in relation to the values for *Amburana acreana*.

All selected variables were used in the models, except Landsat Band 7, which, after analyzing Pearson's coefficient of variation, showed a correlation greater than 0.7 with three variables (Band 3, Band 5 and NDVI). Although there is a correlation between band 3 and band 5, the models performed better with the maintenance of these bands (Table 2).

Considering the contributions of the predictor variables to the selected models, altitude was the most

Table 2 – Predictor variables for the best models and their percentage contribution in the Maxent model, by forest species.

Tabela 2 *– Variáveis preditoras para os melhores modelos e sua contribuição percentual no modelo Maxent, por espécie florestal.*

Variable	Species		
	Amburana acreana	*Apuleia Caastilla*	*Castilla ulei*
Altitude	62.4	54.0	62.3
Band 5	7.0	35.6	13.2
Band 3	18.2	0.0	12.7
Band 4	1.5	2.5	4.4
HAND	7.4	6.8	0.4
NDVI	3.5	1.1	7.0

Where: HAND, vertical distance to the nearest drainage; NDVI, vegetation index by normalized difference.

Em que: HAND, distância vertical à drenagem mais próxima; NDVI, índice de vegetação por diferença normalizada.

Table 3 – Values of AUC and omission (%) of the models distributed by species, sampling systems and sampling intensity (%).
Tabela 3 *– Valores de AUC e omissão (%) dos modelos distribuídos por espécie, sistemas de amostragem e intensidade amostral (%).*

		Amburana acreana		*Apuleia leiocarpa*		*Castilla ulei*	
System	**Intensity**	**AUC**	**Omission**	**AUC**	**Omission**	**AUC**	**Omission**
Random	0.5	0.724	23.2	0.746	18.9	0.792	20.3
	0.8	0.725	23.7	0.747	14.3	0.78	18.8
Conglomerate	0.5	0.517	11.4	0.754	20.3	0.793	20.5
	0.8	0.814	18.1	0.747	12.7	0.803	16.1
Systematic	0.5	0.531	3.9	0.737	19.0	0.783	19.3
	0.8	0.642	30.0	0.765	22.4	0.804	16.9

important predictor variable for the three species, *Amburana acreana*, *Apuleia leiocarpa* and *Castilla ulei*, a result commonly found in research that use this variable as a predictor for the models (Higgins et al., 2012; Figueiredo et al., 2015; Pérez Chaves et al., 2018).

Band 5 (short wave infrared) was the second most important variable for the *Castilla ulei* and *Apuleia leiocarpa* model, for *Amburana acreana* it was Band 3. due to the fact that NDVI incorporates information from bands 3 and 4 it is often used as a general vegetation vigor index. However, NDVI and Band 4 were not the most important variables for any of the models. For the three species, the contributions of bands 3 and 4, if added to the models, contributed more than the NDVI, a result similar to that found in the study by Pérez Chaves et al. (2018).

From the models generated with the variables, altitude, HAND, spectral bands of the TM sensor (Bands 3, 4, 5) and the NDVI, we considered the values of AUC, omission rate and sampling effort in order to select the models of each species. We adopted this method, as there was no statistically significant difference using the Student's t test to compare the AUC scores.

For all species, the best model was the one that considered the sampling by systematic conglomerate and sampling intensity of 0.8%, the *Amburana acreana* model had AUC (0.814) and omission rate (18.1%). The *Apuleia leiocarpa* model had AUC (0.747) and an omission rate (12.7%). And *Castilla ulei* at AUC (0.803) and default (16.1%) (Table 3).

The omission of the models selected for the three species ranged from 12.7% to 18.1%, values above that found in Figueiredo and Figueiredo (2019), with an omission rate of 8.8%, evaluating the modeling of tree species by diametric class. In our study there were areas of effective presence of the species, but that the model was unable to predict. These unforeseen occurrence points were located close to water courses and most of the occurrence records used to generate the models were at higher points in the terrain.

From the maps of probability of occurrence of the selected models, we obtain the wood value zones by species for the ASF (Figure 1B).

In the map of areas of wood value by species, the areas of lesser environmental suitability formed the area of low wood value, which corresponds to areas that are already open, mainly west of the ASF, water courses and in its surroundings and areas of permanent preservation (APP). In general, areas of low and high wood value occupied most of the map.

As a result of the sum of the probability maps of the selected models, the wood value zones for *Amburana acreana*, *Apuleia leiocarpa* and *Castilla ulei* were obtained in the area where the occurrence points that generated the models were distributed.

In general, the largest area on the map is made up of the area with high timber value, that is, the area in which there is a greater probability of finding the three wood species and that effectively have forests. Based on these zones of wood value for the species, the APUs were defined for the area with the highest probability of occurrence (Figure 1C).

The limitations of annual production units coincided with the natural dividers, which are the watercourses, water dividers (slope line) as recommended by Souza and Soares (2013), the roads were also considered as limiters of the APUs. In the areas with the highest timber value, the APUs had a smaller area, considering that in these areas a higher density of individuals is expected and thus a greater volume of wood. Unlike this, for areas of lower wood

Table 4 – Volume of wood of the species *Amburana acreana*, *Apuleia leiocarpa* and *Castilla ulei* in the APUs defined from the species distribution models, Antimary State Forest, Bujari, Acre, 2020.
Tabela 4 *– Volume de madeira das espécies* ***Amburana acreana****,* ***Apuleia leiocarpa*** *e* ***Castilla ulei*** *nas UPAs definidas a partir dos modelos de distribuição de espécies, Floresta Estadual do Antimary, Bujari, Acre, 2020.*

	Amburana acreana	*Apuleia leiocarpa*	*Castilla ulei*
		Volume (m³)	
APU 01	907.3	4,602.3	2,129.1
APU 02	1,875.1	4,811.9	3,682.8
APU 03	1,401.1	5,176.7	3,160.4
APU 04	1,628.3	5,995.8	4,584.4
APU 05	824.0	3,318.1	1,986.8
APU 06	852.6	3,317.7	2,763.5
Total	7,488.4	27,222.4	1,8307
Average	1,248.1	4,537.1	3,051.2
Standard deviation	411.0	965.2	897.0
Coefficient of variation (%)	32.9	21.3	29.4

value, APUs had a larger area, so that the final volume is as proportional as possible between them (Table 4).

We observed a significant variation in the volume of species in the original APUs in the management area in relation to the APUs that were divided with the aid of the models. The difference between the volume of species per APU can be related both to the size of the production units as well as the density and DBH of the species.

The variation coefficient of the original APUs was greater than 50%, only for the *Apuleia leiocarpa* species, which was smaller (CV = 48.3%), which denotes a large variation in volume between the production units, while the CV of the defined APUs was between 21.3 to 32.9%.

4. DISCUSSION

The AUC values for the models were on average (AUC = 0.7), this result corresponds to the study by Figueiredo and Figueiredo (2019), which analyzed the prediction of the distribution of timber forest species, on a local scale (approximately 35,000 ha), using occurrence data grouped by diametric class. Here, this relationship was not true only for *Amburana acreana* (AUC = 0.6), a species that had a lower number of occurrences to generate the models.

As there was no significant difference when different sampling systems and sampling intensity were used for *Apuleia leiocarpa* and *Castilla ulei*, any of the systems and intensities can be used to generate predictive models for these species, as they are species of greater density in the area.

For *Amburana acreana*, it is necessary to have a greater number of occurrence records for the construction of the models, that is, greater sampling intensity. Because the increase in the sampling intensity and a greater number of occurrences better distributed in the geographic space can provide improvement in the prediction of the models (Figueiredo et al., 2015).

The smaller number of occurrence records to generate the *Amburana acreana* models reflected in the decrease in AUC in the systematic and systematic conglomerate system, however just increasing the number of occurrences may not be efficient as in the work of Aguirre Gutiérrez et al. (2013), in which the increase in the number of records significantly decreased the AUC values, although this result is expected when using pseudo-absences or historical data, this is not our case.

We follow the same criteria as Doninck et al. (2019) to maintain or remove predictive variables, they removed the variable when the correlation was greater than 0.78.

The HAND was the third most important variable for the models of *Amburana acreana*, probably because this variable is indirectly related to the depth of the water table, indicating the availability of water in the soil (Rennó et al., 2008).

Despite the correlation considered high (0.77) between the NDVI and band 3, but below the criterion used, the maintenance of the NDVI is important because with the inclusion of this variable in the models, the deforested areas are not predicted as areas of environmental suitability for the species not being necessary to mask deforested areas (Prates Clark et al., 2008; Figueiredo et al., 2015).

Band 3 and 4 contributed more than NDVI, as well as in the work of Pérez Chaves et al. (2018),

where the authors report that the importance of these bands is related to the fact that photosynthesis of vegetation absorbs a large proportion of incident red radiation and reflects a high proportion of near infrared radiation.

We verified that variables from remote sensing are good predictors for most models of species distribution, as in the study by Pérez Chaves et al. (2018). Landsat and SRTM data, in combination with the rate-based rapid field inventory, should serve as a basis for planning research and conservation of vast and remote forests (Higgins et al., 2012). The addition of Landsat-based layers of environmental data improves the discriminatory capacity of models compared to models that use only layers of climatic data (Doninck et al., 2019).

Considering that the occurrences for the modeling were obtained according to three sampling systems and two sampling intensities. The omission of the models probably occurred because the distribution of the occurrence points that generated the models was not able to capture the gradient of altitude variation, the most important variable for the three species, reflecting the importance of the better distribution of data in the area and not just an increase in the number of points.

It was possible to verify, for all species in the omission area, lower altitude values than for the occurrence points used in the calibration of the models, consequently there were problems in the transferability of the models, as they were unable to predict the occurrence of the species for the areas with lower altimetric height. One way to reduce the omission for the models in this case is to previously stratify the environment by topographic gradient and distribute plots in order to cover the altimetric amplitude of the area.

Despite the problems related to the omission and transferability of the models to other areas, knowing the potential distribution of timber forest species can assist both in the formation of forest management areas as well as in the definition of more equiproductive annual production units at the species level. In this way, the definition of production units based on models of potential distribution of species and zones of wood value becomes yet another alternative to optimize the definition of APUs in management areas, and thus regulate volume production (Silva et al., 2018a).

Despite a significant decrease in the difference in wood volume between the APUs defined with the aid of the models, it was not possible to provide equiproductive units by species. This shows that only the modeling of the potential distribution of species and the formation of production units based on the models is not capable of promoting, by itself, the equiproductive distribution of volume by species.

Even so, knowing the potential distribution of timber forest species can assist both in the formation of forest management areas as well as in the definition of more equiproductive annual production units at the species level.

5. CONCLUSION

The best results in the prediction of the potential distribution of *Amburana acreana*, *Apuleia leiocarpa* and *Castilla ulei* were obtained with the occurrence data from the forest inventory using the systematic conglomerate system with a sample intensity of 0.8%.

The definition of the annual production units with the help of species distribution modeling contributed to reduce the difference in the volume of wood between the APUs, when compared with conventional methods of forest planning, but they were not equiproductive.

6. AUTHOR CONTRIBUTIONS

Menezes AB et al. contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript. All authors discussed the results and contributed to the final manuscript.

7. REFERENCES

Aguirre-Gutiérrez J, Carvalheiro LG, Polce C, Van Loon EE, Raes N, Reemer M, Biesmeijer JC. Fit-for-purpose: species distribution model performance depends on evaluation criteria–Dutch hoverflies as a case study. PLoSOne. 2013;8(5): e63708. doi: 10.1371 / journal.pone.0063708

Binoti, DHB, da Silva Binoti, MLM, Leite, HG, Gleriani, JM, de Oliveira, MLR, de Souza, AL, et al.

Regulation of even-aged forest with constraint of age classes per compartment. Cerne. 2014; 20(4):567-573. doi: 10.1590/01047760201420041183

Carvalho KHA, Silva ML, Leite HG, Binoti DHB. Influência da taxa de juros e do preço da madeira em modelos de regulação florestal. Pesquisa Florestal Brasileira. 2015;35(82):143-151. doi: 10.4336/2015.pfb.35.82.554

Cavalcanti FDB, Machado SDA, Osokawa RT, Da Cunha US. Comparison of values estimated by sampling to characterize the structure of a forest area in the Amazon based on the information registered in the Forestry Census. Revista Árvore. 2011;35(5):1061-1068. doi: 10.1590/S0100-67622011000600012

De Araujo HJB. Espécies florestais madeireiras tradicionais consideradas em processo de escassez no estado do Acre. In: Embrapa Acre-Artigo em anais de congresso (ALICE). In: CONGRESSO BRASILEIRO DE REFLORESTAMENTO AMBIENTAL. Anais. CEDAGRO; CREA-ES; Florestas para a Vida; Reflorestar; GEF; Incaper; iema, Vitória: 2014.

De Marco PJr, Siqueira MF. Como determinar a distribuição potencial de espécies sob uma abordagem conservacionista. Megadiversidade. 2009; 5(1-2):65-76.

Doninck JV, Jones MM, Zuquim G, Ruokolainen K, Moulatlet GM, Sirén A, Tuomisto H. Multispectral canopy reflectance improves spatial distribution models of Amazonian understory species. Ecography. 2019; 43(1):128-137. doi: 10.1111/ecog.04729

Elith J, Phillips SJ, Hastie T, Dudik M, Chee YE, Yates CJ. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions. 2011;17(1): 43-57. doi: 10.1111/j.1472-4642.2010.00725.x

Fernandes APD, Guimarães PP, Braz EM, Hoeflich VA, Arce JE. Alternativas de planejamento para a exploração florestal. Embrapa Florestas-Artigo em periódico indexado (ALICE), Floresta. 2013;43(3): 339-350. doi: 10.5380/rf.v43i3.30955

Fielding AH, Bell JF. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation. 1997;24(1):38-49. doi: 10.1017/S0376892997000088

Figueiredo SMM, Figueiredo EO. Modeling the distribution of trees species by diameter class in southwestern Amazonia. Scientia Forestalis. 2019;(124):644-654. doi: 10.18671/scifor.v47n124.06

Figueiredo SMM, Venticinque EM, Figueiredo EO, Ferreira EJL. Predicting the distribution of forest tree species using 50 topographic variables and vegetation index in eastern Acre, Brazil. Acta Amazonica. 2015;45(2):67-174. doi: 10.1590/1809-4392201402834

Freeman EA, Moisen GG. A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. Ecological Modelling. 2008;217(1-2):48-58. doi: 10.1016/j.ecolmodel.2008.05.015

Giannini TC, Siqueira MF, Acosta AL, Barreto FCC, Saraiva AM, Santos IAD. Current challenges of species distribution predictive modelling. Rodriguésia. 2012;63(3):733-749. doi: 10.1590/S2175-78602012000300017

Guisan A, Tingley R, Baumgartner JB, Naujokaitis-Lewis I, Sutcliffe PR, Tulloch AI, Martin TG. Predicting species distributions for conservation decisions. Ecology Letters. 2013;16(12):1424-1435. doi: 10.1111/ele.12189

Higgins MA, Asner GP, Perez E, Elespuru N, Tuomisto H, Ruokolainen K, Alonso A. Use of Landsat and SRTM data to detect broad-scale biodiversity patterns in Northwestern Amazonia. Remote Sensing. 2012;4(8):2401-2418. doi: 10.3390/rs4082401

Martinhago AZ. Otimização para alocação de pátios de estocagem para exploração de impacto reduzido na Amazônia Brasileira. Lavras. Tese [Doutorado em Manejo e Colheita Florestal]. Universidade Federal de Lavras; 2012.

Mateo RG, Gastón A, Aroca-Fernández MJ, Saura S, García-Viñas JI. Optimization of forest sampling strategies for woody plant species distribution modelling at the landscape scale. Forest Ecology and Management. 2018;410:104-113. doi: 10.1016/j.foreco.2017.12.046

Neta SRA, Bias EDS, Brites RS, Dos Santos CAM. Aplicação de um Modelo de NDVI para Detecção Multitemporal de Mudanças no Uso e Cobertura do Solo. Anuário do Instituto de Geociências. 2018;41(3):592-604. doi: 10.11137/2018_3_592_604

Pérez Chaves P, Ruokolainen K, TuomistoH. Using remote sensing to model tree species distribution in Peruvian lowland Amazonia. Biotropica. 2018;50(5):758-767. doi: 10.1111/btp.12597

Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecological Modelling. 2006;190(3-4):231-259. doi: 10.1016/j.ecolmodel.2005.03.026

Prates-Clark CDC, Saatchi SS, Agosti D. Predicting geographical distribution models of high-value timber trees in the Amazon Basin using remotely sensed data. Ecological Modelling. 2008;211(3-4):309-323. doi: 10.1016/j.ecolmodel.2007.09.024.

Rennó CD, Nobre AD, Cuartas LA, Soares JV, Hodnett MG, Tomasella J, Waterloo MJ. HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. Remote Sensing of Environment. 2008;112(9):3469-3481. doi: 10.1016/j.rse.2008.03.018

Secretaria do Meio Ambiente — SEMA. Plano de Manejo Florestal Sustentável (PMFS) Madeireiro. 2019. [cited 2021 January 19]. Edition 3. Available from: http://sema.acre.gov.br/wp-content/uploads/sites/20/2020/04/Plano-de-Manejo-Florestal-Antimary.pdf

Silva EF, Silva RF, Vieira GC, Leite CCC, Aguiar MO, Figueiredo EÓ, Silva MLM, Silva GF. Planning of production units for native forest management areas in the Amazon. Revista Brasileira de Ciências Agrárias. 2018;13(1):1-8. doi: 10.5039/agraria.v13i1a5502

Silva PH, Gomide LR, Figueiredo EO, Carvalho LMT, Ferraz Filho AC. Optimal selective logging regime and log landing location models: a case study in the Amazon forest. Acta Amazonica. 2018;48(1):18-27. doi: 10.1590/1809-4392201603113

Souza AL, Soares CPB. Florestas Nativas: estrutura, dinâmica e manejo. Viçosa, Brasil: Editora UFV; 2013. 322p. ISBN: 978-85-7269-463-6.

Yang XQ, Kushwaha SPS, Saran S, Xu JE, Roy PS. Maxent modeling for predicting the potential distribution of medicinal plant, Justicia adhatoda L. in Lesser Himalayan foothills. Ecological engineering. 2013;51:83-87. doi: 10.1016/j.ecoleng.2012.12.004