Caderno de Geografia (2023) v.33, n.72
ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

ORIGINAL ARTICLE

PROJECTION OF CLIMATE CHANGE IMPACTS ON NET PRIMARY PRODUCTIVITY OF THE LEGAL AMAZON – BRAZIL

Projeção dos impactos das mudanças climáticas na Produtividade Primária Líquida da Amazônia Legal – Brasil

Lucas Augusto Pereira da Silva
PhD Student in Geography, Universidade Federal de Uberlândia – UFU, Brazil
lucasagusto@ufu.br

Cristiano Marcelo Pereira de Souza
PhD in Geography, Universidade Estadual de Montes Claros – UNIMONTES, Brazil
cmpsgeografia@gmail.com

Claudionor Ribeiro da Silva
PhD in Geodetic Sciences, Universidade Federal de Uberlândia – UFU, Brazil
crs@ufu.br

Édson Luís Bolfe
PhD in Geography, Embrapa Agricultura Digital and Universidade Estadual de Campinas – UNICAMP, Brazil
ebolfe@unicamp.br

Andre Medeiros Rocha
PhD Student in Geography, Universidade de São Paulo – USP, Brazil
andremedeiros197@usp.br

Recebido: 07.03.2022
Aceito: 23.01.2023

Abstract

The Amazon Rainforest is one of the main carbon sinks (CO2) on Earth. However, recently, owing to anthropogenic activities and climate change, it has lost its stability in CO2 absorption. Therefore, understanding the dynamics of future climate change scenarios is essential. We assessed the influence of future climate change scenarios on NPP (biomass) levels in the Amazon Forest using ML models. The tested models were Bayesian, linear, and random forest models. The current scenario was evaluated using 19 bioclimatic covariates (WorldClim dataset). Future scenarios were based on RCPs 2.6 and 8.5 (based on the MIROC5 and HadGEM2-ES models). Random Forest had the best performance statistics (R^2 = 0.71 in training and 0.68 in the holdout-test). These climate change scenarios imply an increase in the average NPP for the Amazon forest, especially with the greater intensification in RCP 2.6 (10 and 12 % for the HadGEM2-ES and MIROC5 models, respectively). Forests (evergreen broadleaf forest areas) will have a greater carbon fixation capacity. In general, the Amazon forest will have an increased carbon fixation capacity by the end of the century.

Keywords: Random Forest, Machine Learning, Carbon sink, Amazon Forest.

Caderno de Geografia (2023) v.33, n.72
ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

Resumo
A Floresta Amazônica é um dos principais sumidouros de carbono (CO2) do planeta. No entanto, recentemente, devido a atividades antrópicas e mudanças climáticas, perdeu sua estabilidade na absorção de CO2. Portanto, entender a dinâmica dos cenários futuros das mudanças climáticas torna-se essencial. Avaliamos a influência de cenários futuros de mudanças climáticas nos níveis de NPP (biomassa) na Floresta Amazônica usando modelos de ML. Os modelos testados foram Bayesiano, Modelo Linear e Floresta Aleatória. O cenário atual foi avaliado usando 19 covariáveis bioclimáticas (conjunto de dados *worldclim*). Enquanto os cenários futuros foram baseados nos RCPs 2.6 e 8.5 (baseados nos modelos dos modelos *MIROC5* e *HadGEM2-ES*). A Floresta Aleatória teve o melhor desempenho estatístico (R^2 = 0,71 no treinamento e 0,68 no teste *holdout*). Os cenários de mudanças climáticas implicarão em aumento da NPP média para a floresta amazônica, especialmente com maior intensificação no RCP 2.6 (10 e 12 % para os modelos *HadGEM2-ES* e *MIROC5*, respectivamente). As florestas (áreas de florestas de folhas largas perenes) terão maior capacidade de fixação de carbono. Em geral, a floresta amazônica terá maior capacidade de fixação de carbono até o final do século.

Palavras–chave: Floresta Aleatória, Aprendizagem de Máquina, Sumidouro de Carbono, Floresta Amazônica.

1. INTRODUCTION

Anthropogenic actions have threatened to raise (double) the emission levels of greenhouse gases on a global scale, especially CO2, thereby accelerating the implications of climate change (LINDNER *et al.*, 2010). One alternative, outlined as a solution, is the maintenance of forests as natural carbon reservoirs because they occupy large areas of the world (~30–40%) and are responsible for the sequestration of 359 × 106 t of atmospheric carbon (ALLEN *et al.*, 2010; HUI *et al.*, 2017). In this context, the domains of tropical forests stand out because they represent approximately 25% of the forested areas on Earth and have a high potential for carbon fixation (15–25%) (POORTER *et al.*, 2015; RAMMIG; LAPOLA, 2021).

The Amazon rainforest is the most extensive tropical system on the planet, covering 3% of Earth's surface; it is one of the central biodiversity repositories with a high capacity to provide ecosystem services (HEINRICH *et al.*, 2021). Among the ecosystem services, the absorption of CO2 is remarkable, storing approximately 10% of global forest carbon (120,000 Tg C) (AVITABILE *et al.*, 2016; LAPOLA *et al.*, 2018). Despite this, the Amazon rainforest has experienced several processes associated with the loss of natural vegetation due to systematic deforestation events (JUNIOR *et al.*, 2021; SOARES-FILHO *et al.*, 2006; ZEFERINO *et al.*, 2021). According to the National Institute for Space Research (INPE), the Amazon lost 10,129 km² of its forested area in 2019, an increase of 34% compared to the previous year. In 2020, deforestation rates in the Amazon rainforest

ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

were 182% above that allowable by law (JUNIOR *et al.*, 2021). Therefore, this situation inverts the function of the forest, i.e., it becomes a CO2 emitter rather than a sink (HEINRICH *et al.*, 2021; SONG *et al.*, 2015).

In addition to the occurrence of deforestation over the last four decades, the Amazon forest has been affected by climate change. Recent studies have shown that climate change intensifies the dry season, creating stress on the ecosystem, especially in the eastern Amazon, a region with the highest rates of deforestation (GATTI *et al.*, 2021). The consequence of this dynamic is a negative balance of CO2 absorption by the forest; in contrast, in future scenarios of CO2 increase, studies show that there is an effect from CO2 fertilization that contributes to biomass gain (LYRA *et al.*, 2017). However, there are still uncertainties concerning this carbon fertilization capacity for long-term climate change mitigation, possibly due to the limiting effects of plant nutrients, such as nitrogen and phosphorus, which may inhibit increases in biomass increase (Wang *et al.*, 2020).

Biomass dynamic projections in climate change scenarios that contemplate the extensiveness of the Amazon and adjacent areas are at a global scale (coarse-scale: 0.5° resolution) (YU *et al.*, 2019). Some of these studies are based on modeling (e.g., integrated models of land surface processes) (LYRA *et al.*, 2017) and field measurement data (FLEISCHER *et al.*, 2019). Therefore, studies must use more detailed scales focused on regional contexts, prioritizing data, and open-access software. From this perspective, two branches of scientific study have emerged as viable alternatives: remote sensing and machine learning. Studies have shown the efficiency of combining these methods (ZHANG *et al.*, 2020), but there is still a gap in biomass predictions involving future scenarios and those at a regional scale in the Brazilian Amazon.

The main advantage of remote sensing involves compensation for insufficient field data for some environmental variables, which are undoubtedly costly in regional contexts with difficult accessibility (e.g., the Legal Amazon) (PIAO *et al.*, 2012; SUN; MU, 2018; YU *et al.*, 2019). Among these products, MOD17A3HGF has been widely used (RUNNING; ZHAO, 2019) because it represents the Net Primary Productivity (NPP), which defines the levels of carbon fixed by vegetation during photosynthesis (POTTER; KLOOSTER; GENOVESE, 2012).

For machine learning (ML), environmental prediction studies focus on a methodological framework that combines open-access algorithms from well-developed branches of statistics and a dataset of auxiliary covariates for prediction. ML is essential in climate change studies because the dynamics of most environmental variables of this

ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

nature involve nonlinear relationships that require robust algorithms for explanations (KESKIN; GRUNWALD; HARRIS, 2019). Furthermore, the ML methodology supports the insertion of covariates; for climate change studies, covariates from Global Climate Models (GCMs) are freely available. GCMs are climate projections formulated by the Intergovernmental Panel on Climate Change (IPCC) (VAN VUUREN *et al.*, 2011). Therefore, these scenarios consider the influence of human action on climate change (VAN VUUREN *et al.*, 2011). These scenarios are known as Representative Concentration Routes (RCP), which vary according to the intensity of the radioactive forces (RCPs 2.6, 4.5, 6, and 8.5).

Several studies have used RCPs to measure the effects of climate change on vegetation biomass levels and, consequently, on the capacity to absorb atmospheric carbon. Yao, Piao and Wang (2018) analyzed forest areas in China using different phases of the Coupled Model Intercomparison Project (CMIP5) scenarios. They found that climate change will (slightly) increase carbon sequestration by 0.52–0.60 Pg C by 2040. In contrast, Sung *et al.*, (2016) used RCPs for forests in South Korea. They estimated that if these ecosystems are not properly managed, the NPP will decrease by 90%. If such changes occur, there will be a 50% increase in their capacity to sequester carbon. Considering CO2 levels between 421 and 936 ppm and temperature rise ranging from 1.1 to 2.6 °C, Berberoglu, Donmez, and Cilek (2021) projected a substantial rise in the NPP in Turkey. Overall, the use of RCPs has intrinsically helped in understanding the impacts of climate change scenarios on vegetation productivity in different parts of the world.

Therefore, the objective of this study was to evaluate the influence of future climate change scenarios on NPP (biomass) levels in the Amazon Forest using ML models. To this end, i) machine learning models were evaluated to explain the spatial NPP distribution in the AML in future scenarios using covariates from the RCPs; ii) based on this modeling, the covariates with the highest level of importance for the NPP dynamics were ranked; and iii) the NPP in future scenarios was extracted and analyzed at the vegetation domain level.

2. MATERIAL AND METHODS

2.1. STUDY AREA

The study area consisted of the Legal Amazon (LA) do Brasil political unit (Fig. 1), covering 59% of the Brazilian territory (5 × 106 km2). The LA involves nine states of the federation, has a population of 27.5 million inhabitants, and has a low population density.

Caderno de Geografia (2023) v.33, n.72
ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

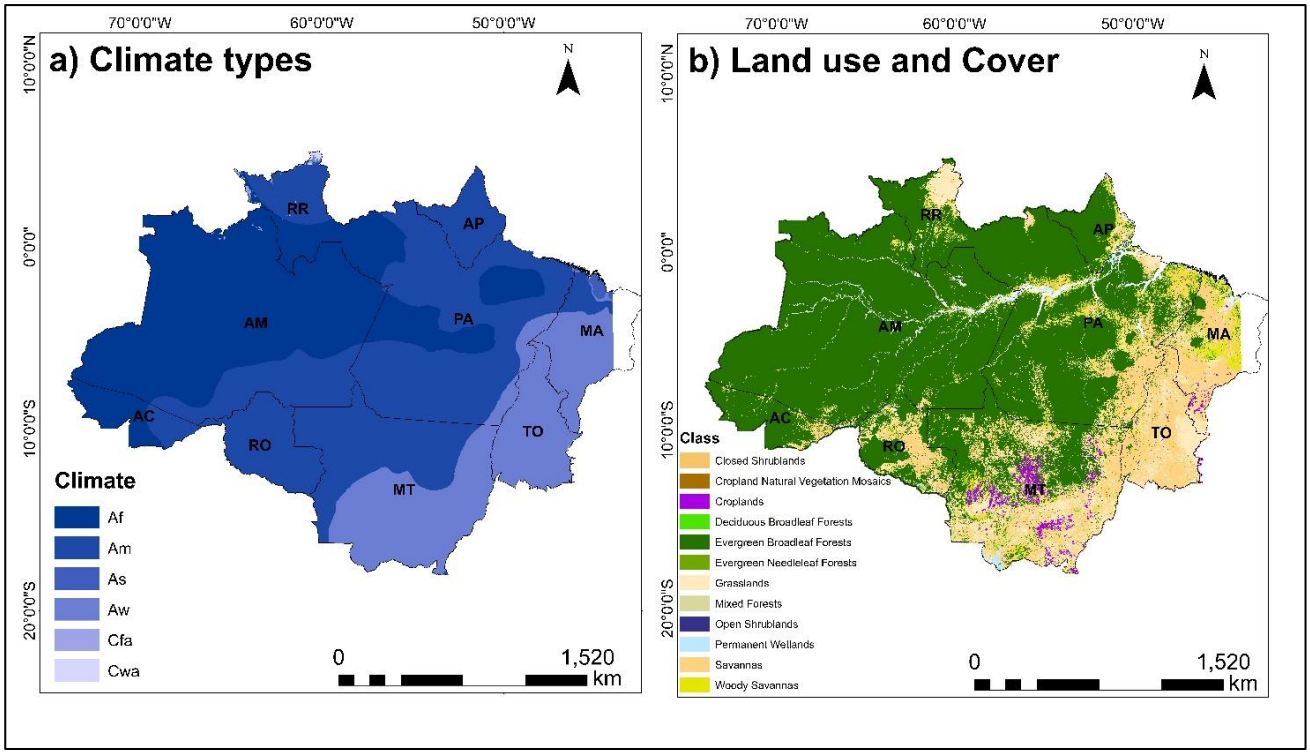

Figure 1 - a) Climate types according to Köppen classification. b) land use and land cover classes in the Amazon.

The LA has three climatic domains, with the most significant territorial expressiveness framed by the Köppen classification (Fig. 1a). The Tropical monsoon climatic type (Am) occupies 43% of the LA and occurs mainly in the central part of the SW–NE extension. It is characterized by two seasons: one is typically hot and rainy while the other is milder and drier. The typical tropical rainforest climate (Af) occupies 38% of the area, dominates the SW area of the LA, and does not have a dry season, with precipitation during the driest month of at least 60 mm. In the SE extension, the dominant climate type is tropical wet-dry (Aw), occupying 19% of the LA. The Aw climate is markedly seasonal, with rainfall above 250 mm per month in the austral summer and a dry winter.

For vegetation, the Amazon has diversified mosaics of land use and cover, such as closed shrublands, croplands, cropland natural vegetation mosaics, deciduous broadleaf forests, evergreen broadleaf forests, grasslands, mixed forests, open shrublands, savannas, and woody savannas (SULLA-MENASHE; FRIEDL, 2019) (Fig. 1b). This varied mosaic gives the Amazon a unique importance in a regional and global context regarding ecosystem services. For example, forest composition plays a fundamental role in distributing moisture (rainfall inputs) to the midwestern portion of Brazil, a critical region for the Brazilian agribusiness sector, mainly due to large-scale global grain exports. The Amazon also has a rich biodiversity of fauna and flora (BULLOCK; WOODCOCK, 2021).

Caderno de Geografia (2023) v.33, n.72
ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

In addition, forests are one of the leading CO2 inhibitors globally (BULLOCK; WOODCOCK, 2021), mainly based on their carbon fixation capacity (through the NPP).

2.2. METHODOLOGICAL STRUCTURE

We constructed the following methodological structure: i) obtained the annual NPP and calculated the average for the current scenario, ii) selected predictor covariates to explain the NPP in future scenarios, iii) removed covariates with high correlation, iv) trained and validated the machine learning models, and v) spatial prediction (Fig. 2).

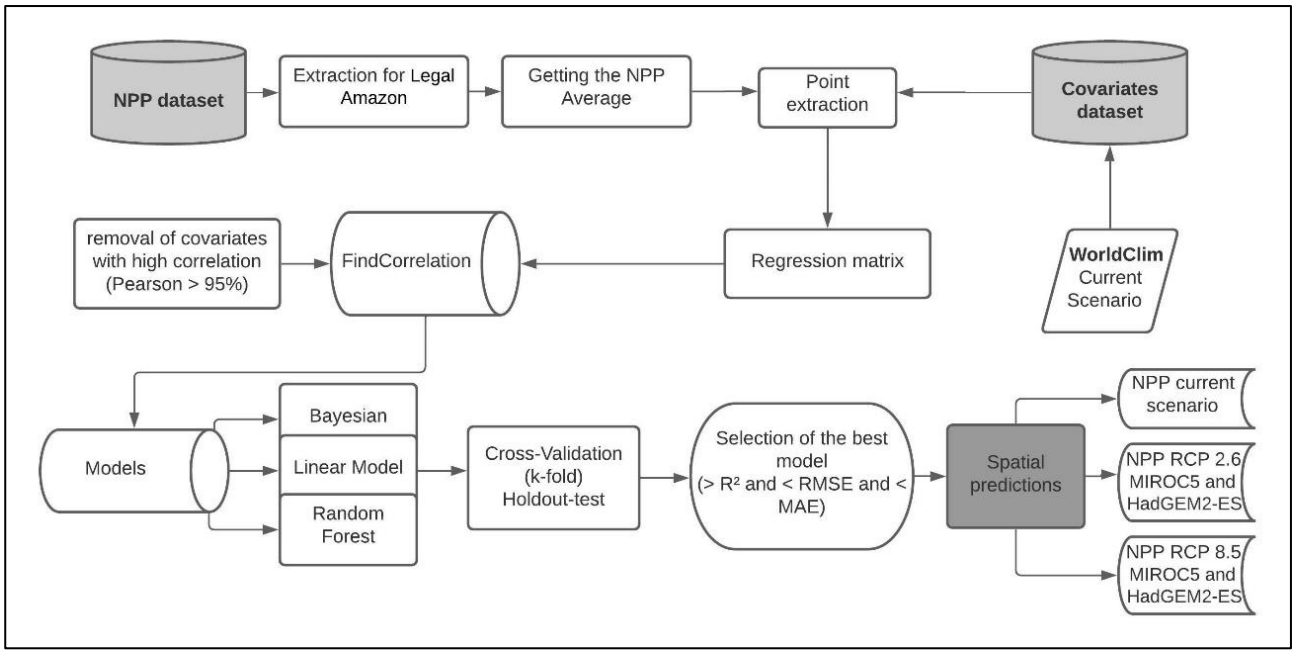

Figure 2 - Flowchart of the methodological structure to obtain NPP spatial predictions in current and future scenarios (RCP 2.6 and 8.5 by models MIROC5 and the HadGEM2-ES) from the Bayesian, Linear Model and Random Forest Machine Learning models. NPP: Net Primary Productivity; R^2: R-Squared; RMSE: root mean squared error; MAE: mean absolute error.

2.3. DATABASE FOR THE DEPENDENT VARIABLE

The dependent variable in the analysis was the NPP, which indicates the capacity of vegetation to fix carbon through the photosynthetic process (YU *et al.*, 2019). The NPP was obtained from the Terra Moderate-Resolution Imaging Spectroradiometer (MODIS), specifically MOD17A3HGF version 6 (RUNNING; ZHAO, 2019). This product has a spatial resolution of 500 m, but we resampled it to 1 × 1 km using the bilinear interpolation method to obtain equivalence with the predictive covariates of the modeling. The time range of the NPP data included the entire MODIS orbital monitoring period (21 years: 2000–2020). Subsequently, we built a randomly distributed grid of points with a minimum

ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

distance of 1 km for the entire area of the LA, extracting the average value for the NPP per point.

2.4. COVARIATE DATABASE

We established a database of covariates to support NPP prediction in the study area for the current and future scenarios. The covariates were derived from WordClim Data and included 19 rasters with a 1 × 1 km resolution covering the globe with various weather information (HIJMANS *et al.*, 2005). WorldClim is a verified data source that provides readily available information on current and future scenarios. Therefore, we divided the analysis into two parts: i) NPP prediction for the current scenario (1960–1990) and ii) future climate change scenarios (2061–2080).

In the current scenario, the covariates were derived from meteorological stations distributed around the globe, including thermal and rainfall aspects (Table 1). For future climate change scenarios, we used a set of covariates with the same nomenclature as the variables used in the current scenario; however, they were derived from two greenhouse gas emissions from the Coupled Model Intercomparison Project 5 (CMIP5) (HIJMANS *et al.*, 2005).

These new scenarios are known as representative concentration pathways (RCPs); therefore, we selected two RCP scenarios (RCPs 2.6 and 8.5) to assess the NPP. RCP 2.6 is the most optimistic scenario considering climate change: by the mid-21st century radioactive forcings are projected to increase by up to 3.1 W m^{-2}, with a rapid decline to 2.6 W m^{-2} by the end of the century (VAN VUUREN *et al.*, 2011). In contrast, RCP 8.5 is the most pessimistic, where radioactive forcing will reach 8.5 W m^{-2} by the end of the 21st century (RIAHI *et al.*, 2011). These scenarios were selected because they represent extreme contexts. We considered two GCMs, MIROC5 and HadGEM2-ES, which are both widely used in South America (CAVALCANTI; SHIMIZU, 2012; DERECZYNSKI *et al.*, 2020).

ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

Table 1 - Covariates used for training the Machine Learning models and NPP spatial prediction in the Legal Amazon in current (1969) and future (2061 - 2080) scenarios.

Abbreviation	Covariates of current and future scenarios	Abbreviation	Covariates of current and future scenarios
Bioc 01	Mean annual temperature	Bioc 11	Mean temperature of coldest quarter
Bioc 02	Mean diurnal range	Bioc 12	Annual precipitation
Bioc 03	Isothermality	Bioc 13	Precipitation of wettest month
Bioc 04	Temperature seasonality	Bioc 14	Precipitation of driest month
Bioc 05	Maximum temperature of warmest month	Bioc 15	Precipitation seasonality
Bioc 06	Minimum temperature of coldest month	Bioc 16	Precipitation of wettest quarter
Bioc 07	Temperature annual range	Bioc 17	Precipitation of driest quarter
Bioc 08	Mean temperature of wettest quarter	Bioc 18	Precipitation of warmest quarter
Bioc 09	Mean temperature of driest quarter	Bioc 19	Precipitation of coldest quarter
Bioc 10	Mean temperature of warmest quarter		

2.5. NPP SPATIAL MODELING

In the prediction step using the algorithms, we used the grid points with the NPP value and sequentially extracted the values of the covariates in the scenarios (Current, RCP 2.6, and RCP 8.5). After extraction, we obtained a regression matrix to train the machine learning models using the R software (TEAM, 2018).

As the covariates used were from the same data source (WorldClim), redundancy between the variables was expected, increasing the uncertainty of the predictions. Therefore, we applied the FindCorrelation function (Caret package) (KUHN *et al.*, 2017) to discard covariates with a high correlation using Pearson > 0.95. This process ensures the fluidity of the model and the principle of parsimony while avoiding overestimated projections (SOUZA *et al.*, 2018).

Subsequently, we randomly divided the database: 80% of the samples for training and 20% for the holdout-test. For the training process (80% of the data), we used three ML models: Random Forest (RF), linear regression model (LM), and Bayesian regularized neural network (BRNN). The RF is an ensemble model in which a set of random trees is built so that the training process occurs and the subsequent prediction of the dependent variable is based on the covariates (BREIMAN, 2002). The final prediction was based on the average of all constructed trees. This model, as well as most MLs, understands non-linear relationships in space and time. Unlike RF, LM analyzes the linear relationships between the dependent variable (in this case NPP) and predictor covariates (DOTTO *et al.*, 2018). Thus, the linear patterns imposed by the covariates have a greater explanation for the variation in the NPP in view of the spatial predictions. The BRNN has been widely

ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

used for spatial prediction analyses, especially considering its ability to prevent overfitting (GARG; MISHRA, 2018). This model was developed from probabilistic interpretations of networks; therefore, BRNN is a linear combination of ANN and Bayesian methods to define/determine the regularization of parameters (GARG; MISHRA, 2018). All models are available through the Caret package.

The adjustment, training, and validation (Cross-Validation) of the models were performed from the "trainControl" function in the Caret package. Cross-validation provides statistical indices for training evaluations, such as R^2, RMSE, and MAE (Eqs. (1), (2), and (3), respectively). The remaining 20% of the samples were used to calculate the overall performance of the models, as these data were not observed in the training process, but as an external validation (holdout-test). Comparing the statistical indices obtained in the Cross-Validation and Holdout-test is crucial for assessing overfitting in the predictions (GOMES *et al.*, 2019).

$$RMSE=\sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i}-X_{mod,i})^2}{n}} \quad \text{Equation 01}$$

$$R^2=1-\frac{\sum_{n}(X_{obs,i}-X_{mod,i})^2}{\sum_{n}(X_{obs,i}-\bar{X})^2} \quad \text{Equation 02}$$

$$MAE=\frac{\sum_{i=1}^{n}|X_{obs,i}-X_{mod,i}|}{n} \quad \text{Equation 03}$$

where xobs represents the observed values for the NPP and xmod are the values predicted by the models. Training was carried out only with the covariates of the current scenario, whereas the spatial predictions were carried out with datasets that represent climate change in future scenarios in the MIROC5 and HadGEM2-ES models.

2.6. EXTRACTION OF NPP VALUES FOR CURRENT AND FUTURE SCENARIOS DOMAINS OF USE AND LAND COVER IN AMAZON RAINFOREST

After training and the holdout test, we obtained the best model to perform spatial predictions based on the statistical indices (> R^2, < RMSE, and < MAE). We extracted the average values of Land Use and Coverage for the Amazon rainforest region using spatial predictions.

We use the 2019 (most recent) MCD12Q1 (SULLA-MENASHE; FRIEDL, 2019) product to represent Land Use and Coverage (LUC). For the Amazon rainforest, the MCD12Q1 originally had 12 classes; however, we removed categories (e.g., urban area,

ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

barren, and water bodies) because they have inconsistent NPP values. The final mapping presented 11 classes: closed shrublands, croplands, cropland natural vegetation mosaics, deciduous broadleaf forests, evergreen broadleaf forests, grasslands, mixed forests, open shrublands, savannas, and woody savannas (Fig. 1b).

3. RESULTS

3.1. PERFORMANCE OF NPP PREDICTION MODELS

The methodological framework developed to predict NPP prioritized the fluidity of the modeling; five covariates (Bioc1, Bioc7, Bioc9, Bioc16, and Bioc17) with high correlations identified by the FindCorrelation function were removed from the database with 19 covariates. In the performance evaluation of the tested ML models, statistical indices (R^2, RMSE, and MAE) were provided during the training and holdout-test processes. Among the tested models, RF was the most efficient at predicting the NPP. Therefore, in the training phase, the RF showed high R^2 (0.71) and low error levels (RMSE = 0.16 and MAE = 0.08 Kg C/m²) with respect to the other models (Fig. 4). In the holdout-test, the metrics were similar to the training, indicating low overfitting in the modeling, with an R^2 of 0.68 and low RMSE and MAE (Table 2).

During the training process, the RF selected the covariates that best explained the spatial distribution of the NPP in the LA. The covariates were ranked according to the degree of importance (%overall); Bioc 19, Bioc 04, and Bioc 14 were ranked with the highest level of importance (Fig. 4).

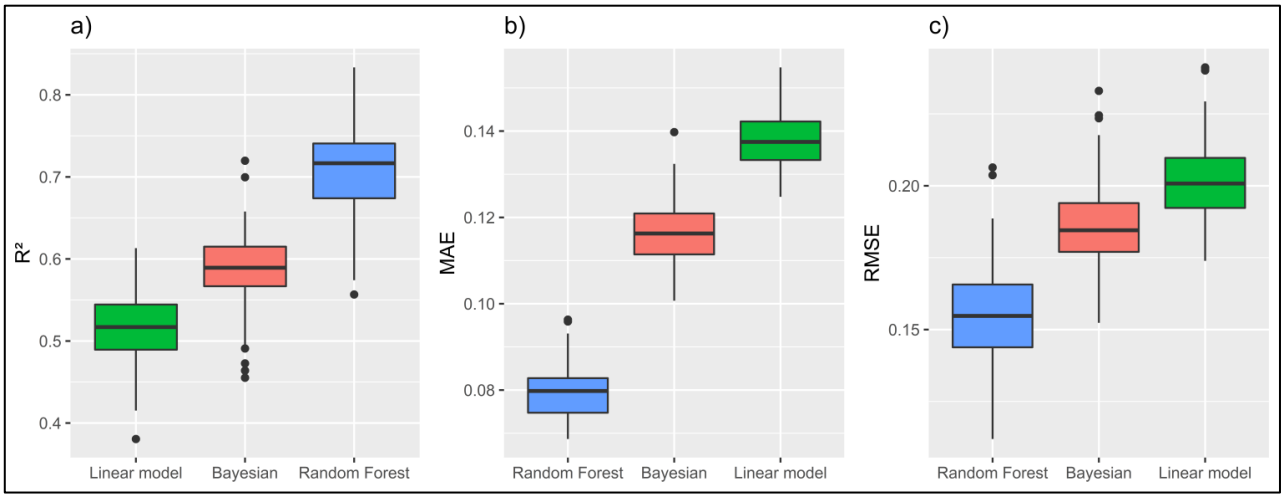

Figure 3 - Machine Learning model performance in the training phase (Bayesian, Linear Model, and Random Forest). a) R-squared (R^2), b) mean error absolute (MAE), and c) root-mean-square error (RMSE).

Caderno de Geografia (2023) v.33, n.72
ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

Table 2. Training and holdout-test for the tested models (Bayesian, Linear Model, and Random Forest). R^2: R-Squared; MAE: mean absolute error; and RMSE: root-mean-square error.

	Training			Holdout-test		
Models	R^2	MAE	RMSE	R^2	MAE	RMSE
Bayesian	0.59	0.12	0.19	0.47	0.14	0.20
Linear Model	0.51	0.14	0.20	0.56	0.12	0.19
Random Forest	0.71	0.08	0.16	0.68	0.08	0.16

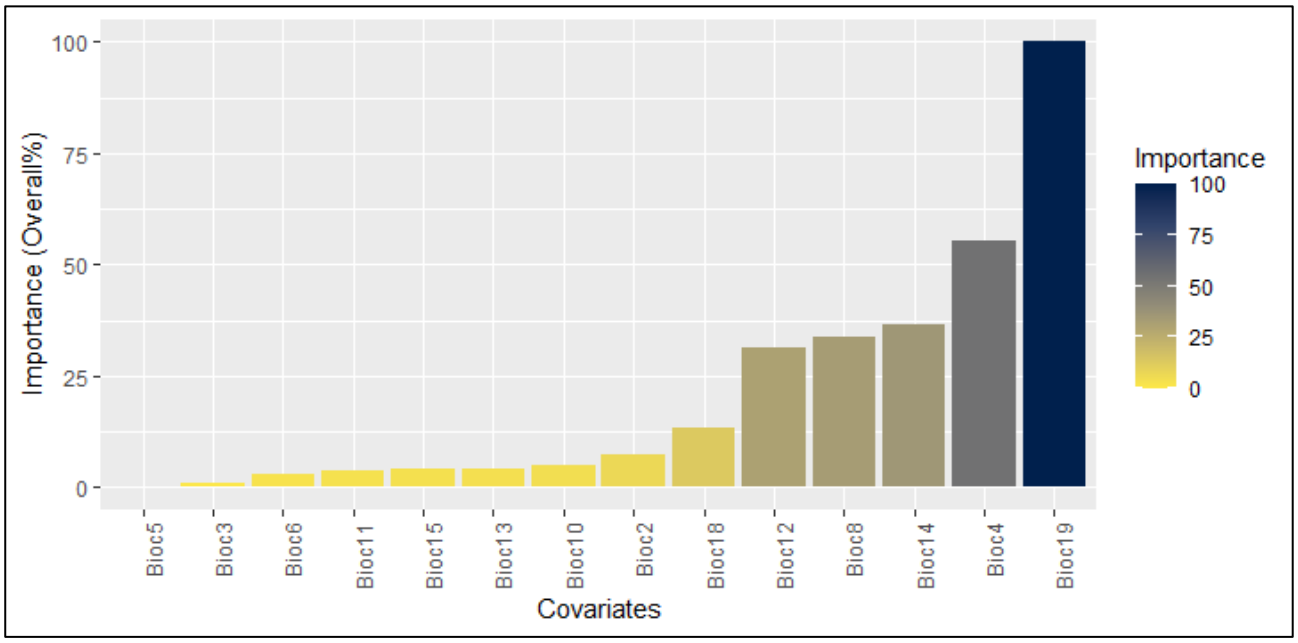

Figure 4 - Ranking of the most important covariates selected by the Random Forest model to explain the spatial distribution of NPP. y-axis degree of importance of covariates (Worldclim dataset) and x-axis: predictor covariates.

3.2. NPP SPATIAL DISTRIBUTION FOR CURRENT AND FUTURE SCENARIOS

From the prediction made by RF, the mean NPP will increase slightly in future scenarios (RCP 2.6 and 8.5) in the Global Climate Model used (ie HadGEM2-ES and MIROC5) (Figure 5). Furthermore, in RCP 2.6, the increase in mean NPP is more significant (10 and 12% for the HadGEM2-ES and MIROC5 models, respectively), while in RCP 8.5, the intensity is relatively low (6 and 10%).

The spatial distribution pattern of the NPP was similar between the HadGEM2-ES and MIROC5 models in RCPs 2.6 and 8.5 (Fig. 5). In general, there will be an expansion of carbon fixation cores in the western portion (> 1.50 Kg C/m²). In contrast, there will be an intensification of carbon emissions source cores in the southern areas of the LA (< 0.71

Caderno de Geografia (2023) v.33, n.72
ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

Kg C/m²); part of this region covers an area with a higher rate of deforestation (deforestation arc).

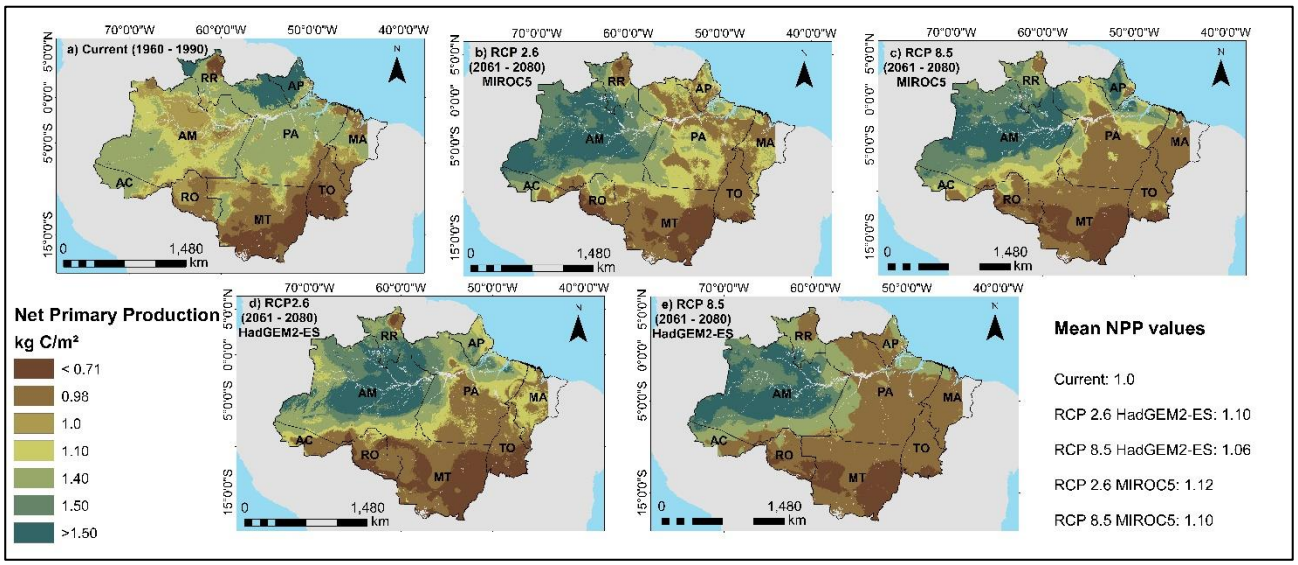

Figure 5 - Net Primary Productivity (NPP) Spatial Distribution for the Amazon Forest in Current and Future Scenarios. a) Current; b) RCP 2.6 MIROC5; c) RCP 8.5 MIROC5; d) RCP 2.6 HadGEM2-ES; e) RCP 8.5 MIROC5.

The behavior of the NPP during LUC showed variability with the climate change scenarios (RCPs 2.6 and 8.5) (Fig. 6). Considering the average NPP values, we highlight a substantial increase in carbon fixation for evergreen broadleaf forest areas because they comprise most of the Amazon rainforest (~65%). For mixed forests, closed shrublands, and woody savanna classes, a decrease in NPP is expected. In the other LUC, we found that the NPP levels remained stable.

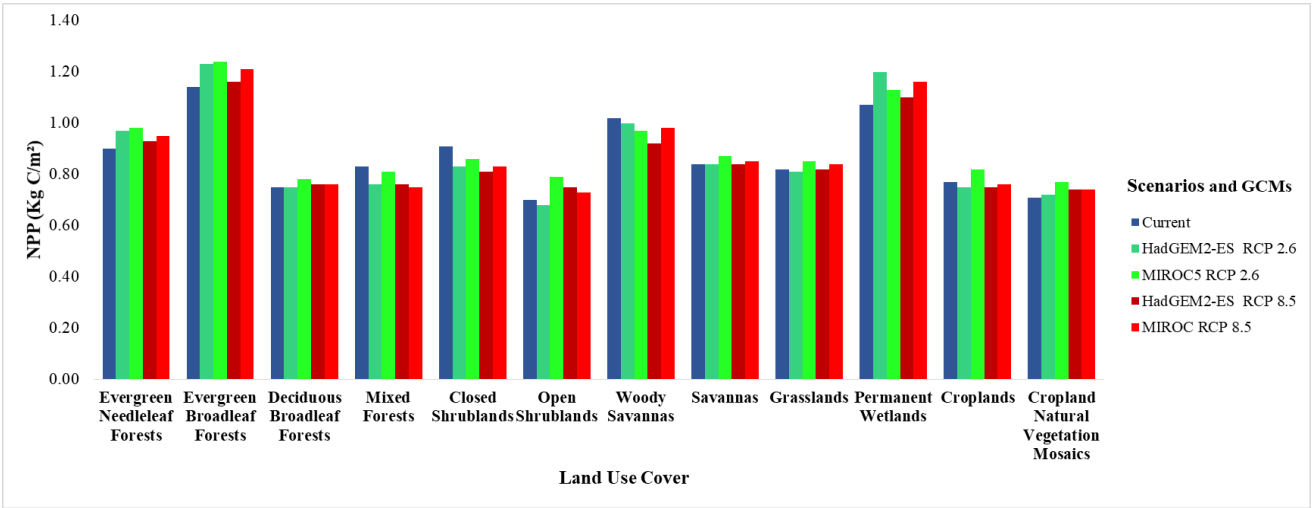

Figure 6 - Average NPP (Net Primary Productivity) for LUC (Land Use Cover) in the current scenarios and RCPs 2.6 and 8.5 for HadGEM2-ES and MIROC5 models.

ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

4. DISCUSSION

4.1. MODELING PERFORMANCE

The RF was the ML with the best performance in explaining the levels of carbon fixation (NPP) in the LA (> R^2 and < RMSE and MAE). The RF model had the greatest capacity to represent nonlinear relationships between the NPP and climatic covariates. RF is widely known for its randomness characteristics and statistical robustness (ROY; LAROCQUE, 2012). It stands out owing to its use of the bootstrap technique (bagging), in which the original training database is resampled, thus creating a new set of random samples (LEE; ULLAH; WANG, 2020; TYRALIS; PAPACHARALAMPOUS; LANGOUSIS, 2019). This new dataset builds uncorrelated trees, thereby decreasing model variance (FERREIRA *et al.*, 2021).

In addition to the bootstrap, we emphasize that RF is an ensemble model, where the final prediction (in regression) is provided by the average of all of the decision trees (aggregation step), which reduces spatial prediction errors and susceptibility to overfitting (LIN *et al.*, 2017). This corroborates the observed results, where the R^2 of the training was similar to that of the test (Fig. 5 and Table 2).

4.2. INFLUENCE OF CLIMATE CHANGE SCENARIOS ON CARBON FIXATION (BY NPP) IN THE BRAZILIAN LEGAL AMAZON

We evaluated the influence of climate change scenarios (RCPs 2.6 and 8.5) on carbon fixation (through NPP) in the Legal Brazilian Amazon using Machine Learning models. The main findings of the study were as follows: (i) there will be an increase in the NPP in the Amazon rainforest based on the RCPs examined and ii) at the LUC level, areas of evergreen broadleaf forest areas will fix larger portions of carbon in future scenarios while classes of mixed forests, closed shrublands, and woody savannas will present lower NPP levels.

The increase in the NPP in the Amazon region due to climate change agrees with other studies conducted in other regions and at a global scale (GANG *et al.*, 2017; MICHALETZ *et al.*, 2014; SUN; MU, 2018). Variations in NPP in different parts of the globe have mainly been associated with climate change, especially fluctuations in temperature and precipitation (water availability) (MICHALETZ *et al.*, 2014). As the covariates included in the modeling of this study were derived from temperature and precipitation, they

ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

certainly helped explain the NPP patterns, e.g., the most important covariate in the NPP distribution (Bioc 19: Precipitation of the Coldest Quarter) (Fig. 5).

The GCMs (i.e., HadGEM2-ES and MIROC5) used in this study project an increase in rainfall inputs in the climate change scenarios (RCPs 2.6 and 8.5) for the LA (HIJMANS *et al.*, 2005). Therefore, the water availability in future scenarios will firmly control the dynamics of carbon fixation in the LA: precipitation favors the photosynthetic activity of vegetation (GANG *et al.*, 2017). Consistent with our study, Azhdari *et al.* (2020) evaluated the impacts of climate change (RCP 4.5) on NPP in southern Iran using the HadGEM2-ES model. They found that the increase in carbon fixation levels was a function of increased precipitation. Similarly, in North American regions, Duveneck and Thompson (2017) used the CCSM4, CESM1, HADGE, and MPIMLR models under RCP 8.5, concluding that the increase in NPP is associated with high rainfall inputs.

Although climatic variables appear to be the main drivers of NPP in the Amazon rainforest (especially precipitation), atmospheric CO2 is also largely responsible for its fluctuations (MICHALETZ *et al.*, 2014). In addition to the projected climate dynamics arising from the RCPs, these scenarios were formulated based on the CO2 levels (SCHWALM; GLENDON; DUFFY, 2020). Vegetation tissues are built from the absorption of CO2 (JIANG *et al.*, 2019), subsequently favoring phenological development and leading to high levels of NPP (PIAO *et al.*, 2012). Thus, there is a fertilization process in which CO2 is essential for plant growth on Earth's surface. Yu *et al.* (2019) projected the global impacts of climate change on the NPP using RCPs 2.6, 4.5, and 8.5. They found that as the concentration of CO2 increased, there was an increase in the NPP. For global scales using the Special Report on Emission Scenarios (SRES), Pan *et al.*, (2014) also noted that the effects of CO2 fertilization will result in a 12–13.9% increase in the NPP by the end of the 21st century. Sun and Mu (2018) used 10 model GCMs with RCP 4.5, observing that climatic variables and CO2 fertilization were the main factors for the increase in the NPP.

Climatic dynamics and CO2 concentration influence aspects at the surface level, such as LUC. Therefore, forests (evergreen broadleaf forest areas) have a favorable structure for carbon fixation, especially considering the high rates of water loss by evapotranspiration, which is proportional to carbon assimilation through the stomata (LIU; WU; WANG, 2017; MCCAUGHEY *et al.*, 2006). Similar to our findings, other studies have also found higher levels of carbon fixation in forested areas (AZHDARI *et al.*, 2020; GANG *et al.*, 2017). Although the increase in the NPP presents a gain in carbon assimilation rates, some aspects in the ecosystem context should be observed. For example, carbon

ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

fixation tends to boost the phenological development of vegetation, which responds rapidly to the loss rates of water by evapotranspiration. Studies have shown that this can affect watershed runoff (reducing), subsequently decreasing water production (TIAN *et al.*, 2016). In contrast, there is a trade-off: even decreasing water flows can contain the effects of soil erosion, as well as hydro-sedimentological flows (BERBEROĞLU *et al.*, 2019). Other theoretical evidence has shown that although an increase in the NPP occurs in forest areas under the effects of climate change, this may have ecosystem implications. By changing overall biodiversity patterns, an increase in NPP can deliberate the transition of species, such as pastures or savannas (low-density phytophysiognomies biomass), in forests (GRAHAM *et al.*, 2016). This does not negate the importance of carbon sequestration through forests; however, caution is required. Overall, we should not indicate erroneous points that can lead to conflicts of interest among different groups.

Another issue requiring analysis in the LA is the context of human activities, especially those regarding the systematic conversion of land use and coverage, particularly in terms of pastures and agricultural land (BARONA *et al.*, 2010; MANN *et al.*, 2014). Although high levels of carbon fixation are projected in future scenarios for forest areas, if the pace of anthropic action is maintained, these projections could be affected, compromising climate change mitigation on a planetary scale. Our results indicate that the lowest NPP levels in future scenarios are intended for LUCs belonging to the deforestation arc, i.e., the region responsible for considerable deforestation in the LA (ALDRICH *et al.*, 2012).

Considering previous advances in anthropic use in the LA, several studies have analyzed the projections of changes for LUC in this region (SOARES-FILHO *et al.*, 2006; ZEFERINO *et al.*, 2021). These studies indicate advances in the loss of forests due to deforestation. Deforestation rates can lead to a decrease in precipitation on the order of 10–20% in this region (MOORE *et al.*, 2007), enhancing drought and fire (FARIA *et al.*, 2017). This is crucial for carbon fixation dynamics, mainly because it reduces the photosynthetic activity of forests, reducing the potential for carbon absorption and intensifying emissions (ARAGÃO *et al.*, 2018). Therefore, if these scenarios occur, the intensification of carbon emissions sources will extend across the region.

In addition to deforestation, the inclusion of land-use classes with low NPP levels can boost CO2 emissions. However, the characteristics of low NPP levels in these classes are possibly associated with the deforestation arc: they present zones with high temperatures and low levels of precipitation (in the current and future scenarios)

ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

(HIJMANS *et al.*, 2005). This tends to imply low levels of soil water availability (low levels), thus altering the vegetation's ability to maintain its metabolic activities. This subsequently decreases plant respiration rates and carbon fixation levels (REICHSTEIN *et al.*, 2002; VAN DER MOLEN *et al.*, 2011), thereby becoming source emitters of CO2. In addition to carbon emissions, in terms of implications, classes with low NPP levels can lead to low evapotranspiration rates, resulting in lower levels of moisture transfer to the upper layers of the atmosphere. This harms the distribution of moisture in Brazil, especially in the Midwest, i.e., the center of agricultural production, which can generate serious economic impacts. The inclusion of classes with lower NPP levels can also affect the dynamics of energy fluxes, increasing the surface temperature (CHAGAS *et al.*, 2019) and boosting sensible heat fluxes, which results in air drying.

In summary, the increase in the NPP in future scenarios for forests is highly positive for ecosystem maintenance in the LA and is crucial for plans to mitigate the impacts of climate change. However, projected advances in deforestation could jeopardize their potential as a carbon sink. Based on our findings, we suggest inserting NPP spatial dynamic projections into the mitigation plans.

5. CONCLUSION

Among the machine learning models tested to assess the influence of climate change on Net Primary Productivity (NPP) in the Brazilian LA, RF had the best performance ($R^2 = 0.71$ for training and $R^2 = 0.68$ for the holdout test).

Climate change will imply an increase in the average NPP for the LA, especially with a greater intensification in RCP 2.6 (10 and 12 % for the HadGEM2-ES and MIROC5 models, respectively). Considering the implications of climate change on the NPP levels for LUC, we found that forests (evergreen broadleaf forest areas) will have the greatest potential for carbon sequestration in RCPs 2.6 and 8.5. We emphasize that information on the future impacts of climate change in the NPP of the Brazilian LA is crucial for decision-making and mitigation plans for future scenarios, especially considering the global importance of the LA in the dynamics of carbon sequestration.

ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

REFERENCES

ALDRICH, S. *et al.* Contentious land change in the Amazon's arc of deforestation. **Annals of the Association of American Geographers**, v. 102, n. 1, p. 103–128, 2012.

ALLEN, C. D. *et al.* A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. **Forest ecology and management**, v. 259, n. 4, p. 660–684, 2010.

ARAGÃO, L. E. O. C. *et al.* 21st Century drought-related fires counteract the decline of Amazon deforestation carbon emissions. **Nature communications**, v. 9, n. 1, p. 1–12, 2018.

AVITABILE, V. *et al.* An integrated pan-tropical biomass map using multiple reference datasets. **Global change biology**, v. 22, n. 4, p. 1406–1420, 2016.

AZHDARI, Z. *et al.* Impact of climate change on net primary production (NPP) in south Iran. **Environmental monitoring and assessment**, v. 192, p. 1–16, 2020.

BARONA, E. *et al.* The role of pasture and soybean in deforestation of the Brazilian Amazon. **Environmental Research Letters**, v. 5, n. 2, p. 24002, 2010.

BERBEROĞLU, S. *et al.* Estimating spatio-temporal responses of net primary productivity to climate change scenarios in the seyhan watershed by integrating biogeochemical modelling and remote sensing. In: **Climate Change Impacts on Basin Agro-ecosystems**. [s.l.] Springer, 2019. p. 183–199.

BERBEROGLU, S.; DONMEZ, C.; CILEK, A. Modelling climate change impacts on regional net primary productivity in Turkey. **Environmental Monitoring and Assessment**, v. 193, n. 5, p. 1–15, 2021.

BREIMAN, L. Manual on setting up, using, and understanding random forests v3. 1. **Statistics Department University of California Berkeley, CA, USA**, v. 1, p. 58, 2002.

BULLOCK, E. L.; WOODCOCK, C. E. Carbon loss and removal due to forest disturbance and regeneration in the Amazon. **Science of The Total Environment**, v. 764, p. 142839, 2021.

CAVALCANTI, I. F. A.; SHIMIZU, M. H. Climate fields over South America and variability of SACZ and PSA in HadGEM2-ES. 2012.

CHAGAS, M. C. *et al.* Gross primary productivity in areas of different land cover in the western Brazilian Amazon. **Remote Sensing Applications: Society and Environment**, v. 16, p. 100259, 2019.

DERECZYNSKI, C. *et al.* Downscaling of climate extremes over South America–Part I: Model evaluation in the reference climate. **Weather and Climate Extremes**, v. 29, p. 100273, 2020.

DOTTO, A. C. *et al.* A systematic study on the application of scatter-corrective and spectral-derivative preprocessing for multivariate prediction of soil organic carbon by Vis-NIR spectra. **Geoderma**, v. 314, p. 262–274, 2018.

ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

DUVENECK, M. J.; THOMPSON, J. R. Climate change imposes phenological trade-offs on forest net primary productivity. **Journal of Geophysical Research: Biogeosciences**, v. 122, n. 9, p. 2298–2313, 2017.

FARIA, B. L. *et al.* Current and future patterns of fire-induced forest degradation in Amazonia. **Environmental Research Letters**, v. 12, n. 9, p. 95005, 2017.

FERREIRA, R. G. *et al.* Machine learning models for streamflow regionalization in a tropical watershed. **Journal of Environmental Management**, v. 280, p. 111713, 2021.

FLEISCHER, K. *et al.* Amazon forest response to CO 2 fertilization dependent on plant phosphorus acquisition. **Nature Geoscience**, v. 12, n. 9, p. 736–741, 2019.

GANG, C. *et al.* Modeling the dynamics of distribution, extent, and NPP of global terrestrial ecosystems in response to future climate change. **Global and Planetary Change**, v. 148, p. 153–165, 2017.

GARG, D.; MISHRA, A. Bayesian regularized neural network decision tree ensemble model for genomic data classification. **Applied Artificial Intelligence**, v. 32, n. 5, p. 463–476, 2018.

GATTI, L. V *et al.* Amazonia as a carbon source linked to deforestation and climate change. **Nature**, v. 595, n. 7867, p. 388–393, 2021.

GOMES, L. C. *et al.* Modelling and mapping soil organic carbon stocks in Brazil. **Geoderma**, v. 340, p. 337–350, 2019.

GRAHAM, V. *et al.* A comparative assessment of the financial costs and carbon benefits of REDD+ strategies in Southeast Asia. **Environmental Research Letters**, v. 11, n. 11, p. 114022, 2016.

HEINRICH, V. H. A. *et al.* Large carbon sink potential of secondary forests in the Brazilian Amazon to mitigate climate change. **Nature communications**, v. 12, n. 1, p. 1–11, 2021.

HIJMANS, R. J. *et al.* Very high resolution interpolated climate surfaces for global land areas. **International Journal of Climatology: A Journal of the Royal Meteorological Society**, v. 25, n. 15, p. 1965–1978, 2005.

HUI, D. *et al.* Climate change and carbon sequestration in forest ecosystems. **Handbook of climate change mitigation and adaptation**, v. 555, p. 594, 2017.

JIANG, M. *et al.* Towards a more physiological representation of vegetation phosphorus processes in land surface models. **New Phytologist**, v. 222, n. 3, p. 1223–1229, 2019.

JUNIOR, C. H. L. S. *et al.* The Brazilian Amazon deforestation rate in 2020 is the greatest of the decade. **Nature Ecology & Evolution**, v. 5, n. 2, p. 144–145, 2021.

KESKIN, H.; GRUNWALD, S.; HARRIS, W. G. Digital mapping of soil carbon fractions with machine learning. **Geoderma**, v. 339, p. 40–58, 2019.

KUHN, M. *et al.* Caret: classification and regression training. 2016. **R package version**, v. 4, 2017.

LAPOLA, D. M. *et al.* Limiting the high impacts of Amazon forest dieback with no-regrets science and policy action. **Proceedings of the National Academy of Sciences**, v. 115, n. 46, p. 11671–11679, 2018.

LEE, T.-H.; ULLAH, A.; WANG, R. Bootstrap aggregating and random forest. In: **Macroeconomic Forecasting in the Era of Big Data**. [s.l.] Springer, 2020. p. 389–429.

LIN, W. *et al.* An ensemble random forest algorithm for insurance big data analysis. **Ieee access**, v. 5, p. 16568–16575, 2017.

LINDNER, M. *et al.* Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. **Forest ecology and management**, v. 259, n. 4, p. 698–709, 2010.

LIU, Z.; WU, C.; WANG, S. Predicting forest evapotranspiration by coupling carbon and water cycling based on a critical stomatal conductance model. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 10, n. 10, p. 4469–4477, 2017.

LYRA, A. *et al.* Projections of climate change impacts on central America tropical rainforest. **Climatic Change**, v. 141, n. 1, p. 93–105, 2017.

MANN, M. L. *et al.* Pasture conversion and competitive cattle rents in the Amazon. **Ecological economics**, v. 97, p. 182–190, 2014.

MCCAUGHEY, J. H. *et al.* Carbon dioxide and energy fluxes from a boreal mixedwood forest ecosystem in Ontario, Canada. **Agricultural and Forest Meteorology**, v. 140, n. 1–4, p. 79–96, 2006.

MICHALETZ, S. T. *et al.* Convergence of terrestrial plant production across global climate gradients. **Nature**, v. 512, n. 7512, p. 39–43, 2014.

MOORE, N. *et al.* Uncertainty and the changing hydroclimatology of the Amazon. **Geophysical Research Letters**, v. 34, n. 14, 2007.

PAN, S. *et al.* Complex spatiotemporal responses of global terrestrial primary production to climate change and increasing atmospheric CO2 in the 21st century. **PloS one**, v. 9, n. 11, p. e112810, 2014.

PIAO, S. *et al.* Impacts of climate and CO2 changes on the vegetation growth and carbon balance of Qinghai–Tibetan grasslands over the past five decades. **Global and Planetary Change**, v. 98, p. 73–80, 2012.

POORTER, L. *et al.* Diversity enhances carbon storage in tropical forests. **Global Ecology and Biogeography**, v. 24, n. 11, p. 1314–1328, 2015.

POTTER, C.; KLOOSTER, S.; GENOVESE, V. Net primary production of terrestrial ecosystems from 2000 to 2009. **Climatic Change**, v. 115, n. 2, p. 365–378, 2012.

RAMMIG, A.; LAPOLA, D. M. The declining tropical carbon sink. **Nature Climate Change**, p. 1–2, 2021.

REICHSTEIN, M. *et al.* Severe drought effects on ecosystem CO2 and H2O fluxes at three Mediterranean evergreen sites: revision of current hypotheses? **Global Change Biology**, v. 8, n. 10, p. 999–1017, 2002.

RIAHI, K. *et al.* RCP 8.5—A scenario of comparatively high greenhouse gas emissions. **Climatic change**, v. 109, n. 1, p. 33–57, 2011.

ROY, M.-H.; LAROCQUE, D. Robustness of random forests for regression. **Journal of Nonparametric Statistics**, v. 24, n. 4, p. 993–1006, 2012.

RUNNING, S.; ZHAO, M. MOD17A3HGF MODIS/Terra Net Primary Production Gap-Filled Yearly L4 Global 500m SIN Grid V006. **NASA EOSDIS Land Processes DAAC.**, v. 6, 2019.

SCHWALM, C. R.; GLENDON, S.; DUFFY, P. B. RCP8. 5 tracks cumulative CO2 emissions. **Proceedings of the National Academy of Sciences**, v. 117, n. 33, p. 19656–19657, 2020.

SOARES-FILHO, B. S. *et al.* Modelling conservation in the Amazon basin. **Nature**, v. 440, n. 7083, p. 520–523, 2006.

SONG, X.-P. *et al.* Annual carbon emissions from deforestation in the Amazon Basin between 2000 and 2010. **PloS one**, v. 10, n. 5, p. e0126754, 2015.

SULLA-MENASHE, D.; FRIEDL, M. MCD12Q1 MODIS/Terra+ Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006. **NASA EOSDIS Land Processes DAAC: Sioux Falls, SD, USA**, 2019.

SUN, G.; MU, M. Assessing the characteristics of net primary production due to future climate change and CO2 under RCP4. 5 in China. **Ecological Complexity**, v. 34, p. 58–68, 2018.

SUNG, S. *et al.* Estimating net primary productivity under climate change by application of global forest model (G4M). **Journal of Korean Society for People, Plants and Environment**, v. 19, n. 6, p. 549–558, 2016.

TEAM, R. C. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http s. **www. R-proje ct. org**, 2018.

TIAN, Y. *et al.* Trade-offs among ecosystem services in a typical Karst watershed, SW China. **Science of the Total Environment**, v. 566, p. 1297–1308, 2016.

TYRALIS, H.; PAPACHARALAMPOUS, G.; LANGOUSIS, A. A brief review of random forests for water scientists and practitioners and their recent history in water resources. **Water**, v. 11, n. 5, p. 910, 2019.

VAN DER MOLEN, M. K. *et al.* Drought and ecosystem carbon cycling. **Agricultural and Forest Meteorology**, v. 151, n. 7, p. 765–773, 2011.

VAN VUUREN, D. P. *et al.* The representative concentration pathways: an overview. **Climatic change**, v. 109, n. 1, p. 5–31, 2011.

ISSN 2318-2962
DOI 10.5752/p.2318-2962.2023v33n.72p.110

YAO, Y.; PIAO, S.; WANG, T. Future biomass carbon sequestration capacity of Chinese forests. **Science Bulletin**, v. 63, n. 17, p. 1108–1117, 2018.

YU, D. *et al.* Projecting impacts of climate change on global terrestrial ecoregions. **Ecological Indicators**, v. 103, p. 114–123, 2019.

ZEFERINO, L. B. *et al.* Environmental conservation policy can bend the trend of future forest losses in the oriental Amazon. **Regional Environmental Change**, v. 21, n. 2, p. 1–11, 2021.

ZHANG, Y. *et al.* An evaluation of eight machine learning regression algorithms for forest aboveground biomass estimation from multiple satellite data products. **Remote Sensing**, v. 12, n. 24, p. 4015, 2020.

Recebido: 07.03.2022
Aceito: 23.01.2023