The predictive contributions of environmental damage to the public management of conservation units: a case study of the Alto do Mucuri APA in Minas Gerais

As contribuições preditivas do dano ambiental para gestão pública de unidades de conservação: estudo de caso da APA do Alto do Mucuri em Minas Gerais

Diego da Silva Passos[1], Orcid: https://orcid.org/0000-0002-0571-7727; **João Paulo Calembo Batista Menezes**[2], Orcid: https://orcid.org/0000-0002-4739-839X; **Bruno de Almeida Vilela**[3], Orcid: https://orcid.org/0000-0003-2388-1034; **Jorge Fulgencio Silva Chaves**[4], Orcid: https://orcid.org/0000-0002-1653-3069

1. Mestre em Administração Pública e graduado em Ciências Contábeis pela Universidade Federal dos Vales do Jequitinhonha e Mucuri - UFVJM. Atua na área Administrativa e Financeira como Analista Ambiental - Núcleo de Ciências Gerenciais do Instituto Estadual de Florestas - IEF em Minas Gerais.. Universidade Federal dos Vales do Jequitinhonha e Mucuri – UFVJM, Diamantina/MG, Brasil. E-mail: diegosilvapassos@hotmail.com
2. Graduação em Contabilidade (2004) e Administração (2016) pela Pontifícia Universidade Católica de Minas Gerais - PUC Minas, graduação em Ciências Econômicas pela Faculdade Arnaldo (2020), especialista em finanças pelo IBMEC-MG (2006), especialista em Ciência de dados e Big Data pela PUC Minas (2023), mestrado em Contabilidade, Fiscalidade e Finanças pela Universidade Técnica de Lisboa (2011) e doutorado em Administração pela Universidade Federal de Minas Gerais (2019). Universidade Federal dos Vales do Jequitinhonha e Mucuri – UFVJM, Diamantina/MG, Brasil. E-mail; joao.calembo@gmail.com
3. Doutor pelo Centro de Pesquisas e Pós-Graduação em Administração da Universidade Federal de Minas Gerais. Doutorado sanduíche pela Technische Universität Braunschweig, no Institut für Controlling & Unternehmensrechnung. Mestrado pelo Centro de Pesquisas e Pós-Graduação em Administração da Universidade Federal de Minas Gerais (2015), na linha de Mercadologia, Estratégia e Operações. Universidade Federal do Espírito Santo, Vitória/ES, Brasil. E-mail: bruno.vilela@ufes.br
4. Graduação em Direito pela Pontifícia Universidade Católica de Minas Gerais (1998), especialização em Controle Externo (Administração Pública) pela Escola de Contas Professor Pedro Aleixo do Tribunal de Contas do Estado de Minas Gerais (2001) e mestrado em Direito Público pela Pontifícia Universidade Católica de Minas Gerais(2004). Doutorado em Direito Público Internacional pela Pontifícia Universidade Católica de Minas Gerais. Universidade Federal dos Vales do Jequitinhonha e Mucuri – UFVJM, Diamantina/MG, Brasil. E-mail: jorgefulgencio@yahoo.com.br

Abstract

The relationship between governments and citizens can be the solution to get around the environmental crisis in which the world is inserted, mainly in the reformulation of the idea that natural resources are inexhaustible and can be deliberately exploited. The article sought to analyze the predictive contributions of environmental damage to public management in combating forest fires in the APA do Alto do Mucuri from 2014 to 2018. The Random Forest algorithm was used as a predictive technique in 138 Fire Occurrence Records (ROIs) of the conservation unit. The results show 14,729.92 hectares of burned vegetation in the region, predominantly of the Semideciduous Seasonal Forest type and analysis with better predictive capacity for low and moderate environmental damage. Thus, it is concluded the need to reformulate the ROI and the adoption of the Random Forest algorithm in decision-making in fighting forest fires.

Keywords: conservation unit, forest fires, record of fire occurrence, APA Alto do Mucuri

Resumo

O relacionamento entre os governos e os cidadãos pode ser a saída para contornar a crise ambiental na qual o mundo está inserido, principalmente na reformulação da ideia de que os recursos naturais são inesgotáveis e podem ser explorados deliberadamente. Buscou-se analisar as contribuições preditivas do dano ambiental para a gestão pública no combate a incêndios florestais da APA do Alto do Mucuri, no período de 2014 a 2018. Foi utilizado o algoritmo *Random Forest* como técnica preditiva em 138 Registros de Ocorrências de Incêndios (ROI) da unidade de conservação. Os resultados evidenciaram 14.729,92 hectares de vegetação queimada na região, predominantemente do tipo Floresta Estacional Semidecidual e análise com melhor capacidade preditiva para danos ambientais do tipo baixo e moderado. Dessa forma, conclui-se pela necessidade de reformulação do ROI e a adoção do algoritmo *Random Forest* na tomada de decisões no combate a incêndios florestais.

Palavras-chave: unidade de conservação, incêndios florestais, registro de ocorrência de incêndio, APA do Alto do Mucuri

Diego da Silva Passos, João Paulo Calembo Batista Menezes, Bruno de Almeida Vilela, Jorge Fulgencio Silva Chaves

1 Introduction

Arising mainly from erroneous ideologies like inexhaustible natural resources and deliberate exploitation, the world's environmental crisis began to be debated with the emergence of environmental movements. The new vision highlighted the difficulty of nature's stability in the face of disordered human extraction and the need for sustainable development for society's well-being (Praes, 2012). Consequently, environmental policies focused on ecological protection grew.

In Brazil, the intensification of environmental policies occurred in the twentieth century, starting in the 1960s with regulations that promoted sustainable interaction between humans and nature, lasting until today (Peccatiello, 2011). Among them, Law no. 9985/00, known as the National System of Conservation Units (SNUC), regulates the creation and management of conservation units in Brazil, defining characteristics for each type to be preserved. Articles 8 and 14 of Law 9.985/00 establish 12 categories of environmental protection, among which the Environmental Preservation Area (APA) stands out.

The Environmental Preservation Area typology is listed in the SNUC as a sustainable use conservation unit endowed with natural wealth and responsible for the sustainable integration of humans and nature (Brasil, 2000). Based on the relevance of the APA for society, a case study was carried out in the Alto do Mucuri APA, an area located in northeastern Minas Gerais and of great importance for the socio-environmental and economic development of the region.

The Alto do Mucuri APA, as well as other conservation units, faces challenges that put its survival at risk. Forest fires are one of these problems, responsible for generating annually inestimable losses to public coffers and biodiversity (Anderson et al., 2019). According to Santos, Soares, and Batista (2006), 85,735.02 hectares of vegetation are estimated to have been lost, verified in 19,377 records of fires in environmental preservation units in Brazil between 1998 to 2002. In recent years, forest fires have been increasingly highlighted in the media, usually caused by humans, conveying the idea of impunity and setbacks in environmental policies.

One of the ways to combat forest fires is to know the profile of fire occurrences to aid prevention, control, and the public management decision-making process (Ibama, 2009). One of the tools for achieving these results is the Fire Occurrence Register (ROI). ROIs are forms filled out by firefighters or conservation unit managers containing information ranging from the identification of the forest fire to its suppression (Bontempo et al., 2011).

Based on the facts elucidated, this study proposed to analyze the predictive contributions that environmental damage can generate for public management in fighting forest fires, using the ROIs from the Alto do Mucuri APA from 2014 to 2018. For this, the research was based on the Random Forest machine learning algorithm as a technique and potential tool for administering conservation units. In addition to the results, descriptive statistics of the elements obtained from the ROIs and the determining factors for forest fire occurrence in the Alto do Mucuri APA were also addressed.

2 The SNUC and its importance in the environmental context

Throughout human history, the loss of fauna and flora can be seen as urbanization and industrialization have expanded worldwide. According to Godoy and Leuzinger (2015), many scientists advocate and report the need for creating legally protected areas to contain this environmental depredation, as is the case of conservation units in Brazil. The country has

abundant natural wealth and beauty, with one of the world's largest biodiversities, which is still growing as discoveries are made.

> The country has 56 thousand species of plants, with 22% of the total of 250 thousand plants existing on the entire planet. The animal wealth consists of 524 species of mammals, 70 species of birds of the order Psittaciformes (macaws, parrots, and parakeets), more than three thousand species of freshwater fish, and something between 10 and 15 million species of insects (the vast majority have not yet been described by science) (Cunha, Oliveira, & Barbalho, 2017, p. 2).

An important instrument to consolidate conservation unit regulation in Brazil is the Brazilian Forest Code, which had three enactments: 1934, 1965, and 2012 (currently in effect). According to Medeiros (2006), the 1934 Forest Code was the first environmental instrument that clearly defined typologies of areas with preservation needs, classified as protective, remnant, model, and yield.

> Protective forests aim to conserve water systems, avoid erosion, fix dunes, assure public health, protect sites, shelter fauna specimens, or protect, from a military point of view, the borders.
> Remnant forests were those declared by the Federal, State, or Municipal Public Authorities, seeking to create parks, including small forests for public use and biological or aesthetic protection reserves.
> The model and yield forests could be subject to natural resource management aimed at timber extraction. These included the National Forests (Medeiros, 2006, p. 10).

Under strong criticism, mainly related to favoring the timber market and inefficiency of national regulation, the Forest Code of 1934 was replaced by the Forest Code of 1965 (RORIZ; FEARNSIDE, 2015). This, in turn, established the typologies of National Park and National Forest (MEDEIROS, 2006), created the concept of Permanent Preservation Area (APP), and strengthened the regulation of the Legal Reserve.

> Law 4,771 established two crucial mechanisms for environmental protection: a) the Permanent Preservation Area (APP), intended to protect soil and water, whose use is limited and depends on government authorization; b) the Legal Reserve, as a percentage of the property that must be covered by natural vegetation and that can be explored with sustainable forest management (Sauer & França, 2012, p. 287).

During the validity of the 1965 Forest Code, other critical legal devices were instituted, such as the National Environmental Policy, Law no. 6.938/81, aiming at the preservation, improvement, and environmental recovery, assuring conditions for sustainable and economic development, national security, and protection of the human life and dignity (Brasil, 1981); and the 1988 Federal Constitution, which emphasized, in its article 225, that:

> Everyone has the right to an ecologically balanced environment, an asset for common use by the people and essential to a healthy quality of life, and it is the duty of the public authorities and the community to defend and preserve it for present and future generations (Brasil, 1988, art. 225).

During the same period, another important instrument for environmental preservation was sanctioned. Seeking to conserve Brazil's natural resources, Law No. 9.985 was instituted on July 18, 2000, known as the National System of Nature Conservation Units (SNUC), aiming to preserve biological diversity, protect endangered species, conserve and restore the variety of

Diego da Silva Passos, João Paulo Calembo Batista Menezes, Bruno de Almeida Vilela, Jorge Fulgencio Silva Chaves

natural ecosystems, promote sustainable development, recover natural resources, and enhance biodiversity's socioeconomic value (Brasil, 2000).

One novelty that SNUC brought with it was the division of each conservation unit's characteristics into two large blocks, being:

> a) Full Protection - allows only the indirect use of natural resources, except in the cases prescribed by law, classified as Ecological Station, Biological Reserve, National Park, National Monument, and Wildlife Refuge;
> b) Sustainable Use - related to the balance between the sustainable use of part of the natural resources and the conservation of the environment's biodiversity, comprising Environmental Protection Area, Area of Relevant Ecological Interest, National Forest, Extractive Reserve, Fauna Reserve, Sustainable Development Reserve, and Private Natural Heritage Reserve (Brasil, 2000).

The SNUC is coordinated by the Ministry of Environment (MMA) and executed by the Chico Mendes Institute for Biodiversity Conservation (ICMBio) and the Brazilian Institute of Environment and Renewable Natural Resources (Ibama). Its management is based on registered conservation units at the municipal, state, and national levels (Brasil, 2000).

According to Leite, Araújo, and Duarte (2018), for better control and transparency in conservation unit management, the MMA created the National Registry of Conservation Units (CNUC), which gathers data on conservation units registered in Brazil, including the implementation of government actions in the locations. This data is made broadly available to interested citizens in the federal government's virtual environment.

The following section addresses the Environmental Protection Area, which represents approximately 51.86% of Brazil's registered units (MMA, 2019). At the core of the research, data from the Alto do Mucuri APA were used to deepen the knowledge about the object of study.
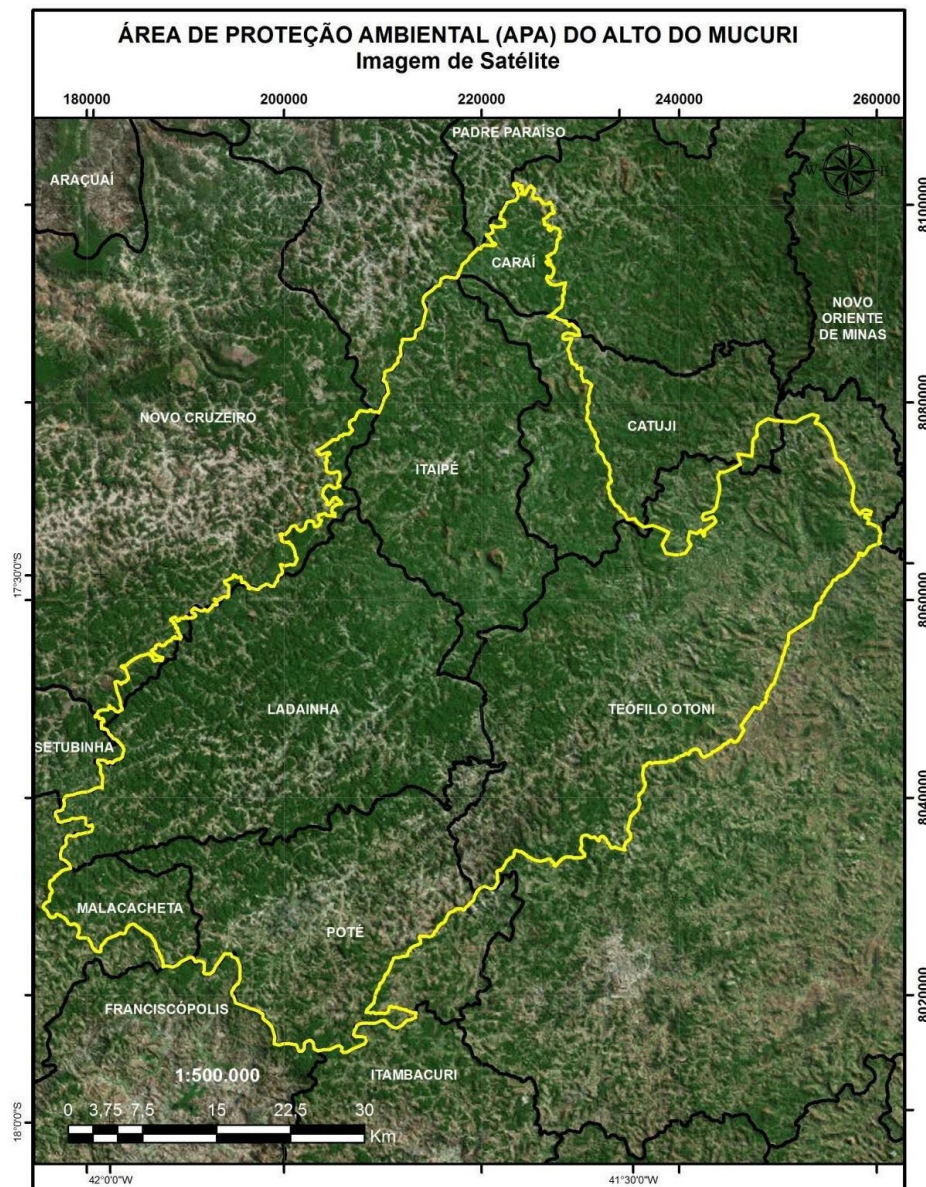
3 The Alto do Mucuri APA

The Alto do Mucuri APA was created by Decree No. 45,877 on December 30, 2011, and is classified as a conservation unit of the Sustainable Use type. Located in the Mucuri Valley, the northeastern region of Minas Gerais, it covers an area corresponding to 325,148.8883 ha (Figure 1). Its territory comprises eight municipalities, including Caraí, Catuji, Novo Cruzeiro, Malacacheta, Poté, Teófilo Otoni, Itaipé, and Ladainha, the last two being fully inserted in the unit's boundaries (Minas Gerais, 2011).

The region is characterized by a large concentration of the Atlantic Forest biome, especially the Seasonal Semideciduous Forest, a vegetation known for foliage loss during droughts and a large amount of nutrients deposited in the soil (Pereira, 2016). It is primarily responsible for the water recharge of the Mucuri river basin, with an extension of approximately 14,000 km² in Minas Gerais, and contributes to the regional water supply (Felippe, Bueno, & Costa, 2009). Regarding the socioeconomic aspect, the Alto do Mucuri APA's main activity is farming and ranching on rural properties, a relevant source for small family farmers' subsistence (Pereira, 2016). According to Fahel, Leite, and Teles (2014), the Jequitinhonha/Mucuri region is considered one of the poorest in Minas Gerais, highlighting the lack of public policies aimed at the local population.

Figure 1 – Localization of the Alto do Mucuri APA

Source: Pereira (2016).

With the creation of the Alto do Mucuri APA, its management and implementation were directed to the State Forestry Institute (IEF), responsible for assisting objectives such as the protection and recovery of surface and underground water resources in the region; the protection of soil, flora, and fauna; the recovery of degraded areas; the connectivity of forest fragments; and the promotion of economic activities compatible with environmental sustainability (Minas Gerais, 2011).

For Pereira (2018), Alto do Mucuri APA's Management Plan was vital to achieve its goals and fortify, which began in August 2016 and was finalized in September 2017. The Management Plan is a document that defines the rules of the conservation unit, such as how to administer the natural resources and its structure. The Alto do Mucuri APA's Management Plan was built in a participatory manner, with the participation of several civil society agents, municipal government representatives, non-governmental organizations, unions, and associations, among others (Pereira, 2018).

Despite the natural riches in the Alto do Mucuri APA, some challenges hinder sustainable development in the region. According to Pereira (2018), illegal environmental interventions, ownership and distribution of private properties, inspection process problems, and the influence of local culture and tradition, among others, can be identified in the area. These challenges suggest the need for the population's more active participation in maintaining the conservation unit, which involves overcoming cultural and communication obstacles, raising environmental awareness, and reconciling conflicts of interest regarding land ownership in the conservation unit's circumscription.

Setting fires in the Alto do Mucuri APA is an illegal environmental intervention. According to its Management Plan, one of the conservation unit's strategic objectives is reducing fires generated in its region. To achieve this, some objectives can be listed, including daily environmental monitoring in critical fire periods, providing guidance to the local population on controlled burning, promoting environmental education and awareness actions for residents to prevent fires, creating fire breaks at vulnerable points, and preparing instructional signs and educational material (IEF, 2017).

4 Forest Fire Indexing Systems and the Fire Occurrence Register

Forest fires are an environmental disaster in several countries, and their consequences affect not only nature but also the economy and local human life. In this context, some countries stand out for the proportion and frequency of environmental fires, for example, Australia and the USA:

> The Australian bushfires (September 2019 to February 2020) stood out, with the deaths of 32 people and 500 million animals, over six million hectares of forest and bush burned, thousands of homes destroyed, and hundreds of thousands of people displaced. In California, USA, forest fires occurred in October and, despite their low human impact (five deaths), they caused $25 billion in economic loss [...] (Cred, 2020 apud Parizzi, 2022, p. 9).

In Europe, Portugal can be cited as one of the countries with the highest annual incidence of forest fires on the continent. According to Beighley and Hyde (2018), climate studies estimate that forest fires in Portugal and other southern European countries will worsen in the coming years with increasing temperatures and reduced precipitation.

Some indexers are used for forest fire intervention in these countries, which guide the fire profiles. In the U.S., the National Fire Danger Rating (NFDR) was updated in 1978, responsible for measuring components involving ignition, propagation, and release of energy, and meteorological calculations, among other characteristics (Macedo et al., 1987 apud Rodrigues, 2009). In Australia, due to its climatological diversity, several types of forest fire indexation are used according to each region of the country, with emphasis on the drought index and the fire danger index (Vélez et al., 2000; Wagner, 1987; Macedo et al., 1987; Soares, 1972 apud Rodrigues, 2009). In Portugal, the indexing is done by the Institute of Meteorology, which initially used data on relative humidity and temperature and later added new characteristics in its analysis, such as wind speed and direction, accumulated precipitation, and soil occupation, among others (Igeo, 2008; Viegas et al., 1999; Macedo et al., 1987 apud Rodrigues, 2009).

In Brazil, the frequency of forest fires has increased, especially in the Amazon and Cerrado biomes between 2003 and 2017 and in the Pantanal biome in 2020 (Jesus et al., 2020). According to data from the Burning Program of the National Institute for Space Research (INPE), 22,116 forest fire outbreaks were verified in the Pantanal region in the year 2020, a growth of approximately 120.61% compared to 2019 (10,025 occurrences) (INPE, 2022). The

INPE Burning Program is an essential tool to understand and combat forest fires as it daily publishes satellite images detecting vegetation burning outbreaks in Brazil, among other data involving the theme (INPE, 2020).

According to Lima et al. (2018), it is first essential to know the forest fire's profile and propose prevention and combat strategies to fight them. The author mentions the Record of Fire Occurrence (ROI) as a mechanism to achieve these objectives because it can obtain a variety of information during an occurrence, such as "[...] the location of the fire, geographic data of the terrain, meteorological data during the event, the polygon of the burned area, combat data, origin, cause, and damage" (Lima et al., 2018, p. 114).

The ROIs are filled out by the individuals involved in fighting forest fires, usually the conservation unit manager or the firefighter working in the event. To avoid information loss, it is recommended that the ROI be prepared as soon as possible, which includes data collected at the beginning of the forest fire. Therefore, the ROI's quality depends on careful filling to avoid incomplete data and lack of clarity in the document (Bontempo et al., 2011).

According to Bontempo et al. (2011), the use of ROIs in protected areas has been supported by public policies and has improved over time, adjusting to the reality of fire incidents. According to IBAMA (2010), adopting ROIs in conservation units has shown positive results in reducing forest fires, although they are still a major environmental problem in Brazil.

5 Methodological Procedures

The data collected from the ROIs of the Alto do Mucuri APA was carried out through a request sent to the State Forest Institute through the Access to Information of the Federal Government website, protocol no. 02100000001201970. A total of 138 ROIs were selected between 2014 to 2018, with the following occurrence distribution: 45 (2014), 36 (2015), 13 (2016), 7 (2017), and 37 (2018). The period starts with the tool's adoption in the conservation unit's activities and finishes with the available data.

As a complement, climate data and data on atmospheric pollutants were collected from the Integrated Environmental Health Information System (SISAM). Natural forest cover and regional agriculture were obtained through the Annual Mapping Project of Land Cover and Land Use in Brazil (MapBiomas). These data were made broadly available to the public and were retrieved from the respective websites.

The data analysis technique used in the study was the Random Forest or Random Decision Forest algorithm. Considered a machine learning technique, it uses Decision Trees to make event predictions. In this work, the fire occurrence was predicted to identify variables that determine environmental damage. The R Studio software, version 1.3.1073, which uses the R language to generate statistical calculations and graphs, was used to apply the technique.

5.1 Random Forest Algorithm

According to Breiman (2001), responsible for disseminating the technique, the Random Forest can quickly generate robust results with good performance compared to other predictive algorithms. For Oshiro (2013, p. 14-15), Random Forest is configured as a method that:

> [...] builds many decision trees that will be used to rank a new example by majority vote. Each decision tree uses a subset of attributes randomly selected from the original set containing all attributes.

Diego da Silva Passos, João Paulo Calembo Batista Menezes, Bruno de Almeida Vilela, Jorge Fulgencio Silva Chaves

To apply the algorithm, 20 variables were selected and treated, with one dependent variable called environmental damage and 19 independent variables (nine obtained from SISAM, eight obtained from ROI, and two obtained from MapBiomas).

Once the data treatment was completed, the sample was applied to two bases, 70% for training (97 observations) and 30% for testing (41 observations). Thus, the algorithm was first applied to the training observations and next to the testing to identify the confidence level and confirm the results found in the training phase. In other words, the training of the sample identified the pattern of the analyzed data, creating a calculation model later used to verify its reliability with tests.

The sample was further used for the comparison of results based on two application models of the Random Forest algorithm, being:

    a) Model I - random analysis, i.e., both the training and the test could present data at any date from 2014 to 2018;

    b) Model II - chronological analysis, i.e., the training presented variables between 05/19/2014 to 09/07/2017 and the test between 09/08/2017 to 10/15/2018.

The support for the results of Models I and II was based on the evaluation of the confusion matrix, predictive performance indicators, and indexes for assessing variable importance.

5.2 Environmental Damage - Dependent Variable

The environmental damage is identified as the quotient between the attribute Burnt Area (AQ), which is the amount of vegetation burned in hectares, and Combat Time (TC), indicated by the total number of minutes used per firefighter to extinguish the fire. The data were obtained from the ROIs and informed that the higher the ratio between AQ and TC, the lower the firefighting efficiency and, consequently, the greater the environmental damage.

Combining the two attributes aims to provide an alternative for measuring the efficiency of forest fire fighting. Therefore, Lima and Soares (1992) and Santos, Soares, and Batista (2006) support the formulation of the distribution classes presented in Table 1.

Table 1 - Environmental damage classes of the sample forest fires

Class	Intensity	AQ/TC (hectares/minute)
I	Low	$Y \leq 0.0015$
II	Moderate	$0.0015 < Y \leq 0.0333$
III	High	$0.0333 < Y \leq 0.0833$
IV	Extreme	$0.0383 < Y$

Source: Research data (2021).

The intensities of 30.43% - low (42), 48.55% - moderate (67), 12.32% - high (17), and 8.70% - extreme (12) were found in the ROIs analyzed after identifying the classes of environmental damage.

5.3 Independent Variables

The 19 independent variables describe the conditions of the forest fire on the date of its occurrence and help to understand the impact factors for the performance of each event, such as the indication of climatic data and the characterization of the fire location. They were obtained due to the reliability of the historical data and website accessibility. Table 2 presents the origin of each of them.

Table 2 – Independent Sample Variables

Origin	Variables	Description
ROI	Municipality	Municipality affected by forest fire
	Location	Type of locality affected by forest fire
	Cause	Identified cause for the forest fire occurrence
	Identification	Informs whether the owner of the site affected by the forest fire has been identified
	Vegetation	Type of vegetation affected by forest fire
	Month	Month of forest fire occurrence
	Detection	Type of forest fire detection
	Detection time	Time when the forest fire was detected
MapBiomas	Agriculture and Livestock Coverage	Territorial portion identified as agriculture and cattle ranching (in hectares)
	Natural forest Coverage	Land area identified as natural forest formation (in hectares)
SISAM	CO Concentration	Carbon monoxide concentration (part per billion - ppb)
	NO_2 Concentration	Nitrogen dioxide concentration (part per billion - ppb)
	O_3 Concentration	Ozone concentration (part per billion - ppb)
	$PM_{2,5}$ Concentration	The concentration of fine inhaled particles (microgram per cubic meter - µgm3)
	SO_2 Concentration	Sulfur dioxide concentration (microgram per cubic meter - µgm3)
	Temperature	Air temperature (in degree Celsius)
	Humidity	Relative humidity (in percent)
	Wind Direction	Wind direction (in degrees)
	Wind speed	Wind speed (in meters per second - m/s)

Source: Research data (2021).

The dependent and independent variables are essential for obtaining concise and reliable results. These elements have some connection with the forest fire (direct or indirect), either from climatic and topographic conditions or from the characteristics found in the fire occurrence.

6 Results Analysis

This section presents the results obtained after descriptive statistical analysis of the ROIs and application of the Random Forest algorithm to the predictive models.

6.1 Condensed Descriptive Statistics of the ROIs

Diego da Silva Passos, João Paulo Calembo Batista Menezes, Bruno de Almeida Vilela, Jorge Fulgencio Silva Chaves

The descriptive statistics of the variables used in the methodological application of the study allow us to understand the characteristics of the Alto do Mucuri APA and, consequently, to trace the profile of forest fire occurrences in this conservation unit. Initially, this fact can be evidenced by the attributes of environmental damage. A total of 14,729.92 hectares of vegetation burned was verified, equivalent to 20,630.14 soccer fields (0.714 ha/field). This burned area represented 4.53% of the territory of the Alto do Mucuri APA. Regarding Combat Time, 9,864.61 hours (591,877 minutes) were employed from 2014 to 2018.

The highest amounts of fire occurrences were in 2014 (45), 2018 (37), and 2015 (36), with the municipalities of Ladainha (56.52%) and Itaipé (23.19%) holding the highest percentages of registered ROIs. About 90.58% (125) of the records analyzed occurred in the municipalities' rural areas. According to Medeiros and Fiedler (2004) and Mistry and Bizerril (2011), fires in rural areas may be related to pasture renewal, clearing for agriculture, animal hunting, illegal extractivism, wood smuggling, or use in charcoal manufacturing. This situation can be observed in Figure 2, which shows that the agricultural and livestock activity had a relevant result (28.26%), along with the unknown causes (55.07%).

Figure 2 - Fire causes in the Alto do Mucuri APA, by ROI, from 2014 to 2018
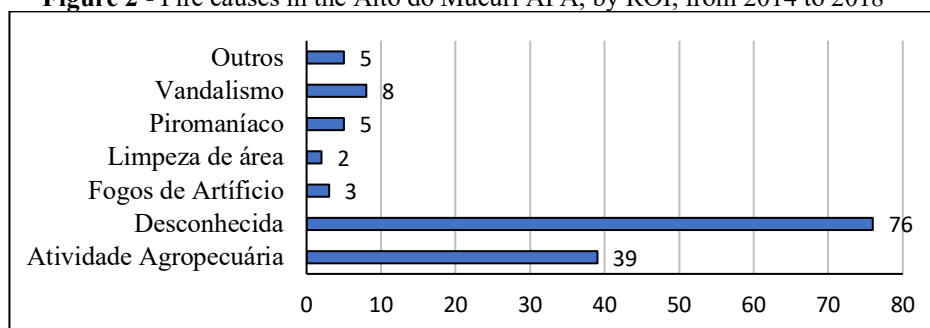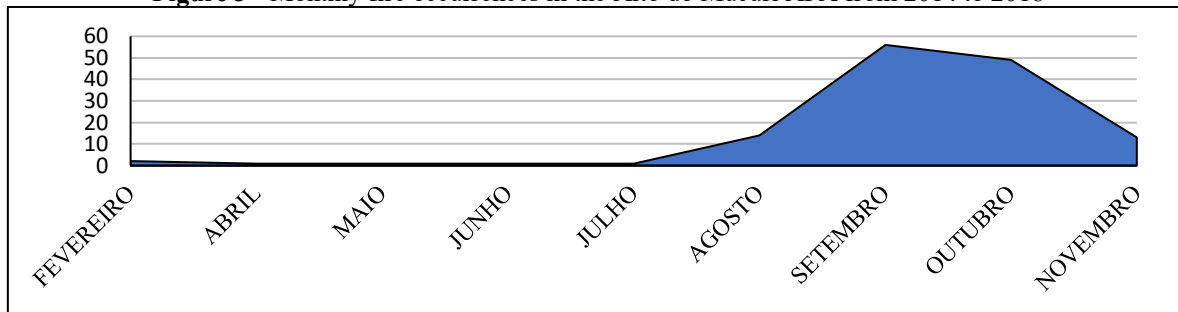

Source: Research data (2021).

Regarding the deforested vegetation, 94.93% (131) were identified as Semideciduous Seasonal Forest, a typology belonging to the Atlantic Forest biome and reported by Pereira (2016) as an abundant vegetation species in the region. During the five years analyzed, 44.98% of landowners where fires occurred were identified, which indicates a representative amount of preventive measures applied to avoid reoccurrences, such as fines, awareness, search for offenders, etc.

The monthly history of forest fire occurrences coincided with the critical period reported in scientific literature. For Santos, Souto, and Souto (2018), this critical period occurs between August and September, predominantly the dry season, which favors fire incidence. In Figure 3, it is possible to notice higher rates of forest fires between August to November, with emphasis on September and October, which reached the percentages of 40.58% (56) and 35.51% (49), respectively, concerning the occurrences of total fires.

The predictive contributions of environmental damage to the public management of conservation units: a case study of the Alto do Mucuri
APA in Minas Gerais
As contribuições preditivas do dano ambiental para gestão pública de unidades de conservação: estudo de caso da APA do Alto do Mucuri
em Minas Gerais

Figure 3 - Monthly fire occurrences in the Alto do Mucuri APA from 2014 to 2018

Source: Research data (2021).

We observed an evolution in the local population's participation in identifying forest fires throughout the period. The growth of this participation represents a positive aspect of the residents' engagement in the Alto do Mucuri APA's environmental issues. This growth is reflected in the annual rates of resident participation in identifying fires: 16.67% (2015), 15.38% (2016), 28.57% (2017), and 35.13% (2018).

6.2 Application of the Random Forest algorithm

To apply the Random Forest algorithm, it was necessary to treat the data, dividing them into two sample cuts, 70% for training (97 observations) and 30% for testing (41 observations). Once this step was concluded, some adjustments had to be made to define the number of random trees (ntree) and the number of attributes used in each decision tree (mtry) in the algorithm. Through R Studio, the best configuration obtained with the adjustments was ntree, with 300 trees and mtry, with four types of attributes. Compiling the organized data allowed for applying and interpreting the predictive models demonstrated in the following topics.

6.2.1 Indicators for interpreting Models I and II

Three classification indicators were considered for interpreting Models I (random) and II (chronological) in applying the Random Forest algorithm, as shown in Table 3.

Table 3 - Predictive performance indicators

Indicator	Equation	Description
Accuracy	$A = \dfrac{VP + VN}{VP + FN + FP + VN}$	Hit rate generated by the model.
Sensitivity	$S = \dfrac{VP}{VP + FN}$	Probability of detecting a characteristic that actually exists in the analyzed object.
Specificity	$E = \dfrac{VN}{FP + VN}$	Probability of detecting the absence of a characteristic in an object that does not really have a certain characteristic

Source: Adapted from Barbosa (2017) and Barros (2019).

The indicators are extracted from the confusion matrix, this being "[...] represented by the cross-tabulation of observed and predicted classes for observations in the test data set [...]" (SANTOS, 2018, p.30). A binary confusion matrix is depicted in Table 4.

Diego da Silva Passos, João Paulo Calembo Batista Menezes, Bruno de Almeida Vilela, Jorge Fulgencio Silva Chaves

Table 4 - Confusion matrix for the analysis of binary classification problems

		Prediction	
		Yes	No
Situação real	Sim	True positive (VP)	False positive (FP)
	Não	False negative (FN)	True negative (VN)

Source: Adapted from Santos (2018).

For the analysis presented, the situations considered as true positives (VP) and true negatives (VN) are desirable, i.e., the predictive hits generated in the model. In the following topics, these concepts are developed to understand the results.

6.2.2 Model I – Random Random Forest

The results for Model I are shown in Table 5, based on a sample cut-off of 70% for training and 30% for testing. The shaded values of the confusion matrix indicate the predictive hits of the model. In the training sample, a total of 49 (50.51%) classifications were identified as correct, of which 43 (82.7%) predicted moderate environmental damage (relative to the other values predicted as moderate).

Table 5 - Confusion matrix and predictive indicators - Model I

		Reference							
		Training				Test			
		Low	Moderate	High	Extreme	Low	Moderate	High	Extreme
Prediction	Low	6	21	0	0	4	0	0	1
	Moderate	5	43	3	1	11	15	7	2
	High	0	9	0	0	0	0	1	0
	Extreme	3	6	0	0	0	0	0	0

Predictive indicators				
	Low	Moderate	High	Extreme
Sensitivity	0.2667	1.0000	0.1250	0.0000
Specificity	0.9615	0.2308	1.0000	1.0000
Balanced Accuracy	0.6141	0.6154	0.5625	0.5000

Source: Research data (2021).

To validate the results, the values were trained, resulting in 20 correct classifications. Regarding these classifications, 48.78% of the values were accurate, with a hit rate of 80% for predicting low environmental damage and 44.12% for predicting moderate environmental damage. The training was better for predicting low environmental damage, and no prediction was considered correct for extreme damage.

Regarding the predictive indicators, the results indicate that their sensitivity is satisfactory for moderate damage (100%), and their specificity is low (96.15%), high (100%), and extreme (100%) damage. The high levels of specificity are associated with the high assertiveness index regarding variables considered absent in the model. The balanced accuracy (average of the sensitivity and specificity results) indicates the low and moderate damage types as the best predictive performances.

6.2.3 Model II – Chronological Random Forest

The results of Model II can be considered the closest to reality since they are temporal; that is, the algorithm performs the calculation as the events are registered. The results can be seen in Table 6.

Table 6 - Confusion matrix and predictive indicators - Model II

		Training				Test			
		Low	Moderate	High	Extreme	Low	Moderate	High	Extreme
Prediction	Low	9	15	0	1	11	9	0	1
	Moderate	9	34	5	1	6	9	3	2
	High	0	12	2	0	0	0	0	0
	Extreme	3	6	0	0	0	0	0	0

Predictive indicators				
	Low	Moderate	High	Extreme
Sensitivity	0.6471	0.5000	0.0000	0.0000
Specificity	0.5833	0.5217	1.0000	1.0000
Balanced Accuracy	0.6152	0.5109	0.5000	0.5000

Source: Research data (2021).

The confusion matrix obtained 45 (46.49%) correct predictions during the training phase. The highest percentage of accurate predictions was for moderate environmental damage (69.39%). In the testing phase, the model generated 20 correct predictions, which resulted in an accuracy rate of 48.78%, also found in model I. Of these accurate predictions, 52.38% were valid. Of these, 52.38% were predicted for low damage and 45% for moderate damage (percentages based on the total predicted values of each class).
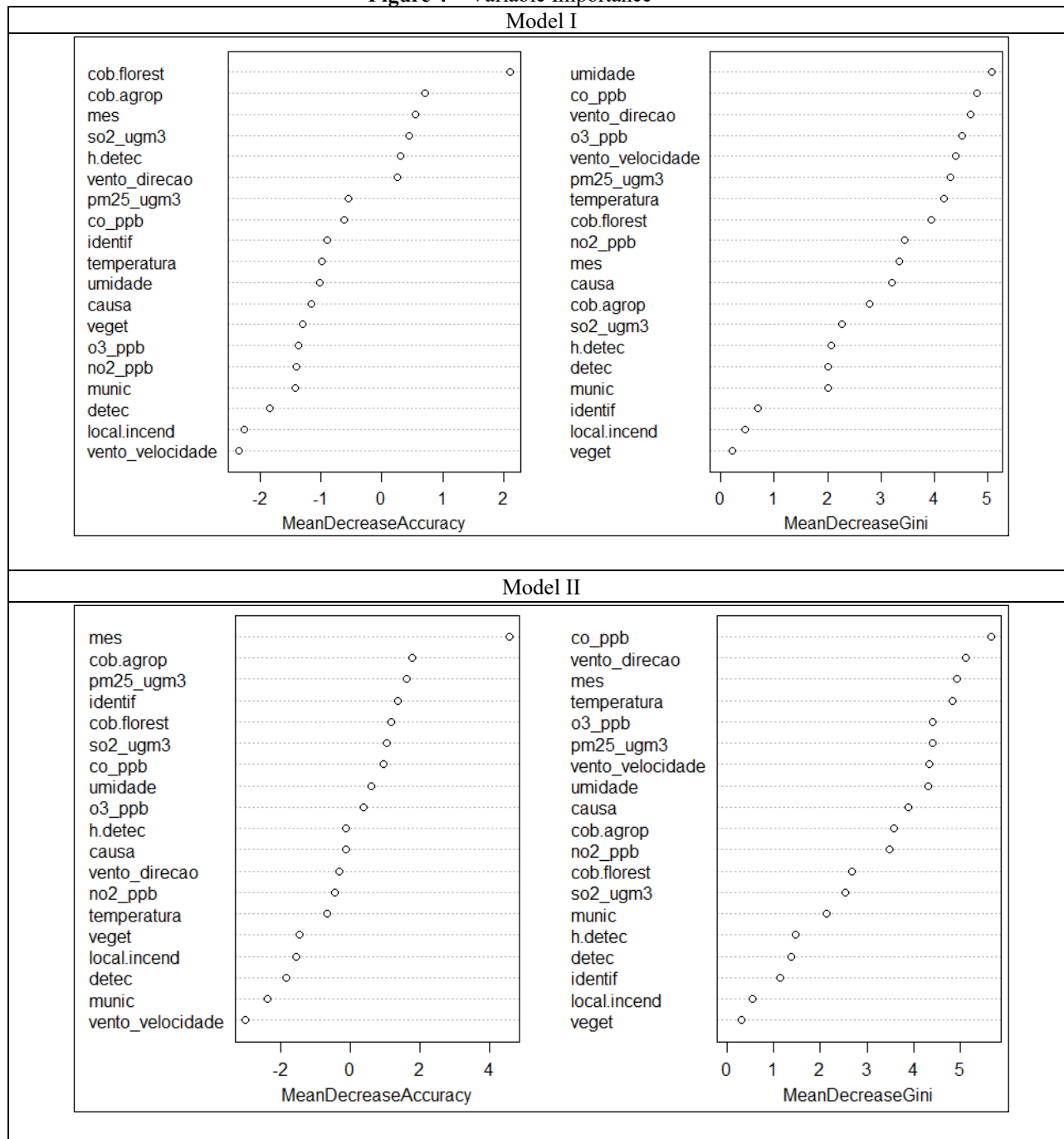
The prediction robustness indicators revealed better-balanced accuracy results for the low (61.52%) and moderate (51.09%) damage types, i.e., the referred environmental damage classes presented higher predictive strengths in relation to the others. Notably, the sensitivity generated results at 0% for the high and extreme damage types; therefore, the predicted combination may have influenced the high specificity indexes for the same classes. This situation demonstrates a specificity with a correct prediction of 100% for the absence of class values.

6.3 Variable Importance

The analysis sought to relate the variables with the greatest influences for predicting environmental damage. Two parameters were used for performance evaluation: the mean decrease accuracy (MDA) and the mean decrease Gini (MDG) (Figure 4). The MDA is defined as the loss of accuracy when the values of a given variable undergo category changes. The MDG uses the Gini impurity criterion to prepare random decision trees in Random Forest (Palczewska et al., 2014).

Diego da Silva Passos, João Paulo Calembo Batista Menezes, Bruno de Almeida Vilela, Jorge Fulgencio Silva Chaves

Figure 4 – Variable Importance

Source: Elaborated and calculated by the R *software* Studio version 1.3.1073 (2021).

It is possible to see in Figure 4 that the variables with the most significant influence on the models have markers positioned on the right side of the table. When comparing the first five variables of each parameter between Models I and II, the following determination of importance is verified:

a) MDA – Agriculture Cover (cob.agrop), Month (mes) and Natural Forest Cover (cob.florest);

b) MDG – Carbon Monoxide (co_ppb), Wind Direction (wind_direction), and Ozone Concentration (o3_ppb).

The predictive contributions of environmental damage to the public management of conservation units: a case study of the Alto do Mucuri
APA in Minas Gerais
As contribuições preditivas do dano ambiental para gestão pública de unidades de conservação: estudo de caso da APA do Alto do Mucuri
em Minas Gerais

The variables mentioned above have the greatest influence among the others in predicting the classes of environmental damage. Thus, the result oscillations of these variables will have a more significant impact on defining a particular type of environmental damage than the other variables with lower importance levels.

6.4 Discussion of Models I and II

Models I and II converged characteristics necessary for understanding and adopting measures in the Alto do Mucuri APA management. Both models achieved 48.78% accuracy in predicting environmental damage in the test phase, reflecting a better predictive ability for low and moderate damage. Thus, if the Random Forest algorithm were to classify a given fire occurrence as low or moderate damage, there is a greater chance that the prediction will be correct than if it is categorized as high or extreme damage.

In practice, this means that by knowing the characteristics of a forest fire, the conservation unit manager can use the Random Forest algorithm to predict the possible extent of environmental damage and allocate the necessary resources for fighting the forest fire. In this way, the prediction will help make efficient decisions since knowing the forest fire's impact will enable the distribution of human resources, materials, equipment, and transportation with greater effectiveness and less public spending.

It is noteworthy that the Random Forest algorithm, a machine learning technique, tends to generate increasingly assertive results regarding predictive power as its database is fed. Machine learning executes decisions based on previous successful experiences (Monard & Baranauskas, 2003). This means that the accuracy of Models I and II will evolve positively with the continuous addition of ROIs in the data applied by the Random Forest algorithm.

As for investigating the variable importance, it was found that the variables that stand out are climatic conditions, air pollutants, and agricultural and natural forest cover data. It is important to mention that this information is not found in the current ROI forms. Therefore, it suggests reformulating the fire occurrence record to incorporate these relevant elements into the profile.

7 Final Considerations

The Random Forest algorithm was used as a potential tool for managing the Alto do Mucuri APA to analyze the predictive contributions of environmental damage for public management in fighting forest fires. Therefore, the ROIs presented by the conservation unit between 2014 to 2018 were used as a sample base.

The forest fires in the Alto do Mucuri APA were profiled utilizing the ROIs. 14,729.92 hectares of burned vegetation were registered, mainly of the Seasonal Semideciduous Forest type (Atlantic Forest biome). Ladainha (56.52%) and Itaipé (23.19%) were the municipalities with the highest fire incidence.

During the five years studied, the critical forest fire period was observed between August and November. The leading causes reported were: unknown (55.07%) and agricultural activity (28.26%). According to Ramos, Fonseca, and Morello (2016), firefighters are usually hired to relieve the demand during the critical period.

Regarding the application of Random Forest in Models I (random) and II (chronological), both verified an accuracy of 48.78%, with better predictive capabilities for low and moderate environmental damage. The determining factors for fire occurrence were measured in two parameters: the MDA, which resulted in the most relevant variables

Agricultural Cover, Month, and Natural Forest Cover; and the MDG, with the variables Carbon Monoxide, Wind Direction, and Ozone Concentration, as the most important in the model.

It was observed that the most important variables for predicting environmental damage are not found in the Alto do Mucuri APA's current ROI forms. These elements are necessary to understand the incidents better; therefore, it is suggested that the ROI form be reformulated to include these characteristics in the occurrences.

Using the Random Forest algorithm in public administration will assist in defining better decisions concerning forest fires. In a hypothetical situation between two incidents, the conservation unit manager could allocate more firefighters, equipment, and transportation to the most damaging situation after the Random Forest algorithm classification. In this way, the tendency is to distribute available resources better and adopt more efficient firefighting logistics.

The results achieved are positive and tend to be improved as the database is fed. It is noteworthy that the environmental damage prediction analysis can be extended to other protected areas, demonstrating flexibility in using more or fewer variables. As a delimitation of the study, no reliable data was found about using other factors that could collaborate in the predicted models' performance, such as precipitation and temperature data.

Finally, we propose adopting the machine learning technique for the public administration of conservation units, especially for fighting forest fires. For this, it is recommended that the ROIs be continuously registered in a systematic online database with broad access for interested parties, linked to easy-to-use software for the Random Forest algorithm.

References

Anderson, L. O., et al. (2019). Conceptual model of disaster risk management and warning system associated with wildfires and public policy challenges in Brazil. *Territorium: Revista portuguesa de riscos, prevenção e segurança, 1*(26), 43–61. Link

Barbosa, J. H. de F. (2017). Early Warning System para Distress Bancário no Brasil (Tese de Doutorado). Faculdade de Economia, Administração e Contabilidade (FACE), Universidade de Brasília. Link

Barros, C. B. (2019). Classificadores de regressão logística, Naive Bayes e Random Forest na análise do Tropismo do HIV-1 de subtipo B (Dissertação de Mestrado). COPPE, Universidade Federal do Rio de Janeiro. Link

Beighley, M., & Hyde, A. C. (2018). Portugal Wildfire Management in a New Era: Assessing Fire Risks, Resources and Reforms. *Instituto Superior de Agronomia da Universidade de Lisboa*. Link

Bontempo, G. C., et al. (2011). Registro de Ocorrência de Incêndio (ROI): evolução, desafios e recomendações. *Biodiversidade Brasileira, n.*(2), 247-263. Link

Brasil. (1988). Constituição da República Federativa do Brasil. *Diário Oficial da União*, 1, 05 out. 1988. Link

The predictive contributions of environmental damage to the public management of conservation units: a case study of the Alto do Mucuri APA in Minas Gerais

As contribuições preditivas do dano ambiental para gestão pública de unidades de conservação: estudo de caso da APA do Alto do Mucuri em Minas Gerais

Brasil. (1981). Lei nº 6.938 de 31 de agosto de 1981. Dispõe sobre a Política Nacional do Meio Ambiente, seus fins e mecanismos de formulação e aplicação, e dá outras providências. *Diário Oficial da União*, 16509, 02 set. 1981. Link

Brasil. (2000). Lei 9.985, de 18 de julho de 2000. Regulamenta o art. 225, § 1º, incisos I, II, III e VII da Constituição Federal, institui o Sistema Nacional de Unidades de Conservação da Natureza e dá outras providências. *Diário Oficial da União*, 1, 19 Jul. 2000. Link

Breiman, L. (2001). Random Forests. *Machine Learning, 45*(1), 5–32. Link

Cunha, A. L., Oliveira, A. P. C., & Barbalho, M. G. S. (2017). Unidade de Conservação Ambiental. *Anais SNCMA, 8*(1). Link

Fahel, M. C. X., Leite, G. P., & Teles, L. R. (2014). Pobreza multidimensional no estado de Minas Gerais: uma mensuração para além da renda. *Revista Brasileira de Monitoramento e Avaliação, 8*, 50-69. Link

Felippe, M. F., Bueno, J., & Costa, A. (2009). Desmatamento na bacia do Rio Mucuri (MG), Brasil no período de 1989 a 2008: uso de imagens Cbers e Landsat na espacialização dos remanescentes florestais. *Simpósio Brasileiro De Sensoriamento Remoto (SBSR), 16*, 2713–2720.

Godoy, L. R. C., & Leuzinger, M. D. (2015). O financiamento do Sistema Nacional de Unidades de Conservação no Brasil: características e tendências. *Revista de Informação Legislativa, 52*(206), 223. Link

IBAMA. (2010). Perfil dos Incêndios Florestais Acompanhados pelo IBAMA. Prevfogo. Link

IBAMA. (2009). Relatório de ocorrências de incêndios em Unidades de Conservação Federais 2005-2008. Prevfogo. Link

IEF. (2017). Plano de Manejo da Área de Proteção Ambiental do Alto do Mucuri. Encarte 2 – planejamento. Belo Horizonte, MG.

INPE. (2022). Monitoramento dos focos ativos por Estado. Link

INPE. (2020). Perguntas frequentes. Programa Queimadas. Link

Jesus, J. B., et al. (2020). Análise da incidência temporal, espacial e de tendência de fogo nos biomas e unidades de conservação do Brasil. *Ciência Florestal, 30*(1), 176-191. Link

Leite, J. P. R., Araújo, D. L. S., & Duarte, M. D. D. C. (2018). Reflexos e considerações sobre a implementação do cadastro nacional de unidades de conservação do estado do Piauí. *Sustentare, 2*(1), 20–31. Link

Lima, G. S., et al. (2018). Avaliação da eficiência de combate aos incêndios florestais em unidades de conservação brasileiras. *Floresta, 48*(1), 113–122. Link

Lima, G. S., & Soares, R. V. (1992). Avaliação de eficiência de combate aos incêndios florestais no Brasil. *Floresta, 22*(1/2), 25-38.

Medeiros, M. B. DE, & Fiedler, N. C. (2004). Incêndios florestais no parque nacional da Serra da Canastra: desafios para a conservação da biodiversidade. *Ciência Florestal, 14*(2), 157–168. Link

Medeiros, R. (2006). Evolução das tipologias e categorias de áreas protegidas no Brasil. *Ambiente & Sociedade, 9*(1), 41–64.

Minas Gerais. (2011). Decreto no 45.877, de 30 de dezembro de 2011. Fica criada como Área de Proteção Ambiental – APA – do Alto do Mucuri a área situada nos Municípios de Caraí, Catuji, Itaipé, Ladainha, Novo Cruzeiro, Malacacheta, Poté e Teófilo Otoni. Diário do Executivo. Link

Mistry, J., & Bizerril, M. (2011). Por que é importante entender as inter-relações entre pessoas, fogo e áreas protegidas? *Biodiversidade Brasileira, n.*(2), 40-49. Link

Monard, M. C., & Baranauskas, J. A. (2003). Conceitos sobre aprendizado de máquina. *Sistemas Inteligentes - Fundamentos e Aplicações, 1*(1), 32.

MMA. (2019). Consulta - Relatórios de UC. Link

Oshiro, T. M. (2013). Uma abordagem para a construção de uma única árvore a partir de uma Random Forest para classificação de bases de expressão gênica (Dissertação de Mestrado). Universidade de São Paulo. Link

Palczewska, A., et al. (2014). Interpreting random forest classification models using a feature contribution method. In: *Integration of reusable systems*, 193–218. Link

Parizzi, M. G. (2022). Panorama dos desastres climatológicos, hidrológicos, meteorológicos e geológicos durante o período de 2010 – 2019. *Derbyana, 43*. Link

Peccatiello, A. F. O. (2011). Políticas públicas ambientais no Brasil: da administração dos recursos naturais (1930) à criação do Sistema Nacional de Unidades de Conservação (2000). *Desenvolvimento e Meio Ambiente, 24*. Link

Pereira, J. M. (2016). Políticas públicas florestal e de proteção à biodiversidade em prol da APA do Alto do Mucuri (Dissertação de Mestrado). UFVJM. Link

Pereira, R. S. (2018). Avanços e desafios em áreas protegidas: um estudo da APA do Alto do Mucuri no município de Ladainha - MG (Dissertação de Mestrado em Administração Pública). UFVJM. Link

Praes, E. O. (2012). Código florestal brasileiro: evolução histórica e discussões atuais sobre o novo código florestal. *VI Colóquio Internacional "Educação e Contemporaneidade"*. Link

Ramos, R. M., Fonseca, R. L., & Morello, T. F. (2016). Unidades de conservação e proteção contra incêndios florestais: relação entre focos de calor e ações articuladas pelas

Ramos, R. M., Fonseca, R. L., & Morello, T. F. (2016). Unidades de conservação e proteção contra incêndios florestais: relação entre focos de calor e ações articuladas pelas brigadas contratadas. *Revista Biodiversidade Brasileira*, *n*(2). Link

Rodrigues, M. M. T. (2009). Integração das variáveis de natureza social na avaliação do risco de incêndio florestal na região de Trás-os-Montes e Alto Douro (Dissertação de Mestrado). Universidade de Açores, Bragança. Link

Roriz, P. A. C., & Fearnside, P. M. (2015). A construção do Código Florestal Brasileiro e as diferentes perspectivas para a proteção das florestas. *Novos Cadernos NAEA, 18*(2), 28 dez. 2015. Link

Santos, H. G. (2018). Comparação da performance de algoritmos de machine learning para a análise preditiva em saúde pública e medicina (Tese de Doutorado). Universidade de São Paulo, São Paulo. Link

Santos, J. F., Soares, R. V., & Batista, A. C. (2006). Perfil dos incêndios florestais no Brasil em áreas protegidas no período de 1998 a 2002. *Floresta, 36*(1), 93-100. Link

Santos, W. S. S., Souto, P. C., & Souto, J. S. (2018). Riscos de incêndios florestais em unidades de conservação no bioma Mata Atlântica, Paraíba, Brasil. *Agropecuária Científica no Semiárido, 14*(4). Link

Sauer, S. N., & França, F. C. (2012). Código Florestal, função socioambiental da terra e soberania alimentar. *Caderno CRH, 25*(65). Link