

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**  
**CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS AMBIENTAIS**

**SILVIA CRISTINA DE JESUS**

**APRENDIZADO DE MÁQUINA PARA A CONSERVAÇÃO DA  
BIODIVERSIDADE: adequabilidade de habitat nas Unidades de  
Conservação do estado de São Paulo**

**SÃO CARLOS -SP**  
**2023**

SILVIA CRISTINA DE JESUS

APRENDIZADO DE MÁQUINA PARA A CONSERVAÇÃO DA  
BIODIVERSIDADE: adequabilidade de habitat nas Unidades de  
Conservação do estado de São Paulo

Tese apresentada ao Programa de Pós-Graduação em  
Ciências Ambientais da Universidade Federal de São  
Carlos, como parte dos requisitos para a obtenção do título  
de Doutora em Ciências Ambientais<sup>1</sup>.

Orientadora: Profa. Dra Adriana Maria Zalla Catojo

SÃO CARLOS-SP  
2023

---

<sup>1</sup> Apoio: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

Jesus, Silvia Cristina de

Aprendizado de Máquina para a Conservação da Biodiversidade: adequabilidade de habitat nas Unidades de Conservação do estado de São Paulo / Silvia Cristina de Jesus -- 2023.  
363f.

Tese de Doutorado - Universidade Federal de São Carlos, campus São Carlos, São Carlos  
Orientador (a): Adriana Maria Zalla Catojo  
Banca Examinadora: Luciano Elsinor Lopes, Sérgio Henrique Vannucchi Leme de Mattos, Diego Furtado Silva, Simone Rodrigues de Freitas  
Bibliografia

1. Gestão da paisagem. 2. Aprendizado de Máquina. 3. Áreas protegidas. I. Jesus, Silvia Cristina de. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática (SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Ronildo Santos Prado - CRB/8 7325



**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

Centro de Ciências Biológicas e da Saúde  
Programa de Pós-Graduação em Ciências Ambientais

---

**Folha de Aprovação**

---

Defesa de Tese de Doutorado da candidata Silvia Cristina de Jesus, realizada em 27/02/2023.

**Comissão Julgadora:**

Profa. Dra. Adriana Maria Zalla Catojo (UFSCar)

Prof. Dr. Luciano Elsinor Lopes (UFSCar)

Prof. Dr. Sérgio Henrique Vannucchi Leme de Mattos (UFSCar)

Prof. Dr. Diego Furtado Silva (USP)

Profa. Dra. Simone Rodrigues de Freitas (UFABC)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Ciências Ambientais.

À minha família dedico.

## AGRADECIMENTOS

À minha mãe pelo amor, apoio e confiança.

À Prof. Dra. Adriana Catojo, pela orientação, amizade, confiança e paciência.

Ao Instituto Socioambiental pela cessão da base de dados geoespaciais de Unidades de Conservação, em especial a Cicero Cardoso Augusto.

Ao Vinicius José de Oliveira Freitas, técnico administrativo do PPGCAm, pelo apoio e amizade.

Aos docentes do PPGCAm, especialmente ao Dr. Luciano Elsinor Lopes e ao Dr. Frederico Yuri Hanai, pelo apoio, confiança e amizade.

Ao Prof. Dr. Sérgio Henrique Vannucchi Leme de Mattos pela confiança e incentivo.

Às PyLadies São Carlos pelo apoio, incentivo e amizade.

Ao João Gabriel Viana Hirasawa e à Ana Dulce Padovan Torres pela amizade e pela ajuda na codificação da etapa de filtragem espacial.

Aos docentes que ministraram as disciplinas que cursei remotamente como aluna especial: Dr. Diego Furtado Silva, Dr. Edmur Azevedo Pugliesi, Dr. Miguel Petrere Jr., Dr. Rafael Duarte Coelho dos Santos, Dr. Marcos de Oliveira Lage Ferreira, Dra. Katti Faceli e Dr. Alexandre Plastino de Carvalho.

Aos alunos do curso de graduação em Gestão e Análise Ambiental, pela amizade e pelo retorno positivo nos estágios de docência e projetos de extensão.

Aos colegas do PPGCAm pela amizade.

À Alexandra Elbakyan, pelo desenvolvimento do Sci-Hub.

A todas as mulheres que lutaram para que eu pudesse estar ocupando esse espaço.

À Savannah e ao Simon Le Bon, que foram praticamente co-autores do que eu fiz durante o home office.

Principalmente, agradeço muito a Deus pela minha vida, saúde e discernimento.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Código de Financiamento 001.

"Da mata, do seu mistério, vinha o medo para o coração dos homens. Quando eles chegaram, numa tarde, através dos atoleiros e dos rios, abrindo picadas, e se defrontaram com a floresta virgem, ficaram paralisados pelo medo. A noite vinha chegando e trazia nuvens negras com ela, chuvas pesadas de junho. Pela primeira vez o grito dos corujões foi, nesta noite, um grito agoureiro de desgraça. Ressoou com voz estranha pela mata, acordou os animais, silvaram as cobras, miaram as onças nos seus ninhos escondidos, morreram andorinhas nos galhos, os macacos fugiram. E, com a tempestade que desabou, as assombrações despertaram na mata. Em verdade teriam elas chegado com os homens, na rabada da sua comitiva, junto com os machados e as foices, ou já estariam elas habitando na mata desde o início dos tempos?"

- **Jorge Amado, Terras do Sem-Fim**

"Of course, a great deal of our onslaught on Mother Nature is not really lack of intelligence but a lack of compassion for future generations and the health of the planet: sheer selfish greed for short-term benefits to increase the wealth and power of individuals, corporations and governments. The rest is due to thoughtlessness, lack of education, and poverty. In other words, there seems to be a disconnect between our clever brain and our compassionate heart. True wisdom requires both thinking with our head and understanding with our heart."

- **Jane Goodall, The Book of Hope: A Survival Guide for Trying Times**

"Biodiversidade é a biblioteca das vidas."

- **Thomas Lovejoy**

"Science, my lad, is made up of mistakes, but they are mistakes which it is useful to make, because they lead little by little to the truth."

- **Jules Verne, A Journey to the Center of the Earth**

"Humanity today is like a waking dreamer, caught between the fantasies of sleep and the chaos of the real world. The mind seeks but cannot find the precise place and hour. We have created a Star Wars civilization, with Stone Age emotions, medieval institutions, and godlike technology. We thrash about. We are terribly confused by the mere fact of our existence, and a danger to ourselves and to the rest of life."

- **Edward O. Wilson, The Social Conquest of Earth**

"Ecologia sem luta de classes é jardinagem."

- **Chico Mendes**

JESUS, S. C. - **Aprendizado de máquina para a conservação da biodiversidade:** adequabilidade de habitat nas Unidades de Conservação do estado de São Paulo. 2023. 363 p. Tese (Doutorado em Ciências Ambientais) – Centro de Ciências Biológicas e da Saúde, Universidade Federal de São Carlos, São Carlos, 2023.

## RESUMO

Modelos de Adequabilidade de Habitat (MAH) são modelos estatísticos que relacionam a localização de espécies e variáveis ambientais que restringem a distribuição destas espécies. Podem ser desenvolvidos por algoritmos de Aprendizado de Máquina, visando descobrir padrões nos dados a partir do treinamento e teste em pontos amostrais de presença da espécie. O objetivo geral desta pesquisa é avaliar a Adequabilidade de Habitat (AdH) nas Unidades de Conservação estaduais e federais no estado de São Paulo com base nos Modelos de Adequabilidade de Habitat, como uma forma de medir o potencial dessas áreas protegidas para a manutenção de espécies de mamíferos. A hipótese desta pesquisa é que as categorias mais restritivas de UCs apresentam características ambientais mais favoráveis à ocorrência de mamíferos terrestres. Os modelos usaram como variáveis preditoras: Precipitação, Temperatura, Luzes Noturnas Artificiais, Altitude, Inclinação do Terreno e Uso e Cobertura da Terra. Os dados sobre a ocorrência de espécies de mamíferos foram extraídos do GBIF - *Global Biodiversity Information Facility*. Para gerar pontos de Pseudo-ausência, foi feita a classificação OC-SVM a partir dos pontos de Presença; a área de não-presença foi usada como pano de fundo para gerar aleatoriamente os pontos de Pseudo-ausência. Então os pontos de Presença e de Pseudo-ausência foram usados para o treinamento e teste dos modelos regressão Árvore de Decisão, Floresta Aleatória, KNN, SVR, XGBoost e LightGBM. Os resultados mostraram que os métodos de Aprendizado de Máquina baseados em árvore (Floresta Aleatória e Árvore de Decisão) tiveram melhor desempenho do que os métodos baseados em distâncias (KNN) e métodos de maximização de margens (SVR). A modelagem de processos ecológicos parece ajustar-se melhor à separação de grupos com divisões sucessivas do que à tentativa de discriminar estes grupos em um espaço de atributos. As variáveis preditoras que tiveram maior importância na modelagem foram Temperatura Mínima, Luzes Noturnas Artificiais e Precipitação (variáveis numéricas) e Mosaico de Agricultura e Pastagem e Formação Florestal (variáveis categóricas). No Estado de São Paulo, as Áreas de Proteção Ambiental (APAs), por serem uma categoria com menos restrições de uso do solo, têm AdH média similares às de áreas não protegidas por Unidades de Conservação. As demais UCs de Uso Sustentável, embora tenham algum grau de antropismo, têm boa AdH média para a mastofauna; as UCs de Proteção Integral são as unidades com melhor AdH. Os modelos apontam três principais áreas no Estado de São Paulo com alta AdH: 1) Noroeste do estado, onde as paisagens são fortemente marcadas pelas agricultura; 2) Nordeste do estado, onde a produção de cana-de-açúcar é dominante, mas onde ainda há remanescentes de Cerrado não protegidos por UCs; e 3) domínio geomorfológico da Província Costeira Sul do Estado, onde há justaposição de UCs de diferentes categorias. A ocorrência da espécie sinaliza mas não define a Adequabilidade do Habitat. Ecossistemas agrícolas (e até mesmo urbanos) podem ser usados para o trânsito de indivíduos ou populações de mamíferos que, eventualmente, podem ser observados nesses locais. Contudo, esses ecossistemas não são suficientes para suprir todos os recursos para manutenção da vida silvestre, em termos de alimentação, segurança e conforto térmico, principalmente no caso de espécies especialistas. Esse uso de paisagens agrícolas é decorrente da alta fragmentação florestal que suprimiu o habitat original da mastofauna no estado de São Paulo.

**Palavras-chave:** Modelos de Adequabilidade de Habitat; Gestão da Paisagem; Unidades de Conservação; Aprendizado de Máquina; Áreas Protegidas; Mamíferos.



JESUS, S. C. **Machine Learning for biodiversity: habitat suitability in Conservation Units in São Paulo State, Brazil.** 2023. 363 p. Thesis (Doctorate in Environmental Sciences) – Biological and Health Sciences Center, Federal University of São Carlos, São Carlos - Brazil, 2023.

## ABSTRACT

Habitat Suitability Models (HSMs) are statistical models that relate the location of species and environmental variables that restrict their distribution. They can be developed by Machine Learning algorithms, aiming to discover patterns in the data from training and testing in sample points of species occurrence. The general objective of this research is to evaluate Habitat Suitability (HS) in state and federal Conservation Units in São Paulo State based HSMs, as a way of measuring the potential of these protected areas for the maintenance of mammalian species. The hypothesis of this research is that the most restrictive categories of Conservation Units present environmental characteristics more favorable to the occurrence of terrestrial mammals. The models used as predictor variables: Precipitation, Temperature, Artificial Nightlights, Altitude, Terrain Slope and Land Use Land Cover. Data on the occurrence of mammalian species were extracted from GBIF - Global Biodiversity Information Facility. To generate Pseudo-absence points, the OC-SVM classification was performed based on the Presence points; the non-presence area was used as a background to randomly generate the Pseudo-absence points. Then the Presence and Pseudo-Absence points were used for training and testing the Decision Tree, Random Forest, KNN, SVR, XGBoost and LightGBM regression models. Results showed that tree-based Machine Learning methods (Random Forest and Decision Tree) performed better than distance-based methods (KNN) and margin-maximizing methods (SVR). The modeling of ecological processes seems to be better adjusted to the separation of groups with successive divisions than to the attempt to discriminate these groups in an attribute space. The predictor variables that were most important in the modeling were Minimum Temperature, Artificial Nightlights and Precipitation (numerical variables) and Mosaic of Agriculture and Pasture and Forest Formation (categorical variables). In São Paulo State, Environmental Protection Areas (EPAs), as they are a category with less restrictions on land use, have an mean HS similar to areas not protected by Conservation Units. The other Sustainable Use Conservation Units, although they have some degree of anthropism, have a good mean HS for mastofauna; Integral Protection Conservation Units presents the highest mean HS. The models indicate three main areas in São Paulo State with high AdH: 1) Northwest of the state, where the landscapes are strongly marked by agriculture; 2) Northeast of the state, where sugarcane production is dominant, but where there are still remnants of Cerrado not protected by Conservation Units; and 3) geomorphological domain of the Southern Coastal where there is a overlapping of Conservation Units. The occurrence of the species signals but does not define the habitat suitability. Agricultural (and even urban) ecosystems can be used for the transit of individuals or populations of mammals that, eventually, can be observed in these places. However, these ecosystems are not enough to supply all the resources for the maintenance of wild life, in terms of food, safety and thermal comfort, especially in the case of specialist species. The use of agricultural landscapes is due to the high forest fragmentation that suppressed the original habitat of mastofauna in the São Paulo state.

**Keywords:** Habitat Suitability Models; Landscape Management; Conservation Units; Machine Learning; Protected Areas; Mammals

## LISTA DE FIGURAS

Figura 1 - Linha do tempo da criação de categorias de UCs .....	30
Figura 2 - Unidades de Conservação de Proteção Integral e atividades permitidas / proibidas.....	31
Figura 3 - Unidades de Conservação de Uso Sustentável e atividades permitidas / proibidas .....	32
Figura 4 - Trecho do artigo de George Box publicado no Journal of the American Statistical Association em 1976.....	44
Figura 5 - Etapas da política ambiental que podem se beneficiar da modelagem .....	45
Figura 6 - Diagrama BAM .....	50
Figura 7 - Visão hierárquica das três principais influências na ocorrência de espécies.....	51
Figura 8 - Hierarquia clássica de Aprendizado de Máquina.....	54
Figura 9 - Espectro da Reprodutibilidade.....	58
Figura 10 - Espectro da Reprodutibilidade adaptado para projetos de GIScience.....	58
Figura 11 - Localização da área de estudo.....	59
Figura 12 - Mesorregiões do estado de São Paulo.....	60
Figura 13 - Diagrama de Sankey com os percentuais de vegetação nativa remanescente e outras classes de cobertura da terra em relação à área do estado de São Paulo .....	62
Figura 14 - Províncias zoogeográficas da América do Sul, segundo a proposição de Fittkau (1969).....	64
Figura 15 - Fluxograma de trabalho .....	67
Figura 16 - Abordagem do Modelo de Distribuição de Espécies aplicada nesta pesquisa .....	68
Figura 17- Grade de blocos do WorldClim para dados com resolução de 30s.....	76
Figura 18 - Folhas na escala 1:250.000 que cobrem o estado de São Paulo.....	83
Figura 19 - Comparação dos valores de inclinação do terreno em graus e em porcentagem .....	84
Figura 20 - Representação dos principais métodos de filtragem espacial .....	89
Figura 21 - Agrupamentos de pontos pelo algoritmo DBSCAN .....	93
Figura 22 - Centro do agrupamento (medóide x centróide).....	94
Figura 23 - Efeito da rasterização no tamanho amostral de Presença.....	95
Figura 24 - Síntese gráfica dos métodos de geração de pseudo-ausência por: (a) Zaniewski et al. (2002); (b) Senay et al. (2013); e (c) Coro et al. (2016) .....	97
Figura 25 - Classificação OC-SVM.....	98
Figura 26 - Geração do plano de fundo para os pontos de Pseudo-ausência .....	100
Figura 27 - Representação da conversão das camadas raster das variáveis ambientais em Numpy Array e formação do dado tabular para treinamento e testes dos modelos.....	101
Figura 28 - Estrutura de uma árvore de decisão .....	103
Figura 29 - Árvore de decisão para regressão.....	103
Figura 30 - Algoritmo KNN para regressão .....	104
Figura 31 - Definição do hiperplano no SVM.....	105
Figura 32 - Definição do hiperplano em modelos linear e não linear do SVR .....	106
Figura 33 - Construção da árvore XGBoost a partir de diversas iterações para minimizar os resíduos da predição .....	107
Figura 34 - Crescimento da árvore de decisão (a) pelo LightGBM; e (b) por outros algoritmos .....	107
Figura 35 - Proporção entre UCPIs e UCUSs em termos de (a) número de UCs; (b) área de UCs.....	112
Figura 36 - Mapa de árvore das proporções de área das categorias de UCs estaduais .....	115
Figura 37 - Mapa de árvore das proporções de área das categorias de UCs federais .....	116
Figura 38 - Criação de UCs estaduais e federais de 1934 a 2021 .....	120
Figura 39 - Percentual de formação natural florestal nas UCs no período de 1985 a 2022 (I).....	123
Figura 40 - Percentual de formação natural florestal nas UCs no período de 1985 a 2022 (II).....	124
Figura 41 - Percentual de formação natural florestal nas UCs no período de 1985 a 2022 (III) .....	125
Figura 42 - Percentual de formação natural florestal nas UCs no período de 1985 a 2022 (IV) .....	126
Figura 43 - Proporção das classes de cobertura da terra nas UCs em 2020 (I).....	129
Figura 44 - Proporção das classes de cobertura da terra nas UCs em 2020 (II) .....	130
Figura 45 - Proporção das classes de cobertura da terra nas UCs em 2020 (III) .....	131
Figura 46 - Proporção das classes de cobertura da terra nas UCs em 2020 (IV).....	132
Figura 47 - Percentual de coberturas antrópicas com vegetação nas UCs (I).....	135
Figura 48 - Percentual de coberturas antrópicas com vegetação nas UCs (II).....	136
Figura 49 - Percentual de coberturas antrópicas com vegetação nas UCs (II).....	137
Figura 50 - Percentual de coberturas antrópicas com vegetação nas UCs (IV).....	138
Figura 51 - Distribuição dos fragmentos de Formações Florestais no estado de São Paulo.....	140
Figura 52 - Proporção dos remanescentes de vegetação nativa no estado de São Paulo .....	141

Figura 53 - Localização de grandes remanescentes (> 15 mil ha) de Formações Florestais protegidos por UCs e TIs.....	143
Figura 54 - Localização dos fragmentos de formações nativas e inclinação do terreno na APA Corumbataí, Botucatu, Tejujá - Perímetro Corumbataí.....	145
Figura 55 - Localização dos fragmentos de formações nativas e inclinação do terreno na APA Corumbataí, Botucatu, Tejujá - Perímetro Botucatu .....	145
Figura 56 - Distribuição espacial da (a) cana de açúcar e (b) bovinos no Estado de São Paulo e (c) cartograma bivariado destas variáveis.....	146
Figura 57 - Diagrama de espalhamento para o percentual de área plantada com cana-de-açúcar .....	148
Figura 58 - Diagrama de espalhamento para o efetivo de bovinos.....	148
Figura 59 - Índice LISA para área cultivada com cana-de-açúcar.....	150
Figura 60 - Índice LISA para efetivo de bovinos .....	150
Figura 61 - Uso de agrotóxicos (toneladas) e o valor bruto da produção agrícola (VBPA) (bilhões de Reais) no estado de São Paulo .....	154
Figura 62 - Transição entre a Serra da Cantareira e a área urbana da Região Metropolitana de São Paulo .....	159
Figura 63 - Localização da Reserva da Biosfera do Cinturão Verde de São Paulo e seu zoneamento.....	161
Figura 64 - Represas Guarapiranga (1) e Billings (2) na Região Metropolitana do município de São Paulo .....	163
Figura 65 - Adensamento rodoviário na Serra do Mar .....	164
Figura 66 - Supressão de áreas de mangue no Município de São Vicente nos anos de 1996, 1998 e 2022 .....	165
Figura 67 - Variação do número de ocorrências de espécies com a aplicação de filtros de metadados .....	170
Figura 68 - Variação do número de clusters em função da variável eps do algoritmo DBSCAN (I).....	176
Figura 69 - Variação do número de clusters em função da variável eps do algoritmo DBSCAN (II).....	177
Figura 70 - Variação do número de clusters em função da variável eps do algoritmo DBSCAN (III) .....	178
Figura 71 - Variação do número de clusters em função do eps considerando todas as espécies.....	178
Figura 72 - Pontos de ocorrência antes e depois das filtragens por metadados e espacial.....	179
Figura 73 - Esquema da geração de pixels amostrais nas variáveis ambientais a partir dos pontos de ocorrência de espécies.....	180
Figura 74 - Quantidade de pixels para cobrir uma área de 250 mil km <sup>2</sup> .....	180
Figura 75 - Planos de informação das variáveis ambientais preditoras (I).....	182
Figura 76 - Planos de informação das variáveis ambientais preditoras (II).....	183
Figura 77 - Planos de informação das variáveis ambientais preditoras (IV) .....	184
Figura 78 - Planos de informação das variáveis ambientais preditoras (IV) .....	185
Figura 79 - Histograma das classes de uso da terra antes (30 metros) e depois (250 metros) da reamostragem do tamanho do pixel em toda a área de estudo .....	186
Figura 80 - Histograma de ocorrências por espécies, mostrando a quantidade de todos os pontos e dos pontos incidentes em área urbana (área hachurada) (I).....	196
Figura 81 - Histograma de ocorrências por espécies, mostrando a quantidade de todos os pontos e dos pontos incidentes em área urbana (área hachurada) (II).....	197
Figura 82 - Histograma de ocorrências por espécies, mostrando a quantidade de todos os pontos e dos pontos incidentes em área urbana (área hachurada) (III) .....	198
Figura 83 - Pontos de Presença por classe de cobertura da terra (todas as classes).....	199
Figura 84 - Pontos de Presença por classe de cobertura da terra (exceto Área Urbanizada) .....	199
Figura 85 - Estatísticas da variável Temperatura em áreas não urbanas e áreas urbanas .....	200
Figura 86 - Estatísticas da variável Precipitação em áreas não urbanas e áreas urbanas.....	201
Figura 87 - Estatísticas da variável Luzes Noturnas Artificiais em áreas não urbanas e áreas urbanas .....	201
Figura 88 - Estatísticas da variável Altitude em áreas não urbanas e áreas urbanas.....	202
Figura 89 - Estatísticas da variável Inclinação do Terreno em áreas não urbanas e áreas urbanas.....	202
Figura 90 - Correlograma das variáveis para os pontos de Presença.....	203
Figura 91 - Variação na quantidade de pontos amostrais de Presença ao longo do fluxo de processamento.....	204
Figura 92 - Classificação OC-SVM.....	206
Figura 93 - Valores médios de Temperatura Média (x10°C) dos pontos amostrais de Presença e de Pseudo-ausência .....	210
Figura 94 - Valores médios de Temperatura Mínima (x10°C) dos pontos amostrais de Presença e de Pseudo-ausência .....	210
Figura 95 - Valores médios de Temperatura Máxima (x10°C) dos pontos amostrais de Presença e de Pseudo-ausência .....	211
Figura 96 - Valores médios de Precipitação (mm) dos pontos amostrais de Presença e de Pseudo-ausência .....	211
Figura 97 - Valores médios de Luzes Artificiais Noturnas (nanoWatts/cm <sup>2</sup> /sr) dos pontos amostrais de Presença e de Pseudo-ausência.....	212
Figura 98 - Valores médios de Altitude (metros) dos pontos amostrais de Presença e de Pseudo-ausência .....	212

Figura 99 - Valores médios de Inclinação do Terreno (graus) dos pontos amostrais de Presença e de Pseudo-ausência .....	213
Figura 100 - Proporção de pontos amostrais de Presença e de Pseudo-ausência nas variáveis categóricas (Cobertura da terra) .....	214
Figura 101 - Quantidade absoluta de pontos amostrais por classe de cobertura da terra .....	215
Figura 102 - Correlograma das variáveis para as amostras de Pseudo-ausência .....	217
Figura 103 - Importância relativa das variáveis no modelo Árvore de Decisão .....	219
Figura 104 - Valores SHAP para a Árvore de Regressão .....	220
Figura 105 - Estrutura da Árvore de Decisão .....	221
Figura 106 - Distribuição dos resíduos do modelo Árvore de Decisão .....	221
Figura 107 - Importância relativa das variáveis no modelo Floresta Aleatória .....	223
Figura 108 - Valores SHAP para a Floresta Aleatória .....	224
Figura 109 - Distribuição dos resíduos do modelo Floresta Aleatória .....	225
Figura 110 - Distribuição dos resíduos do modelo KNN .....	226
Figura 111 - Valores SHAP para o modelo KNN .....	227
Figura 112 - Valores SHAP para o modelo SVR .....	229
Figura 113 - Distribuição dos resíduos do modelo SVR .....	230
Figura 114 - Importância relativa das variáveis para o modelo XGBoost .....	231
Figura 115 - Valores SHAP para o modelo LightGBM .....	232
Figura 116 - Distribuição dos resíduos do modelo XGBoost .....	233
Figura 117 - Distribuição dos resíduos do modelo LightGBM .....	234
Figura 118 - Importância relativa das variáveis para o modelo LightGBM .....	234
Figura 119 - Valores SHAP para o modelo LightGBM .....	235
Figura 120 - Diagrama de Coordenadas Paralelas da importância das variáveis nos modelos de regressão .....	236
Figura 121 - Principais variáveis para os seis modelos de regressão .....	237
Figura 122 - Mapas de Adequabilidade de Habitat .....	242
Figura 123 - Configuração do GridSearchCV para Árvore de Decisão .....	243
Figura 124 - Tempo de processamento do GridSearchCV .....	243
Figura 125 - AdH média das UCs Federais e Estaduais .....	244
Figura 126 - Histograma de AdH por esfera administrativa das UCs (I) .....	246
Figura 127 - Histograma de AdH por esfera administrativa das UCs (II) .....	247
Figura 128 - AdH média das UCs por uso .....	248
Figura 129 - Histograma de AdH por uso das UCs (I) .....	249
Figura 130 - Histograma de AdH por uso das UCs (II) .....	250
Figura 131 - Histograma de AdH por uso das UCs (III) .....	251
Figura 132 - AdH média em áreas protegidas e não protegidas .....	252
Figura 133 - AdH média das UCs por categoria .....	253
Figura 134 - Histograma de AdH em APAs .....	254
Figura 135 - Histograma de AdH em ARIEs .....	255
Figura 136 - Histograma de AdH em ESECs .....	256
Figura 137 - Histograma de AdH em FEs .....	257
Figura 138 - Histograma de AdH em FLONAs .....	258
Figura 139 - Histograma de AdH em MONAT .....	259
Figura 140 - Histograma de AdH em PARNA .....	260
Figura 141 - Histograma de AdH em PEs .....	261
Figura 142 - Histograma de AdH em RDSs .....	262
Figura 143 - Histograma de AdH em REBIOS .....	263
Figura 144 - Histograma de AdH em RESEXs .....	264
Figura 145 - AdH nas APAs Jundiá, Cabreúva e Cajamar .....	265
Figura 146 - Moda da ADH em UCs por categoria .....	265
Figura 147 - AdH na ESEC Santa Bárbara .....	267
Figura 148 - AdH na ESEC de Angatuba .....	267
Figura 149 - AdH no PES do Morro do Diabo .....	268
Figura 150 - AdH na APA Mata Iguatemi .....	270
Figura 151 - AdH na FE Avaré I e II .....	272
Figura 152 - AdH na FE Piraju .....	272
Figura 153 - AdH na APA Corumbataí, Botucatu, Tejuπά (Perímetro Corumbataí) e na APA dos Rios Piracicaba e Juqueri-Mirim .....	273
Figura 154 - AdH média nas UCs (I) .....	274
Figura 155 - AdH média nas UCs (II) .....	275

Figura 156 - AdH média nas UCs (III).....	276
Figura 157 - ADH média nas UCs (IV).....	277

## LISTA DE TABELAS

Tabela 1 - Categorias de Unidades de Conservação propostas e aprovadas no SNUC .....	29
Tabela 2 - Categorias de áreas protegidas conforme IUCN e SNUC .....	34
Tabela 3 - Classes de UCs organizadas em função das atividades permitidas .....	36
Tabela 4 - Métodos de modelagem de Adequabilidade de Habitat .....	52
Tabela 5 - Evolução da população urbana no Brasil e no estado de São Paulo .....	61
Tabela 6 - Características gerais dos dados referentes às variáveis preditoras dos MAHs.....	75
Tabela 7 - Variáveis bioclimáticas do WorldClim .....	77
Tabela 8 - Parâmetros de configuração no toolkit do MapBiomias.....	85
Tabela 9 - Parâmetros definidos na busca de registros de ocorrência indexados no GBIF .....	91
Tabela 10 - Recodificação das classes do MapBiomias .....	109
Tabela 11 - Número de UCs por categoria e percentual em relação ao número total.....	113
Tabela 12 - Área de UCs por categoria e percentual em relação à área protegida e ao estado de São Paulo .....	114
Tabela 13 - Sobreposição entre UCs estaduais de Uso Sustentável e de Proteção Integral.....	117
Tabela 14 - Sobreposição entre UCs Estaduais e Federais .....	118
Tabela 15 - Intervalos das classes de tamanho dos fragmentos de Formações Florestais .....	140
Tabela 16 - Índice de Moran e resultado do teste de hipótese .....	149
Tabela 17 - Municípios com maior área plantada com cana-de-açúcar no estado de São Paulo.....	151
Tabela 18 - Municípios com maior efetivo de bovinos no estado de São Paulo .....	152
Tabela 19 - Ocorrências por valor do atributo ‘basisOfRecord’.....	167
Tabela 20 - Ocorrências por valor do atributo ‘order’ .....	169
Tabela 21 - Resultado da filtragem por metadados .....	170
Tabela 22 - Variação por espécie na quantidade de medóides em função do eps.....	171
Tabela 23 - Espécies que compõem a amostra de pontos de Presença .....	188
Tabela 24 - Pontos de Presença de espécies em todas as classes de cobertura da terra e apenas na área urbana .....	193
Tabela 25 - Hiperparâmetros da classificação OC-SVM.....	204
Tabela 26 - Métricas de avaliação da classificação OC-SVM.....	204
Tabela 27 - Comparação das estatísticas dos pontos de Presença e de Pseudo-ausência .....	207
Tabela 28 - Amostras de Presença e de Pseudo-ausência nas UCs .....	216
Tabela 29 - Hiperparâmetros do modelo Árvore de Decisão .....	218
Tabela 30 - Importância absoluta das variáveis no modelo Árvore de Decisão .....	219
Tabela 31 - Hiperparâmetros do modelo Floresta Aleatória.....	222
Tabela 32 - Hiperparâmetros do modelo KNN.....	226
Tabela 33 - Hiperparâmetros do modelo SVR.....	228
Tabela 34 - Hiperparâmetros do modelo XGBoost .....	231
Tabela 35 - Hiperparâmetros do modelo LightGBM.....	233
Tabela 36 - Métricas de avaliação dos modelos de regressão .....	239
Tabela 37 - Estatísticas das variáveis Temperatura Mínima e Precipitação em todo o PE do Rio Peixe e nas amostras de Presença e de Pseudo-ausência.....	266
Tabela 38 - Exemplos na literatura sobre MAH.....	278

## LISTA DE ABREVIATURAS E SIGLAS

AdH	Adequabilidade de Habitat
AEDP	Análise Exploratória dos dados de Presença
ANA	Agência Nacional de Águas
APA	Área de Proteção Ambiental
API	<i>Application Programming Interface</i>
APP	Área de Preservação Permanente
ARIE	Área de Relevante Interesse Ecológico
AUC	<i>Area Under the Curve</i>
CAAP	Campo Alagado e Área Pantanosa
CDB	Convenção sobre Diversidade Biológica
CNPM	Centro Nacional de Pesquisa de Monitoramento por Satélite
CNUC	Cadastro Nacional de Unidades de Conservação
COP	Conferência das Partes da Convenção de Diversidade Biológica
DBSCAN	<i>Density Based Spatial Clustering of Application with Noise</i>
EML	<i>Ecological Metadata Language</i>
ENFA	<i>Ecological niche factor analysis</i>
EOG	<i>Earth Observation Group</i>
ESEC	Estação Ecológica
ESRI	<i>Environmental Systems Research Institute</i>
FAIR	<i>Findability, Accessibility, Interoperability, Reusability</i>
FE	Floresta Estadual
FF	<i>Favourability Function</i>
FLONA	Floresta Nacional
FTP	<i>File Transfer Protocol</i>
GAM	<i>Generalized Linear Model</i>
GADM	<i>Global Administrative Areas</i>
GARP	<i>Genetic Algorithm for Rule Set Production</i>
GBIF	<i>Global Biodiversity Information Facility</i>
GBM	<i>Gradient Boosting Machine</i>
GLM	<i>Generalized Additive Model</i>

GPS	<i>Global Positioning System</i>
HCPS	<i>Hantavirus Cardiopulmonary Syndrome</i>
HS	<i>Habitat Suitability</i>
Ibama	Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis
IBGE	Instituto Brasileiro de Geografia e Estatística
ICMBio	Instituto Chico Mendes de Conservação da Biodiversidade
ICMS	Imposto sobre Circulação de Mercadorias e Serviços
IF	Instituto Florestal
ISA	Instituto Socioambiental
ISO	<i>International Organization for Standardization</i>
IUCN	<i>International Union for Conservation of Nature</i>
LiDAR	<i>Light Detection and Ranging</i>
LISA	<i>Local Indicator of Spatial Association</i>
LPVN	Lei de Proteção da Vegetação Nativa
JBRJ	Jardim Botânico do Rio de Janeiro
JSDM	<i>Joint Species Distribution Models</i>
KNN	<i>K-Nearest Neighbors</i>
LNA	Luzes Noturnas Artificiais
MaB	<i>Man and Biosphere</i>
MAH	Modelagem de Adequabilidade de Habitat
MAP	Mosaico de Agricultura e Pastagem
MAPA	Ministério da Agricultura, Pecuária e Abastecimento
MARS	<i>Multivariate Adaptive Regression Splines</i>
MaxEnt	<i>Maximum Entropy</i>
MaxLike	<i>Maximum Likelihood</i>
MDE	Modelo de Distribuição de Espécies
MDME	Modelo de Distribuição de Multiespécies
MONAT	Monumento Natural
NASA	<i>National Aeronautics and Space Administration</i>
NIMA	<i>National Imagery and Mapping Agency</i>
NOAA	<i>National Oceanic and Atmospheric Administration</i>
OANV	Outras Áreas não Vegetadas



OBIS	<i>Ocean Biodiversity Information System</i>
OC-SVM	<i>One Class Support Vector Machine</i>
OFNF	Outras Formações não Florestais
OLP	Outras Lavouras Perenes
OTP	Outras Lavouras Temporárias
PARNA	Parque Nacional
PE	Parque Estadual
PL	Projeto de Lei
PNUC	Plano Nacional de Unidades de Conservação
RBCV	Reserva da Biosfera do Cinturão Verde de São Paulo
RDS	Reserva de Desenvolvimento Sustentável
REBIO	Reserva Biológica
RESEX	Reserva Extrativista
REST	<i>REpresentational State Transfer</i>
RF	<i>Random Forest</i>
RL	Reserva Legal
RMSP	Região Metropolitana de São Paulo
ROC	<i>Receiver Operating Characteristic</i>
RPPN	Reserva Particular do Patrimônio Natural
RVS	Refúgio de Vida Silvestre
SAR	<i>Synthetic-aperture radar</i>
SFB	Serviço Florestal Brasileiro
SHAP	<i>SHapley Additive exPlanations</i>
SIG	Sistema de Informações Geográficas
SIRGAS	Sistema de Referência Geodésico para as Américas
SNUC	Sistema Nacional de Unidades de Conservação
SRTM	<i>Shuttle Radar Topography Mission</i>
SSDM	<i>Stacked Species Distribution Models</i>
SVM	<i>Support Vector Machine</i>
SVR	<i>Support Vector Regression</i>
TDWG	<i>Taxonomic Databases Working Group</i>
TI	Terra Indígena

UC	Unidade de Conservação
UCPI	Unidade de Conservação de Proteção Integral
UCUS	Unidade de Conservação de Uso Sustentável
UNESCO	Organização das Nações Unidas para a Educação, a Ciência e a Cultura
VIIRS	<i>Visible Infrared Imaging Radiometer Suite</i>
VBPA	Valor Bruto da Produção Agrícola
WGS	<i>World Geodetic System</i>
XGBoost	<i>eXtreme Gradient Boosting</i>

## SUMÁRIO

1 INTRODUÇÃO	22
1.1. DEFINIÇÃO DO PROBLEMA	22
1.2 MOTIVAÇÃO	25
1.3 OBJETIVOS	26
2 REVISÃO BIBLIOGRÁFICA	27
2.1 UNIDADES DE CONSERVAÇÃO	27
<b>2.1.1 Breve histórico e sistema atual</b>	27
<b>2.1.2 Eficácia de Unidades de Conservação</b>	35
<b>2.1.3 Serviços Ecosistêmicos fornecidos pelas Unidades de Conservação</b>	39
2.2 MODELAGEM PARA O PLANEJAMENTO AMBIENTAL	42
2.3 SÍNTESE ECOLÓGICA	46
2.4 CIÊNCIA CIDADÃ PARA A BIODIVERSIDADE	47
2.5 MODELOS DE DISTRIBUIÇÃO DE ESPÉCIES E NICHOS ECOLÓGICO	49
2.6 MODELOS DE DISTRIBUIÇÃO DE ESPÉCIES E ADEQUABILIDADE DE HABITAT	51
2.7 APRENDIZADO DE MÁQUINA	53
2.7 <i>BIG DATA</i> E APRENDIZADO DE MÁQUINA NA MODELAGEM AMBIENTAL	55
2.8 PESQUISA REPRODUTÍVEL	56
3 METODOLOGIA	59
3.1. CARACTERIZAÇÃO DA ÁREA DE ESTUDO	59
3.2. AMBIENTE DE PROCESSAMENTO	65
3.3 MODELAGEM DE ADEQUABILIDADE DE HABITAT	67
3.4 VARIÁVEIS PREDITORAS	69
<b>3.3.1 Precipitação</b>	69
<b>3.3.2 Temperatura</b>	70
<b>3.3.3 Luzes Noturnas Artificiais</b>	71
<b>3.3.4 Altitude</b>	72
<b>3.3.5 Inclinação do terreno</b>	73
<b>3.3.6 Uso e cobertura da terra</b>	73
3.5 FONTE E CARACTERÍSTICAS DOS DADOS	74
<b>3.5.1 Variáveis preditoras</b>	74
3.5.1.1 Precipitação e temperatura	75
3.5.1.2 Luzes Noturnas Artificiais	78
3.5.1.3 Uso e cobertura da terra	78
3.5.1.4 Altitude	78
<b>3.5.2 Limites das UCs federais e estaduais</b>	79
<b>3.5.3 Limites do estado de São Paulo</b>	80
<b>3.5.4 Ocorrência de espécies</b>	80
3.6. PROCESSAMENTO DOS DADOS	82
<b>3.6.1. Geração dos planos de informação das variáveis preditoras</b>	82

3.6.1.1 Variáveis climáticas (WorldClim)	82
3.6.1.2 Luzes Noturnas Artificiais	82
3.6.1.3 Altitude	82
3.6.1.4 Inclinação do terreno	83
3.6.1.5 Uso e cobertura da terra (MapBiomas)	84
<b>3.6.2 Organização dos planos de informação das variáveis preditoras</b>	84
3.6.2.1 Variáveis numéricas	85
3.6.2.2. Variáveis categóricas	86
3.6.2.3. Combinação das variáveis preditoras	87
<b>3.6.3 Filtragem de dados de ocorrência de espécies</b>	87
3.6.3.1 Filtragem por metadados	90
3.6.3.2 Filtragem espacial	92
3.7. EXTRAÇÃO DOS VALORES DAS VARIÁVEIS PREDITORAS NOS PONTOS DE PRESENÇA DE ESPÉCIES	95
3.8 PONTOS DE PSEUDO-AUSÊNCIA	96
<b>3.8.1 One-Class Support Vector Machine para inferência da pseudo-ausência</b>	98
<b>3.8.2 Geração do plano de fundo para os pontos de Pseudo-ausência</b>	99
3.9 MODELOS DE ADEQUABILIDADE DE HABITAT BASEADOS EM ALGORITMOS DE APRENDIZADO DE MÁQUINA	100
<b>3.9.1 Árvores de Decisão e Florestas Aleatórias</b>	102
<b>3.9.2 Regressão KNN</b>	104
<b>3.9.3 Regressão de Vetor Suporte</b>	104
<b>3.9.4 XGBoost</b>	106
<b>3.9.5 LightGBM</b>	107
3.10 MAPAS DE ADEQUABILIDADE DE HABITAT	108
3.11 ESTATÍSTICAS ZONAIS	108
3.12 ANÁLISE DO USO E COBERTURA DA TERRA NAS UCS DO ESTADO DE SÃO PAULO	108
3.13 ANÁLISE DA COBERTURA DA TERRA NO ESTADO DE SÃO PAULO	110
4 RESULTADOS E DISCUSSÕES	112
4.1 USO E COBERTURA DA TERRA E SUAS IMPLICAÇÕES NA OCORRÊNCIA DE MAMÍFEROS	112
<b>4.1.1 Panorama das Unidades de Conservação no estado de São Paulo</b>	112
<b>4.1.2 Cobertura da terra nas Unidades de Conservação do estado de São Paulo</b>	120
<b>4.1.3. Cobertura da terra no estado de São Paulo</b>	139
4.1.3.1 Remanescentes de formações nativas no Estado de São Paulo	139
4.1.3.2 Agricultura e Pecuária no Estado de São Paulo	145
4.1.3.3 Urbanização	156
4.2 OCORRÊNCIA DE MAMÍFEROS NO ESTADO DE SÃO PAULO	165
<b>4.2.1 Filtragem por metadados</b>	165
<b>4.2.2 Filtragem espacial</b>	170
4.3 VARIÁVEIS PREDITORAS	180
4.4 ANÁLISE EXPLORATÓRIA DOS DADOS DOS PONTOS DE PRESENÇA	186
4.5 CLASSIFICAÇÃO OC-SVM	204
4.6 ANÁLISE EXPLORATÓRIA DOS DADOS DE PSEUDO-AUSÊNCIA	206

4.7 MODELOS DE ADEQUABILIDADE DE HABITAT	218
<b>4.7.1 Árvore de Decisão</b>	218
<b>4.7.2 Floresta Aleatória</b>	222
<b>4.7.3 KNN</b>	225
<b>4.7.4 SVR</b>	227
<b>4.7.5 XGBoost</b>	230
<b>4.7.6 LightGBM</b>	233
4.8 COMPARAÇÃO DOS MODELOS DE REGRESSÃO	235
<b>4.8.1. Importância das variáveis</b>	235
<b>4.8.2 Métricas de acurácia dos modelos</b>	238
4.9 ADEQUABILIDADE DE HABITAT NAS UNIDADES DE CONSERVAÇÃO	243
4.10 CONSIDERAÇÕES SOBRE A ABRANGÊNCIA ESPACIAL EM MAHs	278
5 CONCLUSÕES	281
REFERÊNCIAS	287
ANEXO I: Bibliotecas Python usadas nesta pesquisa	333
ANEXO II: Variáveis dos registros de ocorrência do GBIF	343
ANEXO III: Lista das Unidades de Conservação que estão na base de dados do Instituto Socioambiental e foram consideradas neste estudo	359

## 1 INTRODUÇÃO

### 1.1. DEFINIÇÃO DO PROBLEMA

As Unidades de Conservação (UC) são áreas protegidas criadas com o propósito de salvaguardar a biodiversidade, os recursos naturais e paisagísticos, sustentar territórios representativos dos habitats e comunidades biológicas, além de garantir às comunidades tradicionais o uso racional dos recursos naturais. Estas funcionalidades dependem do uso (Proteção Integral e Uso Sustentável) e, mais especificamente, da categoria da UC, que estabelece suas permissões e restrições de uso.

O Sistema Nacional de Unidades de Conservação (SNUC), Lei 9.985 de 2000<sup>2</sup>, foi o primeiro instrumento de ordenamento jurídico para as UCs, que estabelece “a obrigatoriedade de elaboração prévia de estudos técnicos”. Isto significa que existem critérios técnicos e legais para a criação de áreas protegidas para definir a localização, a dimensão e os limites mais adequados para a UC (LEME MACHADO, 2022). Condições como grau de ocupação humana, características da biota e dos elementos paisagísticos, nível de preservação, entre outros aspectos sociais, econômicos e ambientais devem ser considerados, portanto, na delimitação espacial dessas áreas protegidas.

O SNUC trata também sobre a Zona de Amortecimento, que é a área no entorno da unidade e para onde as restrições de uso se estendem, em vistas de minimizar impactos negativos sobre a unidade. As normas que versam sobre tais restrições devem fazer parte do Plano de Manejo, que é o documento técnico que rege os objetivos da unidade, seu zoneamento e normas de uso, e diretrizes para gestão.

Brodie et al. (2018) orientam que as estratégias de conservação devem garantir que as espécies mantenham papéis-chave nas interações interespecíficas e na manutenção de comunidades e ecossistemas, e que estas funções devem ser consideradas nas diretrizes para uma gestão eficaz. Os mamíferos são considerados espécies-chave, termo introduzido por Paine (1969) para aquelas espécies essenciais para manter a organização e a diversidade de suas comunidades ecológicas. Nativos ou introduzidos, controlam presas (como predadores), quando não estão no topo da cadeia trófica, regulam a população de predadores (como presas) e podem eliminar estruturas/materiais que afetam o tipo de habitat e o fluxo de energia (como modificadores do habitat) (MILLS et al., 1993; QUEIROZ et al., 2016; RUSSELL et al. 2018;

---

<sup>2</sup> Lei 9.985, de 18 de julho de 2000. Regulamenta o art. 225, § 1º, incisos I, II, III e IV da Constituição Federal, institui o Sistema Nacional de Unidades de Conservação da Natureza e dá outras providências. Disponível em <[http://www.planalto.gov.br/ccivil\\_03/leis/19985.htm](http://www.planalto.gov.br/ccivil_03/leis/19985.htm)>. Acesso em 08 ago. 2022.

SMITH et al., 2018; LACHER JR et al., 2019; BOGONI et a., 2020; VILLAR et al., 2020; VILLAR et al. 2021; BALLARINI et al., 2021). Na década de 1990, já era proposto que o manejo de espécies-chave individuais deveria ser um foco para o manejo de comunidades inteiras (MILLS et al., 1993).

A classe Mammalia também inclui espécies consideradas espécies guarda-chuva, que são aquelas mais rigorosas quanto à demanda de recursos e condições ambientais e, portanto, encapsulam os requisitos das espécies simpátricas (LAMBECK, 1997). Embora as interações entre espécies não sejam constantes no espaço e tempo, estes conceitos são convenientes para a constatação de que algumas espécies têm interações mais robustas, e facilita a comunicação entre equipes técnicas e de gestão e o público (MILLS et al., 1993). Roberge e Algelstam (2004) analisaram sistematicamente o conceito e explicam que, embora seja improvável que os requisitos de uma espécie incorporem os de todas as outras espécies, as estratégias de guarda-chuva multi-espécies podem ser mais robustas e mais apropriadas nas estratégias de conservação. A estratégia multi-espécies também é apontada como uma alternativa mais acertada por outros autores que avaliaram se os planos de conservação que consideram apenas uma espécie são suficientes para proteger as espécies co-ocorrentes (DINIZ et al., 2018; SHEN et al., 2019; VASUDEVA et al., 2022).

A abordagem de espécies guarda-chuva tem sido incorporada a Modelos de Distribuição de Espécies (MDEs) como estratégia de conservação (BOU et al., 2019; LINERO et al., 2020; SMITH et al., 2021; SUMBH; HOF, 2022). Tratam-se de modelos estatísticos que relacionam espécies e ambiente, baseados em dados de localização de espécies e variáveis ambientais que restringem a distribuição de espécies. São frequentemente usados para fazer previsões espaciais, e resultam em um mapa que indica a probabilidade de ocorrência da espécie em uma determinada área (ELITH; FRANKLIN, 2013).

Estes modelos também indicam a Adequabilidade do Habitat para uma espécie, sendo também chamados Modelos de Adequabilidade de Habitat (MAH); são relevantes para priorizar ou definir áreas protegidas, e para avaliar ameaças a essas áreas. Su et al. (2021) conceituam “habitat” como “o tipo de ambiente natural em que uma determinada espécie vive ou pode encontrar comida, abrigo, proteção e parceiros para reprodução”.

“Adequabilidade de Habitat”, por sua vez, é definida por Muhammed et al. (2022) como “a capacidade do habitat de suportar uma espécie silvestre ou aquática, relacionando variáveis ambientais à probabilidade de ocorrência de uma espécie”. No entanto, é frequente a observação de animais silvestres em ambientes antrópicos como áreas agrossilvopastoris e urbanas, que são

áreas de trânsito. Muhammed et al. (2022) acrescentam que pressões antrópicas como a expansão urbana e a agricultura são fatores chave para o declínio da adequabilidade de habitat.

Os MAHs têm sido aplicados para analisar a eficácia de áreas protegidas quanto ao seu potencial de proteção da biodiversidade (URBINA-CARDONA; LOYOLA, 2008; LI et al., 2018; SOBRAL-SOUZA, 2018; MALAKOUTIKHAH et al., 2020; SCRIVEN et al., 2020; CAMPOS-CERQUEIRA et al., 2021; CONDRO et al., 2021), considerando uma única espécie ou gênero ou pontos aleatórios na área de estudo. Além do planejamento da conservação da biodiversidade, os MAHs têm sido usados na avaliação de impactos (CLAUZEL et al., 2013; LARANJEIRO et al., 2018; TARABON et al., 2019; BAKER et al., 2020), gestão de recursos naturais (BENNETSEN et al., 2016; BITTNER et al., 2020; SÁNCHEZ et al., 2021; AI et al., 2022), restauração ecológica (GASTÓN; GARCÍA-VIÑAS, 2013; GASTÓN et al., 2014; ZELLMER et al., 2019; HU et al., 2020; BELLIS et al., 2021), avaliação de riscos e impactos de espécies invasoras (FERNANDES et al., 2019; RODRIGUEZ-REY et al., 2019; SRIVASTAVA et al., 2019; BLANCO et al., 2021; RUNQUIST et al., 2021) e efeitos do aquecimento global na biodiversidade e nos ecossistemas (SOBRAL-SOUZA, 2018; MALAKOUTIKHAH et al., 2020; CAMPOS-CERQUEIRA et al., 2021; CONDRO et al., 2021; DESJONQUÈRES et al., 2022).

Os MAHs podem ser desenvolvidos por algoritmos de Aprendizado de Máquina, que visam principalmente descobrir padrões nos dados a partir de dados de ocorrência de espécies em dimensões geográficas e temporais, usando dados rotulados binários (presença e ausência / pseudo-ausência de espécies) e variáveis ambientais (GOBEYN et al., 2019). Como explicado por Beery et al. (2021), a estatística tradicional, pelas suposições de dependência linear de covariáveis e de variáveis aleatórias e independente e identicamente distribuídas, pode não atender às relações complexas entre as espécies e seu ambiente.

Os Sistemas de Informações Geográficas (SIGs) viabilizam a modelagem espacial de espécies por meio do armazenamento das variáveis ambientais independentes e da integração dessas variáveis em um mapa que mostra a adequação de um determinado local como habitat para uma espécie específica (CORSI et al., 2000). Entretanto, a modelagem espacial existe em um contexto em que um imenso volume de dados é coletado e depende de uma infraestrutura computacional para desenvolvimento de modelos ecológicos em várias escalas espaço-temporais. A automatização de análises espaciais mais sofisticadas vai ao encontro da Ciência de Dados e *Big Data* e, como consequência, a Ciência de Dados Espaciais existe como uma disciplina com maior interatividade com tecnologias de código aberto e nuvem (HUNT et al., 2009). Uma característica importante da modelagem espacial é que, por ser baseada em



ferramentas computacionais, os estudos são potencialmente reproduzíveis, dado o mesmo código e dados de entrada e execução em uma máquina compatível (IVIE; THAIN, 2018). A Ecologia Informática (MICHENER; JONES, 2012; LIN, 2020) e a Computação Ecológica (PETROVSKII; PETROVSKAYA, 2012; WANG; REN, 2021) também são disciplinas que têm avançado sob essa perspectiva.

## 1.2 MOTIVAÇÃO

As circunstâncias no interior e no entorno das UCs não são imutáveis e a pressão da antropização no seu entorno, principalmente no caso das categorias menos restritivas como as Áreas de Proteção Ambiental (APAs), pode desviar o propósito da unidade. Ao longo do tempo as circunstâncias podem ser alteradas desde a sua criação, principalmente em relação à conversão da cobertura e uso da terra.. Isto posto, avaliar a eficácia das UCs é essencial, principalmente em territórios altamente antropizados como o estado de São Paulo, dominado pelo cultivo de cana-de-açúcar e com baixo percentual de áreas protegidas por UCs de Proteção Integral, especialmente nas porções norte e centro-oeste do estado. Há ocorrências de mamíferos de médio e grande porte em áreas não protegidas, já que estas são locais de trânsito de indivíduos. Além disso, é relevante fazer um diagnóstico da cobertura da terra em todo o estado de São Paulo para identificar locais que seriam apropriados para o estabelecimento de UCs.

Neste trabalho, os modelos valem-se de dados de Presença de várias espécies de mamíferos que são potencialmente guarda-chuvas; ou seja, parte-se da premissa de que, se o habitat é adequado para pelo menos uma das espécies consideradas, é apropriado para manter populações de outros taxa.

A modelagem integra diversos componentes que podem afetar a Adequabilidade do Habitat. A pesquisa é pautada no conceito de Reprodutibilidade, que é um dos princípios do método científico e sustenta que, se um estudo é replicado, os resultados devem ser reiterados com alto grau de confiabilidade (POPPER, 1959). A aquisição, processamento e visualização dos dados foram executados pelas linguagens de programação Python e R, e os principais detalhes para reproduzir o trabalho são documentados, incluindo a ordem em que os scripts de análise devem ser executados, seguindo o padrão *Silver* proposto por Heil et al. (2020).

### 1.3 OBJETIVOS

O objetivo geral desta pesquisa é avaliar a Adequabilidade de Habitat nas Unidades de Conservação estaduais e federais no estado de São Paulo com base nos Modelos de Adequabilidade de Habitat, como uma forma de medir o potencial das Unidades para a manutenção de espécies nativas da mamíferos de mamíferos silvestres.

Os objetivos específicos, que também orientam na avaliação da hipótese de pesquisa, são:

- Avaliar os MAHs como instrumentos orientadores de gestão e planejamento de UCs;
- Comparar métodos de Aprendizado de Máquina na modelagem de Adequabilidade de Habitat;
- Discutir sobre o uso e cobertura da terra no estado de São Paulo e suas implicações na Adequabilidade do Habitat de mamíferos;
- Quantificar a cobertura da terra nas UCs considerando sua categoria e nível de restrição.

A hipótese desta pesquisa é que as categorias mais restritivas apresentam condições ambientais mais favoráveis à ocorrência de mamíferos terrestres e sua cobertura seja predominantemente de formações nativas. A avaliação da adequabilidade de habitat seria, portanto, uma forma de orientar o planejamento de novas áreas protegidas e na adoção de estratégias de restauração da paisagem nas UCs já existentes e do seu entorno. Ou seja, mais do que inferir sobre a ocorrência de mamíferos, os modelos procuram reconhecer locais no estado de São Paulo que tenham potencial para abrigar áreas protegidas, ainda que sejam para categorias de uso pouco restritivo, e também identificar UCs que podem não ser efetivamente protetoras da biodiversidade.

## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 UNIDADES DE CONSERVAÇÃO

#### 2.1.1 Breve histórico e sistema atual

Dean (1996) relata que as áreas protegidas no estado de São Paulo foram criadas em um contexto de necessidade de conter a supressão da vegetação nativa decorrente do avanço das atividades agrícolas e industriais. Nos anos 1950 e 1960, as madeiras de lei eram usadas principalmente para a construção de residência e geração de energia elétrica, além dos volumes destinados à exportação. A expansão da rede rodoviária deu acesso às áreas de Cerrado, mais remotas porém menos densas e mais fáceis para extração da madeira. Mesmo com o alto custo de transporte, esta estratégia compensava o alto custo da plantação de eucalipto. No início da década de 1950, a Secretaria Estadual da Agricultura levantou que o Estado de São Paulo mantinha menos de 14% de cobertura florestal, o que originalmente correspondia a 81% do território estadual. Nesse período, o nacionalismo econômico impulsionou a criação de indústrias altamente poluentes como a mineração, siderurgia e eletroquímica, e as reservas sob proteção estadual e federal já eram cobiçadas para a exploração do seu espaço e recursos naturais, já que eram mais baratas de explorar do que as terras privadas que requeriam aluguel ou desapropriação (DEAN, 1996). Diante desse cenário, a gestão dos recursos naturais pelo Estado tornou-se questão nacional e as primeiras áreas protegidas foram estabelecidas (GUERRA; COELHO, 2012).

O primeiro Parque Estadual de São Paulo, o PE de Campos do Jordão, foi criado em 1941 na Serra da Mantiqueira, à época de expectativa de crescimento econômico regional no eixo São Paulo Rio de Janeiro, e também da emergência de manifestações de apelo conservacionista. As serrarias eram abundantes na região a pretexto do potencial madeireiro, com destaque para a *Araucaria angustifolia*. O desmatamento e a demanda por madeira justificaram a restauração e reflorestamento da região, contexto no qual foi criado o parque. Em conformidade com o Código Florestal de 1934, foi criado para o manejo da fauna e flora, além do reflorestamento para proteção dos terrenos com alta inclinação, mas sua concepção teve finalidade sobretudo produtiva (DEAN, 1996; FUNDAÇÃO FLORESTAL, 2015).

Também em 1941, foi decretada a maior das reservas na região do Pontal do Paranapanema, no extremo oeste do estado, com cerca de 3 mil km<sup>2</sup>. Originalmente, o Pontal foi proposto como reserva por Alberto Loefgren em 1905. À época do decreto de criação, o

território teve sua propriedade abdicada pelo governo estadual e foi invadida por grileiros que contrabandeavam madeira. Em 1952, a imprensa denunciou funcionários públicos envolvidos em ilegalidades e reclamou a retirada dos invasores das terras protegidas, com o apoio de associações privadas e conselhos municipais de partidos da oposição. Alegando que o poder público não era eficiente na proteção das reservas, a Assembleia Legislativa do estado criou um projeto de lei que propôs o arrendamento de toda a região do Pontal a particulares com a condição de que fosse conservada a cobertura florestal de um quarto da área arrendada. As florestas foram então convertidas em pastagens, já que as condições pedológicas e climáticas eram impróprias para as lavouras. Uma porção da reserva original, o Morro do Diabo, persistiu, com pouco mais da metade da sua área intacta, apesar da invasão constante por grileiros durante a década de 1970 e do parco orçamento para a sua manutenção (DEAN, 1996). Outras reservas sob administração estadual foram decretadas nos anos 1950 e 1960, incluindo o PES Turístico do Alto do Ribeira, em 1958 (35.902 km<sup>2</sup>) e o PES da Ilha do Cardoso, em 1962 (13.155 km<sup>2</sup>) (CNUC, 2023) .

No Brasil, as áreas onde já havia concentração populacional e de atividades humanas foram priorizadas para a conservação de ecossistemas remanescentes, com o intuito de proteger esses espaços de impactos já existentes. Ou seja, as UCs já foram criadas em circunstâncias de conflitos territoriais e de acesso a recursos, o que repercutiu na estrutura da sua gestão e na necessidade de um sistema nacional de áreas protegidas. No fim dos anos 1980 e na década seguinte, a criação e a gestão das UCs se consolidaram; com a redemocratização do país, as questões ambientais passaram a ser mais amplamente discutidas, já que durante o regime militar, a política ambiental era coordenada pelo Estado. A criação do Ministério do Meio Ambiente em 1985, a Constituição Federal de 1988, que trouxe um capítulo específico sobre meio ambiente, e a criação do Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais (Ibama) em 1989, reforçaram a política e a estrutura de gestão da conservação no país. O Ibama passou a ser o gestor e fiscalizador das unidades de conservação sob administração federal, e estruturou a implantação de unidades em todas as esferas administrativas (GUERRA; COELHO, 2012). Em 1992, o Ibama encaminhou ao Congresso Nacional, um estudo elaborado pela Funatura, que atualizava o Plano Nacional de Unidades de Conservação (PNUC), que foi a primeira tentativa de sistematizar as categorias de UCs. O PNUC foi conduzido em duas fases distintas: a primeira, em 1979, propôs 16 categorias de unidades de conservação, organizadas em cinco grupos; e a segunda, publicada em 1982, sem alteração nas categorias mas com maior embasamento técnico-científico. O estudo foi encaminhado ao Congresso mediante o

Anteprojeto de Lei nº 2.892/1992 e manteve-se em discussão durante oito anos, dadas as discordâncias quanto às populações tradicionais, às desapropriações e indenizações, e à participação popular na criação e na gestão da UC (PUREZA et al., 2015). O resultado dos debates sobre o anteprojeto foi a promulgação do Sistema Nacional de Unidades de Conservação (SNUC), pela Lei 9.985/2000<sup>3</sup>. A Tabela 1 mostra as categorias propostas no Anteprojeto de Lei nº 2.892/1992, as alterações durante a tramitação e as categorias aprovadas no SNUC. A Figura 1 mostra a linha do tempo com o ano de criação das categorias previstas no SNUC.

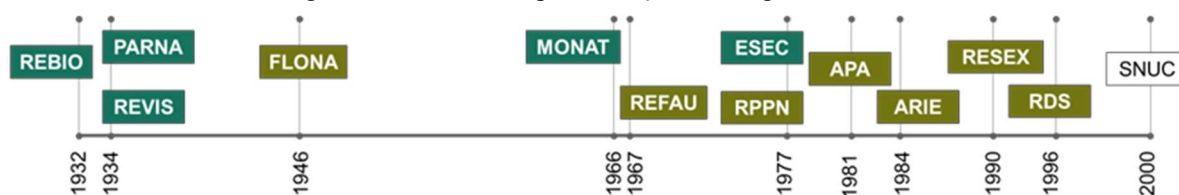
Tabela 1 - Categorias de Unidades de Conservação propostas e aprovadas no SNUC

<b>Categorias propostas no Anteprojeto 2.892/1992</b>		<b>Categorias aprovadas no SNUC</b>	
Unidades de Proteção Integral	Reserva Ecológica Parque Nacional Monumento Natural Refúgio de Vida Silvestre	Unidades de Proteção Integral	Estação Ecológica (ESEC) Reserva Biológica (REBIO) Parque Nacional (PARNA) Monumento Natural (MONAT) Refúgio de Vida Silvestre (RVS)
Unidades de Manejo Provisório	Reserva de Recursos Naturais	-	-
Unidades de Manejo Sustentável	Reserva de Fauna Área de Proteção Ambiental Floresta Nacional Reserva Extrativista	Unidades de Uso Sustentável	Área de Proteção Ambiental (APA) Área de Relevante Interesse Ecológico (ARIE) Floresta Nacional (FLONA) Reserva Extrativista (RESEX) Reserva de Fauna (REFAU) Reserva de Desenvolvimento Sustentável (RDS) Reserva Particular do Patrimônio Natural (RPPN)

Fonte: PUREZA et al., 2015

<sup>3</sup> 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. Disponível em <[https://www.planalto.gov.br/ccivil\\_03/leis/19985.htm](https://www.planalto.gov.br/ccivil_03/leis/19985.htm)>. Acesso em 10 dez. 2022.

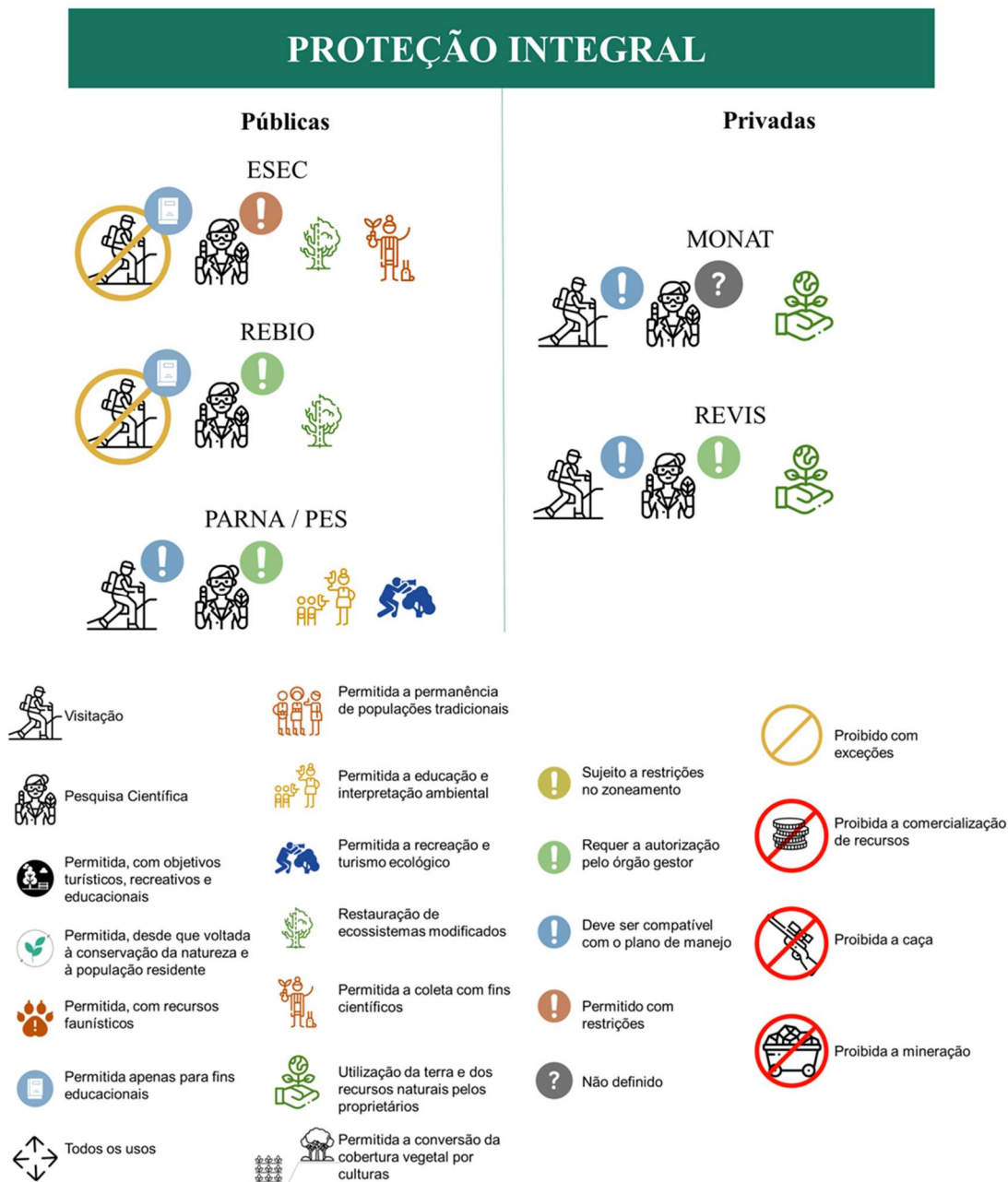
Figura 1 - Linha do tempo da criação de categorias de UCs



Fonte: PUREZA et al., 2015. Organizado pela autora.

Até então, a gestão das UCs não era padronizada e não havia uma centralização na sua gestão, visto que os objetivos e estratégias de gestão eram diferentes para cada categoria (GUERRA; COELHO, 2012). O SNUC estabelece que as UCs estão classificadas em dois tipos de Uso: Proteção Integral e Uso Sustentável. As unidades de Proteção Integral permitem o uso indireto dos seus recursos naturais, isto é, que não implica no seu consumo, coleta, dano ou destruição. As unidades de Uso Sustentável, por outro lado, admitem coleta e uso, comercial ou não, de parte dos seus recursos naturais. As atividades permitidas dependem da categoria em que a UC se enquadra (Figuras 2 e 3).

Figura 2 - Unidades de Conservação de Proteção Integral e atividades permitidas / proibidas



Fonte: Organizado pela autora.

Figura 3 - Unidades de Conservação de Uso Sustentável e atividades permitidas / proibidas



Fonte: Organizado pela autora.



O Plano de Manejo, principal instrumento de gestão da UC, descreve o objetivo da sua criação, seu zoneamento e normas de uso da terra e do manejo de recursos naturais, formalizado com a participação de todas as partes interessadas na gestão da unidade (SILVA; MELLO, 2019). De acordo com o Cadastro Nacional de Unidades de Conservação (CNUC), 66% das UCs estaduais paulistas não têm Plano de Manejo; a mesma proporção não têm Conselho Gestor. Nas unidades federais, 86% carecem do Plano de Manejo e 81% não dispõem de Conselho Gestor, ainda que o SNUC estabeleça que o documento deva ser elaborado no prazo de cinco anos a partir da sua data de criação. Brasil e Carvalho (2021) identificam que, embora tenha havido um esforço por parte do ICMBio e dos governos estaduais em elaborar os roteiros metodológicos de elaboração dos Planos de Manejo, que padroniza, agiliza e facilita sua estruturação, a escassez de recursos humanos, financeiros e estruturais ainda constitui a principal adversidade para a implantação dos planos. De acordo com Santana et al. (2020), a inexistência do Plano de Manejo compromete a gestão da biodiversidade e também o gerenciamento e democratização nas tomadas de decisão na gestão das UCs. Barros e Leuzinger (2018) apontam que o sistema jurídico ambiental é acertado para subsidiar a elaboração dos planos de manejo e que a efetividade da UC está atrelada ao planejamento e execução da sua gestão.

O Plano de Manejo define também o zoneamento da unidade, que estabelece as normas que regem o uso das terras e o manejo dos recursos naturais, incluindo a implantação da infraestrutura necessária à gestão da unidade. O zoneamento organiza espacialmente as zonas no interior e entorno da UC, conforme regras de uso e níveis de restrição (SANTANA et al., 2020). Ou seja, partes da unidade são classificadas como tendo diferentes objetivos de manejo (IUCN, 2022)<sup>4</sup>.

Em todo o mundo, existem centenas de categorias com denominações e definições distintas. Então, desde 1962, a União Internacional para Conservação da Natureza (IUCN - *International Union for Conservation of Nature*) tem agrupado as áreas protegidas de todos os países em classes com objetivos de manejo e características compatíveis (DOUROJEANNI; PÁDUA, 2007). A Tabela 2 mostra as seis categorias da IUCN e as categorias do SNUC correspondentes.

---

<sup>4</sup> IUCN. Guidelines for applying protected area management categories. Disponível em <<https://portals.iucn.org/library/sites/library/files/documents/pag-021.pdf>>. Acesso em 01 dez. 2022

Tabela 2 - Categorias de áreas protegidas conforme IUCN e SNUC

<b>Categoria IUCN</b>	<b>Definição</b>	<b>Uso</b>	<b>Categorias SNUC</b>
<b>I. Áreas naturais selvagens</b>			
Ia. Reserva natural restrita	Áreas naturais manejadas principalmente com fins científicos	Indireto Terra pública	Estação Ecológica Reserva Biológica
Ib. Área natural silvestre	Áreas naturais manejadas principalmente para a proteção da natureza		
<b>II. Parque nacional</b>			
II. Parque nacional	Áreas naturais manejadas principalmente para a proteção de ecossistemas e atividades recreativas	Indireto Terra pública	Parque Nacional / Estadual / Natural Municipal
<b>III. Monumento natural</b>			
III. Monumento natural	Áreas protegidas manejadas para conservar características naturais específicas	Indireto Terra pública ou privada	Monumento Natural
<b>IV. Área de manejo de espécies e habitats</b>			
IV. Área de manejo de espécies e habitats	Áreas manejadas principalmente para a conservação, com intervenção ao nível de gestão	Direto Terra pública ou privada	Refúgio de Vida Silvestre Área de Relevante Interesse Ecológico
<b>V. Paisagem terrestre ou marinha protegida</b>			
V. Paisagem terrestre ou marinha protegida	Áreas terrestres ou marinhas, manejadas principalmente para conservar paisagens e recreação	Direto Terra pública ou privada	Área de Proteção Ambiental
<b>VI. Reserva de recursos manejados</b>			
VI. Reserva de recursos manejados	Áreas de reserva de recursos manejados de forma sustentável	Direto Terra pública ou comum	Reserva Extrativista Reserva de Fauna Reserva de Desenvolvimento Sustentável Floresta Nacional / Estadual

No Brasil, as unidades de Uso Sustentável foram criadas para expandir os sistemas federais e estaduais de UCs, sendo as APAs a categoria dominante na ampliação das UCs brasileiras (CNUC, 2023; DOUROJEANNI e PÁDUA, 2007). Dourojeanni e Pádua (2007) levantam, de modo bastante oportuno, que as categorias de Uso Sustentável não são comparáveis entre si; por exemplo, florestas nacionais e estaduais e reservas extrativistas, embora permitam o uso dos seus recursos, podem garantir a conservação da biodiversidade que abrigam. Contudo, as APAs são absolutamente diferentes, haja vista a possibilidade das terras serem públicas ou privadas, com sobreminência destas últimas. Ainda que o Plano de Manejo - frequentemente inexistente - contemple o zoneamento da área protegida, a dominância de áreas privadas, no caso das APAs, pode implicar no seu descumprimento. Então, é fundamental esclarecer que aumento no número e na área de UCs de Uso Sustentável dão a falsa impressão de que a proteção da biodiversidade é garantida nessas áreas. Chape et al. (2005) complementam que o indicador de extensão de área não informa se áreas protegidas estão atingindo seus objetivos de conservação e que é fundamental considerar, além da eficácia de gestão, quanto e qual biodiversidade está incluída nas áreas protegidas.

### **2.1.2 Eficácia de Unidades de Conservação**

Gurgel et al. (2009) reclassificaram as UCs de acordo com os principais tipos de uso previstos no SNUC (Tabela 3). Os autores explicam, mostrando as principais atividades permitidas nas categorias de UCs, que o uso das terras e dos recursos naturais nestas áreas protegidas podem gerar emprego e renda e, conseqüentemente, o desenvolvimento do país, sem comprometimento à conservação ambiental. As UCs, seriam, deste modo, um meio de promover a organização territorial e não implicam no antagonismo entre conservação e desenvolvimento. A ordem das categorias, em correspondência às classes definidas por Gurgel et al. (2009) é semelhante àquela apresentada por Terborgh e van Schaik (1997) apud Dourojeanni e Pádua (2007) que compara as categorias de manejo quanto à sua eficiência.

Sendo assim, espera-se que as atividades praticadas no interior das UCs estejam restritas àquelas previstas no SNUC e mais especificadas por Gurgel et al. (2009) e que alterações na cobertura da terra estejam relacionadas à ineficiência das áreas protegidas em cumprir com os objetivos das sua categoria de manejo. E, ainda, o zoneamento da Unidade, descrito no Plano de Manejo, deve contemplar essas atividades. Ou seja, é suposto que REBIOS, ESECs e Parques

contenham o avanço da antropização e da conversão da cobertura vegetal nativa, e que APAs, RESEXs e RDSs comportem atividades agrossilvopastoris.

Tabela 3 - Classes de UCs organizadas em função das atividades permitidas

<b>Classe</b>	<b>Principais atividades contemplados no SNUC</b>	<b>Categoria de manejo</b>
Classe 1 – Pesquisa científica e educação ambiental	Desenvolvimento de pesquisa científica e de educação ambiental	Reserva Biológica; Estação Ecológica
Classe 2 – Pesquisa científica, educação ambiental e visitação	Turismo em contato com a natureza	Parque Nacional/ Estadual; Reserva Particular do Patrimônio Natural
Classe 3 – Produção florestal, pesquisa científica e visitação	Produção florestal	Floresta Nacional/Estadual
Classe 4 – Extrativismo, pesquisa científica e visitação	Extrativismo por populações tradicionais	Reserva Extrativista
Classe 5 – Agricultura de baixo impacto, pesquisa científica, visitação, produção florestal e extrativismo	Áreas públicas e privadas onde a produção agrícola e pecuária é compatibilizada com os objetivos da UC	Reserva de Desenvolvimento Sustentável; Refúgio de Vida Silvestre; Monumento Natural
Classe 6 – Agropecuária, atividade industrial, núcleo populacional urbano e rural	Terras públicas e particulares com possibilidade de usos variados visando a um ordenamento territorial sustentável	Área de Proteção Ambiental; Área de Relevante Interesse Ecológico

Fonte: GURGEL et al., 2009.

Para Rodrigues e Cazalis (2020), as áreas protegidas eficazes são aquelas que cumprem o objetivo da conservação da natureza a longo prazo, pelo menos dentro de seus limites; entretanto, como afirmam os autores, a eficácia não é uma característica binária de sim ou não, mas existe em um gradiente de vários fatores que medem a contribuição dessas áreas para a conservação. Os autores afirmam também que a eficácia das áreas protegidas depende tanto das definições no momento do seu planejamento quanto das tomadas de decisão subsequentes.

A eficiência das UCs deve ser avaliada de forma compatível com seus objetivos de criação. Sua avaliação é definida, segundo a IUCN<sup>5</sup>, como a análise de quão bem as áreas protegidas estão sendo geridas, principalmente quanto a metas e objetivos alcançados. Esta reflexão a respeito da gestão das áreas protegidas tem relevância local, regional e nacional, e também internacional. A Convenção sobre Diversidade Biológica (1992)<sup>6</sup>, por exemplo, propôs a apresentação de relatórios regulares à Conferência das Partes, composta pelos 160 países que assinaram o acordo. O documento publicado pela IUCN discute a eficácia das áreas protegidas no sentido da sua boa gestão e compreende a adequabilidade de habitat como um dos valores conservados pelas áreas protegidas.

O gestor deve definir o que deve ter prioridade no planejamento, gestão e avaliação das unidades. Não seria viável, por exemplo, planejar o manejo para garantir a sobrevivência de todas as espécies e a avaliação não poderia mensurar tudo; usualmente, este planejamento enfatiza a conservação de espécies ameaçadas, endêmicas ou espécies chave, e a avaliação da eficácia das áreas protegidas considera indicadores representativos de alguns valores. A eficácia tem um componente individual, isto é, em relação ao planejamento e gestão da unidade em si, e um componente global, que trata da contribuição da unidade em uma rede de áreas protegidas, e que depende da sua extensão total, localização, conectividade e representatividade (RODRIGUES e CAZALIS, 2020). De fato, as revisões sobre eficácia de áreas protegidas (HOCKINGS, 2003; LEVERINGTON et al, 2010; COAD et al., 2015) remetem à sua gestão de forma mais marcada do que em relação aos resultados de proteção da biodiversidade. Em casos mais específicos, versam sobre a perda de habitat e declínios populacionais (GELDMANN et al., 2013).

---

<sup>5</sup> IUCN ESARO. The state of protected and conserved areas in Eastern and Southern Africa. State of Protected and Conserved Areas Report Series No. 1. Nairobi, Kenya: IUCN ESARO, 2020. Disponível em <<https://rris.biopama.org/sites/default/files/2021-01/IUCN2020%20ESA%20SoPACA%20PAME%20Chapter.pdf>>. Acesso em 18 set. 2020.

<sup>6</sup> A Convenção sobre Diversidade Biológica - CDB. Disponível em <<https://www.gov.br/mma/pt-br/textoconvenoportugus.pdf>>. Acesso em 09 set. 2022.

Gonçalves-Souza et al. (2021) avaliaram a eficiência das UCs – neste caso, em relação à manutenção da cobertura da vegetação nativa - em todos os biomas brasileiros, considerando seu tamanho, época de designação e tipo de uso. Os autores concluíram que unidades mais antigas e maiores, e aquelas governadas por indígenas, são menos propensas à conversão de cobertura da terra. Mostraram também que, nas UCs de Uso Sustentável, o reconhecimento das comunidades locais como partes interessadas na gestão das áreas protegidas, e sua integração no planejamento e gestão podem melhorar o desempenho de indicadores de biodiversidade. Garcia et al. (2017), por sua vez, estimaram a supressão da vegetação em áreas protegidas e não protegidas no ecótono Amazônia – Cerrado, na Bacia do Rio Araguaia, e mostraram que as áreas protegidas atenuam a fragmentação, dependendo da sua localização e nível de proteção, embora a aptidão para a agricultura também influencie a conversão da cobertura da terra. Laurance et al. (2012) alertam que metade das áreas protegidas nos trópicos sofre com a destruição do habitat, caça e exploração de recursos florestais, no seu interior e proximidades, que resultam em uma erosão da biodiversidade. Outros trabalhos apontam atividades conflitantes no interior de UCs (BELLÓN et al., 2020; CONCEIÇÃO et al., 2010; MATOS et al., 2021; MORAES et al., 2017; SILVA et al., 2021a).

Por vezes, os conflitos relacionados ao uso das terras, à utilização dos recursos naturais e à permanência de populações tradicionais, ou ainda, a necessidade de fortalecer a proteção de algum recurso natural, justificam a alteração da categoria da UC (SILVA et al., 2021b). Payés et al. (2013) avaliaram as mudanças no uso das terras na região do Despraiado, no litoral sul do Estado de São Paulo, uma área que passou por três categorias de manejo diferentes ao longo dos anos, e concluíram que as restrições ambientais nunca foram integralmente cumpridas. Embora tenha havido uma recuperação da floresta, as atividades agrícolas continuaram existindo mesmo que em baixa proporção, mesmo no período de quase vinte anos (1982-2001) quando a área foi administrada como Estação Ecológica.

O SNUC estabelece que o processo de criação de uma UC contemple a fase preliminar de estudos técnicos e de consulta pública, com exceção das ESECs e REBIOS, de modo a viabilizar seu estabelecimento e gestão. Os objetivos, interesses e motivações para a criação da

unidade devem estar descritos no documento legal de criação e no Plano de Manejo<sup>7,8</sup>. A determinação legal da UC pressupõe sua relevância natural, que não precisa ser, necessariamente, de riqueza paisagística ou particularidade notável. Sua localização, extensão e categoria de manejo devem considerar critérios socioecológicos como: a) alta sensibilidade à pressão humana; b) elevado nível de estresse; c) elevado endemismo; e d) alta diversidade de espécies e habitats (BENJAMIN, 2006)<sup>9</sup>. As duas primeiras estão relacionadas a conflitos já existentes à época da criação da UC. Fendrich et al. (2019) mostrou que, no Brasil, quase 20% das UCs federais e mais de 70% das estaduais estão localizadas em locais não prioritários, o que, em termos de área, representa respectivamente 0,5% e 61% dessas áreas protegidas.

### **2.1.3 Serviços Ecosistêmicos fornecidos pelas Unidades de Conservação**

Além da proteção direta da biodiversidade e dos recursos naturais, as UCs são provedoras de serviços ecossistêmicos, como provisão de água, regulação do clima regional, controle hidrológico, dispersão de sementes, polinização, serviços culturais, entre outros (CATOJO; JESUS, 2022). Os serviços ecossistêmicos são os benefícios que as pessoas obtêm dos ecossistemas, isto é, representam os produtos e resultados de inter-relações ecológicas complexas que são úteis e necessárias para o bem-estar humano (MÜLLER et al., 2015). Além disso, as funções ecossistêmicas referem-se de várias formas ao habitat, às propriedades biológicas ou do sistema, aos processos dos ecossistemas, aos bens do ecossistema (como alimentos) e aos serviços (como a assimilação de resíduos). Ou seja, representam os benefícios que as populações humanas obtêm, direta ou indiretamente, das funções do ecossistema (CONSTANZA et al. 1997). Sendo assim, os investimentos em conservação, restauração e o uso sustentável dos ecossistemas geram benefícios ecológicos, sociais e econômicos substanciais, não se tratando apenas de um compromisso ético entre desenvolvimento e meio ambiente (DE GROOT et al., 2010). Maass et al. (2005) classificam os serviços ecossistêmicos em provisão ou abastecimento (água doce, bens agrícolas e pastoris, recursos diversos e opções

---

<sup>7</sup> ICMBIO / MMA. Roteiro metodológico para elaboração e revisão de planos de manejo das unidades de conservação federais. Disponível em <[https://www.gov.br/icmbio/pt-br/centrais-de-conteudo/publicacoes/roteiros/roteiro\\_metodologico\\_elaboracao\\_revisao\\_plano\\_manejo\\_ucs.pdf](https://www.gov.br/icmbio/pt-br/centrais-de-conteudo/publicacoes/roteiros/roteiro_metodologico_elaboracao_revisao_plano_manejo_ucs.pdf)>. Acesso em 01 dez. 2022.

<sup>8</sup> GOVERNO DE SÃO PAULO. Roteiro metodológico: planos de manejo das Unidades de Conservação do Estado de São Paulo. Disponível em <<https://smastr16.blob.core.windows.net/home/2020/01/roteiro-metodologico-em-baixa.pdf>>. Acesso em 01 dez. 2022.

<sup>9</sup> Benjamin, A. H. V. O regime brasileiro de Unidades de Conservação. Disponível em <<https://core.ac.uk/download/pdf/16015846.pdf>>. Acesso em 01 dez. 2022.

futuras), de regulação (regulação do clima, manutenção da fertilidade do solo, controle de inundações e biorregulação), cultural (beleza cênica) e serviços de suporte.

A biodiversidade, por sua vez, é um dos principais agentes da produtividade, estabilidade e dinâmica dos nutrientes dos ecossistemas e é necessária para manter vários tipos de processos ecossistêmicos. Sendo assim, comunidades com mais espécies tendem a manter maiores níveis de multifuncionalidades do ecossistema (TILMAN et al., 2014). A conservação de áreas de especial interesse para a biodiversidade e para os serviços ecossistêmicos, assim como o aumento de benefícios de biodiversidade e serviços ecossistêmicos foram um dos objetivos do Plano Estratégico de Biodiversidade. Esses objetivos foram estabelecidos nas Metas de Aichi<sup>10</sup> para a Biodiversidade, propostas durante a 10ª Conferência das Partes da Convenção de Diversidade Biológica (COP 10) em 2002 e que deveriam ter sido implementadas até 2020. A Meta 11 estabeleceu percentuais de áreas protegidas em ecossistemas terrestres, aquáticos e costeiros que, como apontado por Wenceslau et al. (2020), foram cumpridos no Brasil, mas de forma desigual entre os biomas. A Meta 14 propõe a restauração e preservação de ecossistemas provedores de serviços essenciais. No Brasil, foram criados mecanismos financeiros para incentivar a conservação da biodiversidade e dos serviços ecossistêmicos como transferência fiscal ecológica, cotas de reserva ambiental, pagamentos por serviços ambientais, concessões turísticas e concessões florestais, tendo a maioria vistas ao envolvimento do setor privado (YOUNG; CASTRO, 2021).

As Unidades de Conservação, no seu papel de salvaguardar a biodiversidade, compõem um cenário de proteção dos serviços ecossistêmicos, principalmente por meio da regulação do uso das terras. Petroni et al. (2022) concluíram que as UCs de Proteção Integral podem reprimir o avanço de usos antrópicos do solo que, por sua vez, afetam o abastecimento de água, a produção de alimentos, o fornecimento de recursos pesqueiros e a regulação hidroclimática. Medeiros et al. (2011) mostraram que as contribuições econômicas das UCs, quando monetizadas, excedem as quantias destinadas à sua manutenção. Os benefícios incluem a produção de água: 80% da hidreletricidade do país vem de fontes geradoras que têm pelo menos um tributário a jusante de unidade de conservação; 9% da água para consumo humano é diretamente captada em unidades de conservação e 26% é captada em fontes a jusante de unidade de conservação; 4% da água utilizada em agricultura e irrigação é captada de fontes dentro ou a jusante de unidades de conservação. Os autores mencionam também os lucros com

---

<sup>10</sup> CBD – Convention on Biological Diversity. Aichi Biodiversity Targets. Disponível em < <https://www.cbd.int/sp/targets/>>. Acesso em 03 dez. 2022.



a produção de madeira e borracha nas UCs de Uso Sustentável, o turismo nos Parques, o controle da emissão de carbono, economia no tratamento de água onde os mananciais têm alta cobertura florestal, e a receita do ICMS Ecológico. Resende et al. (2021) mostraram que, além das UCs, as Terras Indígenas também são altamente eficazes no resguardo da biodiversidade e de serviços ecossistêmicos, principalmente quanto à produção de água, retenção de sedimentos e proteção de espécies ameaçadas.

As UCs são, portanto, imprescindíveis para atividades como a agricultura e a pecuária, que consomem altos volumes de água para a irrigação e a produção de carne e leite. A Instrução Normativa nº 62, de 29 de dezembro de 2011<sup>11</sup>, que aprova o Regulamento Técnico de Produção, Identidade e Qualidade do Leite, prevê que a fonte de abastecimento de água deve assegurar o volume correspondente a 100 litros de água por animal a ordenhar e 6 litros para cada litro de leite produzido. Na pecuária de corte, a ingestão para dessedentação de bovinos é de até 78 litros de água por dia por animal (PALHARES, 2013), além do consumo para limpeza e pós-abate. No entanto, a pegada hídrica da cadeia de produção de bovinos varia conforme o manejo do rebanho, características dos animais, dieta e variáveis ambientais (DICK et al., 2021; PALHARES et al., 2021; PAS; PAS, 2021; PEREIRA et al., 2018). Rodrigues Junior e Dziedzic (2021) obtiveram os valores de pegada hídrica para bovinos de corte de 13.074 L.kg<sup>-1</sup> de carne produzida na Amazônia, em comparação com a média mundial de 21.829 L.kg<sup>-1</sup>. Em relação, à produção agrícola, Bordonal et al. (2018) estimaram o consumo de aproximadamente 1,85 m<sup>3</sup> Mg<sup>-1</sup> de cana-de-açúcar, sendo que a pegada hídrica do etanol de cana-de-açúcar é de 2.245 litros de água por litro de etanol. Oliveira et al. (2022) calculam que o consumo extra de água devido à expansão da cana-de-açúcar no estado de São Paulo foi de 2.150 hm<sup>3</sup> ano<sup>-1</sup> no período de 1985 a 2017.

Outros serviços ecossistêmicos fornecidos por Unidades de Conservação foram avaliados por Coelho-Junior et al. (2021), Gonçalves et al. (2021), Lima (2019), Oliveira et al. (2017), Plácido et al. (2019), Ribeiro et al. (2018), Ricketts et al. (2010) e Soares-Filho et al. (2010).

---

<sup>11</sup> Ministério da Agricultura, Pecuária e Abastecimento – MAPA. Instrução Normativa nº62, de 29 de dezembro de 2011. Disponível em < <https://wp.ufpel.edu.br/inspleite/files/2018/06/IN62.pdf>>. Acesso em 03 dez. 2022.

## 2.2 MODELAGEM PARA O PLANEJAMENTO AMBIENTAL

A criação das UCs deve seguir critérios técnicos que indiquem as áreas prioritárias para conservação da biodiversidade e de outros recursos naturais, e esse processo pode ser orientado pela modelagem ambiental. A conservação da biodiversidade aplicada está relacionada à tomada de decisões, para a qual há um vasto arsenal de técnicas para lidar com a complexidade do gerenciamento de recursos naturais. Nesse sentido, a modelagem é uma abordagem proveitosa para o manejo e monitoramento da biodiversidade e para o desenvolvimento de metas e políticas significativas (BOLAM et al., 2018; NICHOLSON et al., 2018; PERINO et al., 2021).

O desenvolvimento, implementação e avaliação de políticas públicas têm se beneficiado dos modelos computacionais, pois podem facilitar e agilizar a formulação de políticas. Gilbert et al. (2018) apresentam algumas recomendações em relação à modelagem de políticas públicas: a) muitas vezes o principal benefício de projetar e usar um modelo é que ele fornece uma compreensão do domínio da política, em vez dos números que ele gera; b) é preciso tomar cuidado para que os modelos sejam projetados em um nível apropriado de abstração; c) embora os dados apropriados para calibração e validação possam às vezes ser escassos, a modelagem ainda é valiosa; d) a modelagem colaborativa, que envolve uma série de partes interessadas desde o início, aumenta a probabilidade de que o modelo seja usado e adequado para o propósito; e) deve ser dada atenção à comunicação eficaz entre modeladores e partes interessadas; e f) a modelagem para políticas públicas envolve questões éticas que precisam ser cuidadosamente consideradas.

Os modelos podem ser conceituais qualitativos, descrevendo relacionamentos entre elementos de um sistema, ou quantitativos, baseados na análise de padrões observados nos dados (NICHOLSON et al., 2018). Silva e Bergamasco (2001) organizam os modelos em conceituais, físicos e estatísticos, podendo estes últimos ser classificados como estatísticos, estocásticos ou probabilísticos, e dinâmicos.

Os modelos representam entendimentos simplificados e idealizados de um sistema e podem descrever ou prever resultados de conservação sob uma variedade de cenários alternativos (NICHOLSON et al., 2018). Abordagens de serviços ecossistêmicos, por exemplo, organizam a complexa relação entre as pessoas e a natureza para a tomada de decisões, que, por sua vez, demandam resultados concretos e mensuráveis para políticas e planejamento. À medida em que o conceito de serviços ecossistêmicos vai além da aplicação como um dispositivo de

política e comunicação, as suposições de serviços ecossistêmicos e a realidade de sistemas dinâmicos e complexos torna-se mais problemática (EVANS, 2019). Neste contexto, o progresso do poder computacional e das capacidades de modelagem computacional viabilizaram as previsões operacionais para a gestão do meio ambiente (BEVEN, 2009).

As relações causais podem ser inferidas por meio da redução da complexidade e da omissão de detalhes que compreendem a idealização e a abstração da realidade (EVANS, 2018; NIAZ, 1999). Shou et al. (2015) apontam que é preciso existir um equilíbrio entre abstração (aspectos qualitativos) e realismo (aspectos quantitativos), o que depende da questão abordada. Neste sentido, simplificações podem estruturar um modelo altamente abstrato e revelar características gerais de um objeto ou sistema, embora esse tipo de modelo não possibilite previsões quantitativas. Por outro lado, um modelo altamente detalhado não facilita a compreensão de um sistema ou processo, mas pode fazer previsões mais confiáveis. Segundo Hong et al., (2017), as tomadas de decisão para o planejamento ambiental tendem a priorizar os modelos mais simples por serem de mais fácil compreensão.

O princípio da parcimônia, ou Navalha de Ochham (ENGSTRÖM et al., 2016; KIMMINS et al., 2008; WEIJS; RUDELL, 2020; YOUNG et al., 1996), faz parte do método científico e propõe que dentre múltiplas premissas que explicam um determinado fenômeno, deve ser priorizada aquela constituída pelo menor número de variáveis e hipóteses. Esta simplificação tange a colocação de Rol (2008) de que, quanto maior a simplificação, uma afirmação se torna mais próxima da verdade e da generalização. Por outro lado, Coelho et al. (2019) argumentam que não existe evidência empírica que relacione parcimônia e credibilidade: em Ecologia, a parcimônia é importante nos modelos preditivos em que os dados têm uma função decisiva na estrutura do modelo (abordagem instrumentalista). No entanto, a parcimônia pode ser prejudicial para os modelos mecanicistas que consideram mecanismos naturais para representar as causas dos fenômenos biológicos, pressupondo que a natureza seja simples. Os autores enfatizam que a estatística moderna estruturou a ciência como uma investigação de fenômenos naturais baseada em dados, e não mais fundamentada na descrição e categorização de observações. Para exemplificar um inconveniente da simplificação, Hong et al. (2017) mencionam a definição de limiares rígidos na modelagem, que podem gerar superfícies de resposta muito erráticas. Beven (2002) aponta que um modelo “ótimo” é equivalente a mapear um único ponto no espaço do modelo; esta afirmação conduz à ideia de modelos sobreajustados que, devido à presença de ruído, tamanho limitado do conjunto de treinamento e à complexidade dos algoritmos, impedem a generalização dos modelos para ajustar os dados de treinamento e

de teste (YING, 2019). Giere (2004) complementa que os cientistas podem reconhecer semelhanças entre o modelo e aquele aspecto da realidade que está sendo representado; sendo assim, não é o modelo em si que representa esse aspecto da realidade, mas é o cientista que usa o modelo que está fazendo a representação, por meio da seleção de variáveis e ajuste de parâmetros.

A propriedade de simplificação dos modelos remete ao aforismo "*All models are wrong, but some are useful*" enunciado pelo estatístico George Box (BOX, 1976). O primeiro período da frase expressa que todos os modelos estão errados porque nunca representarão o comportamento real exato de um sistema ou objeto. No entanto, apesar das suas limitações, os modelos podem ser úteis, fazendo previsões e/ou facilitando a compreensão da realidade que está sendo representada. E, ainda, é importante entender até que ponto o modelo pode, de fato, ser útil, isto é, é necessário ter a noção da sua incerteza (Figura 4). Para Beven (2002), considerar as incertezas inerentes à modelagem ambiental implicaria na demanda de recursos computacionais e também questões na sociologia da ciência, política e tomada de decisão; sendo assim, não seria necessário considerar a incerteza de forma mais realista a menos que os formuladores de políticas e tomadores de decisão solicitem essa averiguação.

Figura 4 - Trecho do artigo de George Box publicado no Journal of the American Statistical Association em 1976

### **2.3 Parsimony**

Since all models are wrong the scientist cannot obtain a "correct" one by excessive elaboration. On the contrary following William of Occam he should seek an economical description of natural phenomena. Just as the ability to devise simple but evocative models is the signature of the great scientist so overelaboration and overparameterization is often the mark of mediocrity.

### **2.4 Worrying Selectively**

Since all models are wrong the scientist must be alert to what is importantly wrong. It is inappropriate to be concerned about mice when there are tigers abroad.

A modelagem ambiental possibilita a projeção dos impactos das políticas e planos de manejo e, segundo Nicholson et al. (2018), podem subsidiar as políticas de conservação nas etapas de 1) definição da agenda; 2) formulação de metas; 3) implementação; e 4) revisão de objetivos, diretrizes e metas (Figura 5). Pressey et al. (2021) explicam que, no caso da previsão temporal, os modelos devem ser periodicamente re-executados com parâmetros atualizados para reduzir as incertezas do modelo e alterar as prioridades de conservação caso necessário.

Para Nicholson et al. (2018), formuladores de políticas e pesquisadores devem trabalhar de forma colaborativa para aprimorar conhecimentos, métodos e lacunas de implementação, por meio da melhor compreensão das informações e ferramentas utilizadas pelos tomadores de decisão e, a partir daí, da adaptação dos modelos nas escalas temporais e espaciais apropriadas por quem desenvolve os modelos. Esta premissa é reiterada por Chopin et al. (2019) que constatou que modelos que envolvem a participação das partes interessadas tendem a ser mais aceitos pelos tomadores de decisão, visto que o processo participativo facilita a transferência de conhecimento sobre o sistema e sobre conceitos que estão sendo modelados.

Figura 5 - Etapas da política ambiental que podem se beneficiar da modelagem



Fonte: Nicholson et al. (2018), Organizado pela autora.

Diversas aplicações da modelagem no domínio do Planejamento Ambiental têm sido reportadas, tais como conservação da biodiversidade (FERRIER; DRIELSMA, 2010; JETZ et al., 2019; POLLOCK et al., 2020; SEQUEIRA et al., 2018; SPIERENGURG, 2012; VILLERO

et al., 2017), uso e cobertura da terra (CUNHA et al., 2021; MALAWSKA et al. 2014; SILVA et al., 2018; SOHL; CLAGGETT, 2013), agronomia (DURY et al., 2010; LI et al., 2019; SILVA et al., 2018; VIANNA et al., 2020), serviços ecossistêmicos (EVANS, 2019; GENELETTI et al., 2020; GUILLEM et al., 2015; STRAND et al., 2018) e energia (CARVALHO et al., 2021a; DAGOUMAS; KOLTSAKLIS, 2019; LOPION et al., 2018).

### 2.3 SÍNTESE ECOLÓGICA

Os estudos em Ecologia frequentemente combinam dados de diferentes fontes, teorias e ferramentas da ecologia e de outras disciplinas. Este é o conceito de Síntese Ecológica, usada em meta-análises, revisões e na integração de teoria e modelos conceituais, assim como dados espaciais e outras agregações de dados (HALPERN et al., 2020).

A Ecologia é uma ciência que se sobrepõe com várias disciplinas caracterizadas pela dinamicidade e que tem passado por um período de mudanças no modo como é feita, desde a concepção da pesquisa, passando pela coleta e análise dos dados e publicação dos resultados (FARLEY et al., HUNT et al., 2009). Modelos preditivos correlativos, como os de Adequabilidade de Habitat, por exemplo, têm suporte tanto da Ecologia quanto da Geografia e têm se popularizado com o avanço das tecnologias relacionadas aos Sistemas de Informações Geográficas e tecnologias relacionadas (MILLER, 2010).

Os grandes conjuntos de dados (*Big Data*) para o estudo de sistemas ecológicos incluem sensores *in situ* e remotos, recursos de dados comunitários, bancos de dados de biodiversidade, ciência cidadã e estações permanentes (FARLEY et al., 2018). Como resultado, um imenso volume de dados é coletado e a infraestrutura computacional é imprescindível para o desenvolvimento de modos ecológicos em várias escalas espaço-temporais. Na ciência ecológica os dados são coletados por vários grupos e indivíduos que usam diferentes metodologias de amostragem ou simulação e padrões de dados. Só então esses dados são compartilhados e demandam grande esforço de compilação e análise, o que representa um desafio para infraestrutura computacional. Boa parte dos dados históricos coletados em campo estão sendo digitalizados e, ainda que não sistematizados na forma de coleta e que não sejam regulares na série temporal, esses dados são críticos para a ciência (HUNT et al., 2009).

A Síntese integra dados e conhecimento para resolver problemas práticos e muitas questões ambientais atuais estão tornando sua necessidade ainda maior (HALPERN et al.,

2020). As abordagens computacionais dão suporte à Síntese Ecológica, já que viabilizam a aquisição de grandes volumes de dados e a formulação matemática de problemas ecológicos (POISOT et al., 2019) como os Modelos de Adequabilidade de Habitat (MAHs). Essas abordagens computacionais compreendem tanto a Ecologia Computacional, que trata da simulação de ecologias, quanto a Ecoinformática, que tem a ver com a coleta e a análise de informação ecológica (TOLLE et al., 2009).

Poisot et al. (2019) complementam que a Ecologia Computacional é a aplicação da Ciência da Computação a problemas ecológicos, o que reconhece os sistemas ecológicos como complexos e adaptativos. Os autores apontam os MAHs como exemplo de práticas computacionais que geram conhecimento ecológico e concluem que, embora já existisse parte do conhecimento ecológico, métodos estatísticos e dados, as simulações computacionais proporcionaram a abstração e automatização de sistemas para gerar previsões. Em relação à Ecoinformática, Rosenheim e Gratton (2017) associam o termo à aplicação de métodos de *Big Data* em Ecologia, que são oportunos quando o foco são os processos ecológicos que ocorrem em escalas espaciais ou temporais que são muito grandes para serem abordados com experimentação. Essas aplicações com uso intensivo de dados podem ser usadas para fazer previsões (por exemplo, revelando causalidade ambiente-organismo), classificação, ordenação e caracterização (LIN, 2020; PARK et al., 2007).

## 2.4 CIÊNCIA CIDADÃ PARA A BIODIVERSIDADE

Embora grande parte dos dados usados em Sínteses Ecológicas seja proveniente de sensores de campo ou orbitais, principalmente nos estudos geoespaciais, o volume de dados coletados pela Ciência Cidadã é significativo (HUNT et al., 2009). No campo da biodiversidade, espécimes têm sido coletados e registrados durante séculos como observações que permitem investigações científicas baseadas na biodiversidade e na modelagem preditiva. Muitas fontes complementares de observações e medições fornecem perspectivas contemporâneas, incluindo, entre outras, pesquisa ecológica e monitoramento de campo, ciência cidadã e iniciativas comunitárias locais, estudos moleculares, câmeras e sensores automatizados e imagens de satélite (HOBERN et al., 2019).

A Ciência Cidadã ou Participativa é conceituada por Haklay (2021) como “trabalho científico realizado por membros do público em geral onde os dados gerados têm um aspecto

geográfico deliberado e explícito, como capturar uma observação ecológica registrando suas coordenadas no Sistema de Posicionamento Global”. Goodchild (2007) apresenta o termo “*Citizens as sensors*” no contexto em que todo ser humano pode atuar como um sensor inteligente, equipado com dispositivos simples como GPS e instrumentos de medição de variáveis ambientais. De acordo com Frigerio et al. (2021), os projetos de ciência cidadã para o monitoramento da biodiversidade convidam os cidadãos a contribuir com a coleta de dados em diferentes habitats e locais por um longo período de tempo, em pesquisas nacionais e até globais.

A emergência da Internet como infraestrutura de comunicação global e o uso de dispositivos móveis fomentaram o compartilhamento de informações geográficas, inclusive no formato de mapas interativos (HAKLAY, 2021; KAMILARIS; OSTERMANN, 2018), que facilitam tomadas de decisão participativas e colaborativas (BOROUSHAKI; MALCZEWSKI, 2010). Para Matutini et al., (2020), os dados provenientes da Ciência Cidadã têm grande potencial porque (a) grandes quantidades de dados podem ser coletadas em grandes áreas, o que seria difícil e caro para os pesquisadores coletarem; (b) os dados podem ser coletados durante longos períodos de tempo, o que é especialmente útil para estudos em larga escala; (c) os cidadãos estão envolvidos no processo de pesquisa, ganhando conhecimento, e seu envolvimento pode levar a uma melhor implementação da ação de conservação da biodiversidade.

As pesquisas de monitoramento da biodiversidade estão entre os projetos de ciência cidadã mais comuns nas Ciências da Vida (FRIGERIO et al., 2021). Esses dados contribuíram para o avanço dos MAHs por meio de iniciativas que agregam conjuntos de dados em uma única base (MATUTINI et al., 2020) como as plataformas virtuais *Global Biodiversity Information Facility* (GBIF)<sup>12</sup>, *speciesLink*<sup>13</sup>, *Catalogue of Life*<sup>14</sup> e *Ocean Biodiversity Information System* (OBIS)<sup>15</sup>.

No entanto, esses conjuntos de dados de biodiversidade não seguem o mesmo desenho amostral e nem sempre são veiculados por especialistas e podem seguir um viés de observação. A imprecisão geográfica, mais do que incertezas taxonômicas, podem implicar em falsos positivos. O desenvolvimento de projetos interativos, treinamento de voluntários, validação de

---

<sup>12</sup> GBIF - Global Biodiversity Information Facility. Disponível em < <https://www.gbif.org/>>. Acesso em 29 jan. 2023.

<sup>13</sup> speciesLink. Disponível em < <https://specieslink.net/>>. Acesso em 29 jan. 2023.

<sup>14</sup> Catalogue of Life. Disponível em < <https://www.catalogueoflife.org/>>. Acesso em 29 jan. 2023.

<sup>15</sup> OBIS - Ocean Biodiversity Information System. Disponível em < <https://obis.org/>>. Acesso em 29 jan. 2023.



dados de especialistas e melhoria da modelagem estatística são estratégias para melhorar a precisão desses dados (MALDONADO et al., 2015; MATUTINI et al., 2020).

## 2.5 MODELOS DE DISTRIBUIÇÃO DE ESPÉCIES E NICHOS ECOLÓGICOS

Os Modelos de Distribuição de Espécies (MDEs) quantificam a correlação entre fatores ambientais e a distribuição de espécies (MILLER, 2010). São importantes para compreender a relação entre uma espécie e seu ambiente biótico e abiótico com base em observações para fins de inferência ecológica, ou para testar hipóteses ecológicas ou biogeográficas sobre a distribuição das espécies (FRANKLIN, 2009). Embora existam extensos bancos de dados de biodiversidade que informam sobre a localização da ocorrência de espécies, esses modelos valorizam as observações diretas e podem contribuir para a interpretação e compreensão da distribuição de espécies (FRANKLIN, 2009). Outras denominações são Modelagem Preditiva de Distribuição de Habitat (GUISAN; ZIMMERMANN, 2000; RAHMANIAN et al., 2022), Modelagem de Habitat (RATHER et al., 2020; WEGE et al., 2020), Modelos de Adequabilidade de Habitat (OLIVEIRA-SILVA, 2022; TOURINHO, 2022) e Modelagem de Nicho (RODGERS et al., 2019; SIMÕES et al., 2019).

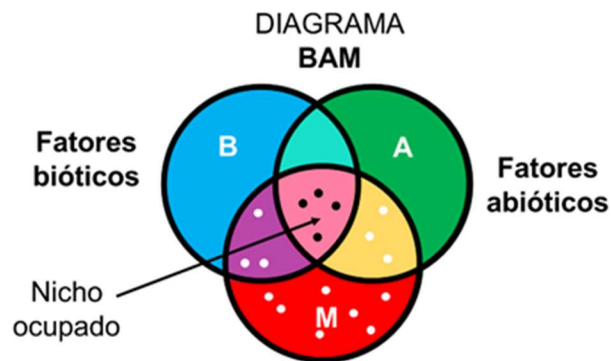
Vale ressaltar que esta última designação refere-se, na maioria dos modelos, à modelagem do Nicho Fundamental, já que considera como variáveis os fatores abióticos do ecossistema. O Nicho Realizado descreve uma porção mais restrita do nicho pois inclui as relações interespecíficas (fatores bióticos), isto é, consiste no conjunto das  $n$  variáveis que permitem a sobrevivência da espécie e que limitam sua distribuição e abundância. Este é o conceito de “Nicho Ecológico n-Dimensional” proposto por Hutchinson (1957), que é idealizado como um hipervolume com  $n$  eixos representativos das variáveis biofísicas. Sendo assim, esta pesquisa não envolve o conceito de “Nicho” pois envolve apenas algumas variáveis abióticas.

Quando condicionados apenas a variáveis climáticas, o modelo caracteriza o "Nicho Climático", que é o caso do Bioclim, baseado na modelagem de envelope bioclimático e que seus resultados são reportados como limiares geográficos de predição de espécies (FRANKLIN, 2009).

Sóberon e Nakamura (2009) apresentam o diagrama BAM (Biotic - Abiotic - Movement) que incorpora a componente “Movimento” ao conceito de Nicho, ou seja, uma espécie só pode ocupar um local se for possível a sua dispersão (Figura 6). A sobreposição das três componentes

é o que os autores chamam de Nicho Ocupado, onde a espécie está presente (pontos pretos). Os locais para onde as espécies podem se deslocar mas que não dispõem dos elementos bióticos e abióticos requeridos para a espécie seriam as verdadeiras ausências (pontos brancos). Nota-se que parte dessas verdadeiras ausências correspondem a parte do Nicho Fundamental, de acordo com o diagrama BAM.

Figura 6 - Diagrama BAM



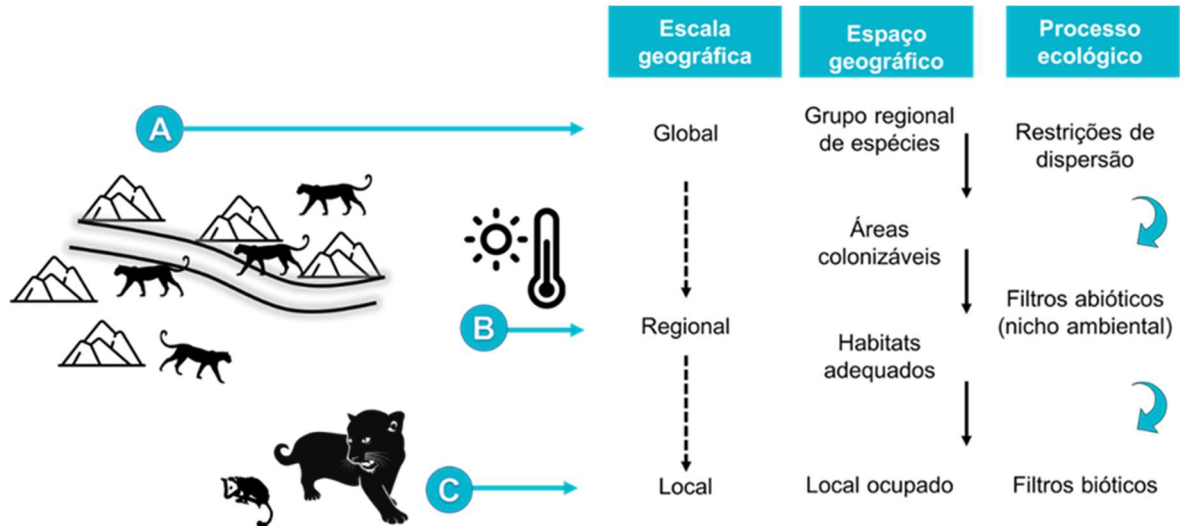
Fonte: SÓBERON; NAKAMURA, 2009. Adaptado

De forma semelhante, Guisan et al. (2017) descrevem três condições principais que precisam ser atendidas para que uma espécie ocupe um local e mantenha as populações: 1) A espécie tem que chegar ao local, ou seja, acessar a região e aí se dispersar: esta condição é uma questão de capacidade de dispersão da espécie a partir daquelas áreas anteriormente ocupadas pela espécie; 2) As condições ambientais abióticas devem ser ecofisiologicamente adequadas para a espécie: trata-se da adequação do habitat abiótico para a espécie-alvo, o que implica que a combinação de variáveis ambientais abióticas no local - muitas vezes referida como adequação ambiental - seja incluída nas condições ambientais que uma espécie precisa para crescer e manter populações viáveis, ou seja, seu ambiente agradável e constitui a base da abordagem de modelagem de adequação de habitat; 3) O ambiente biótico (interações) deve ser adequado para a espécie: tratam-se de interações bióticas, ou seja, interações com outros organismos, sejam positivas (comensalismo, mutualismo) ou negativas (competição, predação), que são ditadas pelo ambiente por meio de sua influência em todos os organismos da comunidade local.

A Figura 7 mostra uma visão hierárquica desses determinantes da ocorrência de espécies e o espaço geográfico correspondente nas escalas global a regional, embora estes processos não ocorram necessariamente seguindo a sequência apresentada; processos estocásticos ou neurais

podem neutralizar alguns desses processos determinísticos e adicionar aleatoriedade aos conjuntos (GUISAN et al., 2017).

Figura 7 - Visão hierárquica das três principais influências na ocorrência de espécies



Fonte: GUISAN et al., (2017). Adaptado pela autora.

Os Modelos de Distribuição Multiespécies (MDMEs), por sua vez, abordam a modelagem do ponto de vista da comunidade biológica e ponderam as relações interespecíficas e que as espécies não são distribuídas de forma independente (CHAPMAN; PURSE, 2011; KISSLING et al., 2011; VERMEIREN et al., 2020; CHEN et al., 2017; NORBERG et al., 2019; POGGIATO et al., 2021; van EE et al., 2022). A modelagem de comunidades pode ser feita a partir de Modelos Conjuntos de Distribuição de Espécies (*Joint Species Distribution Models - JSDM*) que abordam a co-ocorrência de espécies (POLLOCK et al., 2014) e pelos Modelos Empilhados de Distribuição de Espécies (*Stacked Species Distribution Models - SSDM*) que combinam vários modelos individuais para produzir um modelo em nível de comunidade (SCHMITT et al., 2017).

## 2.6 MODELOS DE DISTRIBUIÇÃO DE ESPÉCIES E ADEQUABILIDADE DE HABITAT

Os habitats podem ser descritos com base em um conjunto de variáveis que expressam várias características do ambiente, como clima, uso da terra e solo para plantas, ou clima, recursos alimentares e estrutura de habitat para animais, em um ou uma série de locais (GUISAN et al., 2017). Avaliar suas condições é indispensável para a formulação da conservação da biodiversidade (SU et al., 2021) e os fatores que comprometem sua

adequabilidade podem ser identificados por meio da modelagem (SANI, 2017). Como já mencionado anteriormente, os Modelos de Distribuição de Espécies também são chamados de Modelos de Adequabilidade de Habitat (MAH), que é um termo mais apropriado para esta pesquisa.

Em questões mais práticas, como o manejo da vida silvestre, restauração ecológica, reintrodução de espécies ou avaliação de impacto, os MAHs podem ser usados para prever o valor da adequação do habitat em locais pontuais não pesquisados (onde não houve observação em campo) ou para um período de tempo diferente (FRANKLIN, 2009). Miller (2010) explica que, apesar da designação, o que é modelada ou mapeada é a distribuição dos fatores ambientais adequados. Cayuella et al. (2009) apresentam, de modo similar, que é a caracterização das condições ambientais adequadas para uma espécie que estima sua distribuição geográfica potencial.

Alterações no habitat, principalmente aquelas relacionadas a atividades humanas, como desmatamento, degradação e fragmentação do habitat, modificam o microclima e afetam diretamente as espécies de mamíferos que, por serem espécies-chave, são essenciais para a manutenção da função e da estrutura de ecossistemas (QUINTERO et al., 2023; RAMOS; SANTOS, 2006). Diniz et al. (2022) mostraram que alterações não planejadas da cobertura da terra podem prejudicar a quantidade, qualidade e conectividade do habitat de espécies ameaçadas, comprometendo sua persistência a longo prazo.

Os MAHs podem ser baseados em perfil (que usam apenas pontos de Presença), modelos de regressão e métodos de Aprendizado de Máquina (ambos requerem pontos de Presença e de Pseudo-ausência). Estes últimos dispõem de métodos que servem tanto para classificação quanto para regressão. A Tabela 4 exemplifica alguns desses métodos.

Tabela 4 - Métodos de modelagem de Adequabilidade de Habitat

	<b>Método</b>	<b>Exemplo</b>
Perfil	BioClim	BOOTH (2022)
	Domain	MOURA JÚNIOR et al. (2021)
	ENFA ( <i>Ecological niche factor analysis</i> )	LOURENÇO-DE-MORAES et al. (2019)

	<b>Método</b>	<b>Exemplo</b>
	Distância de Mahalanobis	ETHERINGTON (2021)
Regressão	GLM ( <i>Generalized Linear Model</i> )	SOUZA; PREVEDELLO (2019)
	GAM ( <i>Generalized Additive Model</i> )	BENEMANN et al., 2022
	MARS ( <i>Multivariate Adaptive Regression Splines</i> )	AMIRI et al., 2020
	MaxLike ( <i>Maximum Likelihood</i> )	SCHMIDT et al. (2020)
	<i>Favourability Function</i> (FF)	COSTA et al. (2021)
Aprendizado de Máquina	MaxEnt (Máxima Entropia)	MERCHANT et al. (2023)
	Redes Neurais Artificiais	TAGLIARI et al. (2021)
	GARP ( <i>Genetic Algorithm for Rule Set Production</i> )	GONZÁLEZ et al. (2020)
	GBM ( <i>Gradient Boosting Machines</i> )	VILLA-MACHÍO et al. (2020)
	Floresta Aleatória	COSTA et al. (2022)
	SVM	MIRANDA et al. (2019)
	XGBoost	DOBSON et al. (2022)

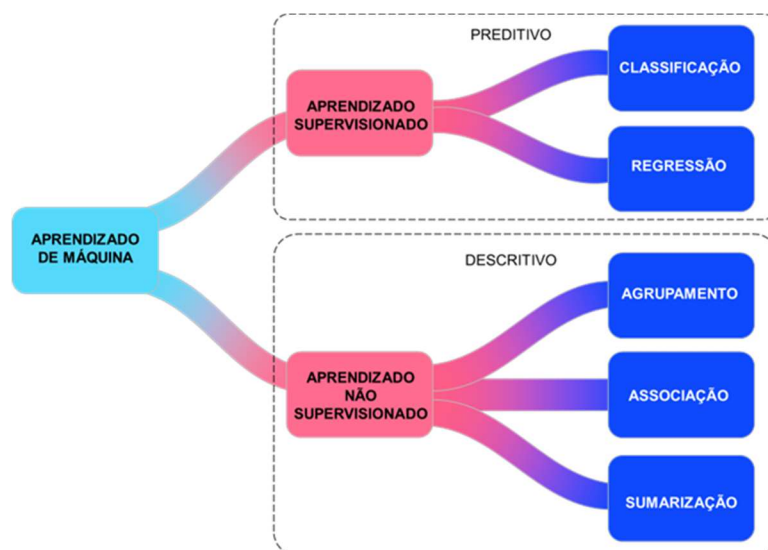
## 2.7 APRENDIZADO DE MÁQUINA

A velocidade de geração e o volume de dados e a complexidade dos problemas impulsionaram o aprimoramento de técnicas computacionais que não fossem tão dependentes da mediação humana para geração de conhecimento (FACELI et al., 2021). A maior parte dessas técnicas é baseada em Aprendizado de Máquina que, segundo Morettin e Singer (2023), consiste na "utilização de modelos estatísticos acoplados a algoritmos computacionais desenvolvidos para extrair informação de conjuntos de dados contendo, em geral, muitas unidades amostrais e muitas variáveis". Mitchell (1997) complementa que esses algoritmos são aprimorados automaticamente por meio da experiência.

Os algoritmos de Aprendizado de Máquina são aplicados em tarefas que podem ser Preditivas ou Descritivas (Figura 8). Nas tarefas preditivas, os algoritmos são aplicados a conjuntos de dados de treinamento rotulados para ensinar o modelo como deve ser feita a predição; o modelo então é aplicado a novos objetos representados pelos valores dos seus atributos (ou variáveis), seguindo o paradigma de aprendizado supervisionado. A supervisão refere-se à interferência de um “supervisor externo” que orienta o que o modelo deve aprender, ou seja, a qual rótulo um objeto deve pertencer de acordo com os valores dos seus atributos. O objetivo é gerar uma função a partir dos dados coletados e classificados que calcula o valor da classe para um novo elemento de acordo com os valores nas variáveis consideradas no aprendizado. Nas tarefas de classificação, o rótulo é discreto, e o algoritmo divide os objetos de características semelhantes em um número finito de classes. Nas tarefas de regressão, os rótulos são contínuos e assumem valores infinitos dentro de um limite definido pelo conjunto de treinamento. Nas tarefas descritivas, os algoritmos reconhecem padrões dos valores preditivos de um conjunto de dados sem a orientação do “supervisor externo”, seguindo o paradigma de aprendizado não supervisionado, como é executado no agrupamento de dados similares e a definição de regras de associação (ERTEL, 2017; FACELI et al., 2021).

A capacidade de generalização do algoritmo é essencial e previne que seu desempenho seja sobreajustado aos dados de treinamento. Os dados de treinamento têm o conhecimento que o algoritmo de aprendizado deve extrair e aprender e devem ser representativos para a tarefa a ser aprendida. Os dados de teste servem para avaliar se o agente treinado pode generalizar bem para novos dados a partir dos dados de treinamento (ERTEL, 2017).

Figura 8 - Hierarquia clássica de Aprendizado de Máquina



Fonte: FACELI et al. (2021). Adaptado pela autora.

## 2.7 BIG DATA E APRENDIZADO DE MÁQUINA NA MODELAGEM AMBIENTAL

Os termos “*Data Science*” (Ciência de Dados) e “*Big Data*” (Megadados) são usados desde a década de 1950 (DONOHO, 2017), embora sua aplicação tenha se tornado mais frequente com a crescente produção de dados, como é o caso dos dados geoespaciais, que incluem dados de censo, imagens de satélite, dados meteorológicos, dados de telefones celulares, imagens desenhadas e dados de mídia social (KAMILARIS; OSTERMANN, 2018).

O grande volume de dados, o aumento do uso de computação em nuvem e de alto desempenho e o acesso a dados abertos (*open data* e *open access*) propiciaram a tomada de decisões orientada por dados em diversas áreas das Ciências Ambientais (LOCKERS et al., 2016) como recursos hídricos (SUN; SCANLON, 2019), agricultura (WEERSINK et al., 2018), planejamento urbano (BABAR et al., 2019), resíduos sólidos (YI; MA, 2020), mudanças climáticas (PAPADOPOULOS; BALTA, 2022), serviços ecossistêmicos (WILLCOCK et al., 2018) e conservação da biodiversidade (RUNTING et al., 2020).

Farley et al. (2018) associam as características de *Big Data* (V's do *Big Data*) ao tipo de fonte de dados ecológicos: (1) Sistemas de Sensoriamento Remoto geram grandes *volumes* de dados; (2) Esforços de ciência cidadã, por indivíduos com graus variados de especialização, envolvem a *veracidade* do dado coletado; (3) Redes de sensores quase em tempo real fornecem dados contínuos em alta *velocidade*; (4) Observações de campo e experimentos por cientistas apresentam ampla *variedade* de medidas, sistemas e escalas.

No caso dos dados geoespaciais, os sistemas de Sensoriamento Remoto orbital, aéreo e terrestre, adquiridos em várias resoluções espacial, temporal e espectral, processados por ferramentas com alto poder computacional e armazenados em sistemas em nuvem, permitem simulações em escala global com precisão cada vez melhor (NATIVI et al., 2015). Algumas aplicações como o monitoramento de queimadas, desmatamento e desastres naturais, requerem o processamento de dados em tempo real ou quase real e métodos computacionais mais escaláveis para melhor aproveitamento dos recursos de armazenamento em nuvem (ANTUNES, 2019; YANG et al., 2017). Técnicas de computação distribuída como Hadoop (ELDAWY; MOKBEL, 2014; WANG et al., 2019), MapReduce (ALKATHIRI et al., 2019), Spark (GUO et al., 2022; YU et al., 2019) e Dask (RAWSON et al., 2021) são ferramentas de computação paralela que têm sido empregadas no processamento de dados geoespaciais. Yue e Jiang (2014) apresentam o conceito de BigGIS, o Sistema de Informações Geográficas habilitado para o *Big Data* e que engloba aspectos do WebGIS, CloudGIS e CyberGIS.

A descrição de sistemas como os abordados pelas Ciências Ambientais (por exemplo, a relação entre organismos vivos e seu ambiente) requer informações complexas e técnicas, como as de Aprendizado de Máquina, que permitam lidar com a magnitude e a heterogeneidade dos dados, reconhecer padrões, identificar correlações desconhecidas e descobrir tendências (LOCKERS et al., 2016; WU et al., 2016). Ou seja, as técnicas de Aprendizado de Máquina são bastante oportunas quando uma análise geoespacial traz muitos planos de informação. Neste contexto estão inseridas aplicações como classificação de uso e cobertura da terra (SOUZA JR et al., 2020; ZHANG; LI, 2022), mapeamento de risco de enchentes (DARABI et al., 2019) e de deslizamentos (BOURAMTANE et al., 2022) e modelos de adequabilidade de habitat (BEERY et al., 2021; CHEN; DING, 2022; GOBEYN et al., 2019; SILVA et al., 2022; ZHANG; LI, 2017).

Nestes últimos, o aprendizado supervisionado é normalmente aplicado para prever a ocorrência de espécies em dimensões geográficas e/ou temporais; o treinamento dos modelos é feito a partir de dados rotulados com apenas duas classes, presença e ausência (ou pseudo-ausência) com dados ambientais que são as variáveis preditoras (GOBEYN et al., 2019).

## 2.8 PESQUISA REPRODUTÍVEL

A modelagem ambiental, por utilizar-se de ferramentas computacionais, a torna potencialmente reprodutível. Uma pesquisa (ou análise) é reprodutível quando outros podem reproduzir os resultados de um estudo científico, a partir dos dados, código e documentação originais (ESSAWY et al., 2020). As observações e teorias não dependem de quem as relata, mas da noção de que as mesmas descobertas são obtidas por qualquer pessoa que realize procedimentos semelhantes (COVENEY et al., 2021).

O critério de reprodutibilidade é um dos pilares do método científico (POPPER, 1959), que abrange as seguintes etapas: 1) Formulação da hipótese; 2) Projeção do estudo; 3) Execução dos estudos e coleta dos dados; 4) Análise dos dados; 5) Divulgação do estudo. Cada uma dessas etapas deve ser relatada com clareza, fornecendo documentação clara e aberta, tornando o estudo transparente e reproduzível (BEZJAK et al., 2018).

Ivie e Tain (2018) elencam alguns dos benefícios da Pesquisa Reprodutível que são: 1) Verificar (ou refutar) resultados de outros pesquisadores; 2) Verificar os próprios resultados; 3) Melhorar a produtividade; 4) Resistir à evolução tecnológica, isto é, em condições de



atualização de versões, uma memória descritiva pode ajudar a recriar a pesquisa; 3) Viabilizar o suporte da comunidade científica por meio de sistemas colaborativos.

Dado que nem sempre é possível replicar um estudo na sua completude, a reprodutibilidade é definida como um espectro de possibilidades entre uma investigação que não pode ser validada e outra que é totalmente replicável (PENG, 2011). A Figura 9 mostra as condições de reprodutibilidade de um estudo que envolve a Computação Científica. Zaragozi et al. (2020) complementaram e adaptaram o espectro de reprodutibilidade para projetos de GIScience (Figura 10), que apresenta três zonas principais de acordo com os recursos disponíveis: 1) apenas publicação (não reproduzível na maioria dos casos); 2) publicação com alguns elementos que promovem a reprodutibilidade (ou seja, código, dados, ambientes, ativos, plataformas, práticas, etc.); e 3) pesquisa totalmente replicável.

É oportuna a diferenciação entre os termos “Reprodutibilidade” e “Replicabilidade”. O primeiro termo refere-se à capacidade de utilizar o código e os dados de uma publicação anterior, executar o código novamente e obter os mesmos resultados. O segundo termo refere-se à capacidade de executar novamente um experimento e obter resultados condizentes com o estudo original usando novos dados (CASSEY; BLACKBURN, 2006; ESSAWY et al., 2020).

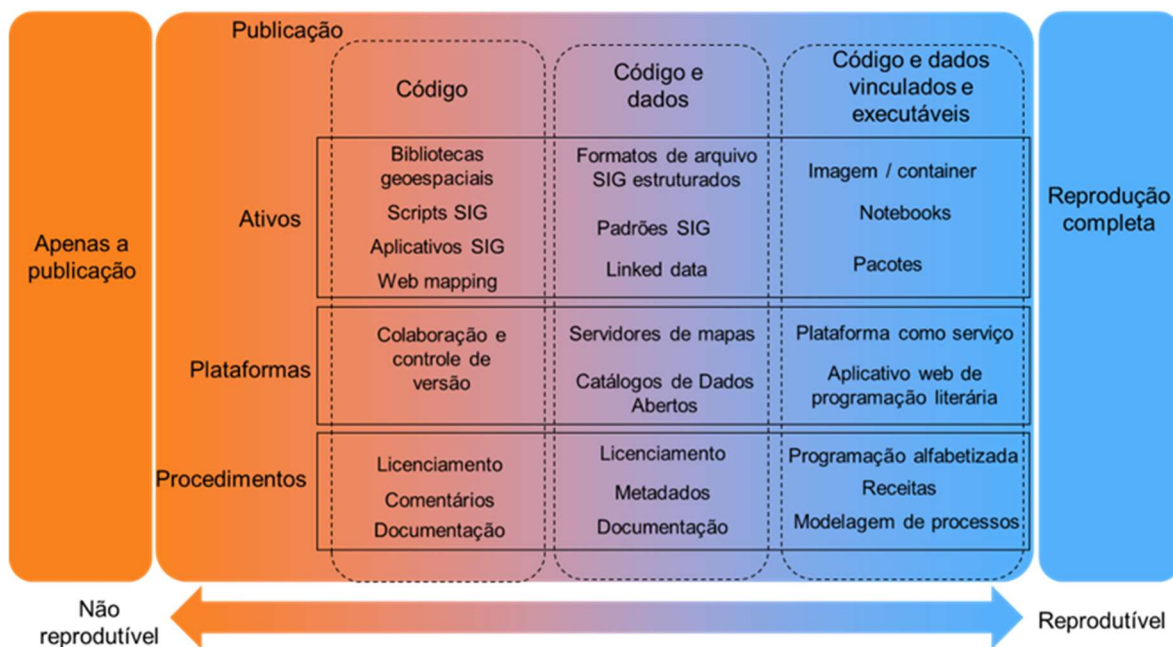
Uma das premissas da reprodutibilidade é o acesso aos dados que devem seguir os princípios FAIR (*Findability, Accessibility, Interoperability, Reusability*) de dados que anunciam que os dados devem ser localizáveis (a partir da criação de metadados que apoiem sua localização), acessíveis (os mecanismos de acesso aos dados sejam claros), interoperáveis (os dados devem ser estruturados para que possam ser integrados a outros dados, aplicativos e fluxos de trabalho) e reutilizáveis (devem ser aplicadas licenças que permitam a reutilização dos dados) (LAMPRECHT et al., 2020). Perino et al. (2021) apontam que a aplicação dos princípios FAIR é essencial para facilitar a transparência e preencher a lacuna de dados e, assim, viabilizar a estruturação de sistemas nacionais de monitoramento da biodiversidade.

Figura 9 - Espectro da Reprodutibilidade



Fonte: ZARAGOZI et al. (2020). Adaptado pela autora.

Figura 10 - Espectro da Reprodutibilidade adaptado para projetos de GIScience



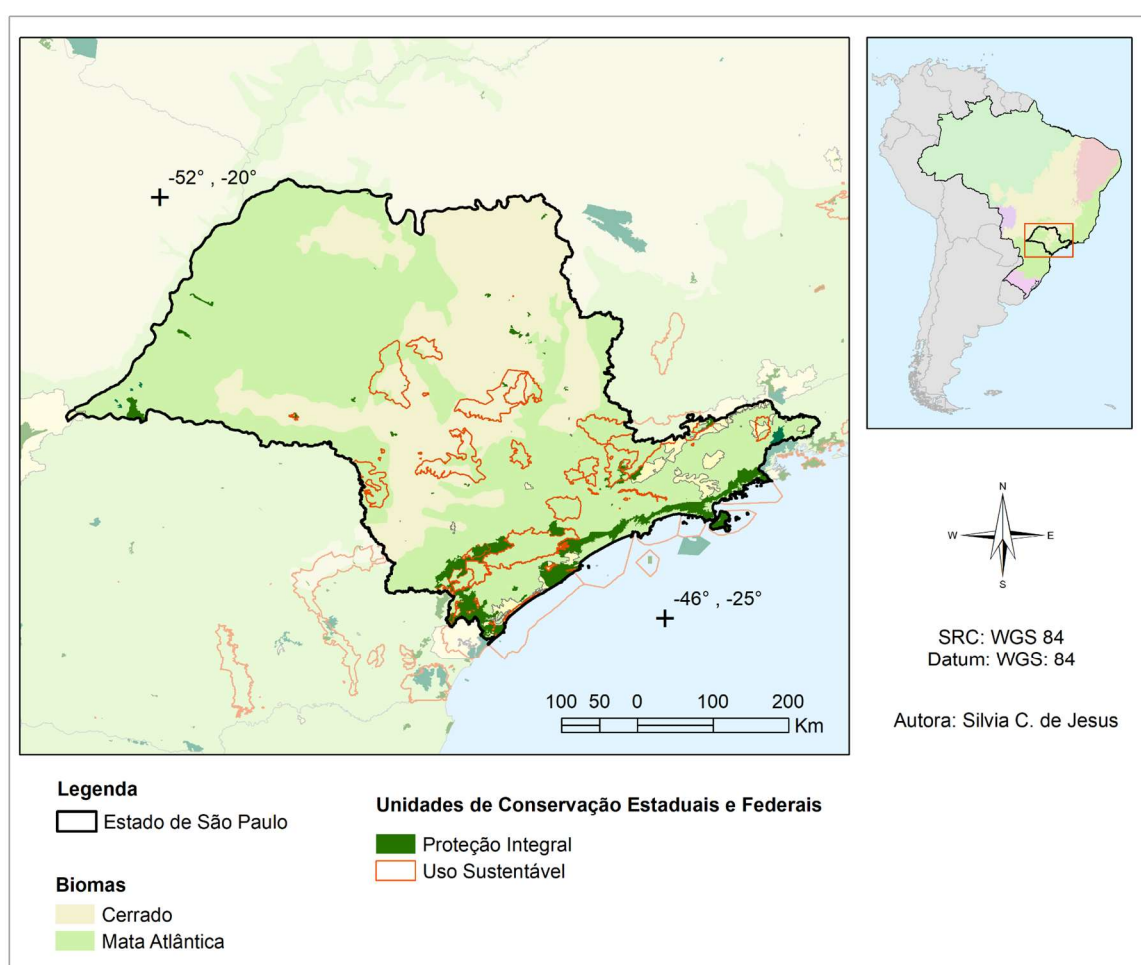
Fonte: ZARAGOZI et al. (2020). Adaptado pela autora.

### 3 METODOLOGIA

#### 3.1. CARACTERIZAÇÃO DA ÁREA DE ESTUDO

O estado de São Paulo é o mais populoso e economicamente desenvolvido do país<sup>16</sup>. Tem aproximadamente 248.219 km<sup>2</sup> de extensão, sendo 77% deste território consolidadamente antropizado (MAPBIOMAS, 2022). Neste estudo foram selecionadas as 106 UCs estaduais e 15 UCs federais que compõem o banco de dados de Unidades de Conservação do Instituto Socioambiental (ISA, 2018) que cobrem 35.286 km<sup>2</sup> (14,2% do território)<sup>17</sup> (Figura 11) .

Figura 11 - Localização da área de estudo

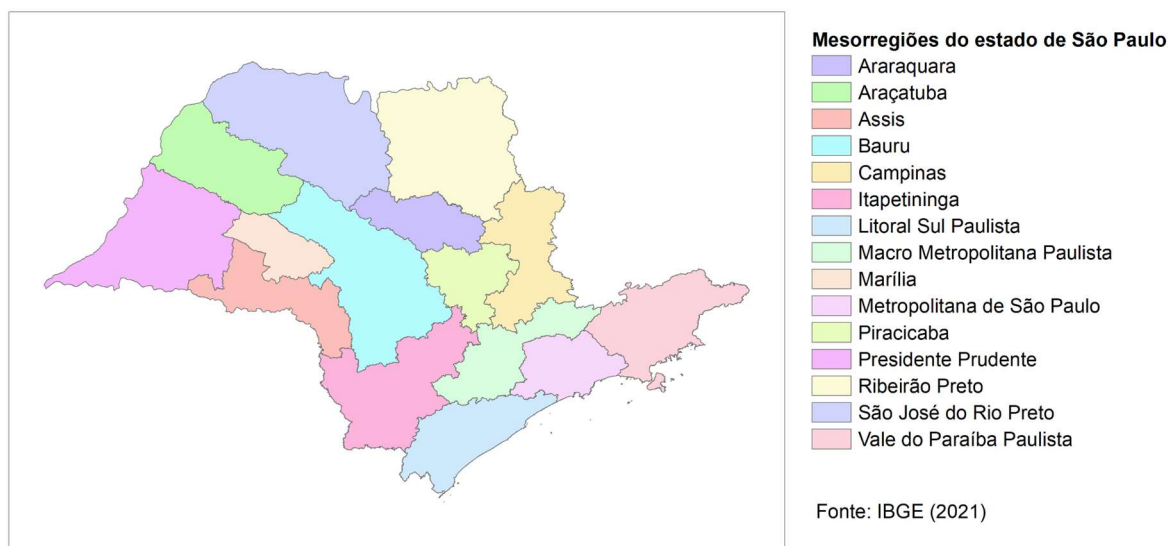


<sup>16</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. **Estimativas da População (2021)**. Disponível em <[https://ftp.ibge.gov.br/Estimativas\\_de\\_Populacao/Estimativas\\_2021/estimativa\\_dou\\_2021.xls](https://ftp.ibge.gov.br/Estimativas_de_Populacao/Estimativas_2021/estimativa_dou_2021.xls)>. Acesso em 24 out. 2022.

<sup>17</sup> A descrição detalhada sobre a organização dos dados vetoriais de Unidades de Conservação está na seção 3.5.2

O estado de São Paulo foi dividido geograficamente pelo IBGE<sup>18</sup> em 15 mesorregiões (Figura 12). É o maior produtor de cana-de-açúcar do Brasil, com 55% da extensão da área plantada no país<sup>19,20</sup>. A sacaricultura é predominante no estado de São Paulo, ocupando 23% do território (SOUZA JR et al., 2020; MAPBIOMAS, 2022) e está presente em quase todo o centro-oeste.

Figura 12 - Mesorregiões do estado de São Paulo



Fonte: IBGE (2021)<sup>21</sup>. Organizado pela autora.

O relevo contribuiu no processo de ocupação antrópica em todo o Brasil; até meados do século XX, as práticas agrícolas eram estabelecidas nas regiões onde predominam relevos planos ou pouco inclinados, como os planaltos e depressões. Nos anos 1970, o incremento tecnológico no setor agropecuário expandiu o eixo de desenvolvimento para o Norte do país e atualmente grande parte do país é dominada por atividades antrópicas. As características da paisagem associadas a fatores técnico-científicos e político-econômicos definiram a configuração atual do território brasileiro. Em diferentes momentos da economia brasileira, essa

<sup>18</sup> IBGE - Malhas Territoriais. Disponível em <[https://geofp.ibge.gov.br/organizacao\\_do\\_territorio/malhas\\_territoriais/malhas\\_municipais/municipio\\_2021/Brasil/BR/](https://geofp.ibge.gov.br/organizacao_do_territorio/malhas_territoriais/malhas_municipais/municipio_2021/Brasil/BR/)>. Acesso em 31 mar 2022.

<sup>19</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. Censo agropecuário. Disponível em <<https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/21814-2017-censo-agropecuario>>. Acesso em 16 abr. 2022

<sup>20</sup> UNICA - Associação Brasileira da Indústria da Cana-de-Açúcar. Observatório da Cana. Disponível em <<https://observatoriodacana.com.br/>>. Acesso em 16 abr. 2022.

<sup>21</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. Mesorregiões. Disponível em <[https://geofp.ibge.gov.br/organizacao\\_do\\_territorio/malhas\\_territoriais/malhas\\_municipais/municipio\\_2021/Brasil/BR/BR\\_Mesorregioes\\_2021.zip](https://geofp.ibge.gov.br/organizacao_do_territorio/malhas_territoriais/malhas_municipais/municipio_2021/Brasil/BR/BR_Mesorregioes_2021.zip)>. Acesso em 14 abr. 2022.

relação aconteceu de forma variada por meio de programas, projetos e planos para uma política desenvolvimentista que não considerou as implicações sociais e ambientais (ROSS, 2006).

O Estado de São Paulo sempre manteve uma taxa de urbanização acima da nacional (Tabela 5). A região metropolitana de São Paulo é ocupada por 10,3% de toda a população brasileira<sup>22</sup>, sendo um dos maiores aglomerados urbanos do mundo (VICTOR et al., 2004). No Brasil, o processo de urbanização se intensificou no início do século 20, com uma guinada na segunda metade do século, como resultado da industrialização.

Tabela 5 - Evolução da população urbana no Brasil e no estado de São Paulo

<b>Brasil</b>	<b>População total aproximada</b>	<b>População urbana</b>
1940**	70 milhões	18,8 milhões (26,3%)
1960**	70.992.343	32.004.817 (45%)
1970**	94.508.583	52.904.744 (56%)
1980**	121.150.573	82.013.375 (68%)
1991**	146.917.459	110.875.826 (75%)
2000**	169.590.693	137.755.550 (81%)
2010**	190.755.799	160.925.792 (84%)

<b>Estado de São Paulo</b>	<b>População total aproximada</b>	<b>População urbana</b>
1960**	8.149.979	12.974.699 (63%)
1970**	14.432.244	17.958.693 (80%)
1980**	22.494.328	25.375.199 (89%)
1991**	29.272.927	31.546.473 (93%)
2000**	34.531.635	36.969.476 (93%)
2010**	39.585.251	41.262.199 (96%)

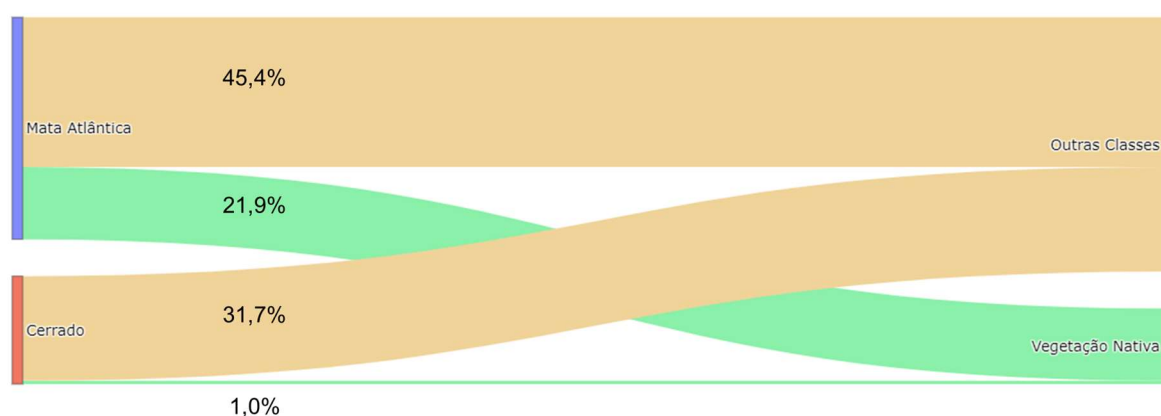
Fonte: \*VICTOR et al. (2004); \*\* IBGE (2010)<sup>23</sup>. Organizado pela autora.

<sup>22</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. Estimativas da População (2021). Disponível em <[https://ftp.ibge.gov.br/Estimativas\\_de\\_Populacao/Estimativas\\_2021/estimativa\\_dou\\_2021.xls](https://ftp.ibge.gov.br/Estimativas_de_Populacao/Estimativas_2021/estimativa_dou_2021.xls)>. Acesso em 24 out. 2022.

<sup>23</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. Sinopse do Censo Demográfico 2010. Disponível em <<https://censo2010.ibge.gov.br/sinopse/index.php?dados=8>>. Acesso em 24 out. 2022.

O Inventário Florestal do Estado de São Paulo<sup>24</sup> mapeou a cobertura vegetal nativa do estado, classificando e quantificando as fitofisionomias. A vegetação nativa remanescente totaliza 56.705 km<sup>2</sup>, ou 22,9% da área do estado. Os remanescentes mantêm 32,6% da superfície original da Mata Atlântica e apenas 3,0% do Cerrado no estado de São Paulo. A Figura 13 mostra os percentuais de vegetação nativa remanescente e outras classes de cobertura da terra em relação à área do estado de São Paulo nos dois biomas.

Figura 13 - Diagrama de Sankey com os percentuais de vegetação nativa remanescente e outras classes de cobertura da terra em relação à área do estado de São Paulo



Fonte: IF, 2020 (organizado pela autora).

As fitofisionomias mais expressivas em termos de área são a floresta ombrófila densa (estágio médio e avançado), que corresponde a 44,3% da vegetação nativa, concentrada na região da Serra do Mar, e se estende à porção norte da mesorregião Macro Metropolitana Paulista. A Floresta Estacional Semidecidual (estágio médio e avançado) equivale a 30,8% da vegetação nativa no estado e está distribuída em fragmentos por todo a porção oeste de São Paulo até a porção classificada por Almeida (1964) como Depressão Periférica, com destaque para o fragmento protegido pelo Parque Estadual do Morro do Diabo, na Mesorregião de Presidente Prudente. Trata-se da maior área contínua remanescente de floresta estacional semidecidual na porção ocidental do estado de São Paulo<sup>25</sup>.

<sup>24</sup> IF - Instituto Florestal. Inventário Florestal do Estado de São Paulo: Mapeamento da cobertura vegetal nativa. 40 páginas. 2020. Disponível em <<https://smastr16.blob.core.windows.net/home/2020/07/inventarioflorestal2020.pdf>>. Acesso em 08 mar. 2022

<sup>25</sup> IF - Instituto Florestal. Plano de Manejo do Parque Estadual do Morro do Diabo. 312 páginas. 2006. Disponível em <<https://smastr16.blob.core.windows.net/fundacaoflorestal/2012/01/morrododiabo.pdf>>.

Ross (2006) classifica a vegetação e os sistemas naturais no estado de São Paulo como fortemente transformados em campos de pastagens cultivadas, campos de agricultura e silvicultura, sendo:

a) Domínio da floresta tropical semidecidual, onde originalmente a floresta subcaducifolia tropical apresentava vários estratos vegetais que perdem as folhas parcialmente durante a época de estiagem. São áreas em que as condições pedoclimáticas viabilizaram a agricultura mecanizada e a pecuária de cria e corte. Os índices pluviométricos variam em torno de 1.300 mm/ano com estiagem de três meses e temperaturas médias entre 18°C e 22°C. O relevo é caracterizado por amplas colinas de topos convexos com declividades médias que variam entre 10% e 20% e altitudes entre 400 e 800 metros. Ainda persistem alguns fragmentos florestais que, em áreas particulares, são intensamente degradados pela exploração florestal seletiva. Corresponde à região classificada por Fernandes (2006) como Setor do Planalto, onde a vegetação segue o padrão geral das formações escleromorfias. Ocorrem o Cerrado, o Cerradão e encaves vegetacionais da floresta atlântica. As feições do Cerrado (Campo Cerrado e Cerrado) cobrem amplas áreas, reproduzindo o padrão savânico mais generalizado. O Cerradão é mais comum nas encostas das chapadas ou às margens dos rios.

b) Domínio da floresta tropical atlântica, onde o clima predominantemente quente e úmido e o relevo montanhoso direcionaram o desenvolvimento da floresta perenifolia higrófila com cinco ou mais estratos e dossel emergente de até 30 metros. O relevo é bastante dissecado, com alta densidade de drenagem, vales entalhados e vertentes inclinadas, sendo caracterizado pelos morros e serras que esculpem os planaltos e serras do planalto atlântico, e pelas escarpas de grande envergadura como as das serras do Mar e Mantiqueira. Apresenta altitudes acima de 700 metros, ultrapassando, em algumas áreas, os 2.000 metros. Os índices pluviométricos ultrapassam 4.000 mm/ano, com chuvas menos intensas no inverno. A vegetação nativa foi convertida em pastagens, cafeicultura e silvicultura, mineração e urbanização. Há fragmentos de matas secundárias que são resultado dos processos de auto-regeneração, principalmente nas regiões de relevos inacessíveis e nas unidades de conservação. Corresponde à região classificada por Fernandes (2006) como Província Atlântica, com composição florística bem diversificada à qual é atribuída a maior riqueza taxonômica no Brasil. Há um padrão vegetacional de gênese orográfica pelo alto volume de precipitação caracterizado pela floresta pluvial.

O estado de São Paulo está na região Zoogeográfica Neotropical. Marques Neto (2018) descreve as províncias zoogeográficas propostas por Fittkau (1969) que, no estado de São Paulo, equivalem às Províncias Tupi e Guaraní (Figura 14). A Província Tupi apresenta alto grau de endemismo, como é o caso do *Callithrix aurita* (sagui-da-serra-escuro) em perigo de extinção<sup>26</sup>. Por outro lado, a Província Guaraní representa os corredores de transição entre o domínio tropical atlântico e o Cerrado, o que reflete na falta de endemismo de mamíferos. Há abundância de didelídeos, generalistas e de alta vagueabilidade, o que garante ampla distribuição geográfica e manifesta-se em um único provável endemismo (*Monodelphis theresa*). Os primatas têm representação nos gêneros *Brachyteles*, *Leontopithecus*, *Alouatta*, *Callicebus*, *Cebus* e *Callithrix*, sendo que melhor se adaptam a ambientes alterados as espécies mais generalistas, que podem habitar ambientes perturbados, e as espécies onívoras, de dieta mais ampla. O estado de São Paulo abriga grande parte dos mamíferos carnívoros brasileiros, com destaque para as famílias Felidae, Canidae e Mustelidae. Os felinos são representados pela *Panthera onca* e pelos gêneros *Puma* e *Leopardus* que são mais generalistas e mantêm-se em paisagens fragmentadas e antropizadas. Outras espécies terrestres de grande porte são representadas pela anta (*Tapirus terrestris*), pelos taiassuídeos (*Pecari tajacu* e *Tayassu pecari*), pelos javalis introduzidos (*Sus scrofa*) e por cervídeos do gênero *Mazama*.

Figura 14 - Províncias zoogeográficas da América do Sul, segundo a proposição de Fittkau (1969)



Fonte: MARQUES NETO (2018).

<sup>26</sup> Ministério do Meio Ambiente. Portaria MMA de 7 de junho de 2022. Disponível em <[https://www.icmbio.gov.br/cepsul/images/stories/legislacao/Portaria/2020/P\\_mma\\_148\\_2022\\_altera\\_anexo\\_s\\_P\\_mma\\_443\\_444\\_445\\_2014\\_atualiza\\_especies\\_amecadas\\_extincao.pdf](https://www.icmbio.gov.br/cepsul/images/stories/legislacao/Portaria/2020/P_mma_148_2022_altera_anexo_s_P_mma_443_444_445_2014_atualiza_especies_amecadas_extincao.pdf)>. Acesso em 10 set. 2022



Vivo et al. (2011) registraram 231 espécies de mamíferos no estado de São Paulo, que são espécies de ampla distribuição na América do Sul tropical. Segundo os autores, a diversidade de biomas no estado também justifica esta riqueza de espécies, considerada alta em relação à extensão do território. Entretanto, Bovo et al. (2018) apontam que pequenos fragmentos florestais não sustentam populações viáveis de mamíferos de pequeno a grande portes mas desempenham um papel importante na conectividade em paisagens antropicamente modificadas. Dallapicolla et al. (2021) identificaram que, na Mata Atlântica, apenas 12% da área de endemismo de pequenos mamíferos não voadores estão cobertos por remanescentes florestais e menos de 11% estão sob proteção legal, o que reforça a necessidade de que as UCs sejam representativas nas áreas de alta endemicidade.

### 3.2. AMBIENTE DE PROCESSAMENTO

Os estudos que envolvem Análise Espacial, Geoprocessamento e/ou Sensoriamento Remoto, tanto aqueles que envolvem plataformas com interface gráfica quanto os baseados em códigos e linguagem de programação, são análises auxiliadas por computador (*computer aided analysis*) (MUENCHOW et al., 2019) e, portanto, são passíveis de serem totalmente reprodutíveis (KONKOL, 2019; ZARAGOZÍ et al., 2020; NÜST; PEBESMA, 2021). Esta pesquisa é baseada no conceito de reprodutibilidade, isto é, seu fluxo de trabalho e resultados podem ser reproduzidos com base nos dados, códigos e documentação originais (ALSTON; RICK, 2021). Zaragozí et al. (2020) explicam que a reprodutibilidade existe em um espectro de possibilidades que variam desde a situação em que apenas a publicação do artigo está acessível, o que não permite a reprodução do estudo, até a reprodutibilidade total em que os códigos executáveis e dados estão disponíveis. O compartilhamento destes recursos (dados, códigos e documentação) é viabilizado por ferramentas colaborativas, como os notebooks, que são ambientes computacionais estruturados como listas de células de entrada e saída que podem conter códigos, textos, gráficos e outros elementos. Sendo assim, toda a sequência de passos é registrada, facilitando a reprodutibilidade interativa e detalhada (DANCHEV, 2021; MOOERS et al., 2021). No caso dos notebooks web, o processamento dos dados é feito na nuvem e a ferramenta não requer configuração de uso (QUARANTA et al., 2021).

Nesta pesquisa, foi usada a ferramenta Google Colab<sup>27</sup>, um ambiente de notebook acessível a partir do navegador web e uma conta gratuita do Google, sem a necessidade de instalação de softwares no computador do usuário e do controle de versão de bibliotecas. Os arquivos referenciados no notebook do Colab ficam armazenados no Google Drive<sup>28</sup>, que é o ambiente virtual de armazenamento de arquivos, e podem ser compartilhados com outros usuários. Os notebooks foram estruturados nas linguagens Python e R, e referem-se à aquisição e preparação dos dados, pré-processamento, análise exploratória dos dados e modelos. A sequência dos processamentos é apresentada na Figura 15. Todos os dados e códigos estão disponíveis na plataforma de hospedagem GitHub<sup>29</sup>.

Na computação, uma biblioteca é uma coleção de arquivos, programas, rotinas, scripts ou funções que podem ser referenciados no código de programação. Uma vantagem é que, se o usuário precisar escrever um programa complexo, pode usar a biblioteca em vez de implementar o algoritmo do zero. A linguagem Python dispõe, além da sua biblioteca padrão<sup>30</sup>, de diversas bibliotecas para funções variadas, sendo algumas de uso mais comum para ciência de dados e visualização e outras mais específicas para lidar com dados geoespaciais. A lista das bibliotecas Python e R usadas nesta pesquisa, módulos e suas respectivas funções estão no Anexo I.

---

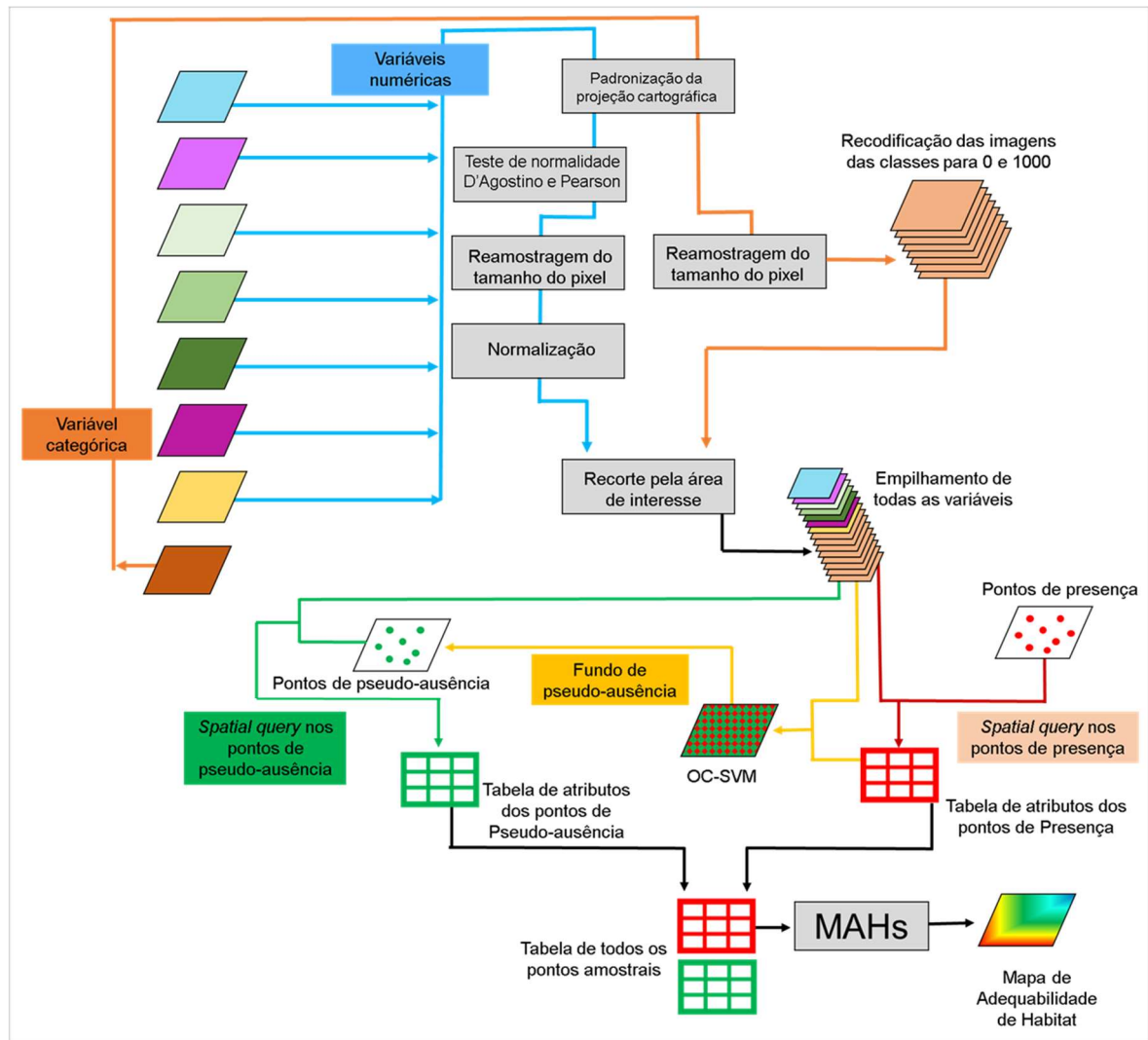
<sup>27</sup> Google Colab. Disponível em <<https://colab.research.google.com/>>. Acesso em 16 abr. 2022.

<sup>28</sup> Google Drive. Disponível em <<https://drive.google.com/>>. Acesso em 16 abr. 2022.

<sup>29</sup> GitHub. Disponível em <<https://github.com/silvia-jesus/>> a partir da data de depósito da tese no Repositório Institucional da UFSCar.

<sup>30</sup> A Biblioteca Padrão do Python. Disponível em <<https://docs.python.org/pt-br/3/library/>>.

Figura 15 - Fluxograma de trabalho



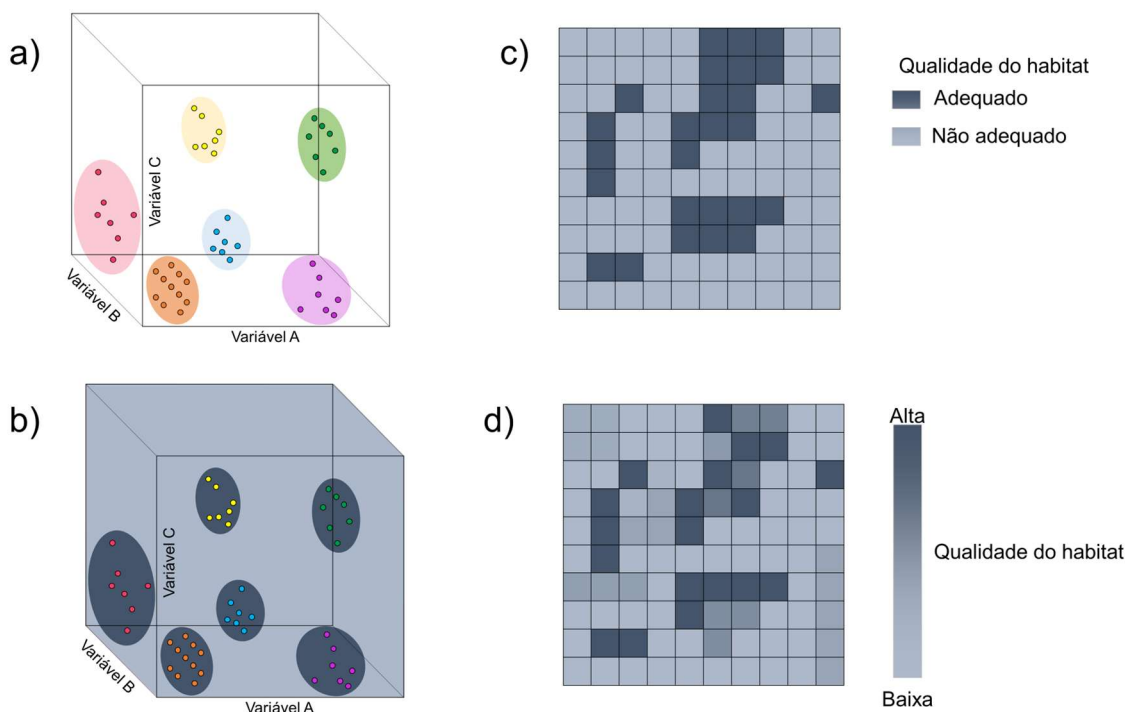
### 3.3 MODELAGEM DE ADEQUABILIDADE DE HABITAT

Usualmente, os MAHs consideram a ocorrência de uma única espécie ou gênero e, a partir da amplitude da variável em que aquela espécie ocorre, são definidos envelopes ambientais no espaço de atributos (Figura 16a), onde o modelo entende como a caracterização do nicho fundamental e aplica a função para toda a área de estudo. Aqui é considerada uma abordagem em que os grupos de envelopes ambientais definem áreas adequadas ou inadequadas (Figura 16b) e podem ser representados em dados raster com apenas dois valores de pixels (Figura 16c); no entanto, como são usados métodos de regressão, o resultado não é a informação binária (adequado / não adequado) mas um gradiente de adequabilidade (Figura 16d). Então, se o local é apropriado o suficiente para abrigar a espécie guarda-chuva que modela um

determinado pixel, então outras espécies co-ocorrentes podem ser conservadas nesse local. O propósito da modelagem não é mapear a distribuição das espécies, mas verificar se o local é adequado para a manutenção das espécies consideradas. Considerar várias espécies na modelagem aumenta o tamanho de amostras para treinamento e teste, mesmo após a filtragem espacial, feita para remover registros muito próximos e evitar a autocorrelação espacial na modelagem.

Um dos aspectos relevantes para a modelagem da adequabilidade de habitat é o tamanho amostral ou o tamanho da amostra composta por presenças e pseudo-ausências que será usada para treinar o modelo. Diversos estudos têm comparado o poder de predição de modelos variados e sua resposta em relação a um tamanho amostral absoluto (STOCKWELL; PETERSON, 2002; HERNANDEZ et al., 2006; BEAN et al., 2012) e de modo geral, mostram que quanto maior o tamanho amostral, melhor a capacidade preditiva. Esta é uma vantagem de considerar diversas espécies na modelagem de adequabilidade de habitat, pois isso torna o tamanho amostral mais significativo que ao usar apenas uma espécie.

Figura 16 - Abordagem do Modelo de Distribuição de Espécies aplicada nesta pesquisa



a) Representação de envelopes ambientais no espaço de atributos tridimensional; b) Representação de envelopes ambientais generalizados no espaço de atributos tridimensional; c) Representação rasterizada da classificação de adequabilidade (adequado e não adequado); d) Gradiente de adequabilidade resultante de um modelo de regressão.

### 3.4 VARIÁVEIS PREDITORAS

Para Marques Neto (2018), o clima é a principal variável reguladora do funcionamento dos sistemas ambientais. Temperatura e precipitação interferem nos padrões zoogeográficos, tanto indiretamente, regulando a fenologia das plantas que servem de habitat e recursos alimentícios para muitos animais, quanto diretamente, dadas as exigências fisiológicas das espécies animais para manter seu metabolismo. As variações da composição florística em função da altitude também têm implicações na fauna, tanto pela vegetação quanto pela diversidade de geoambientes: quanto maior a altitude, menores são a produtividade primária e as temperaturas médias. Segundo o autor, essa relação pode implicar na existência de espécies endêmicas decorrentes da especiação alopátrica nos geossistemas montanhosos, ou, ainda, na distribuição disjunta de espécies congêneres. Por outro lado, muitas espécies animais têm maior capacidade de dispersão e valem-se da vagueabilidade para transpor barreiras. O movimento da fauna está relacionado à estrutura da paisagem, que pode favorecer ou impossibilitar a locomoção (GUISAN et al., 2017; MARQUES NETO, 2018).

Os modelos consideraram como variáveis preditoras: Precipitação, Temperatura Mínima, Temperatura Média, Temperatura Máxima, Precipitação, Luzes Noturnas Artificiais, Altitude, Inclinação do Terreno e Uso e Cobertura da Terra. A seguir é descrita a importância dessas variáveis como fatores limitantes no padrão de distribuição de espécies, seja como um elemento que estabelece a ocorrência da espécie em um determinado habitat, ou como uma condição para a persistência das populações. De modo geral, as variáveis preditoras têm implicações sobre a disponibilidade de recursos, ou interferem mais diretamente nos indivíduos, como na manutenção de processos fisiológicos e comportamentais. São apresentados alguns exemplos da literatura que relacionam as consequências da exposição das espécies a um gradiente ambiental.

#### 3.3.1 Precipitação

A precipitação está indiretamente relacionada à disponibilidade de recursos (MAESTRI; MARINHO, 2014). A água é um recurso essencial para a manutenção fisiológica de todos os organismos e compõe alto percentual da massa corpórea, especialmente nos vertebrados terrestres. Em todos os mamíferos terrestres, a ingestão de água se dá por três vias: água ingerida nos alimentos consumidos, água metabólica resultante da oxidação de macronutrientes e água

bebida. A água ingerida é eliminada por meio da excreção, ingestão ou evaporação através da pele ou do trato respiratório. É esperado que os mamíferos herbívoros obtenham maior volume de água por meio de sua dieta do que os onívoros e carnívoros, visto que os itens vegetais têm uma composição de água maior do que a dos animais. Porém, os elementos vegetais podem variar quanto à composição química e influenciam na sua confiabilidade como fonte de água, o que interfere na necessidade dos animais de beber água (CHAVES et al., 2021).

Altos volumes de precipitação são associados a taxas mais altas de reprodução (DUBOST; HENRY, 2017) e de mudança na densidade populacional de mamíferos (MAESTRO; MARINHO, 2014). Kay e Madden (1997) encontraram correlações positivas significativas entre a precipitação e o número de espécies de mamíferos em várias guildas (mamíferos não voadores, frugívoros, consumidores primários, arborícolas, primatas e espécies com peso corporal entre 100 gramas a 10 quilogramas). Udy et al. (2021) concluíram que, nos Neotrópicos, a riqueza máxima de espécies ocorre em faixas de precipitação ligeiramente abaixo dos valores máximos, o que os autores atribuem à inclusão de áreas de baixa umidade que são pobres em espécies nas amostras que apresentam faixas de precipitação muito altas. Todavia, Dambros et al. (2022) compararam a distribuição de roedores e marsupiais na Mata Atlântica e concluíram que a riqueza de espécies aumenta principalmente em áreas com baixa precipitação e grandes fragmentos florestais.

### **3.3.2 Temperatura**

A temperatura tem efeito sobre aspectos fisiológicos e comportamento dos animais (MOTA-ROJAS et al., 2021). Os mamíferos são endotérmicos, isto é, são capazes de manter a temperatura corporal elevada e estável em condições flutuantes. Esta característica possivelmente foi fundamental para o sucesso ecológico e evolutivo de mamíferos (LOVEGROVE, 2012). A endotermia, no entanto, é energeticamente onerosa: um mamífero endotérmico consome cinco vezes mais alimento do que um tetrápode ectotérmico de tamanho semelhante (BASTOS et al., 2021). A regulação da temperatura corporal em endotérmicos envolve processos dinâmicos que levam a perturbações em outros processos fisiológicos, sendo uma prioridade em relação a outras funções fisiológicas (HANSEN, 2009). Em condições de baixa disponibilidade de recursos (como, por exemplo, em casos de perda de habitat e redução do dossel florestal), os mamíferos podem interromper temporariamente essa termorregulação (HANSEN, 2009; DAUSMANN, 2014; BASTOS et al., 2021; GEISER; COOPER, 2023).

Thonis et al. (2020) observaram que as características termorreguladoras de *Tupaia minor* (Ordem Scandentia) permitem a economia de energia e água em temperaturas bem superiores às temperaturas ambientais atuais e concluíram que os endotérmicos tropicais podem ser pouco sensíveis ao clima. No entanto, outros trabalhos relatam mudanças comportamentais e fisiológicas de mamíferos em situação de variabilidade térmica. Hansen (2009) explica que o estresse térmico tem consequências na reprodução dos mamíferos, desde a produção de gametas, desenvolvimento embrionário e fetal, até a lactação, o que é resultado da hipertermia associada ao estresse por calor ou dos ajustes fisiológicos para regular a temperatura corporal. Attias et al. (2018) modelaram o comportamento ao longo do dia de duas espécies de clamiforídeos, *Euphractus sexcinctus* (tatu-peba) e *Tolypeutes matacus* (tatu-bola). Os autores constataram que a atividade de *E. sexcinctus* durante o dia diminui à medida que a temperatura do ar aumenta. Por outro lado, *T. matacus* apresenta menor variação e mantém um padrão de atividade predominantemente noturno, embora seu período de atividade seja de menor duração e atinja o pico no início do dia, à medida que a temperatura do ar diminui. Beaudrot et al. (2019) constataram que a distribuição de mamíferos tropicais está associada a mudanças na temperatura local, mas que a composição da comunidade interfere nessa relação, ou seja, as interações entre as espécies influenciam a dinâmica de colonização ou extinção de mamíferos tropicais. Nord e Giroud (2020) explicam que as condições térmicas no início da vida ajustam os fenótipos termorreguladores, isto é, vivenciar um ambiente térmico diferente daquele ao qual a população está adaptada pode restringir a qualidade fenotípica. Ou seja, o desvio sustentado do ambiente térmico ao qual a população está adaptada parece restringir a qualidade fenotípica. Dambros et al. (2022) constataram que, na Mata Atlântica, espécies com maior razão cauda/corpo e arborealidade são encontradas em temperaturas mais quentes.

### 3.3.3 Luzes Noturnas Artificiais

As espécies são classificadas como diurnas ou noturnas dependendo da proporção do tempo de repouso e atividade ao longo do dia. A identificação desse nicho temporal é essencial para avaliar a pressão do ecossistema sobre as populações. As variações fisiológicas ao longo do dia e sazonais são reguladas pela melatonina, da qual a síntese e a liberação são bloqueadas pela luz. No entanto, as Luzes Noturnas Artificiais (LNAs) interferem nos ciclos naturais de luz escura e interrompem os ciclos circadianos e sazonais, alterando a fisiologia e o comportamento de mamíferos, principalmente de pequenos roedores. Alguns habitats enfrentam a iluminação

contínua durante toda a noite (por edifícios, estradas e rodovias, sistemas de segurança, publicidade e iluminação decorativa) ou a luz intermitente como a luz da estrada e os faróis dos veículos (Singhal et al., 2021). A antropização contínua dos ecossistemas e a expansão urbana alteram de forma significativa os habitats silvestres, especialmente quanto à poluição luminosa e ruído de tráfego (HOFFMANN et al., 2018; KHAN et al., 2020; ISSAD et al., 2021).

As LNAs modificam o componente abiótico do ecossistema, como um elemento chave para o aquecimento global dada a produção de gases do efeito estufa por meio da emissão das luzes (Singhal et al., 2021). Issad et al. (2021) elencam uma série de efeitos nocivos da LNA em roedores: alteração no padrão de locomoção, temperatura corporal e hormônios, aceleração do envelhecimento, gênese tumoral, desenvolvimento de transtorno depressivo e ansiedade. Hoffmann et al. (2018) complementam que mesmo baixos níveis de poluição luminosa, como os que ocorrem em áreas remotas, podem levar a mudanças no comportamento animal e podem ter consequências para as interações das espécies.

### **3.3.4 Altitude**

A heterogeneidade da topografia implica na maior diversidade de componentes bióticos em decorrência das condições microclimáticas que são variáveis nas pequenas áreas geográficas. Por exemplo, a captação da radiação solar é maior nos topos, enquanto os vales são usualmente mais secos; como resultado, ocorre um mosaico de manchas de vegetação às quais são associadas faunas distintas (NOGUERA-URBANO, 2017).

Pequenas mudanças na altitude implicam em grandes variações no microclima e na distribuição de recursos; a altitude é, portanto um preditor importante na modelagem de adequabilidade de habitat e sua omissão pode reduzir o desempenho do modelo (OKE; THOMPSON, 2015). Há pesquisas que apontam essa relevância da altitude na abundância e a composição das comunidades de pequenos mamíferos (PERCEQUILLO et al., 2017; RICHARD et al. 2022). Geise et al. (2004) constataram que existe um padrão de ordenação relacionado à altitude nas comunidades de marsupiais, primatas e roedores, com extremos de riquezas de espécies em altitudes entre 500 m e 1.500 m acima do nível do mar na floresta montana. Udy et al. (2021) encontraram um padrão em que, nos Neotrópicos, a riqueza de espécies tem correlação negativa com a altitude.



### 3.3.5 Inclinação do terreno

Além da altitude, a variação do relevo influencia os padrões climáticos e, portanto, está relacionada com a composição da vegetação e a disponibilidade de alimentos, que podem variar significativamente ao longo dos gradientes de altitude (MENA et al., 2020). Em macroescala, o relevo acidentado e outras características topográficas importantes têm grande importância na riqueza de espécies pela maior diversidade de habitats, efeitos altitudinais, formação de refúgios e outros fatores (MARES, 1997). A topografia influencia também o deslocamento e o uso do habitat pelos mamíferos terrestres, especialmente na época de chuvas, quando podem ser formadas ilhas nas várzeas, que podem ser usadas como sítio de alimentação e descanso (ALVARENGA et al., 2018). Arteaga e Venticinque (2008) constataram que os dasipodídeos constroem tocas preferencialmente em terrenos de baixa altitude e inclinados, provavelmente porque nessas áreas o esforço de escavação é menor e os corpos d'água estão mais próximos.

### 3.3.6 Uso e cobertura da terra

Para Corlett (2020), na região Neotropical, mais do que as condições climáticas atuais (temperatura média anual e precipitação anual), os impactos humanos históricos (a partir de 1700 EC) e pré-históricos têm maior influência na estrutura da comunidade de mamíferos. Torres-Romero e Giordano (2022) também apontam que as alterações na paisagem durante o Antropoceno alteraram fortemente os padrões espaciais de pequenos carnívoros no mundo todo. Existe ainda a seriedade da exposição da fauna a pesticidas, produtos farmacêuticos, plásticos, material particulado associado a incêndios florestais e nanopartículas (CHAPMAN et al., 2022).

De Marco Jr. et al. (2020) encontraram uma relação espacial entre áreas climaticamente adequadas para mamíferos e paisagens altamente impactadas, com menos de 50% da vegetação remanescente. Os autores afirmam que a agricultura e a infraestrutura de transporte são as atividades que mais afetam as paisagens no Cerrado. Dotta e Verdade (2011) constataram que áreas de pastagem apresentam menor riqueza de espécies e os canais a maior frequência de ocorrência, sendo que a maior parte das espécies encontradas é constituída por espécies que se adaptam às alterações ambientais. Os habitats florestais são essenciais para a regulação térmica dos animais, especialmente para aqueles de baixo metabolismo como *Myrmecophaga tridactyla* (Myrmecophagidae) e são importantes fontes de recursos em períodos de escassez nos campos

abertos (DESBIEZ, 2009). Entretanto, a supressão dos fragmentos florestais interrompe estas funções do ecossistema.

Além da riqueza e abundância, a antropização da paisagem também interfere no comportamento; a perda de vagilidade, por exemplo, pode implicar na exclusão de espécies de longo alcance, além de prejudicar a persistência de metapopulações, as interações predador-presa, ciclagem de nutrientes e interferir nos padrões de transmissão de doenças. Essa redução na capacidade de deslocamento pode ser resultado da mudança de habitat e da fragmentação, e também da redução de recursos necessários para a movimentação, como fontes de alimento e de água (TUCKER et al., 2022).

Outros riscos, além das alterações na paisagem e da perda de habitat, são o desmatamento e a degradação, incêndios, cercas e poços artificiais, caça ilegal e poluição multissensorial da luz noturna artificial e ruídos (DESBIEZ, 2009; DOMINONI et al., 2020; SCHUTTLER et al., 2021).

### 3.5 FONTE E CARACTERÍSTICAS DOS DADOS

#### 3.5.1 Variáveis preditoras

Os dados obtidos por sensores remotos orbitais podem ser usados em MAHs como alternativa aos métodos de campo que exigem mais recursos e tempo (WILSON et al., 2013). Além disso, os dados orbitais cobrem áreas extensas e inacessíveis e não interferem no objeto ou área de análise. Esta seção apresenta as características dos dados referentes às variáveis preditoras dos Modelos de Adequabilidade de Habitat gerados nesta pesquisa. A Tabela 6 resume as principais características das variáveis preditoras dos MAHs.

Tabela 6 - Características gerais dos dados referentes às variáveis preditoras dos MAHs

Variável	Fonte	Resolução espacial do dado original	Unidade	Tipo da variável
Precipitação anual	WorldClim <sup>31</sup>	0,5°	mm	numérica
Temperatura média anual			°C * 10	numérica
Temperatura máxima no mês mais quente				
Temperatura mínima no mês mais frio				
Luzes Noturnas Artificiais	Earth Engine Data Catalog <sup>32</sup>	463,83 metros	nanoWatts/cm <sup>2</sup> /sr	numérica
Uso e cobertura da terra	MapBiomas <sup>33</sup>	30 metros	Mapa temático em que cada nível de cinza representa uma classe da legenda	categórica
Altitude	Embrapa Territorial <sup>34</sup>	90 metros	metros	numérica
Inclinação do terreno	Derivado da altitude	90 metros	graus	numérica

### 3.5.1.1 Precipitação e temperatura

As variáveis climáticas, precipitação e temperatura, são provenientes do projeto WorldClim (FICKS; HIJMANS, 2017), que consiste em um banco de dados climáticos e meteorológicos globais. São dados interpolados de estações meteorológicas combinados a dados orbitais, que refinam a precisão da previsão, particularmente para áreas com baixa

<sup>31</sup> WordClim. Disponível em <<https://worldclim.org/>>. Acesso em 24 set 2021.

<sup>32</sup> VIIRS Stray Light Corrected Nighttime Day/Night Band Composites Version 1. Disponível em <[https://developers.google.com/earth-engine/datasets/catalog/NOAA\\_VIIRS\\_DNB\\_MONTHLY\\_V1\\_VCMSLCFG](https://developers.google.com/earth-engine/datasets/catalog/NOAA_VIIRS_DNB_MONTHLY_V1_VCMSLCFG)>. Acesso em 14 ago 2021.

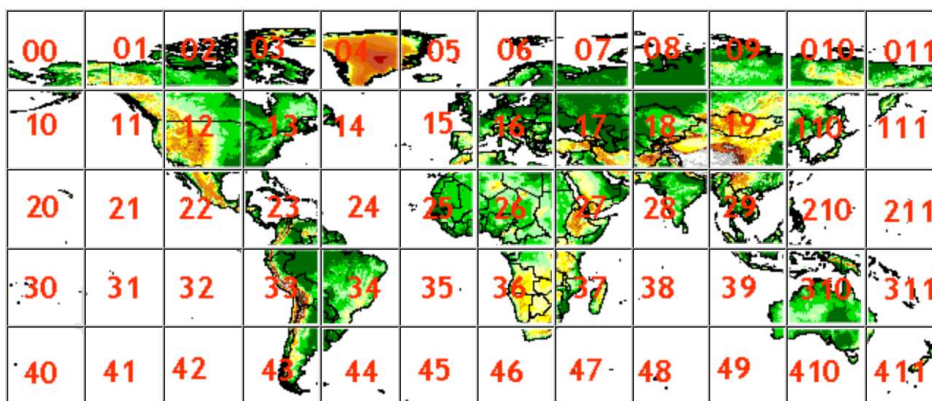
<sup>33</sup> MapBiomas. Disponível em <[mapbiomas.org](http://mapbiomas.org)>. Acesso em 01 out 2021.

<sup>34</sup> Brasil em Relevô. Disponível em <<https://www.cnpm.embrapa.br/projetos/relevobr/>>. Acesso em 12 nov 2020.

densidade de estações. O intervalo temporal para os dados da estação é de 1970 a 2000. Mais detalhes sobre a metodologia de geração dos dados podem ser encontrados em Weiss et al. (2014) e Ficks e Hijmans (2017).

São 19 variáveis bioclimáticas que estão disponíveis para as áreas terrestres globais, exceto a Antártida. Os dados estão no sistema de referência de coordenadas de latitude/longitude e o datum é WGS84. Os dados estão disponíveis em resoluções espaciais de 10min., 5min., 2,5min. e 30s. Aqueles com resolução de 30s estão disponíveis em blocos (*tiles*) de 30 x 30 graus organizados em uma grade como mostrado na Figura 17. O arquivo com as variáveis bioclimáticas contém 19 camadas de dados (Tabela 7). Nesta pesquisa foram consideradas as variáveis Bio1 (Temperatura média anual), Bio5 (Temperatura máxima do mês mais quente), Bio6 (Temperatura mínima do mês mais frio) e Bio12 (Precipitação anual). Foram descartadas as variáveis que envolvem mais de uma característica climática (Amplitude diurna média, Isothermalidade, Sazonalidade da Temperatura, Amplitude térmica anual, Sazonalidade da precipitação, Temperatura média do mês mais úmido, Temperatura média do mês mais seco, Precipitação do trimestre mais quente e Precipitação do trimestre mais frio). Para temperatura, foram selecionadas as variáveis relacionadas aos extremos.

Figura 17- Grade de blocos do WorldClim para dados com resolução de 30s



Fonte: WorldClim (2022).

Tabela 7 - Variáveis bioclimáticas do WorldClim

<b>Camada</b>	<b>Nome da variável</b>	<b>Camada</b>	<b>Nome da variável</b>
Bio1	Temperatura média anual	Bio11	Temperatura média do mês mais frio
Bio2	Amplitude diurna média	Bio12	Precipitação anual
Bio3	Isotermalidade	Bio13	Precipitação do mês mais úmido
Bio4	Sazonalidade da temperatura	Bio14	Precipitação do mês mais seco
Bio5	Temperatura máxima do mês mais quente	Bio15	Sazonalidade da precipitação
Bio6	Temperatura mínima do mês mais frio	Bio16	Precipitação do trimestre mais úmido
Bio7	Amplitude térmica anual	Bio17	Precipitação do trimestre mais seco
Bio8	Temperatura média do mês mais úmido	Bio18	Precipitação do trimestre mais quente
Bio9	Temperatura média do mês mais seco	Bio19	Precipitação do trimestre mais frio
Bio10	Temperatura média do mês mais quente	Fonte: WordClim ( <a href="https://worldclim.org/">https://worldclim.org/</a> )	

### 3.5.1.2 Luzes Noturnas Artificiais

As informações sobre Luzes Noturnas Artificiais é proveniente do dado *VIIRS Stray Light Corrected Nighttime Day/Night Band Composites Version 1* fornecido pelo *Earth Observation Group (EOG)*<sup>35</sup> e pode ser acessado pela plataforma Google Earth Engine<sup>36</sup>.

São imagens de radiância média mensal que usam dados da passagem noturna do sensor VIIRS (*Visible Infrared Imaging Radiometer Suite*), que é um dos instrumentos de imageamento do satélite NOAA-20<sup>37</sup>. Os dados estão em coordenadas geográficas e o datum é WGS84. A resolução espacial nominal é de 15 arcseg (MILLS et al., 2013).

### 3.5.1.3 Uso e cobertura da terra

Os dados de uso e cobertura da terra foram obtidos do Projeto de Mapeamento Anual da Cobertura e Uso do Solo do Brasil (MapBiomias), que é uma iniciativa de monitoramento colaborativo criada em 2015. As imagens utilizadas no mapeamento são das séries históricas dos satélites Landsat, com resolução espacial de 30 metros. O método de classificação das imagens orbitais pode ser conferido em MapBiomias (2022). A acurácia geral é de 74,9% para o Cerrado e 85,5% para a Mata Atlântica, considerando o nível 3 de mapeamento.

Foi considerado o mapa anual de uso e cobertura de 2020, coleção 6, que pode ser acessado pela plataforma Google Earth Engine.

### 3.5.1.4 Altitude

Os dados de altitude são do projeto “Brasil em Relevo”, coordenado pela Embrapa Territorial, que publicou os mosaicos da missão SRTM (*Shuttle Radar Topography Mission*) de todo o Brasil.

A missão SRTM (VAN ZYL, 2001) foi coordenada pela *National Imagery and Mapping Agency* (NIMA) e pela *National Aeronautics and Space Administration* (NASA), com objetivo

---

<sup>35</sup> Earth Observation Group (EOG). Disponível em <<https://eogdata.mines.edu/products/vnl/>>. Acesso em 14 ago 2021.

<sup>36</sup> Google Earth Engine. Disponível em <<https://code.earthengine.google.com/>>. Acesso em 14 ago 2021.

<sup>37</sup> JPSS-1/NOAA-20. Disponível em <<https://www.eoportal.org/web/eoportal/satellite-missions/n/noaa-20>>. Acesso em 14 ago 2021.

de obter um Modelo Digital de Elevação da superfície continental entre os paralelos 56°S e 60°N. Os dados foram adquiridos por interferometria SAR (*Synthetic-aperture radar*) em um sistema de radar a bordo da nave espacial Endeavour, em fevereiro de 2000. Os dados são organizados em blocos de 1° x 1° com 3 arcseg de resolução e têm acurácia vertical absoluta de 16 metros.

O projeto Brasil em Relevo organiza as imagens de acordo com a articulação sistemática de cartas na escala 1:250.000 do IBGE (CARVALHO, 1980).

### 3.5.2 Limites das UCs federais e estaduais

Os limites das UCs federais e estaduais foram concedidos pelo Instituto Socioambiental (ISA)<sup>38</sup> em 2018 por correio eletrônico. O ISA compila dados do ICMbio e dos órgãos estaduais responsáveis pelo gerenciamento das UCs nesta esfera. Os dados foram fornecidos em formato Esri Shapefile, no Sistema de Coordenadas de Referência GCS SIRGAS 2000. Os arquivos (UCs estaduais e federais) compreendem os dados vetoriais das UCs terrestres e marinhas de todo o Brasil, com exceção das RPPNs (Reserva Particular do Patrimônio Natural). As tabelas de atributos contêm o nome da UC, área em hectares, uso, link para as informações da UC no site do ISA e fonte.

Por conta da diferença de escala e generalização cartográfica, os limites das UCs nem sempre são coincidentes com os limites do Estado, mesmo das unidades estaduais. Para as unidades que fazem intersecção com a área de estudo, foram aplicados os seguintes critérios: a) UCs estaduais com pelo menos 40% da sua extensão na área de estudo; b) UCs federais com pelo menos 25% da sua extensão na área de estudo. Este procedimento elimina as UCs que têm apenas porções pequenas no Estado de São Paulo e também as UCs marinhas.

---

<sup>38</sup> Instituto Socioambiental. Fwd: [INSTITUTO SOCIOAMBIENTAL] Mensagem enviada pelo formulário de contato do website do ISA. Mensagem recebida por: <silviacrisj@gmail.com> em 11 abr. 2018. .

### 3.5.3 Limites do estado de São Paulo

Os limites do estado de São Paulo, versão atualizada em 2021, foram acessados pelo FTP (*File Transfer Protocol*) do IBGE<sup>39</sup>.

Os polígonos insulares foram removidos, sendo considerada como área de interesse apenas a porção continental do território paulista e o município de Ilhabela.

### 3.5.4 Ocorrência de espécies

Os dados sobre a ocorrência de espécies foram extraídos do GBIF - *Global Biodiversity Information Facility*, que é uma rede e infraestrutura de dados internacional que funciona como um sistema centralizado para que pesquisadores e instituições possam contribuir com seus dados para um banco de dados maior e de fácil acesso. O GBIF é a maior rede de dados sobre biodiversidade (OTEGUI et al., 2013) e, além de registrar espécimes de coleções de museus e herbários, o GBIF também coleta dados de outras fontes, como o iNaturalist<sup>40</sup>, que é uma plataforma de ciência cidadã para registro e compartilhamento de observações biológicas. O banco de dados possui coleções biológicas, listas de espécies, dados espaciais, eventos de amostragem e imagens multimídia, georreferenciadas ou não.

É possível acessar o banco de dados pelo site do GBIF ou por meio de uma API (*Application Programming Interface*) disponível para as linguagens Python<sup>41</sup> e R<sup>42</sup>. Trata-se de um dado tabular com 165 colunas que incluem informações sobre a taxonomia da espécie, localização, data e responsável pela coleta do dado. Os dados georreferenciados têm as colunas de longitude e latitude e podem ser espacializados em um Sistema de Informações Geográficas.

Em princípio, o compartilhamento dos dados segue os padrões *Biodiversity Information Standards* (TDWG - *Taxonomic Databases Working Group*)<sup>43</sup>, o *Darwin Core Controlled*

---

<sup>39</sup> IBGE: Organização do Território. Disponível em <[https://geoftp.ibge.gov.br/organizacao\\_do\\_territorio/malhas\\_territoriais/malhas\\_municipais/municipio\\_2021/UFs/SP/SP\\_UF\\_2021.zip](https://geoftp.ibge.gov.br/organizacao_do_territorio/malhas_territoriais/malhas_municipais/municipio_2021/UFs/SP/SP_UF_2021.zip)>. Acesso em 20 abr. 2021.

<sup>40</sup> iNaturalist. Disponível em <<https://www.inaturalist.org/>>. Acesso em 15 fev. 2021.

<sup>41</sup> pygbif: Python client for the GBIF API. Disponível em <<https://www.gbif.org/tool/OlyoYyRbKCSCkMKIi4oIT/pygbif-gbif-python-client>> . Acesso em 11 fev. 2021.

<sup>42</sup> rgbif: An interface to the GBIF API for the R statistical programming environment. Disponível em <<https://www.gbif.org/tool/81747/rgbif>>. Acesso em 11 fev. 2021.

<sup>43</sup> Biodiversity Information Standards (TDWG). Disponível em <[www.tdwg.org](http://www.tdwg.org)>. Acesso em 15 fev. 2021.



*Vocabulary Scheme*<sup>44</sup> (WIECZOREK et al., 2012) e, para os metadados, o *Ecological Metadata Language* (EML)<sup>45</sup>. No entanto, os dados publicados, eventualmente, apresentam inconsistências nas coordenadas de localização e algumas das observações das espécies são replicadas. Os atributos do dado estão no Anexo II.

O GBIF tem um nó brasileiro que é o Sistema Brasileiro de Informações sobre Biodiversidade. Este sistema integra, compartilha, publica e sintetiza dados de biodiversidade gerados e processados no âmbito do Ministério do Meio Ambiente e suas instituições afins (Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis - Ibama, Instituto Chico Mendes de Conservação da Biodiversidade - ICMBio, Jardim Botânico do Rio de Janeiro - JBRJ, Agência Nacional de Águas - ANA e Serviço Florestal Brasileiro - SFB). A arquitetura básica de interoperabilidade inclui um provedor de dados coletando, armazenando e disponibilizando dados de biodiversidade e seus metadados (SILVA et al., 2015).

Os dados podem ser aplicados à conservação da biodiversidade (LIMA et al., 2020; MORAES et al. 2020; CONCEIÇÃO et al., 2022), mudanças climáticas (EVANGELISTA-VALE et al., 2021; FREITAS-OLIVEIRA et al., 2021; LIMA et al., 2022), saúde (AMADO et al., 2021; HEINRICH et al., 2021; LI et al., 2022), sequenciamento de DNA e metagenômica (CHUA et al., 2021; KENNEDY et al., 2022; ROGER et al. 2022), agricultura e segurança alimentar (GRAEFE et al., 2013; VINCENT et al., 2013; VINCENT et al., 2019) e a projetos no setor privado (PATARROYO CAMARGO; OCHOA SANABRIA, 2021; LIZCANO et al., 2022).

No GBIF, o termo “ocorrência” refere-se a qualquer registro no conjunto de dados publicados. O status de ocorrência de uma espécie em um determinado local pode ter os valores “presente”, “comum”, “rara”, “irregular”, “duvidosa”, “ausente” ou “excluída”<sup>46</sup>. Tanto na API como na busca pelo portal, há a opção para a seleção do status de ocorrência. Para esta pesquisa, foram consideradas apenas ocorrências que constam como “presença”. Sendo assim, aqui “ocorrência” e “presença” assumem o mesmo significado.

---

<sup>44</sup> List of Darwin Core terms. Disponível em <<https://dwc.tdwg.org/list/>>. Acesso em 15 fev. 2021.

<sup>45</sup> Ecological Metadata Language (EML). Disponível em <<https://eml.ecoinformatics.org/>>. Acesso em 15 fev. 2021.

<sup>46</sup> Darwin Core Vocabulary: Occurrence Status GBIF Vocabulary. Disponível em <[http://rs.gbif.org/vocabulary/gbif/occurrence\\_status/](http://rs.gbif.org/vocabulary/gbif/occurrence_status/)>. Acesso em 11 nov. 2022.

## 3.6. PROCESSAMENTO DOS DADOS

### 3.6.1. Geração dos planos de informação das variáveis preditoras

#### 3.6.1.1 Variáveis climáticas (WorldClim)

Os dados do WorldClim foram acessados pela função *'getData'* da biblioteca raster do R. É necessário fornecer o tipo de variável e a resolução espacial, sendo configuradas, respectivamente, as variáveis bioclimáticas derivadas da temperatura mínima, média e máxima e precipitação, e resolução de  $0,5^\circ$ <sup>47</sup>. Os argumentos relacionados às coordenadas geográficas (lon = -48, lat = -22), necessários para os dados nesta resolução, remetem ao grid *'bio\_34'* (Figura 17). Os dados em formato raster foram delimitados por um buffer<sup>48</sup> de  $1^\circ$  em relação aos limites da área de interesse.

#### 3.6.1.2 Luzes Noturnas Artificiais

Os dados de Luzes Noturnas Artificiais (LNA) foram acessados pela API *earthengine-api* que permite a interação do Google Earth Engine (GORELICK et al., 2017) usando a linguagem Python. Uma *ImageCollection* é um conjunto de imagens que pode ser acessado pelo catálogo de dados do Google Earth Engine. Para carregar a coleção de LNA, a função *ee.ImageCollection* requer o nome da coleção, o filtro temporal e a banda espectral. Foram consideradas todas as imagens adquiridas no período de 01 de janeiro a 31 de dezembro de 2020 do dado *'NOAA/VIIRS/DNB/MONTHLY\_V1/VCMSLCFG'*, banda *'avg\_rad'*. Então foi calculada a média de cada posição na imagem (pixel a pixel) para compor uma imagem única de luzes noturnas artificiais.

#### 3.6.1.3 Altitude

As imagens da missão SRTM que cobrem a área de interesse foram acessadas na página do projeto Brasil em Relevo e armazenadas no Google Drive, ou seja, não foram obtidas por uma requisição em rede pelo próprio código. As imagens que cobrem a área de interesse com a

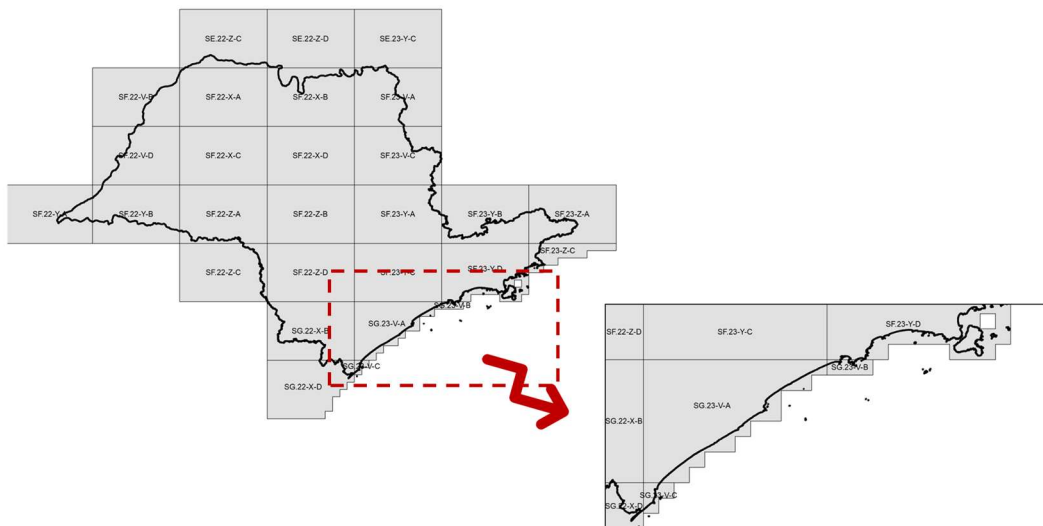
---

<sup>47</sup> WorldClim: Data format. Disponível em <<https://worldclim.org/data/v1.4/formats.html>>. Acesso em 23 abr. 2022.

<sup>48</sup> Buffer é uma região em torno de um elemento vetorial (pontos, linhas ou polígonos) a partir de uma certa distância, útil em análises de proximidade geográfica.

respectiva nomenclatura das folhas estão representadas na Figura 18, que mostra que nem todas as ilhas na região costeira do estado de São Paulo são cobertas pelos dados SRTM. O mosaico dos blocos de imagens foi construído pela função ‘merge’ da biblioteca ‘Rasterio’<sup>49</sup>.

Figura 18 - Folhas na escala 1:250.000 que cobrem o estado de São Paulo



Fonte: IBGE (2021)<sup>50</sup>. Organizado pela autora.

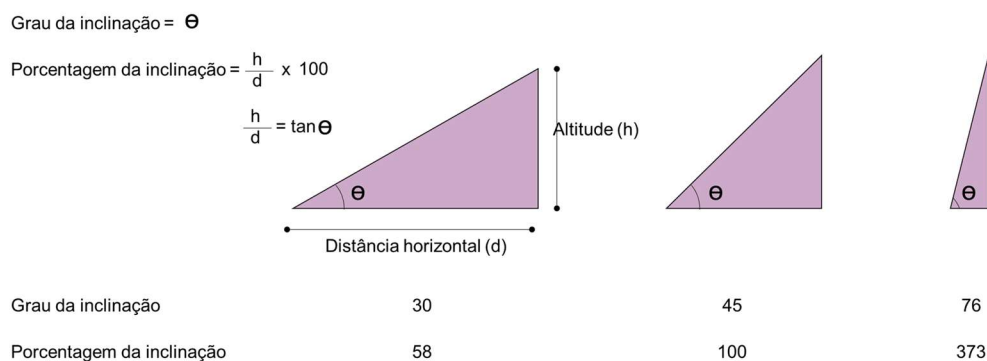
#### 3.6.1.4 Inclinação do terreno

A inclinação do terreno é a relação entre a distância horizontal no terreno e a altitude e pode ser gerada a partir do modelo de elevação SRTM. Usualmente os algoritmos de geração de inclinação do terreno calculam a variação entre um pixel e seus vizinhos. Quanto maior o valor da inclinação, que pode ser representada em graus ou porcentagem, mais íngreme é o terreno (Figura 19).

<sup>49</sup> Rasterio: Access to geospatial raster data. Disponível em <<https://rasterio.readthedocs.io/>>. Acesso em 25 abr. 2022.

<sup>50</sup> Bases cartográficas contínuas - Brasil. Disponível em <<https://www.ibge.gov.br/geociencias/cartas-e-mapas/bases-cartograficas-continuas/15759-brasil>>. Acesso em 24 abr. 2022.

Figura 19 - Comparação dos valores de inclinação do terreno em graus e em porcentagem



Fonte: ESRI (2022)<sup>51</sup>. Adaptado pela autora.

Aqui foi considerada a inclinação em graus, calculada pela biblioteca RichDEM<sup>52</sup>.

### 3.6.1.5 Uso e cobertura da terra (MapBiomass)

Os dados de uso e cobertura da terra foram acessados pelo toolkit do MapBiomass<sup>53</sup> organizados no Google Earth Engine. Na Tabela 8 estão os parâmetros de configuração para acesso e download do dado. O resultado é armazenado diretamente no Google Drive.

### 3.6.2 Organização dos planos de informação das variáveis preditoras

Para viabilizar os processamentos baseados em distância, como cálculo de áreas e definição do tamanho do pixel, todos os planos de informação foram convertidos para a projeção South America Albers Equal Area (ESRI:102033) que é uma projeção plana e é adequada para grandes áreas.

<sup>51</sup> How Slope works. Disponível em <<https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-analyst/how-slope-works.htm>>. Acesso em 01 ago. 2022.

<sup>52</sup> RichDEM — High-Performance Terrain Analysis. Disponível em <[richdem.readthedocs.io](http://richdem.readthedocs.io)>. Acesso em 27 abr. 2022.

<sup>53</sup> Google Earth Engine: MapBiomass. Disponível em <<https://code.earthengine.google.com/ed86f98d0dade2e5143ff47f8f14ff66>>. Acesso em 01 jan. 2022.

Tabela 8 - Parâmetros de configuração no toolkit do MapBiomias

<b>Download parameters Script</b>	mapbiomas-user-toolkit-lulc.js
<b>Region</b>	mapbiomas-brazil
<b>Collection</b>	collection-6.0
<b>Tables</b>	state
<b>Properties</b>	name_en
<b>Features</b>	Sao Paulo
<b>Data Type</b>	Coverage
<b>Buffer</b>	None
<b>Layers</b>	Sao Paulo 2020

### 3.6.2.1 Variáveis numéricas

A padronização é o processo de transformação dos dados para que eles passem a apresentar uma distribuição normal, com média zero e desvio padrão igual a um. Isso é importante porque muitas técnicas de análise estatística assumem distribuição normal, como alguns testes de média. Por outro lado, a normalização – embora esse termo possa erroneamente remeter à distribuição normal - é usada para definir o processo de aplicação de uma transformação nos dados para produzir um conjunto de dados no intervalo 0-1.

Para alguns algoritmos de Aprendizado de Máquina, especialmente aqueles baseados em distância, as variáveis devem ser normalizadas ou padronizadas. A padronização é importante porque os recursos que possuem intervalos maiores – ou em maior grandeza - podem dominar os resultados da análise estatística. Ambas as técnicas têm suas desvantagens. A normalização é muito sensível a outliers; se eles estiverem presentes nos dados, a normalização dos dados certamente irá dimensionar os dados “normais” para um intervalo muito pequeno. Por outro lado, quando a padronização é usada, os novos dados padronizados não são limitados a um intervalo (KORDON et al., 2020; JAFARI, 2022). Por conta da autocorrelação espacial,

espera-se que não haja outliers nos planos de informação das variáveis preditoras (considerando toda a área de estudo e não apenas os pontos amostrais de Presença e de Pseudo-ausência).

Para verificar a normalidade das variáveis numéricas, foi aplicado o teste de D'Agostino e Pearson (D'AGOSTINO; PEARSON, 1973), adequado para amostras com  $n > 100$  e que avalia a distribuição com base no achatamento e na curtose do histograma, considerando a área de interesse como máscara. Para tanto, foi considerada a função '*normaltest*' da biblioteca *scipy.stats*<sup>54</sup>.

Em todas as variáveis numéricas foi aplicada a normalização Max-Min que transforma os dados linearmente. Assumindo que  $max_a$  e  $min_a$  são, respectivamente, os valores máximo e mínimo do atributo  $a$ , o cálculo para a normalização segue a Equação 1 (CASTRO; FERRARI, 2016). Para manter os dados no formato inteiro, a normalização foi feita no intervalo 1-1000.

$$a' = \frac{a - min_a}{max_a - min_a} \quad \text{Equação 1}$$

### 3.6.2.2. Variáveis categóricas

Uso e cobertura da terra é a única variável categórica, onde os valores dos pixels referem-se a uma classe da legenda do mapa e não estão associados a uma grandeza ou quantidade. Sendo assim, para cada valor foi gerada uma nova imagem com os valores 0 ou 1000, que representam, respectivamente, a ausência ou presença da classe naquele pixel. Por exemplo, para a classe 'Formação Florestal', o valor do pixel é igual a 3. Então foi gerado plano de informação para esta classe a partir da recodificação (Equação 2):

$$\begin{aligned} \text{Se } i = 3 \text{ então } i' &= 1000 \\ \text{Se } i \neq 3 \text{ então } i' &= 0 \end{aligned} \quad \text{Equação 2}$$

onde  $i$  é o valor do pixel na imagem original e  $i'$  é o valor do pixel na imagem recodificada.

<sup>54</sup> *scipy.stats.normaltest*. Disponível em <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.normaltest.html#r7bf2e556f491-1>. Acesso em 06 mai. 2022.

Não foi considerado o uso da operação codificação *one-hot (one-hot encoding)*<sup>55</sup> porque este método é mais apropriado para dados estruturados (tabelas, planilhas e bancos de dados) e aqui todas as variáveis são apresentadas no formato matricial.

### 3.6.2.3. Combinação das variáveis preditoras

Todos os planos de informação das variáveis preditoras passaram pelos processos de 1) reamostragem do tamanho do pixel; 2) recorte do dado matricial pelos limites da área de interesse. Entretanto, estes processamentos foram realizados em ordens distintas, caso a variável seja numérica ou categórica.

As variáveis numéricas foram padronizadas para o intervalo '0-1000', como descrito a seguir. A reamostragem do tamanho do pixel foi anterior à etapa da normalização para evitar que esse processo de reamostragem eventualmente resultasse em pixels com valores fora do intervalo definido.

Os pixels dos planos de informação da variável categórica (uso e cobertura da terra) foram recodificados para os valores 0 ou 1000, como descritos a seguir. Para contar com a informação espacialmente mais precisa, esta recodificação foi feita antes da reamostragem do pixel. Como os valores já estão na escala desejada, não foi necessária a normalização das variáveis categóricas.

O intervalo 0 a 1000 é apropriado para manter os arquivos das imagens das variáveis ambientais preditoras no formato inteiro, o que ocupa menos espaço de armazenamento.

Todas as variáveis preditoras foram agrupadas em um único arquivo, ou seja, foram empilhadas em 31 planos de informação, 7 das variáveis numéricas e 24 referentes às classes de uso e cobertura da terra.

### 3.6.3 Filtragem de dados de ocorrência de espécies

As amostras para MAHs devem ser independentes, o que significa que cada unidade amostral traz novas informações ao modelo. Essa premissa é frequentemente violada em

---

<sup>55</sup> sklearn.preprocessing.OneHotEncoder. Disponível em <<https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.OneHotEncoder.html>>. Acesso em 02 jul. 2022.

conjuntos de dados compilados de várias fontes, onde é comum encontrar registros duplicados em uma amostragem local ou local muito densa (ELITH; FRANKLIN, 2013). Quando esses conjuntos de dados de amostra com desvios causados por erros de medição são usados como entrada em modelos preditivos ou de classificação, como modelos de distribuição de espécies ou outros modelos espaciais, eles devem passar pela fase de limpeza de dados. No caso de coordenadas incorretas, os registros podem ser removidos ou novos valores podem ser inseridos. Por outro lado, no caso de observações iguais ou semelhantes (vários pontos na mesma coordenada ou muito próximos), esses registros podem enviesar ou sobreajustar o modelo.

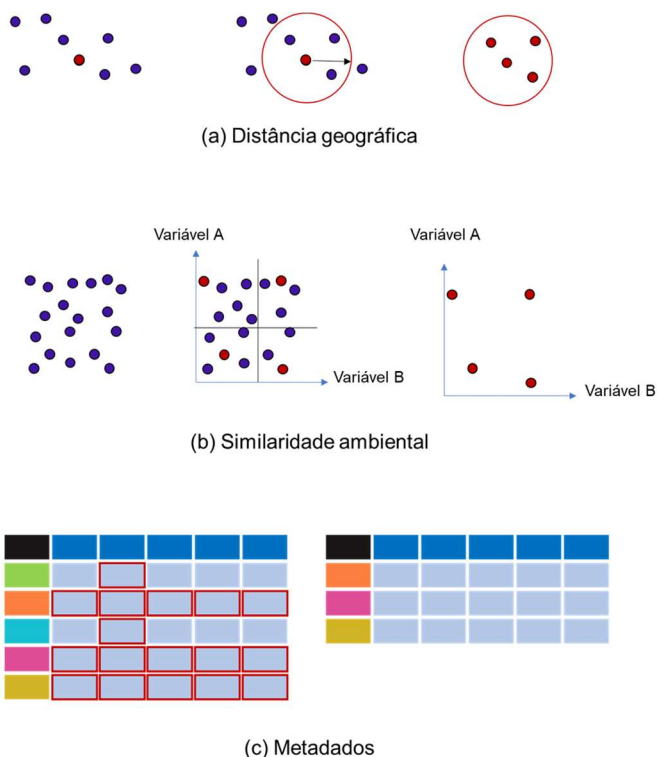
Os dados da ciência cidadã são cada vez mais usados para modelar as distribuições de espécies porque oferecem ampla cobertura espaço-temporal de observações locais (Steen et al., 2019). A aplicação de MAHs progrediu a partir da década de 1970 em parte porque os dados ambientais geoespaciais tornaram-se amplamente disponíveis (ELITH; FRANKLIN, 2013). As observações de espécies têm sido registradas e processadas em escalas e taxas sem precedentes como dados de programas de ciência cidadã, atlas, museus e pesquisas planejadas, dos quais a estrutura e os vieses de amostragem podem variar consideravelmente (FLETCHER et al., 2019). Esses dados geralmente não são coletados de forma sistemática e têm limitações como vieses espaciais, resoluções temporais e espaciais imprecisas e sub ou super-relato de espécies, o que aumenta a chance de que os modelos de previsão usando esses dados sejam comprometidos por viés e ruído (STEEN et al., 2019). Eles não são apenas de grande escala e distribuídos em diferentes instituições, mas também são heterogêneos e dinâmicos e muitas vezes são questionáveis quanto à sua precisão e integridade (ZHANG, 2017).

A literatura científica relata que a limpeza de dados interfere positivamente nos MAHs. Beck et al. (2014) avaliaram a adequabilidade de habitat para a borboleta eurásiana (*Aglais urticae*) e concluíram que a precisão diminuiu devido ao agrupamento espacial dos registros de distribuição no GBIF, ou seja, uma amostra parcial pode ser vantajosa para reduzir o impacto do viés espacial em MAHs. Gueta e Carmel (2016) concluíram que a limpeza de conjuntos de dados de mamíferos pode melhorar significativamente o desempenho do SDM em diferentes escalas; os autores explicam que a limpeza de dados é importante quando são usados grandes bancos de dados como o GBIF, que são propensos a erros que ocorrem na edição, publicação e/ou coleta dos dados.



A redução de viés e ruído pode ser alcançada por meio de técnicas baseadas em filtros que removem observações duvidosas, como discrepâncias, ou por técnicas estatísticas que ajustam modelos que abordam viés de amostragem e heterogeneidade de observação (STEEN et al., 2019). Uma estratégia para melhorar a confiabilidade desses dados é filtrar ou limpar os dados (VAN EUPEN et al., 2021), pela detecção de inconsistências (SERRA-DIAZ et al., 2017), removendo outliers (LI et al., 2020), por filtragem geográfica e/ou ambiental (INMAN et al., 2021), ou com base em metadados ou atributos de dados (VAN EUPEN et al., 2021; ZIZKA et al., 2020). A filtragem também pode ser baseada na distância geográfica, como feito por Aiello-Lammens et al. (2015) que desenvolveram um algoritmo para remover aleatoriamente registros que não atendem a uma distância mínima do vizinho mais próximo. A filtragem também pode ser feita por similaridade ambiental, removendo ocorrências repetidas em condições ambientais semelhantes, como fizeram Varela et al. (2014). Castellanos et al. (2019) explicam que os filtros geográficos e ambientais são conceitualmente equivalentes, pois ambos usam uma medida de distância para determinar o tamanho do filtro; no entanto, a medida de distância geralmente não leva em consideração questões biológicas. A Figura 20 ilustra estes principais tipos de filtros.

Figura 20 - Representação dos principais métodos de filtragem espacial



Fonte: Organizado pela autora.

O objetivo da filtragem é melhorar a qualidade dos dados para reduzir o viés e o erro. No entanto, existe a desvantagem de reduzir o tamanho amostral que pode influenciar no desempenho do modelo, sendo necessário considerar um trade-off entre qualidade dos dados e tamanho amostral (VAN EUPEN et al., 2021). Parece não haver consenso sobre a importância do tamanho da amostra na precisão dos modelos; no entanto, essas contradições podem ser explicadas por diferenças nos dados e características do modelo (GÁBOR et al., 2020).

Como os MAHs são modelos preditivos, a amostra (pontos de ocorrência da espécie) servirá como dados de treinamento e dados de teste. Se observações repetidas  $n$  vezes forem usadas como dados de treinamento, esse elemento terá um peso  $n$  vezes maior no ajuste de dados, e o modelo de previsão será enviesado para esses registros em detrimento de observações não repetidas.

Neste trabalho, com o objetivo de eliminar a redundância de dados de ocorrência de espécies, foi feita a filtragem em duas etapas: 1) filtragem por metadados; e 2) filtragem espacial com base no algoritmo de agrupamento DBSCAN (*Density-based spatial clustering of applications with noise* - Agrupamento Espacial de Aplicativos com Ruído Baseado em Densidade) (ESTER et al., 1996) e na definição do medóide dos clusters.

### 3.6.3.1 Filtragem por metadados

Foram organizados dois códigos para aquisição e filtragem dos dados de ocorrência de espécies. Um deles acessa os dados via API RESTful<sup>56</sup> (REST: *REpresentational State Transfer* - Transferência de Representação de Estado) do GBIF que trata os registros e os disponibiliza por meio de serviço web em formato JSON (*JavaScript Object Notation*) (OTEGUI; GURALNICK, 2016). A API fornece o serviço de busca de registros de ocorrência indexados pelo GBIF por meio de filtros definidos pelo usuário<sup>57</sup>. O outro código acessa o arquivo .csv descarregado do portal do GBIF<sup>58</sup> e armazenado no Google Drive. Com o intuito de manter o mesmo conjunto de dados ao longo do desenvolvimento e correção de todos os códigos, foi considerado o dado adquirido no portal do GBIF, dado que o acesso pela API fornece os

---

<sup>56</sup> Occurrence API. Disponível em <<https://www.gbif.org/pt/developer/occurrence>>. 22 abr. 2022.

<sup>57</sup> GBIF Vocabulary. Disponível em <<https://gbif.github.io/gbif-api/apidocs/>>. Acesso em 23 abr. 2022.

<sup>58</sup> GBIF.org (05 May 2022) GBIF Occurrence Download <https://doi.org/10.15468/dl.h5g93p>

conjuntos de dados atualizados até a data da busca. A lista de todos os parâmetros do GBIF está no Anexo II. Como filtro de busca, foram usados os parâmetros descritos na Tabela 9.

Tabela 9 - Parâmetros definidos na busca de registros de ocorrência indexados no GBIF

Parâmetro	Descrição	Valor
classKey	Chave de classificação de classe	359 (mamíferos)
hasCoordinate	Limita as pesquisas a registros de ocorrência que contêm um valor em latitude e longitude	true (limita a registros de ocorrência com valores de coordenadas)
hasGeospatialIssue	Inclui/exclui registros de ocorrência que contêm problemas espaciais	false (inclui apenas registros sem problemas espaciais)
country	O código do país em que a ocorrência foi registrada, conforme ISO-3166-1 <sup>59</sup> .	BR (Brasil)
gdam_gid	Um identificador geográfico GADM <sup>60</sup> em qualquer nível.	BRA.25_1 (Estado de São Paulo)

Para a variável ‘*BasisOfRecord*’, foram removidos os registros que constam como: amostra fóssil, amostra de material (amostras retiradas de outros espécimes ou do meio ambiente), amostra preservada, espécime vivo (animais manejados em um zoológico ou plantas cultivadas em um jardim ou estufa) e citação de material. Foram mantidos os registros de observação humana, observação por máquina, citação de material (referência em publicações acadêmicas).

Foram removidos os registros das ordens Chiroptera, Cetacea e Didelphimorphia. Os quirópteros e didelfídeos foram removidos por serem generalistas e serem encontrados em

<sup>59</sup> ISO 3166: Country Codes. Disponível em <<https://www.iso.org/iso-3166-country-codes.html>>. Acesso em 22 abr. 2022.

<sup>60</sup> Global Administrative Areas. Disponível em <<https://gadm.org/metadata.html>>. Acesso em 22 abr. 2022.

habitats diversificados. Os cetáceos foram removidos porque são consideradas nesta pesquisa apenas as espécies predominantemente terrestres. No nível de espécies, foram removidos os registros de *Arctocephalus tropicalis*, *Bos taurus*, *Bubalus bubalis*, *Canis lupus*, *Capra hircus*, *Equus caballus*, *Felis catus*, *Lepus europaeus*, *Mirounga leonina*, *Mus musculus* e *Sus scrofa* que são espécies exóticas e/ou domésticas, e de *Hydrochoerus hydrochaeris* por ser uma espécie generalista que habita qualquer lugar que tenha pasto e água. Também foram eliminados registros com duplicidade simultânea nos parâmetros de coordenadas geográficas e nome da espécie. Foi verificada então a grafia dos nomes das espécies.

O GBIF, além da informação sobre presença da espécie, também dispõe de informações sobre ausência. Estes registros foram removidos. Foram mantidas apenas as ocorrências registradas a partir de 2010, assumindo que não houve alteração nas variáveis ambientais a partir desta data. Por último, foram removidos os pontos que podem ter sido equivocadamente registrados no GBIF como sendo do estado de São Paulo, mas que espacialmente estão fora do seu perímetro.

### 3.6.3.2 Filtragem espacial

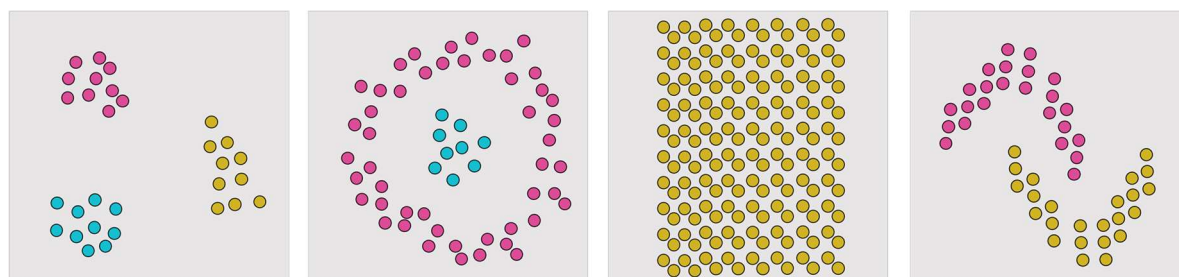
Para filtragem espacial dos pontos de ocorrência, foi considerado o algoritmo DBSCAN (*Density-based space clustering of applications with noise*) desenvolvido por Ester et al. (1996). O DBSCAN encontra agrupamentos de diferentes formatos e ruídos, com base na densidade de objetos no espaço, sendo a noção de densidade relacionada à quantidade de objetos dentro de um raio de vizinhança (ESTER et al., 1996; CASTRO; FERRARI, 2016).

Um agrupamento ou cluster é um conjunto de objetos próximos ou que satisfazem alguma relação espacial<sup>61</sup>. Pode ser baseado em densidade, isto é, uma região de alta densidade de pontos, que está separada de outras regiões de alta densidade, por outras regiões com baixa densidade de pontos (FACELI et al., 2021). A Figura 21 mostra esta definição de agrupamentos por densidade, onde os pontos coloridos representam regiões de alta densidade e o espaço cinza representa as regiões de baixa densidade de pontos; pontos próximos fazem parte do mesmo agrupamento.

---

<sup>61</sup> O termo “espacial” não se refere necessariamente ao sentido geográfico, mas pode estar associado a um espaço de atributos qualquer.

Figura 21 - Agrupamentos de pontos pelo algoritmo DBSCAN



Fonte: Organizado pela autora.

Este algoritmo foi escolhido porque a) é um método baseado em distância; b) não é necessário definir previamente o número de clusters; c) permite definir um número mínimo de pontos por cluster. Aqui o DBSCAN não é usado para agrupar dados com características semelhantes, mas para agrupar pontos de presença que estão espacialmente próximos com base em coordenadas geográficas (latitude e longitude). Registros espacialmente isolados não constituem outliers, e já estão filtrados espacialmente.

DBSCAN requer apenas dois parâmetros:  $\text{minPts}$  e  $\epsilon$ . O mínimo de pontos foi definido como 1, pois pode haver agrupamentos com apenas um ponto; são casos de observações não repetidas, sem outras observações próximas e que não são consideradas ruídos. O raio  $\epsilon$  foi definido para as distâncias: 30m, 90m, 250m, 500m, 1km e 2km. As coordenadas geográficas foram convertidas para radianos, pois o algoritmo trabalha com distâncias.

A biblioteca `sklearn.cluster.DBSCAN`<sup>62</sup> (PEDREGOSA et al., 2011) permite configurar argumentos adicionais; o algoritmo para indexar o cálculo da vizinhança foi o ‘*ball tree*’ e a métrica a ser utilizada no cálculo da distância foi ‘*haversine*’.

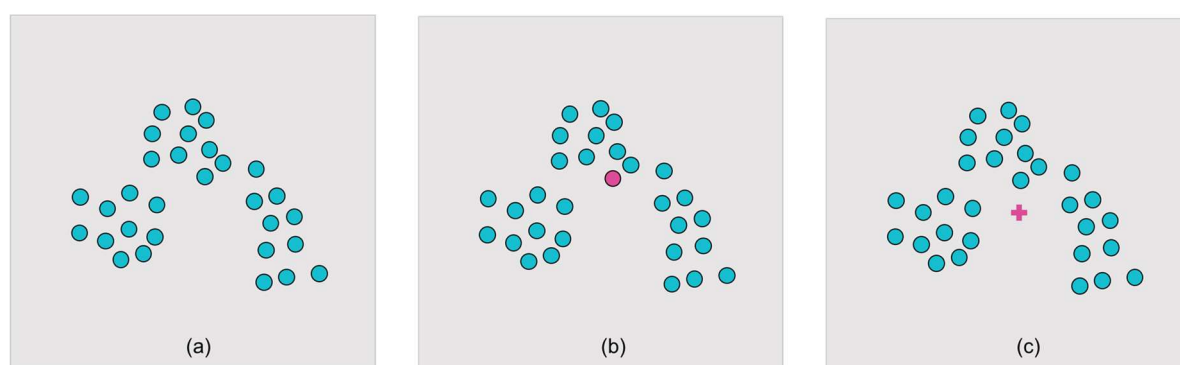
O algoritmo *Ball Tree* é uma árvore métrica que define como os dados são particionados no espaço, isto é, como os agrupamentos são separados, e é específica para consultas espaciais que utilizam distância euclidiana (OMOHUNDRO, 1989; DOLATSHAH et al., 2015). A Fórmula *Haversine* encontra a distância entre dois pontos em uma superfície não plana, usando parâmetros de longitude e latitude (PRASETYA et al., 2020).

Boeing (2018) desenvolveu um algoritmo semelhante para reduzir conjuntos de dados espaciais que lida com vários pontos por país, o que é basicamente um problema de visualização de dados. No caso de agrupamento de pontos considerando países como rótulo, não há

<sup>62</sup> scikit learn: DBSCAN. Disponível em <<https://scikit-learn.org/stable/modules/generated/sklearn.cluster.DBSCAN.html>>. Acesso em 05 mai. 2022.

possibilidade de sobreposição de territórios. No entanto, no caso de dados de biodiversidade, é possível que haja sobreposição espacial entre agrupamentos de espécies diferentes devido à ocorrência de espécies. Portanto, neste trabalho os agrupamentos foram definidos por espécies, sendo que uma mesma espécie pode estar em vários agrupamentos espacialmente distantes. Pontos próximos que são de espécies diferentes estarão em clusters diferentes. Outra diferença em relação ao trabalho de Boeing (2018) é que aqui consideramos o medóide do cluster e não o centróide. Isso porque o medóide refere-se a uma coordenada na qual uma espécie foi realmente observada, enquanto o centróide do agrupamento não é necessariamente um dos pontos e pode indicar uma localização com características (valores das variáveis preditivas) diferentes daquelas realmente observadas (Figura 22). Como apontado em Faceli et al. (2021), o centro de um agrupamento pode ser um centróide, como a média aritmética da distância entre os pontos do agrupamento (ou centro de gravidade), ou o medóide, que é o ponto mais representativo do agrupamento.

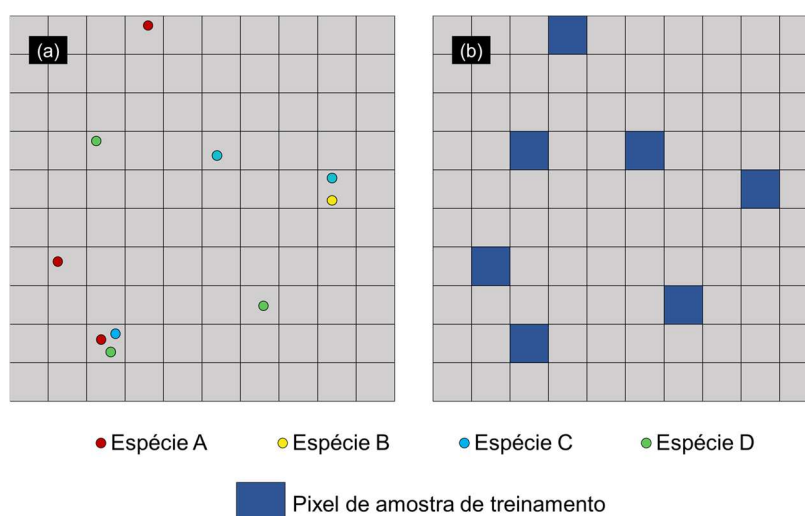
Figura 22 - Centro do agrupamento (medóide x centróide)



(a) Agrupamento original; (b) Representação pelo medóide (c) Representação pelo centróide.  
Fonte: Organizado pela autora.

Os agrupamentos são separados por espécies para a análise exploratória de dados para verificar o quanto os pontos estão agrupados por espécie. Contudo, para a modelagem em si, o filtro acaba por ser por ocorrência. Isto porque pontos de ocorrência de espécies distintas, mesmo sendo próximos são mantidos, mas como ponto amostral para o MAH serão contabilizados como uma única amostra pois estarão representando um mesmo pixel (Figura 23).

Figura 23 - Efeito da rasterização no tamanho amostral de Presença



(a) Localização dos pontos de ocorrência de espécies na matriz de pixels dos planos de informação das variáveis preditoras. (b) Pixels de amostra de treinamento.

Fonte: Organizado pela autora.

Os dados tabulares foram então convertidos para formato ESRI shapefile na projeção South America Albers Equal Area Conic (ESRI:102033). Este produto foi usado na análise exploratória dos dados e como amostra de treinamento para os modelos.

### 3.7. EXTRAÇÃO DOS VALORES DAS VARIÁVEIS PREDITORAS NOS PONTOS DE PRESENÇA DE ESPÉCIES

Nos planos de informação das variáveis preditoras, os pixels com posição coincidente aos pontos de presença foram obtidos pela função *'points\_query'* da biblioteca *'rasterstats'*<sup>63</sup>. Foi configurado o método de interpolação *'nearest'* para não haver alteração nos valores obtidos em relação ao valor real do pixel. Os resultados foram organizados de forma tabular para a Análise Exploratória dos Dados, seleção das variáveis e nova filtragem dos pontos de presença.

Os pontos de presença, gerados em formato vetorial, foram então convertidos para o formato raster para alimentação dos modelos.

<sup>63</sup> rasterstats, Disponível em <<https://pythonhosted.org/rasterstats/>>. Acesso em 10 jul. 2022.

### 3.8 PONTOS DE PSEUDO-AUSÊNCIA

Os MAHs são categorizados em dois grupos: os que precisam apenas de dados de presença e os que requerem dados de presença e de ausência. O modelo Bioclim usa apenas dados de presença; GARP e MaxEnt requerem dados de fundo ou de pseudo-ausência. A ausência é, por vezes, de difícil comprovação, principalmente para espécies que se locomovem e demandam maior esforço de amostragem para garantir a confiabilidade dos dados (BARBET-MASSIN et al., 2012).

Gerar pontos aleatórios por toda a área de estudo é a forma mais simples de criar os pontos de pseudo-ausência. Este é o caso da geração dos pontos de pseudo-ausência no modelo MaxEnt, que considera um plano de fundo em toda a área de interesse. Porém, existe a possibilidade de gerar ausência em uma área que é favorável à espécie, o que pode ser ainda mais inconveniente no caso de espécies raras ou pouco observadas, pois a quantidade de registros de presença não neutralizaria tantas falsas ausências (ENGLER et al., 2004).

Barbet-Massin et al. (2012) recomendam que os pontos de pseudo-ausência sejam definidos aleatoriamente, especialmente quando a alta especificidade (percentual de verdadeiros negativos) é priorizada em relação à alta sensibilidade (percentual de verdadeiros positivos), ou seja, a distribuição prevista de uma espécie é apenas a área onde a espécie é altamente provável de estar presente. É o caso, por exemplo, do planejamento de áreas protegidas. Os autores explicam que a alta especificidade garante que a proporção de ausências verdadeiras previstas como presenças será minimizada. Por outro lado, a alta sensibilidade garante que a proporção de verdadeiras presenças preditas como ausências seja minimizada. Também é aconselhável que o número de pontos de pseudo-ausência seja balanceado ao de pontos de presença. Diferentes métodos de geração dos pontos de pseudo-ausência baseados em aleatoriedade têm sido reportados na literatura.

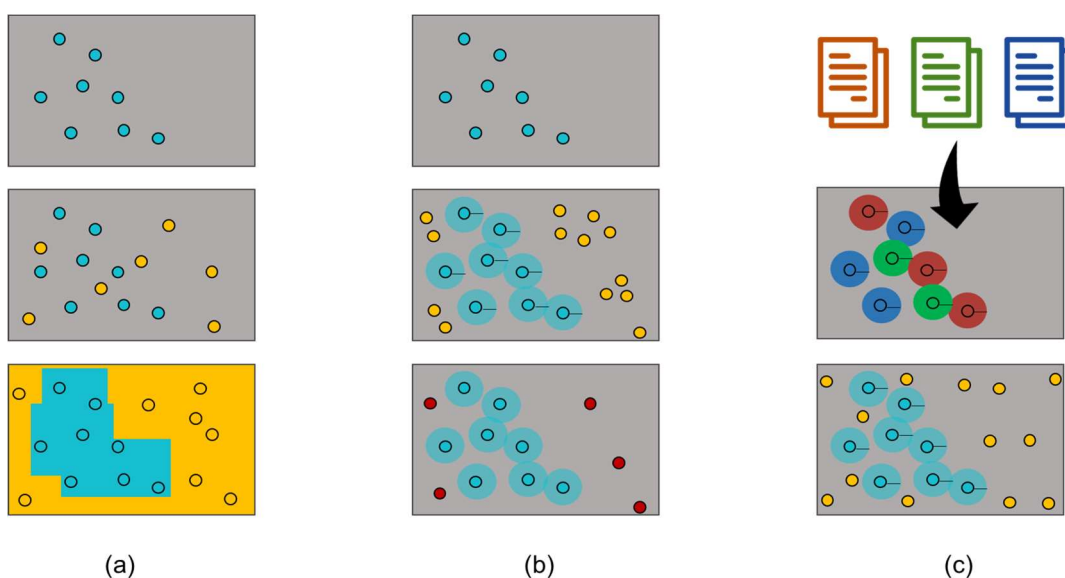
Zaniewski et al. (2002) consideraram dois métodos para gerar os pontos de pseudo-ausência para alimentar o modelo GAM (*Generalized additive model*): 1) gerar aleatoriamente pontos por toda a área de estudo; e 2) ponderar uma distribuição aleatória de pseudo-ausências em favor de áreas conhecidas por conterem ausências verdadeiras. Os autores criaram um mapa de adequabilidade de habitat de todas as espécies de samambaias com pontos totalmente aleatórios de pseudo-ausência; então usaram o resultado do primeiro modelo para produzir um segundo conjunto de pontos de pseudo-ausências para as espécies individuais.



Senay et al. (2013) desenvolveram um método de três etapas que compreende: 1) o estabelecimento de uma extensão geográfica em torno dos pontos de presença das espécies, a partir dos quais os pontos de pseudo-ausência são selecionados; 2) definição dos locais que são ambientalmente diferentes dos pontos de presença dentro da distância especificada na primeira etapa; 3) agrupamento K-means para redução do número de pseudo-ausências potenciais ao conjunto desejado, tomando os centróides dos agrupamentos na classe mais ambientalmente diferente identificada no segundo passo.

Coro et al. (2016) apresentaram um método para estimar locais de pseudo-ausências confiáveis para espécies marinhas, que são ausências que se aproximam bem das ausências verdadeiras. O método considera o *Ocean Biogeographic Information System* (OBIS) que é um banco de dados de dados de ocorrências de espécies, e para cada espécie, e segue as etapas: 1) coleta de informações sobre levantamentos que tiveram a espécie em sua lista de alvos; 2) cruzamento e processamento dos locais dos levantamentos para produzir locais de presença e ausências de amostragem; 3) seleção de locais de ausência de amostragem, que estão espacialmente distantes dos locais de presença, ou seja, não se sobrepõem aos locais de presença de acordo com um limite de distância definido pelo usuário. A Figura 24 resume os três métodos descritos.

Figura 24 - Síntese gráfica dos métodos de geração de pseudo-ausência por: (a) Zaniwski et al. (2002); (b) Senay et al. (2013); e (c) Coro et al. (2016)



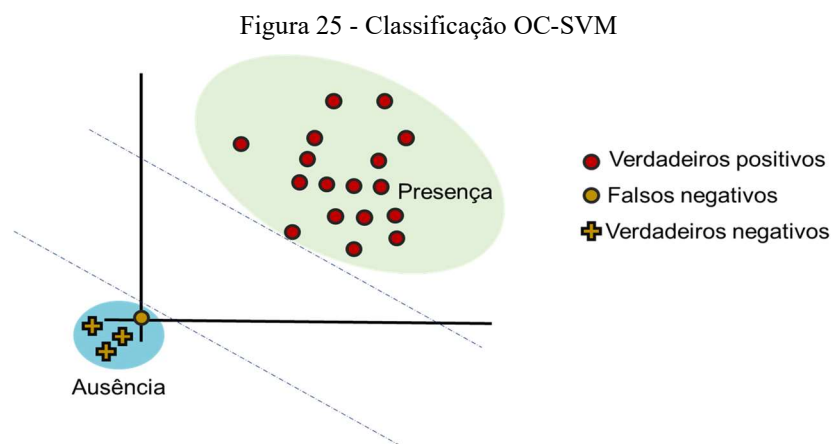
Fonte: ZANIEWSKI et al. (2002); SENAY et al. (2013); CORO et al. (2016).  
Organizado pela autora.

### 3.8.1 One-Class Support Vector Machine para inferência da pseudo-ausência

Neste trabalho, para definir os pontos de pseudo-ausência, foi considerado um algoritmo de classificação para uma única classe, ou seja, a classe de presença. O objetivo foi obter um resultado binário indicando locais de presença e pseudo-ausência. A partir dos pontos de presença e das variáveis preditoras, foi gerado um modelo a partir do algoritmo *One-Class Support Vector Machine* (OC-SVM), semelhante ao estudo de Senay et al. (2013).

Usualmente, os problemas de classificação lidam com várias classes. Porém, no caso da classificação de apenas uma classe, os elementos do conjunto são separados em uma única categoria. Esta abordagem é útil para detecção de outliers e detecção de novidades. O SVM (*Support Vector Machine*, explicado mais detalhadamente na seção sobre os modelos de distribuição de espécies baseados em Aprendizado de Máquina) pode ser adaptado para esta finalidade; os pontos de dados no espaço de atributos são separados de modo que a distância do hiperplano em relação aos conjuntos de pontos seja maximizada. Se os pontos estiverem em uma região densa, então a função atribui o valor +1 a estes pontos; caso estejam em uma região de baixa densidade, não fazem parte da classe e são considerados outliers e a eles são atribuídos o valor -1. Sendo assim, o OC-SVM requer apenas informações positivas para o treinamento (SCHÖLKOPF et al., 2001).

Isto significa que, considerando o espaço de atributos com as variáveis ambientais preditoras, é possível “encapsular” os pontos de presença no espaço e, eventualmente, alguns pontos de presença podem ficar de fora dessa região mais densa de pontos. O ponto mais distante em relação à região mais densa é considerado o vetor ou limiar de definição da classe (Figura 25).



Fonte: Organizado pela autora.

O algoritmo One-Class Support Vector Machine está implementado na biblioteca *Scikit Learn*<sup>64</sup>. Na região da área de interesse indicada pelo resultado do One-Class SVM como ausência das espécies, foi gerada uma amostra de pontos aleatórios de mesmo tamanho do número de presenças.

As camadas raster de cada variável ambiental foram convertidas em Numpy Array em formato Altura, Largura e Canal. Os valores dos pixels onde estão os pontos de ocorrência de espécies foram convertidos para o formato raster, sendo que os pixels coincidentes com os pontos têm valor 1 e os demais têm valor 0. O modelo de classificação OC-SVM foi treinado e testado neste conjunto amostral de Presença.

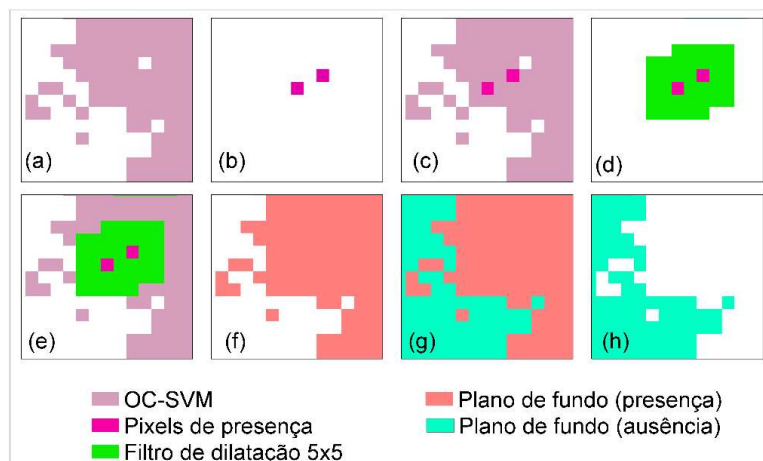
### 3.8.2 Geração do plano de fundo para os pontos de Pseudo-ausência

Foi aplicado um filtro morfológico de dilatação com vizinhança de dimensão 5 x 5 nos pixels de presença com o objetivo de criar uma zona de buffer nos locais de presença das espécies. O resultado deste procedimento foi combinado ao resultado do modelo OC-SVM para preparação do plano de fundo no qual foi gerada a amostra de pontos aleatórios de Pseudo-ausência de mesmo tamanho da amostra de presença para que as classes fiquem balanceadas. As Figuras 26a e 26b representam, respectivamente, o resultado da classificação OC-SVM e pixels de Presença, combinados na Figura 26c. Na Figura 26d está representado o filtro de dilatação em torno dos pixels de presença. As três camadas associadas à informação de Presença estão na Figura 26e que, quando somados, compõem o plano de informação de Presença (Figura 26f). O complemento desta (Figura 26g) é o plano de fundo de Pseudo-ausência (Figura 26h), onde foram gerados os pontos aleatórios de Pseudo-ausência.

---

<sup>64</sup> Scikit Learn: OneClassSVM. Disponível em <<https://scikit-learn.org/stable/modules/generated/sklearn.svm.OneClassSVM.html>>. Acesso em 01 jul. 2022.

Figura 26 - Geração do plano de fundo para os pontos de Pseudo-ausência

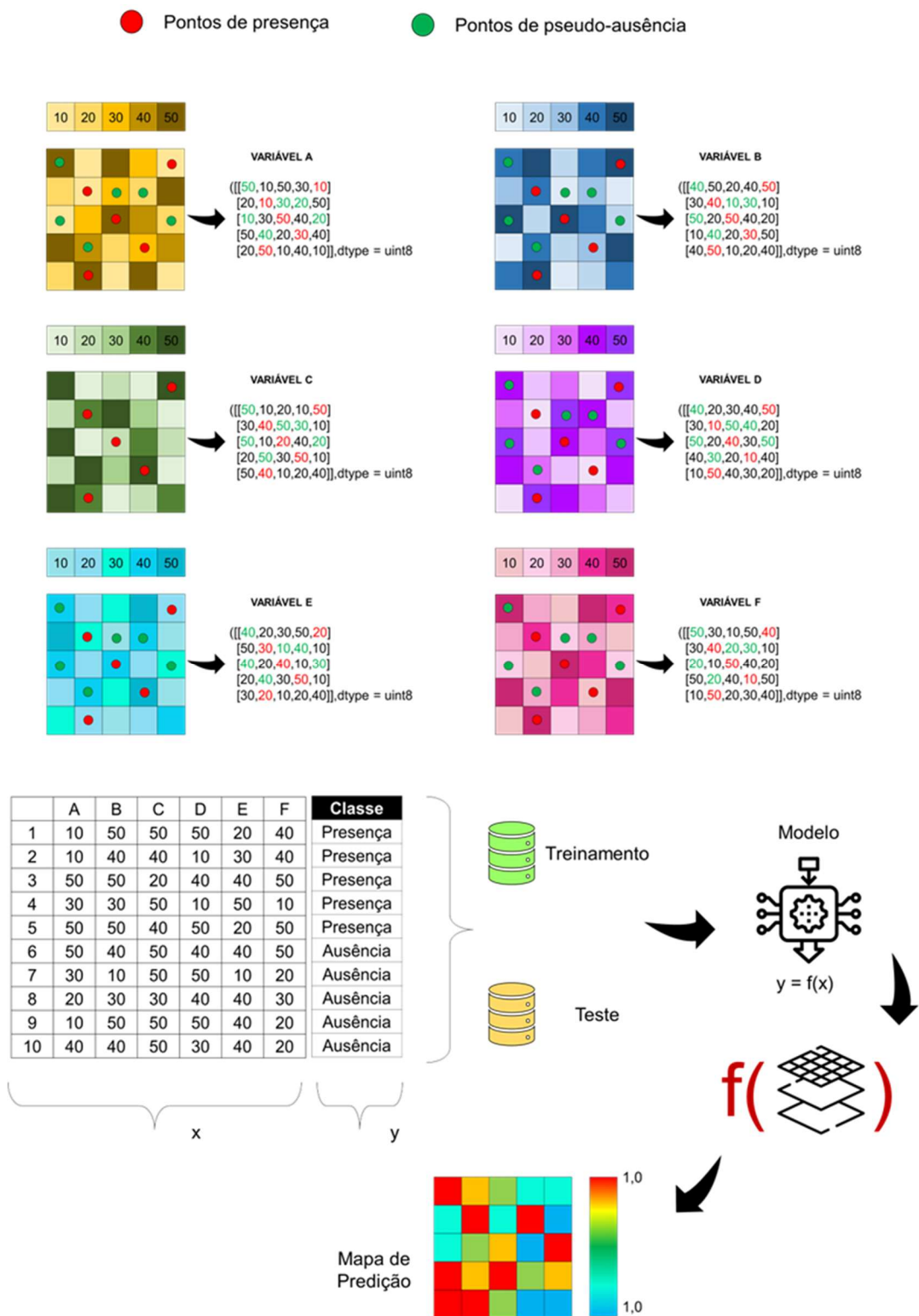


Fonte: Organizado pela autora.

### 3.9 MODELOS DE ADEQUABILIDADE DE HABITAT BASEADOS EM ALGORITMOS DE APRENDIZADO DE MÁQUINA

As camadas raster de cada variável ambiental foram convertidas em Numpy Array em formato Altura, Largura e Canal. Os pontos de Pseudo-ausência e Presença foram convertidos para o formato raster com os valores 1 e 2, respectivamente. Os modelos foram treinados e testados nestes pixels de amostras e estendidos para toda a área de estudo (Figura 27).

Figura 27 - Representação da conversão das camadas raster das variáveis ambientais em Numpy Array e formação do dado tabular para treinamento e testes dos modelos



Fonte: Organizado pela autora.

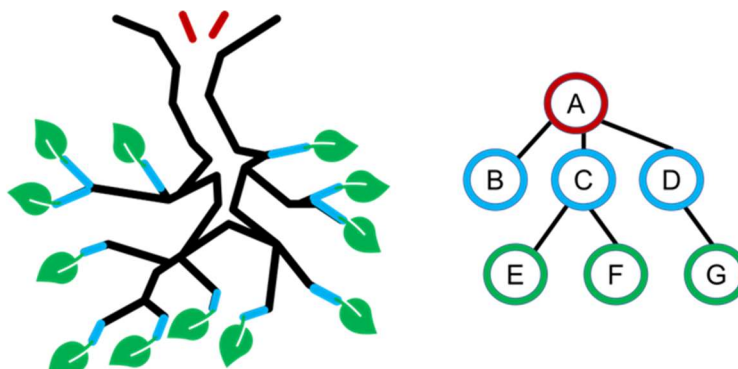
Para todos os modelos, foi aplicada a ferramenta GridSearchCV, que ajusta os hiperparâmetros de um modelo a partir do conjunto de treinamento usando a validação cruzada (PEDREGOSA et al., 2011; BERGSTRA; BENGIO, 2012; GOEL et al., 2021). Foram considerados os valores dos hiperparâmetros do modelo que obteve maior valor de coeficiente de determinação ( $r^2$ ).

A homocedasticidade dos resíduos é um dos pressupostos da regressão, ou seja, a variância deve ser a mesma para todos os valores preditos, e os erros devem ter distribuição normal. Quanto mais variáveis um modelo contém, mais complexo ele é e os dados são mais esparsos, isto é, o modelo é mais sensível a erros devido à variação. Essa condição não é essencial para obter os estimadores de mínimos quadrados, mas é indispensável para definir intervalos de confiança e testes de significância. Isto significa que os estimadores não são tendenciosos na ausência de normalidade; porém, os testes não têm validade (SOUZA et al., 2005; HARRISON, 2020). Para todos os modelos de regressão foi realizado o teste Breusch-Pagan para heterocedasticidade para verificar se a variância dos resíduos se comporta nos valores previstos (BREUSCH; PAGAN, 1979).

### 3.9.1 Árvores de Decisão e Florestas Aleatórias

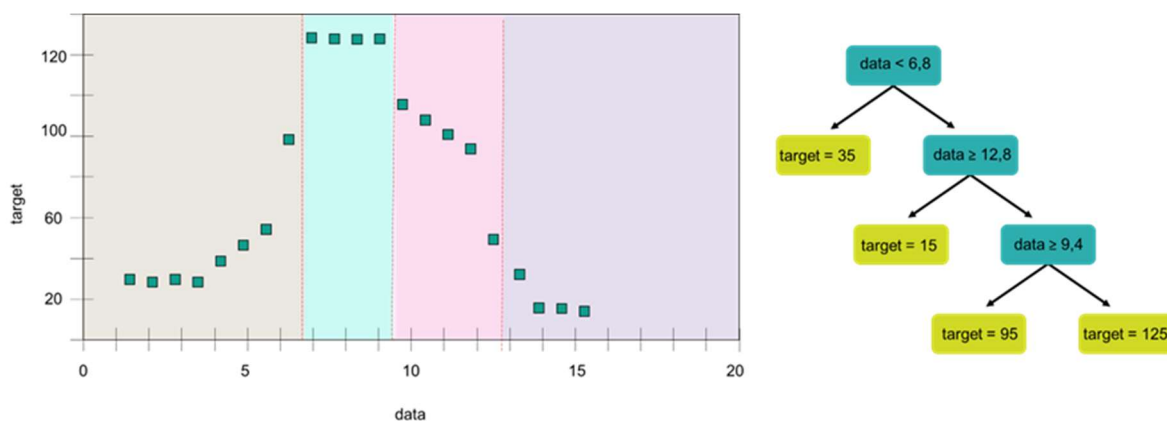
A Árvore de Decisão (*Decision Tree*) constrói modelos de regressão ou classificação seguindo a estrutura de árvore, dividindo um conjunto de dados em subconjuntos cada vez menores. Uma árvore de decisão tem o formato de uma árvore invertida. A Figura 28 mostra a estrutura de uma árvore de decisão onde os círculos são os nós da árvore, cada linha é uma aresta. A é a raiz; B, C e D são descendentes de A (o que é aplicado para os demais nós); B, C e D são nós internos e contêm atributos; E, F e G são folhas e possuem classes (ou valores contínuos, no caso da regressão). O objetivo é dividir os atributos que criam os nós filhos mais puros possíveis, o que mantém, no mínimo, o número de divisões que precisam ser feitas para classificar todas as instâncias do conjunto de dados (Figura 29). Ou seja, cada nó é uma condição de como dividir valores em uma única variável, para que valores semelhantes da variável preditora fiquem no mesmo conjunto após a divisão. A pureza é medida pelo conceito de ganho de informação relacionado a quanto precisaria ser conhecido sobre uma instância não vista anteriormente para que ela fosse classificada adequadamente (BREIMAN et al., 1984).

Figura 28 - Estrutura de uma árvore de decisão



Fonte: Organizado pela autora.

Figura 29 - Árvore de decisão para regressão



Fonte: Organizado pela autora.

A Floresta Aleatória é um conjunto de árvores de decisão, na qual o vetor de entrada é executado por meio de várias árvores de decisão. Apenas um subconjunto aleatório de atributos participa da subdivisão de um nó. Para regressão, o valor de saída de todas as árvores é calculado; para classificação é usado um esquema de votação para determinar a classe final. São adequadas para reduzir a variação em relação a uma única árvore de decisão e atenuam sobreajustes (BREIMAN, 2001).

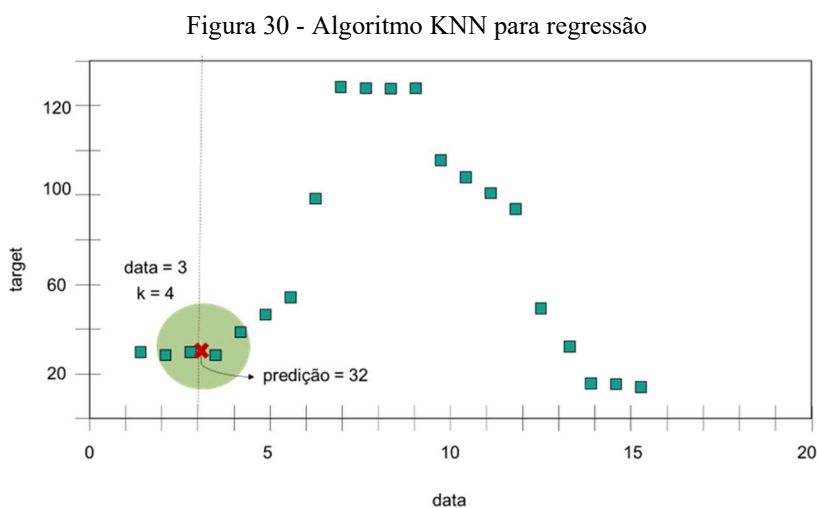
As Árvores de Decisão são adequadas para aprender relacionamentos complexos e altamente não lineares. Entretanto, são propensas a sobreajustes e a conter estruturas desnecessárias, embora isso possa ser atenuado por podas de árvores adequadas e conjuntos florestais aleatórios maiores. Florestas aleatórias maiores são mais lentas e exigem mais memória computacional. Os sobreajustes ocorrem quando o modelo se ajusta demais aos dados

de treinamento e não consegue prever os dados de teste não vistos; ou seja, o modelo não tem capacidade de generalização a dados adicionais porque força a construção de um padrão e ruídos específicos nos dados de treinamento.

### 3.9.2 Regressão KNN

A regressão KNN (*K-Nearest Neighbor Regressor*) é um algoritmo de interpolação não paramétrico que usa k-vizinhos para estimar a variável alvo. Não requer nenhuma correlação (relação linear) entre os recursos e a variável de destino, o que é um requisito para a regressão linear. O alvo é previsto pela interpolação local dos alvos associados aos vizinhos mais próximos no conjunto de treinamento (ALTMAN, 1992). Na Figura 30, considerando a variável dependente igual a 3, o valor predito é a média dos k valores mais próximos (neste caso,  $k = 4$ ); este procedimento é, então, realizado para todos os valores dos dados de treinamento para geração do modelo.

A regressão KNN pode ser lenta para implementar, pois depende do cálculo da distância entre todas as instâncias de vetor, o que pode consumir muito tempo para grandes conjuntos de dados.



Fonte: Organizado pela autora.

### 3.9.3 Regressão de Vetor Suporte

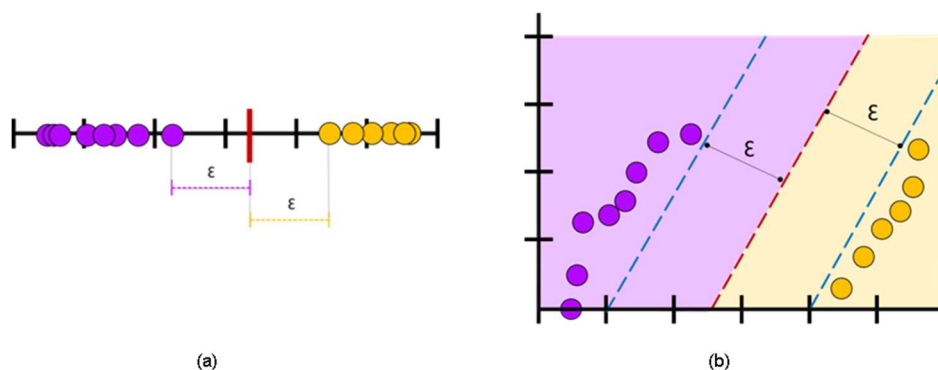
O objetivo de um algoritmo de Máquina de Vetores de Suporte (*Support Vector Machine - SVM*) é encontrar um hiperplano em um espaço n-dimensional para predição. Os pontos de dados em ambos os lados do hiperplano que estão mais próximos do hiperplano são chamados



de Vetores de Suporte. Estes influenciam a posição e orientação do hiperplano e, assim, ajudam a construir o SVM (VAPNIK, 1995).

Na Figura 31a, o limiar representado pelo traço vermelho é o hiperplano que melhor separa os dois conjuntos de dados (roxos e laranjas), isto é, a distância  $d$  entre o hiperplano e os vetores suporte é maximizada. Isso melhora a garantia de que novas observações sejam preditas com valores mais próximos dos reais. Quando o espaço de atributos é bidimensional, o vetor de suporte é uma linha (pontilhado vermelho na Figura 31b). As linhas azuis paralelas mostram onde os pontos estão em relação ao hiperplano.

Figura 31 - Definição do hiperplano no SVM

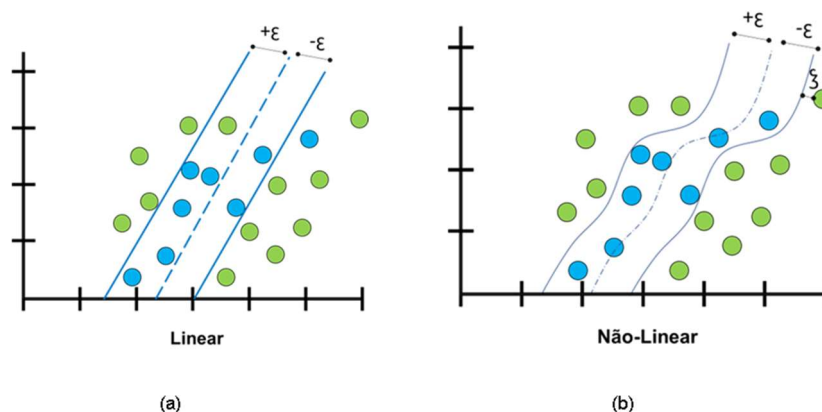


Fonte: Organizado pela autora.

Drucker et al. (1996) propuseram uma versão do SVM para regressão, nomeada Regressão de Vetor Suporte (*Support Vector Regression – SVR*). O SVR prevê valores contínuos e usa o mesmo princípio dos SVMs, sendo que a linha de melhor ajuste é o hiperplano que possui o número máximo de pontos. Ao contrário de outros modelos de regressão que visam minimizar o erro entre o valor observado e o previsto, o SVR tenta ajustar a melhor linha dentro de um limiar. O limiar é a distância entre o hiperplano e a linha de limite.

A Figura 32a mostra a separação dos pontos por um hiperplano linear; por outro lado, o hiperplano não linear discrimina melhor os pontos e consegue fazer uma predição mais acurada (Figura 32b). O hiperparâmetro  $\xi$  suaviza a linha da margem e flexibiliza a variação dos pontos.

Figura 32 - Definição do hiperplano em modelos linear e não linear do SVR



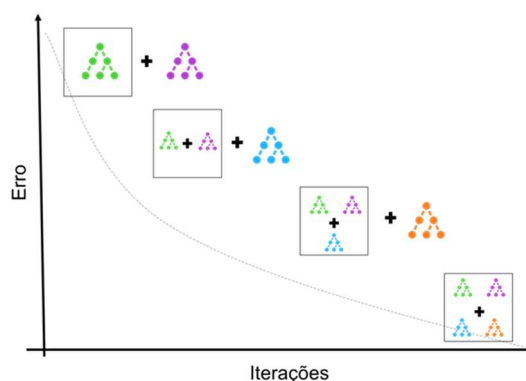
Fonte: Organizado pela autora.

O SVR tem boa capacidade de generalização, com alta precisão de previsão, e é resistente na predição de outliers. Entretanto, nos casos em que o número de atributos exceda o número de amostras de dados de treinamento, o SVR pode ter um desempenho insuficiente. Também não funciona muito bem quando as classes de destino estão sobrepostas, ou seja, quando o conjunto de dados tem mais ruído.

### 3.9.4 XGBoost

O XGBoost (*eXtreme Gradient Boosting*) é um algoritmo baseado em árvore de decisão e que utiliza uma estrutura de *gradient boosting*. Neste caso, cada árvore é ponderada de forma a compensar fraquezas de aprendizado do modelo (erros residuais) na árvore gerada anteriormente. As florestas são construídas de forma iterativa e o XGBoost faz novas previsões baseadas nestas taxas de aprendizado e na função de perda, que é uma penalidade por previsões ruins. A construção da árvore é baseada em iterações na tentativa de diminuir os resíduos das previsões até que estes sejam menores que um determinado limiar ou até atingirem o número máximo de árvores (Figura 33) (CHEN; GUESTRIN, 2016).

Figura 33 - Construção da árvore XGBoost a partir de diversas iterações para minimizar os resíduos da predição



Fonte: Organizado pela autora.

### 3.9.5 LightGBM

LightGBM é uma estrutura de *gradient boosting* baseada em árvores, com agrupamento de atributos para lidar com grandes quantidades de variáveis (KE et al., 2017). A árvore do LightGBM cresce verticalmente (crescimento por folha, Figura 34a), enquanto em outros algoritmos cresce horizontalmente (crescimento por nível, Figura 34b). Diferente do XGBoost, que usa árvores de decisão para dividir em uma variável e explorar diferentes podas nessa variável, o LightGBM concentra-se em uma divisão e continua dividindo a partir dela para obter um melhor ajuste (que é a folha estratégia de crescimento de árvores). Assim, o LightGBM alcança mais rapidamente um bom ajuste dos dados e gera soluções alternativas em comparação com o XGBoost (BOSCHETTI; MASSARON, 2018).

Outra diferença é que o XGBoost requer a codificação das variáveis categóricas, ao passo que para o LightGBM é suficiente declarar variáveis são categóricas e o algoritmo executa uma divisão equitativa dos valores (OWEN, 2022). Nesta pesquisa, as variáveis ambientais passaram pelo mesmo pré-processamento para todos os modelos e, portanto, o LightGBM foi alimentado com as variáveis categóricas (classes de cobertura da terra) já codificadas.

Figura 34 - Crescimento da árvore de decisão (a) pelo LightGBM; e (b) por outros algoritmos



Fonte: Organizado pela autora.

### 3.10 MAPAS DE ADEQUABILIDADE DE HABITAT

Os modelos de regressão, por serem treinados em pixels de valores 1 (Pseudo-ausência) ou 2 (Presença), têm resultados que variam neste intervalo. Para manter os valores de adequabilidade entre 0 e 1, os dados raster resultantes dos modelos foram convertidos para este novo intervalo.

Os resultados dos MAHs podem ser interpretados como mapas de Adequabilidade de Habitat (AdH), com pixels que variam de 0 (menor adequabilidade) a 1 (maior adequabilidade).

### 3.11 ESTATÍSTICAS ZONAIIS

Para cada categoria de UC e para as UCs individuais foram calculadas as estatísticas zonais do nível de Adequabilidade de Habitat gerado por cada modelo. São os valores máximo, mínimo, média, mediana e moda dos pixels nos MAHs. A Adequabilidade de Habitat Média (AdH média) é a média dos valores dos pixels que representam uma UC no mapa de Adequabilidade.

### 3.12 ANÁLISE DO USO E COBERTURA DA TERRA NAS UCS DO ESTADO DE SÃO PAULO

Os dados geospaciais das Unidades de Conservação estaduais e federais do estado de São Paulo são provenientes do banco de dados do Instituto Socioambiental (ISA). Foi feita uma caracterização das UCs quanto à proporção de Unidades por esfera e categoria de manejo e suas sobreposições.

Para análise do uso e cobertura da terra, foram usados os dados do MapBiomias (2022), Coleção 6, referentes aos anos de 1985, 1995, 2005, 2015 e 2020. Estas datas foram selecionadas pois compõem intervalos de tempo nos quais foram criados números semelhantes de Unidades de Conservação no estado. A mesma base de dados de Unidades de Conservação foi usada para todas as datas.

Foi contabilizada a área das classes de cobertura da terra nas UCs a partir do número de pixels de cada classe mapeada pelo MapBiomias e conversão em hectares (Equação 3).

área em hectares = número de pixels \*30<sup>2</sup>/100

Equação 3

A análise da cobertura da terra nas UCs será feita a partir dos percentuais de formações florestais naturais e de coberturas antrópicas com vegetação. Para tanto, os pixels associados a estas classes foram recodificados conforme a Tabela 10 para compor cinco classes gerais: 1) formações naturais florestais; 2) formações naturais não florestais; 3) coberturas antrópicas com vegetação; 4) áreas antrópicas não vegetadas; 5) Corpos d'Água.

Tabela 10 - Recodificação das classes do MapBiomias

Classe	Valor original do pixel	Valor do pixel recodificado
Formação Florestal	3	1
Formação Savânica	4	
Mangue	5	
Restinga Arborizada	49	
Campo Alagado e Área Pantanosa	11	2
Formação Campestre	12	
Outras Formações não Florestais	13	
Praia, Duna e Areal	23	
Afloramento Rochoso	29	
Apicum	32	3
Silvicultura	9	
Pastagem	15	
Cana	20	
Mosaico de Agricultura e Pastagem	21	

Classe	Valor original do pixel	Valor do pixel recodificado
Soja	39	
Outras Lavouras Temporárias	41	
Café	46	
Citrus	47	
Outras Lavouras Perenes	48	
Área Urbanizada	24	
Outras Áreas não Vegetadas	25	4
Mineração	30	
Aquicultura	31	5
Rio, Lago e Oceano	33	

### 3.13 ANÁLISE DA COBERTURA DA TERRA NO ESTADO DE SÃO PAULO

A análise da cobertura em todo o estado de São Paulo se faz importante no sentido de permitir a comparação da distribuição espacial de algumas classes de cobertura da terra com os resultados dos modelos de adequabilidade. Foram avaliadas três principais coberturas no estado de São Paulo: 1) Remanescentes de formações florestais, quantificados a partir das classes Formação Florestal e Formação Savânica dos dados do MapBiomass (2022), Coleção 6, com pixels de 30 metros; 2) Agricultura e Pecuária, quantificada a partir do Censo Agropecuário<sup>65</sup> e da Pesquisa da Pecuária Municipal<sup>66</sup>, ambos do IBGE, já que o MapBiomass mapeia parte dessas classes como Mosaico de Agricultura e Pastagem; e 3) Urbanização,

<sup>65</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. Censo agropecuário. Disponível em <<https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/21814-2017-censo-agropecuario>>. Acesso em 16 abr. 2022.a

<sup>66</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. Pesquisa da pecuária municipal. Disponível em <<https://sidra.ibge.gov.br/tabela/3939>>. Acesso em 16 abr. 2022.b

avaliada nas proximidades de algumas UCs a partir de imagens dos satélites Sentinel 2-A e 2-B, Landsat-4 e 9.

No caso das atividades Agricultura e Pastagem, o Índice de Moran e o Índice LISA (*Local Indicator of Spatial Association*) foram aplicados aos municípios paulistas para caracterizar o padrão espacial a localização de agrupamentos de municípios que se destacam nessas atividades (ANSELIN, 2010; CLIFF; ORD, 1973; MORAN, 1948).

## 4 RESULTADOS E DISCUSSÕES

### 4.1 USO E COBERTURA DA TERRA E SUAS IMPLICAÇÕES NA OCORRÊNCIA DE MAMÍFEROS

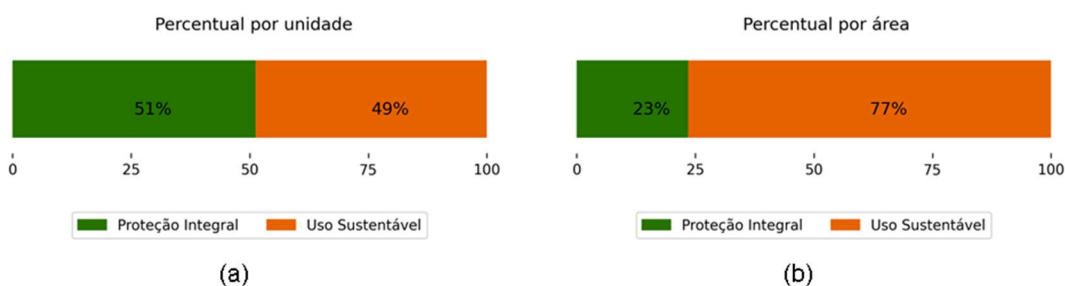
#### 4.1.1 Panorama das Unidades de Conservação no estado de São Paulo

No estado de São Paulo são 35.286 km<sup>2</sup> (14,2% do estado) de áreas protegidas, sendo 106 UCs estaduais e 15 UCs federais (ISA, 2018).<sup>67</sup> Este quantitativo considera que há sobreposição entre unidades de diferentes usos e esferas. As UCs estaduais estão sob coordenação, principalmente, do Instituto Florestal/Fundação Florestal e totalizam 30.694 km<sup>2</sup>. O ICMBio coordena as UCs federais que totalizam 6.524 km<sup>2</sup> no estado de São Paulo; algumas destas são interestaduais (APA da Serra da Mantiqueira, em São Paulo e Minas Gerais; PARNA da Serra da Bocaina, em São Paulo e Rio de Janeiro; APA das Ilhas e Várzeas do Rio Paraná, em São Paulo, Mato Grosso do Sul e Paraná).

São 62 unidades de Proteção Integral e 59 de Uso Sustentável (ISA, 2018); entretanto, em termos de área, as UCs de Uso Sustentável somam uma extensão mais expressiva (77%) (Figura 35). As Tabelas 11 e 12 mostram os quantitativos por categorias de UCs em termos de número e área, respectivamente.

Apenas 8% da área classificada pelo IBGE<sup>68</sup> como Cerrado é protegida por UCs (0,3% como UCPI e 7,8% como UCUS, com sobreposições entre unidades de diferentes usos). Na Mata Atlântica, esse percentual é de 17% (5,4% como UCPI e 12,9% como UCUS, com sobreposições entre unidades de diferentes usos).

Figura 35 - Proporção entre UCPIs e UCUSs em termos de (a) número de UCs; (b) área de UCs



<sup>67</sup> Na preparação do arquivo da área de estudo, com a remoção de algumas ilhas, não estão sendo consideradas a RVS do Ilhas do Abrigo e Guararitama e a RVS Arquipélago dos Alcatrazes, tampouco as UCs marinhas.

<sup>68</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. Biomas do Brasil 1:250 000. Disponível em <[http://www.geoservicos.ibge.gov.br/geoserver/wms?service=WFS&version=1.0.0&request=GetFeature&typeName=CREN:biomas\\_5000&outputFormat=SHAPE-ZIP](http://www.geoservicos.ibge.gov.br/geoserver/wms?service=WFS&version=1.0.0&request=GetFeature&typeName=CREN:biomas_5000&outputFormat=SHAPE-ZIP)>. Acesso em 11 nov. 2022.



Tabela 11 - Número de UCs por categoria e percentual em relação ao número total

Uso	Categoria	Quantidade de UCs	% por categoria
Proteção Integral	ESEC	24	19,8
	MONAT	1	0,8
	PARNA	1	0,8
	PE	33	27,3
	REBIO	3	2,5
Uso Sustentável	APA	30	24,8
	ARIE	9	7,4
	FE	7	5,8
	FLONA	3	2,5
	RDS	7	5,8
	RESEX	3	2,5

ESEC = Estação Ecológica; MONAT = Monumento Natural; PARNA = Parque Nacional; PE = Parque Estadual; REBIO = Reserva Biológica; APA = Área de Proteção Ambiental; ARIE = Área de Relevante Interesse Ecológico; FE = Floresta Estadual; FLONA = Floresta Nacional; RDS = Reserva de Desenvolvimento Sustentável; RESEX = Reserva Extrativista

As categorias de UCs mais representativas em termos de extensão são as APAs e PE, que representam 12,1% e 3,12% do estado de São Paulo, respectivamente (Tabela 12). Quando essa representatividade é avaliada em termos de esfera administrativa, essas ainda são as categorias de maior extensão. Os mapas de árvore nas Figuras 36 e 37 mostram, respectivamente, a proporção da área das categorias de UCs estaduais e federais. Nas duas esferas administrativas, as APAs têm a maior contribuição na extensão de áreas protegidas por UCs, no todo e também considerando apenas as UCs de Uso Sustentável. Essa categoria é constituída por terras públicas ou privadas e, neste último caso, não requer a desapropriação, o que viabiliza a criação de Unidades.

Quanto às UCPI, os parques são a categoria de maior proporção em termos de área, nas esferas estadual e federal.

Tabela 12 - Área de UCs por categoria e percentual em relação à área protegida e ao estado de São Paulo

<b>Uso</b>	<b>Categoria</b>	<b>Área (km<sup>2</sup>)</b>	<b>% de área em relação à extensão das UCs</b>	<b>% de área em relação à extensão do estado de São Paulo</b>
Proteção Integral	ESEC	1.217	3,0	0,5
	MONAT	33	0,1	0,01
	PARNA	364	0,9	0,15
	PE	7.752	19,4	3,12
	REBIO	31	0,1	0,01
Uso Sustentável	APA	30.022	75,1	12,1
	ARIE	153	0,4	0,06
	FE	103	0,3	0,04
	FLONA	98	0,3	0,04
	RDS	181	0,5	0,07
	RESEX	38	0,1	0,02

Figura 36 - Mapa de árvore das proporções de área das categorias de UCs estaduais

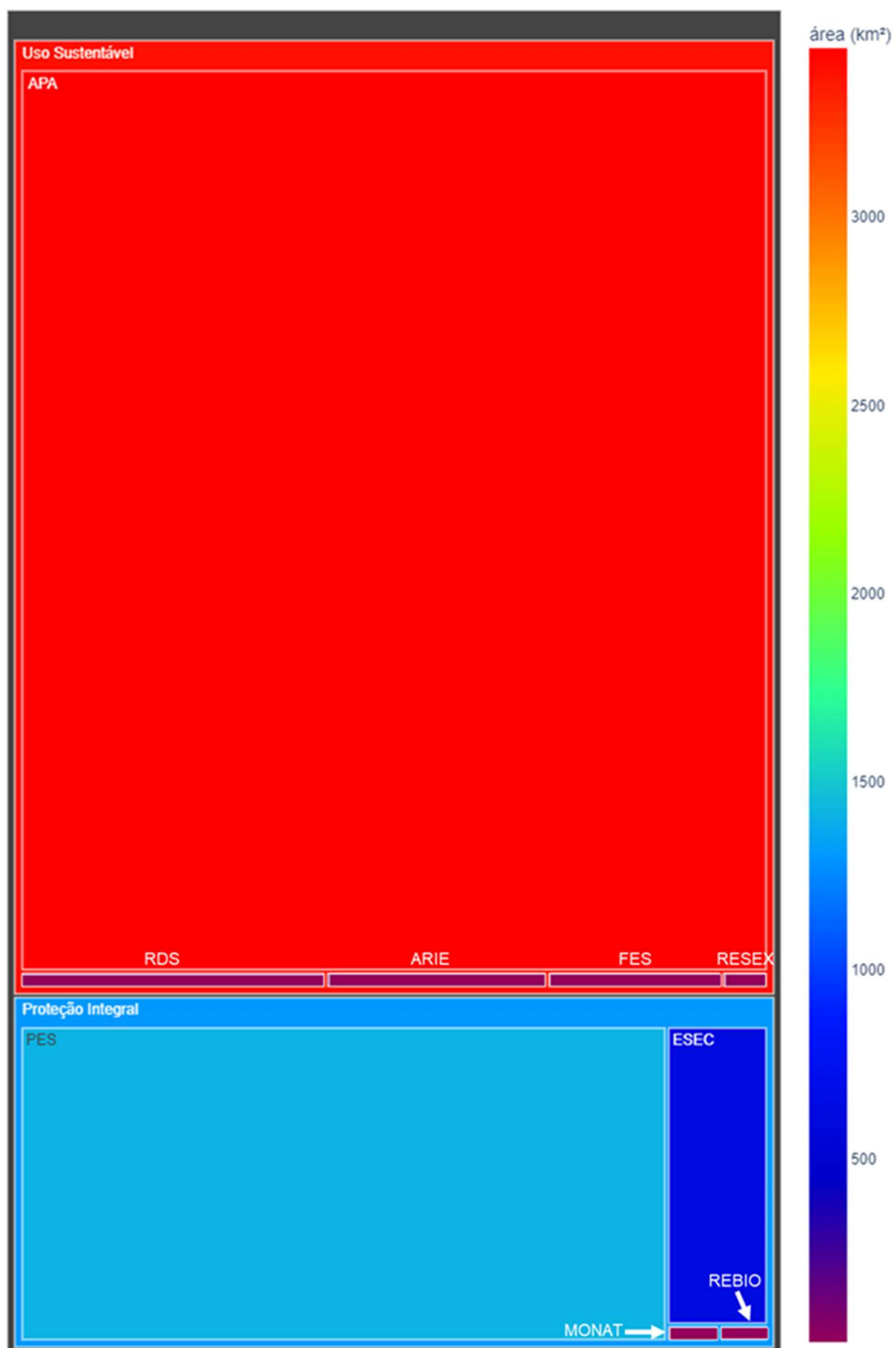
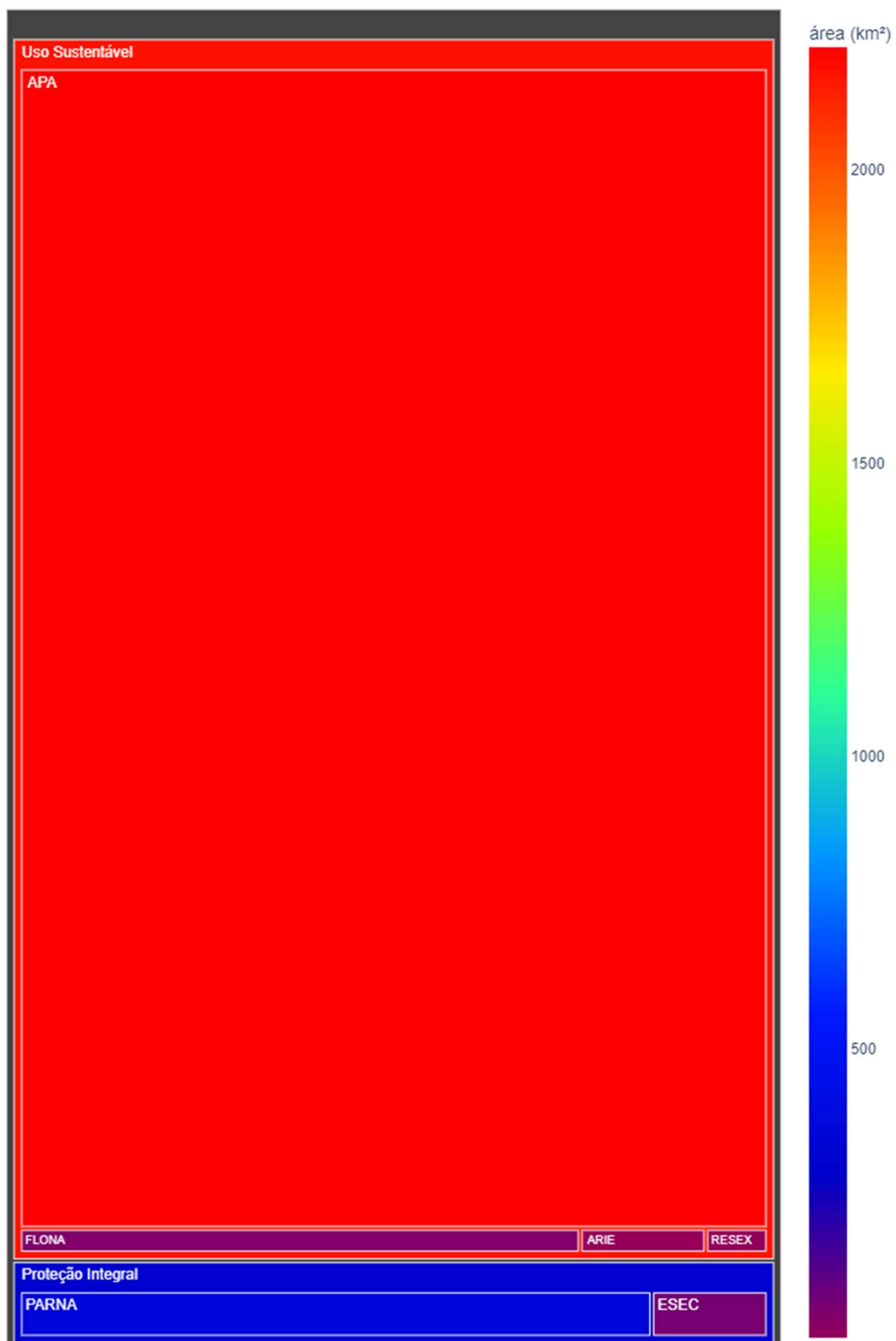


Figura 37 - Mapa de árvore das proporções de área das categorias de UCs federais



Há sobreposição de 438 km<sup>2</sup> entre UCs estaduais de usos diferentes, como mostra a Tabela 13. É comum que as operações geoespaciais de interseção identifiquem certa sobreposição entre polígonos adjacentes. No entanto, há casos como o do PE Nascentes do Paranapanema que está em sua totalidade no interior da APA Serra do Mar.

Tabela 13 - Sobreposição entre UCs estaduais de Uso Sustentável e de Proteção Integral

<b>UC Uso Sustentável</b>	<b>UC Proteção Integral</b>	<b>Sobreposição (km<sup>2</sup>)</b>
APA da Serra do Mar	PE Nascentes do Paranapanema	208
APA Campos do Jordão	PE de Campos do Jordão	79
APA Sistema Cantareira	PE de Itapetinga	71
APA Corumbataí, Botucatu e Tejuπά	ESEC de Itirapina	23
APA dos Rios Piracicaba e Juqueri-Mirim	PE de Itapetinga	16
APA Sistema Cantareira	MONAT da Pedra Grande	14
APA dos Quilombos do Médio Ribeira	PE Intervalas	10
APA Campos do Jordão	PE dos Mananciais de Campos do Jordão	5
APA da Serra do Mar	PE Caverna do Diabo	5
FES Assis	ESEC de Assis	4
APA Rio Batalha	ESEC Sebastião Aleixo da Silva (Bauru)	3
ARIE de São Sebastião	PE da Serra do Mar	0,4
APA dos Rios Piracicaba e Juqueri-Mirim	MONAT da Pedra Grande	0,05
FE de Guarulhos	PE de Itaberaba	0,02
APA dos Quilombos do Médio Ribeira	PE Caverna do Diabo	0,002

A Tabela 14 mostra as sobreposições entre UCs estaduais e federais; as células verdes representam as UCPIs e as amarelas as UCUSs. Os Planos de Manejo dessas áreas protegidas

devem contemplar a gestão compartilhada pelas duas esferas administrativas, envolvendo o Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) e o Instituto Florestal (IF).

Tabela 14 - Sobreposição entre UCs Estaduais e Federais

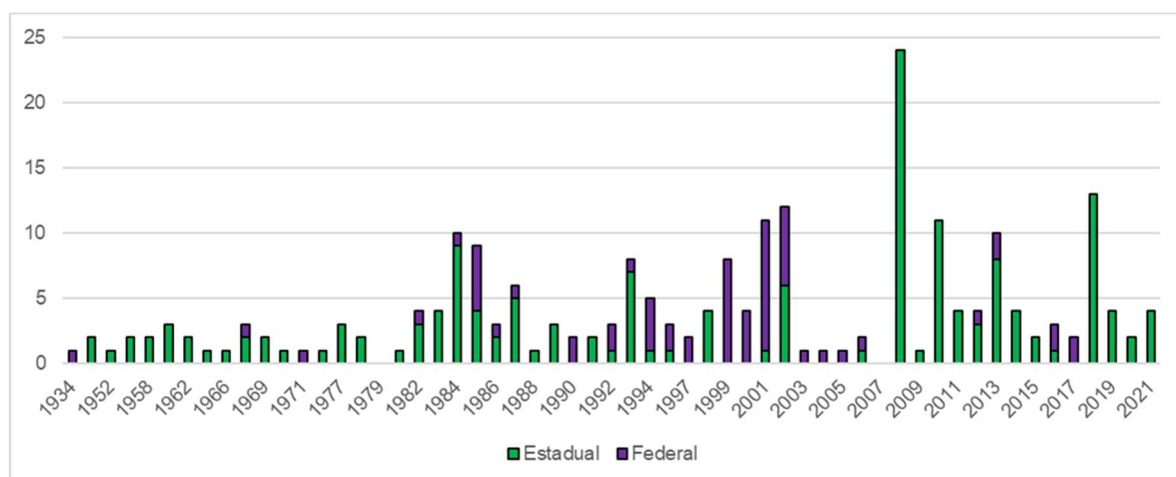
UC Estadual	UC Federal	Sobreposição (km <sup>2</sup> )
ESEC Juréia-Itatins	APA Cananéia - Iguape - Peruíbe	655,05
APA Sapucaí-Mirim	APA da Serra da Mantiqueira	262,30
PE do Lagamar de Cananéia	APA Cananéia - Iguape - Peruíbe	202,31
APA Silveiras	APA Mananciais da Bacia do Rio Paraíba do Sul	162,52
APA Campos do Jordão	APA da Serra da Mantiqueira	161,91
PE de Itaberaba	APA Mananciais da Bacia do Rio Paraíba do Sul	129,92
APA São Francisco Xavier	APA Mananciais da Bacia do Rio Paraíba do Sul	109,85
PE da Serra do Mar	PARNA da Serra da Bocaina	107,24
PE de Campos do Jordão	APA da Serra da Mantiqueira	78,77
RDS do Despraiado	APA Cananéia - Iguape - Peruíbe	39,60
RESEX Taquari	APA Cananéia - Iguape - Peruíbe	16,48
PE da Serra do Mar	APA Mananciais da Bacia do Rio Paraíba do Sul	16,46
RDS de Itapanhapima	APA Cananéia - Iguape - Peruíbe	12,41
RESEX da Ilha do Tumba	APA Cananéia - Iguape - Peruíbe	11,11
PE do Itinguçu	APA Cananéia - Iguape - Peruíbe	10,55
PE de Vassununga	ARIE Cerrado Pé-de-Gigante	9,65

UC Estadual	UC Federal	Sobreposição (km <sup>2</sup> )
ESEC Juréia-Itatins	APA Cananéia - Iguape - Peruíbe	655,05
APA Sapucaí-Mirim	APA da Serra da Mantiqueira	262,30
APA da Ilha Comprida	APA Cananéia - Iguape - Peruíbe	8,47
APA Sistema Cantareira	APA Mananciais da Bacia do Rio Paraíba do Sul	4,76
APA dos Rios Piracicaba e Juqueri-Mirim	APA Mananciais da Bacia do Rio Paraíba do Sul	4,63
RDS Barra do Una	ARIE Ilha do Ameixal	4,33
ARIE da Ilha Comprida	APA Cananéia - Iguape - Peruíbe	3,33
PE dos Mananciais de Campos do Jordão	APA da Serra da Mantiqueira	2,57
ARIE Guará	APA Cananéia - Iguape - Peruíbe	2,07
FE de Guarulhos	APA Mananciais da Bacia do Rio Paraíba do Sul	0,98
RDS Barra do Una	APA Cananéia - Iguape - Peruíbe	0,55
APA Sapucaí-Mirim	APA Mananciais da Bacia do Rio Paraíba do Sul	0,37
APA Campos do Jordão	APA Mananciais da Bacia do Rio Paraíba do Sul	0,06
PE de Campos do Jordão	APA Mananciais da Bacia do Rio Paraíba do Sul	0,06
RESEX Taquari	RESEX do Mandira	0,02
PE do Lagamar de Cananéia	RESEX do Mandira	0,017
ESEC Juréia-Itatins	ARIE Ilha do Ameixal	0,005
PE do Prelado	APA Cananéia - Iguape - Peruíbe	0,001

\* As células verdes representam as UCPIs e as amarelas as UCUSs.

Embora as categorias RDS, RESEX e MONAT estejam previstas no SNUC (2020), as primeiras unidades estaduais nestas categorias foram criadas em 2008 (RESEX Ilha do Tumba, RESEX Taquari, RDS Barreiro Anhemas, RDS dos Pinheirinhos, RDS Itapanhapima, RDS Lavras e RDS Quilombos de Barra do Turvo) e 2010 (MONAT da Pedra do Baú e MONAT da Pedra Grande). A Figura 38 mostra o número de UCs estaduais e federais criadas no Estado de São Paulo. Em 2008 foram 24 unidades estaduais criadas, sendo 21 de Uso Sustentável (incluindo cinco RPPNs) e 3 de Proteção Integral).

Figura 38 - Criação de UCs estaduais e federais de 1934 a 2021



Fonte: CNUC, 2022<sup>69</sup>. Organizado pela autora.

No estado de São Paulo, a criação de UCs é incentivada pelo ICMS Ecológico<sup>70</sup> que concede benefícios fiscais aos municípios que abrangem territórios protegidos, sendo este um critério para repasse da quota municipal do ICMS.

#### 4.1.2 Cobertura da terra nas Unidades de Conservação do estado de São Paulo

Embora todas as variáveis tenham sua importância na adequabilidade de habitat, a cobertura da terra está relacionada, de alguma forma, às demais variáveis. Essa relação é uma ação mútua, seja na regulação do microclima (DUFEK; AMBRIZZI, 2007; DODONOV et al., 2013; SOUZA et al., 2016), na demanda para instalação de luzes noturnas artificiais (KIM; HONG, 2015; SMALL; SOUSA, 2016; MA, 2018), ou na dependência da configuração do relevo para o estabelecimento de atividades e usos (JASINSKI et al., 2005; GRECCHI et al.,

<sup>69</sup> CNUC - Cadastro Nacional de Unidades de Conservação. Disponível em <<http://sistemas.mma.gov.br/cnuc/index.php?ido=relatorioparametrizado.exibeFormularioPortal>>. Acesso em 15 jul. 2022.

<sup>70</sup> Lei Estadual nº 8.510, de 29 de dezembro de 1993. Disponível em <<https://smastr16.blob.core.windows.net/cpla/2011/12/1993-Lei-8510.pdf>>. Acesso em 17 nov. 2022.



2014; SPERA et al., 2014; MOLIN et al., 2017; ESPINDOLA et al., 2021). As mudanças no uso e cobertura da terra são significativas para a perda de biodiversidade e de serviços ecossistêmicos (HAINES-YOUNG, 2009) e essa relação é essencial para a gestão da biodiversidade em qualquer táxon (WOOD et al., 2017). Variáveis ambientais restringem ou viabilizam a ocorrência de uma espécie, dada suas características fisiológicas e comportamentais e capacidade de adaptação e dispersão. Entretanto, ainda que um ambiente tenha condições climáticas e de relevo adequadas, a cobertura da terra é um fator determinante para a adequabilidade de habitat para mamíferos de médio e grande porte.

O SNUC institui que nas UCPIs é admitido apenas o uso indireto dos recursos naturais e que nas UCUs o uso de parte dos recursos deve ser compatível com a conservação da natureza. Ou seja, espera-se que os ecossistemas que abrigam esses recursos naturais sejam minimamente conservados. As APAs incluem áreas de domínios público e privado e têm como principal objetivo regulamentar o processo de ocupação; neste caso, a manutenção da cobertura vegetal nativa é uma exigência legal no caso das Áreas de Preservação Permanente (APPs) e das Reservas Legais (RLs) e, portanto, não é garantido que essas UCs tenham alto percentual de formações naturais florestais. As RPPNs, não contempladas nesta pesquisa, embora sejam de domínio privado assim como as APAs, têm como objetivo a conservação da biodiversidade e, conforme o SNUC, são “gravadas com perpetuidade” por meio da averbação do imóvel. As unidades dessa categoria são administradas pelos proprietários interessados na conservação do local.

A cobertura da terra em UCs deve ser consoante as normas especificadas no manejo e de acordo com a categoria à qual pertence. As Figuras 39 a 42 mostram o percentual de cobertura da terra associado a formações naturais florestais nas UCs do estado de São Paulo, com a indicação da data de criação da unidade (marcador preto), quando consta no CNUC. As células em cinza referem-se a pixels ‘*no data*’ no dado do MapBiomas. Ressalta-se que grandes percentuais de formações naturais florestais não são indicativos da qualidade dos remanescentes, que podem ter certo grau de antropização e sentir os efeitos de borda e da fragmentação de habitat. Como mostrado anteriormente, vários pequenos fragmentos, quando sua extensão é somada, podem resultar em grandes quantitativos.

As APAs, em geral, têm cerca de 40% de formações naturais florestais. A principal exceção é a APA Serra do Mar, que faz parte de um mosaico de UCs de outras categorias (ARIE, ESEC, PE, RDS, RESEX, RVS, PARNA e REBIO) no litoral paulista e sobrepõe-se a

territórios indígenas (TI Amba Porã, TI Djaiko-aty, TI Peguaoty e RI Takuari Eldorado), mantendo 88% de formação nativa no período analisado. A APA Mata do Iguatemi é um pequeno fragmento de Formação Florestal entremeadado à área densamente urbanizada no Município de São Paulo e que tem sofrido a pressão da urbanização ao longo do período de análise, com diminuição do percentual de formações naturais florestais. As APAs Tietê (Tietê, criada em 1983), Ibitinga (Ibitinga, criada em 1987), Represa do Bairro da Usina (Atibaia, criada em 1986) e Rio Batalha (Centro-Oeste do estado, criada em 2001) já eram dominadas por atividades agrossilvopastoris na época da sua criação e seu percentual de formações naturais florestais permaneceu abaixo de 25% no período analisado.

O perímetro da ARIE Buriti de Vassununga (Santa Rita do Passa Quatro) não segue limites geográficos ou políticos<sup>71</sup>. Trata-se de uma área que pertencia à Usina de Vassununga mas que, após a falência da fábrica, foi desapropriada pelo Estado para preservar remanescentes do Cerrado em 1990, quando então foi criada a UC. ARIE é uma categoria de Uso Sustentável que, segundo o SNUC, apresenta “características naturais extraordinárias ou que abriga exemplares raros da biota regional”, mas que admite a exploração dos recursos naturais em consonância com a conservação da natureza. Embora seja localizada em uma área bem antropizada dominada por paisagens agrícolas (61% por cana e 26% por Mosaico de Agricultura e Pastagem), seu percentual de formações naturais florestais vem aumentando desde o início do período de análise.

---

<sup>71</sup> Área de Relevante Interesse Ecológico Buriti de Vassununga. Disponível em <<https://uc.socioambiental.org/arp/894>>. Acesso em 16 nov. 2022.

Figura 39 - Percentual de formação natural florestal nas UCs no período de 1985 a 2022 (I)

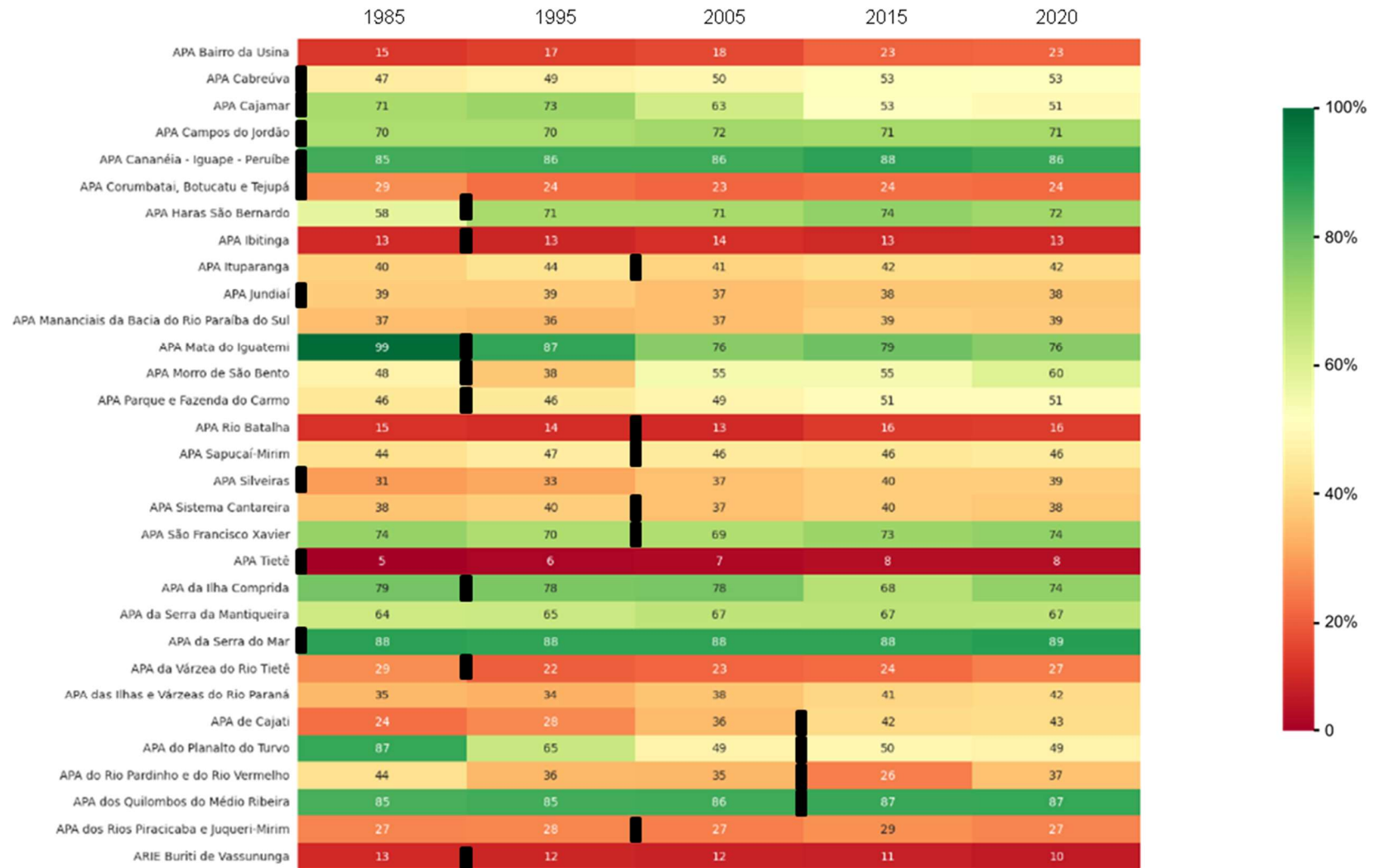


Figura 40 - Percentual de formação natural florestal nas UCs no período de 1985 a 2022 (II)

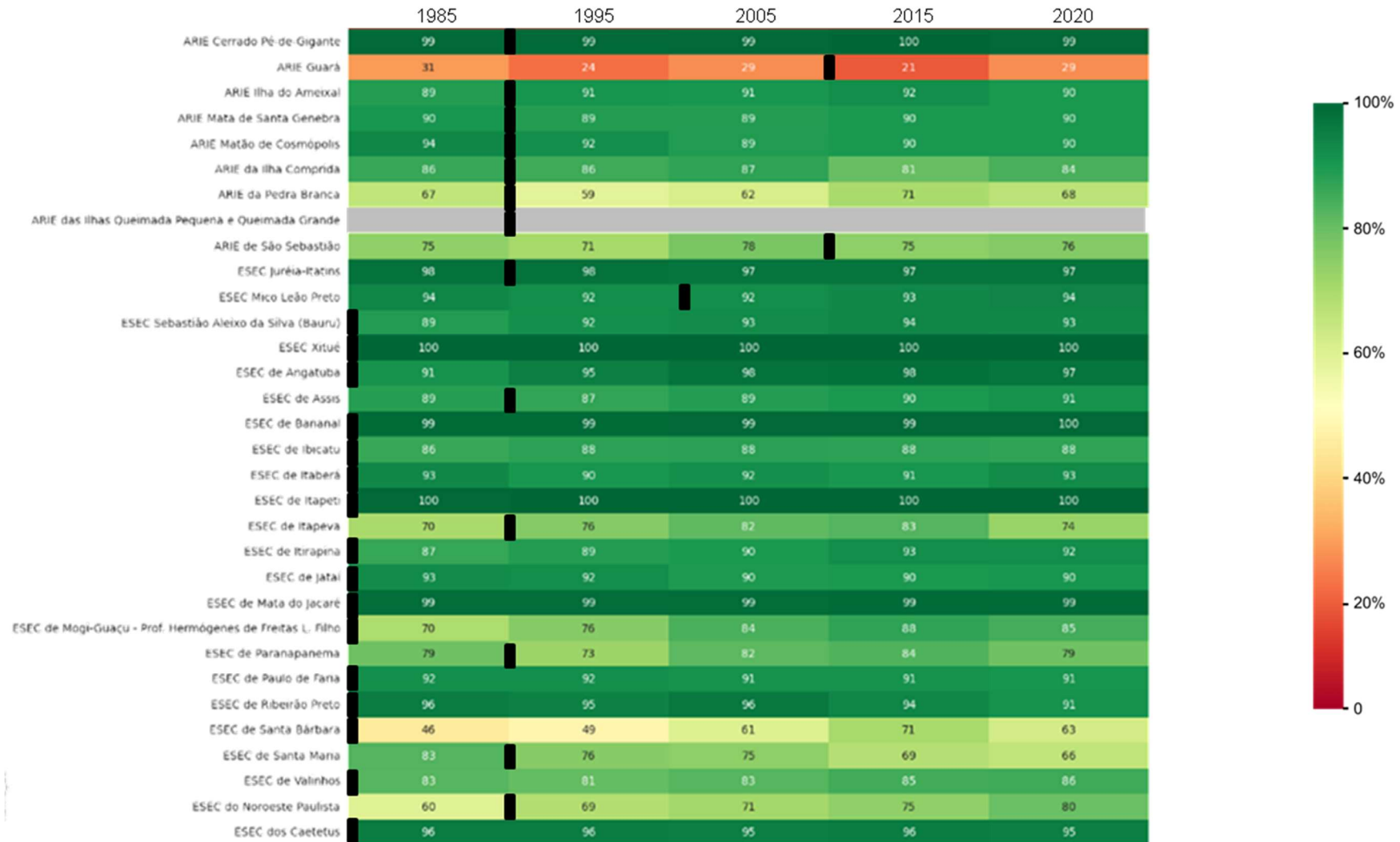


Figura 41 - Percentual de formação natural florestal nas UCs no período de 1985 a 2022 (III)

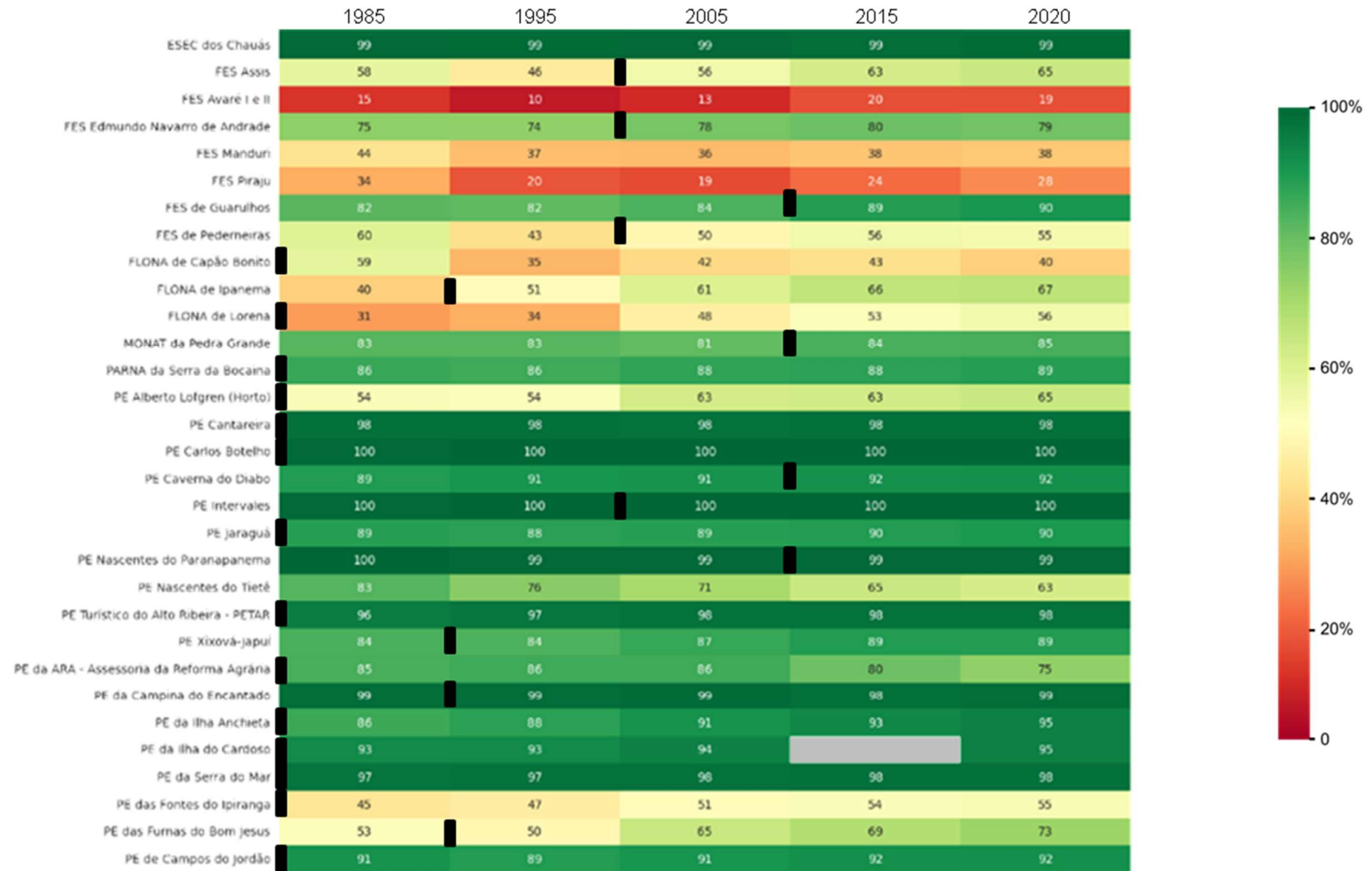
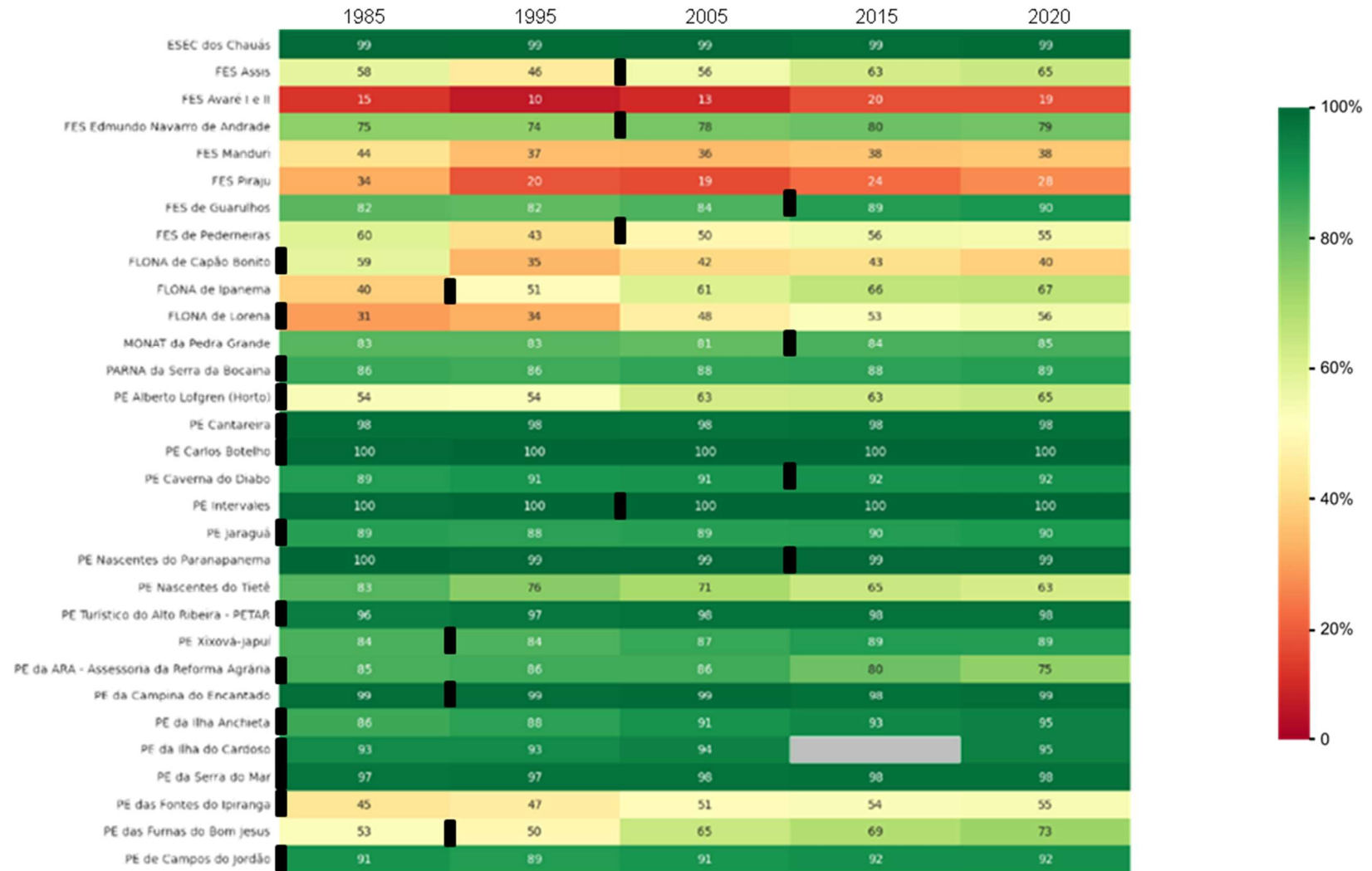


Figura 42 - Percentual de formação natural florestal nas UCs no período de 1985 a 2022 (IV)



A ESEC de Santa Bárbara (Santa Bárbara) também tem baixo percentual de formações naturais florestais. Criada em 1984, tem 35% de atividades agrossilvopastoris e seu Plano de Manejo relata como atividades conflitantes os reflorestamentos comerciais, a criação extensiva de gado, caça, pesca e impactos das rodovias SP-330 e SP-261<sup>72</sup>. Há contradições em relação ao mapeamento de cobertura da terra pelo MapBiomias (2020) e a caracterização apresentada no Plano de Manejo (2011). O dado do MapBiomias classifica como Pastagem áreas que no Plano de Manejo constam como Cerrado típico e Campo Cerrado. Essa contradição acontece porque essas classes de cobertura da terra são espectralmente similares e as classificações de imagens orbitais estão sujeitas a erros. Para o bioma Cerrado, a coleção 6 do MapBiomias tem acurácia geral de 74,9%. Essas inconsistências refletem na acurácia da modelagem caso haja pontos amostrais onde a cobertura da terra não está correta. Há contradições no mapeamento de outras UCs. O PE do Juquery (Franco da Rocha e Caieiras) está localizado na Região Metropolitana de São Paulo onde encontra-se o único remanescente de Cerrado da Região (GASPARETO, 2014). Segundo Baitello et al. (2013), o PE tem formações campestre, savânica e florestal em região montanhosa, onde o MapBiomias classifica como Pastagem e Mosaico de Agricultura e Pastagem. As RDSs Barreiro Anhemas e dos Pinheirinhos, ambas no município de Barra do Turvo, na Serra do Mar, também têm baixas proporções de formações naturais florestais, respectivamente 28% e 39%.

Há UCs que não apresentam alto percentual de Formação Florestal Nativa, mas que abrangem coberturas que não estão associadas a antropismos. É o caso dos PE do Rio Peixe (Oeste do Estado) e do Aguapeí (Noroeste do Estado) são localizados na várzea de rios onde o MapBiomias classifica corretamente como Campo Alagado e Área Pantanosa; esses campos são entremeados por formações arbustivas classificadas como Mosaico de Agricultura e Pastagem, que podem ser, na verdade, formações campestres.

Os PE Alberto Löfgren e Fontes do Ipiranga são UCs urbanas no município de São Paulo que, apesar da pressão da urbanização, têm predominância de Formação Florestal (92% e 60%, respectivamente) e sua proporção de formações naturais florestais aumentou no período de análise. As FLONAs de Ipanema e de Lorena também incrementaram a extensão coberta por formações naturais florestais, de 40% para 67% e de 31% para 56%.

---

<sup>72</sup> Plano de Manejo - ESEC de Santa Bárbara. Disponível em <[https://www.infraestruturameioambiente.sp.gov.br/institutoflorestal/wp-content/uploads/sites/234/2013/03/Plano\\_de\\_Manejo\\_EEc\\_Santa\\_Barbara.pdf](https://www.infraestruturameioambiente.sp.gov.br/institutoflorestal/wp-content/uploads/sites/234/2013/03/Plano_de_Manejo_EEc_Santa_Barbara.pdf)>. Acesso em 16 nov. 2022.

Para as UCs nas quais a data de criação consta no CNUC e que foram criadas após o início do período de análise, não houve diferenças significativas nos percentuais de formações naturais florestais a partir do seu decreto de criação. As pequenas variações positivas podem referir-se a regeneração ou restauração florestal. O Plano Conceitual de Manejo do PE das Furnas do Bom Jesus (Pedregulho, criado em 1989) redigido por Branco et al. (1991), relata que a porção do parque onde o relevo é mais suave foi substituída por cafeicultura e pastagens. Nos 35 anos de análise, o percentual de formações naturais florestais aumentou de 53% para 73%.

As Figuras 43 a 46 apresentam os histogramas empilhados com a proporção de cada cobertura da terra. As cores da legenda seguem a paleta padrão do Projeto MapBiomias. Como previsto na hipótese, as UCPI têm predominância de formações naturais florestais, principalmente Formação Florestal. Formações naturais não florestais, como Campo Alagado e Formação Campestre também são relevantes para a manutenção da biodiversidade e são marcantes no PE do Rio Peixe, no PE do Aguapeí e na ESEC de Santa Bárbara.



Figura 43 - Proporção das classes de cobertura da terra nas UCs em 2020 (I)

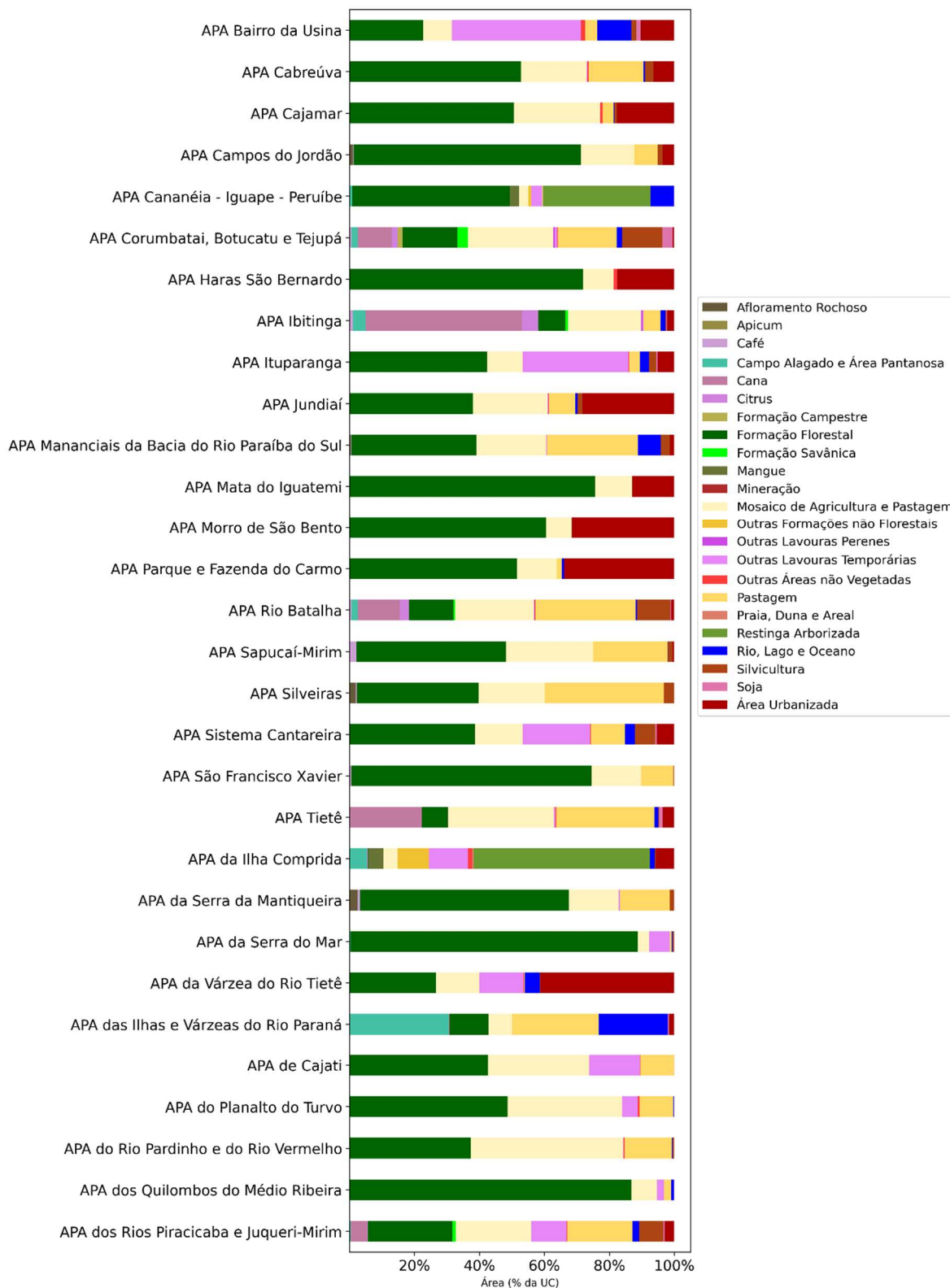


Figura 44 - Proporção das classes de cobertura da terra nas UCs em 2020 (II)

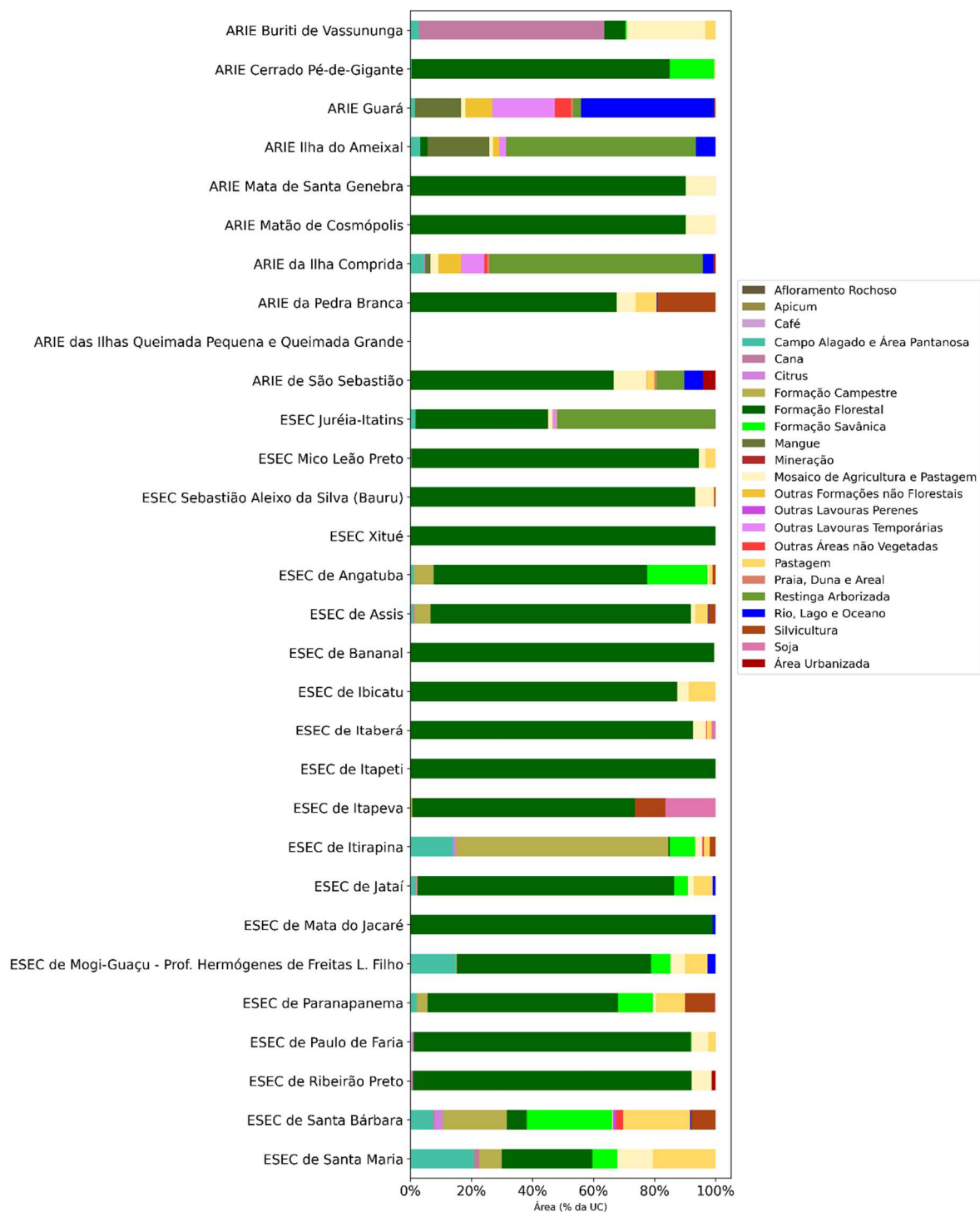


Figura 45 - Proporção das classes de cobertura da terra nas UCs em 2020 (III)

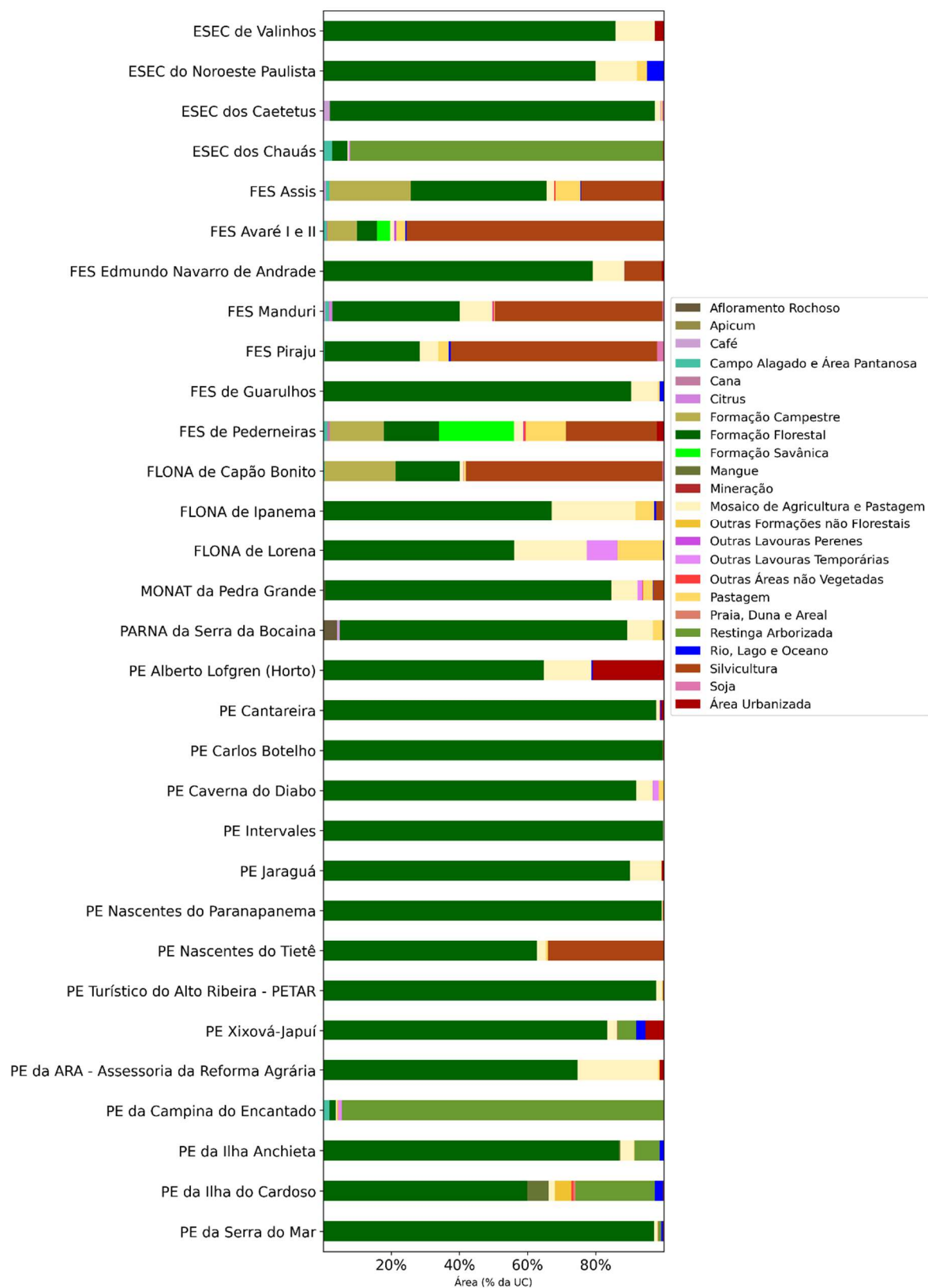
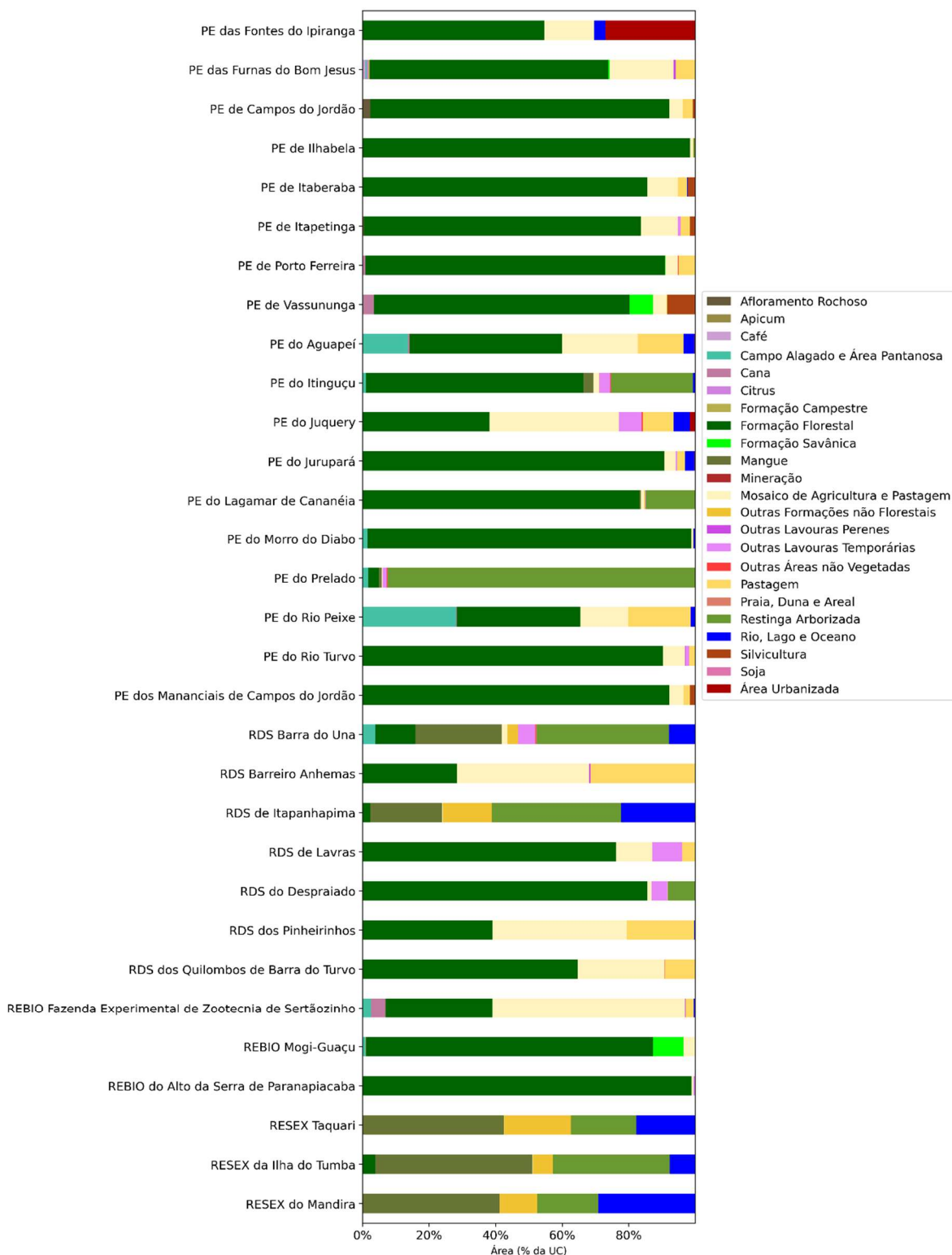


Figura 46 - Proporção das classes de cobertura da terra nas UCs em 2020 (IV)



As Figuras 47 a 50 apresentam o percentual de coberturas antrópicas com vegetação, associadas a atividades agrossilvopastoris, no período de 1985 a 2020, evidenciando a produtividade econômica nas UCs. As APAs, particularmente, são pouco restritivas quanto às atividades permitidas dentro do seu perímetro e no entorno e o percentual de coberturas

antrópicas nesta categoria reflete essa flexibilização. As outras categorias de UCUS também são economicamente produtivas. Por exemplo, na RDS Quilombos de Barra do Turvo, 40% da área total da unidade é usada para criação de gado e/ou monocultivo de pupunha, mandioca e cana (SANTOS; GARAVELLO, 2016). Na RESEX do Mandira destaca-se o extrativismo de ostras e caranguejos no manguezal da região, além do cultivo de arroz, milho, banana e mandioca e produção de recursos não madeireiros destinados à produção de medicamentos caseiros e à alimentação<sup>73</sup>. Na RESEX Taquari, a extração vegetal envolve o uso de recursos para produzir material de insumo às atividades pesqueiras, construção de casas, lenha, alimentação, artesanato e para fins medicinais<sup>74</sup>. Ou seja, a existência de áreas protegidas não é entrave para a produtividade econômica.

Vale ressaltar que as extensas paisagens agrícolas no estado de São Paulo constituem, principalmente, monoculturas de cana-de-açúcar, pastagem e silvicultura e, portanto, não estão diretamente associadas à produção de alimentos. Em escala nacional, 85% da produção de alimentos da cesta básica brasileira é proveniente da agricultura familiar (MUSSOI, 2003), que consiste em apenas 23% da área total das propriedades agropecuárias no Brasil<sup>75</sup>. Sendo assim, a manutenção de áreas protegidas em detrimento dos grandes espaços agrícolas, não é um entrave para a produção de alimentos no país. Vijay e Armsworth (2021) explicam que todas as categorias de áreas protegidas da IUCN contêm terras agrícolas, embora estas sejam mais associadas às categorias menos restritivas (Categorias V e VI da IUCN), áreas protegidas não categorizadas e até mesmo em áreas estritamente protegidas (Categoria IV).

Jesus e Catojo (2018) constataram que os grandes fragmentos florestais no Cerrado paulista (> 1.000 ha) estão concentrados nas Cuestas Basálticas, onde há áreas com altitude superiores a 1.000 metros e declividade superior a 49°, localizadas em parte na APA Corumbataí, Botucatu e Tejuapá. Tratam-se de APPs, onde é vedada a supressão da vegetação nativa, conforme a Lei de Proteção da Vegetação Nativa (Lei 12.651/2012). Por outro lado, a várzea do rio Grande, na divisa de São Paulo e Minas Gerais, que também constitui uma APP, é dominada pelo cultivo de cana-de-açúcar na porção noroeste do estado paulista, onde o terreno

---

<sup>73</sup> Plano de Manejo Participativo - Reserva Extrativista do Mandira. Disponível em <<https://www.gov.br/icmbio/pt-br/assuntos/biodiversidade/unidade-de-conservacao/unidades-de-biomas/mata-atlantica/lista-de-ucs/resex-do-mandira/arquivos/plano-manejo-mandira-2010.pdf>>. Acesso em 17 nov 2022.

<sup>74</sup> Plano de Manejo - Reserva Extrativista Taquari. Disponível em <>. Acesso em 17 nov 2022.

<sup>75</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. **Atlas do espaço rural brasileiro**. Rio de Janeiro: IBGE, 2020. 321 p. 2020. Disponível em <<https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2101773>>. Acesso em 18 nov. 2022.

é mais plano e viabiliza a mecanização agrícola. Ou seja, a restrição ao uso da terra está mais relacionada a aspectos ambientais do que legais. Meirelles Filho e Borges (2016) apontam, de forma análoga, que na Amazônia, as florestas de terra firme são convertidas integralmente para outras coberturas e as várzeas são resguardadas. Segundo os autores, o uso que é feito da terra é baseado no benefício econômico e ao custo de implantação da atividade produtiva e não é justificado pela manutenção de trechos das formações nas propriedades. Ou seja, as variáveis ambientais, além de influenciar na disponibilidade de recursos, também podem definir o uso do solo, e ambos podem configurar a Adequabilidade de Habitat. Esta constatação vai ao encontro do que é proposto por Guisan (2017) quanto à ocorrência de espécies no espaço geográfico.

No contexto da gestão de UCs, mais especificamente quanto à elaboração dos Planos de Manejo, esta questão conduz à importância da inclusão de restrições mais específicas às Zonas de Amortecimento e não apenas mencionar diretrizes abrangentes. A área protegida como Unidade de Conservação deve reforçar a proteção de áreas já salvaguardadas em outras formas como APPs e RLs. Particularmente nas APAs, onde a implantação de empreendimentos potencialmente negativos para o meio ambiente pode ser tão acessível quanto em áreas não protegidas, sua ação efetiva como UC ficou comprometida com o Projeto de Lei 2.159/2021<sup>76</sup>, aprovado na Câmara dos Deputados e em tramitação no Senado, que dispensa do licenciamento ambiental diversas atividades, incluindo a agricultura e a pecuária, além de obras de melhoria em instalações já existentes, mesmo em UCs. O PL ainda confere autonomia aos estados e municípios para a adoção de sistemas de concessão das licenças ambientais.

---

<sup>76</sup> Projeto de Lei nº 2159, de 2021. Disponível em <<https://www25.senado.leg.br/web/atividade/materias/-/materia/148785>>. Acesso em 17 nov. 2022.

Figura 47 - Percentual de coberturas antrópicas com vegetação nas UCs (I)

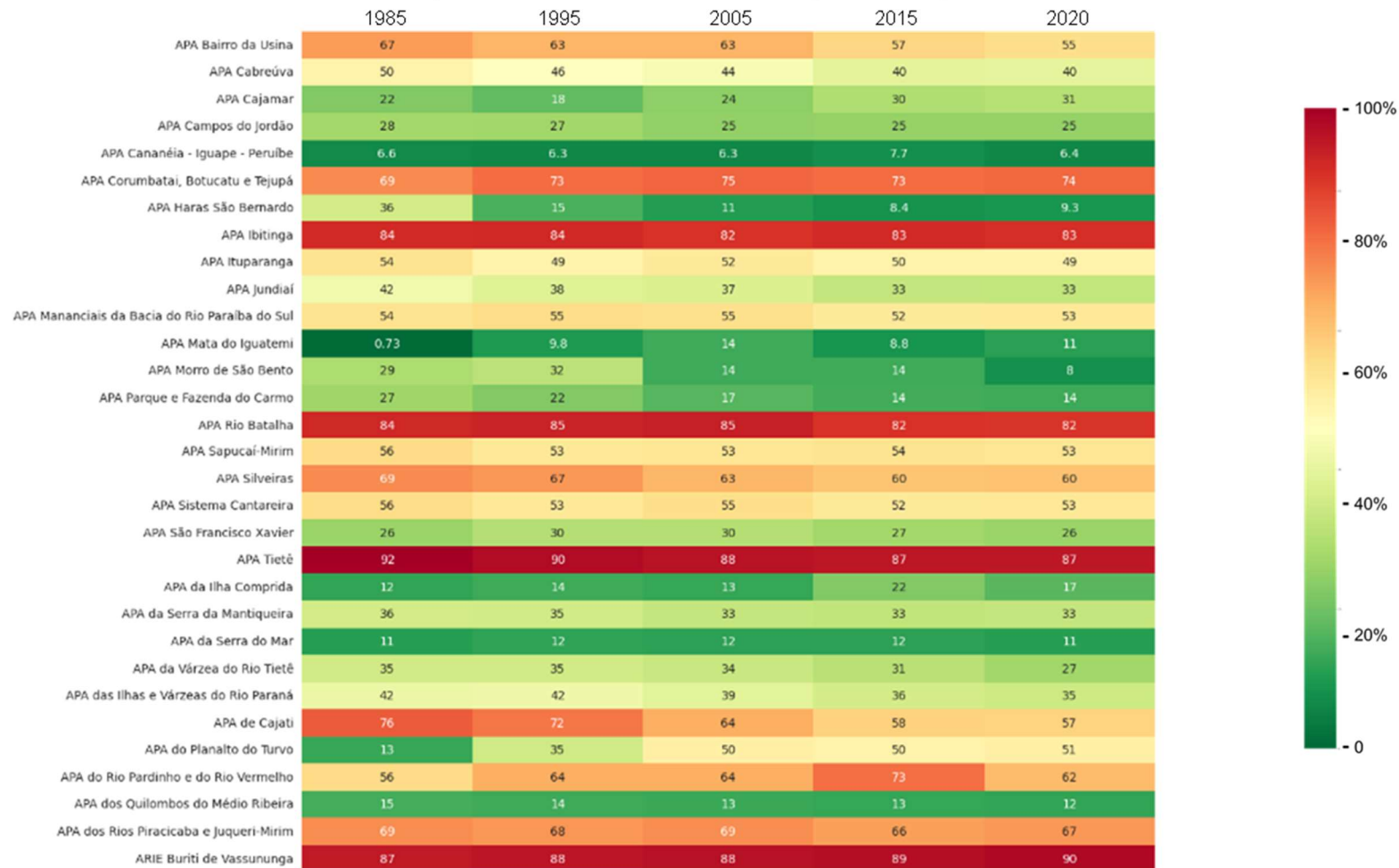


Figura 48 - Percentual de coberturas antrópicas com vegetação nas UCs (II)

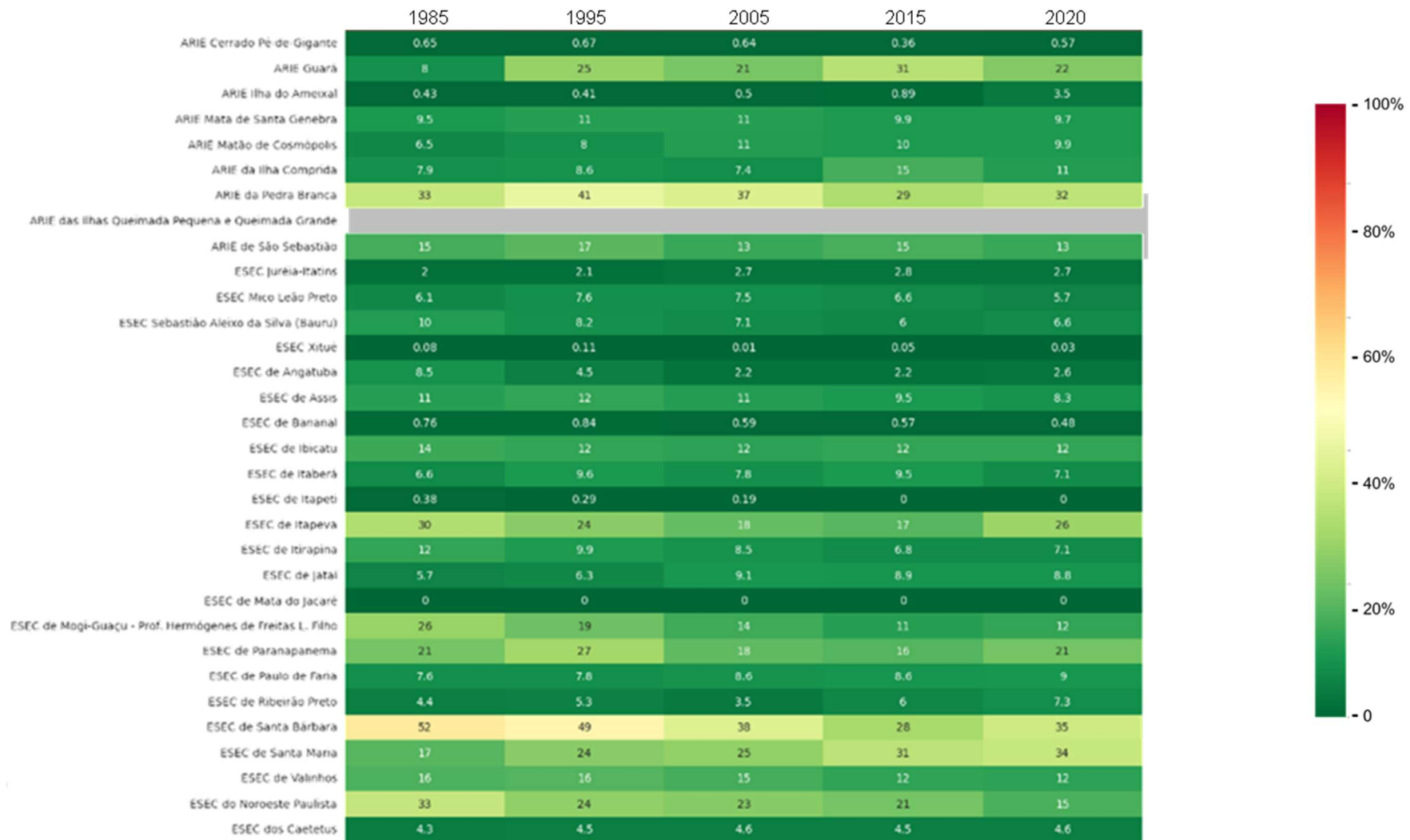




Figura 49 - Percentual de coberturas antrópicas com vegetação nas UCs (II)

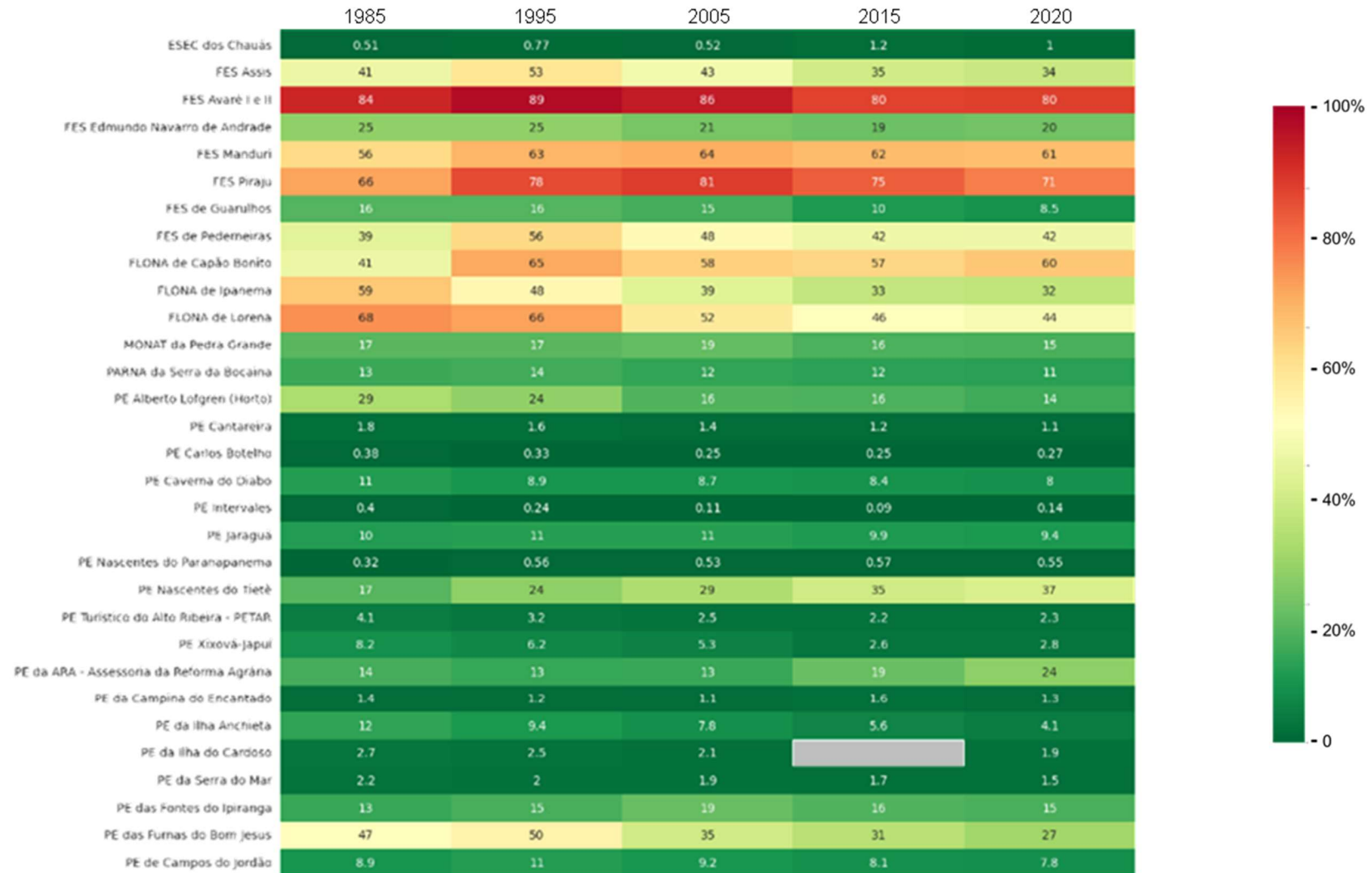
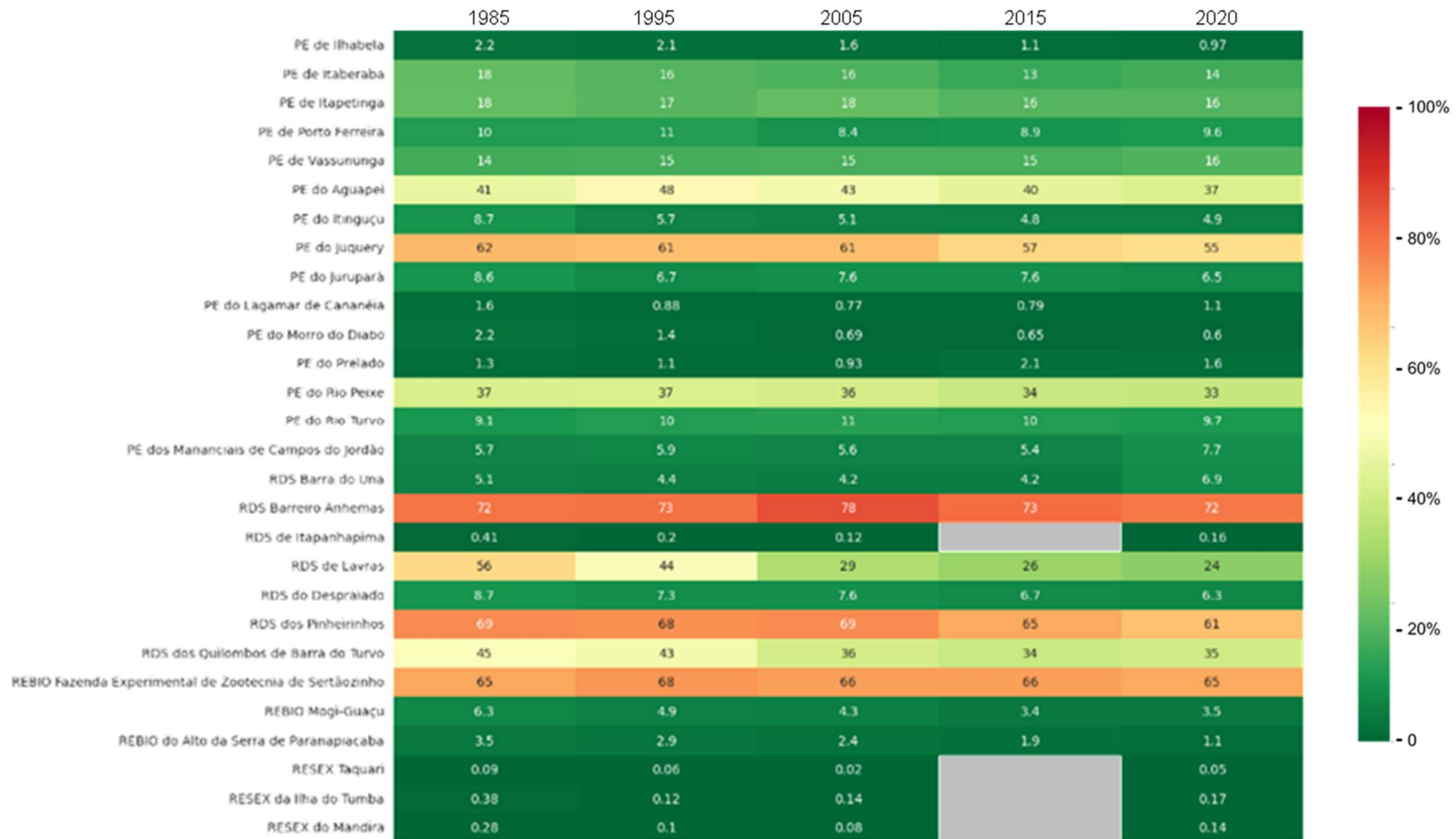


Figura 50 - Percentual de coberturas antrópicas com vegetação nas UCs (IV)



### 4.1.3. Cobertura da terra no estado de São Paulo

#### 4.1.3.1 Remanescentes de formações nativas no Estado de São Paulo

As estratégias que as espécies usam para persistir nos ecossistemas antropizados é um resultado da variação espaço-temporal na estrutura da paisagem; por exemplo a heterogeneidade temporal dos agroecossistemas, com a grande variação da produção de biomassa ao longo do ciclo das culturas, pode promover novos microhabitats que sustentam espécies oportunistas e generalistas de pequenos mamíferos (GHELER-COSTA et al., 2012; TIMO et al., 2014).

Dotta e Verdade (2011) avaliaram os padrões de distribuição e abundância de mamíferos de médio a grande porte em uma paisagem agrícola no sudeste do Brasil, com o registro de espécies nativas, exóticas e domésticas; as áreas de pastagem apresentaram a menor riqueza de espécies e os canaviais a maior frequência de ocorrência, sendo que a comunidade encontrada é predominantemente composta por espécies capazes de se adaptar às alterações ambientais. Isto porque a fauna silvestre usa os espaços agrícolas como rotas entre remanescentes de vegetação nativa, o que reforça a necessidade de manter fragmentos de vegetação nativa entremeados às paisagens agrícolas, visto que estas últimas podem ter um papel importante na conservação de mamíferos.

Em relação aos pequenos mamíferos, Gheler-Costa et al. (2012) constataram que os fragmentos de floresta semidecídua são ocupados por apenas algumas espécies generalistas, dado que no local de estudo predominam fragmentos em forma de tiras que potencializam o efeito de borda. Todavia, algumas espécies conseguem manter alta densidade populacional em paisagens heterogêneas e, portanto, assegurar a base da cadeia trófica, o que, de acordo com Gheler-Costa et al. (2012), confere certo valor de conservação às paisagens agrícolas.

A Figura 51 mostra a distribuição espacial dos remanescentes florestais em São Paulo, com oito classes de tamanho, considerando a classificação pelo método de quebras naturais de Jenks<sup>77</sup>. Os intervalos das classes são apresentados na Tabela 15.

---

<sup>77</sup> O método de otimização de Jenks agrupa dados definindo o melhor arranjo de valores em diferentes classes, minimizando o desvio médio de cada classe em relação à média da classe, e maximizando o desvio de cada classe em relação às médias das outras classes.

Figura 51 - Distribuição dos fragmentos de Formações Florestais no estado de São Paulo

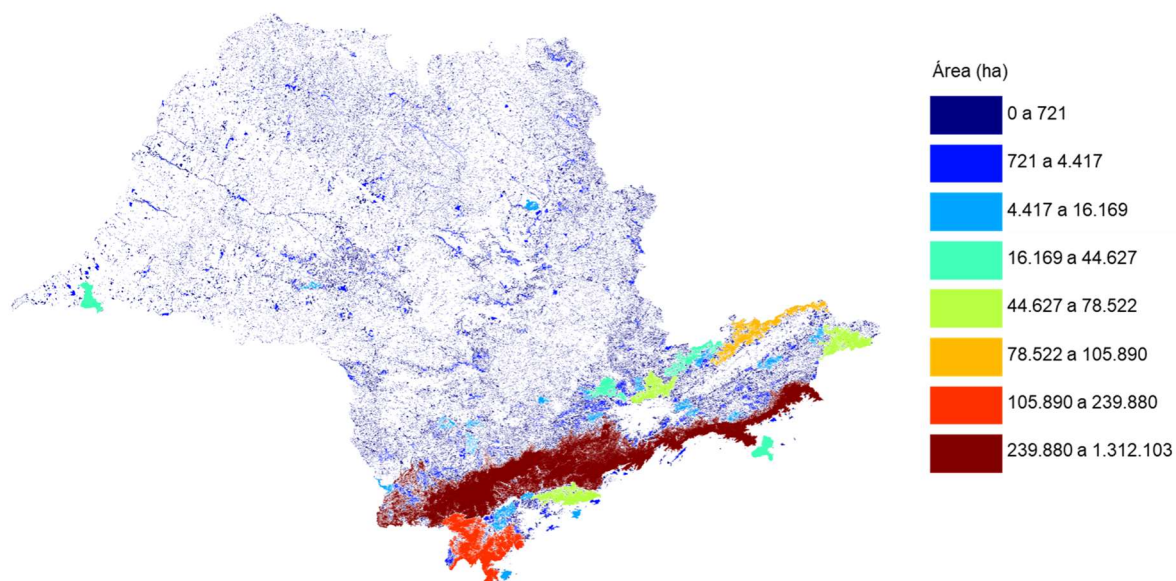


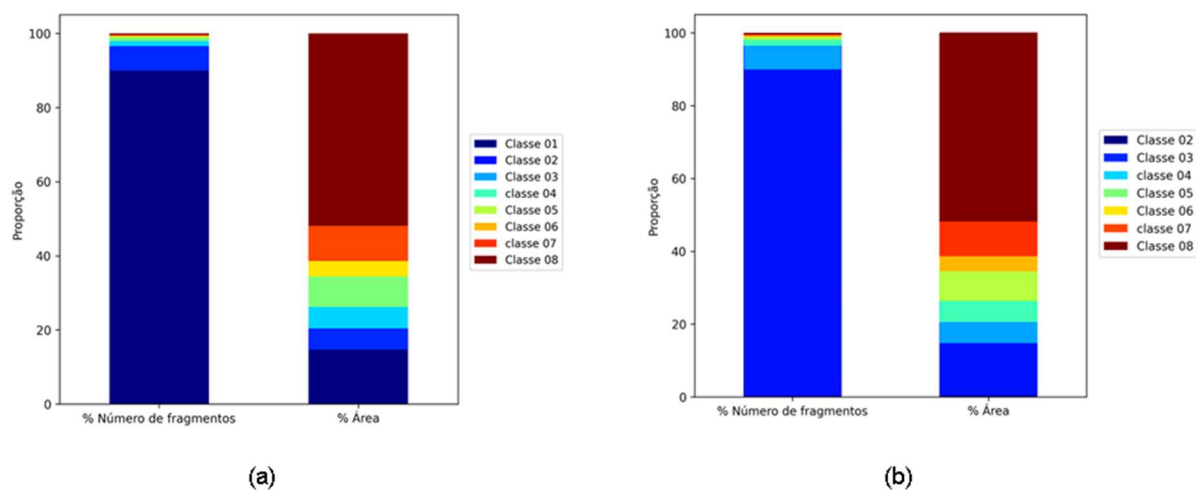
Tabela 15 - Intervalos das classes de tamanho dos fragmentos de Formações Florestais

<b>Classe</b>	<b>Intervalo de tamanho (ha)</b>
1	$\leq 721$
2	( 721 - 4.417 ]
3	( 4.417 - 16.169 ]
4	( 16.169 - 44.627 ]
5	( 44.627 - 78.522 ]
6	( 78.522 - 105.890 ]
7	( 105.890 - 239.880 ]
8	( 239.880 - 1.312.103 ]

A Figura 52a mostra a proporção das classes de tamanho quanto ao número de fragmentos e à área. Esta análise foi feita a partir da vetorização da classe “Formação Florestal” do MapBiomas, com tamanho de pixel de 30 metros (0,09 ha), o que resulta em múltiplos pixels isolados. Sendo assim, a Figura 52b mostra a distribuição das classes desconsiderando os pixels muito pequenos ( $< 721$  ha, que é a primeira quebra calculada pelo método de Jenks). As

proporções mostram que os grandes fragmentos (> 44.627 ha, que é a quarta quebra pelo método de Jenks) contribuem com quase metade da extensão de florestas (40%), embora constituam poucos fragmentos (menos de 1% do número de fragmentos). Por outro lado, os fragmentos muito pequenos (< 721 ha), compõem quase que a totalidade de Formações Florestais em termos de número e contribuem com quase metade (46%) da área desta cobertura. Isto significa que, embora as Formações Florestais cubram 19% do território paulista, grande parte dos remanescentes são muito pequenos e não são suficientes para manter populações de mamíferos, especialmente os carnívoros (VETTER et al., 2011) de grande porte que requerem grandes áreas para sobrevivência (DIAS et al., 2012). Os pequenos mamíferos, por sua vez, também são afetados pela fragmentação da paisagem, pois têm áreas de vida menores e são habituados a percorrer curtas distâncias durante a dispersão. Entretanto, sua estratégia de reprodução, com tempos de geração mais curtos e produção elevada de descendentes a cada ciclo reprodutivo, possibilita que algumas espécies driblem tal adversidade (PIANKA, 1970; GHELIER-COSTA et al., 2012). Vale ressaltar que, mesmo para os grandes remanescentes, a existência da Formação Florestal não significa que não há certo grau de degradação e, portanto, de efeito de borda.

Figura 52 - Proporção dos remanescentes de vegetação nativa no estado de São Paulo



Proporção em termos de: a) número de fragmentos e área total para cada classe de tamanho; e b) número de fragmentos e área total dos fragmentos para cada classe de tamanho com a remoção dos fragmentos da Classe 01.

Crooks et al. (2017) mostraram que mamíferos que vivem em habitats fragmentados correm maior risco de extinção. A fragmentação, além de causar o efeito de borda, também está relacionada à limitação de dispersão e colonização, impedindo o trânsito de animais dado o risco de predação (PRIMACK; RODRIGUES, 2001). No entanto, embora não sustentem populações viáveis de grandes mamíferos, os fragmentos podem servir como trampolins ecológicos que mantêm a conectividade entre os grandes remanescentes (DIAS et al., 2012).

Os pequenos fragmentos estão em todo o Estado, principalmente em áreas de Reserva Legal e Áreas de Preservação Permanente (APPs) ao longo de corpos d'água. A Lei de Proteção da Vegetação Nativa (Lei 12.651/2012, LPVN)<sup>78</sup> proporcionou alguns retrocessos em relação às leis precedentes que versam sobre as APPs e a respeito das sanções conferidas à supressão florestal. O Código Florestal (Lei 4.771/1965)<sup>79</sup> estabelecia como APPs as faixas ao longo dos cursos d'água a partir do leito máximo; então, a LPVN encolheu essa faixa a partir da borda da calha do leito regular. Isso isentou os proprietários rurais de manter ou recuperar a vegetação em uma faixa maior. Além disso, a alteração na definição sobre o que são “áreas consolidadas” oportunizou a suspensão do reflorestamento de áreas que também deveriam ser recuperadas. Nas regiões Centro-Oeste e Noroeste do Estado, onde a cobertura da terra é predominantemente de atividades agrossilvopastoris, a manutenção das APPs e RLs é essencial para manter a conectividade entre os fragmentos de vegetação nativa.

No Pontal do Paranapanema, oeste do Estado, o PE do Morro do Diabo (Figura 53A) abriga um dos maiores remanescentes contínuos da Mata Atlântica (Floresta Estacional Semidecidual), criado na década de 1940 após a invasão e supressão de reservas por grileiros. No entanto, é na Serra do Mar que estão presentes os grandes fragmentos de formação florestal nativa (Figura 53C) dado o relevo acidentado que dificulta a mecanização e instalação da agroindústria. As mesorregiões Litoral Sul Paulista, Campinas e Vale do Paraíba Paulista também abrigam fragmentos de área expressiva, preservados nas UCPI (Figura 53B, D e E). Na divisa com o sul do Estado de Minas Gerais, o mosaico de APAs e PE ainda sustentam

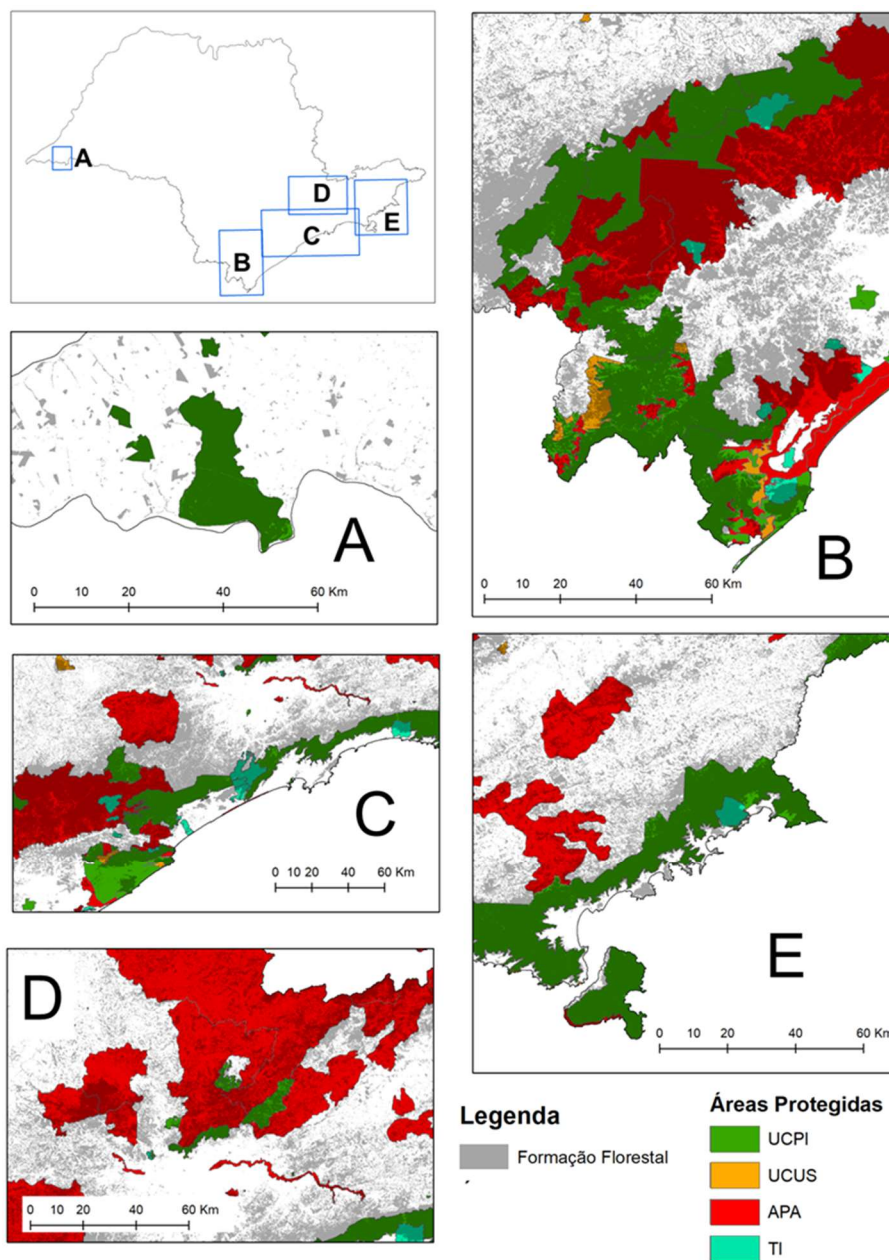
---

<sup>78</sup> Lei 12.651 de 25 de maio de 2012. Dispõe sobre a proteção da vegetação nativa; altera as Leis n°s 6.938, de 31 de agosto de 1981, 9.393, de 19 de dezembro de 1996, e 11.428, de 22 de dezembro de 2006; revoga as Leis n°s 4.771, de 15 de setembro de 1965, e 7.754, de 14 de abril de 1989, e a Medida Provisória n° 2.166-67, de 24 de agosto de 2001; e dá outras providências. Disponível em <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2012/lei/112651](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/112651)>. Acesso em 07 out. 2022.

<sup>79</sup> Lei 4.771 de 15 de setembro de 1965. Institui o novo Código Florestal. Disponível em <[http://www.planalto.gov.br/ccivil\\_03/leis/14771](http://www.planalto.gov.br/ccivil_03/leis/14771)>. Acesso em 07 out. 2022.

fragmentos com mais de 15.000 ha. São áreas protegidas por UCs de Proteção Integral e de Uso Sustentável e também por Terras Indígenas (TIs).

Figura 53 - Localização de grandes remanescentes (> 15 mil ha) de Formações Florestais protegidos por UCs e TIs



53A: PE do Morro do Diabo; 53B: Litoral Sul Paulista; 53C: Serra do Mar; 53d: Mesorregião de Campinas; 53E: Mesorregião do Vale do Paraíba Paulista.

As Cuestas Basálticas também mantêm área florestal nativa, principalmente nos municípios de Analândia, Descalvado, São Carlos, Santa Rita do Passa Quatro, Santa Rosa do Viterbo e Altinópolis (RONQUIM, 2018), e que, em princípio, deveria ser protegida pela APA Corumbataí. Na APA Corumbataí, Botucatu e Tejuπά, os remanescentes de vegetação nativa estão principalmente nas APPs de encosta onde a inclinação do terreno inviabiliza a mecanização agrícola. As Figuras 54 e 55 comparam a localização dos remanescentes nativos e das áreas com alta inclinação nos perímetros Corumbataí e Botucatu, respectivamente.

Os maiores remanescentes de Formação Savânica (>5 km<sup>2</sup>) estão na Bacia Hidrográfica do Tietê/Jacaré e fora de UCs. Quando quantificados por município, o maior está em Bauru (1.615 ha), fora do perímetro da APA Rio Batalha, em uma porção de Cerrado stricto sensu que inclui o Jardim Botânico de Bauru. As APPs ao longo dos rios mantêm a conectividade desse grande remanescente com fragmentos nos municípios de Pederneiras (4.230 ha) e Agudos (104.474 ha). Neste último, os remanescentes estão em uma propriedade privada que inclui extensas glebas de Floresta Plantada<sup>80</sup>. Trata-se de uma das Unidades Florestais da Duratex S/A, que mantém programas ambientais previstos no seu Plano de Manejo Florestal<sup>81</sup>.

Na região das Cuestas Basálticas, os municípios de Altinópolis, Pedregulho e Patrocínio Paulista têm cerca de 7 mil, 1,2 mil e 1 mil hectares de Formações Savânicas não protegidas por qualquer categoria de UC, apesar do seu potencial turístico (VOLTOLINI, 2011). Paranapanema, São Carlos, Ribeirão Bonito, Brotas, Araraquara e Ibaté também abrigam remanescentes savânicos com mais de 500 ha.

A Estação Ecológica de Jataí, localizada no município de Luís Antônio, abriga cerca de 5,5 mil hectares de Cerradão<sup>82</sup> que foram mapeados pelo MapBiomias como Formação Florestal.

---

<sup>80</sup> SFB - Serviço Florestal Brasileiro. SICAR. Disponível em <<https://www.car.gov.br/publico/municipios/downloads>>. Acesso em 06 out. 2022.

<sup>81</sup> Duratex. Plano de Manejo Florestal: Resumo Público São Paulo - Minas Gerais. Disponível em <<https://www.dex.co/Arquivos/Download/Plano-Manejo-Florestal-2017.pdf>>. Acesso em 06 out. 2022.

<sup>82</sup> IF - Instituto Florestal. Estação Ecológica de Jataí: Plano de Manejo. Disponível em <<https://www.infraestruturameioambiente.sp.gov.br/fundacaoflorestal/planos-de-manejo/planos-de-manejo-planos-concluidos/plano-de-manejo-ee-jatai/>>. Acesso em 01 fev. 2023.



Figura 54 - Localização dos fragmentos de formações nativas e inclinação do terreno na APA Corumbataí, Botucatu, Tejujá - Perímetro Corumbataí

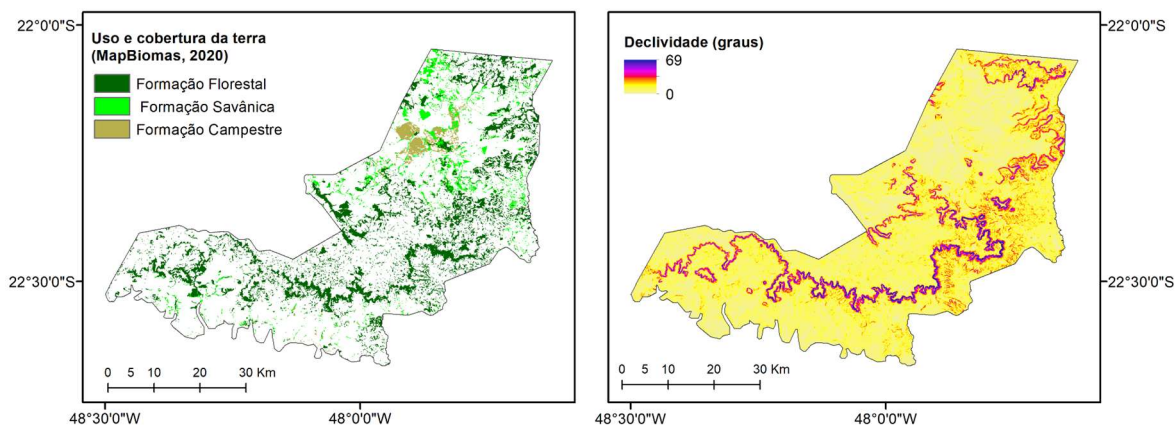
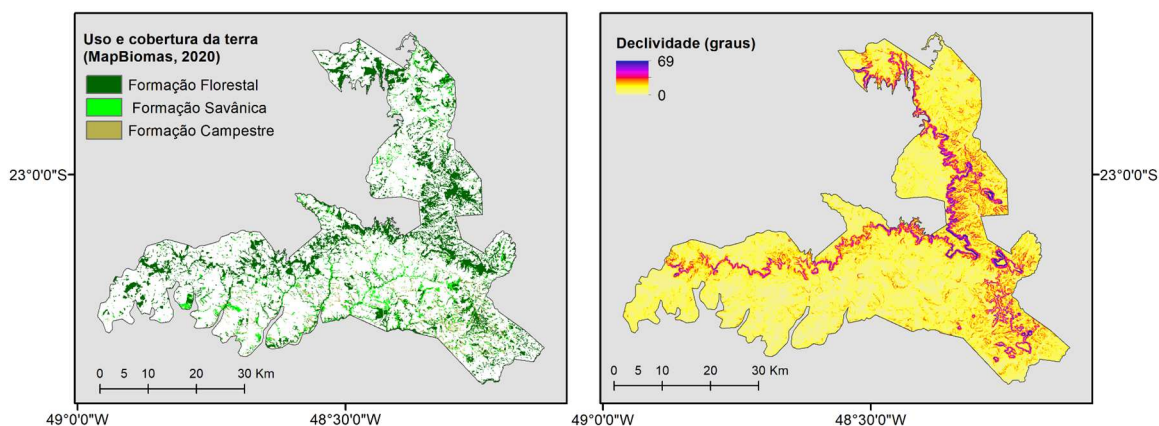


Figura 55 - Localização dos fragmentos de formações nativas e inclinação do terreno na APA Corumbataí, Botucatu, Tejujá - Perímetro Botucatu



#### 4.1.3.2 Agricultura e Pecuária no Estado de São Paulo

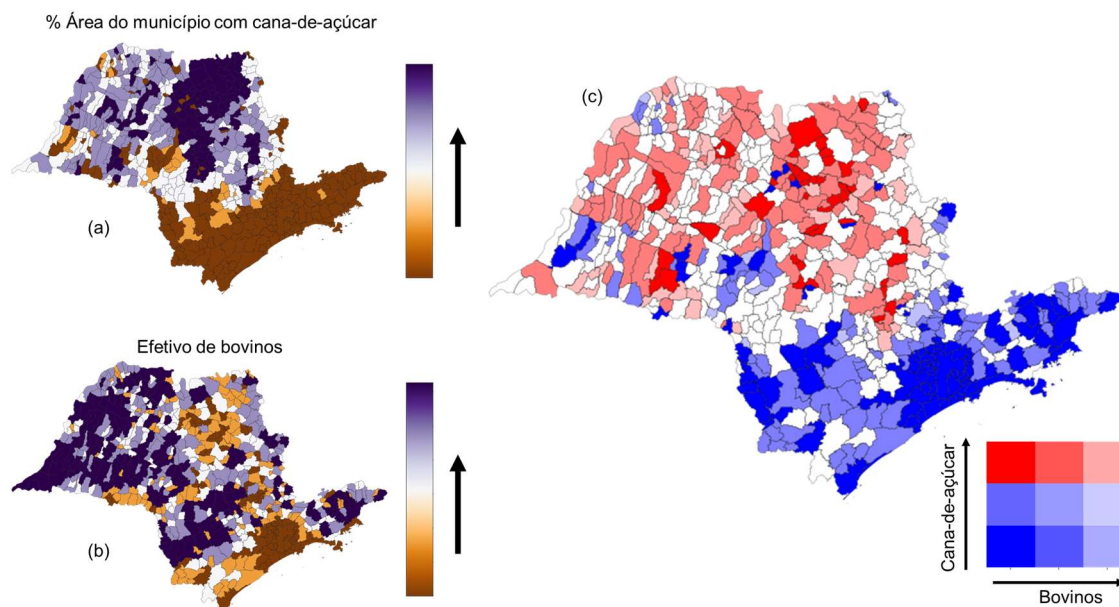
Como referenciado na seção Área de Estudo, a área plantada com cana-de-açúcar e as pastagens ocupam uma área expressiva do estado de São Paulo. Sendo assim, foi feita uma análise exploratória sobre a distribuição espacial da sacaricultura<sup>83</sup> e do efetivo de bovinos<sup>84</sup> no estado. O plantio de cana-de-açúcar é expressivo na porção nordeste do Estado, principalmente na mesorregião de Ribeirão Preto, limitando-se a leste pelas Cuestas Basálticas (Figura 56a), que dificultam a mecanização da agricultura. Nesta região, portanto, a pecuária bovina não tem

<sup>83</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. **Censo agropecuário**. Disponível em <<https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/21814-2017-censo-agropecuario>>. Acesso em 16 abr. 2022.

<sup>84</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. **Pesquisa da pecuária municipal**. Disponível em <<https://sidra.ibge.gov.br/tabela/3939>>. Acesso em 16 abr. 2022

espaço, visto que a região está dominada pela sacaricultura; essa atividade é mais expressiva no extremo noroeste do estado e na macrorregião de Presidente Prudente (Figura 56b).

Figura 56 - Distribuição espacial da (a) cana de açúcar e (b) bovinos no Estado de São Paulo e (c) cartograma bivariado destas variáveis



Fonte: IBGE (2022a<sup>85</sup>; 2022b<sup>86</sup>). Organizado pela autora.

Nas regiões metropolitanas e na Serra do Mar, que são áreas já consolidadas pela urbanização e pela impraticabilidade da agricultura mecanizada, essas atividades são praticamente ausentes. O cartograma bivariado (Figura 56c)<sup>87</sup> mostra a distribuição dessas duas variáveis pelas componentes das cores da legenda; a variação na matiz (de azul a vermelho) quantifica o percentual da área plantada com cana-de-açúcar nos municípios; a variação na saturação indica o número de cabeças de bovinos por quilômetro quadrado nos municípios. Os polígonos em vermelho claro têm valores altos (acima da média) em ambas as variáveis e os polígonos azuis têm valores baixos (abaixo da média). Os polígonos incolores têm valores intermediários em ambas as variáveis. Parte do oeste e nordeste paulista tem alta produção nas duas atividades, pecuária bovina e sacaricultura, com alguns municípios isolados que são

<sup>85</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. Censo agropecuário. Disponível em <<https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/21814-2017-censo-agropecuario>>. Acesso em 16 abr. 2022.a

<sup>86</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. Pesquisa da pecuária municipal. Disponível em <<https://sidra.ibge.gov.br/tabela/3939>>. Acesso em 16 abr. 2022.b

<sup>87</sup> As três representações consideram o método de quantis para classificação dos dados.

exceção. A região da Grande São Paulo, Litoral, Vale do Paraíba e sul do Estado estão associadas à concentração de indústrias automobilísticas, petroquímicas e têxteis.

A seguir é apresentado o resultado da análise de agrupamento espacial das variáveis ‘percentual da área plantada com cana-de-açúcar’ e ‘cabeças de bovinos por quilômetro quadrado’. O diagrama de espalhamento de Moran (ANSELIN, 1996) mostra a dependência espacial de uma variável em um polígono de um mapa coroplético, comparando os valores normalizados com a média dos polígonos vizinhos. Considera-se uma variável  $z$ , dada em desvios em relação à média, que classifica a autocorrelação espacial em quatro categorias. Sendo assim, o diagrama é centrado na média (de zero), e todos os pontos à direita da média têm  $z_i > 0$  e todos os pontos à esquerda têm  $z_i < 0$ , sendo  $z_i$  uma observação no local  $i$ . O diagrama é dividido em quadrantes: Q1, onde estão os valores positivos e médias positivas; Q2, onde estão os valores negativos e médias negativas; Q3, onde estão os valores positivos e médias negativas e Q4, onde estão os valores negativos e médias positivas. Ou seja, Q1 e Q2 indicam que o polígono tem vizinhos com valores similares (autocorrelação espacial positiva), enquanto Q3 e Q4 indicam que o polígono diverge de seus vizinhos (autocorrelação espacial negativa).

As Figuras 57 e 58 mostram o diagrama de espalhamento de Moran para, respectivamente, ‘percentual da área plantada com cana-de-açúcar’ e ‘cabeças de bovinos por quilômetro quadrado’. Os valores são padronizados e apresentados em unidades de desvio padrão (a média é zero e o desvio padrão é um). A forma da nuvem de pontos é influenciada pelos *outliers*; no caso da área plantada com cana-de-açúcar, há valores maiores que três unidades de desvio padrão em relação à média. Por outro lado, no terceiro quadrante, quase não há dispersão nos valores de área plantada. No caso do efetivo de bovinos, há uma maior proporção de outliers no primeiro quadrante, onde estão os valores mais altos, com valores até seis unidades de desvio padrão em relação à média.

Figura 57 - Diagrama de espalhamento para o percentual de área plantada com cana-de-açúcar

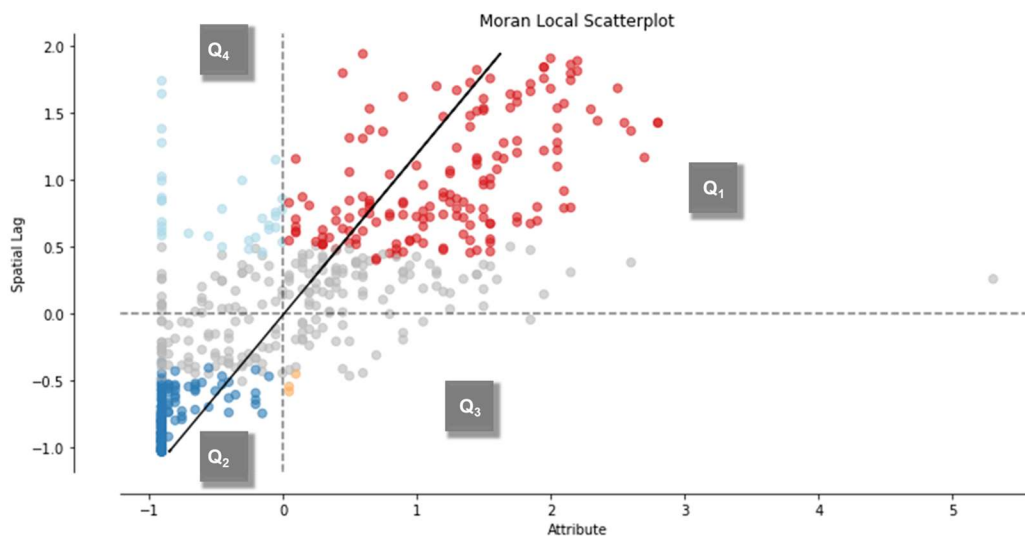
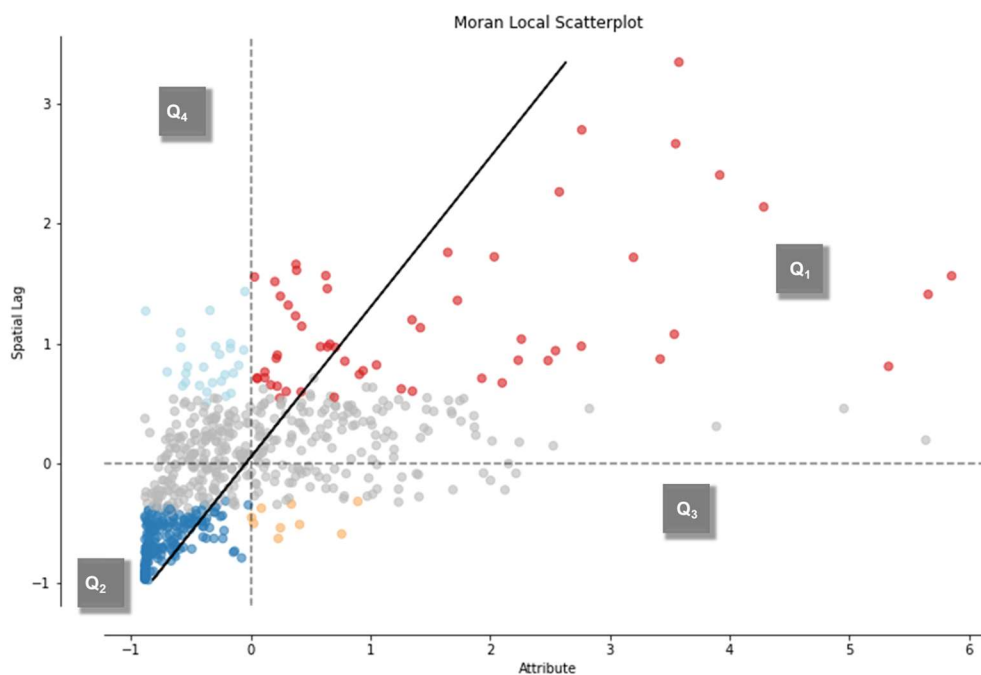


Figura 58 - Diagrama de espalhamento para o efetivo de bovinos



O índice de Moran ( $I$ ) calculado, que indica a inclinação da reta do gráfico de espalhamento (MORAN, 1948; CLIFF; ORD, 1973), foi de 0,41 para as variáveis. Entretanto, os valores de  $I$  são uma estimativa do índice mas não informam sobre a significância do teste, ou seja, se a hipótese nula de aleatoriedade espacial deve ser rejeitada, dando a indicação de agrupamento. A estatística do teste pode ser verificada por permutação aleatória como indicado

por Almeida (2012); todos os valores observados são permutados aleatoriamente para várias regiões e, então, é calculada a estatística para cada uma dessas permutações e gerada uma distribuição de referência. A Tabela 16 mostra os valores calculados e esperados de  $I$  para as variáveis; de acordo com os testes de hipótese, ambas as variáveis apresentam autocorrelação espacial.

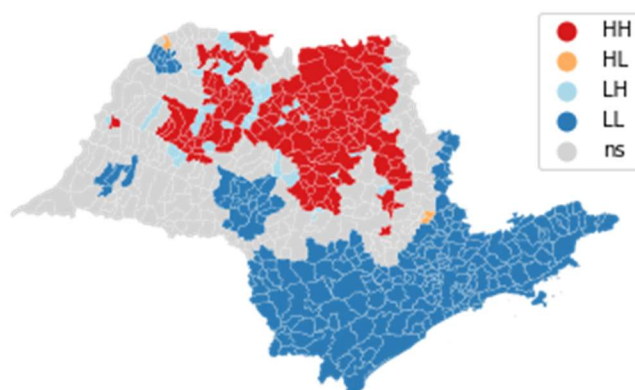
Tabela 16 - Índice de Moran e resultado do teste de hipótese

<b><math>I</math> esperado</b>	<b><math>I</math> calculado</b>	<b>Teste de hipótese</b>
-0,0016	0,41	$I_{\text{esp}} < I_{\text{calc}}$ : Rejeita-se $H_0$ Existe autocorrelação espacial

O índice de Moran caracteriza um padrão espacial geral e não indica a localização dos agrupamentos, ou seja, fornece uma medida única para toda a área de estudo. O Indicador Local de Associação Espacial (*Local Indicator of Spatial Association - LISA*) fornece uma estatística local e estabelece uma relação proporcional entre a soma das estatísticas locais e uma estatística global correspondente (ANSELIN, 2010). No cartograma do indicador LISA, os valores de  $z_i$ , positivos e negativos, são referidos respectivamente como altos e baixos, no sentido de superior ou inferior à média.

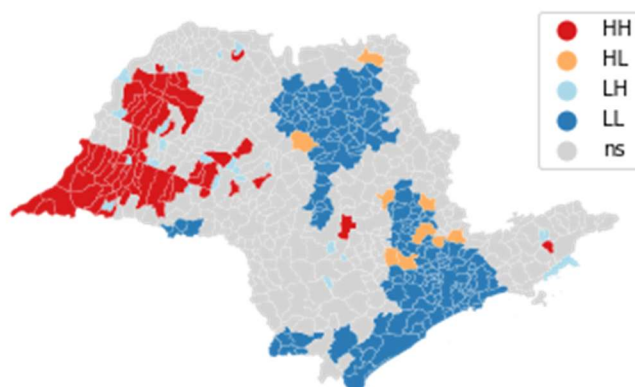
As Figuras 59 e 60 mostram o índice LISA para as variáveis ‘percentual da área plantada com cana-de-açúcar’ e ‘cabeças de bovinos por quilômetro quadrado’. As classes “Alto - Alto” (HH), “Alto - Baixo” (HL), “Baixo - Alto” (LH) e “Baixo-Baixo” referem-se aos quadrantes  $Q_1$ ,  $Q_2$ ,  $Q_3$  e  $Q_4$  do diagrama de espalhamento de Moran. Para a área plantada com cana-de-açúcar, há um aglomerado de altos valores nas mesorregiões de Ribeirão Preto, Araraquara, Araçatuba e São José do Rio Preto. De fato, os municípios com maior extensão cultivada estão nestas mesorregiões; à exceção do município de Paraguaçu Paulista, que está espacialmente isolado (Tabela 17). As mesorregiões de Itapetininga, Macro Metropolitana Paulista, Metropolitana de São Paulo, Vale do Paraíba Paulista e sul da mesorregião de Bauru são aglomerados de municípios com extensão nula ou muito baixa de sacaricultura.

Figura 59 - Índice LISA para área cultivada com cana-de-açúcar



HH = altos valores com vizinhança de altos valores médios; HL = altos valores com vizinhança de baixos valores médios; LH = baixos valores com vizinhança de altos valores médios; LL = baixos valores com vizinhança de baixos valores médios; ns = não significativa.

Figura 60 - Índice LISA para efetivo de bovinos



HH = altos valores com vizinhança de altos valores médios; HL = altos valores com vizinhança de baixos valores médios; LH = baixos valores com vizinhança de altos valores médios; LL = baixos valores com vizinhança de baixos valores médios; ns = não significativa.

Tabela 17 - Municípios com maior área plantada com cana-de-açúcar no estado de São Paulo

Posição	Município	Área plantada (ha)	Mesorregião
1	Morro Agudo	86.110 (62%)	Ribeirão Preto
2	Guaiara	67.529 (54%)	Ribeirão Preto
3	Barretos	66.443 (42%)	Ribeirão Preto
4	Paraguaçu Paulista	48.850 (49%)	Assis
5	Itápolis	47.974 (48%)	Araraquara
6	Valparaíso	45.126 (53%)	Araçatuba
7	Jaboticabal	43.252 (61%)	Ribeirão Preto
8	Novo Horizonte	42.459 (46%)	São José do Rio Preto
9	Guararapes	42.346 (44%)	Araçatuba
10	Ituverava	42.118 (60%)	Ribeirão Preto



Fonte: IBGE, 2022<sup>88</sup>. Os percentuais referem-se à área total do município.

Presidente Prudente, Araçatuba e Marília têm aglomerados de municípios com grande número de cabeças de gado. Já as mesorregiões de Ribeirão Preto, Araraquara, assim como as mesorregiões Macro Metropolitana Paulista, Metropolitana de São Paulo e Litoral Sul Paulista abrigam um número de cabeças de gado bem abaixo da média do Estado (Tabela 18).

<sup>88</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. Censo agropecuário. Disponível em <<https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/21814-2017-censo-agropecuario>>. Acesso em 16 abr. 2022

Tabela 18 - Municípios com maior efetivo de bovinos no estado de São Paulo

<b>Posição</b>	<b>Município</b>	<b>Efetivo (cabeças)</b>	<b>Mesorregião</b>
1	Presidente Bernardes	125.109	Presidente Prudente
2	Mirante do Paranapanema	121.475	Presidente Prudente
3	Cunha	121.100	Vale do Paraíba Paulista
4	Rancharia	115.328	Presidente Prudente
5	Marília	108.403	Marília
6	Presidente Venceslau	95.974	Presidente Prudente
7	Presidente Epitácio	89.147	Presidente Prudente
8	Castilho	88.624	Araçatuba
9	Marabá Paulista	82.817	Presidente Prudente
10	Santo Anastácio	82.305	Presidente Prudente



Fonte: IBGE, 2022<sup>89</sup>

A rápida e intensa expansão da agricultura no cerrado brasileiro foi viabilizada por um conjunto de fatores econômicos, políticos e naturais: a) o valor das terras que era muito baixo

<sup>89</sup> IBGE - Instituto Brasileiro de Geografia e Estatística. Pesquisa da pecuária municipal. Disponível em <<https://sidra.ibge.gov.br/tabela/3939>>. Acesso em 16 abr. 2022



dada a ocupação por pecuária extensiva de baixa rentabilidade; b) a facilidade da conversão da vegetação de baixa biomassa, estrato e densidade de árvores; c) as políticas públicas do governo federal de financiamento agrícola; d) o apoio técnico-científico das pesquisas agronômicas (ROSSI, 2006). Tavares et al. (2019) afirmam que paisagens com matrizes agrícolas resultam na perda de biodiversidade e serviços ecossistêmicos, principalmente quando há intenso uso de agrotóxicos e fertilizantes sintéticos. Os autores afirmam que a agricultura comercial é a principal agente da conversão de terras e perda de biodiversidade, dado o alto grau de mecanização, o plantio de variedades transgênicas e o uso intenso de fertilizantes e agrotóxicos.

No Brasil, a sacaricultura consome 10% dos agrotóxicos e, por isso, há uma tendência de que os produtores rurais sejam contrários a regulações mais restritas quanto aos ingredientes ativos usados nas lavouras (MORAES, 2019). A Figura 61 compara os aumentos no uso de agrotóxicos (toneladas) e no valor bruto da produção agrícola (VBPA) (bilhões de Reais) no estado de São Paulo. O retorno em termos de VBPA não acompanhou o aumento no uso de agrotóxicos, sendo que a relação VBPA/Agrotóxicos (R\$ bilhões / toneladas de agrotóxicos) diminuiu ao longo do tempo. Ou seja, aumentar o consumo de agrotóxicos não aumentou o valor da produção agrícola no estado de São Paulo. Ainda assim, o número de registros anuais de agrotóxicos no Ministério da Agricultura, Pecuária e Abastecimento (MAPA)<sup>90</sup> é crescente e as concentrações de agrotóxicos no Brasil são superiores às previstas na legislação internacional (DIDONÉ et al., 2021; NUNES et al., 2021). Em 2020, segundo o IBAMA<sup>91</sup>, foram protocolados 309 ingredientes ativos, químicos e bioquímicos, sendo os dez mais comercializados: Glifosato; 2,4-D; Mancozebe; Atrazina; Acefato; Clorotalonil; Malationa; Enxofre; Imidacloprido e Clorpirifós. A taxa de registros pode aumentar ainda mais caso seja aprovado o Projeto de Lei 6.299/2002<sup>92</sup> que retornou à tramitação no Senado Federal em 2022 e que flexibiliza a permissão de agrotóxicos no Brasil.

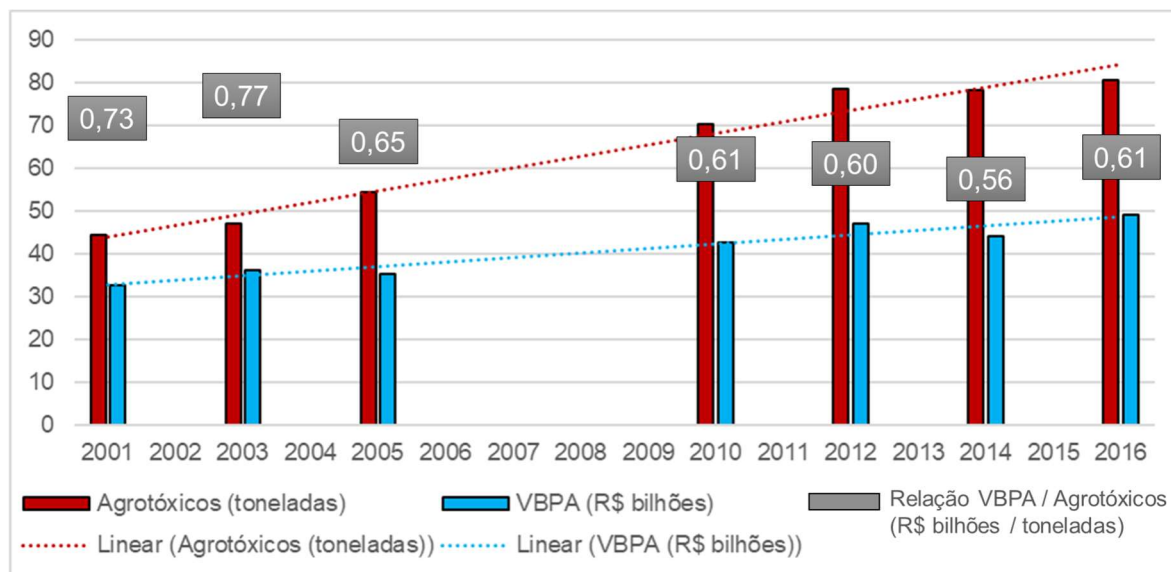
---

<sup>90</sup> AGROFIT. Disponível em <<https://indicadores.agricultura.gov.br/agrofit/index.htm>>. Acesso em 08 out 2022.

<sup>91</sup> IBAMA. Relatórios de comercialização de agrotóxicos. Disponível em <<https://www.ibama.gov.br/agrotoxicos/relatorios-de-comercializacao-de-agrotoxicos#boletinsanuais>>. Acesso em 08 out 2022.

<sup>92</sup> Projeto de Lei 6.299, de 2002. Disponível em <<https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=46249>>. Acesso em 17 nov 2022.

Figura 61 - Uso de agrotóxicos (toneladas) e o valor bruto da produção agrícola (VBPA) (bilhões de Reais) no estado de São Paulo



Fonte: (MORAES, 2019). Organizado pela autora.

Os agrotóxicos induzem a mortalidade de embriões e alterações teratogênicas, além de malformações congênitas, alterações reprodutivas e comportamentais em espécies não-alvo de vertebrados, principalmente de organismos aquáticos, que são os mais sensíveis a esses produtos. As alterações observadas nas diferentes classes de animais são diversas e incluem não apenas malformações morfológicas, mas também alterações reprodutivas e comportamentais (GARCÊS et al., 2019).

Ressalta-se a importância das técnicas de conservação da água e do solo na escala da bacia hidrográfica para reduzir a conectividade dos fluxos de água e sedimentos das áreas agrícolas aos sistemas fluviais e, portanto, a dispersão de agroquímicos (DIDONÉ et al., 2021).

Além do uso intenso de agrotóxicos e contaminação, as lavouras de cana-de-açúcar estão relacionadas a outros vetores de ameaça à biodiversidade, como os riscos de atropelamento em estradas e rodovias, poluição e bioinvasão (BERTASSONI et al., 2019). Existe ainda o risco de queimadas e incêndios decorrentes do uso do fogo na pré-colheita da cana, embora esta prática seja restrita de acordo com o Decreto Estadual 47.700/2003<sup>93</sup>. Hurst et al. (2014) apontam que a monocultura de cana-de-açúcar afeta negativamente as

<sup>93</sup> Decreto Estadual 47.700 de 11 de março de 2003. Regulamenta a Lei nº 11.241, de 19 de setembro de 2002, que dispõe sobre a eliminação gradativa da queima da palha da cana-de-açúcar e dá providências correlatas. Disponível em <[https://www.cetesb.sp.gov.br/licenciamento/documentos/2003\\_Dec\\_Est\\_47700.pdf](https://www.cetesb.sp.gov.br/licenciamento/documentos/2003_Dec_Est_47700.pdf)>. Acesso em 08 out. 2022.

comunidades de pequenos mamíferos em ecossistemas savânicos. Por exemplo, a remoção de granívoros pode alterar a estrutura da vegetação e conduzir à invasão de espécies arbustivas por meio de níveis reduzidos de predação de sementes, ao passo que o aumento de pequenos mamíferos onívoros causa danos às culturas e aumenta a prevalência de doenças transmitidas por vetores.

Na década de 2000, Rossi (2006) apontou que 70% a 80% de todo o cerrado brasileiro estava convertido em pastos plantados com o capim braquiária e campos agrícolas de cultivo mecanizado. A braquiária é originária da África e foi introduzida no Brasil como uma espécie forrageira; é considerada uma espécie invasora que tem alta taxa de proliferação e, por ser altamente competitiva e adaptada ao ambiente em que foi introduzida, é nociva às espécies nativas<sup>94</sup>. O aumento da sua capacidade competitiva é decorrente da produção de substâncias alelopáticas que reduzem a germinação das espécies nativas e potencializam a invasão (BARBOSA et al., 2008). O pastoreio do gado também afeta a composição e a estrutura de uma comunidade vegetal, que estão diretamente relacionadas à qualidade do habitat silvestre, em curto e longo prazo (KRAUSMAN et al., 2009).

No início do século 21, o Brasil tornou-se o sexto país com maior extensão de monoculturas de árvores, que ocupam terras de empresas de papel e celulose, outras propriedades privadas e também territórios de domínio público, sendo, na sua maioria, plantios de espécies exóticas (*Pinus* e *Eucalyptus*) que têm alta produtividade (FONSECA et al., 2009; PAULA et al., 2020). Essas monoculturas de espécies exóticas têm forte potencial invasor em detrimento das espécies de plantas nativas (VILELA et al., 2019; GUZMÁN et al., 2022). Por outro lado, as florestas de espécies exóticas que têm um sub-bosque bem desenvolvido podem sustentar algumas espécies de mamíferos de pequeno a grande porte (STALLINGS, 1990; MARTIN et al., 2012; SIMONETTI et al., 2013; TIMO et al., 2014) e espécies de plantas nativas, estimulando o processo de sucessão secundária (DE PINHO JÚNIOR et al., 2015). De acordo com Timo et al. (2014), as indústrias de celulose e papel são incentivadas a seguir a legislação ambiental, dada a imposição do mercado pelas organizações de certificação como o Forest Stewardship Council (FSC)<sup>95</sup>. As florestas plantadas podem ainda reduzir a pressão sobre as florestas naturais e atender à demanda de madeira e celulose e, quando manejadas de forma correta, podem ser associadas à prática agroflorestal (RAJ et al., 2016) e aprimorar o

---

<sup>94</sup> CBD - Convention on Biological Diversity. **What are Invasive Alien Species?** Disponível em <<https://www.cbd.int/invasive/WhatareIAS.shtml>>. Acesso em 08 mar. 2022.

<sup>95</sup> Forest Stewardship Council. Disponível em <<https://br.fsc.org/br-pt>>. Acesso em 22 out 2022.

funcionamento do ecossistema e o fornecimento de bens e serviços ecossistêmicos (BROCKERHOFF et al., 2013).

A cafeicultura, embora ocupe apenas 0,4% do território paulista, é um exemplo de plantação que é viável em paisagens pouco antropizadas, como os sistemas agroflorestais (SAFs). Trata-se do café sombreado, que consiste no plantio de café em locais com sombra (RICCI et al., 2006; MANCUSO et al., 2013). Esse sistema permite o plantio em regiões com déficit hídrico e sob alta incidência solar, além de proporcionar o cultivo consorciado, isto é, combinado com outras produções, como a de madeira, por exemplo. Os SAFs aumentam a fixação de carbono em relação ao plantio de café a céu aberto (ZARO et al., 2020), melhoram as condições microclimáticas e a drenagem de águas profundas em comparação com os sistemas de café sem sombra (CARVALHO et al., 2021) e são uma alternativa para adaptação do café frente ao aumento da temperatura decorrente das mudanças climáticas (GOMES et al., 2020; ZARO et al., 2023). Caudill et al. (2015) mostraram que a densidade e a riqueza de pequenos mamíferos é maior onde a proporção de café sombreado é maior em relação ao café de sol e que os mamíferos se desenvolvem em áreas adjacentes a fragmentos florestais e à medida que a proporção de áreas florestais na paisagem aumentou. Gomez et al. (2017) concluíram que os agroecossistemas podem ser uma boa estratégia para a conservação de pequenos mamíferos. Ferreira et al. (2020) também mostraram que as agroflorestas podem manter uma alta diversidade de mamíferos de médio e grande porte.

#### 4.1.3.3 Urbanização

A urbanização representa uma das formas mais extremas de alteração do uso da terra e muitas espécies não conseguem lidar com essas mudanças que acontecem em tão pouco tempo, resultando na perda de biodiversidade dada a falta de adaptação para explorar novos recursos e evitar riscos (SANTINI et al., 2019; MAZZA et al., 2020). Todavia, algumas espécies conseguem se desenvolver em habitats urbanos. No caso de mamíferos, são populações com altas densidades e altas taxas de sobrevivência e reprodução, e que conseguem adaptar-se à dieta antropogênica e usar elementos da paisagem urbana - edifícios, bueiros e pontes - como abrigo e local de nidificação. Ou seja, áreas antropizadas com alimentos antropogênicos, alta produção primária líquida e estruturas para abrigo são locais potenciais para servir como hábitat por mamíferos silvestres (McCLEERY, 2010), principalmente para as espécies que apresentam algum grau de plasticidade fenotípica, flexibilidade comportamental, habilidades de dispersão e generalismo de nicho (SANTINI et al., 2019). Por vezes, algumas populações acabam

dependendo dos seres humanos direta ou indiretamente para ter alimentação e abrigo (SARKAR; BHADRA, 2022).

Em áreas urbanas, a relação entre cobertura da terra e ocorrência de mamíferos envolve dois aspectos principais: a) a proximidade de remanescentes de vegetação nativa - dentro ou fora de UCs - e seus riscos e eventuais prejuízos à população humana; e b) riscos das alta antropização para a mastofauna que transita nas proximidades e/ou no interior do perímetro urbano.

A existência de grandes fragmentos nas proximidades de ambientes antropizados pode resultar na predação de animais domésticos rurais (PALMEIRA et al., 2015; BICKLEY et al., 2019; CHINCHILLA et al., 2022), zoonoses (DASZAK et al., 2000; GONÇALVES-OLIVEIRA et al., 2020; EVANS; DRAKE, 2022) e danos nas plantações (FREITAS et al., 2008; HILL, 2017; CARVALHO et al., 2019). No caso das zoonoses, destacam-se as populações de roedores da família Cricetidae, que são abundantes em paisagens agrícolas; Prist et al. (2021) concluíram que a restauração de florestas contribui para diminuir o risco de transmissão da Síndrome Cardiopulmonar por Hantavírus (HCPS - *Hantavirus Cardiopulmonary Syndrome*), que é uma doença com altas taxas de letalidade humana.

Os riscos à mastofauna estão relacionados à caça ilegal (DIAS et al., 2020; PARDO et al., 2022; PRASNIEWSKI et al., 2022), predação por animais domésticos (OLIVEIRA et al., 2008; LACERDA et al., 2009; LESSA et al., 2016; VERÍSSIMO et al., 2022), acidentes em rodovias (CACERES, 2011; FREITAS et al., 2015; FERREGUETTI et al., 2020; ABRA et al., 2021; HILÁRIO et al., 2021; PINTO et al., 2022), choques em redes elétricas (GORDO et al., 2013; ANDRADE et al., 2020; PEREIRA et al., 2020) e transmissão de doenças por vetores (MEDINA et al., 2014; DASZAK et al., 2000; HARVEY; HOLMES, 2022). E, ainda, há os casos de retaliação em decorrência dos danos às populações humanas citadas anteriormente, como o assassinato de felinos pela predação do gado (SILVEIRA et al., 2008) e de primatas que são vistos como risco à saúde pública (BICCA-MARQUES; FREITAS, 2010; HAPPI et al., 2022; TAYLOR, 2022). Estes danos são particularmente nocivos às espécies com algum grau de ameaça de extinção (ARAÚJO et al., 2020). Existe também o risco das comunidades de mamíferos se tornarem domesticadas e dependentes de ambientes antropogênicos (DOTTA; VERDADE, 2011).

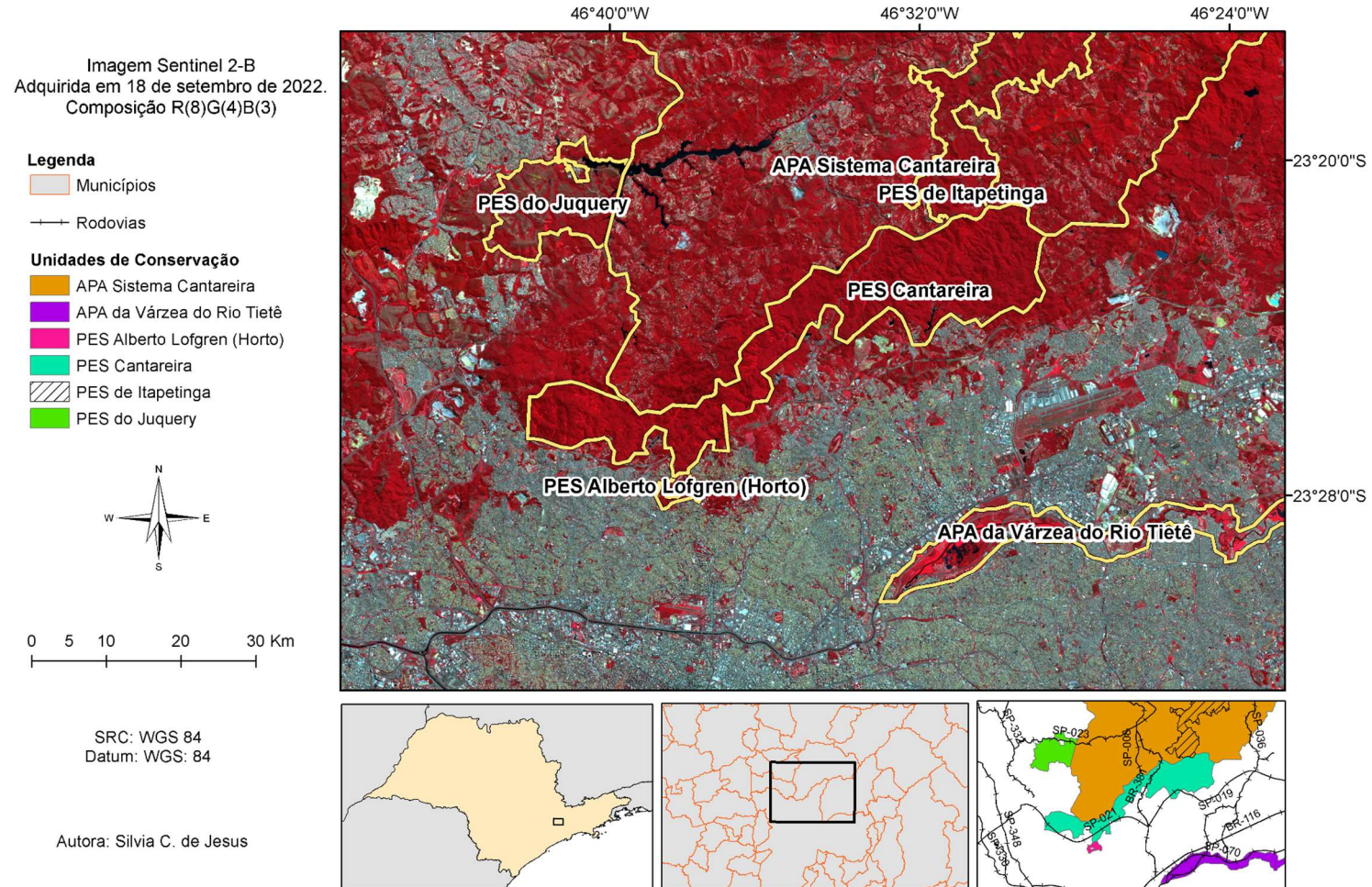
A existência de áreas verdes próximas aos lotes urbanos atrai a aquisição de propriedades nesses locais que, por vezes, são financeiramente mais valorizadas (CHO et al., 2006; HENRIQUE, 2006; FARIAS et al., 2016; ROLI, 2019; TEIXEIRA, 2019; FABIANE et al., 2020). A população tem se deslocado para o interior em busca de melhor qualidade de vida e para viver próxima à natureza, e isso aumenta a demanda por novos lotes nos setores periurbanos dos municípios. Daí a importância da gestão municipal e regional atentarem ao Plano Diretor dos municípios e à Lei Federal 6.766/1979<sup>96</sup>, que versa sobre o parcelamento do solo urbano.

A Área de Proteção Ambiental do Sistema Cantareira (APA Cantareira), por exemplo, é uma região que abriga mananciais de importância regional e que está sujeita à especulação imobiliária, centros residenciais e recreativos insustentáveis ao redor do reservatório e em outras áreas protegidas adjacentes, apesar das restrições legais sobre o uso do solo (HOEFEL et al., 2021). A Figura 62 mostra o espraiamento urbano no município de Guarulhos e zona norte de São Paulo em direção às APAs Sistema Cantareira e da Várzea do Rio Tietê e aos Parques Estaduais Alberto Löfgren, Cantareira, Itapetinga e do Juquery. Nessas áreas periféricas, o crescimento populacional é maior do que na própria cidade e é onde estão concentrados os conflitos ambientais (MELLO-THÉRY; CORREIA, 2009; MELLO-THÉRY, 2011). Couto (2011) reporta que a Serra da Cantareira tem sido desmatada para loteamentos legais e clandestinos, o que resultou em alterações no clima local, alterações no regime fluvial, contaminação da água, descarte ilegal de resíduos sólidos e erosão. No município de Mairiporã, a área urbana está localizada às margens da Represa do Rio Juqueri, que abastece boa parte da região da Grande São Paulo.

---

<sup>96</sup> Lei 6.766, de 19 de dezembro de 1979. Dispõe sobre o Parcelamento do Solo Urbano e dá outras Providências. Disponível em <[http://www.planalto.gov.br/ccivil\\_03/leis/l6766.htm](http://www.planalto.gov.br/ccivil_03/leis/l6766.htm)>. Acesso em 23 out. 2022.

Figura 62 - Transição entre a Serra da Cantareira e a área urbana da Região Metropolitana de São Paulo



A partir da Serra da Cantareira e do entorno da Região Metropolitana de São Paulo (RMSP) até a Serra do Mar e a Baixada Santista se estende a Reserva da Biosfera do Cinturão Verde de São Paulo (RBCV), onde estão os grandes remanescentes da Mata Atlântica - e de vegetação nativa, de um modo geral. A RBCV foi integrada à Reserva da Biosfera da Mata Atlântica em 1994 pela Organização das Nações Unidas (UNESCO). Trata-se de uma área de aproximadamente 21.000 km<sup>2</sup>, que compreende os mananciais que abastecem a área metropolitana e as cabeceiras e afluentes dos rios que permeiam o espaço urbano, e também abriga UCs de Proteção Integral e de Uso Sustentável. Entretanto, está sujeita a perturbações associadas à alta urbanização, como especulação imobiliária, grandes obras de infraestrutura, mineração, extração ilegal de recursos florestais, deposição de resíduos sólidos urbanos e poluição atmosférica<sup>97</sup>. De acordo com Ramos-Ribeiro (2014), 22% da área desmatada na RBCV ocorreu nos municípios da Serra da Cantareira.

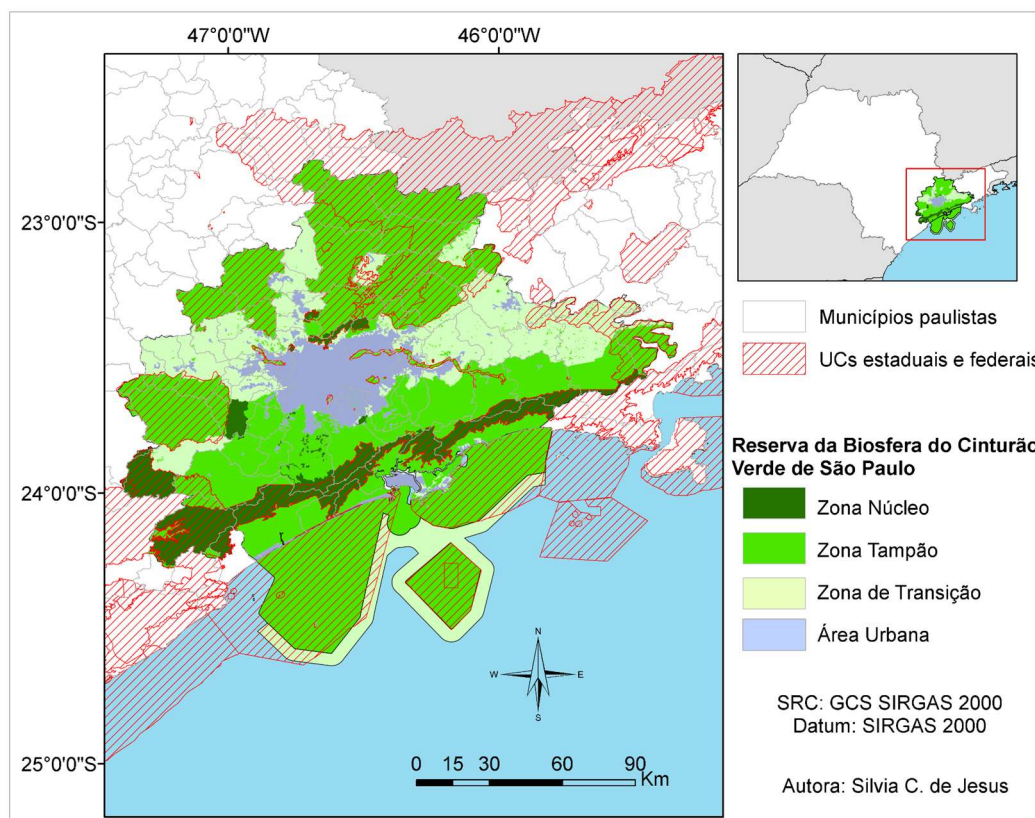
O zoneamento da RBCV segue os princípios do Programa – MaB (*Man and Biosphere* – O Homem e a Biosfera) da UNESCO e consiste em três categorias de zoneamento para o planejamento da ocupação e uso do solo e de seus recursos ambientais: Zona Núcleo (áreas, em sua maioria, compostas por UCs estaduais, englobando principalmente remanescentes da Mata Atlântica e algumas áreas de Cerrado), Zona Tampão (áreas subjacentes às Zonas Núcleo, onde as atividades desenvolvidas, sejam econômicas ou de qualquer outra natureza, devem se adequar às características de cada Zona Núcleo) e Zona de Transição (constituídas pelas áreas externas às Zonas Tampão, onde é permitido o uso mais intensivo, porém não destrutivo, do solo e seus recursos ambientais) (IF, 2022). A Figura 63 mostra o zoneamento da RBCV.

---

<sup>97</sup> IF - Instituto Florestal. Reserva da Biosfera do Cinturão Verde da Cidade de São Paulo. Disponível em: <https://www.infraestruturameioambiente.sp.gov.br/institutoflorestal/o-instituto/rbcv/>. Acesso em 24 out. 2022



Figura 63 - Localização da Reserva da Biosfera do Cinturão Verde de São Paulo e seu zoneamento



Na Região Metropolitana de São Paulo, a melhoria do sistema viário impulsionou a construção de subúrbios e vilas, o que ocasionou o desaparecimento de áreas rurais. Nas bacias hidrográficas Guarapiranga e Billings, os loteamentos destinados à população de baixa renda infringiram as leis de proteção de mananciais, sem considerar a inclinação do terreno, a vegetação existente e a proximidade aos corpos d'água (YOUNG, 2013). Os distritos de Grajaú (360.287 habitantes) e Jardim Ângela (295.434 habitantes), na zona Sul de São Paulo, abrangem parte das represas Billings e Guarapiranga, respectivamente, e são os mais populosos do município<sup>98</sup> (Figura 64). O extremo sul do município está inserido na bacia hidrográfica do Capivari-Monos, que abrange parte da Serra do Mar e onde estão grandes remanescentes da Mata Atlântica. Esta bacia coincide com a APA Capivari-Monos<sup>99</sup>, criada em 2001, que é uma UC municipal e, portanto, não faz parte do escopo desta pesquisa. A APA Capivari-Monos não

<sup>98</sup> São Paulo - Subprefeituras. Dados demográficos dos distritos pertencentes às Subprefeituras. Disponível em <[https://www.prefeitura.sp.gov.br/cidade/secretarias/subprefeituras/subprefeituras/dados\\_demograficos/index.php?p=12758](https://www.prefeitura.sp.gov.br/cidade/secretarias/subprefeituras/subprefeituras/dados_demograficos/index.php?p=12758)>. Acesso em 25 out 2022.

<sup>99</sup> São Paulo. APA Capivari-Monos. Disponível em <[https://www.prefeitura.sp.gov.br/cidade/secretarias/meio\\_ambiente/unid\\_de\\_conservacao/apa\\_capivarimonos/index.php?p=41966](https://www.prefeitura.sp.gov.br/cidade/secretarias/meio_ambiente/unid_de_conservacao/apa_capivarimonos/index.php?p=41966)>. Acesso em 25 out 2022.

se sobrepõe a outras UCs e é de extrema importância para regular o uso do solo em uma área de extrema pressão de urbanização, principalmente por conta dos mananciais na região.

A Serra do Mar mantém os remanescentes mais extensos de Mata Atlântica, apesar do adensamento rodoviário (Figura 65). A presença de rodovias e estradas secundárias está diretamente relacionada à dinâmica de uso da terra e cobertura florestal, visto que a expansão de áreas urbanas e de pastagens são os principais usos após a supressão florestal; além disso, a acessibilidade a fragmentos de vegetação nativa facilita o estabelecimento de outras atividades que debilitam os ecossistemas florestais, como agricultura, extração de madeira e caça (FREITAS et al., 2011). As escarpas da Serra do Mar são intensamente intemperizadas dado o clima quente e úmido local, o que as deixa sujeitas a deslizamentos de terra (CERRI et al., 2018), principalmente onde não há cobertura vegetal nativa e a suscetibilidade aos movimentos de massa é maior (RAMOS et al., 2022).

A zona costeira paulista integra-se social e economicamente à RMSP e desempenha funções complementares às da metrópole, expressas nas atividades portuárias, industriais e turísticas (AFONSO, 2006). A faixa litorânea Santos – São Vicente, Guarujá e Praia Grande é intensamente urbanizada como efeito do turismo e lazer, onde algumas manchas de manguezais ainda persistem (VICTOR et al., 2004), apesar da supressão da vegetação nesse ecossistema na década de 1990 (MAPBIOMAS, 2022). A Figura 66 mostra as imagens orbitais do município de São Vicente dos anos de 1996 e 1998, período em que houve supressão de quase 270 hectares de mangue (MAPBIOMAS, 2022); a figura mostra também as alterações de 1998 a 2022. Em Cubatão, as encostas de morros, manguezais e planícies inundáveis foram ocupadas por moradias para a população de baixa renda, após a implantação do complexo industrial e valorização dos terrenos no centro da cidade; em contrapartida, 60% do município está em áreas protegidas por UCs (MELLO et al., 2011). No litoral norte, região importante para o turismo e setor energético, o desenvolvimento econômico engatou na década de 1970 com o asfaltamento da Rodovia Rio-Santos e melhorias na Rodovia dos Tamoios, que liga a região ao Vale do Paraíba (TORRES et al. 2019; DAUNT et al., 2021). Em 2021, o Porto de São Sebastião movimentou quase 2 bilhões de dólares em exportações<sup>100</sup>.

---

<sup>100</sup> MDIC - Ministério da Indústria, Comércio Exterior e Serviços. Plataforma Comex Stat. Disponível em <<http://comexstat.mdic.gov.br/>>. Acesso em 25 out. 2022.

Figura 64 - Represas Guarapiranga (1) e Billings (2) na Região Metropolitana do município de São Paulo

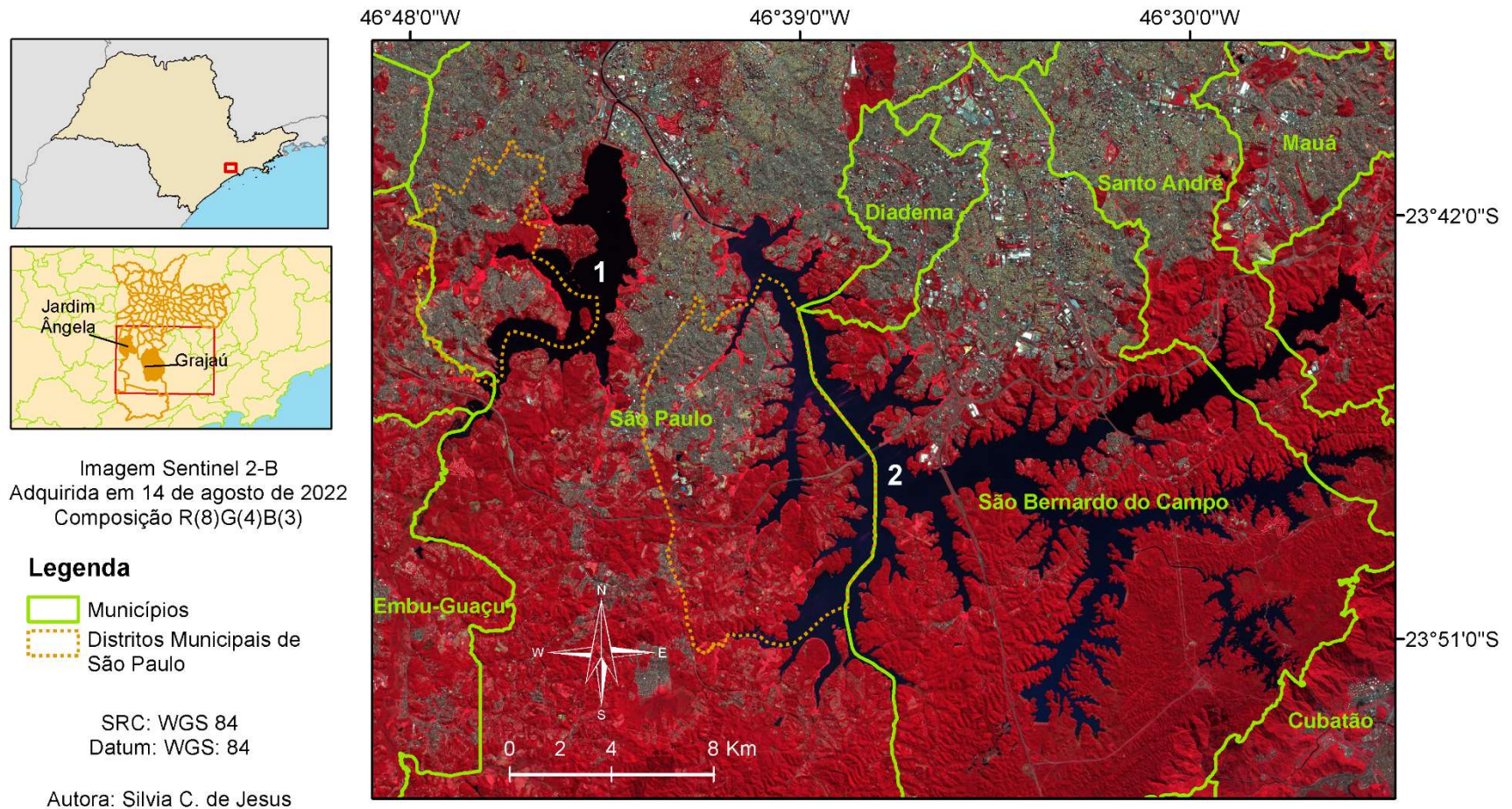


Figura 65 - Adensamento rodoviário na Serra do Mar

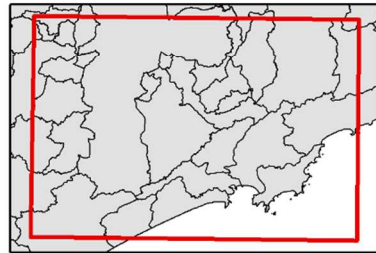
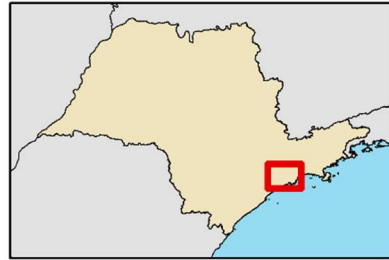




Imagem Sentinel 2B  
Adquirida em 14 de agosto de 2022.  
Composição R(8)G(4)B(3)

**Legenda**

-  Municípios
-  Trecho rodoviário

SRC: WGS 84  
Datum: WGS: 84

Autora: Silvia C. de Jesus

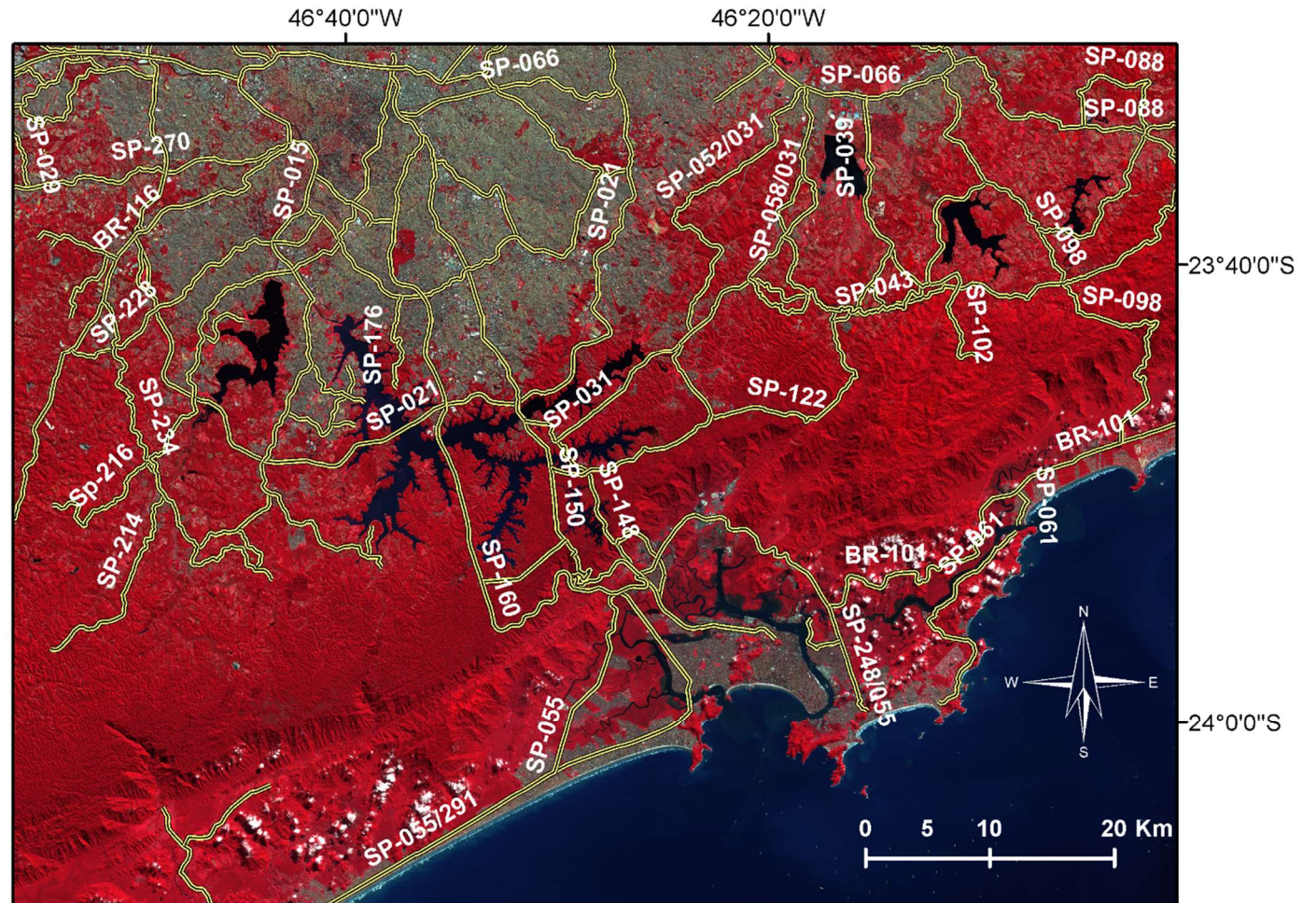
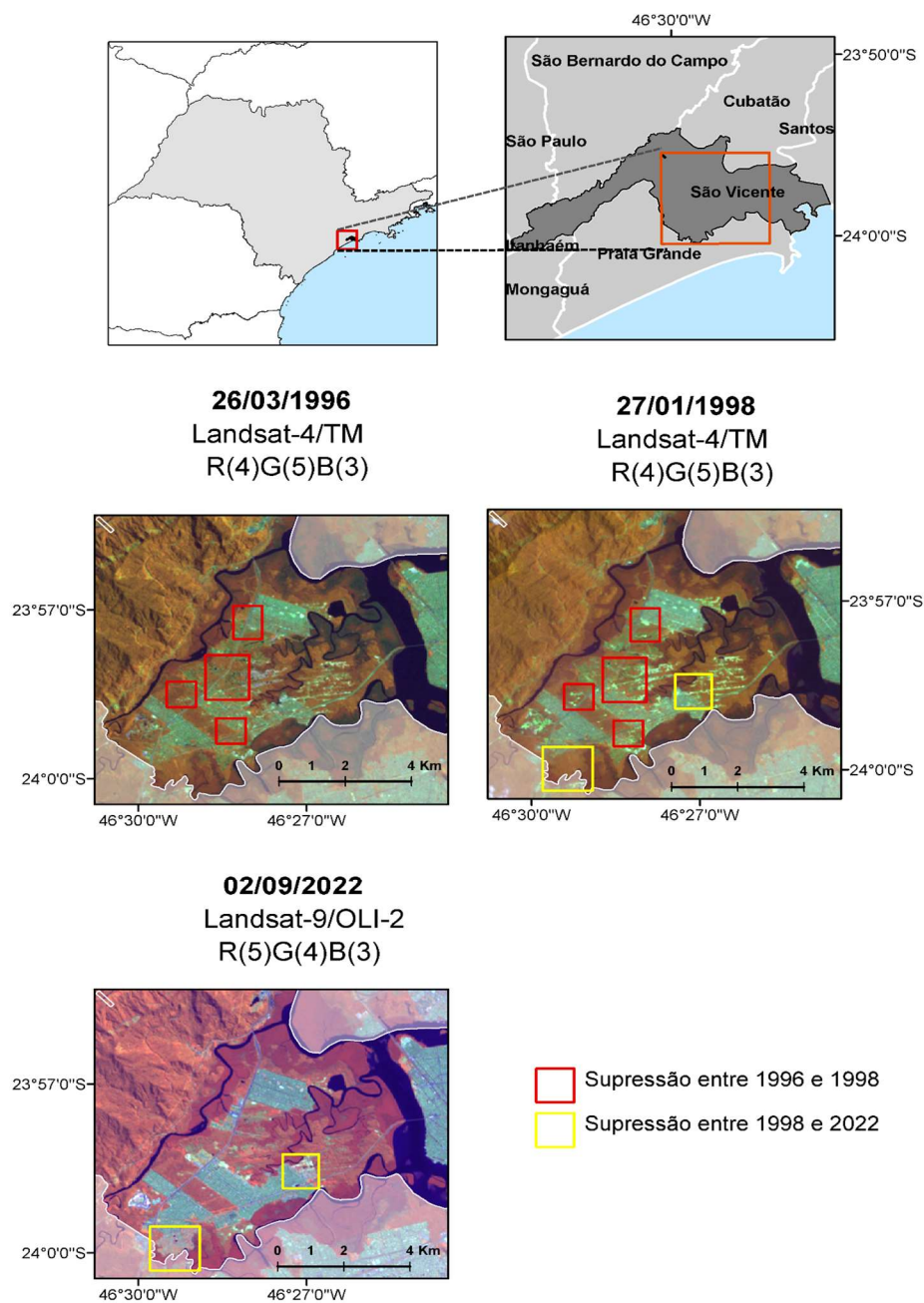


Figura 66 - Supressão de áreas de mangue no Município de São Vicente nos anos de 1996, 1998 e 2022



## 4.2 OCORRÊNCIA DE MAMÍFEROS NO ESTADO DE SÃO PAULO

### 4.2.1 Filtragem por metadados

Zizka et al. (2020) ressaltam a importância de explorar os metadados para identificar registros problemáticos. Foram verificadas as ocorrências para os parâmetros que poderiam comprometer a confiabilidade dos dados. A busca de registros de ocorrências de espécies via

API retornou um dado tabular com 165 variáveis e 8.590 registros<sup>101</sup>, todos com a informação de coordenadas geográficas. Destes, 3.303 (38%) têm valores nulos na coluna ‘species’, que foram removidos.

A Tabela 19 mostra o número de ocorrências em cada valor do atributo ‘basisOfRecord’, que se refere à natureza específica do registro. A maior parte dos registros (57,3%) é composta por espécimes preservados em museus. Embora estas ocorrências não tenham sido consideradas como amostras neste trabalho, estas coleções biológicas constituem repositórios de biodiversidade mantidos em museus de história natural, herbários e outros centros de ciências e são fundamentais para a conservação, pesquisa científica e desenvolvimento tecnológico (SILVA et al., 2020). A conservação *ex situ* de componentes da diversidade biológica é prevista na Convenção da Diversidade Biológica<sup>102</sup>, que estabelece que cada país deve adotar medidas e manter instalações para essa estratégia de conservação de componentes da diversidade biológica, preferencialmente no país de origem de tais componentes. Schmitt et al. (2018) destacam a importância educacional dos espécimes preservados e do envolvimento do público na documentação da biodiversidade por meio da ciência cidadã. Em contrapartida, os registros de ‘MACHINE\_OBSERVATION’ têm baixa frequência (0,02%) na amostra de mamíferos no estado de São Paulo, apesar dos avanços nas tecnologias de sensoriamento *in situ* da biodiversidade (TURNER, 2014; LAHOZ-MONFORT; MAGRATH, 2021).

Telemetria (GROTTA-NETO; DUARTE, 2019; DESBIEZ, 2021; TWISS et al., 2021; BERTASSONI et al., 2022), armadilhas fotográficas (DE BONDI et al., 2010; LEEMPOEL et al.; 2020; AMIOT et al. 2021; ORTMANN; JOHNSON, 2021; VILAS BOAS et al., 2022), dispositivos de gravação de som (SILVA et al., 2013; BURIVALOVA, 2019; DARRAS et al., 2019; LAVOR et al., 2022; STOWELL, 2022), drones (WICH; KOH, 2018; CORCORAN et al., 2021; KAYS et al., 2022; SPAAN et al., 2022) e coleções de DNA ambiental (eDNA) (LEEMPOEL et al.; 2020; CARVALHO et al., 2021b; LYET et al., 2021; OGDEN, 2022; SALES et al., 2022) são ferramentas bem descritas na literatura sobre levantamento de dados de diversidade biológica. A baixa proporção de registros de ‘MATERIAL\_CITATION’ (0,02%) também soa como impropriedade. As monografias taxonômicas, por exemplo, são fontes de dados que compreendem esquemas de citação com ligações externas (FAWCETT et

---

<sup>101</sup> Consulta em 05 de mai. de 2022.

<sup>102</sup> ONU - Organização das Nações Unidas. Convention on Biological Diversity: Text of the Convention. Disponível em <<https://www.cbd.int/convention/text/>>. Acesso em 28 set. 2022.

al., 2022) e contam com sistemas automatizados que facilitam a produção de tratamentos taxonômicos (REGINATO, 2016). Kirschner e Kaplan (2002) destacam a importância das monografias taxonômicas para definir estratégias de conservação da biodiversidade.

Foram suprimidos os registros que constam como 'FOSSIL\_SPECIMEN', 'MATERIAL\_SAMPLE', 'PRESERVED\_SPECIMEN', 'LIVING\_SPECIMEN', o que resultou em 2.019 registros remanescentes.

Tabela 19 - Ocorrências por valor do atributo 'basisOfRecord'

Atributo	Descrição <sup>103 104</sup>	Nº de ocorrências	Percentual em relação ao total
FOSSIL_SPECIMEN	Um espécime preservado que é um fóssil, por exemplo, um corpo fóssil, um coprólito, um gastrólito, um icnofóssil ou um pedaço de árvore petrificada	27	0,5%
PRESERVED_SPECIMEN	Um espécime que foi preservado, por exemplo, uma planta em uma folha de herbário ou um lote catalogado de animais em recipientes	3.049	57,3%
MATERIAL_SAMPLE	Um registro de ocorrência baseado em amostras retiradas de outros espécimes ou do ambiente	133	2,5%
LIVING_SPECIMEN	Um espécime vivo, por exemplo, uma planta viva em um jardim botânico ou um animal vivo em um zoológico	59	1,1%
HUMAN_OBSERVATION	Uma saída do processo de observação humana, por exemplo, evidência de uma ocorrência retirada de notas de campo ou literatura ou registros de uma ocorrência sem evidência física nem evidência capturada com uma máquina	1.985	37,9%

<sup>103</sup> Biodiversity Data Use: Basis of Record. Disponível em <<https://docs.gbif.org/course-data-use/en/basis-of-record.html>>. Acesso em 28 set. 2022.

<sup>104</sup> GBIF API Vocabulary: Basis of Record. Disponível em <<https://gbif.github.io/parsers/apidocs/org/gbif/api/vocabulary/BasisOfRecord.html>>. Acesso em 28 set. 2022.

Atributo	Descrição <sup>103 104</sup>	Nº de ocorrências	Percentual em relação ao total
MACHINE_OBSERVATION	Uma saída de um processo de observação de máquina, por exemplo, uma fotografia, um vídeo, uma gravação de áudio, uma imagem de sensoriamento remoto ou um registro de ocorrência baseado em telemetria	1	0,0%
MATERIAL_CITATION	Uma referência ou citação de um, parte ou vários espécimes em publicações acadêmicas, por exemplo, uma citação de um espécime físico de uma coleção científica em tratamento taxonômico em uma publicação científica ou uma ocorrência mencionada em um livro de notas de campo	1	0,1%
OCCURRENCE	Uma existência de um Organismo <sup>105</sup> em um determinado lugar em um determinado momento	32	0,6%
TOTAL		5.287	

\*Os registros nas células amarelas foram removidos deste estudo.

As ordens 'Chiroptera', 'Cetacea' e 'Didelphimorphia' foram removidas; isto porque são consideradas neste trabalho apenas as espécies predominantemente terrestres e porque as ordens 'Chiroptera' e 'Didelphimorphia' são consideradas generalistas no uso de habitat e são comuns em áreas urbanizadas (KAJIN et al., 2008; BRITO; BOCCHIGLIERI, 2012; SILVA et al., 2014; NUNES et al., 2017). Essas ordens, portanto, não teriam o potencial de definir a adequabilidade de habitat para outros mamíferos. Com esta exclusão, permaneceram 1.749 registros (Tabela 20). Então, foram removidas as espécies domésticas e/ou introduzidas: *Arctocephalus tropicalis*, *Bos taurus*, *Bubalus bubalis*, *Canis lupus*, *Capra hircus*, *Equus caballus*, *Felis catus*, *Hydrochoerus hydrochaeris*, *Lepus europaeus*, *Mirounga leonina*, *Mus musculus*, *Sus scrofa*, restando 1.476 ocorrências.

<sup>105</sup> Term Name dwc:Organism. Disponível em <[https://dwc.tdwg.org/list/#dwc\\_Organism](https://dwc.tdwg.org/list/#dwc_Organism)>. Acesso em 28 set. 2022.



Tabela 20 - Ocorrências por valor do atributo 'order'

<b>Ordem</b>	<b>Nº de ocorrências</b>	<b>Percentual em relação ao total</b>
Artiodactyla	42	2,1%
Carnivora	306	14,9%
Cetacea	16	0,8%
Chiroptera	112	5,5%
Cingulata	51	2,5%
Didelphimorphia	142	6,9%
Lagomorpha	16	0,8%
Perissodactyla	28	1,4%
Pilosa	69	3,4%
Primates	677	33,1%
Rodentia	560	27,3%
<b>TOTAL</b>	<b>2.019</b>	

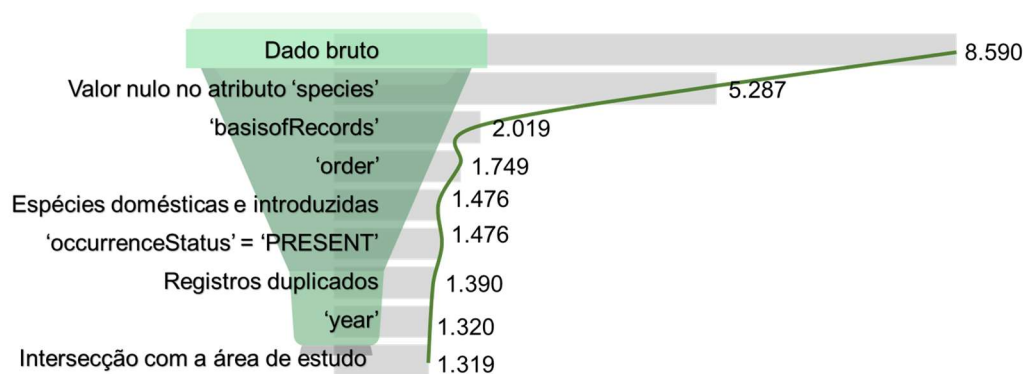
\*Os registros nas células amarelas foram removidos deste estudo.

A Tabela 21 mostra o resultado dos outros filtros aplicados nos dados de ocorrência de espécies, com a permanência de 1.320 ocorrências (15,4% do dado bruto). A variação no tamanho da amostra com a aplicação dos filtros é apresentada na Figura 67.

Tabela 21 - Resultado da filtragem por metadados

Filtro (registros removidos)	Resultado da filtragem (nº de ocorrências)
'occurrenceStatus' ≠ 'PRESENT'	1.476 (nenhuma ocorrência removida)
registros duplicados com os mesmos valores nos atributos 'decimalLatitude',' decimalLongitude' e 'species'	1.390
year ≤ 2010	1.320
intersecção com a área de interesse	1.319

Figura 67 - Variação do número de ocorrências de espécies com a aplicação de filtros de metadados



#### 4.2.2 Filtragem espacial

O algoritmo de clusterização DBSCAN e a definição dos medóides dos clusters foram executados para filtrar espacialmente aglomerados de ocorrências. O DBSCAN requer o parâmetro eps, que especifica a proximidade entre os pontos para que sejam considerados parte de um cluster. Ou seja, se a distância entre dois pontos é menor ou igual a esse valor (eps), esses pontos são considerados vizinhos. Foram gerados clusters para cada espécie considerando as distâncias de 0,03 km, 0,09 km, 0,10 km, 0,5 km, 1 km e 2 km. A Tabela 22 mostra o valor bruto (antes da clusterização) e o número de medóides de clusters gerados para cada espécie para cada valor de eps. Também foi calculado o desvio padrão do número de clusters gerados em cada eps para discriminar para quais espécies houve variação com o aumento desta variável no processamento do DBSCAN. Esta tabela está ordenada pelo desvio padrão. As espécies em que o número de clusters não varia em função do eps (53,2% das espécies) são aquelas em que

os espécimes estão em uma distância maior do que os 2 km, que é o maior valor de eps considerado.

Tabela 22 - Variação por espécie na quantidade de medóides em função do eps

<b>Espécie</b>	<b>Bruto</b>	<b>eps=0,03</b>	<b>eps=0,09</b>	<b>eps=0,25</b>	<b>eps=0,50</b>	<b>eps=1,00</b>	<b>eps=2,00</b>	<b>desvio padrão</b>
<i>Alouatta guariba</i>	129	83	46	36	32	27	25	38,5
<i>Callithrix penicillata</i>	177	137	116	103	90	80	66	37,7
<i>Callithrix jacchus</i>	99	68	61	55	49	43	36	20,8
<i>Sciurus aestuans</i>	129	110	99	91	80	75	73	20,5
<i>Myocastor coypus</i>	72	33	29	27	25	25	24	17,2
<i>Nasua nasua</i>	73	69	58	44	37	35	32	16,9
<i>Sapajus nigritus</i>	145	145	145	145	145	136	112	12,4
<i>Bradypus variegatus</i>	42	29	20	18	15	14	14	10,4
<i>Cavia aperea</i>	33	29	22	18	17	17	17	6,6
<i>Cerdocyon thous</i>	54	49	46	43	41	40	37	5,8
<i>Dasypus novemcinctus</i>	43	43	35	34	33	32	31	5,0
<i>Eira barbara</i>	20	18	15	15	15	15	15	2,0





Espécie	Bruto	eps=0,03	eps=0,09	eps=0,25	eps=0,50	eps=1,00	eps=2,00	desvio padrão
<i>Puma yagouaroundi</i>	6	6	6	6	6	6	6	0,0
<i>Sapajus libidinosus</i>	1	1	1	1	1	1	1	0,0
<i>Sylvilagus brasiliensis</i>	7	7	7	7	7	7	7	0,0
<i>Tamandua tetradactyla</i>	6	6	6	6	6	6	6	0,0
<i>Tapirus terrestris</i>	24	24	24	24	24	24	24	0,0
<i>Tayassu pecari</i>	2	2	2	2	2	2	2	0,0
<i>Thaptomys nigrita</i>	1	1	1	1	1	1	1	0,0

Os agrupamentos por espécie, por meio da variação no número de agrupamentos em função do eps, indica o quanto as observações são espacialmente próximas. Os primatas e os roedores apresentaram as maiores variações no número de pontos representativos dos clusters; ou seja, originalmente era um número alto de ocorrências aglomeradas. Este resultado pode estar associado à vida em grupo e formação de bandos. De acordo com Krebs e Davies (1996), este comportamento tem como principais vantagens o reforço na proteção contra predadores e a busca de alimento; por outro lado, a vida em grupo implica também competição por recursos e o aumento na evidência em relação a predadores. Sendo assim, o tamanho ótimo do grupo promove o equilíbrio entre os benefícios e danos da formação de bandos.

Entre os primatas, os comportamentos cooperativos são mais comuns do que comportamentos agonísticos em todas as espécies de primatas; defesa conjunta de recursos, defesa de alcance, caça cooperativa, formação de alianças, busca e colheita cooperativa de alimentos, limpeza mútua, aglomeração, proximidade espacial e vigilância de predadores são

os benefícios do comportamento afiliativo por primatas (SUSSMAN et al., 2005). As espécies do gênero *Alouatta* são sociais e formam grupos com tamanhos que podem variar conforme a espécie (MIRANDA; PASSOS, 2005) e o local (GONZÁLEZ-SOLÍS et al., 2001). Espécies do gênero *Callithrix* apresentam cuidado parental e compartilhamento de alimentos, além de evidências de aprendizado social e cuidado parental (CALDWELL; WHITEN, 2003).

Os roedores Caviomorpha (MAHER; BURGER, 2011) e Sciurideos (CHIARELLO, 2008) também vivem em grupos. A sociabilidade neste grupo pode estar associada aos riscos de predação, qualidade de tocas e forrageio (EBENSPERGER, 1998; SOLOMON; KEANE, 2016).

Os carnívoros, no entanto, geralmente não vivem em grupos, sendo que apenas parte das espécies se agregam fora do período de reprodução para compartilhar uma área de vida comum, um mesmo recurso alimentar, caça e/ou reprodução (GITTLEMAN, 1989).

Para as espécies em que o número de clusters não se manteve constante em função do eps, a variação é apresentada nas Figuras 68 a 70 . Nos gráficos é possível verificar em qual valor de eps está o ponto de inflexão da curva, ou seja, a partir de qual valor de eps a redução no número de clusters estaciona. Considerando todas as espécies, o ponto de inflexão da curva do eps corresponde a 0,25 (Figura 71). Isto significa que, para cada espécie, as ocorrências estão distantes entre si por pelo menos 250 metros, embora possam ocorrer outras espécies dentro desta distância.

Figura 68 - Variação do número de clusters em função da variável eps do algoritmo DBSCAN (I)

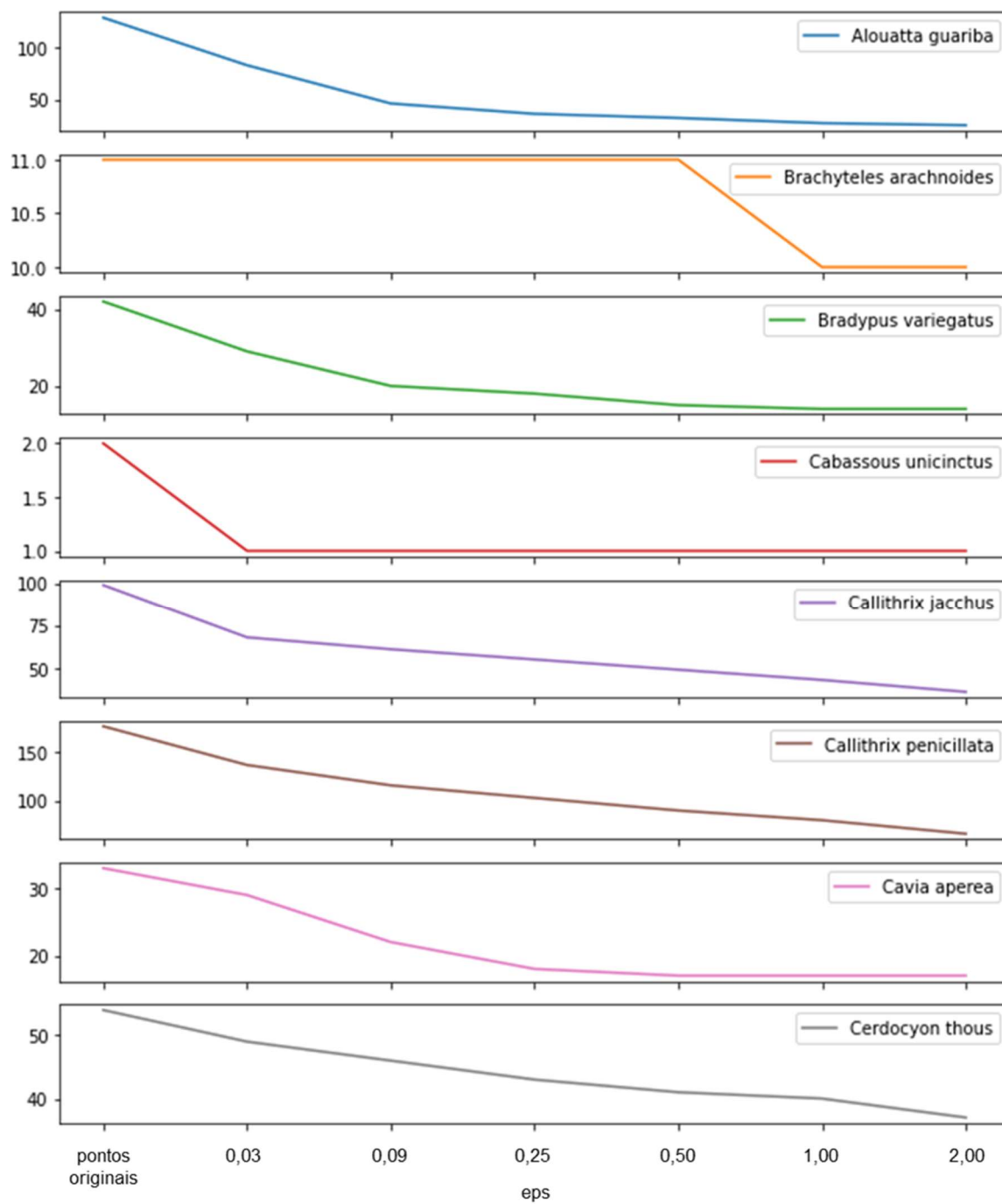




Figura 69 - Variação do número de clusters em função da variável eps do algoritmo DBSCAN (II)

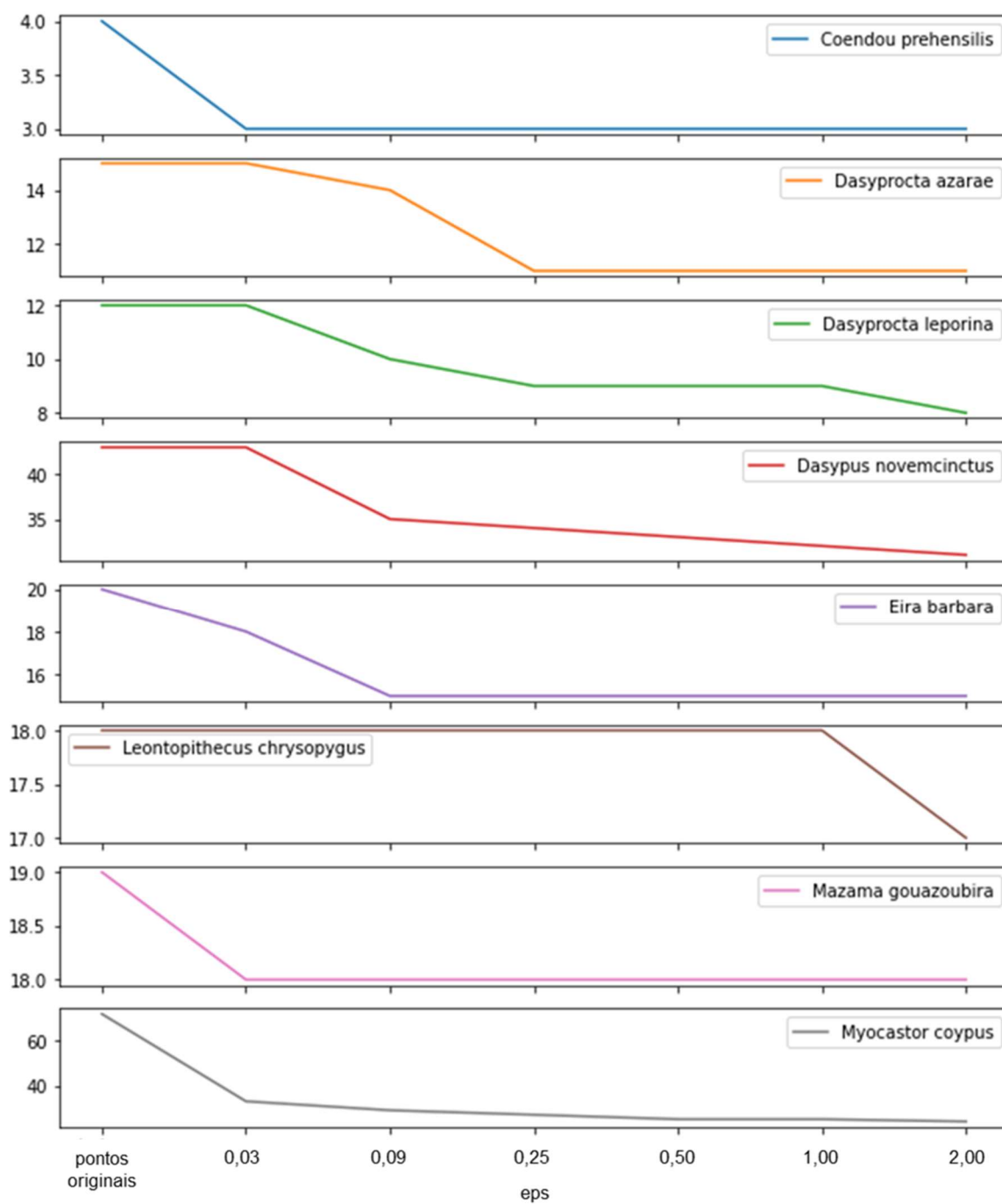


Figura 70 - Variação do número de clusters em função da variável eps do algoritmo DBSCAN (III)

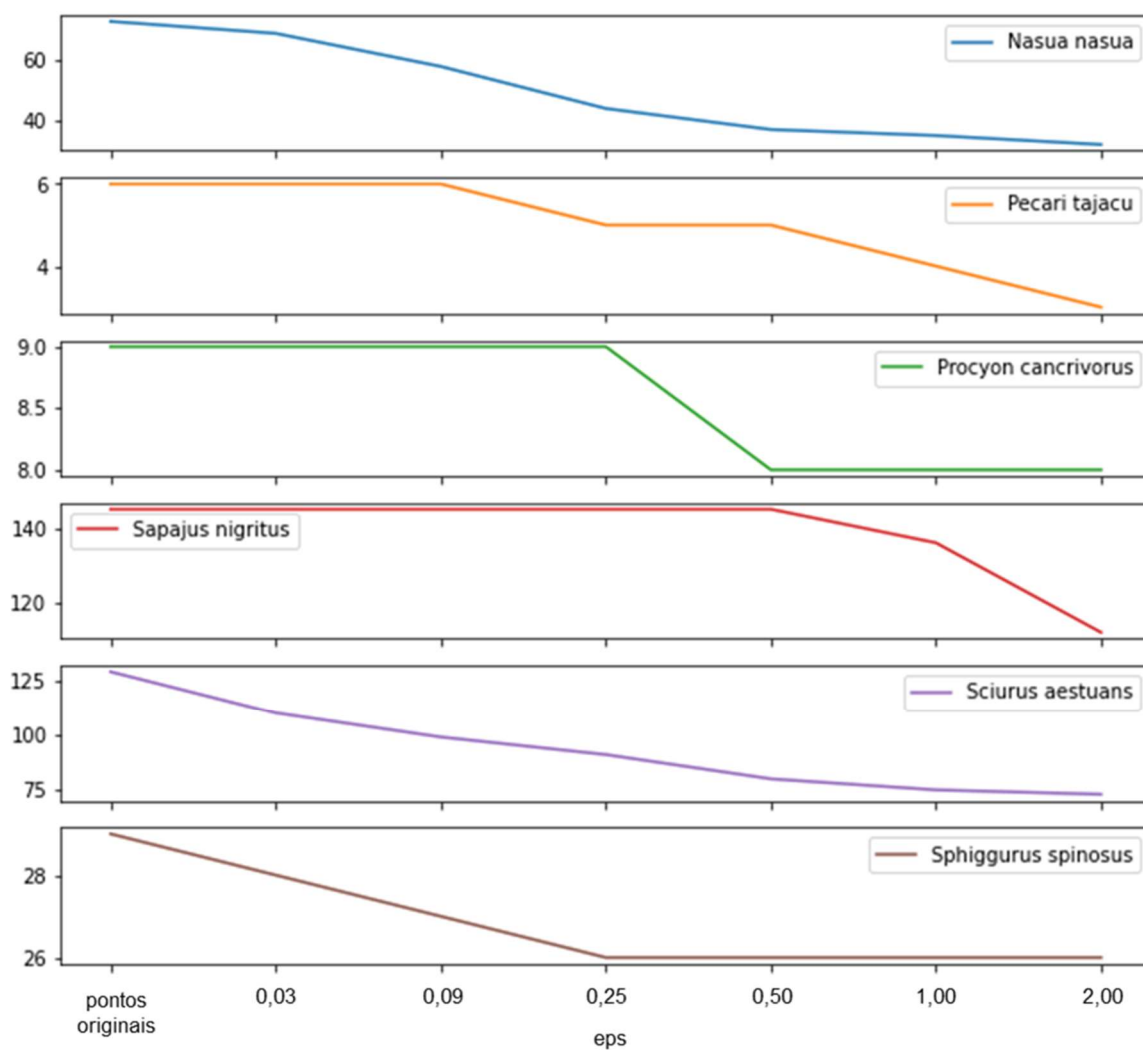
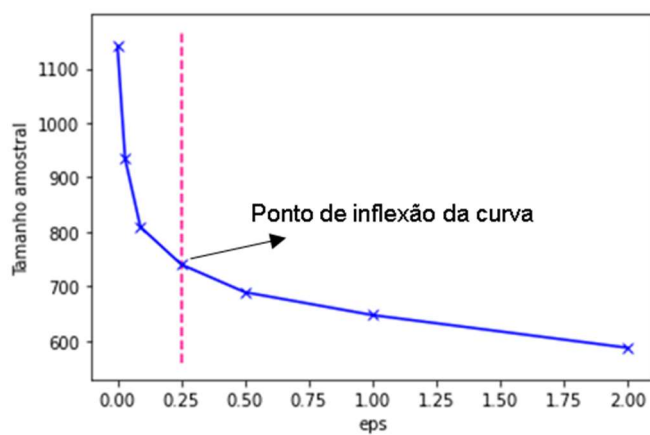
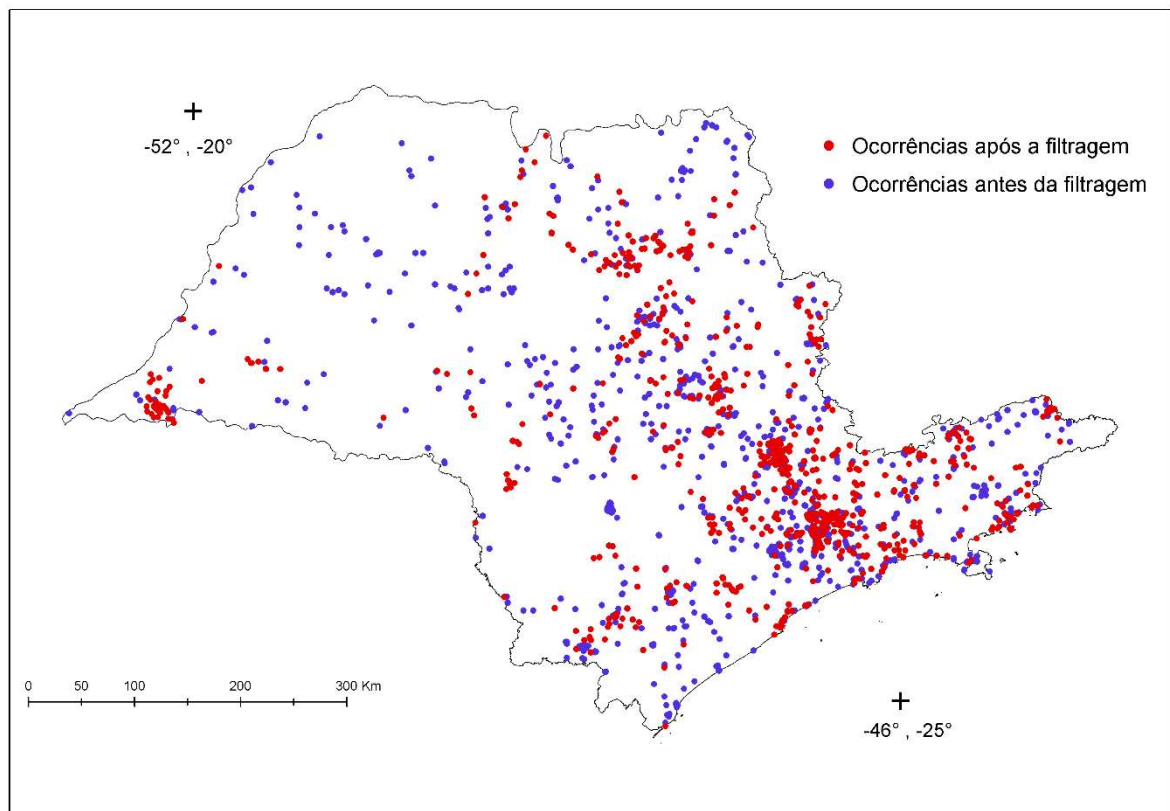


Figura 71 - Variação do número de clusters em função do eps considerando todas as espécies



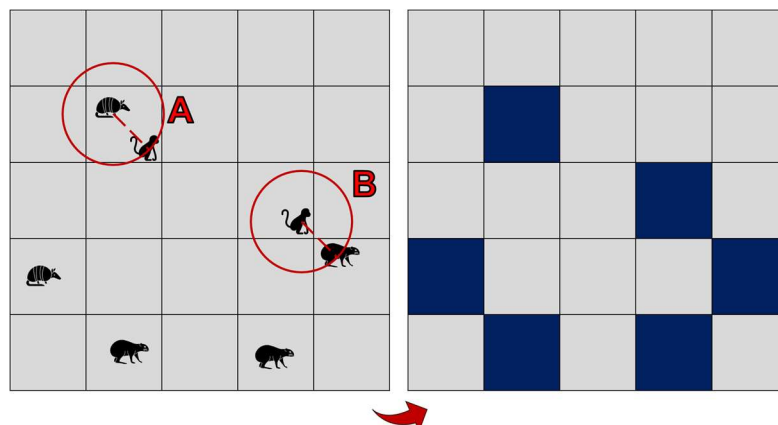
A Figura 72 mostra os pontos de ocorrência de espécies antes e depois das filtrações por metadados e espacial. Os pontos originais sem pontos filtrados nas proximidades referem-se às espécies excluídas na filtração por metadados. Os aglomerados não filtrados nas regiões noroeste e centro-oeste do Estado confirmam o reportado na literatura sobre a abundância de pequenos mamíferos nas paisagens agrícolas (GHELER-COSTA et al., 2012), já que as ordens de pequenos mamíferos generalistas foram removidas.

Figura 72 - Pontos de ocorrência antes e depois das filtrações por metadados e espacial



A discriminação das espécies é considerada nesta pesquisa para orientar as discussões da presença de mamíferos na área de estudo. No entanto, quando tomadas como amostras para treinamento dos modelos, caso duas ou mais espécies estejam localizadas dentro de um mesmo pixel nos planos de informação das variáveis ambientais, será contabilizada uma única amostra. A Figura 73 ilustra duas situações de proximidade entre espécies diferentes: em A, duas ocorrências de espécies diferentes coincidem espacialmente com o mesmo pixel da variável ambiental; em B, embora as ocorrências estejam espacialmente próximas, não coincidem com o mesmo pixel e, portanto, constituem amostras distintas; à direita, os pixels em azul são os que constituem amostras para os modelos.

Figura 73 - Esquema da geração de pixels amostrais nas variáveis ambientais a partir dos pontos de ocorrência de espécies

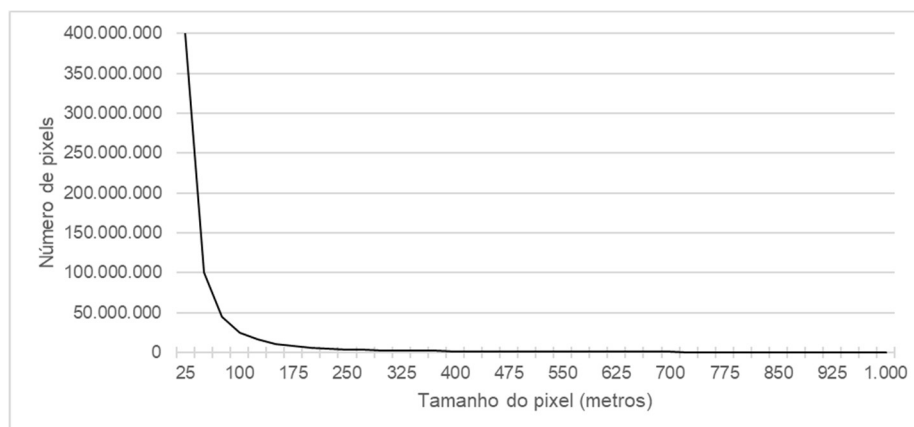


Como pontos representativos das espécies, foram então considerados os medóides dos clusters gerados a partir do eps de 250 metros.

#### 4.3 VARIÁVEIS PREDITORAS

A distância de 250 metros também foi considerada para reamostrar o tamanho dos pixels das imagens das variáveis ambientais. Hengl (2006) explica que o tamanho do pixel pode estar relacionado com o tamanho da área e poder de processamento do nosso computador. No caso do estado de São Paulo, que tem quase 250 mil km<sup>2</sup>, uma imagem com pixels de 250 metros (62.500 m<sup>2</sup>) equivale a 4 milhões de pixels (desconsiderando os pixels de fundo sem valor) (Figura 74).

Figura 74 - Quantidade de pixels para cobrir uma área de 250 mil km<sup>2</sup>



As Figuras 75 a 78 mostram os planos de informação de todas as variáveis preditoras. Nota-se que algumas classes de cobertura da terra (variáveis categóricas) têm baixa frequência. Sendo assim, foi gerado o histograma das classes de cobertura da terra antes e depois da reamostragem do tamanho do pixel (Figura 79) que mostra que, mesmo após esse processamento, a proporção de pixels para cada cobertura se mantém.

Figura 75 - Planos de informação das variáveis ambientais preditoras (I)

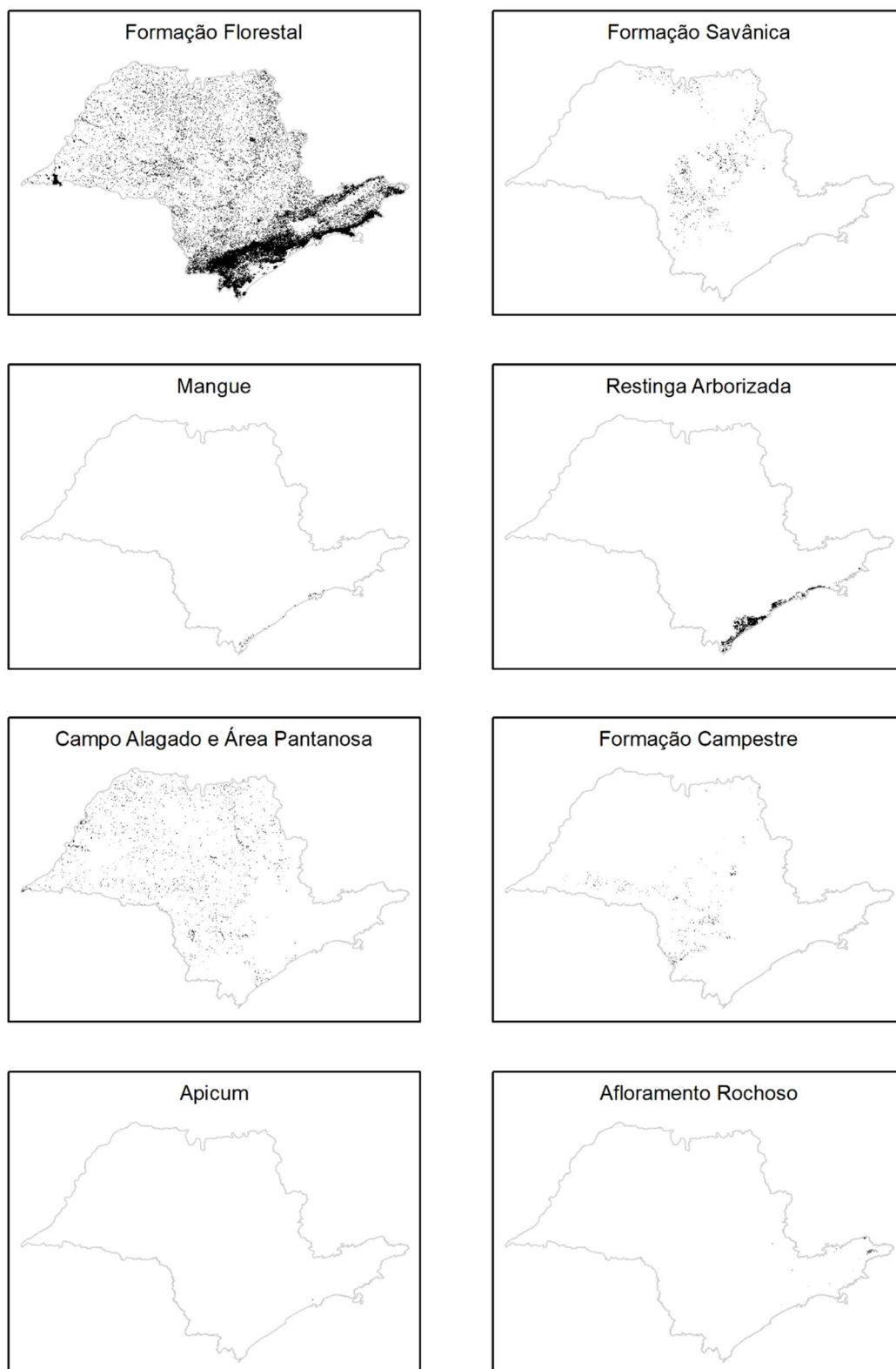


Figura 76 - Planos de informação das variáveis ambientais predictoras (II)

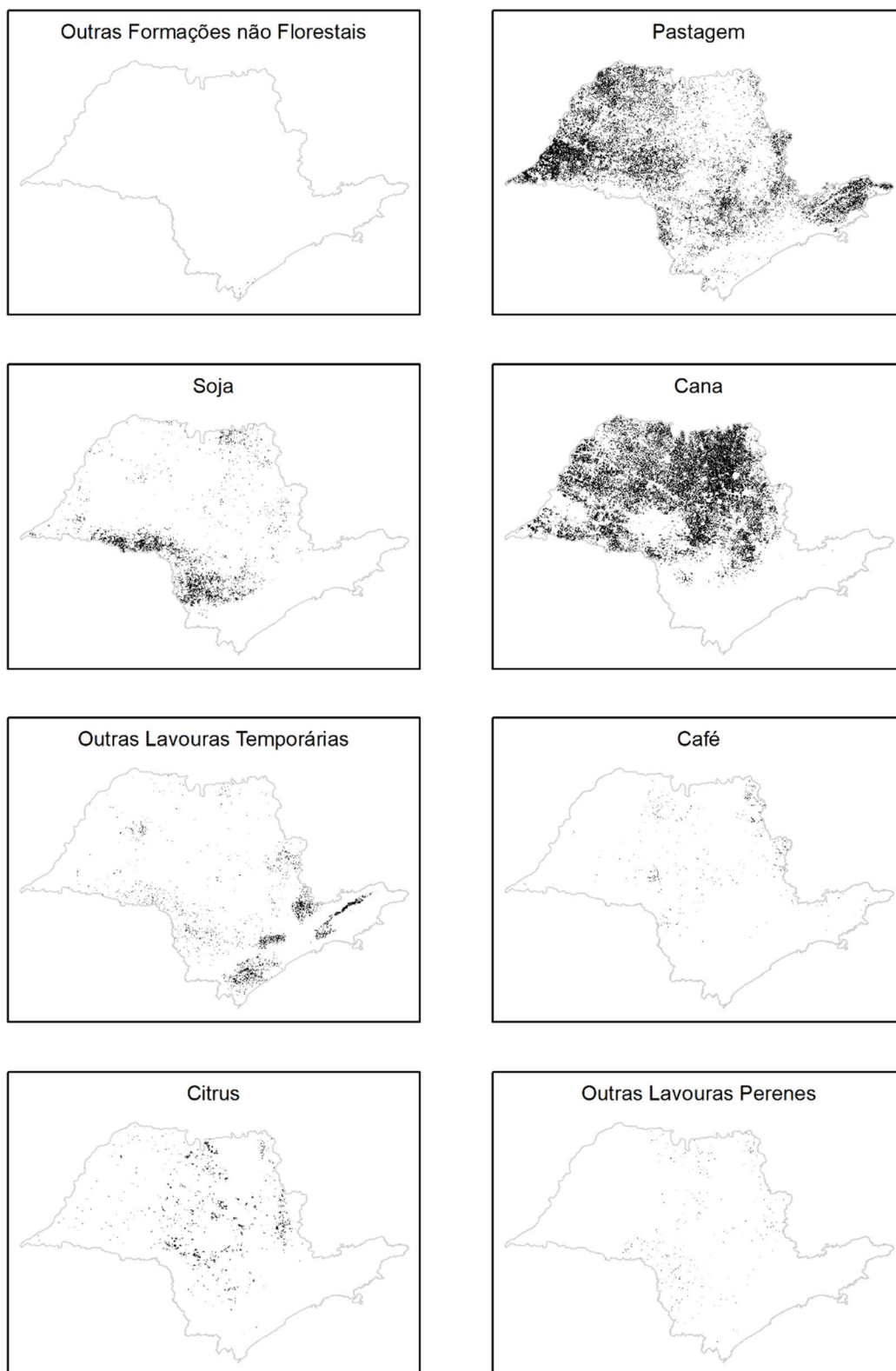


Figura 77 - Planos de informação das variáveis ambientais preditoras (IV)

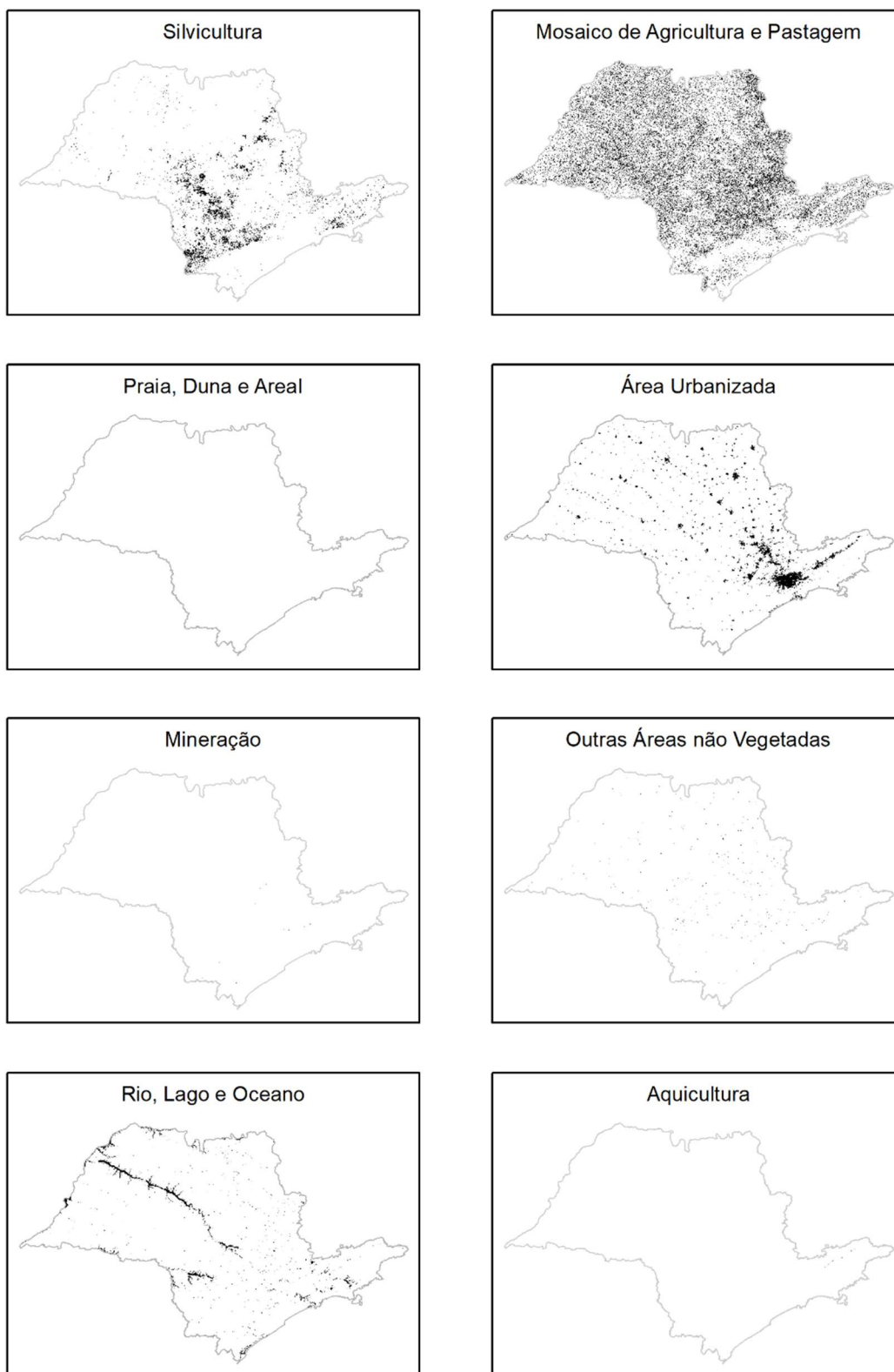




Figura 78 - Planos de informação das variáveis ambientais predictoras (IV)

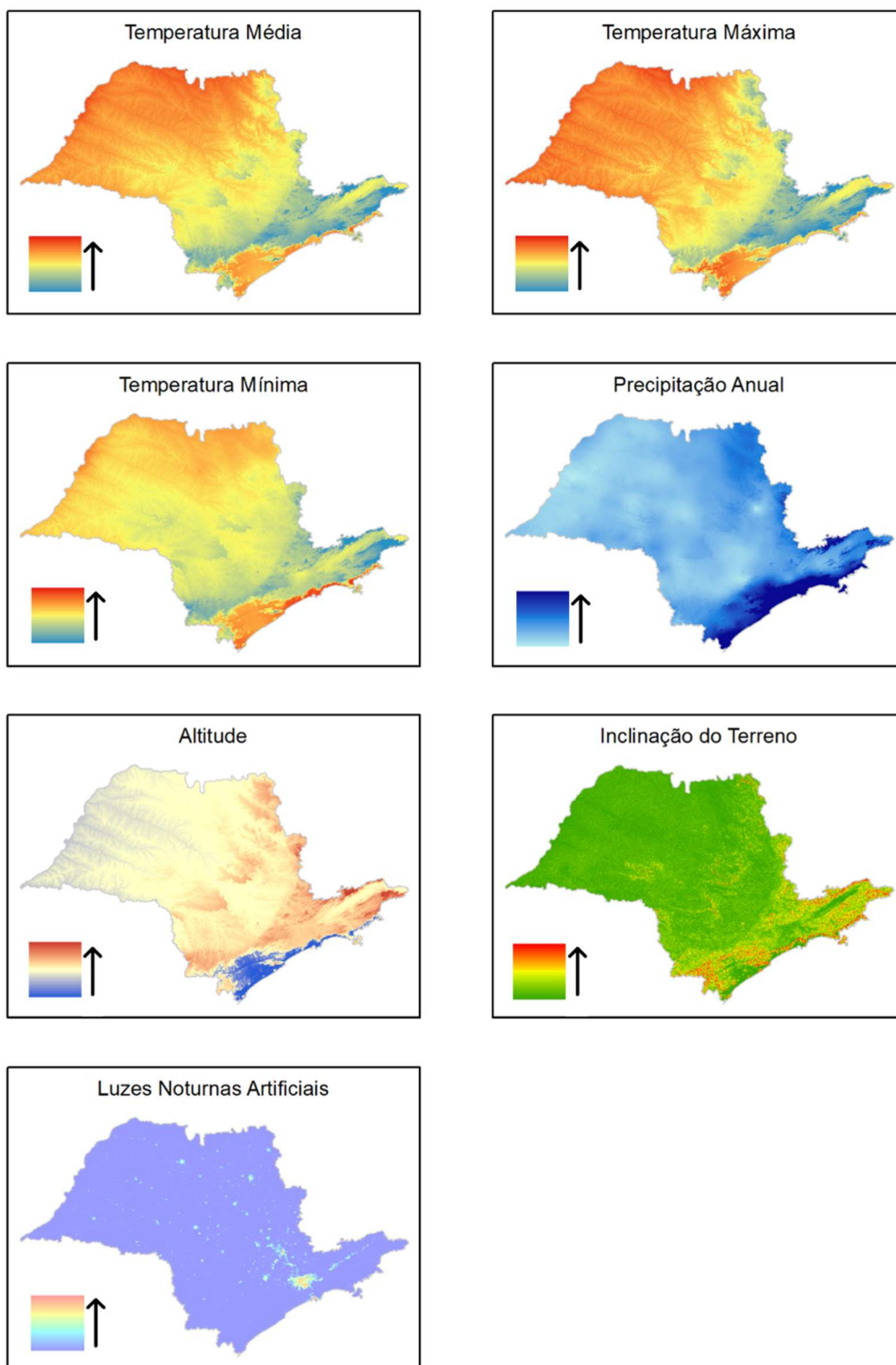
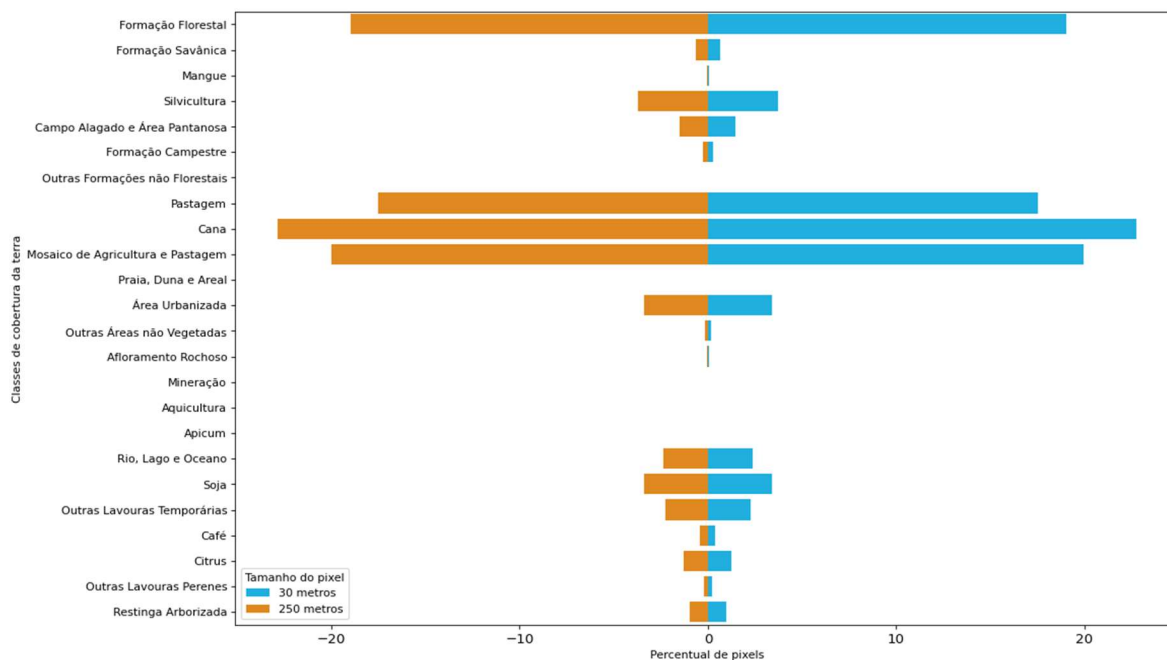


Figura 79 - Histograma das classes de uso da terra antes (30 metros) e depois (250 metros) da reamostragem do tamanho do pixel em toda a área de estudo



#### 4.4 ANÁLISE EXPLORATÓRIA DOS DADOS DOS PONTOS DE PRESENÇA

Foram realizados dois procedimentos de análise exploratória dos dados. Em um primeiro momento, foram analisados os valores das variáveis preditoras nos pontos de Presença para decisão sobre a remoção de pontos e para seleção das variáveis para o processamento da classificação OC-SVM (descrita na Seção 4.6). Após o resultado deste modelo e geração dos pontos de Pseudo-ausência, foi realizada outra análise exploratória para comparar esses pontos com os de Presença (descrita na Seção 4.7). Esta análise exploratória descreve as variáveis nos pontos amostrais de Presença e de Pseudo-ausência, que foram usados para o treinamento e teste dos modelos, e não se referem a toda a área de estudo.

Nos pontos espacialmente filtrados, considerando o valor de eps igual a 250 metros no DBSCAN, foi feita a extração dos valores dos pixels dos planos de informação normalizados (com valores de 0 1 1000). Para as variáveis numéricas climáticas e de relevo, foram removidos os pontos com valor igual a zero, que são pontos das bordas da área de interesse e condizem com o fundo da imagem. Este procedimento não foi feito para Luzes Noturnas Artificiais (LAN) pois, neste caso, os valores zerados devem ser considerados; são as áreas rurais em que não há luzes de estradas e rodovias e/ou edificações isoladas.

Foram verificadas quais classes de cobertura da terra não apresentaram pontos amostrais de ocorrência: Mangue, Outras Formações não Florestais, Outras áreas não Vegetadas, Praia, Duna e Areal, Afloramento Rochoso, Mineração, Aquicultura, Apicum, Outras Lavouras Perenes e Café.

Considerando que, apesar de haver algum grau de adaptabilidade dos mamíferos silvestres em habitar espaços antropogênicos, a ocorrência de animais silvestres no perímetro urbano não é adequada visto os riscos expostos anteriormente. Sendo assim, a remoção das ocorrências de espécies em áreas urbanas foi ponderada e a análise exploratória dos dados de Presença (AEDP) discriminou os pontos incidentes em área urbana.

A Tabela 23 identifica as espécies que compõem a amostra de pontos de Presença e a Tabela 24 mostra a ocorrência das espécies considerando todos os pontos (todas as coberturas da terra) e percentual das ocorrências em área urbana.

Tabela 23 - Espécies que compõem a amostra de pontos de Presença

<b>Espécie</b>	<b>Ordem</b>	<b>Família</b>	<b>Gênero</b>	<b>Estado de Conservação (2016)<sup>106</sup></b>	<b>Estado de Conservação (2022)<sup>107</sup></b>	<b>Nome popular<sup>108</sup></b>	<b>Nome popular em inglês<sup>107</sup></b>
<i>Alouatta caraya</i>	Primates	Atelidae	Alouatta	LC	-	barbado, bugio	Black-and-Gold Howler Monkey
<i>Alouatta guariba</i>	Primates	Atelidae	Alouatta	LC	-	bugio-ruivo, guariba	Brown Howler Monkey
<i>Brachyteles arachnoides</i>	Primates	Atelidae	Brachyteles	EN	EN	muriqui-do-sul, mono-carvoeiro	Southern Muriqui
<i>Bradypus variegatus</i>	Pilosa	Bradypodidae	Bradypus	LC	-	preguiça, bicho-preguiça	Brown-throated Sloth
<i>Cabassous unicinctus</i>	Cingulata	Chlamyphoridae	Cabassous	LC	-	tatu-de-rabo-mole	Southern Naked-tailed Armadillo
<i>Callicebus nigrifrons</i>	Primates	Pitheciidae	Callicebus	NT	-	guigó, sauá	Black-fronted Titi Monkey
<i>Callicebus personatus</i>	Primates	Pitheciidae	Callicebus	VU	VU	guigó, sauá	Northern Masked Titi Monkey
<i>Callithrix aurita</i>	Primates	Callitrichidae	<i>Callithrix</i>	VU	EN	sagui-da-serra-escuro	Buffy-tufted-ear Marmoset

<sup>106</sup> IUCN. The IUCN Red List of Threatened Species. Disponível em <<https://www.iucnredlist.org/>>. Acesso em 10 set. 2022.

<sup>107</sup> Ministério do Meio Ambiente. Portaria MMA de 7 de junho de 2022. Disponível em <[https://www.icmbio.gov.br/cepsul/images/stories/legislacao/Portaria/2020/P\\_mma\\_148\\_2022\\_altera\\_anexos\\_P\\_mma\\_443\\_444\\_445\\_2014\\_atualiza\\_especies\\_ameacadas\\_extincao.pdf](https://www.icmbio.gov.br/cepsul/images/stories/legislacao/Portaria/2020/P_mma_148_2022_altera_anexos_P_mma_443_444_445_2014_atualiza_especies_ameacadas_extincao.pdf)>. Acesso em 10 set. 2022.

<sup>108</sup> Taxeus. Lista de Espécies. Disponível em <<https://www.taxeus.com.br/>>. Acesso em 10 set. 2022.

<b>Espécie</b>	<b>Ordem</b>	<b>Família</b>	<b>Gênero</b>	<b>Estado de Conservação (2016) <sup>106</sup></b>	<b>Estado de Conservação (2022) <sup>107</sup></b>	<b>Nome popular <sup>108</sup></b>	<b>Nome popular em inglês <sup>107</sup></b>
<i>Callithrix jacchus</i>	Primates	Callitrichidae	<i>Callithrix</i>	LC	-	sagui-de-tufos-branco	Common Marmoset
<i>Callithrix penicillata</i>	Primates	Callitrichidae	<i>Callithrix</i>	LC	-	sagui, mico-estrela	Black-tufted-ear Marmoset
<i>Cavia aperea</i>	Rodentia	Caviidae	<i>Cavia</i>	LC	-	preá	Brazilian Guinea Pig
<i>Cerdocyon thous</i>	Carnivora	Canidae	<i>Cerdocyon</i>	LC	-	cachorro-do-mato, graxaim, raposa	Crab-eating Fox
<i>Chrysocyon brachyurus</i>	Carnivora	Canidae	<i>Chrysocyon</i>	NT	VU	lobo-guará, guará	Maned Wolf
<i>Coendou prehensilis</i>	Rodentia	Erethizontidae	<i>Coendou</i>	LC	-	ouriço, porco-espinho	Brazilian Porcupine
<i>Cuniculus paca</i>	Rodentia	Cuniculidae	<i>Cuniculus</i>	LC	-	paca	Spotted Paca
<i>Dasyprocta azarae</i>	Rodentia	Dasyproctidae	<i>Dasyprocta</i>	DD	-	cutia	Azaras Agouti
<i>Dasyprocta leporina</i>	Rodentia	Dasyproctidae	<i>Dasyprocta</i>	LC	-	cutia	Red-rumped Agouti
<i>Dasypus novemcinctus</i>	Cingulata	Dasypodidae	<i>Dasypus</i>	LC	-	tatu, tatu-galinha	Nine-banded Armadillo
<i>Dasypus septemcinctus</i>	Cingulata	Dasypodidae	<i>Dasypus</i>	LC	-	tatu, tatu-mulita, tatuí	Seven-banded Armadillo

<b>Espécie</b>	<b>Ordem</b>	<b>Família</b>	<b>Gênero</b>	<b>Estado de Conservação (2016) <sup>106</sup></b>	<b>Estado de Conservação (2022) <sup>107</sup></b>	<b>Nome popular <sup>108</sup></b>	<b>Nome popular em inglês <sup>107</sup></b>
<i>Eira barbara</i>	Carnivora	Mustelidae	<i>Eira</i>	LC	-	irara, papa-mel	Tayra
<i>Euphractus sexcinctus</i>	Cingulata	Chlamyphoridae	<i>Euphractus</i>	LC	-	tatu-peludo, tatu-peba	Six-banded Armadillo
<i>Galictis cuja</i>	Carnivora	Mustelidae	<i>Galictis</i>	LC	-	furão	Lesser Grison
<i>Leontopithecus chrysopygus</i>	Primates	Callitrichidae	<i>Leontopithecus</i>	EN	EN	mico-leão-preto	Black Lion Tamarin
<i>Leopardus guttulus</i>	Carnivora	Felidae	<i>Leopardus</i>	VU	VU	gato-do-mato	Southern Little Spotted Cat
<i>Leopardus pardalis</i>	Carnivora	Felidae	<i>Leopardus</i>	LC	-	jagatirica	Ocelot
<i>Leopardus wiedii</i>	Carnivora	Felidae	<i>Leopardus</i>	NT	VU	gato-maracajá, maracajá	Margay
<i>Lontra longicaudis</i>	Carnivora	Mustelidae	<i>Lontra</i>	NT	-	lontra	Neotropical Otter
<i>Lycalopex vetulus</i>	Carnivora	Canidae	<i>Lycalopex</i>	LC	VU	raposinha	Hoary Fox
<i>Mazama americana</i>	Artiodactyla	Cervidae	<i>Mazama</i>	DD	-	veado-mateiro	South American Red Brocket

<b>Espécie</b>	<b>Ordem</b>	<b>Família</b>	<b>Gênero</b>	<b>Estado de Conservação (2016) <sup>106</sup></b>	<b>Estado de Conservação (2022) <sup>107</sup></b>	<b>Nome popular <sup>108</sup></b>	<b>Nome popular em inglês <sup>107</sup></b>
<i>Mazama gouazoubira</i>	Artiodactyla	Cervidae	<i>Mazama</i>	LC	-	veado-catingueiro	South American Brow Brocket
<i>Myocastor coypus</i>	Rodentia	Myocastoridae	<i>Myocastor</i>	LC	-	ratão-do-banhado	Coypu
<i>Myrmecophaga tridactyla</i>	Pilosa	Myrmecophagidae	<i>Myrmecophaga</i>	VU	VU	tamanduá-bandeira	Giant Anteater
<i>Nasua nasua</i>	Carnivora	Procyonidae	<i>Nasua</i>	LC	-	quati	South American Coati
<i>Panthera onca</i>	Carnivora	Felidae	<i>Panthera</i>	NT	VU	onça-pintada	Jaguar
<i>Pecari tajacu</i>	Artiodactyla	Tayassuidae	<i>Pecari</i>	LC	-	cateto, caititu	Collared Peccary
<i>Procyon cancrivorus</i>	Carnivora	Procyonidae	<i>Procyon</i>	LC	-	guaxinim, mão-pelada	Crab-eating Raccoon
<i>Puma concolor</i>	Carnivora	Felidae	<i>Puma</i>	LC	-	onça-parda, suçuarana, leão-baio	Cougar
<i>Puma yagouaroundi</i>	Carnivora	Felidae	<i>Puma</i>	LC	-	jaguarundi, gato-mourisco	Jaguarundi
<i>Sapajus libidinosus</i>	Primates	Cebidae	<i>Sapajus</i>	LC	-	macaco-prego	Bearded Capuchin
<i>Sapajus nigritus</i>	Primates	Cebidae	<i>Sapajus</i>	NT	-	macaco-prego	Black-horned Tufted Capuchin

<b>Espécie</b>	<b>Ordem</b>	<b>Família</b>	<b>Gênero</b>	<b>Estado de Conservação (2016) <sup>106</sup></b>	<b>Estado de Conservação (2022) <sup>107</sup></b>	<b>Nome popular <sup>108</sup></b>	<b>Nome popular em inglês <sup>107</sup></b>
<i>Sciurus aestuans</i>	Rodentia	Sciuridae	<i>Sciurus</i>	LC	-	caxinguelê, esquilo	Guianan Squirrel
<i>Sphiggurus spinosus</i>	Rodentia	Erethizontidae	<i>Sphiggurus</i>	LC	-	ouriço-cacheiro	Paraguayan Hairy Dwarf Porcupine
<i>Sylvilagus brasiliensis</i>	Lagomorpha	Leporidae	<i>Sylvilagus</i>	LC	-	coelho, tapeti	Tapeti
<i>Tamandua tetradactyla</i>	Pilosa	Myrmecophagidae	<i>Tamandua</i>	LC	-	tamanduá-de-colete, tamanduá-mirim	Southern Tamandua
<i>Tapirus terrestris</i>	Perissodactyla	Tapiridae	<i>Tapirus</i>	VU	VU	anta	South American Tapir
<i>Tayassu pecari</i>	Artiodactyla	Tayassuidae	<i>Tayassu</i>	VU	VU	queixada, porco-do-mato	White-lipped Peccary
<i>Thaptomys nigrita</i>	Rodentia	Cricetidae	<i>Thaptomys</i>	LC	-	rato-do-chão	Ebony Grass Mouse



Tabela 24 - Pontos de Presença de espécies em todas as classes de cobertura da terra e apenas na área urbana

<b>Espécie</b>	<b>Todas as classes de cobertura da terra</b>	<b>Área não urbana</b>	<b>% de ocorrência da espécie em área urbana</b>
<i>Alouatta caraya</i>	4	1	75%
<i>Alouatta guariba</i>	36	29	19%
<i>Brachyteles arachnoides</i>	11	11	0%
<i>Bradypus variegatus</i>	16	14	13%
<i>Cabassous unicinctus</i>	1	1	0%
<i>Callicebus nigrifrons</i>	21	19	10%
<i>Callicebus personatus</i>	5	4	20%
<i>Callithrix aurita</i>	16	13	19%
<i>Callithrix jacchus</i>	55	27	51%
<i>Callithrix penicillata</i>	101	51	50%
<i>Cavia aperea</i>	18	13	28%
<i>Cerdocyon thous</i>	41	34	17%
<i>Chrysocyon brachyurus</i>	13	13	0%
<i>Coendou prehensilis</i>	3	3	0%
<i>Cuniculus paca</i>	5	2	60%
<i>Dasyprocta azarae</i>	11	7	36%
<i>Dasyprocta leporina</i>	7	4	43%

<b>Espécie</b>	<b>Todas as classes de cobertura da terra</b>	<b>Área não urbana</b>	<b>% de ocorrência da espécie em área urbana</b>
<i>Dasyopus novemcinctus</i>	33	24	27%
<i>Dasyopus septemcinctus</i>	2	2	0%
<i>Eira barbara</i>	15	11	27%
<i>Euphractus sexcinctus</i>	1	0	100%
<i>Galictis cuja</i>	2	1	50%
<i>Leontopithecus chrysopygus</i>	18	18	0%
<i>Leopardus guttulus</i>	5	5	0%
<i>Leopardus pardalis</i>	5	5	0%
<i>Leopardus wiedii</i>	1	1	0%
<i>Lontra longicaudis</i>	9	8	11%
<i>Lycalopex vetulus</i>	5	4	20%
<i>Mazama americana</i>	1	1	0%
<i>Mazama gouazoubira</i>	18	13	28%
<i>Myocastor coypus</i>	27	21	22%
<i>Myrmecophaga tridactyla</i>	16	16	0%
<i>Nasua nasua</i>	44	33	25%
<i>Panthera onca</i>	1	1	0%

<b>Espécie</b>	<b>Todas as classes de cobertura da terra</b>	<b>Área não urbana</b>	<b>% de ocorrência da espécie em área urbana</b>
<i>Pecari tajacu</i>	5	4	20%
<i>Procyon cancrivorus</i>	9	8	11%
<i>Puma concolor</i>	17	16	6%
<i>Puma yagouaroundi</i>	6	6	0%
<i>Sapajus libidinosus</i>	1	1	0%
<i>Sapajus nigritus</i>	145	101	30%
<i>Sciurus aestuans</i>	85	63	26%
<i>Sphiggurus spinosus</i>	24	16	33%
<i>Sylvilagus brasiliensis</i>	7	5	29%
<i>Tamandua tetradactyla</i>	6	6	0%
<i>Tapirus terrestris</i>	24	23	4%
<i>Tayassu pecari</i>	2	2	0%
<i>Thaptomys nigrita</i>	1	0	100%

Para algumas espécies, a ocorrência em áreas urbanas é expressiva, como *Alouatta caraya* (75%), *Cuniculus paca*, (60%), *Galictis cuja* (50%) e *Callithrix penicillata* (50%). As Figuras 80 a 82 mostram o histograma de ocorrências por espécies e o efeito quantitativo na remoção das ocorrências em área urbana, que somam 26% dos pontos. Esse alto percentual pode ser decorrente do fácil acesso por parte de quem avista as ocorrências.

Considerando todas as espécies, as classes de cobertura da terra com maior incidência de pontos foram Formação Florestal (32%), Área Urbanizada (26%) e Mosaico de Agricultura e Pastagem (19%) (Figura 83). Quando removidos os pontos na Área Urbanizada, os

percentuais de Formação Florestal e Pastagem sobem para, respectivamente, 44% e 25% (Figura 84).

Figura 80 - Histograma de ocorrências por espécies, mostrando a quantidade de todos os pontos e dos pontos incidentes em área urbana (área hachurada) (I)

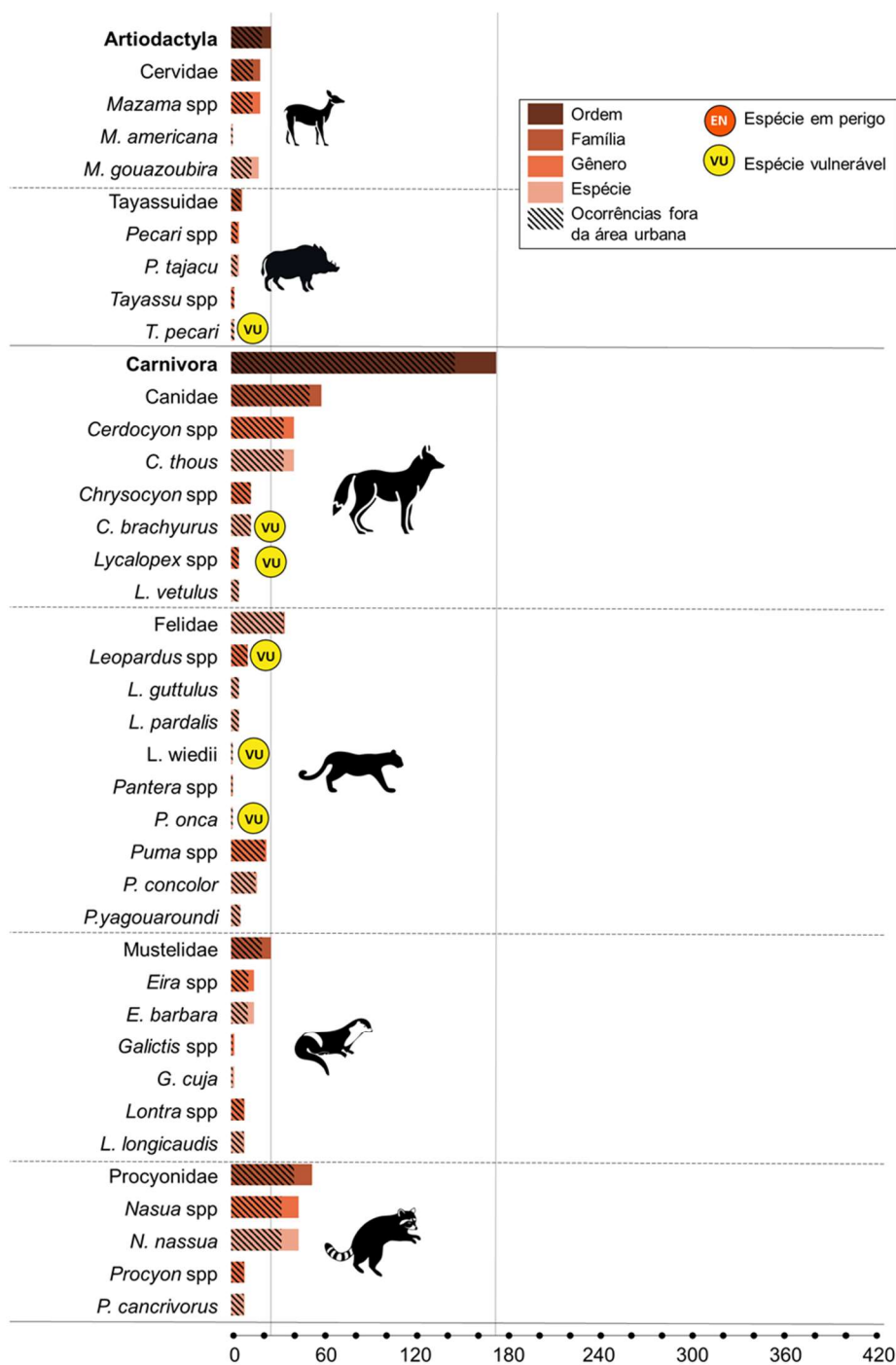


Figura 81 - Histograma de ocorrências por espécies, mostrando a quantidade de todos os pontos e dos pontos incidentes em área urbana (área hachurada) (II)

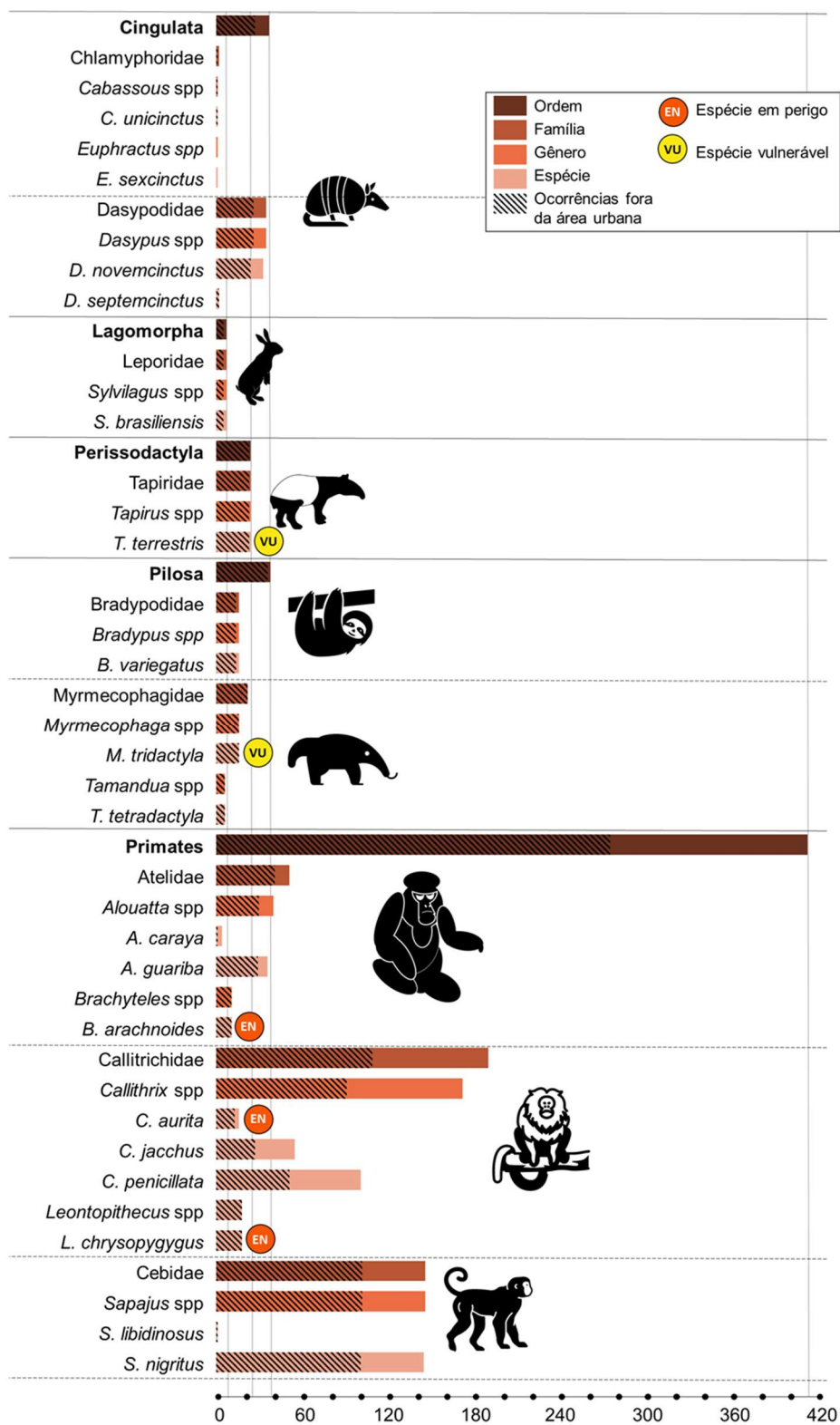


Figura 82 - Histograma de ocorrências por espécies, mostrando a quantidade de todos os pontos e dos pontos incidentes em área urbana (área hachurada) (III)

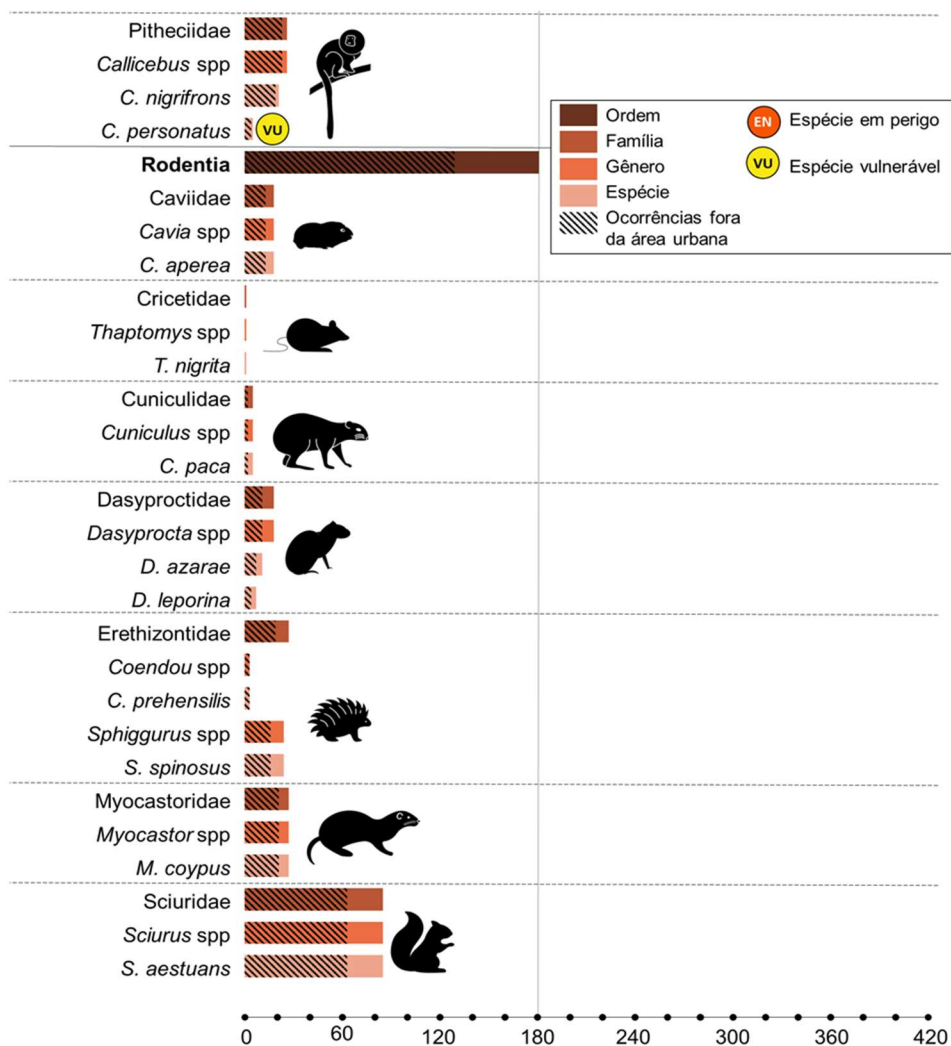


Figura 83 - Pontos de Presença por classe de cobertura da terra (todas as classes)

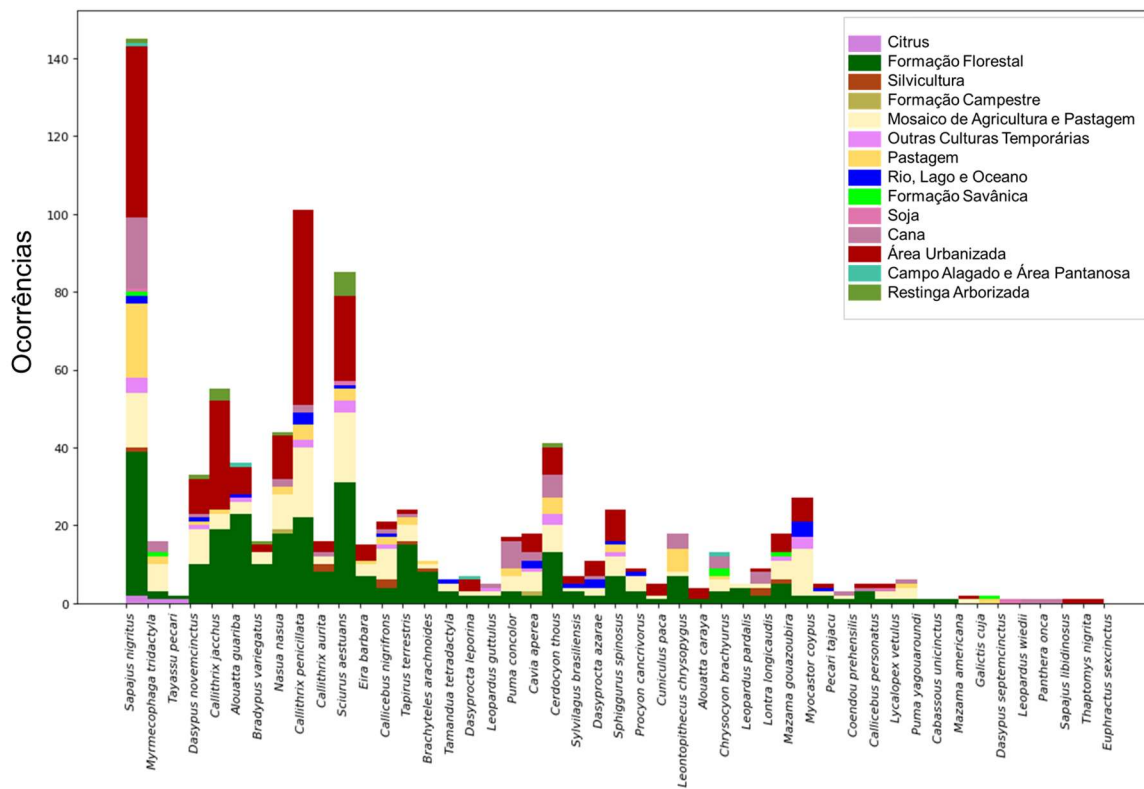
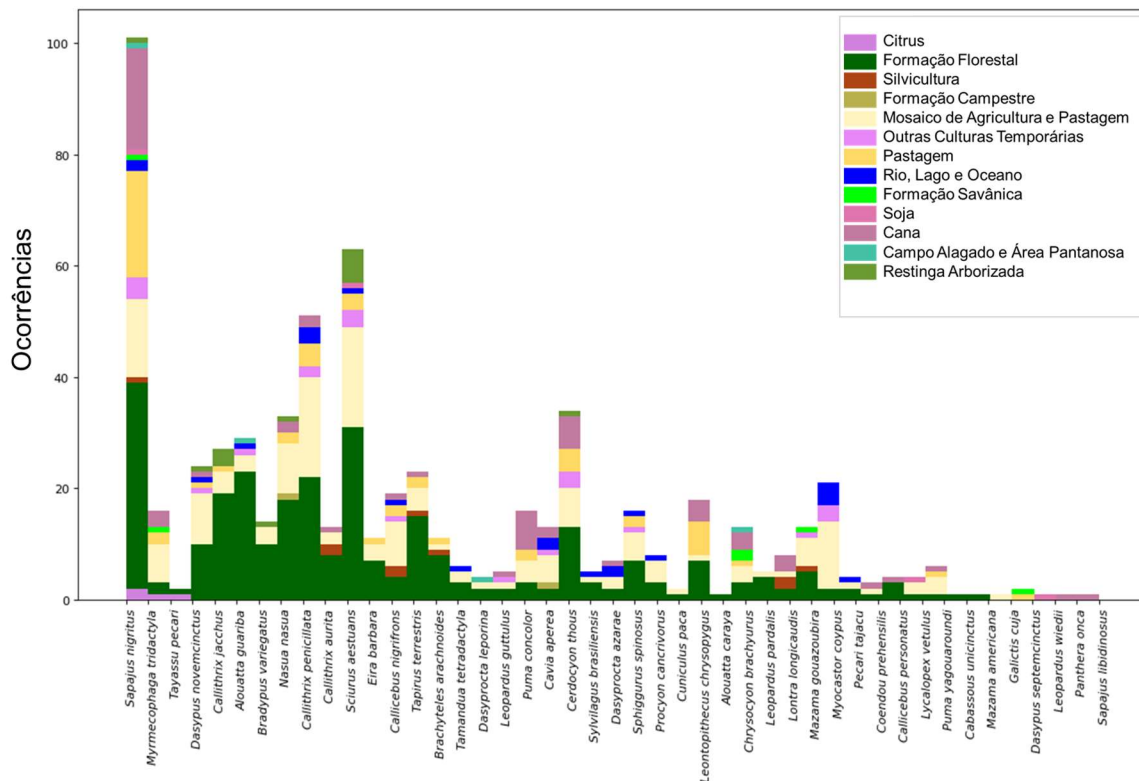


Figura 84 - Pontos de Presença por classe de cobertura da terra (exceto Área Urbanizada)



Na AEDP também foram averiguados os valores das variáveis preditoras nos pontos de amostragem de presença. Foram avaliados os efeitos da remoção dos pontos em áreas urbanas nas variáveis numéricas para os valores originais (não normalizados). Para Temperatura Mínima, Média e Máxima, a supressão dos pontos não afeta a média, embora ocorra uma redução nos outliers, e os valores no entorno da média tornem-se mais frequentes (Figura 85). Para precipitação, não há alteração além da diminuição de outliers (Figura 86). Como seria esperado, para LAN, a remoção dos pontos em áreas urbanas é significativa, com aumento nos valores médios e maior frequência de valores mais altos. Fora da área urbana, a maior frequência está em torno de zero nanoWatts/cm<sup>2</sup>/sr, com valores máximos próximos de 90 nanoWatts/cm<sup>2</sup>/sr. Por outro lado, nos pontos que incidem em áreas urbanizadas, a média da radiação é de aproximadamente 40 nanoWatts/cm<sup>2</sup>/sr, com valores máximos em torno de 20 nanoWatts/cm<sup>2</sup>/sr (Figura 87). A altitude é bimodal, com alta frequência de pontos em altitudes intermediárias, em torno de 750 metros, e certa frequência em baixas altitudes (litoral e várzea de rios) (Figura 88). Para Inclinação do Terreno, não houve mudanças significativas na média e variância dos pontos (Figura 89). Sendo assim, não há perdas em remover os pontos incidentes em áreas urbanas.

Figura 85 - Estatísticas da variável Temperatura em áreas não urbanas e áreas urbanas

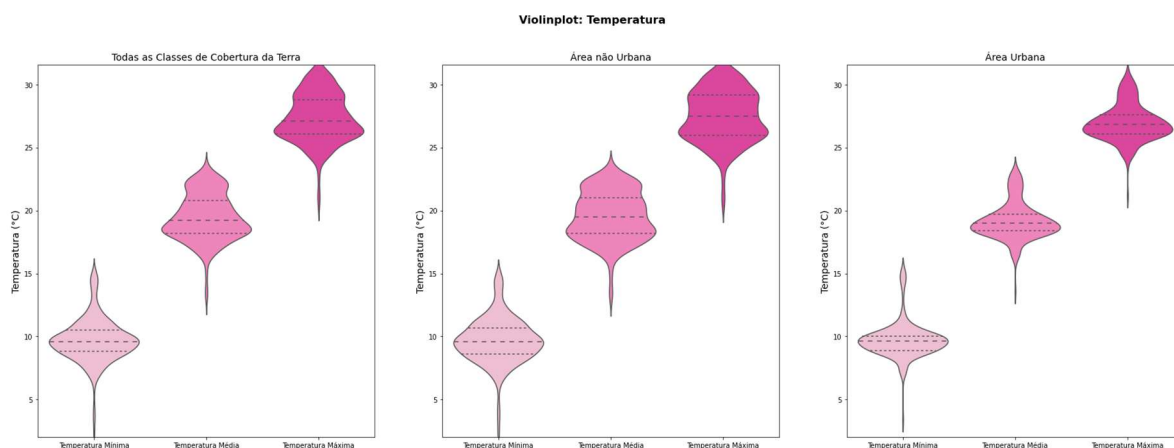




Figura 86 - Estatísticas da variável Precipitação em áreas não urbanas e áreas urbanas

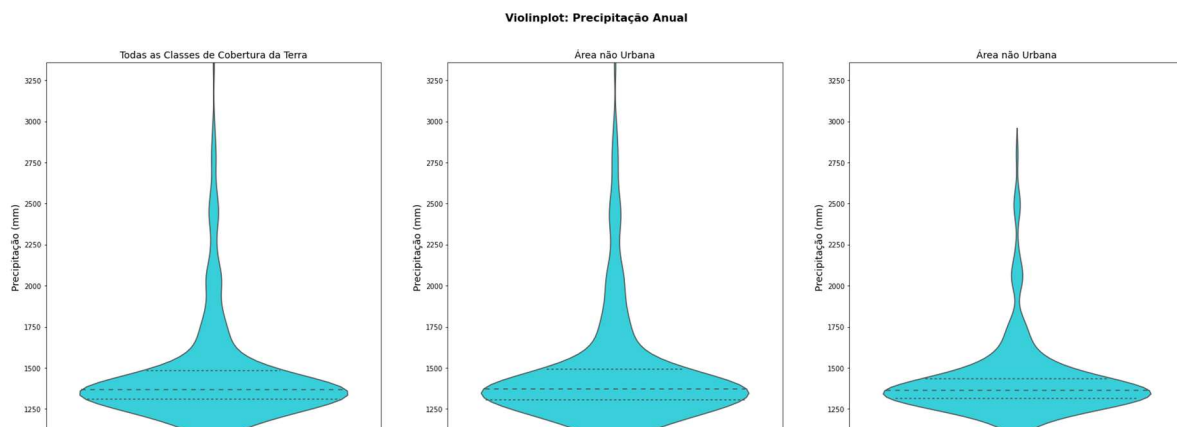


Figura 87 - Estatísticas da variável Luzes Noturnas Artificiais em áreas não urbanas e áreas urbanas

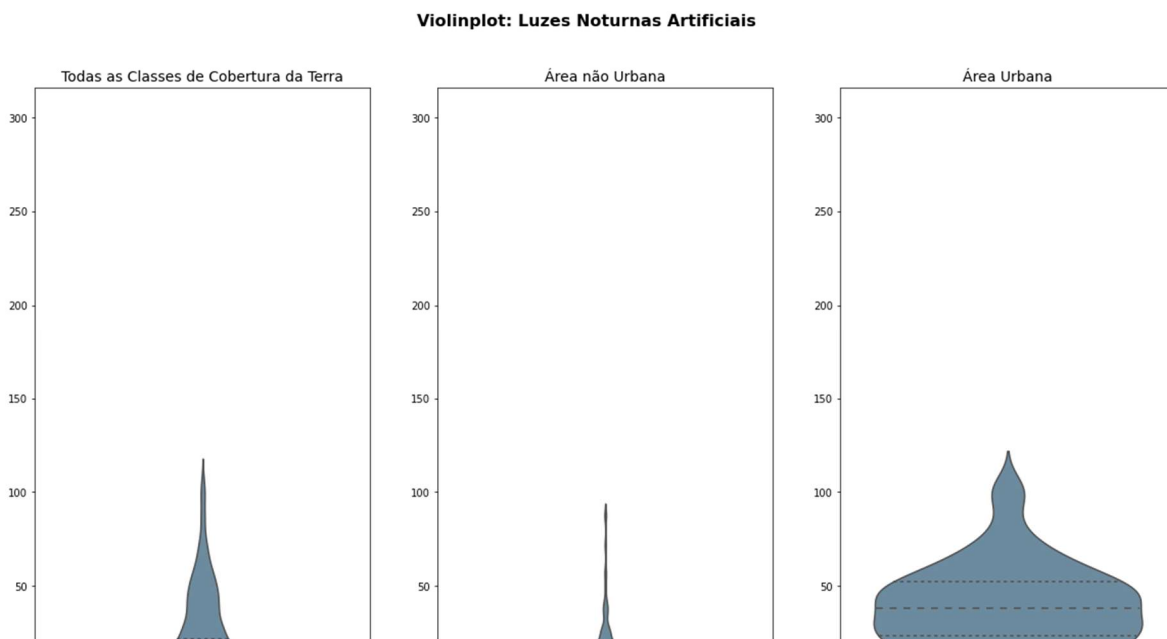


Figura 88 - Estatísticas da variável Altitude em áreas não urbanas e áreas urbanas

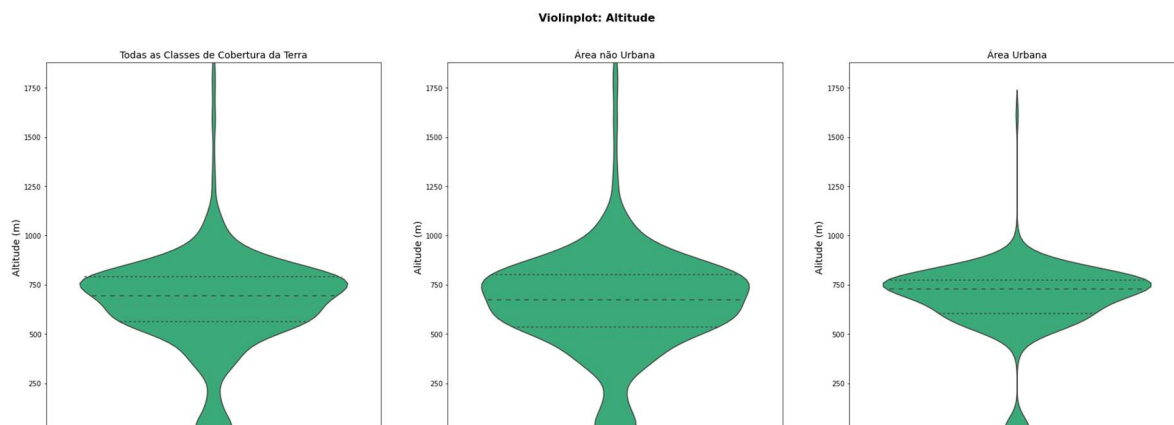
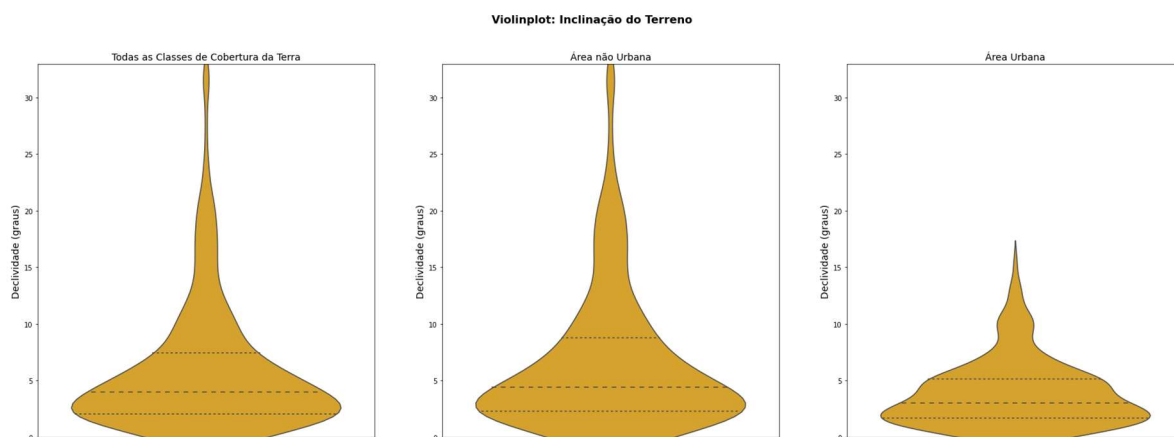


Figura 89 - Estatísticas da variável Inclinação do Terreno em áreas não urbanas e áreas urbanas

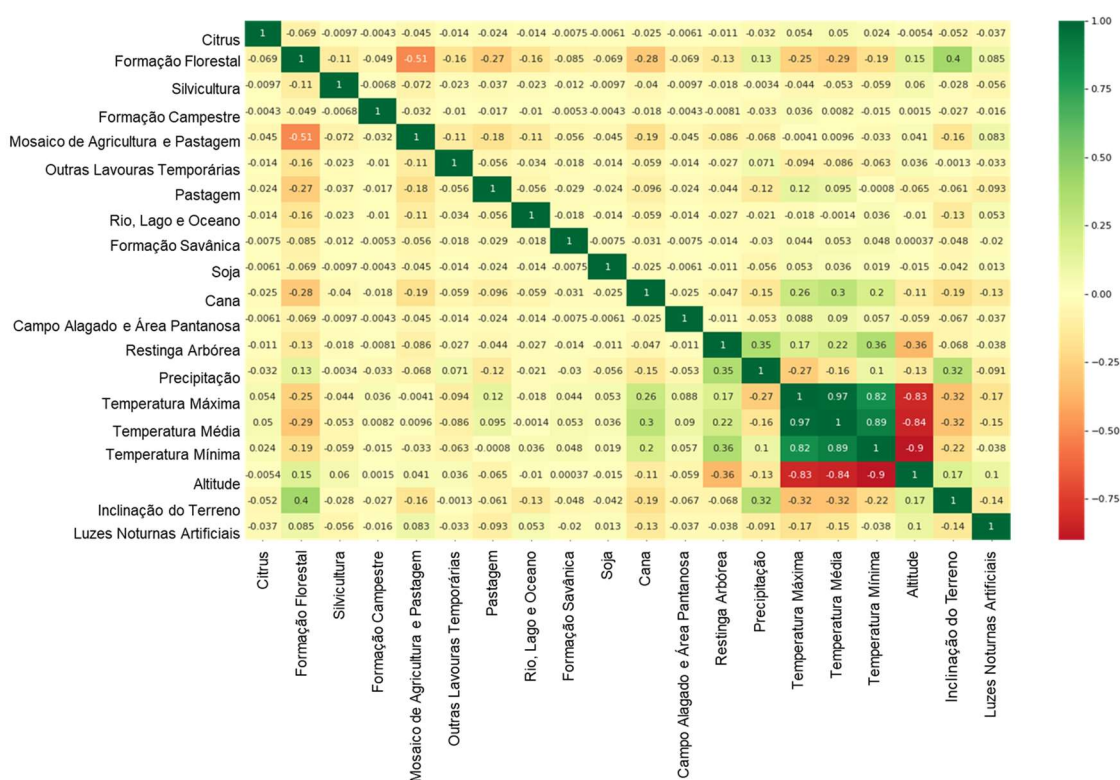


O correlograma na Figura 90 mostra as relações lineares entre as variáveis, tanto as categóricas quanto as numéricas. As variáveis relacionadas à Temperatura são altamente correlacionadas entre si (coeficiente de determinação  $R^2 > 0.80$ ). Isto posto, destaca-se a importância de avaliar como estas variáveis são selecionadas para modelos como o BioClim, que são baseados nas características climáticas sem necessariamente considerar outros aspectos importantes para adequabilidade de habitat. A variação da Temperatura também é fortemente explicada pela variação da Altitude. As demais variáveis não são correlacionadas entre si, com  $R^2$  de até 0,51 (em módulo).

Quando duas ou mais variáveis preditoras têm uma forte relação linear e não podem prever independentemente o valor da variável dependente, pode haver o problema de colinearidade na execução de um modelo de regressão. Então, as variáveis preditoras explicam parte da mesma variância na variável dependente, reduzindo sua significância estatística. De

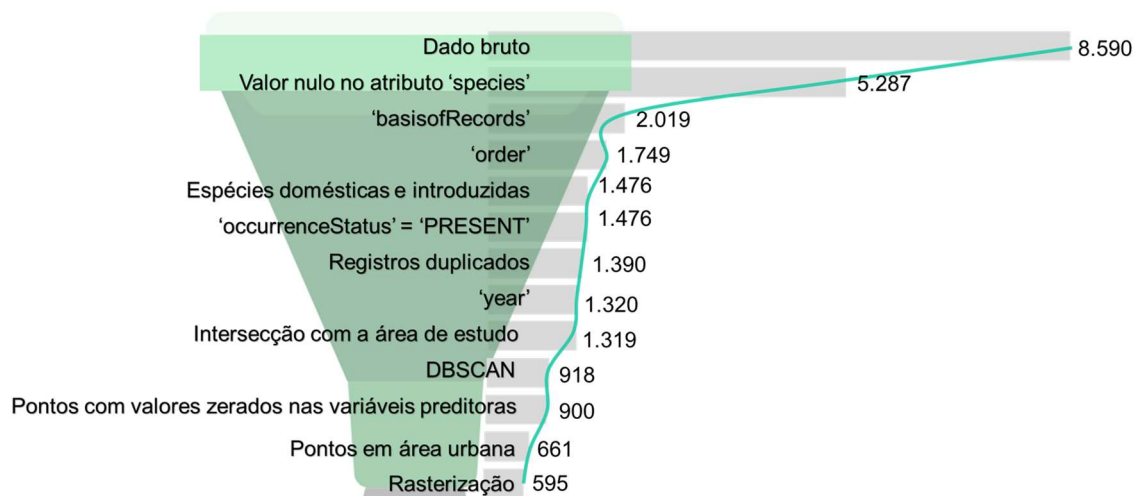
acordo com De Marco Júnior e Nóbrega (2018), a colinearidade entre os preditores diminui a eficácia e aumenta a incerteza dos modelos de distribuição de espécies, o que é mais inconveniente para modelos mais simples. Sendo assim, para o modelo OC-SVM, foram removidas as variáveis Temperatura Máxima, Temperatura Mínima, Altitude e Área Urbanizada, além daquelas zeradas nos pontos de ocorrência (Mangue, Outras Formações não Florestais, Outras áreas não Vegetadas, Praia, Duna e Areal, Afloramento Rochoso, Mineração, Aquicultura, Apicum, Outras Lavouras Perenes e Café).

Figura 90 - Correlograma das variáveis para os pontos de Presença



No processo de conversão dos pontos de presença para o formato raster, a amostra de pontos de presença foi reduzida. Isto porque alguns pontos podem coincidir com o mesmo pixel no dado raster. A Figura 91 mostra a redução do tamanho amostral após a remoção dos pontos com valores zerados nas variáveis numéricas e dos pontos em áreas urbanas. Sendo assim, os MAHs foram alimentados por 595 pontos de presença.

Figura 91 - Variação na quantidade de pontos amostrais de Presença ao longo do fluxo de processamento



#### 4.5 CLASSIFICAÇÃO OC-SVM

Para o OC-SVM foram processados dois modelos: um com os hiperparâmetros com os valores padrão do modelo e outro com os hiperparâmetros ajustados. Os hiperparâmetros padrão e os definidos pelo GridSearchCV estão na Tabela 25; embora os hiperparâmetros sejam distintos, a acurácia de treino e teste dos dois modelos são iguais (Tabela 26). A descrição dos hiperparâmetros do OC-SVM pode ser acessada na sua documentação<sup>109</sup>.

Tabela 25 - Hiperparâmetros da classificação OC-SVM

	Hiperparâmetros padrão	Hiperparâmetros ajustados pelo GridSearchCV
degree	3	1
gamma	scale	scale
kernel	rbf	sigmoid
nu	0,5	0,25

Tabela 26 - Métricas de avaliação da classificação OC-SVM

Hiperparâmetros padrão	Hiperparâmetros ajustados
------------------------	---------------------------

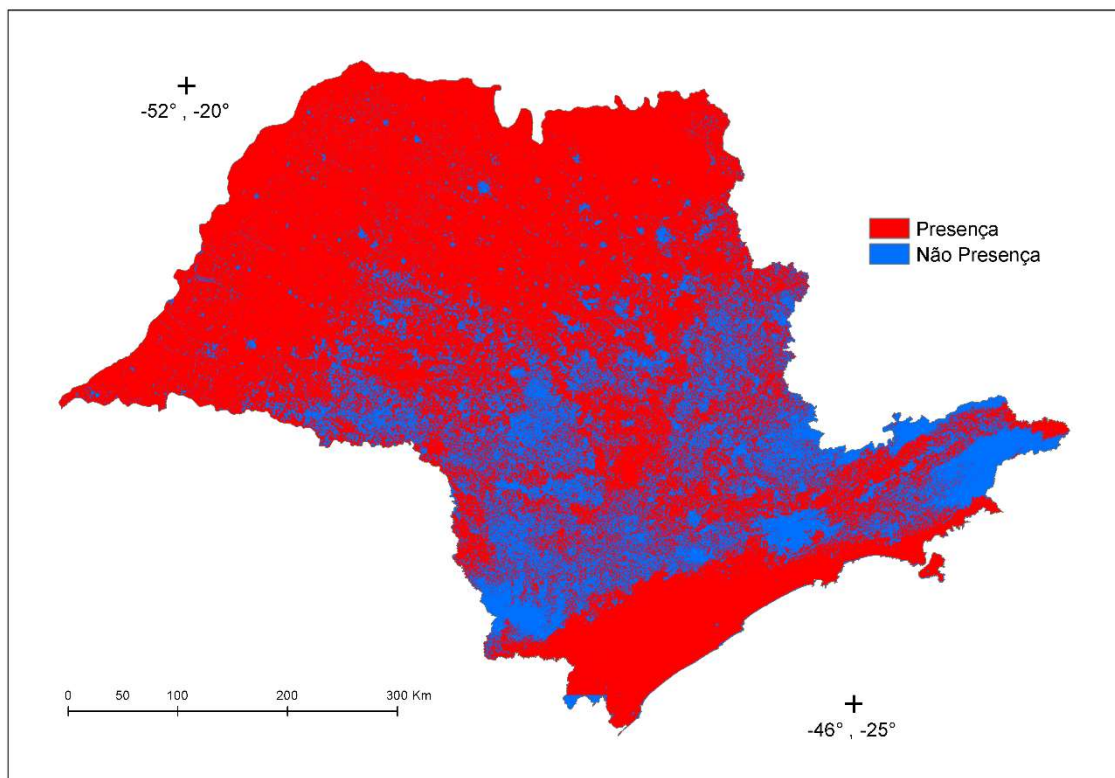
<sup>109</sup> sklearn.svm.OneClassSVM. Disponível em <<https://scikit-learn.org/stable/modules/generated/sklearn.svm.OneClassSVM.html>>. Acesso em 30 out. 2022.

	<b>pelo GridSearchCV</b>	
Acurácia do treinamento	0,50	0,50
Acurácia do teste	0,56	0,56
Revocação	0,56	0,56
Precisão	1,0	1,0
F1	0,72	0,72

Precisão é a proporção de Presença que está correta. Deve ser considerada quando as falsas presenças são mais prejudiciais do que falsas ausências. Seria o caso de considerar um local adequado, quando na verdade não é, e entender que seria apropriado para a criação de uma área protegida. Revocação é a proporção, entre todas as situações de Presença como valor esperado, quantas estão corretas. É importante quando as falsas ausências são mais prejudiciais do que as falsas presenças. Neste caso, um local com boa adequabilidade é tomado como não apropriado para as espécies e acaba sendo desconsiderado para abrigar uma área protegida. No caso de espécies ameaçadas, as falsas ausências são ainda mais desfavoráveis. O F1 é a média harmônica entre precisão e revocação, sendo uma medida única para avaliação, que varia de 0 a 1. A classificação OC-SVM teve uma pontuação F1 de 0,72, que será considerada suficiente para geração dos pontos de Pseudo-ausência.

No modelo de classificação OC-SVM, a adequabilidade de habitat coincide com as áreas onde estão os grandes remanescentes de vegetação nativa, como na Serra do Mar e Cuestas, e também nas áreas agrícolas (Figura 92). As áreas urbanizadas foram classificadas como não presença dado que os pontos de presença nesse locais foram removidos. A partir do resultado da classificação, foram gerados pontos aleatórios de pseudo-ausência sobre a classe “Não Presença”, de mesmo tamanho amostral (n=595).

Figura 92 - Classificação OC-SVM



#### 4.6 ANÁLISE EXPLORATÓRIA DOS DADOS DE PSEUDO-AUSÊNCIA

A Tabela 27 mostra as estatísticas das variáveis numéricas e o resultado dos teste-t de Welch para comparação de médias (WELCH, 1947) e do teste de Levene para comparação de variâncias (LEVENE, 1960). As células em roxo indicam que foi rejeitada a hipótese nula de que não há diferença entre as amostras de Presença e Pseudo-ausência; as células em verde, indicam que a hipótese nula foi aceita. Essa diferença entre as amostras favorece a discriminação destas classes nos modelos.

Tabela 27 - Comparação das estatísticas dos pontos de Presença e de Pseudo-ausência

Variável	Estatísticas	Presença	Pseudo-ausência
Temperatura Média	média	689,2	672,9
	desvio padrão	125,7	118,8
	25%	601	601
	50%	690	703
	75%	785	753
	Teste t de Welch	Estatística t = -2,30 p-value = 0,02 H <sub>0</sub> é rejeitada: há diferença entre as amostras de presença e pseudo-ausência.	
	Teste de Levene	Estatística de Levene = 8,26 p-value = 0,004 H <sub>0</sub> é rejeitada: há diferença entre as amostras de presença e pseudo-ausência.	
Temperatura Máxima	média	695,7	698,0
	desvio padrão	126,0	126,3
	Teste t de Welch	Estatística t = 0,31 p-value = 0,76 H <sub>0</sub> é aceita: não há diferença entre as amostras de presença e pseudo-ausência.	
	Teste de Levene	Estatística de Levene = 1,38 p-value = 0,24 H <sub>0</sub> é aceita: não há diferença entre as amostras de presença e pseudo-ausência..	
	Temperatura Mínima	média	656,0
desvio padrão		112,5	104,0
Teste t de Welch		Estatística t = - 6,90	

Variável	Estatísticas	Presença	Pseudo-ausência
Precipitação		<p>p-value = 8,5e.12</p> <p>H<sub>0</sub> é rejeitada: há diferença entre as amostras de presença e pseudo-ausência.</p>	
	Teste de Levene	<p>Estatística de Levene = 5,15</p> <p>p-value = 0,02</p> <p>H<sub>0</sub> é rejeitada: há diferença entre as amostras de presença e pseudo-ausência.</p>	
	média	146,6	103,4
	desvio padrão	126,5	63,2
	Teste t de Welch	<p>Estatística t = -7,46</p> <p>p-value = 1,71e-13</p> <p>H<sub>0</sub> é rejeitada: há diferença entre as amostras de presença e pseudo-ausência.</p>	
	Teste de Levene	<p>Estatística de Levene = 31,09</p> <p>p-value = 3,05e-08</p> <p>H<sub>0</sub> é rejeitada: há diferença entre as amostras de presença e pseudo-ausência.</p>	
	média	4,90	4,41
	desvio padrão	10,7	11,4
	Teste t de Welch	<p>Estatística t = -0,76</p> <p>p-value = 0,45</p> <p>H<sub>0</sub> é aceita: não há diferença entre as amostras de presença e pseudo-ausência.</p>	
	Teste de Levene	<p>Estatística de Levene = 0,40</p> <p>p-value = 0,53</p> <p>H<sub>0</sub> é aceita: não há diferença entre as amostras de presença e pseudo-ausência.</p>	
Altitude	média	243,2	267,0
	desvio padrão	99,6	89,4



Variável	Estatísticas	Presença	Pseudo-ausência
	Teste t de Welch	Estatística t = 4,88 p-value = 1,23e-06 H <sub>0</sub> é rejeitada: há diferença entre as amostras de presença e pseudo-ausência.	
	Teste de Levene	Estatística de Levene = 4,20 p-value = 0,04 H <sub>0</sub> é rejeitada: há diferença entre as amostras de presença e pseudo-ausência.	
	média	93,3	74,8
	desvio padrão	87,7	69,1
Inclinação do Terreno	Teste t de Welch	Estatística t = -4,02 p-value = 6,08e-05 H <sub>0</sub> é rejeitada: há diferença entre as amostras de presença e pseudo-ausência.	
	Teste de Levene	Estatística de Levene = 21,32 p-value = 4.32e-06 H <sub>0</sub> é rejeitada: há diferença entre as amostras de presença e pseudo-ausência.	

O diagrama BAM (SOBERÓN; NAKAMURA, 2009) mostra que pode haver uma parte do Nicho Fundamental onde estão as verdadeiras ausências da espécie, ou seja, onde embora as condições ambientais permitam a ocorrência dessas espécies, não há ocupação por elas. Essa sobreposição é representada nas Figuras 93 a 99. Contudo, para as variáveis numéricas, a média e a variância nos pontos de Presença e de Pseudo-ausência são distintas, exceto Temperatura Máxima e Luzes Noturnas Artificiais.

Ainda que possa ocorrer maior esforço amostral ao longo de estradas, rodovias e outras áreas antropizadas que são mais acessíveis e, portanto, as observações serem mais frequentes nesses locais. Isto significa que, no caso da variável Luzes Artificiais Noturnas, não houve viés de amostragem por conta da maior facilidade de observações das espécies.

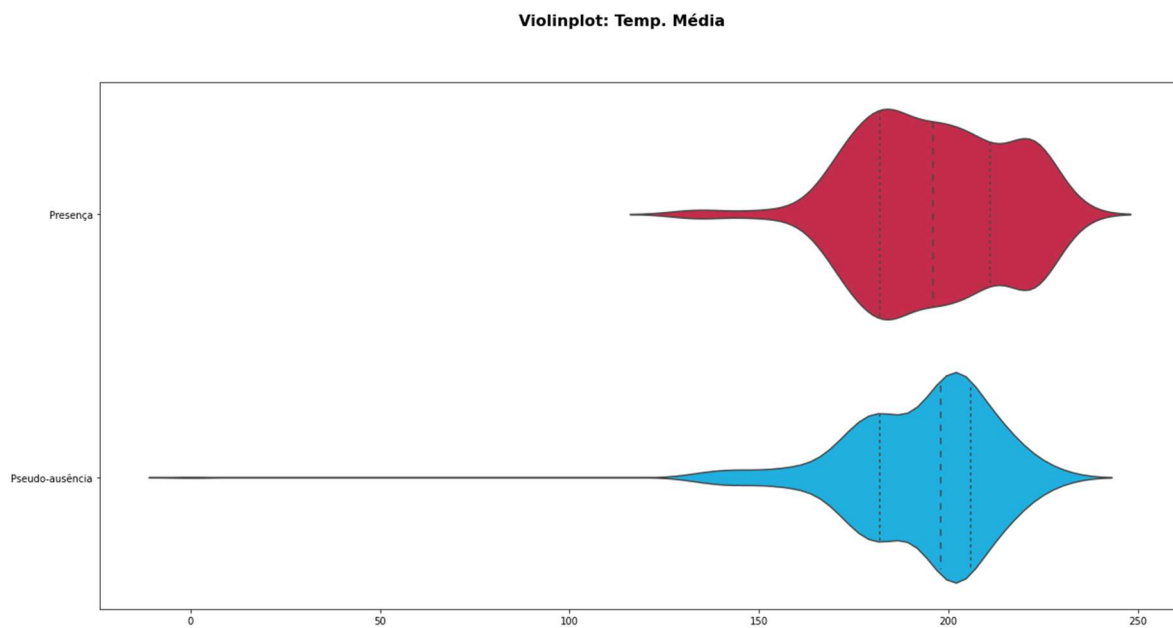
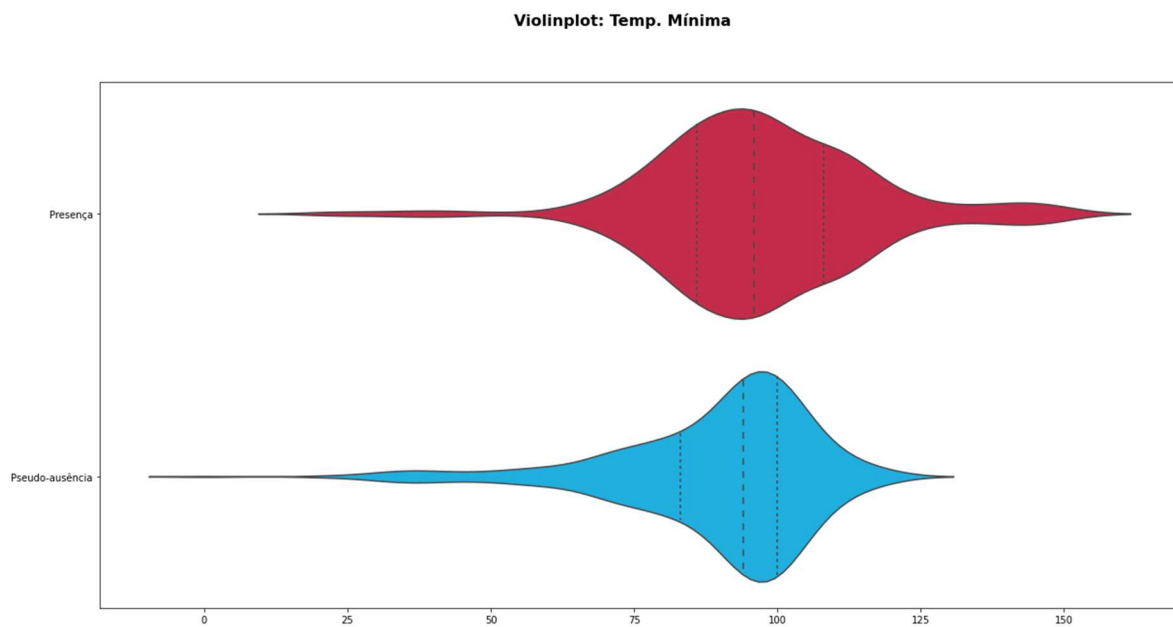
Figura 93 - Valores médios de Temperatura Média ( $\times 10^{\circ}\text{C}$ ) dos pontos amostrais de Presença e de Pseudo-ausênciaFigura 94 - Valores médios de Temperatura Mínima ( $\times 10^{\circ}\text{C}$ ) dos pontos amostrais de Presença e de Pseudo-ausência

Figura 95 - Valores médios de Temperatura Máxima (x10°C) dos pontos amostrais de Presença e de Pseudo-ausência

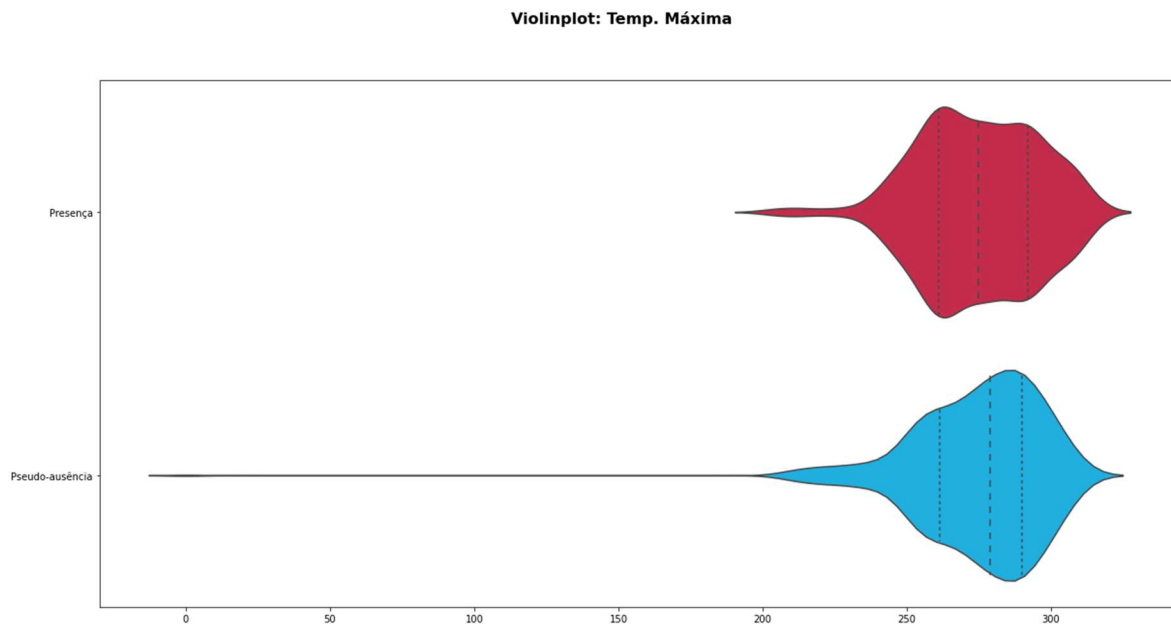


Figura 96 - Valores médios de Precipitação (mm) dos pontos amostrais de Presença e de Pseudo-ausência

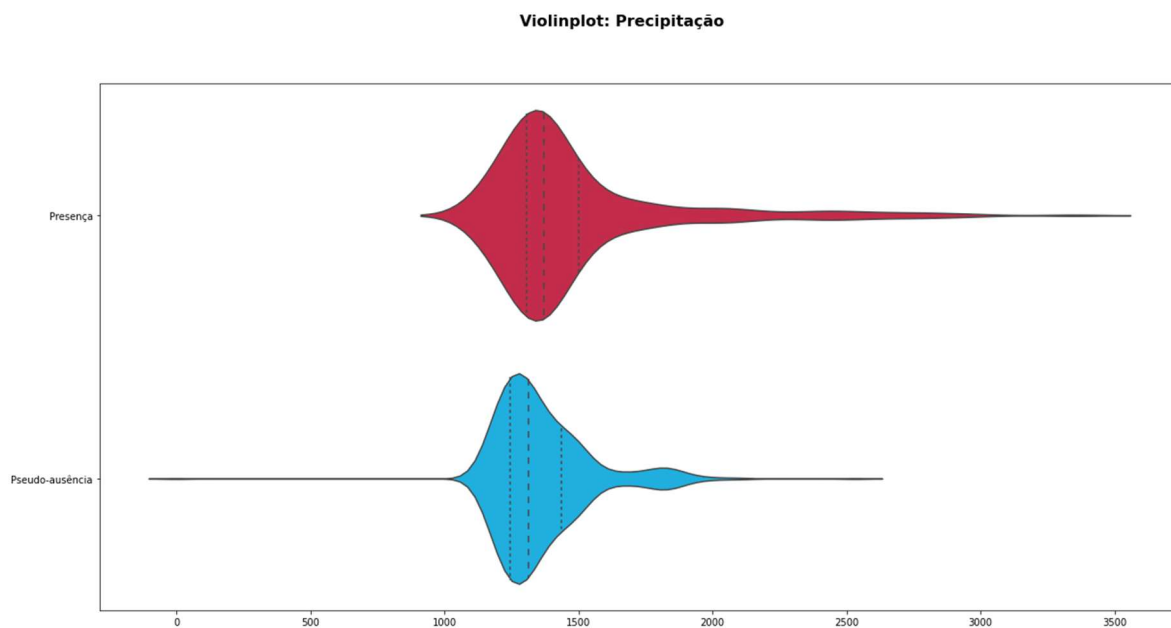


Figura 97 - Valores médios de Luzes Artificiais Noturnas (nanoWatts/cm2/sr) dos pontos amostrais de Presença e de Pseudo-ausência

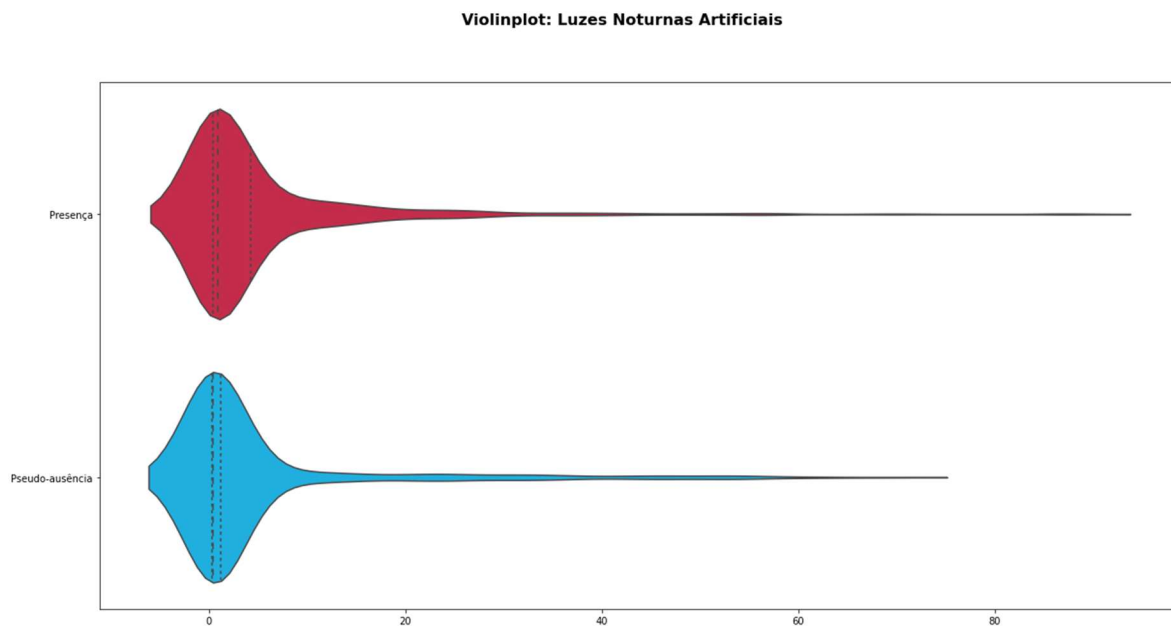


Figura 98 - Valores médios de Altitude (metros) dos pontos amostrais de Presença e de Pseudo-ausência

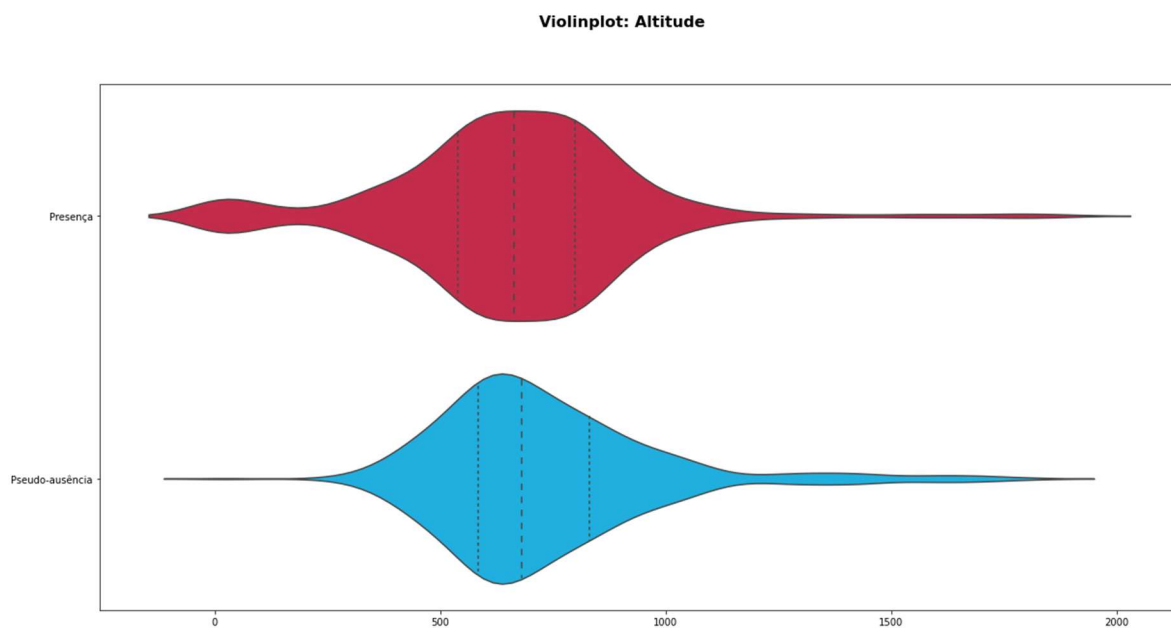
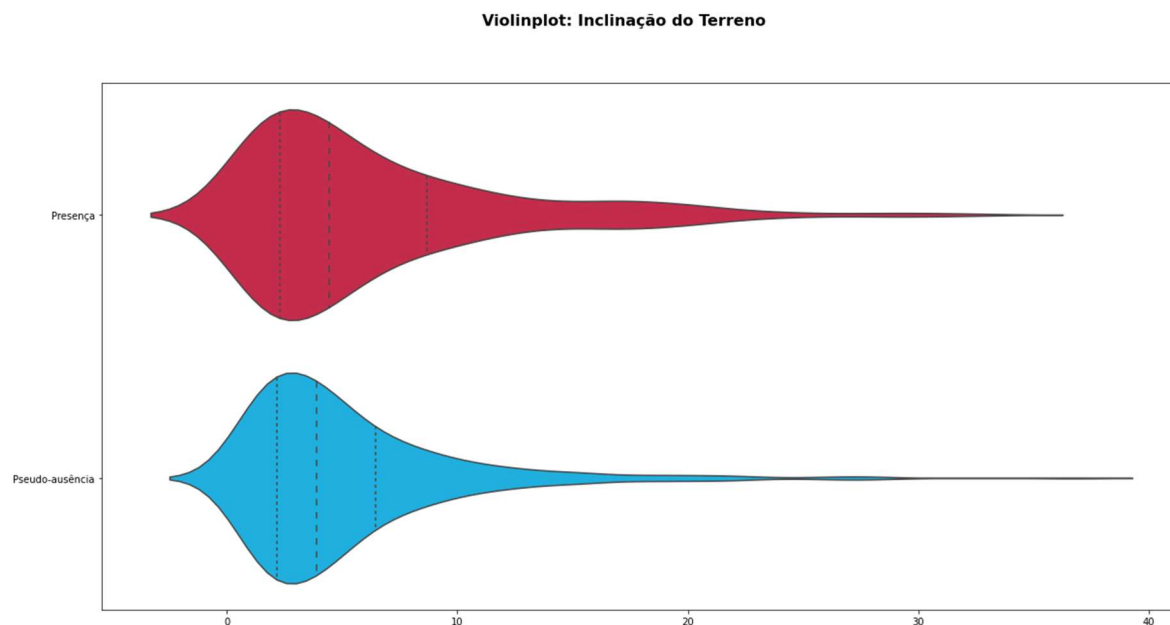


Figura 99 - Valores médios de Inclinação do Terreno (graus) dos pontos amostrais de Presença e de Pseudo-ausência



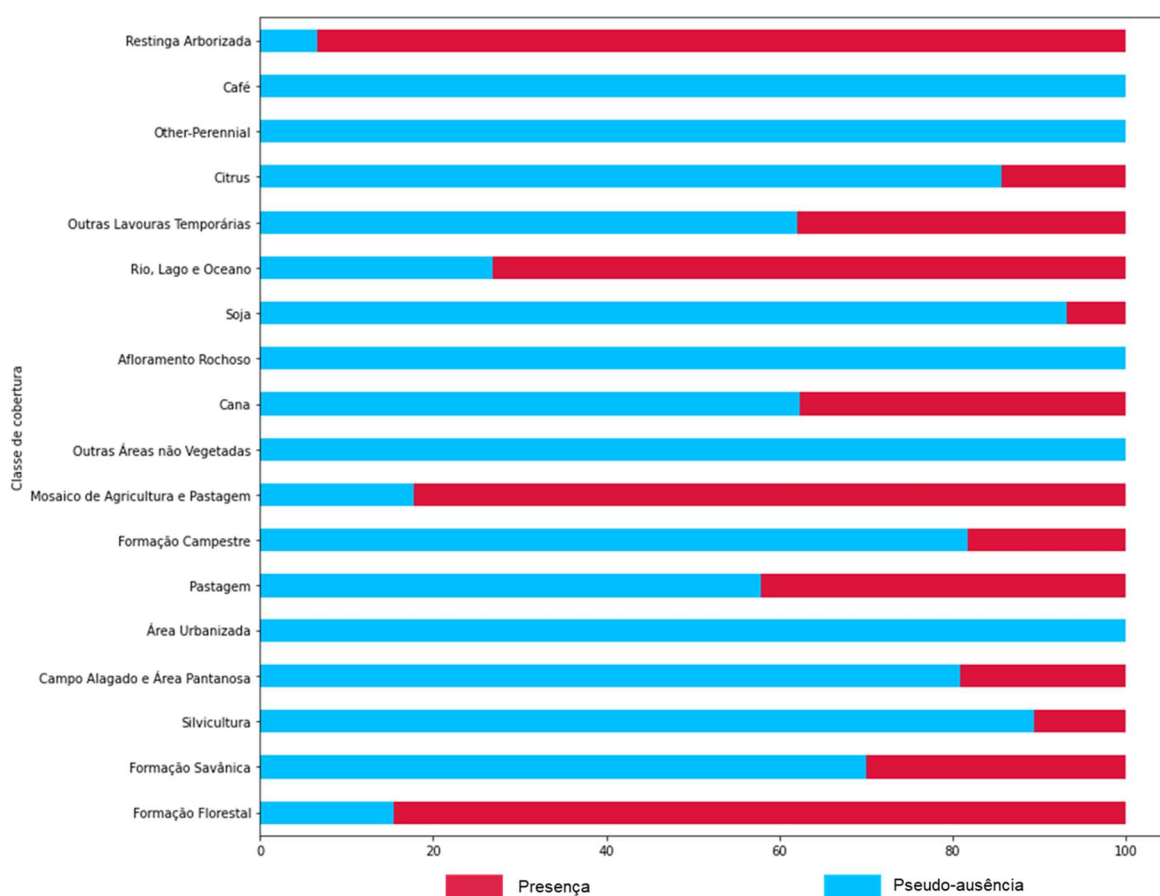
Para as variáveis categóricas, as proporções entre pontos de Presença e de Pseudo-ausência são bem distribuídas; os pontos incidentes nas classes de cobertura da terra Formação Florestal (85,5%), Mosaico de Agricultura e Pastagem (82,2%) e Restinga Arborizada (93,3%) são na maioria da amostra de presença (Figura 100). O alto percentual de amostras de Presença na classe Mosaico de Agricultura e Pastagem indica que “ocorrência” não é “adequabilidade”, especialmente para espécies que se locomovem. Esta classe inclui áreas de cana-de-açúcar mapeadas pós-colheita; entretanto, isso não significa que o indivíduo tenha sido observado nesse período. As espécies de mamíferos observadas em ambientes hostis, como os canaviais, estão sujeitas ao consumo de água e alimento contaminados com agrotóxicos (FERNANDES-SANTOS et al., 2020; ROCHA et al., 2022) e às queimadas (CELERE, 2012; GHELERCOSTA et al., 2013; PHILOGENE; NI-MEISTER, 2021), embora o uso do fogo em canaviais seja proibido por lei no estado de São Paulo<sup>110</sup>. Segundo a análise exploratória dos dados de Presença, as espécies mais frequentes nessa classe são *Callithrix penicillata* (11%) e *Sciurus aestuans* (11%). Isto acontece porque existe uma dinâmica espaço temporal da cobertura de terra que resulta na fragmentação de habitats e as áreas agrícolas continuam sendo zonas de trânsito dessas espécies. A conversão da cobertura da terra altera o microclima, mas esta variação pode não ser detectada na escala em que os dados climáticos são adquiridos. As

<sup>110</sup> Lei nº 11.241, de 19 de setembro de 2002. Dispõe sobre a eliminação gradativa da queima da palha da cana-de-açúcar e dá providências correlatas. Disponível em <<https://www.al.sp.gov.br/repositorio/legislacao/lei/2002/lei-11241-19.09.2002.html>>. Acesso em 23 jan 2023.

variáveis relacionadas ao relevo (Altitude e Inclinação do Terreno) não são alteradas por essas coberturas.

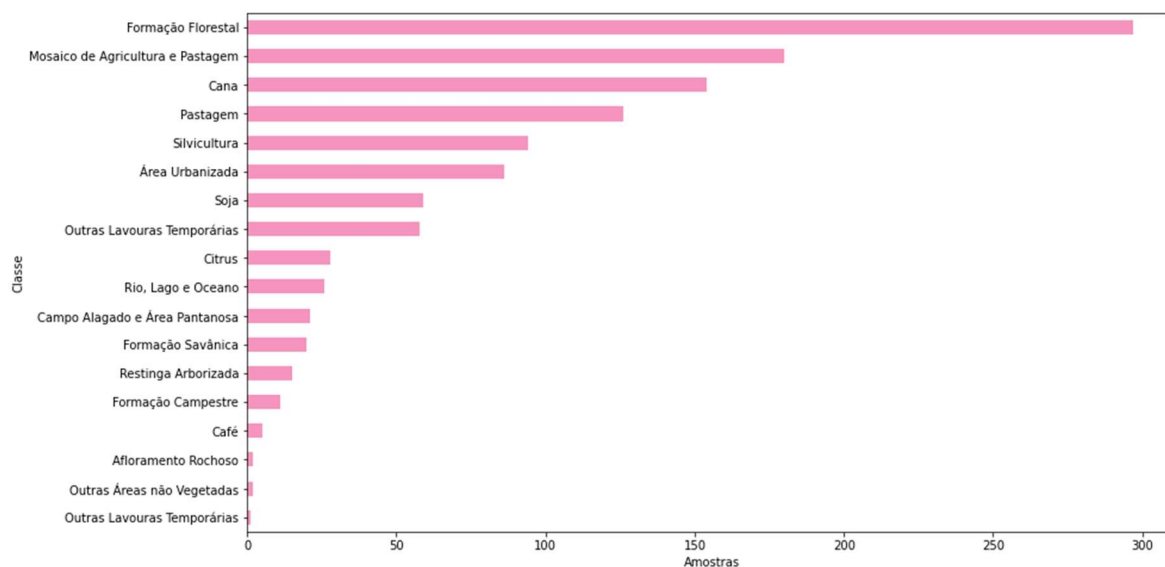
As amostras de Pseudo-ausência, por sua vez, são frequentes em áreas agrossilvopastoris (Café, Soja, Citrus, Pastagem, Silvicultura, Outras Lavouras Perenes). A probabilidade de um ponto aleatório de Pseudo-ausência incidir nessas classes é proporcional à sua extensão no plano de fundo de Pseudo-ausência.

Figura 100 - Proporção de pontos amostrais de Presença e de Pseudo-ausência nas variáveis categóricas (Cobertura da terra)



A Figura 101 mostra o histograma do tamanho absoluto das amostras por cobertura da terra agrupando as amostras de Presença e de Pseudo-ausência. As classes mais frequentes são Formação Florestal (25%), Mosaico de Agricultura e Pastagem (15%) e Cana (13%).

Figura 101 - Quantidade absoluta de pontos amostrais por classe de cobertura da terra



Caso os pontos de Presença e de Pseudo-ausência estejam mal distribuídos dentro e fora de áreas protegidas, os modelos podem ser enviesados. Por exemplo, se houver muitos pontos de presença no interior das UCs de Proteção Integral, os resultados dos modelos podem apontar que estas unidades têm alta adequabilidade ambiental apenas em decorrência da amostragem desbalanceada. A Tabela 28 mostra a quantidade de pontos amostrais de Presença e Pseudo-ausência para cada categoria de UC. No entanto, como existe sobreposição entre UCs de diferentes esferas e/ou categorias, alguns pontos correspondem a mais de uma unidade e estão duplicados na tabela. Considerando esta sobreposição, apenas 204 pontos de Presença (34%) e 117 pontos de Pseudo-ausência (20%) incidem em UCs.

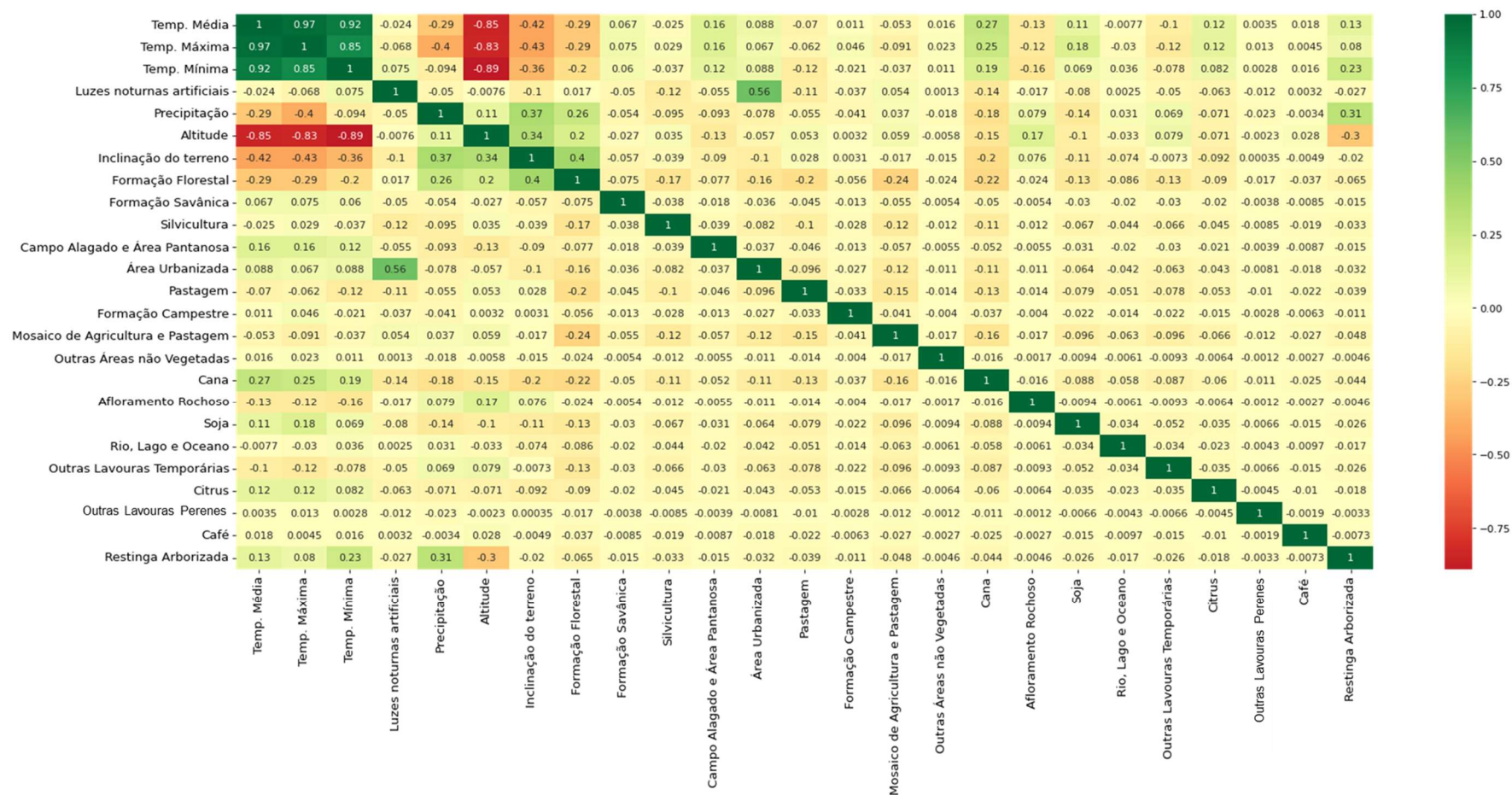
O correlograma de variáveis também foi construído para os pontos de pseudo-ausência (Figura 102) e apresenta a mesma condição dos pontos de presença, com alta correlação entre as variáveis relacionadas à Temperatura, e entre estas e a Altitude. As variáveis com maior coeficiente de correlação entre as demais, Temperatura Máxima, Temperatura Média e Altitude, não foram, portanto, incluídas na alimentação dos modelos de regressão; além destas, as classes de cobertura da terra Mangue, Outras Formações não Florestais, Praia, Duna e Areal, Mineração, Aquicultura e Apicum não têm pontos amostrais e foram removidas dos modelos.

Tabela 28 - Amostras de Presença e de Pseudo-ausência nas UCs

<b>Uso</b>	<b>Categoria</b>	<b>Pontos de Presença</b>	<b>Pontos de Pseudo- ausência</b>
Proteção Integral	ESEC	5	2
	PARNA	1	6
	PE	79	11
Uso Sustentável	APA	133	126
	ARIE	3	0
	FE	3	1
	FLONA	2	0
	RDS	1	0
Total		204 (34%)	117 (20%)



Figura 102 - Correlograma das variáveis para as amostras de Pseudo-ausência



## 4.7 MODELOS DE ADEQUABILIDADE DE HABITAT

### 4.7.1 Árvore de Decisão

A Tabela 29 mostra os valores dos hiperparâmetros, com o valor padrão e os sugeridos pelo GridSearchCV, e respectivos valores de  $R^2$ . A descrição dos hiperparâmetros pode ser acessada na documentação da ferramenta<sup>111</sup>.

Tabela 29 - Hiperparâmetros do modelo Árvore de Decisão

	<b>Hiperparâmetros padrão</b>	<b>Hiperparâmetros ajustados pelo GridSearchCV</b>
critério	squared error	squared error
max_features	None	auto
splitter	best	best
min_samples_leaf	1	1
min_samples_split	2	2
$R^2$	-0,33	-0,30

Os modelos de Árvore de Decisão dispõem um atributo que ordena a importância das variáveis no modelo. A Figura 103 mostra as variáveis ordenadas conforme a variância explicada que cada variável contribui para o modelo com os parâmetros ajustados. A importância relativa é a importância percentual em relação à importância da variável mais importante. As variáveis mais importantes na Árvore de Decisão e respectivas importâncias absolutas no modelo estão na Tabela 30.

<sup>111</sup> sklearn.tree.DecisionTreeRegressor. Disponível em <<https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeRegressor.html>>. Acesso em 31 out. 2022.

Figura 103 - Importância relativa das variáveis no modelo Árvore de Decisão

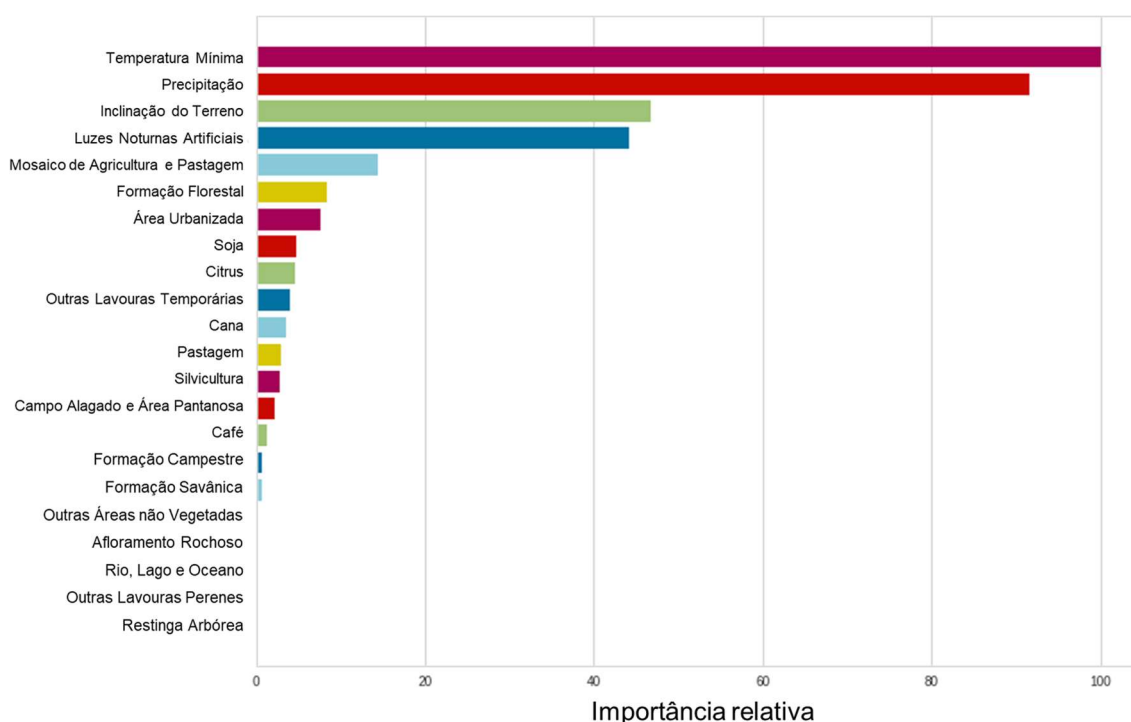


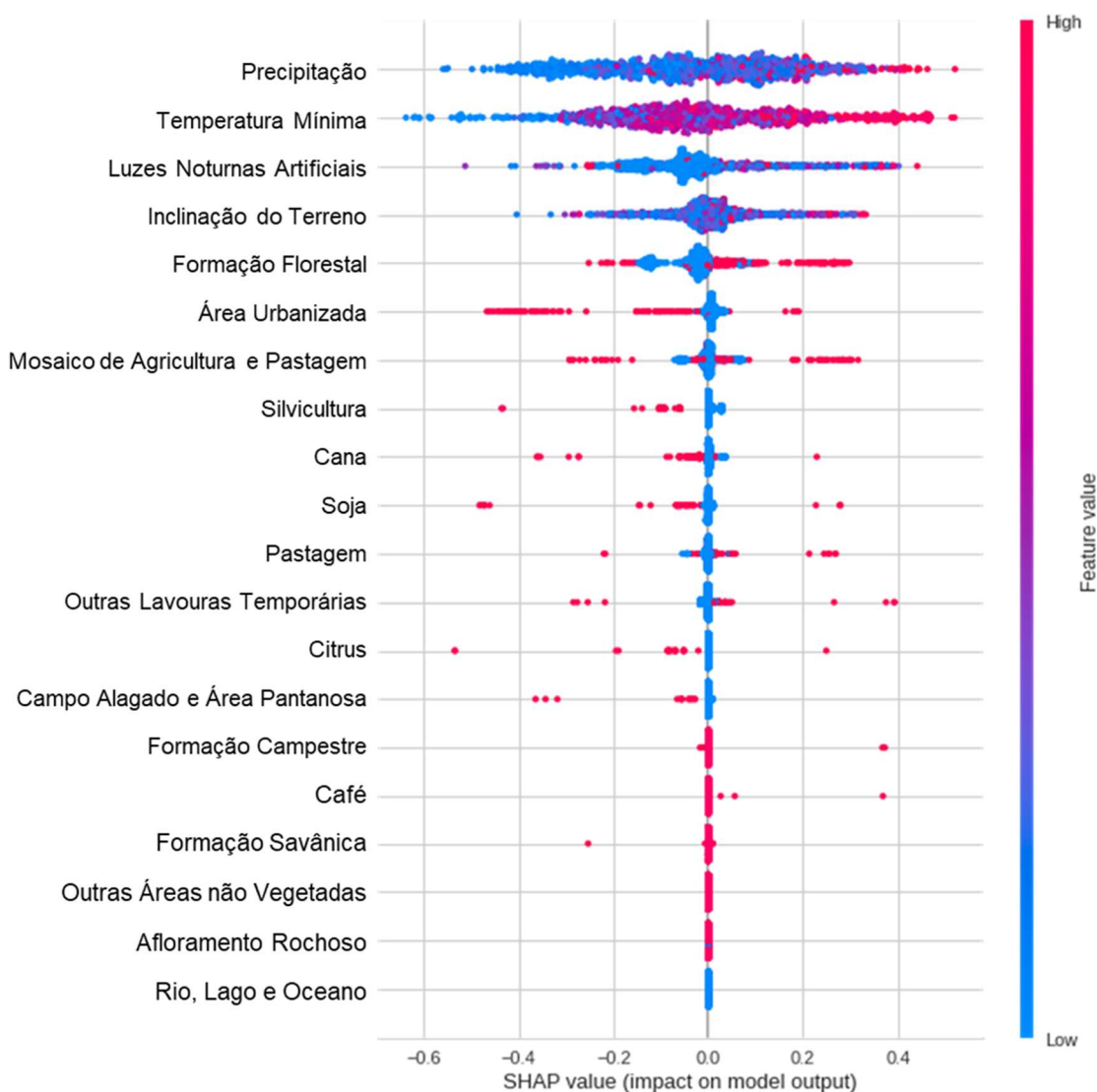
Tabela 30 - Importância absoluta das variáveis no modelo Árvore de Decisão

Variável	Importância absoluta
Temperatura Mínima	0,294
Precipitação	0,269
Inclinação do Terreno	0,137
Luzes Noturnas Artificiais	0,130
Mosaico de Agricultura e Pastagem	0,043

Neste modelo, as variáveis mais importantes são numéricas; a primeira variável categórica (Mosaico de Agricultura e Pastagem) é a segunda mais frequente nas amostras de Presença e Pseudo-ausência. O diagrama dos valores SHAP (*SHapley Additive exPlanations*) (SHAPLEY, 1953; ROTH, 1988) na Figura 104 diferencia a importância das variáveis para definir Presença e Pseudo-ausência. Os pontos no diagrama SHAP correspondem às amostras em cada variável (cada ponto no gráfico é um ponto de Presença ou Pseudo-ausência) e representa a contribuição desse ponto no resultado da predição. Os pontos rosas indicam as amostras de Presença e os azuis de Pseudo-Ausência. Valores de SHAP próximos a zero

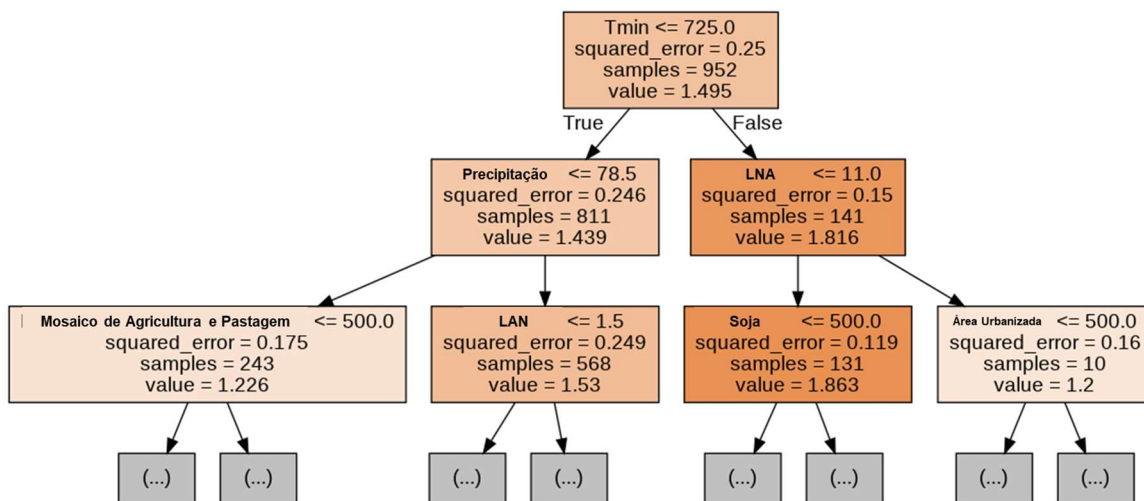
indicam pouca contribuição do ponto para a predição; valores altos em módulo (positivos ou negativos) indicam maior contribuição do ponto para prever a classe positiva (Presença) ou negativa (Pseudo-ausência). Os pontos incidentes nas classes de cobertura Formação Campestre, Café, Formação Savânica, Outras Áreas não Vegetadas, Afloramento Rochoso e Rio, Lago e Oceano não contribuem para a previsão; estas classes são algumas das menos frequentes na amostra de pontos de Presença e Pseudo-ausência.

Figura 104 - Valores SHAP para a Árvore de Regressão



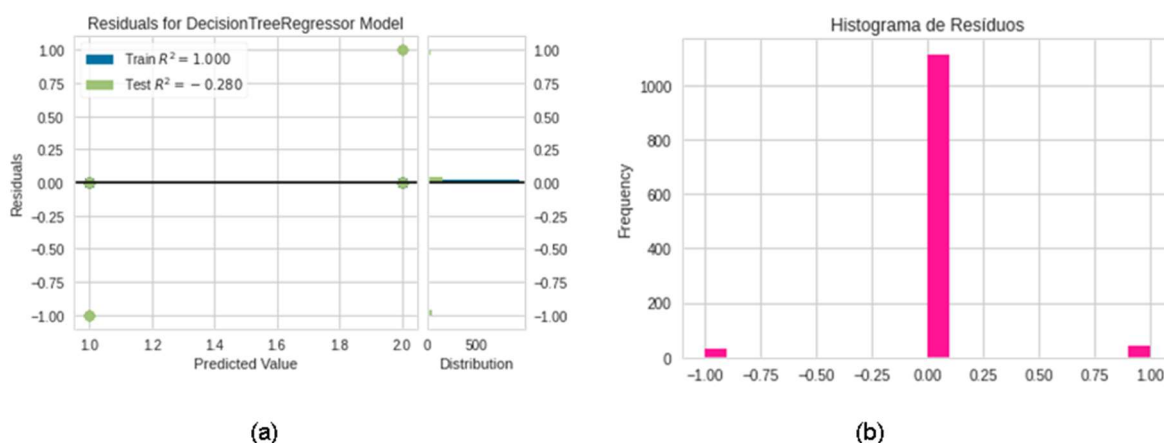
Nas Árvores de Decisão, a geração das folhas é baseada na variância entre elas; a importância da variável está relacionada a quanto essa variável contribui para melhorar a discriminação das amostras nas folhas, ou seja, maximiza a variância entre as folhas e diminui dentro das folhas. A Figura 105 mostra a raiz e os primeiros nós da Árvore de Decisão.

Figura 105 - Estrutura da Árvore de Decisão



A Figura 106 mostra a) a relação entre os valores preditos e os respectivos resíduos, e b) a distribuição de frequência dos resíduos. Os valores dos resíduos não são distribuídos aleatoriamente, mas existe um padrão. No caso da Árvore de Decisão, o modelo gerou valores discretizados na regressão e essa condição pode ter implicado na heterocedasticidade dos resíduos e no baixo valor de  $R^2$  (-0,381) no teste do modelo; porém, quase 100% dos resíduos do modelo é igual a zero. No teste de Breusch-Pagan, a hipótese nula de que os resíduos são homocedásticos e, portanto, de que não há distorções na predição, foi rejeitada (Estatística Breusch-Pagan = 698,55; p-value = 0,0).

Figura 106 - Distribuição dos resíduos do modelo Árvore de Decisão



### 4.7.2 Floresta Aleatória

Os hiperparâmetros, com o valor padrão e os sugeridos pelo GridSearchCV, e respectivos valores de  $R^2$  são apresentados na Tabela 31. A descrição dos hiperparâmetros pode ser acessada na documentação da ferramenta<sup>112</sup>.

Tabela 31 - Hiperparâmetros do modelo Floresta Aleatória

	<b>Hiperparâmetros padrão</b>	<b>Hiperparâmetros ajustados pelo GridSearchCV</b>
critério	squared error	squared error
max_features	None	None
min_samples_leaf	1	1
min_samples_split	2	2
n_estimators	100	20
$R^2$	0,28	0,25

No modelo Floresta Aleatória, o ajuste de hiperparâmetros diminuiu o  $R^2$ . Por tratar-se de um modelo de natureza randômica, esta diferença pode não ser necessariamente uma diminuição na robustez do modelo, mas uma oscilação aleatória.

Para o modelo com os valores padrão de hiperparâmetros, a ordem de importância relativa das variáveis é semelhante à da Árvore de Decisão (Figura 107). As variáveis com maior importância relativa são Temperatura Mínima (0,291), Precipitação (0,248), Inclinação do Terreno (0,168), Luzes Noturnas Artificiais (0,122) e Formação Florestal (0,040). Temperatura e Precipitação também são as variáveis mais importantes de acordo com o diagrama SHAP (Figura 108).

Pelo teste de Breusch-Pagan, os resíduos são heterocedásticos e, portanto, há distorções na predição (Estatística Breusch-Pagan = 641,46; p-value = 0,0). Os resíduos não são aleatórios e apresentam um padrão (Figura 109a) mas o histograma de resíduos aparenta ter distribuição normal (Figura 109b).

<sup>112</sup> sklearn.ensemble.RandomForestRegressor. Disponível em <<https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html>>. Acesso em 01 nov. 2022.

Figura 107 - Importância relativa das variáveis no modelo Floresta Aleatória

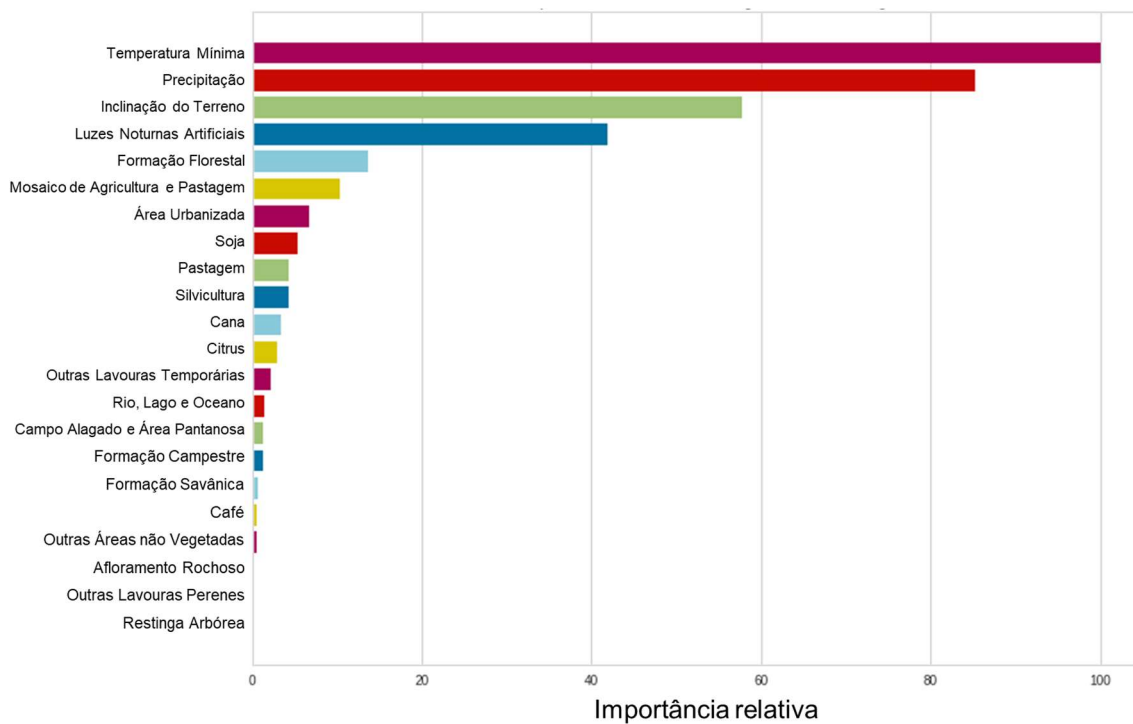


Figura 108 - Valores SHAP para a Floresta Aleatória

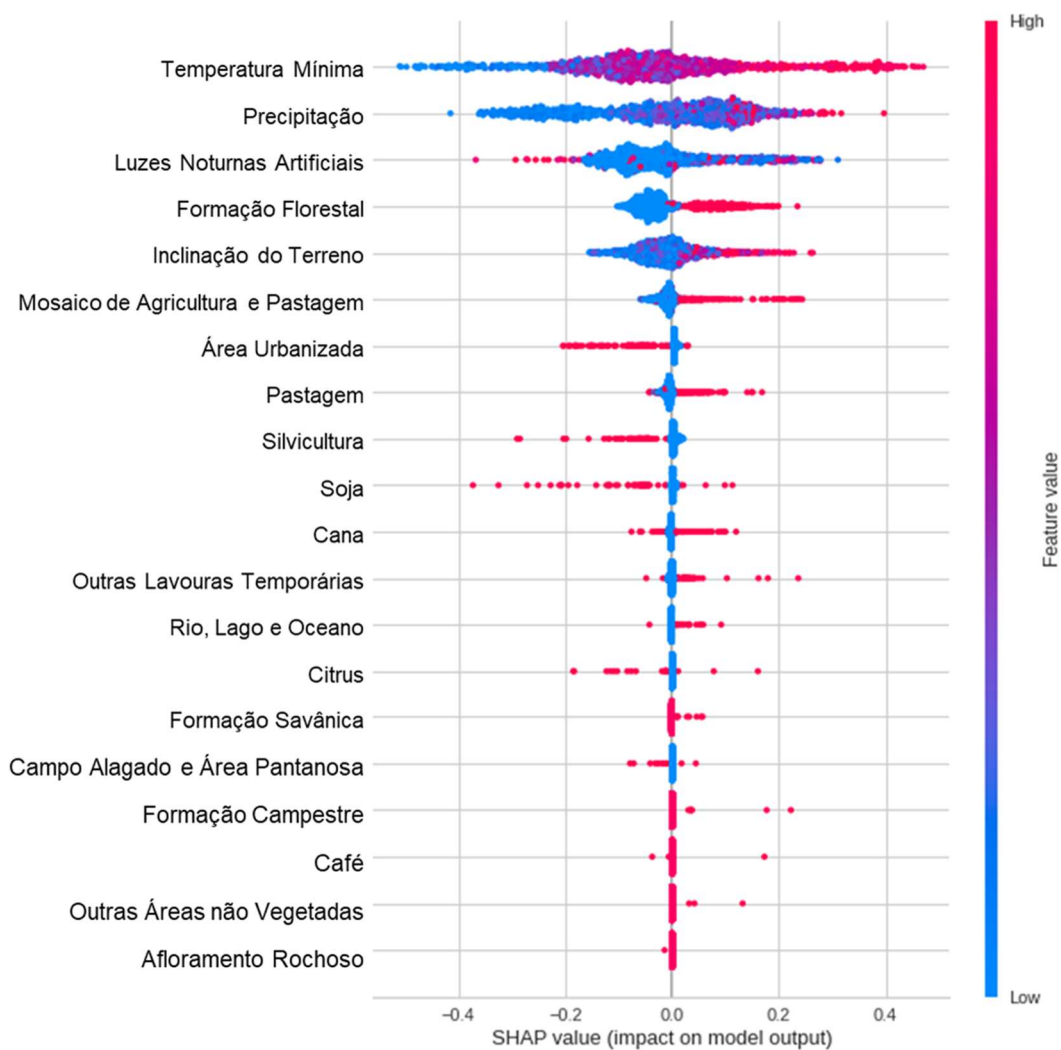
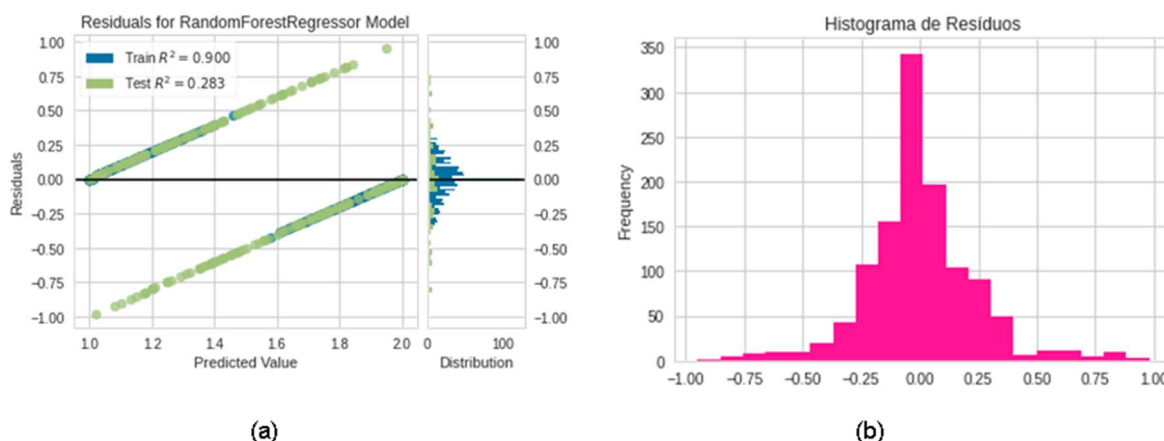




Figura 109 - Distribuição dos resíduos do modelo Floresta Aleatória



#### 4.7.3 KNN

Os hiperparâmetros do KNN, com o valor padrão e os sugeridos pelo GridSearchCV e respectivos valores de  $R^2$  são apresentados na Tabela 32. A descrição dos hiperparâmetros pode ser acessada na documentação da ferramenta<sup>113</sup>. No caso do KNN, a otimização de hiperparâmetros manteve os valores padrão.

No teste Breusch-Pagan, a hipótese nula foi rejeitada, indicando que os resíduos são heterocedásticos e as previsões têm distorções (Estatística Breusch-Pagan = 509,91; p-value = 0,0). Assim como na Árvore de Decisão, os resíduos apresentam um padrão discretizado e não aleatório (Figura 110a). Entretanto, o histograma de resíduos tem certa aparência de distribuição normal (Figura 110b).

O diagrama de valores SHAP (Figura 111) mostra que as variáveis de maior importância absoluta são Formação Florestal, Temperatura Mínima, Mosaico de Agricultura e Pastagem. Precipitação e Inclinação do Terreno, sendo a primeira e a terceira as que melhoram a separabilidade de Presença e Pseudo-ausência, pois tem grande dispersão no eixo horizontal do diagrama e pouca sobreposição de pontos. A Temperatura Mínima tem alto impacto, embora nos baixos valores de SHAP haja sobreposição de pontos de Presença e Ausência.

<sup>113</sup> sklearn.neighbors.KNeighborsRegressor. Disponível em <<https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsRegressor.html>>. Acesso em 01 nov. 2022.

Tabela 32 - Hiperparâmetros do modelo KNN

	Hiperparâmetros padrão	Hiperparâmetros ajustados pelo GridSearchCV
algorithm	auto	auto
metric	minkowski	euclidean
n_neighbors	5	5
weights	uniform	uniform
$R^2$	0,22	0,22

Figura 110 - Distribuição dos resíduos do modelo KNN

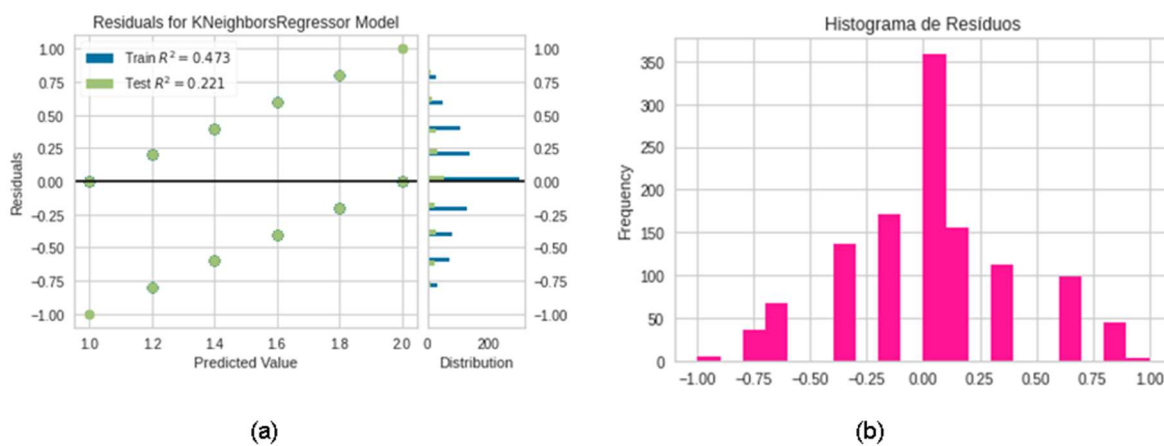
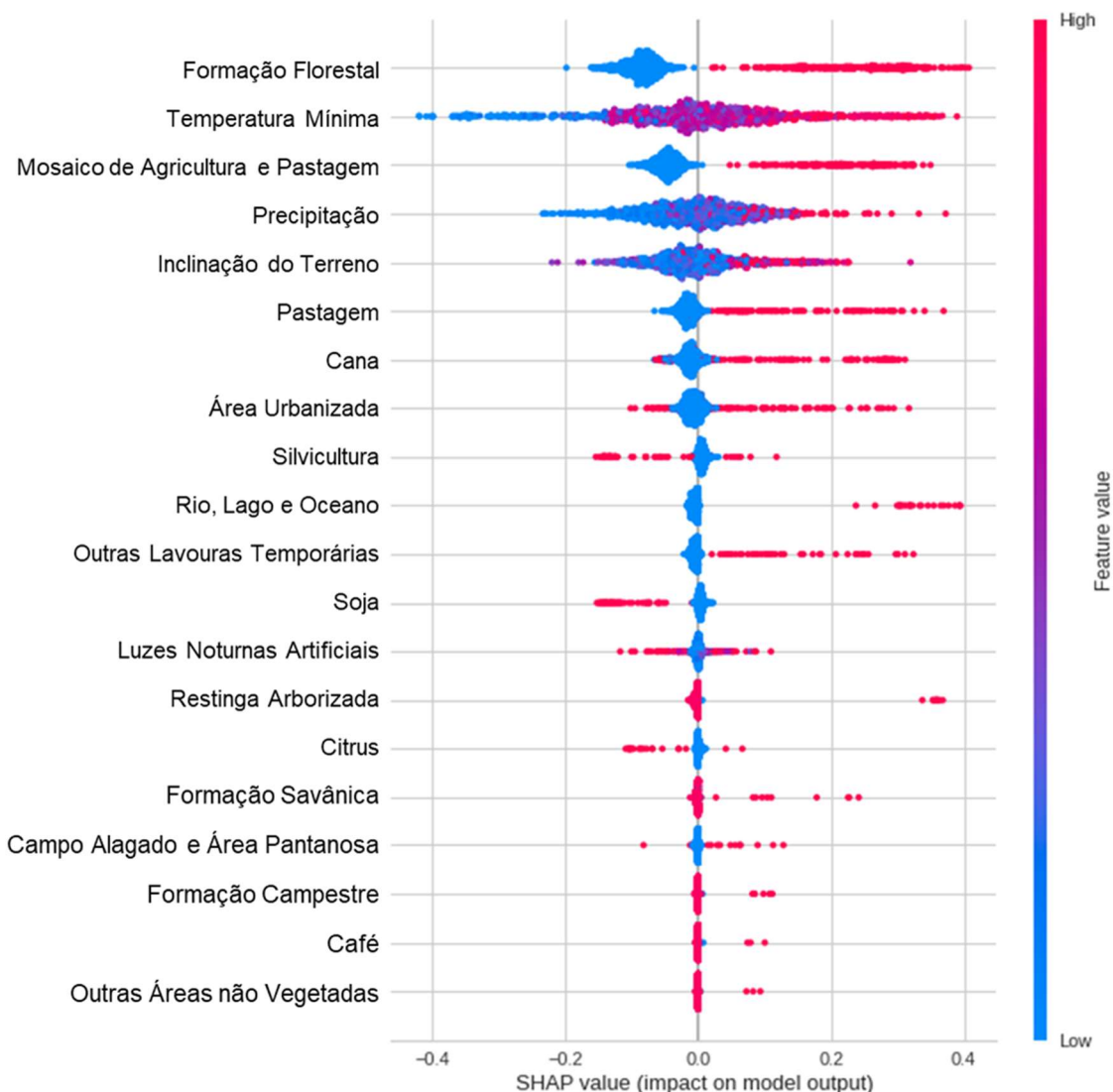


Figura 111 - Valores SHAP para o modelo KNN



#### 4.7.4 SVR

Os hiperparâmetros e valor de  $R^2$  do modelo SVR estão na Tabela 33. A descrição dos hiperparâmetros do SVR pode ser acessada na documentação<sup>114</sup>. O tempo de processamento de busca dos melhores valores para os hiperparâmetros do SVR excedeu 20 mil segundos e, portanto, esta etapa do SVR foi eliminada.

<sup>114</sup> sklearn.svm.SVR. Disponível em <<https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html>>. Acesso em 02 nov 2022.

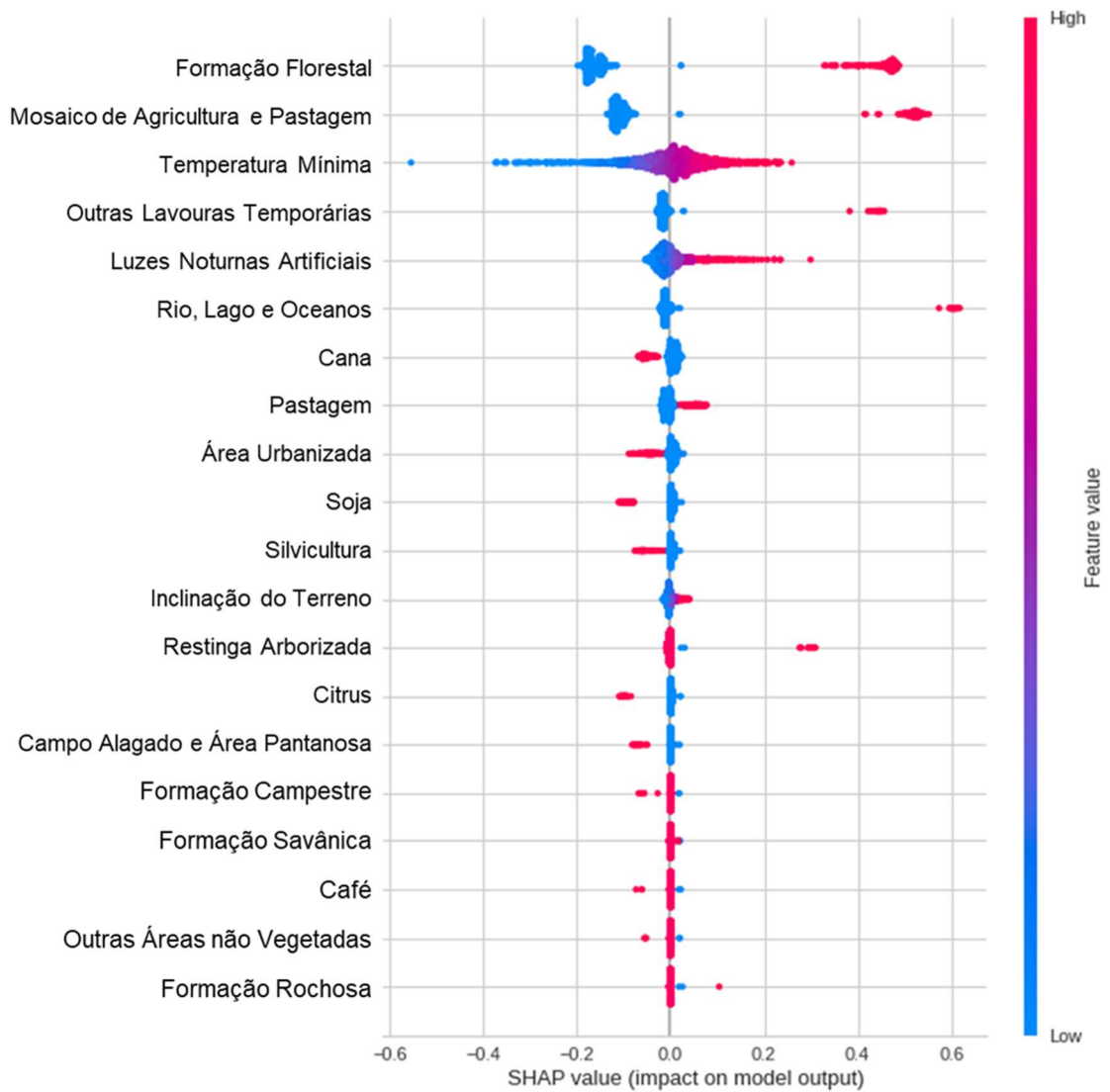
Tabela 33 - Hiperparâmetros do modelo SVR

	Hiperparâmetros padrão	Hiperparâmetros ajustados pelo GridSearchCV
kernel	rbf	
C	1,0	
gamma	scale	Não processado
epsilon	0,1	
R <sup>2</sup>	-0,007	

O SVR não tem o atributo para ordenação da importância de variáveis, mas dispõe do método `.coef_` que elenca os pesos atribuídos às variáveis quando o hiperparâmetro kernel assume o valor “linear”. O SVR foi processado com este valor de kernel e retornou R<sup>2</sup> de -0,089, ainda muito baixo mas maior do que o calculado para os hiperparâmetros padrão. Quando aplicado para toda a área de estudo, o modelo retornou apenas pixels de Presença. Sendo assim, o modelo configurado com kernel linear foi desprezado.

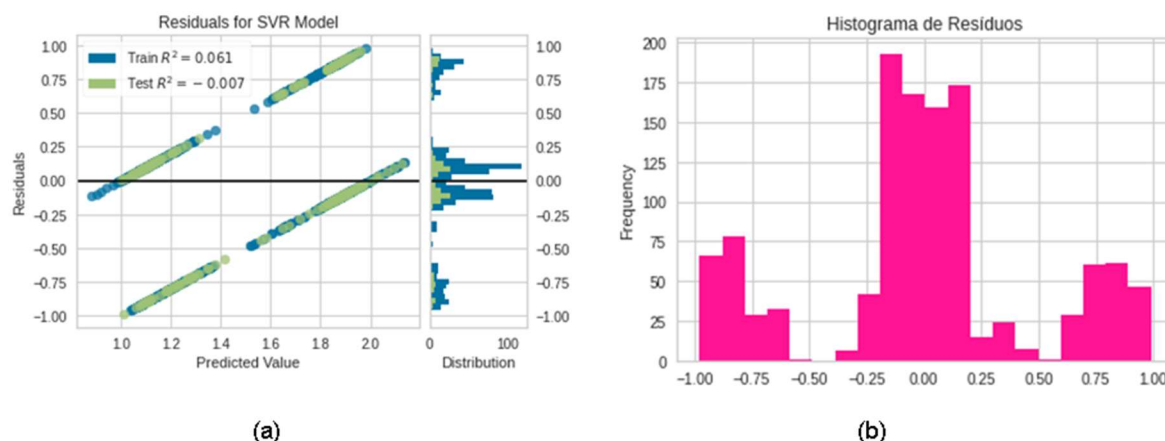
Para o modelo com os hiperparâmetros padrão, as classes de cobertura da terra Formação Florestal e Mosaico de Agricultura e Pastagem, além de terem altos valores absolutos de SHAP em relação às demais variáveis, discriminam bem os pontos de Presença e Pseudo-ausência (Figura 112).

Figura 112 - Valores SHAP para o modelo SVR



Assim como nos modelos anteriores, os resíduos do SVR não são aleatórios e apresentam um padrão de distribuição (Figura 113a) e o histograma dos resíduos não se ajusta à curva normal (Figura 113b). O teste Breusch-Pagan apontou a heterocedasticidade dos resíduos, com distorções nas previsões (Estatística Breusch-Pagan = 700,21; p-value = 0,0).

Figura 113 - Distribuição dos resíduos do modelo SVR



#### 4.7.5 XGBoost

No modelo XGBoost, a seleção dos hiperparâmetros pelo GridSearchCV diminuiu o valor de  $R^2$  (Tabela 34) e os hiperparâmetros padrão foram os usados no modelo. A descrição dos hiperparâmetros pode ser acessada na documentação<sup>115</sup>. Para o modelo com os hiperparâmetros padrão, as variáveis com maior importância absoluta foram Formação Florestal (0,207), Mosaico de Agricultura e Pastagem (0,140), Silvicultura (0,098), Luzes Noturnas Artificiais (0,072) e Temperatura Mínima (0,069).

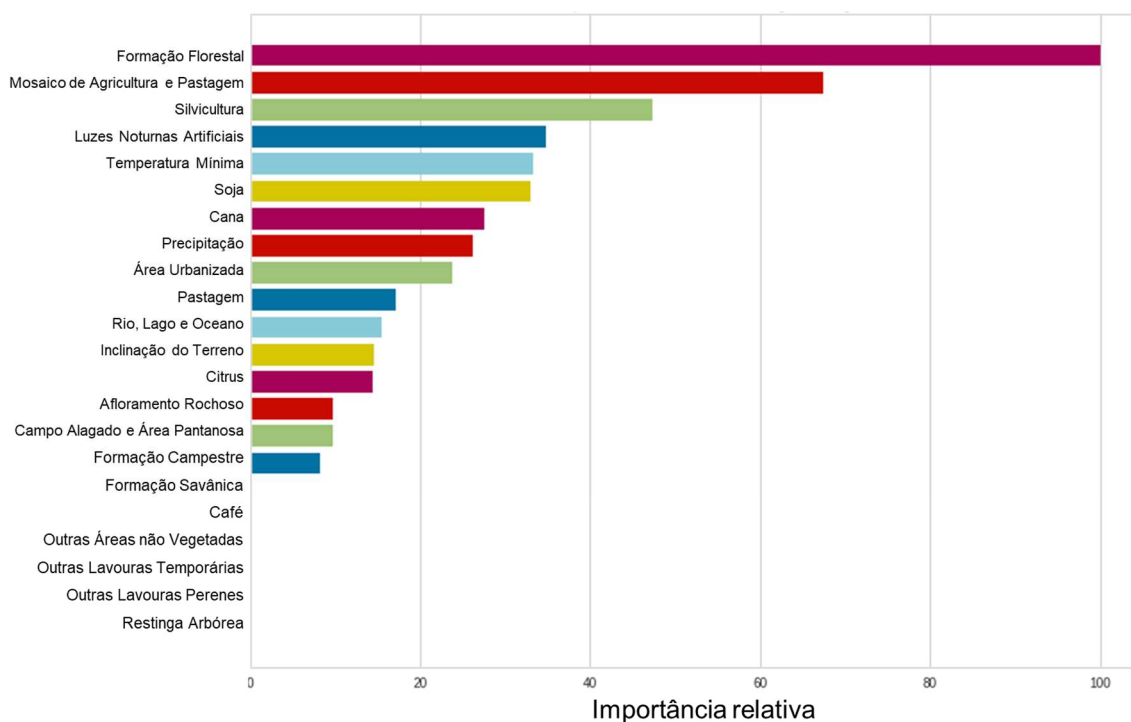
As variáveis com maior importância relativa no modelo XGBoost são Formação Florestal, Mosaico de Agricultura e Pastagem, Silvicultura, Luzes Artificiais Noturnas e Temperatura Mínima (Figura 114).

<sup>115</sup> XGBoost Parameters. Disponível em <<https://xgboost.readthedocs.io/en/stable/parameter.html>>. Acesso em 02 nov. 2022.

Tabela 34 - Hiperparâmetros do modelo XGBoost

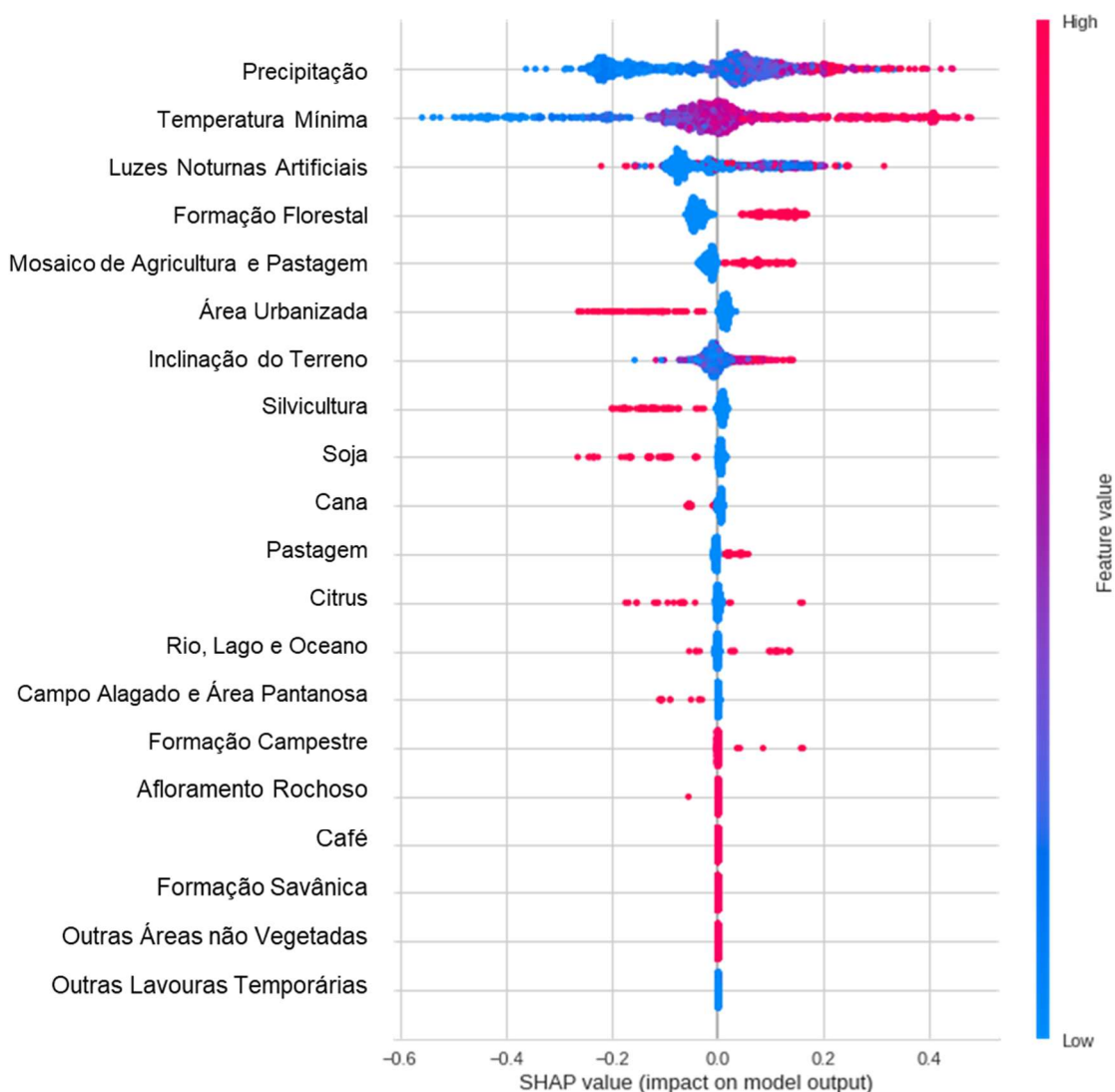
	Hiperparâmetros padrão	Hiperparâmetros ajustados pelo GridSearchCV
objective	reg:squarederror	reg:linear
learning_rate	0,3	0,3
max_depth	6	5
colsample_bytree	1	0,25
min_child_weight	1	1
subsample	1	0,25
R <sup>2</sup>	0,31	-0.19

Figura 114 - Importância relativa das variáveis para o modelo XGBoost



Precipitação e Temperatura Mínima têm altos valores SHAP em relação às demais variáveis, ou seja, têm alto impacto no resultado do modelo, ainda que haja sobreposição de pontos de Presença e Pseudo-ausência no diagrama (Figura 115).

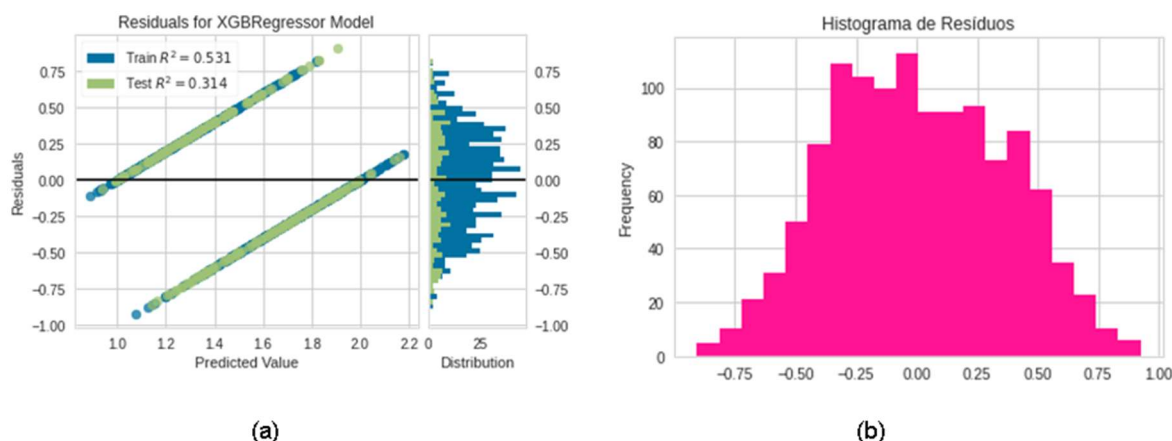
Figura 115 - Valores SHAP para o modelo LightGBM



Os resíduos do XGBoost não são aleatórios e apresentam padrão organizado em relação ao valor predito (Figura 116a). Embora o histograma de resíduos se assemelhe à curva normal (Figura 103b), o teste de Breusch-Pagan rejeitou a hipótese nula, apontando que os resíduos são heterocedásticos e as previsões têm distorções (Estatística Breusch-Pagan = 504,91; p-value = 0,0).



Figura 116 - Distribuição dos resíduos do modelo XGBoost



#### 4.7.6 LightGBM

O ajuste de hiperparâmetros melhorou suavemente o coeficiente de determinação  $R^2$  (Tabela 35). A descrição dos hiperparâmetros pode ser acessada na documentação<sup>116</sup>.

Tabela 35 - Hiperparâmetros do modelo LightGBM

	Hiperparâmetros padrão	Hiperparâmetros ajustados pelo GridSearchCV
num_leaves	31	7
learning_rate	0,1	0,1
max_depth	-1 (sem limite)	-1
n_estimators	100	50
$R^2$	0,26	0,29

Para o modelo com os hiperparâmetros ajustados, os resíduos do LightGBM não são aleatórios (Figura 117a) e o histograma de resíduos tem aparência bimodal (Figura 117b). O teste Breusch-Pagan apontou a heterocedasticidade dos resíduos, com distorções nas previsões (Estatística Breusch-Pagan = 500,97; p-value = 0,0).

As variáveis com maior importância relativa são Temperatura Mínima (95 vezes), Precipitação (70 vezes), Luzes Noturnas Artificiais (42 vezes), Inclinação do Terreno (29 vezes)

<sup>116</sup> lightgbm.LGBMRegressor. Disponível em <<https://lightgbm.readthedocs.io/en/latest/pythonapi/lightgbm.LGBMRegressor.html>>. Acesso em 02 nov. 2022..

e Área Urbanizada (14 vezes) (Figura 118). No diagrama SHAP (Figura 119), a Temperatura Mínima tem grande dispersão, ainda que não discrimine bem Presença e Pseudo-ausência. Formação Florestal, Área Urbanizada e Mosaico de Agricultura e Pastagem são variáveis que caracterizam bem Presença e Pseudo-ausência.

Figura 117 - Distribuição dos resíduos do modelo LightGBM

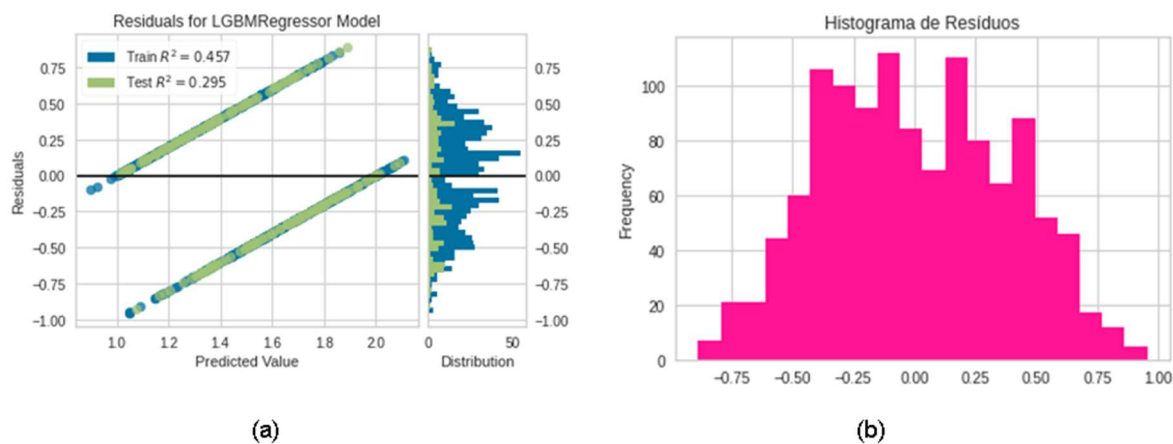


Figura 118 - Importância relativa das variáveis para o modelo LightGBM

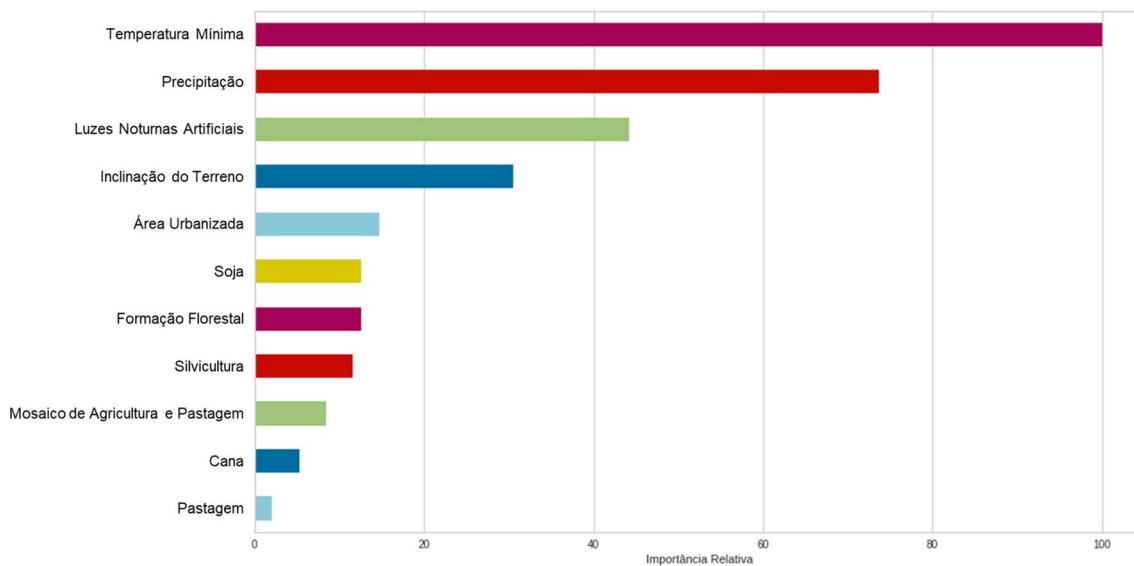
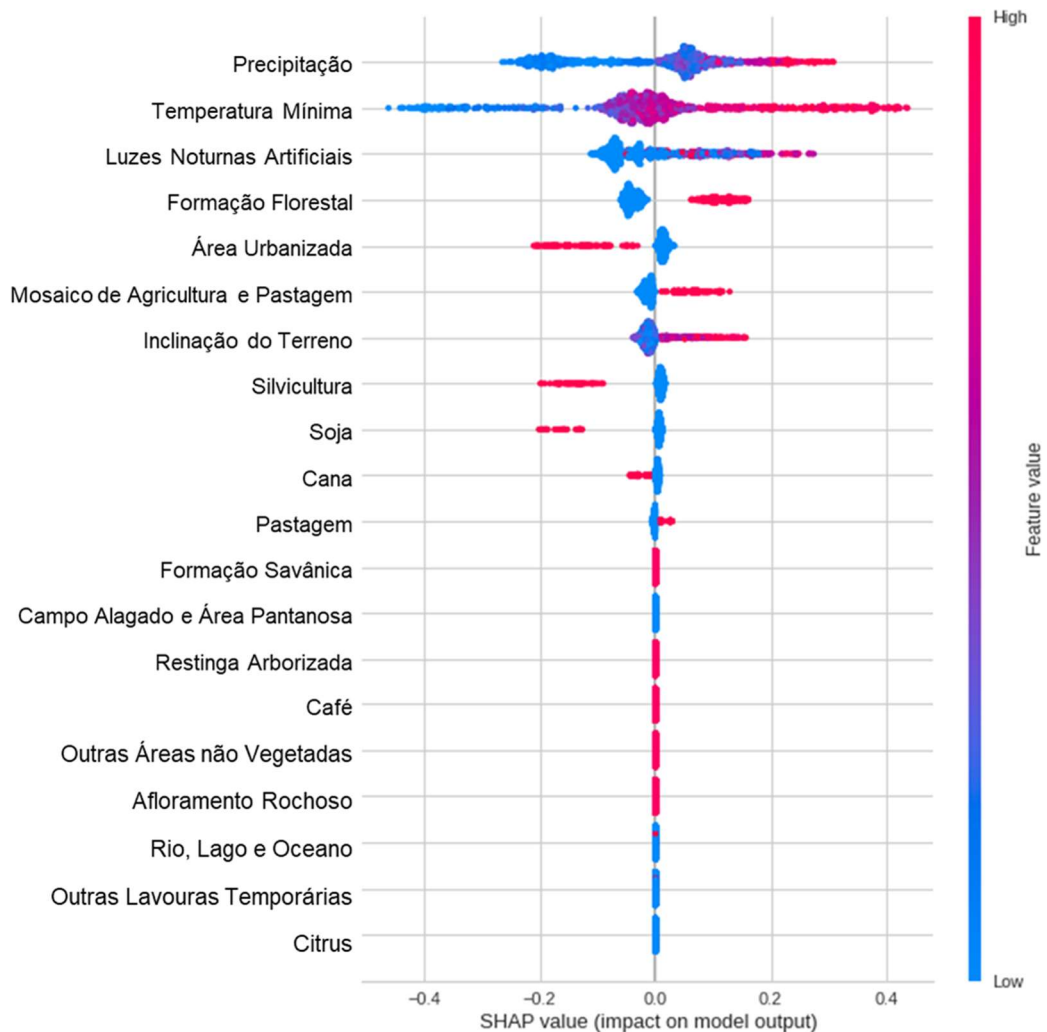


Figura 119 - Valores SHAP para o modelo LightGBM

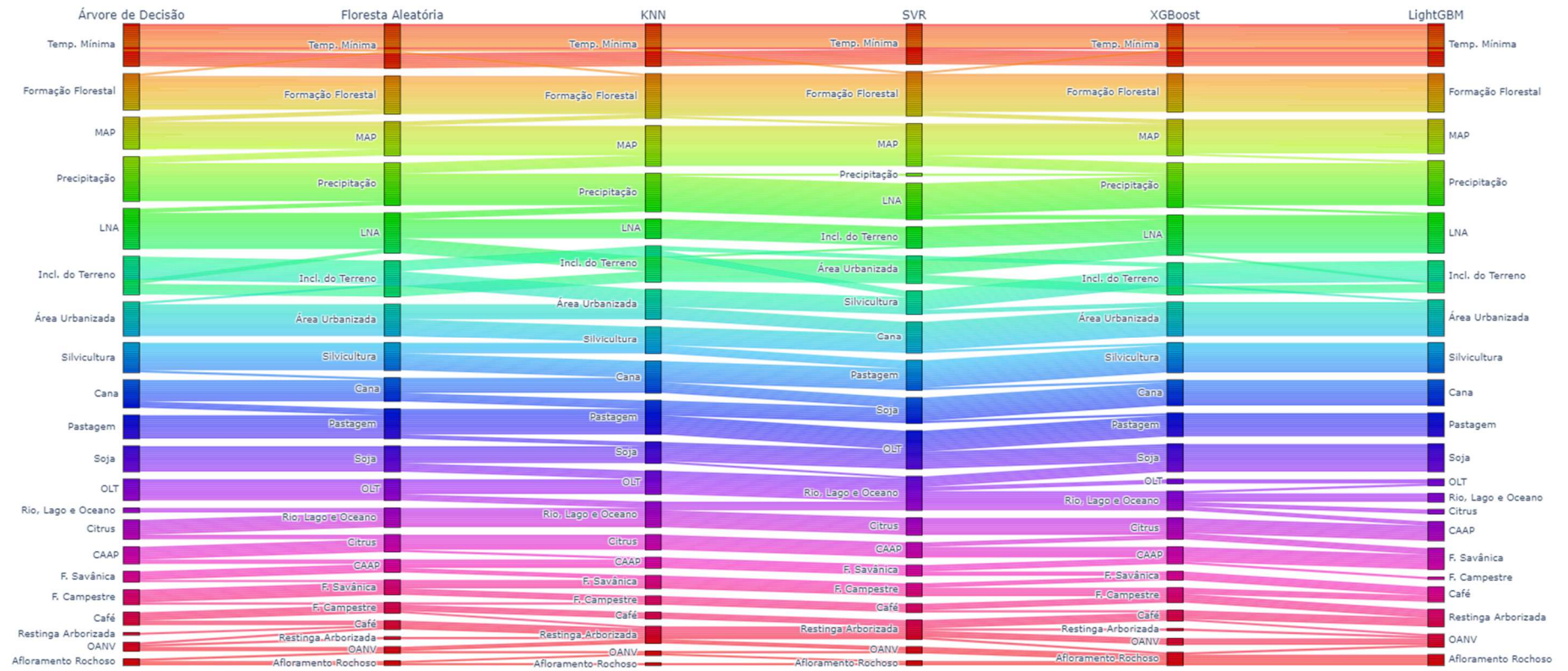


## 4.8 COMPARAÇÃO DOS MODELOS DE REGRESSÃO

### 4.8.1. Importância das variáveis

Para verificar quais variáveis apresentam maior impacto nos modelos, foi calculada uma pontuação para as 21 variáveis indicadas no diagrama de SHAP; para cada modelo, a variável menos frequente recebeu a pontuação 1 e a mais frequente obteve pontuação igual a 21 (Figura 120). Neste diagrama de coordenadas paralelas é possível notar algumas variações acentuadas na importância de algumas variáveis. Por exemplo, a Precipitação, que foi uma das mais usadas na maioria dos modelos, foi a menos usada no SVR; no diagrama, a barra associada a essa variável aparece mais estreita na coluna SVR do que nos demais modelos. Luzes Noturnas Artificiais teve menor importância para o modelo KNN em relação aos outros modelos.

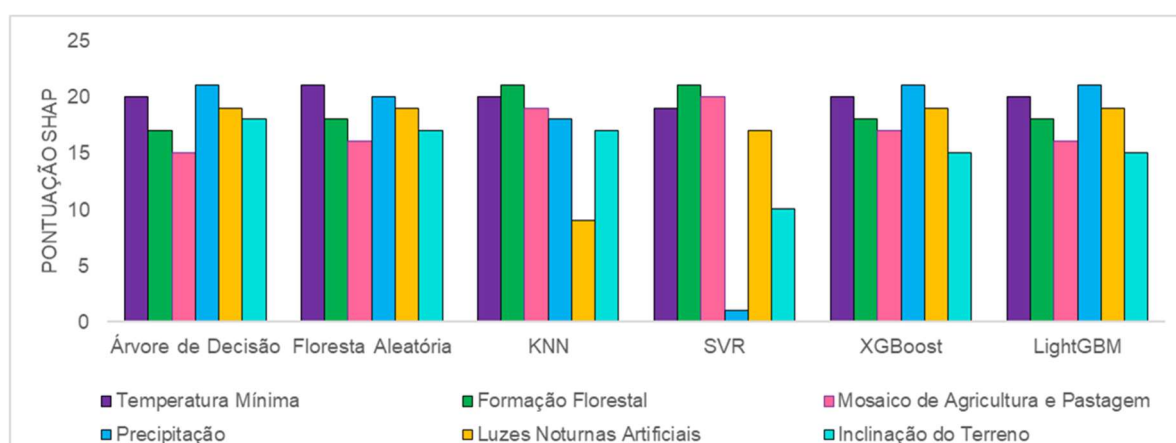
Figura 120 - Diagrama de Coordenadas Paralelas da importância das variáveis nos modelos de regressão



MAP = Mosaico de Agricultura e Pastagem; OLT = Outras Lavouras Temporárias; CAAP = Campo Alagado e Área Pantanosa; OANV = Outras Áreas não Vegetadas

A Figura 121 destaca as cinco variáveis com os maiores valores de SHAP nos modelos de regressão. As variáveis que tiveram os valores mais altos de SHAP foram: Temperatura Mínima, Formação Florestal e Mosaico de Agricultura e Pastagem, Precipitação, Luzes Noturnas Artificiais e Inclinação do Terreno. Estas variáveis foram importantes em todos os modelos estruturados em árvores, nos quais o cálculo da importância da variável considera o quanto cada variável contribui para diminuir a variância nas folhas e nós intermediários. São todas as variáveis numéricas e também as categóricas com as maiores quantidades de pontos amostrais. Para as demais variáveis (outras classes de uso e cobertura da terra), as pontuações SHAP seguem ordem similar à daquela apresentada na Figura 101. As Figuras 94, 96 e 97 mostram que, dentre as variáveis numéricas consideradas na modelagem, a Temperatura Mínima, variável que se sobressai em relação às categóricas, apresenta maior frequência de valores (moda) em torno da média, com baixa frequência de *outliers*, principalmente na amostra de Presença. Por outro lado, para a Inclinação do Terreno, a média se distancia da moda em direção aos *outliers*. Sendo assim, o enviesamento da média por anomalias pode ser uma estratégia de seleção de variáveis em vez de usar apenas a correlação entre elas. Ecologicamente, este padrão pode significar que os espécimes amostrados no GBIF estão em uma amplitude mais restrita de Temperatura Mínima e mais abrangente de Inclinação do Terreno.

Figura 121 - Principais variáveis para os seis modelos de regressão



O SHAP mede o efeito das variáveis considerando a sua interação com as demais, e explicando por que um algoritmo teve um determinado resultado. A forma como a máquina aprende depende dos valores dos hiperparâmetros configurados no modelo. Por exemplo, na Floresta Aleatória, o hiperparâmetro `max_features`, que é o número de variáveis que resultará na melhor partição nos nós, pode ter implicações na importância da variável. Além disso, as

características das variáveis e o tamanho amostral parecem ter maior importância nos modelos do que razões ecológicas, visto que as variáveis categóricas Formação Florestal e Mosaico de Agricultura e Pastagem foram as que tiveram maior pontos amostrais. Nos modelos KNN e SVR, que tiveram desempenho inferior aos demais, houve maior variação na pontuação SHAP entre as variáveis. A Precipitação, com altos valores de SHAP e importância relativa em quase todos os modelos, foi uma das menos usadas no modelo SVR, que considerou mais as variáveis Outras Lavouras Temporárias e Rio, Lago e Oceano, em relação aos outros modelos. Luzes Noturnas Artificiais, também com altos valores de SHAP na maioria dos modelos, foi pouco usada no modelo KNN. Nos modelos baseados em distância e em maximização de margens, tiveram maior importância as variáveis categóricas, que apresentam apenas os valores extremos 0 e 1000 e, portanto, são melhor discriminadas no espaço de atributos e não dependem de definições rígidas de limiares.

#### 4.8.2 Métricas de acurácia dos modelos

O coeficiente de regressão ( $R^2$ ) é uma métrica de regressão, com valores que variam de -1 a 1, que representa o percentual da variância com o qual as variáveis preditoras contribuem. De modo geral, quanto maior esse valor, melhor é o poder de predição do modelo; entretanto, um alto valor de  $R^2$  não quer dizer, necessariamente, que o modelo é robusto (HARRISON, 2020). Então, são apresentadas outras métricas para discussão dos modelos. Além disso, é importante lembrar que essas métricas são calculadas apenas sobre os pontos amostrais de presença e de pseudo-ausência, e o modelo é gerado também para toda a área de interesse. Por isso, é importante que os modelos tenham certo poder de generalização, isto é, de adequar-se a novos dados além daqueles usados no treinamento e teste, visto que são 1.190 pixels que amostram uma área de quase 250 mil km<sup>2</sup>. A Tabela 36 mostra as métricas de avaliação dos seis modelos de regressão. As células mais escuras são as que tiveram melhor desempenho. Os modelos baseados em árvores apresentaram as melhores métricas de predição.

Uma curva ROC (*Receiver Operating Characteristic*) é a representação gráfica do desempenho de um modelo de Aprendizado de Máquina e mostra a relação da Sensibilidade (taxa de verdadeiros positivos) e da Especificidade (taxa de falsos positivos) quando usado um método de classificação. A AUC (*Area Under the Curve*) é a área do gráfico que fica sob a curva ROC e permite comparar modelos; seu valor pode variar de 0 a 1 e, quanto maior o AUC,

melhor o desempenho do modelo. Para os métodos de regressão não é possível calcular a Sensibilidade e a Especificidade. No entanto, a AUC coincide com a probabilidade de um classificador discriminar uma instância positiva (Presença) escolhida aleatoriamente ser maior do que uma instância negativa (Pseudo-ausência) escolhida aleatoriamente (ZWEIG; CAMPBELL, 1993). O sklearn dispõe do atributo `sklearn.metrics.roc_auc_score`<sup>117</sup> que calcula a AUC a partir de pontuações de predição, inclusive para modelos de regressão.

Tabela 36 - Métricas de avaliação dos modelos de regressão

	Árvore de Decisão	Floresta Aleatória	KNN	SVR	XGBoost	LightGBM
R <sup>2</sup>	0,724	0,777	0,422	0,048	0,488	0,425
Variância explicada	0,724	0,777	0,422	0,048	0,488	0,424
Erro absoluto médio	0,069	0,163	0,283	0,354	0,298	0,320
Erro quadrático médio	0,069	0,056	0,144	0,234	0,128	0,143
Log do erro quadrático médio	0,011	0,009	0,024	-	0,021	0,023
<b>AUC</b>	<b>0,93</b>	<b>0,98</b>	<b>0,87</b>	<b>0,76</b>	<b>0,91</b>	<b>0,89</b>

A Análise Exploratória dos Dados mostrou que os pontos de Presença e Pseudo-ausência não se distinguem nas variáveis numéricas e isso, combinado ao grande tamanho amostral (1.190 pontos), compromete a separabilidade dos pontos no hiperplano do SVR. A separabilidade também fica comprometida no modelo KNN, que discrimina Presença e Pseudo-ausência no espaço de atributos, ainda que possa lidar com relações não lineares entre variáveis preditoras e variável resposta.

A distribuição bilinear e não aleatória dos resíduos, com ausência de normalidade e variância inconstante, acontece provavelmente porque a entrada dos modelos têm apenas dois valores (1 = Pseudo-ausência e 2 = Presença). Os gráficos de distribuição de resíduos em função dos valores preditos mostram que existe autocorrelação nos modelos. No entanto, a violação de premissas de homocedasticidade e distribuição normal de resíduos é fundamental nos modelos

<sup>117</sup> `sklearn.metrics.roc_auc_score`. Disponível em <[https://scikit-learn.org/stable/modules/generated/sklearn.metrics.roc\\_auc\\_score.html](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.roc_auc_score.html)>. Acesso em 04 nov 2022.

lineares, que não é necessariamente o caso dos MAHs que usam diversas variáveis ambientais que não têm relação linear com a presença das espécies. Por exemplo, a presença de uma determinada espécie e a temperatura podem ter uma relação quadrática, com um ponto ótimo (ou um intervalo) de temperatura, em que a adequabilidade é máxima e que a partir dele, a adequabilidade começa a diminuir (MITCHELL et al., 2018; ARROYO et al., 2022); caso contrário, temperaturas extremas não seriam um problema para a manutenção de populações ou implicar em alterações metabólicas e comportamentais, mesmo para animais endotérmicos como os mamíferos. Esta observação é válida para as variáveis numéricas que têm um gradiente contínuo, e não para as categóricas, que são binárias (existe ou não existe a classe de cobertura da terra no pixel).

Os modelos baseados em árvores, por serem sucessivas tomadas de decisão, conseguem lidar com essas relações não lineares. As árvores de decisão, no entanto, podem ser mais instáveis porque pequenas variações nos dados podem resultar em modelos completamente diferentes, que não têm capacidade de generalização, ou seja, são mais sujeitos a sobreajustes. Quando as árvores são muito profundas, com vários níveis, a partição vai especificando cada vez mais as características das folhas. O balanceamento dos pontos de Presença e Pseudo-ausência são uma estratégia para que as árvores não criem modelos tendenciosos.

As florestas aleatórias, por mesclarem várias árvores de decisão e apresentarem um maior número de decisões, podem garantir previsões mais precisas com menor risco de sobreajuste. As florestas aleatórias selecionam subconjuntos aleatórios de variáveis e constroem árvores menores (menos profundas) e, portanto, a partição não fica tão específica e o modelo não tende a sobreajustes.

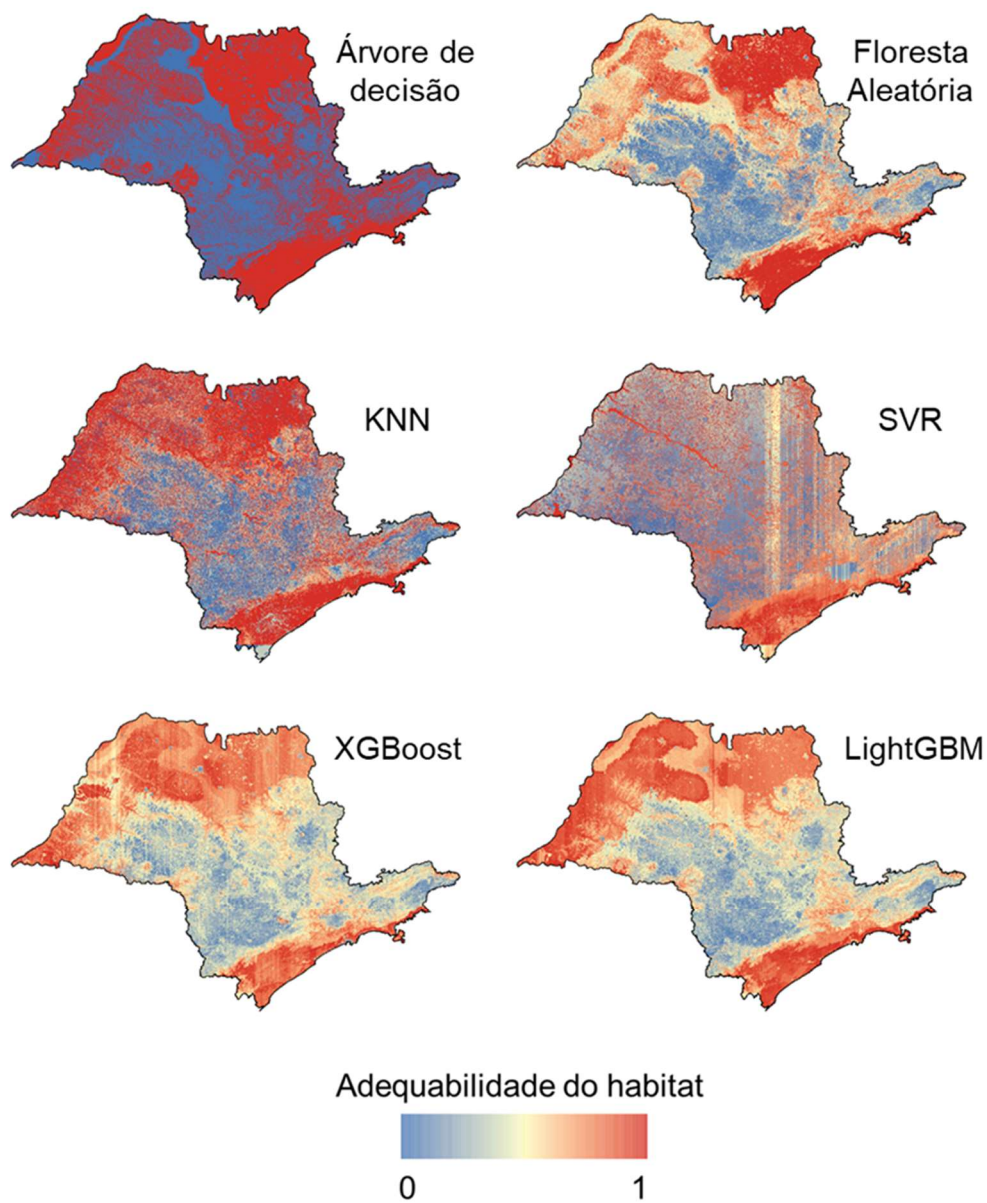
Considerar vários modelos permite verificar se os resultados seguem um determinado padrão. Com exceção do SVR, que teve desempenho inferior aos demais, os modelos apontam três principais áreas no Estado de São Paulo com alta AdH: 1) Noroeste do estado, onde as paisagens são fortemente marcadas pelas agricultura; 2) Nordeste do estado, onde a produção de cana-de-açúcar é dominante, mas onde ainda há remanescentes de Cerrado não protegidos por UCs; e 3) domínio geomorfológico da Província Costeira Sul do Estado, onde há justaposição de UCs de várias categorias. Esse resultado é um efeito da alta proporção de pontos de Presença em áreas de Mosaico de Agricultura e Pastagem e Formação Florestal, que são classes de cobertura da terra dominantes nessas porções do Estado de São Paulo.



As porções noroeste e nordeste do Estado, onde praticamente existem UCs federais e estaduais, podem ser potenciais áreas para o estabelecimento de novas Unidades, mais especificamente de Áreas de Proteção Ambiental (APAs), que não requerem a desapropriação de terras e poderiam reforçar a regulação do uso da terra. O noroeste do Estado, como apresentado no item 4.1.3.2 consiste em uma região dominada pelo cultivo de cana-de-açúcar. Embora não seja o habitat natural dos mamíferos silvestres, as paisagens agrossilvopastoris são locais de passagem desses animais. A alta AdH média se deve à alta frequência de observações das espécies nessas áreas. Sendo assim, essa classe de cobertura da terra acaba tendo maior importância na modelagem, visto que a amostragem nesses locais é mais viável do que em locais de difícil acesso como regiões de alta inclinação do terreno ou de grandes altitudes. A alta frequência de observações em áreas de uso intensivo indica que deve haver iniciativas de conservação mesmo fora das áreas protegidas, além de estratégias eficientes de resgate de fauna e de educação ambiental que tenham como objetivo inibir retaliação por parte de populações humanas. Esta condição reforça a importância da manutenção da vegetação nativa em Áreas de Preservação Permanente e Reservas Legais como trampolins ecológicos.

Por outro lado, a porção central do estado tem vastas extensões onde a AdH é baixa e requerem atenção quanto à regulação do uso do solo para serem proveitosas como áreas de trânsito mastofauna silvestre, especialmente para as espécies que demandam maiores áreas de vida (Figura 122).

Figura 122 - Mapas de Adequabilidade de Habitat



A Árvore de Decisão resultou em um mapa binário de Adequabilidade de Habitat, com valores de 0 e 1, ou seja, não tem valores contínuos como esperado para um modelo de regressão. Os modelos de árvores do tipo *boosting*, XGBoost e LightGBM, são espacialmente similares, seguindo o padrão do plano de informação de Precipitação, que é a variável com pontuações SHAP mais altas nesses modelos.

O GridSearchCV é uma técnica demorada pois percorre todas as combinações possíveis de hiperparâmetros (HSU, 2020). Por exemplo, para Árvore de Decisão, foram considerados cinco hiperparâmetros, com 480 combinações possíveis (Figura 123). O RandomizedSearchCV

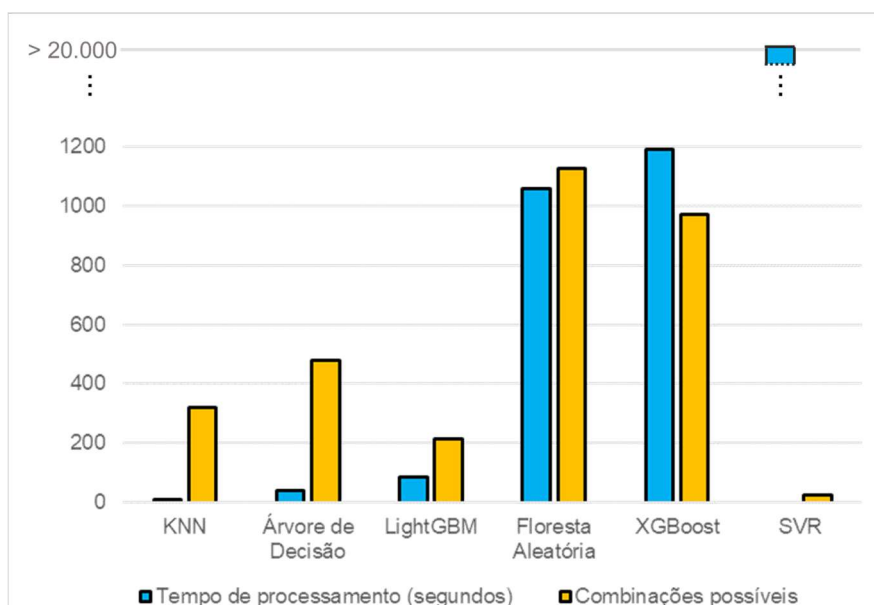
é uma alternativa mais ágil mas que perde em desempenho (KORSTANJE, 2021). A Figura 124 mostra o tempo em segundos para o processamento do GridSearchCV e as combinações possíveis. O SVR, embora tenha sido configurado com apenas 24 combinações, demandou mais do que 20 mil segundos. Floresta Aleatória e XGBoost, embora tenham sido configurados para mais de 1000 combinações, foram processados em tempo satisfatório (até 20 minutos). Entretanto, para estes dois modelos, o GridSearchCV reduziu o valor de  $R^2$ .

Figura 123 - Configuração do GridSearchCV para Árvore de Decisão

```
hyperparameter_space = {
    'criterion': ['squared_error', 'friedman_mse', 'absolute_error', 'poisson'],
    'splitter': ['best', 'random'],
    'max_features':['auto', 'sqrt', 'log2'],

    'min_samples_split':[2,10,15,20],
    'min_samples_leaf':[1,2,3,4,5]
}
```

Figura 124 - Tempo de processamento do GridSearchCV



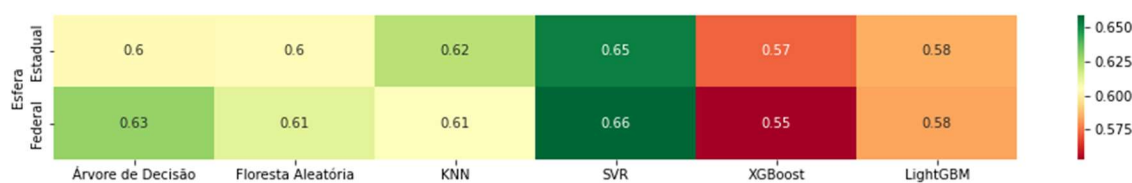
#### 4.9 ADEQUABILIDADE DE HABITAT NAS UNIDADES DE CONSERVAÇÃO

Os chamados Modelos de Distribuição de Espécies também referem-se à Modelagem de Adequabilidade de Habitat (BRADLEY et al., 2012). Como apresentado na Introdução, mais do que prever a ocorrência de uma espécie, esta pesquisa visa identificar se as UCs no Estado

de São Paulo apresentam características ambientais que promovam a manutenção de mamíferos e os resultados dos modelos podem ser interpretados como mapas de adequabilidade ambiental. Sendo assim, para todas as UCs, foram calculadas as estatísticas de Adequabilidade de Habitat (AdH), que são o mínimo, o máximo, a média, a mediana e a maioria dos pixels (250 metros) que cobrem a UC. As estatísticas também foram calculadas por esfera administrativa, uso e categoria.

No SVR, as unidades de ambas as esferas administrativas pontuaram mais do que nos demais modelos. Para todos os modelos, a AdH média foi similar nas UCs federais e estaduais (Figura 125); 91,5% das UCs federais e 71% das UCs estaduais são APAs, e estas proporções podem ser a razão da baixa AdH média (em torno de 0,6) para todos os modelos. Espera-se, de fato, que as APAs, por serem pouco restritivas quanto ao uso do solo, sejam mais antropizadas e a AdH média seja menor, principalmente porque as coberturas Formação Florestal e Mosaico de Agricultura e Pastagem estiveram entre as variáveis mais importantes em todos os modelos, embora as variáveis numéricas também sejam relevantes. As UCs estaduais contemplam maior variedade de categorias; embora a proporção de APAs também seja alta nas unidades sob administração estadual (71% da extensão das áreas protegidas), os parques estaduais, de uso mais restritivo, equivalem a 25% da extensão das UCs estaduais. Esta proporção de unidades mais restritivas aumenta a AdH média das unidades estaduais.

Figura 125 - AdH média das UCs Federais e Estaduais



Os histogramas de AdH por modelo e esferas administrativas são apresentados nas Figuras 126 e 127. Pelos histogramas, é possível observar a posição da média e o valor da maioria dos pixels. Caso trate-se de uma curva assimétrica negativa (que declina para esquerda, com maior frequência de altos valores), a AdH é mais alta. Este é o comportamento esperado para as UCs mais restritivas como os Parques e Estações Ecológicas. Os histogramas que aparentam uma curva assimétrica positiva (que declina para a direita, com frequência de baixos valores) são esperados para as categorias menos restritivas como as APAs.

Como já mencionado, a Árvore de Decisão produziu um mapa de AdH binário, somente com valores 0 e 1. Sendo assim, a frequência nestes valores extremos é alta, já que o número

total de pixels não foi distribuído pelos valores intermediários. Padronizar o eixo y dos histogramas para comparação dos gráficos resultaria em frequências muito baixas nos valores intermediários dos demais modelos. Sendo assim, os histogramas foram gerados com valores variados de frequência e, para comparar os histogramas, é essencial observar a amplitude do eixo y.

A Árvore de Decisão, que gerou o mapa binário, apresenta maior frequência de altos valores de AdH, tanto para a esfera Federal quanto para Estadual. A Floresta Aleatória também apresenta moda no valor máximo de AdH, com valores bem distribuídos para pixels com outros valores. Para as UCs federais, o KNN gerou histograma bimodal, ainda que a maior proporção dos pontos tenha valor máximo; para as UCs federais, quanto maior o valor de AdH, maior a frequência. No SVR, a moda não está no valor máximo, mas em torno de 0,7 e 0,8 para as esferas federal e estadual, respectivamente. No modelo XGBoost, as UCs federais têm moda em torno de 0,85, mas também têm altas frequências no intervalo de 0,3 a 0,5; na esfera estadual, as frequências são bem distribuídas entre baixos e altos valores de AdH. O LightGBM é similar ao XGBoost, mas com ligeiro deslocamento da moda para valores mais altos, de 0,8 para 0,9.

Figura 126 - Histograma de AdH por esfera administrativa das UCs (I)

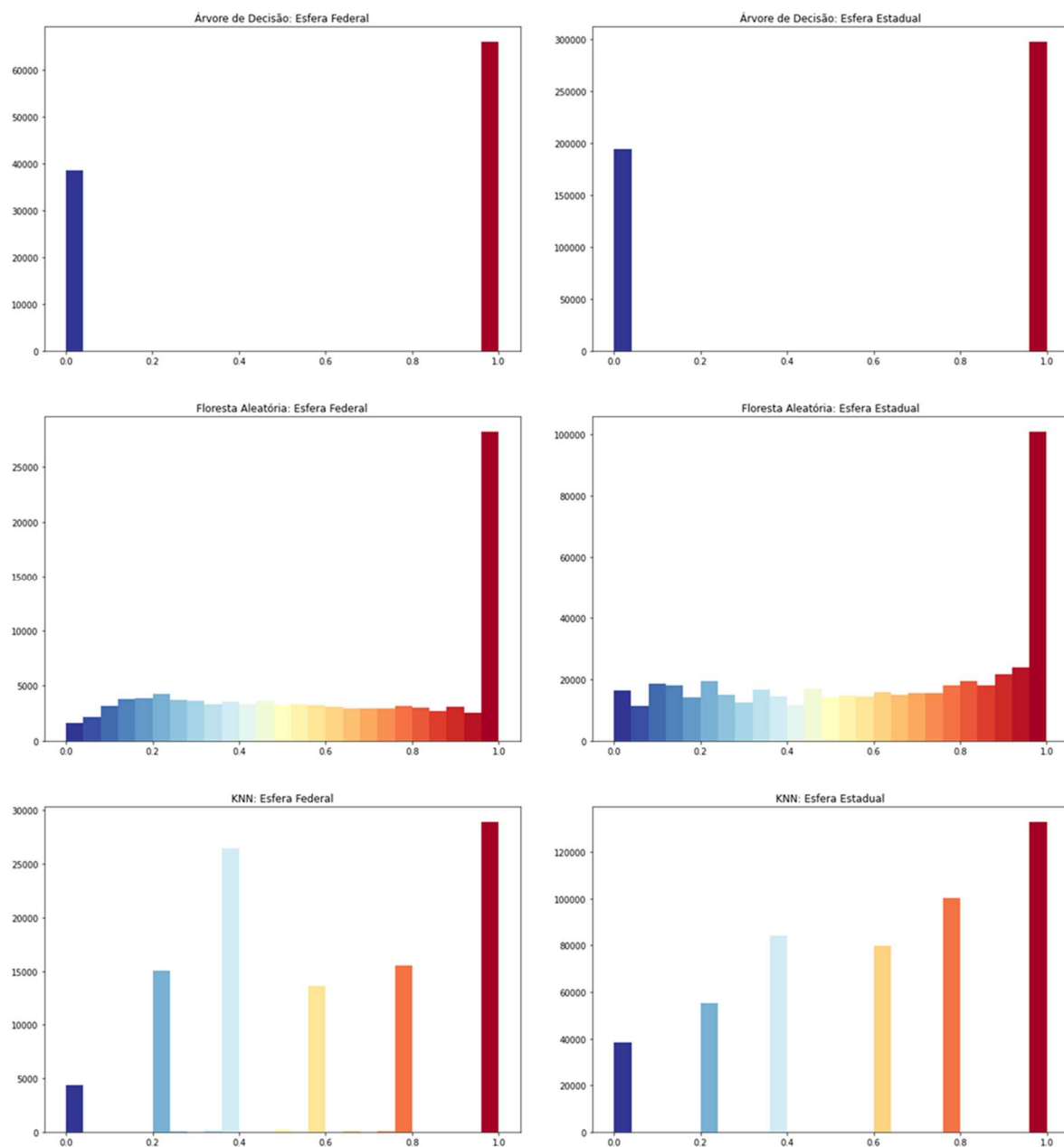
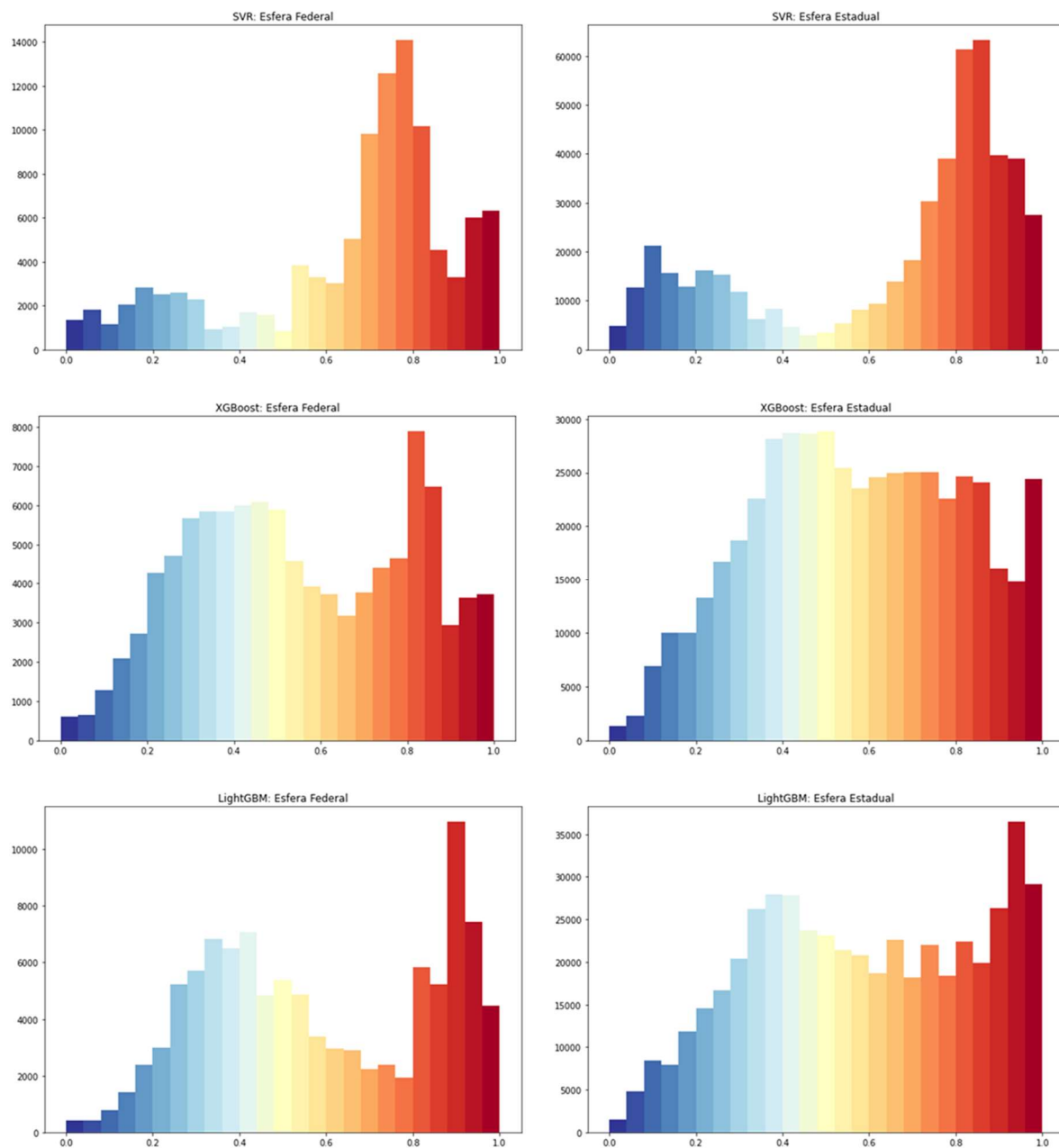
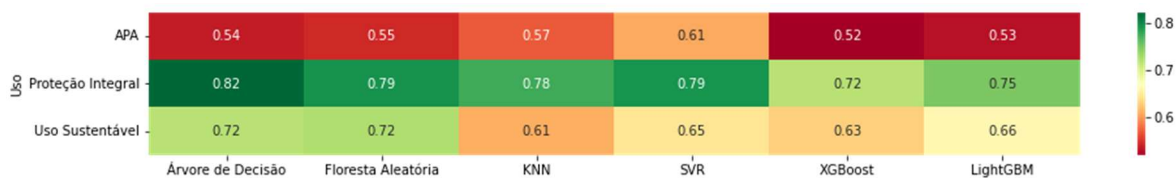


Figura 127 - Histograma de AdH por esfera administrativa das UCs (II)



As APAs são apresentadas à parte das demais unidades de Uso Sustentável, pela alta proporção em termos de extensão quanto por serem menos restritivas quanto ao uso do solo. Para todos os modelos, as APAs têm AdH média inferior às demais UCs de Uso Sustentável e às UCs de Proteção Integral. Essas têm as maiores AdH médias para todos os modelos, com exceção do SVR, que teve o pior desempenho (Figura 128).

Figura 128 - AdH média das UCs por uso



A análise por uso mostra que, para a Árvore de Decisão, tanto UCs de Proteção Integral como APAs e demais UCs de Uso Sustentável têm maior AdH média. Entretanto, a proporção de alta adequabilidade é menor nas APAs em relação às demais UCs. Esta condição se repete na Floresta Aleatória, mas nesta, os valores de AdH se distribuem em valores intermediários, com aumento na frequência de baixa AdH nas APAs. Os histogramas por uso (Figuras 129 a 131) mostram que os demais modelos apresentam padrões semelhantes: as UCs de Proteção Integral têm histogramas de assimetria negativa, com maior frequência de alta AdH; os baixos valores aumentam nas UCs de Uso Sustentável; as APAs têm proporção ainda mais alta de baixa AdH, sendo que nos modelos de boosting, a moda está no intervalo de 0,3 a 0,5.



Figura 129 - Histograma de AdH por uso das UCs (I)

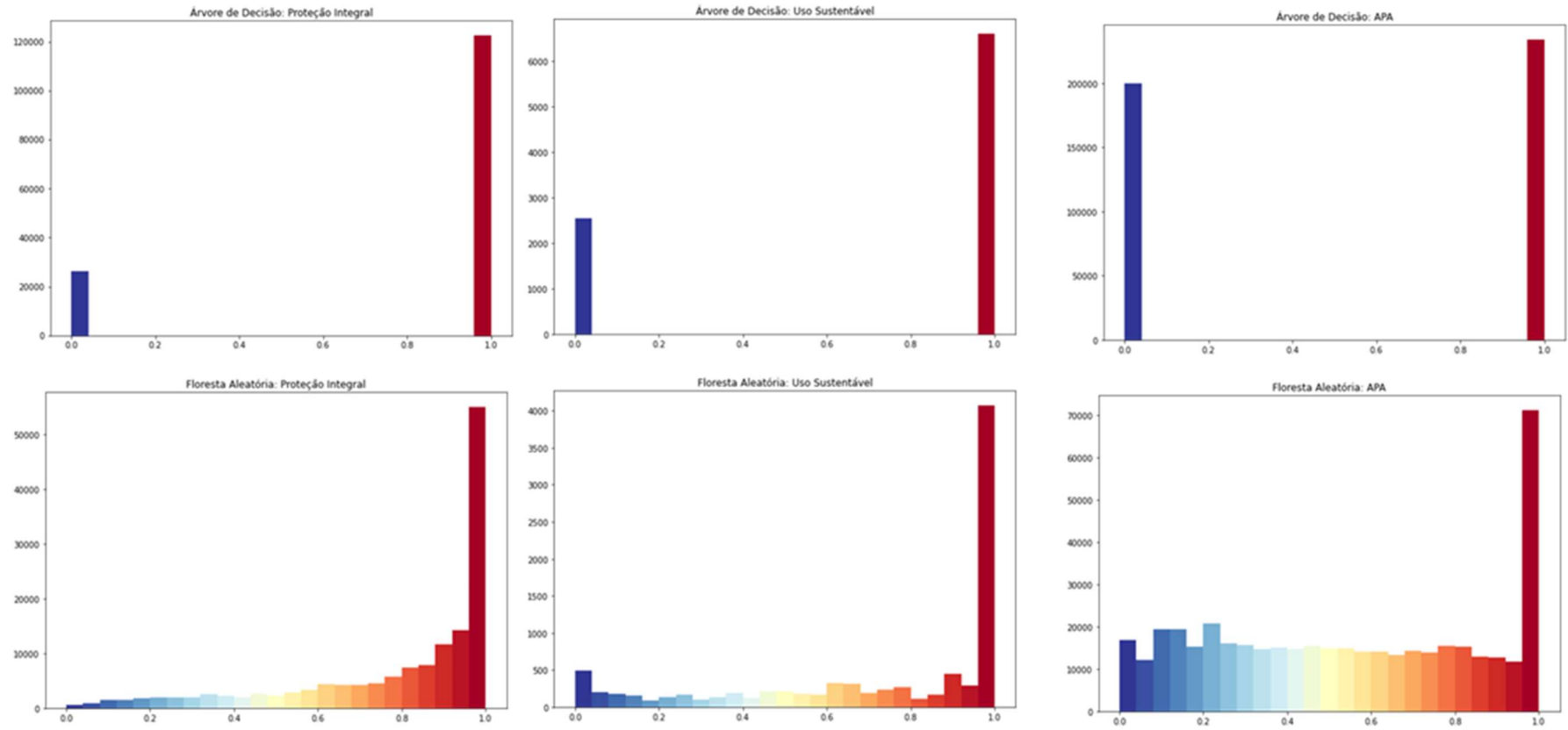


Figura 130 - Histograma de AdH por uso das UCs (II)

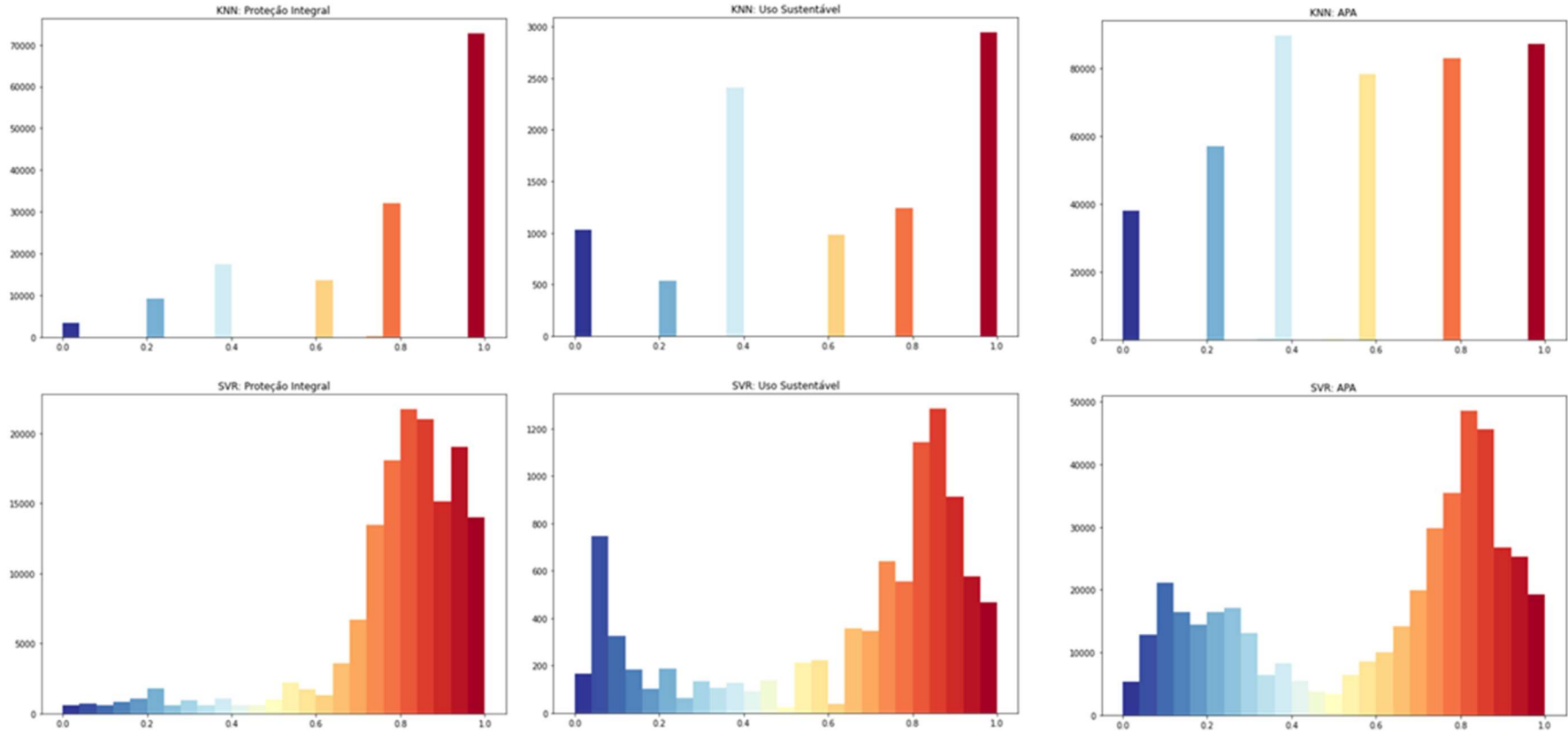
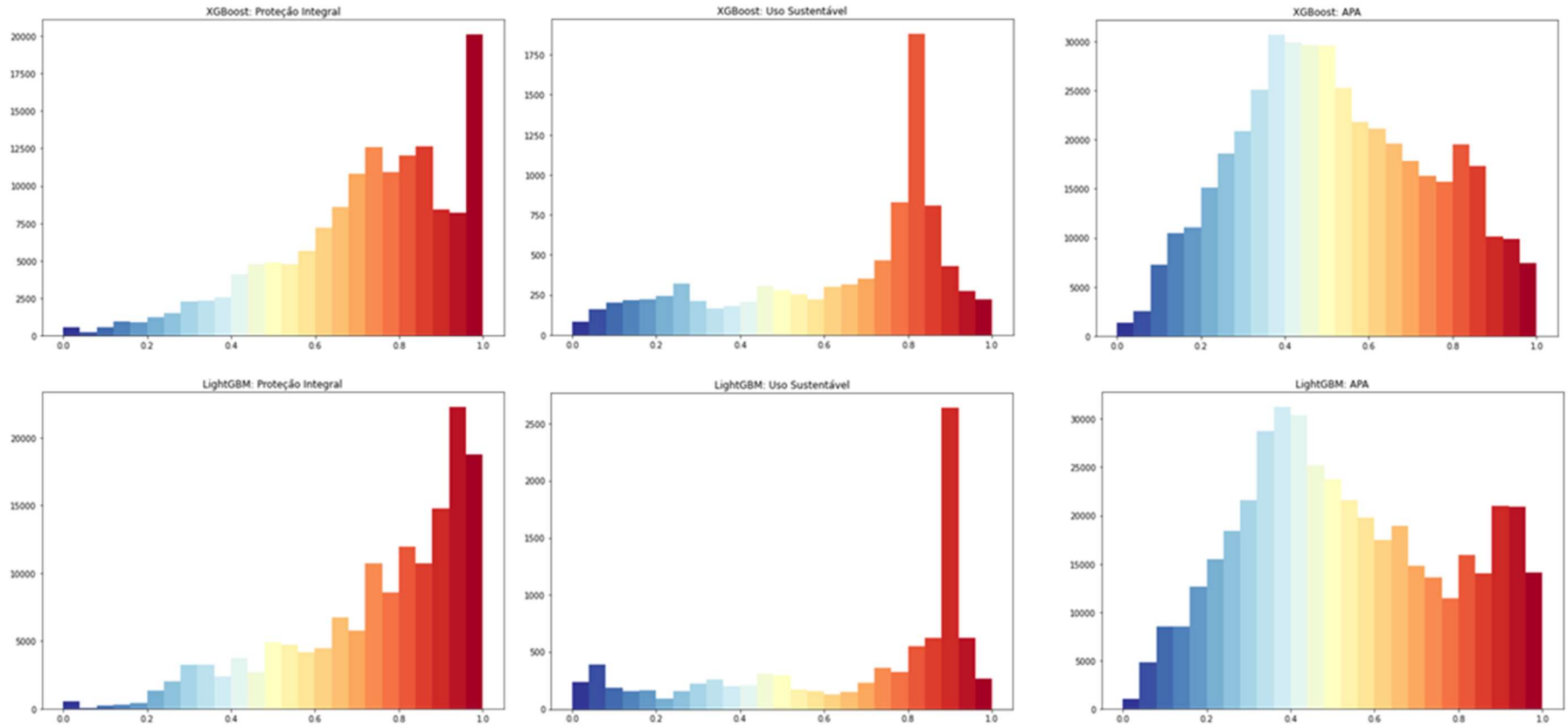
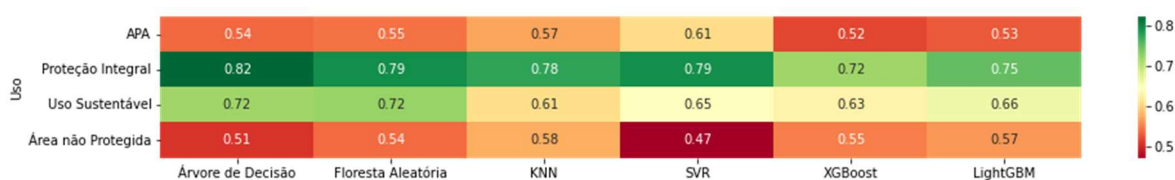


Figura 131 - Histograma de AdH por uso das UCs (III)



A Figura 132 acrescenta a área não protegida por UCs estaduais e federais nas categorias analisadas. Nestas áreas, a AdH média é similar à das APAs para todos os modelos exceto o SVR. Esta similaridade se deve, provavelmente, porque altos percentuais 82% dos pontos incidentes na classe de cobertura da terra Mosaico de Agricultura e Pastagem, predominante no Estado de São Paulo, estão associados a pontos amostrais de Presença. Ainda assim, os modelos vão ao encontro da hipótese de que as UCPIs têm AdH média mais alta. Isto porque outras classes associadas a atividades agrossilvopastoris têm baixo percentual de pontos de Presença e a classificação OC-SVM classificou as áreas associadas a essas classes como Não-presença.

Figura 132 - AdH média em áreas protegidas e não protegidas



No entanto, as APAs não constituem a categoria com a menor AdH. Com exceção do SVR, o modelo menos robusto, todos os modelos resultaram em baixa AdH nas FLONAs e MONAT (Figura 133). Os histogramas destas categorias (Figuras 134 a 144) mostram distribuição com curva assimétrica positiva, isto é, os baixos valores de AdH são mais frequentes. Outras categorias de Uso Sustentável também têm AdH média abaixo de 0,5 na maioria dos modelos. As FEs também têm AdH média de até 0,5. Ressalta-se que esses são valores médios para toda a categoria.

As APAs Jundiáí, Cabreúva e Cajamar, por exemplo, têm áreas de AdH alta (Figura 145), na Serra do Japi, mas também tem locais inviáveis para a mastofauna silvestre, já que estão em uma área altamente urbanizada, a Aglomeração Urbana de Jundiáí<sup>118</sup>. No caso dessas APAs, especificamente onde elas são adjacentes, os modelos que tiveram os melhores resultados (Árvore de Decisão e Floresta Aleatória) indicam que a região dominada por Formação Florestal tem baixa AdH. Isso ocorre porque os modelos consideraram importantes outras variáveis como Inclinação do Terreno e Altitude, que são mais pronunciadas nesse local.

Nos modelos mais robustos, Floresta Aleatória e Árvore de Decisão, as categorias com maior AdH média são RESEX e ARIE (Uso Sustentável) e REBIO e ESEC (Proteção Integral).

<sup>118</sup> Jundiáí, Prefeitura de. Plano Local de Habitação de Interesse Social. Disponível em <<https://fumas.jundiai.sp.gov.br/wp-content/uploads/2017/11/PLHIS-Jundiai-VOLUME-COMPLETO.pdf>>. Acesso em 08 nov. 2022.

Os histogramas dessas categorias apresentam, predominantemente, distribuição com curva assimétrica negativa, ou seja, com maior frequência de altos valores de AdH. Os histogramas do modelo SVR, que teve medidas de acurácia inferiores, são “espelhados” dos demais modelos para as categorias PARNA e APA .

Figura 133 - AdH média das UCs por categoria

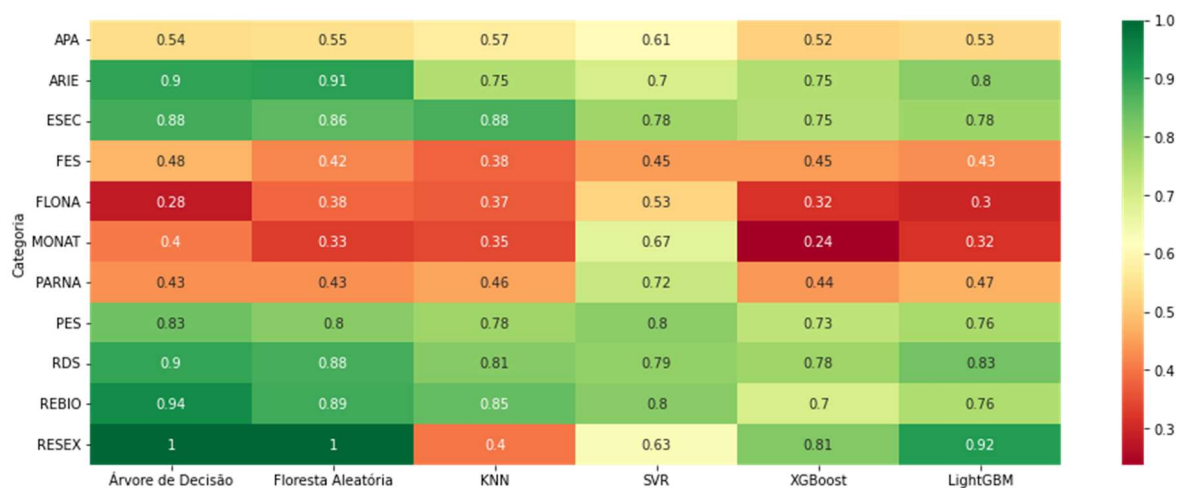


Figura 134 - Histograma de AdH em APAs

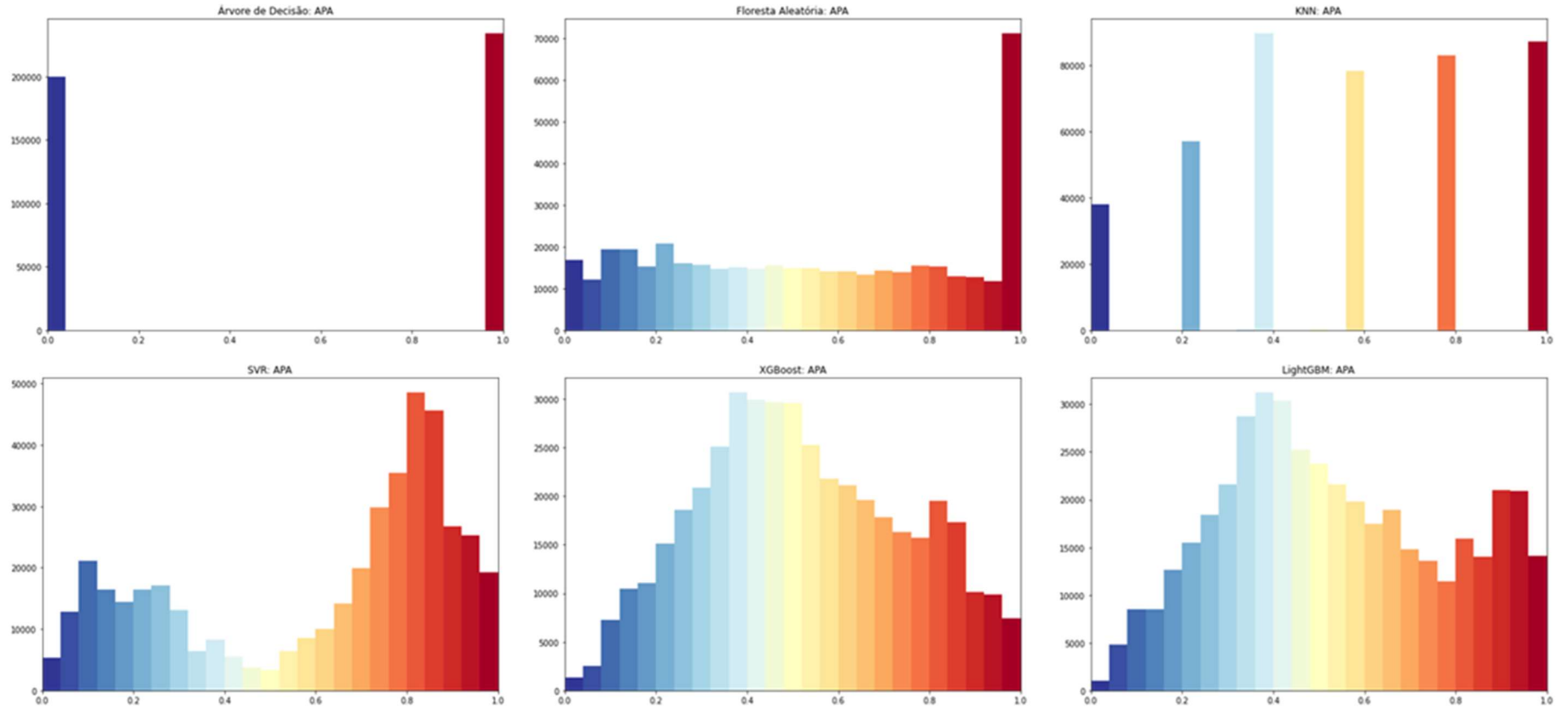


Figura 135 - Histograma de AdH em ARIEs

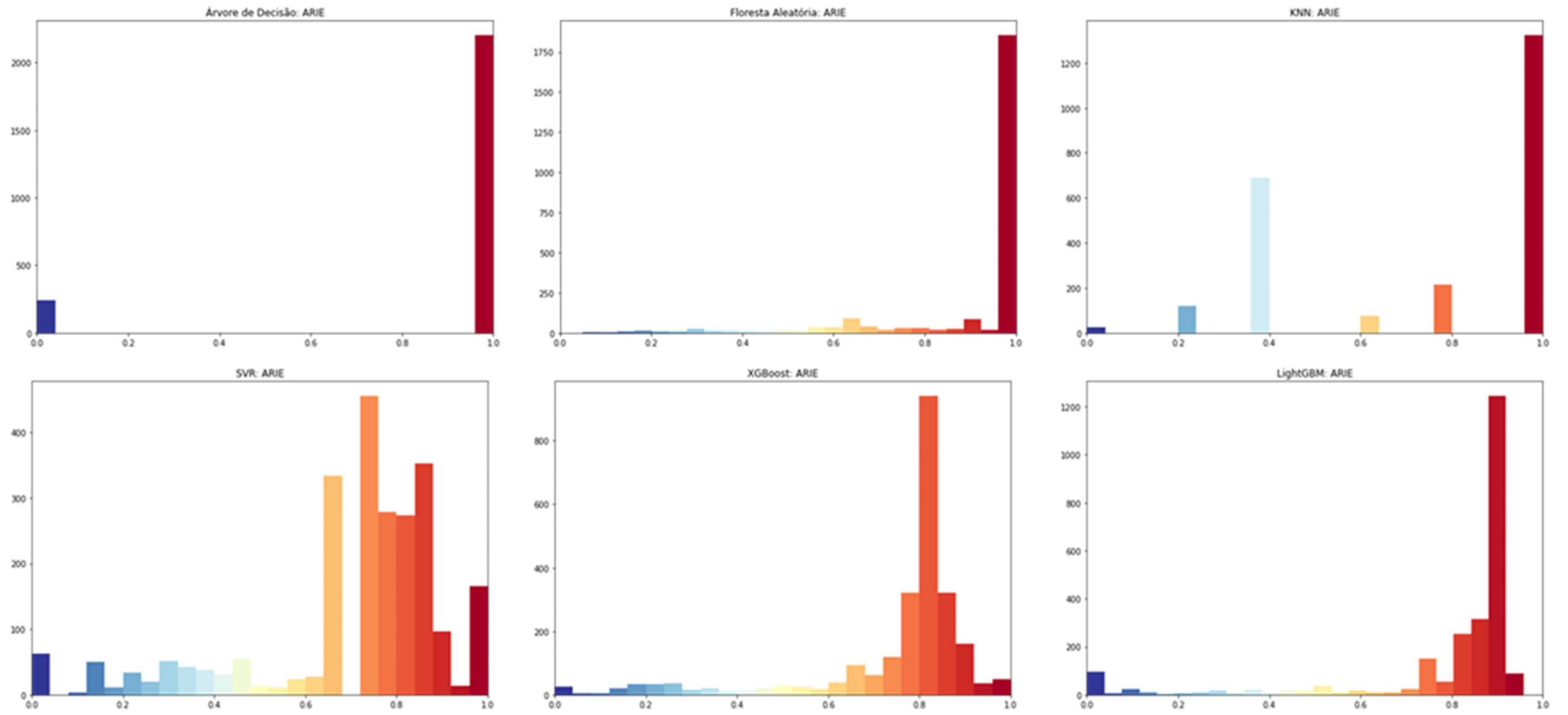


Figura 136 - Histograma de AdH em ESECs

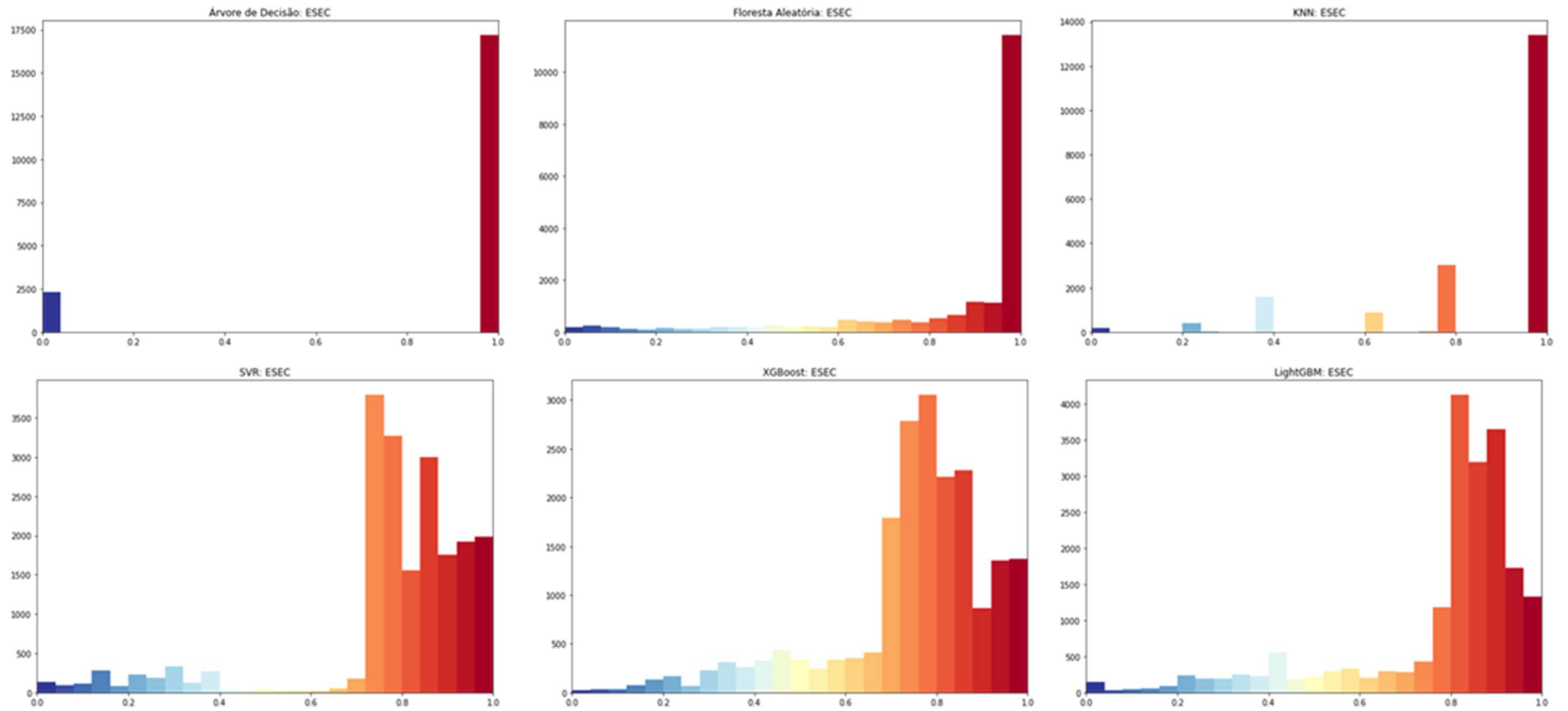




Figura 137 - Histograma de AdH em FEs

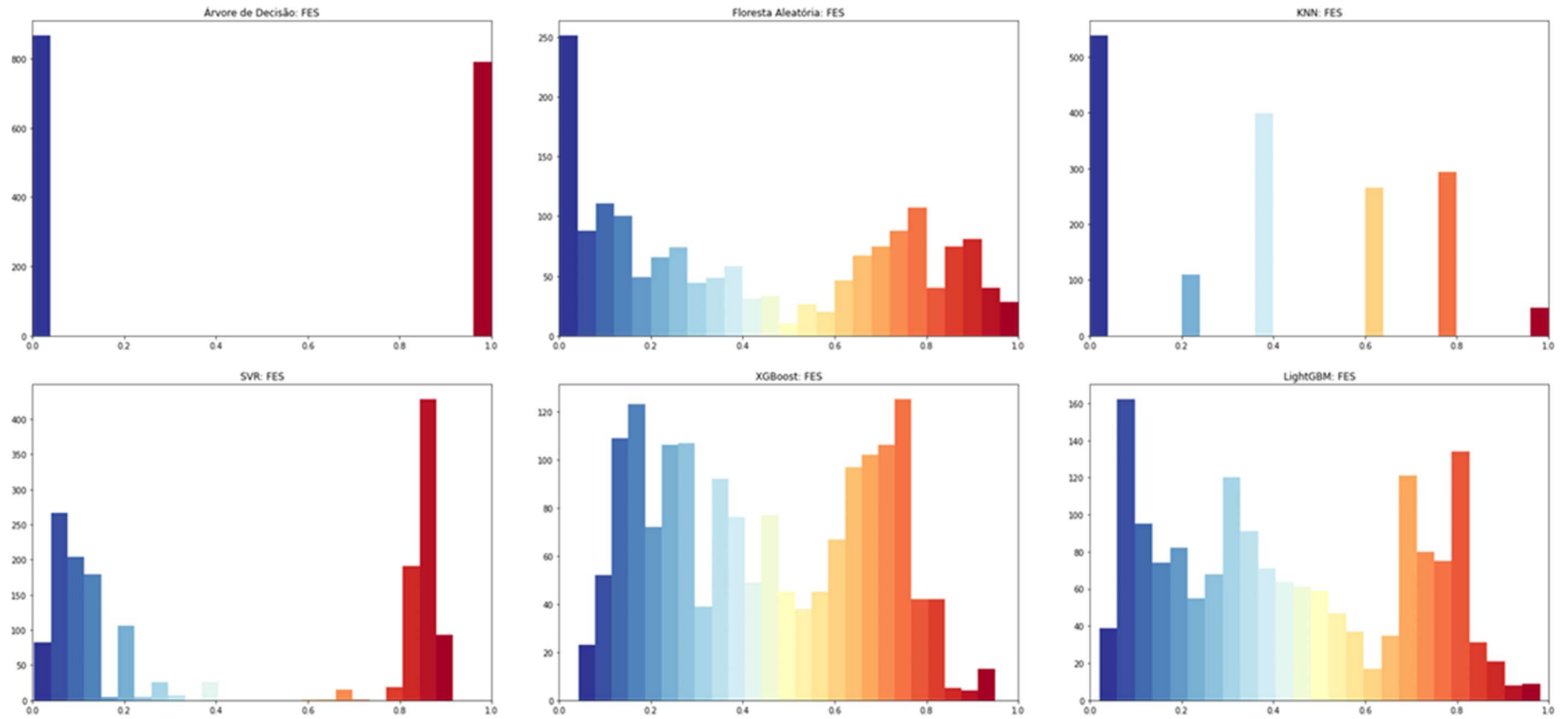


Figura 138 - Histograma de AdH em FLONAs

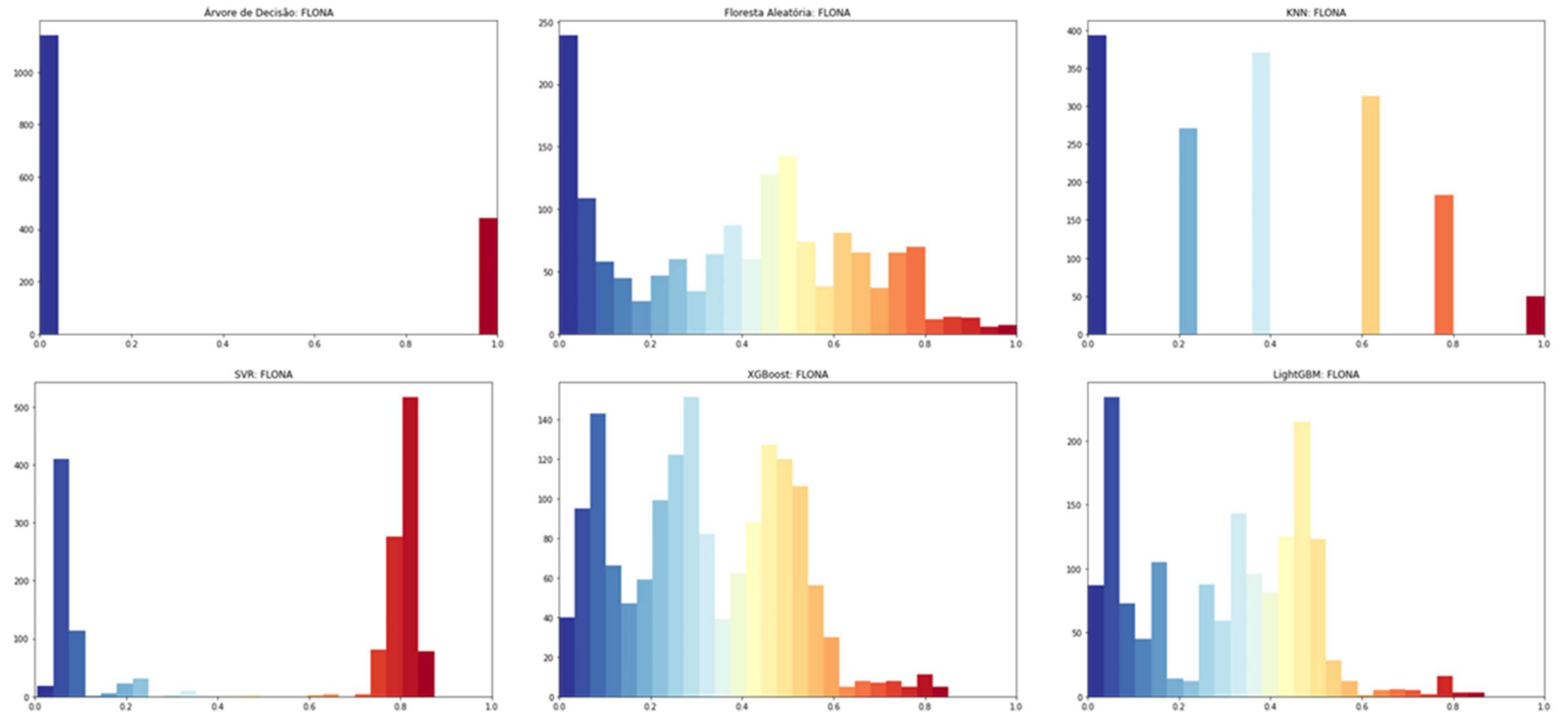


Figura 139 - Histograma de AdH em MONAT

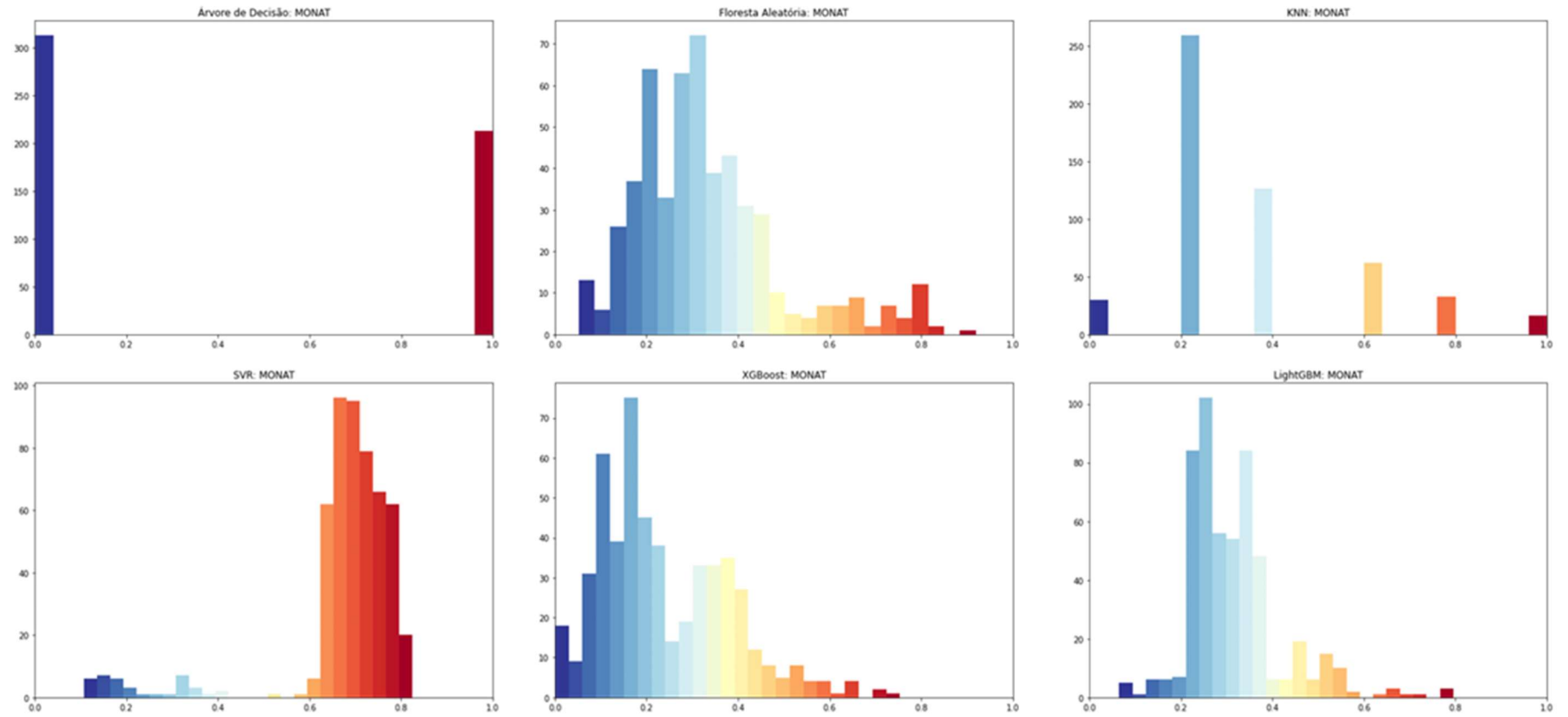


Figura 140 - Histograma de AdH em PARNA

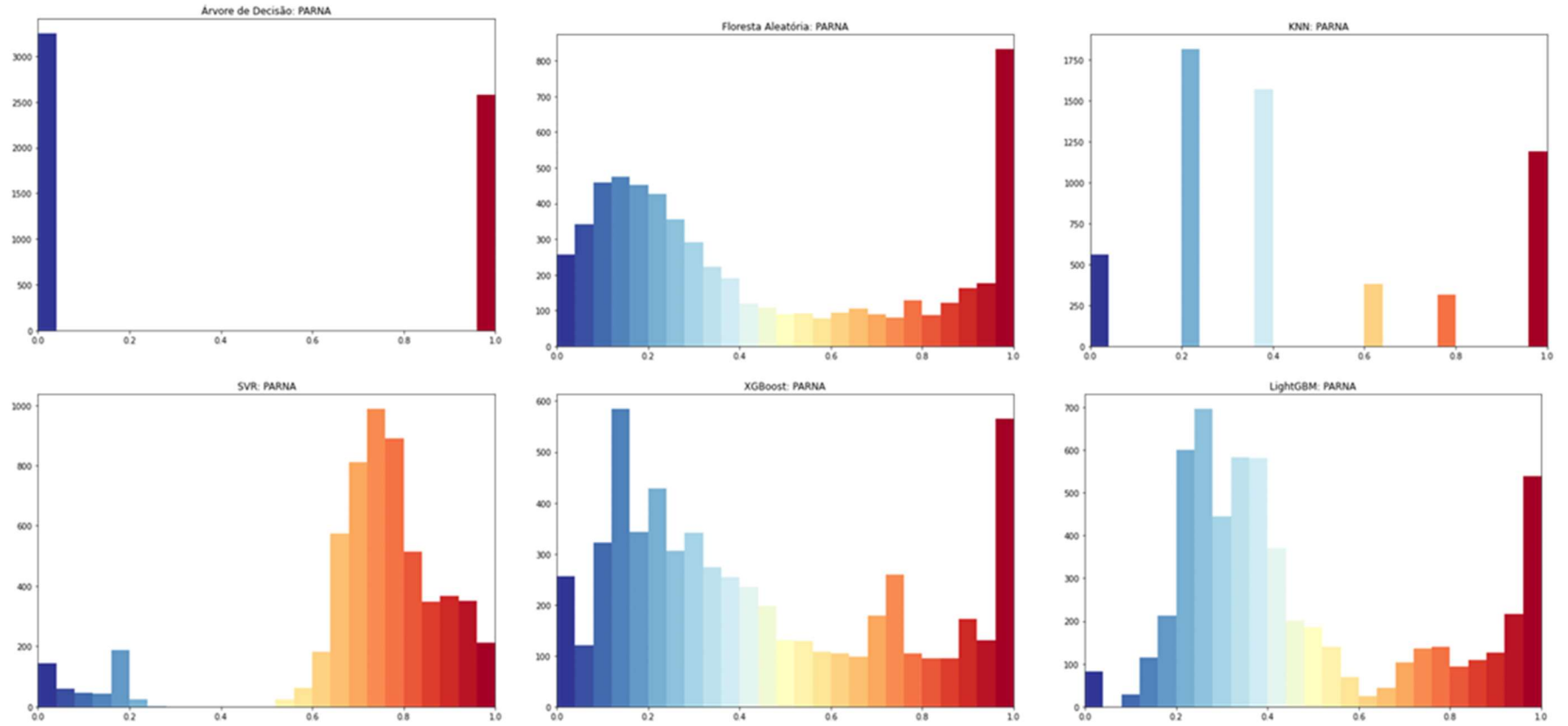


Figura 141 - Histograma de AdH em PEs

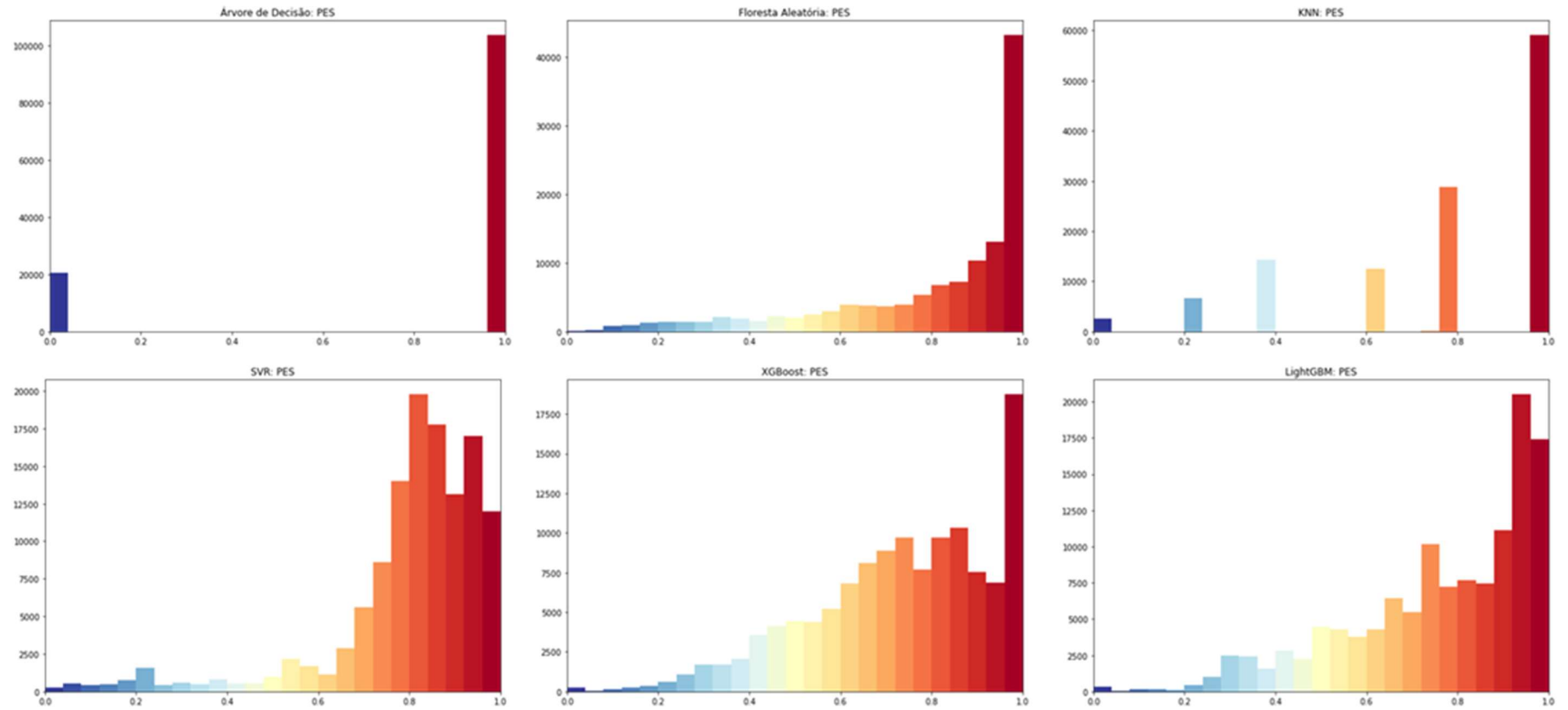


Figura 142 - Histograma de AdH em RDSs

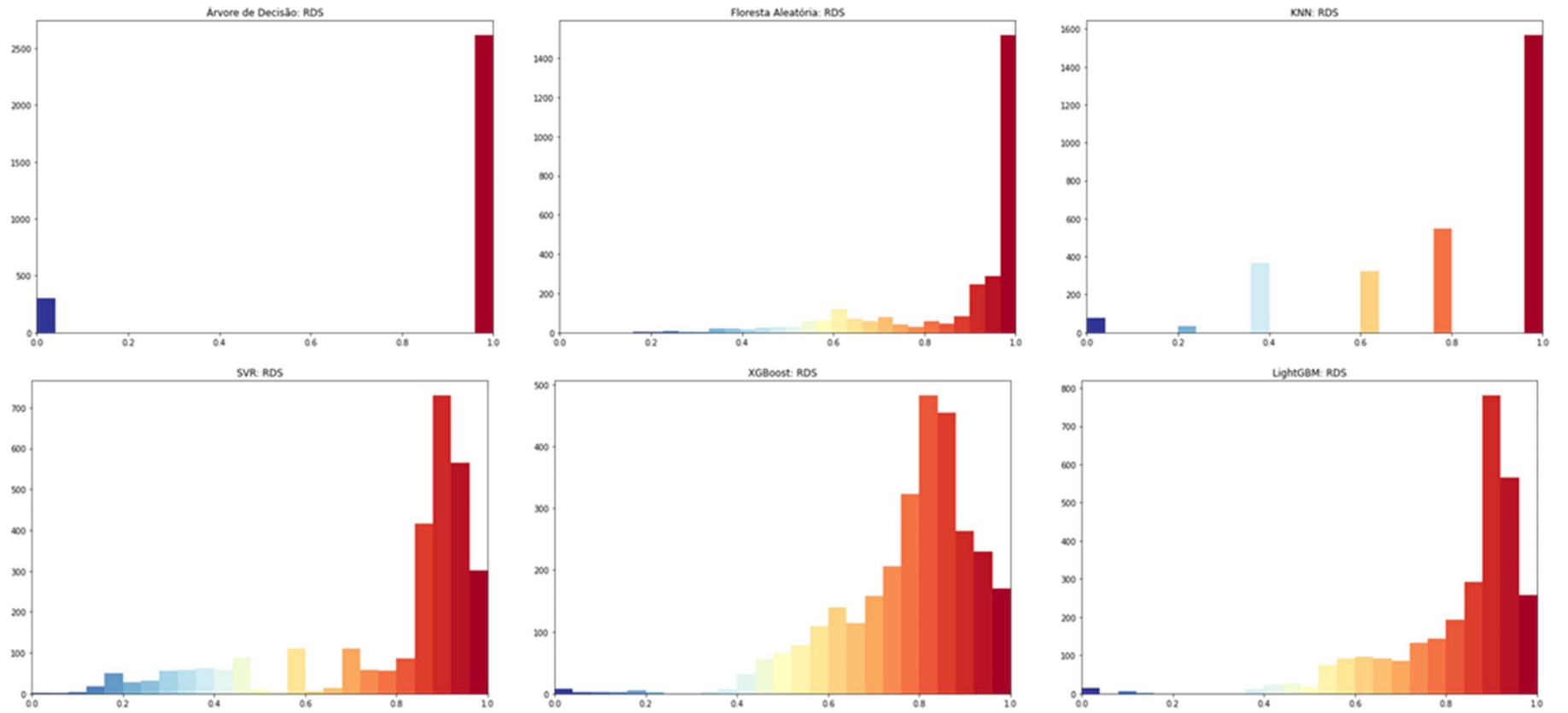


Figura 143 - Histograma de AdH em REBIOS

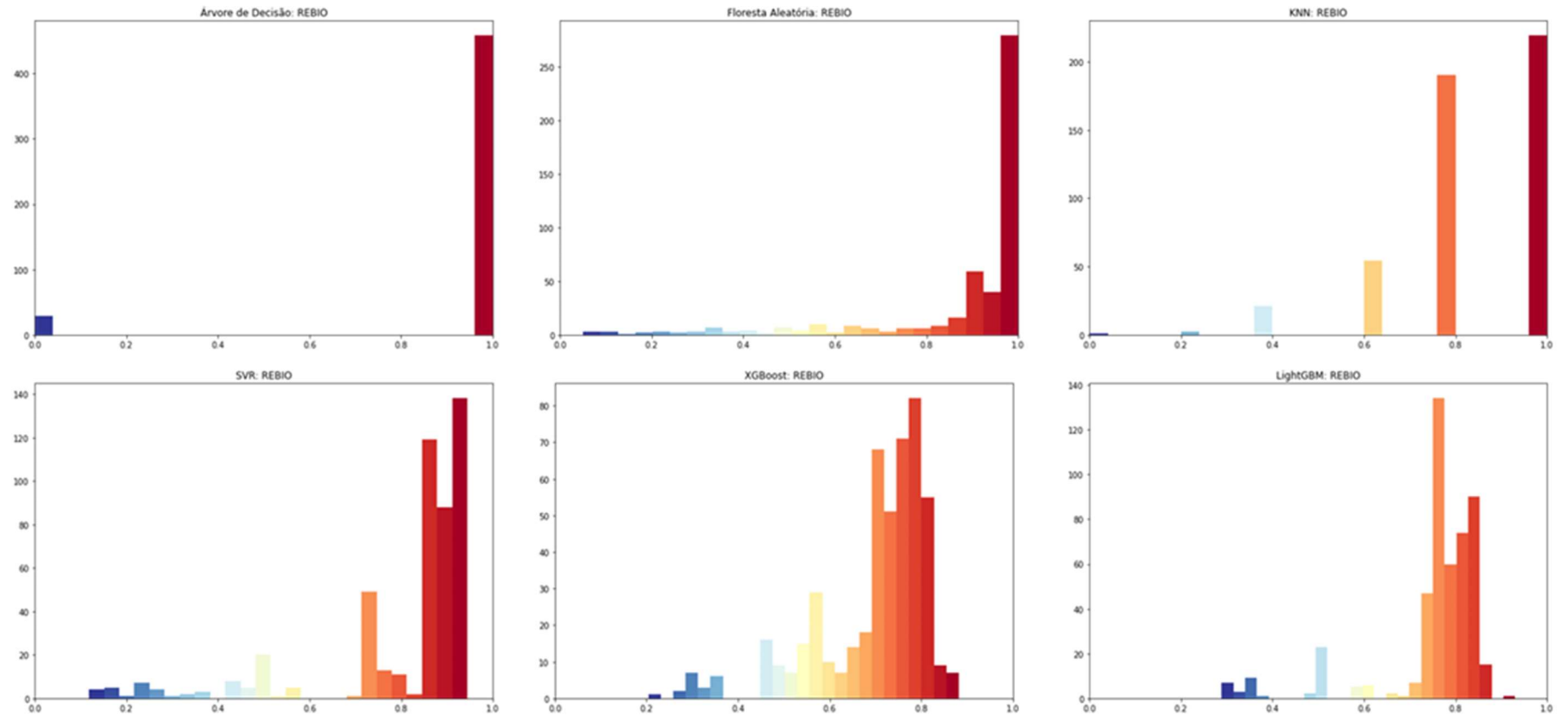


Figura 144 - Histograma de AdH em RESEXs

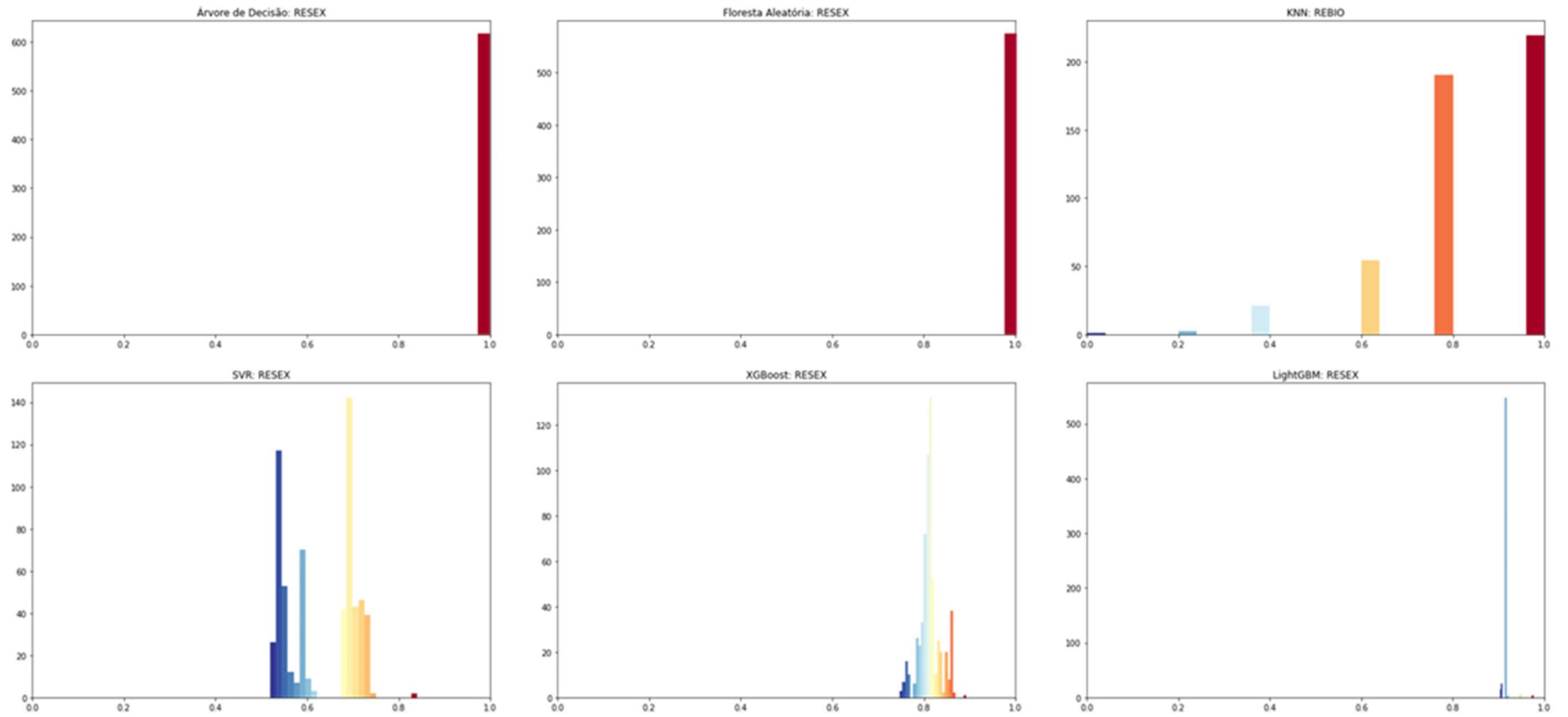
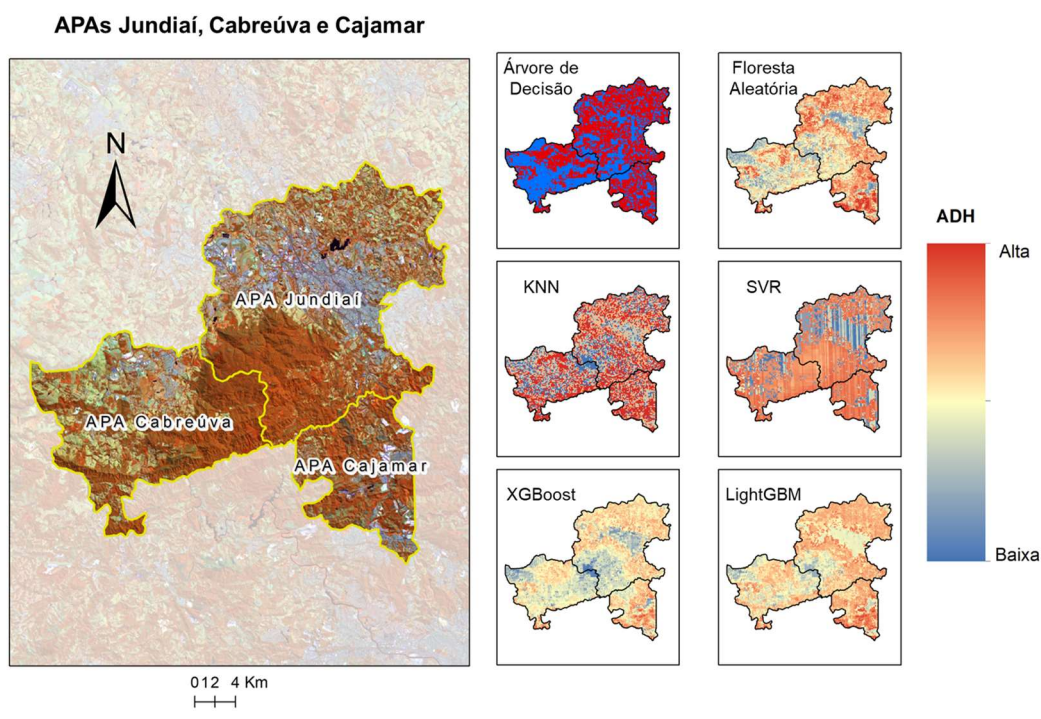


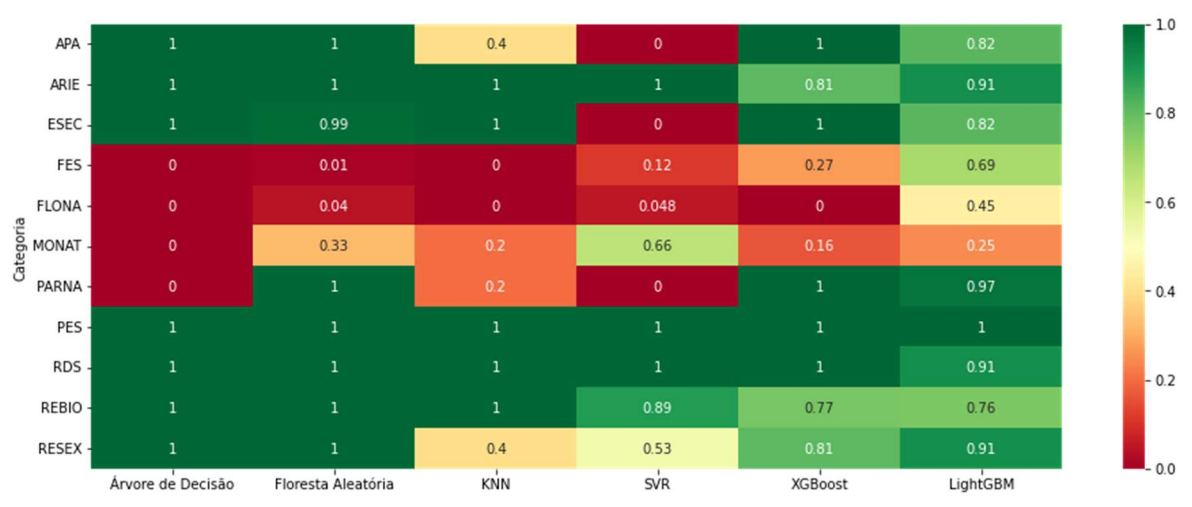


Figura 145 - AdH nas APAs Jundiá, Cabreúva e Cajamar



Para as categorias também são apresentadas as modas (Figura 146), ou seja, o valor mais frequente de AdH nas UCs. Para as APAs, ainda que a AdH fique em torno de 0,6, a maior parte da extensão das unidades tem AdH máxima; este cenário só não é válido para os modelos SVR e KNN. A maior parte da extensão dos PE têm AdH máxima, mesmo que a média esteja em torno de 0,8.

Figura 146 - Moda da ADH em UCs por categoria



Pressupõe-se que as UCs de proteção integral tenham alta AdH nos modelos mais acurados. No entanto, o resultado dos modelos depende da relevância das variáveis. Sendo

assim, as UCs que sejam predominantemente cobertas por Formação Florestal ou Mosaico de Agricultura e Pastagem podem ter os maiores valores de AdH, assim como as unidades que tenham intervalos de Temperatura Mínima, LNA e Precipitação similares à média destas variáveis nos pontos de Presença (ou mais similares aos pontos de Presença do que aos de Pseudo-ausência, ainda que, para estas variáveis o teste de média tenha indicado que os grupos não são diferentes).

O PE do Rio Peixe tem baixa AdH nos modelos Árvore de Decisão (0,41) e Floresta Aleatória (0,32). Nestes modelos, as variáveis mais importantes são Precipitação e Temperatura Mínima, sendo que as estatísticas do PES estão mais próximas da amostra de Pseudo-ausência (Tabela 37). O PE do Rio do Peixe tem 28% da sua extensão coberta por Campo Alagado e Área Pantanosa, variável amostrada com 19% de pontos de Presença e 81% de Pseudo-ausência; os percentuais das variáveis Formação Florestal ou Mosaico de Agricultura e Pastagem são, respectivamente, 37% e 14%. Estas podem ser as razões da AdH deste PE não ter valores muito altos.

Tabela 37 - Estatísticas das variáveis Temperatura Mínima e Precipitação em todo o PE do Rio Peixe e nas amostras de Presença e de Pseudo-ausência

	Temperatura Mínima (°C)			Precipitação (mm)		
	Mínimo	Média	Máximo	Mínimo	Média	Máximo
PE do Rio Peixe	11,4	11,7	11,9	1.129	1.136,7	1.157
Pontos de Presença	2,0	9,7	15,1	1.110	1.485,9	3.360
Pontos de Pseudo-ausência	0,0	9,0	12,1	0,0	1.362,0	2.529

O baixo percentual de pontos de presença também teve impacto na AdH da ESEC de Santa Bárbara (Figura 147), que dispõe de 28% de Formação Savânica e 21% de Formação Campestre variáveis amostradas, respectivamente, com 30% e 18% de pontos de Presença. Situação semelhante ocorre na ESEC de Angatuba (Figura 148), coberta por 20% de Formação Savânica.

Figura 147 - AdH na ESEC Santa Bárbara

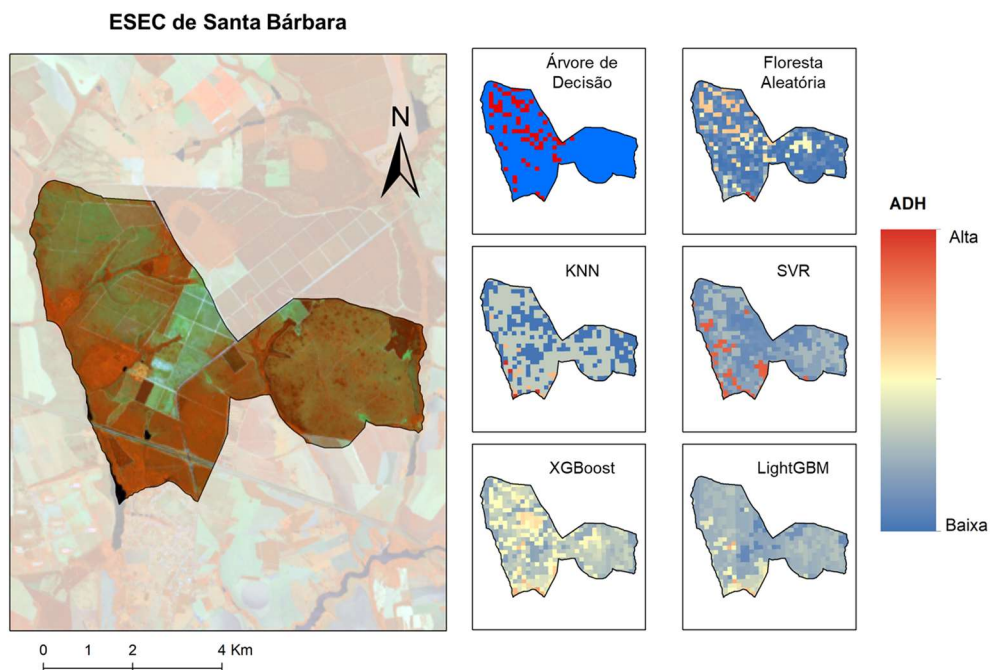
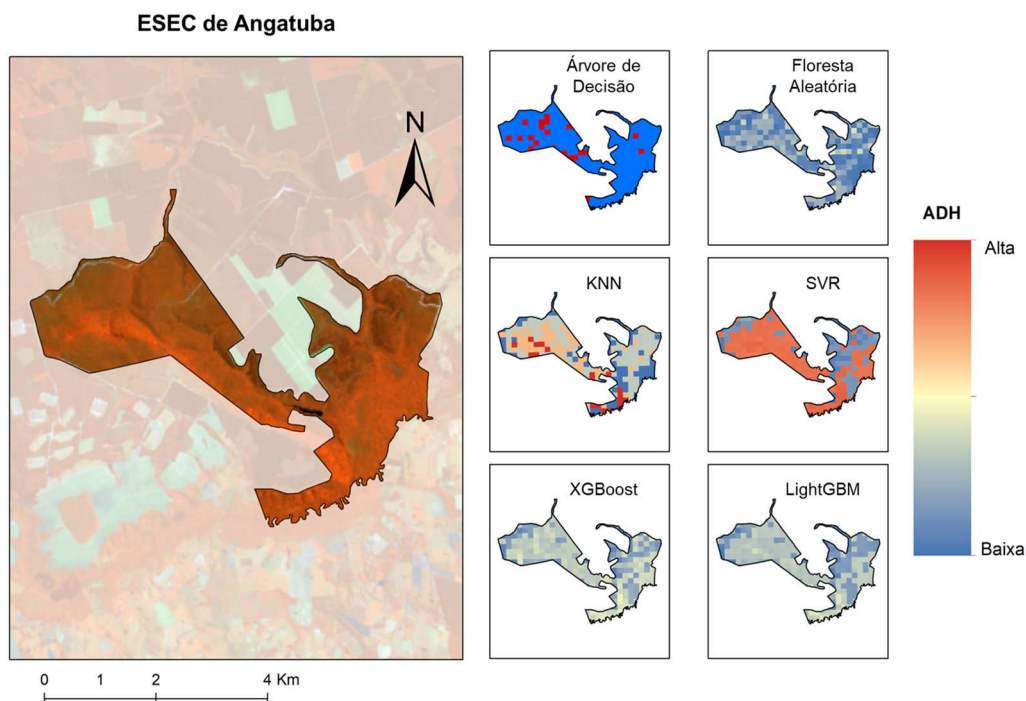


Figura 148 - AdH na ESEC de Angatuba

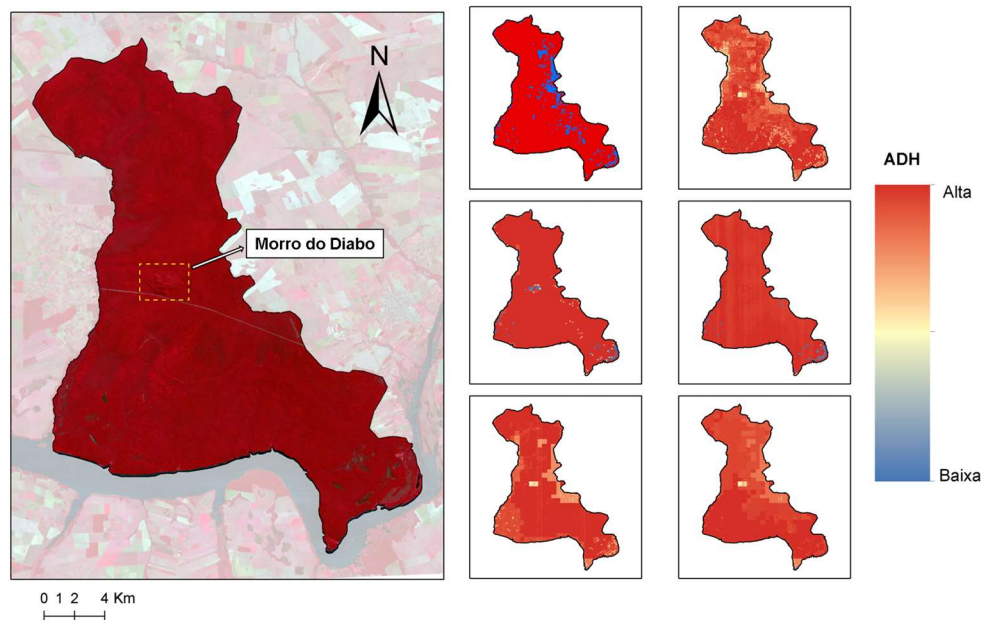


As regiões de morros e montanhas tendem a baixa AdH, o que não necessariamente se estende a toda UC. O MONAT da Pedra Grande, com 159 km<sup>2</sup>, está localizado em parte das Serras de Atibaia, do Una e da Pedra Vermelha, e seu relevo é caracterizado por escarpas e

morros com inclinação do terreno acima de  $25^{\circ}$ <sup>119</sup>. O Parque Estadual de Campos do Jordão, com 83,41 km<sup>2</sup>, está localizado no Planalto de Campos do Jordão, que é a porção mais elevada da Mantiqueira, entre a escarpa da Mantiqueira e as escarpas festonadas do sul de Minas Gerais, com inclinações superiores a  $30^{\circ}$ <sup>120</sup>. Ambos são UCs pequenas em áreas de relevo abrupto e que têm baixa AdH em todos os modelos, exceto o SVR. Embora a Inclinação do Terreno não tenha sido mencionada entre as variáveis mais importantes nos modelos, é a sexta mais relevante na modelagem.

Outro exemplo de importância da Inclinação do Terreno é o Parque Estadual do Morro do Diabo, que é 97% coberto pela classe de cobertura da terra Formação Floresta e contempla um dos maiores remanescentes da Mata Atlântica (Floresta Estacional Semidecidual) no Estado de São Paulo. O local de maior altitude é o Morro do Diabo, um pequeno platô sedimentar de forma escarpada a 600 metros do nível do mar, com inclinação do terreno acima de  $12^{\circ}$ . A Figura 149 mostra que, no morro, a AdH é baixa para todos os modelos, exceto SVR e Árvore de Decisão, destacando-se do seu entorno que tem alta ou média AdH em todo o Parque.

Figura 149 - AdH no PES do Morro do Diabo



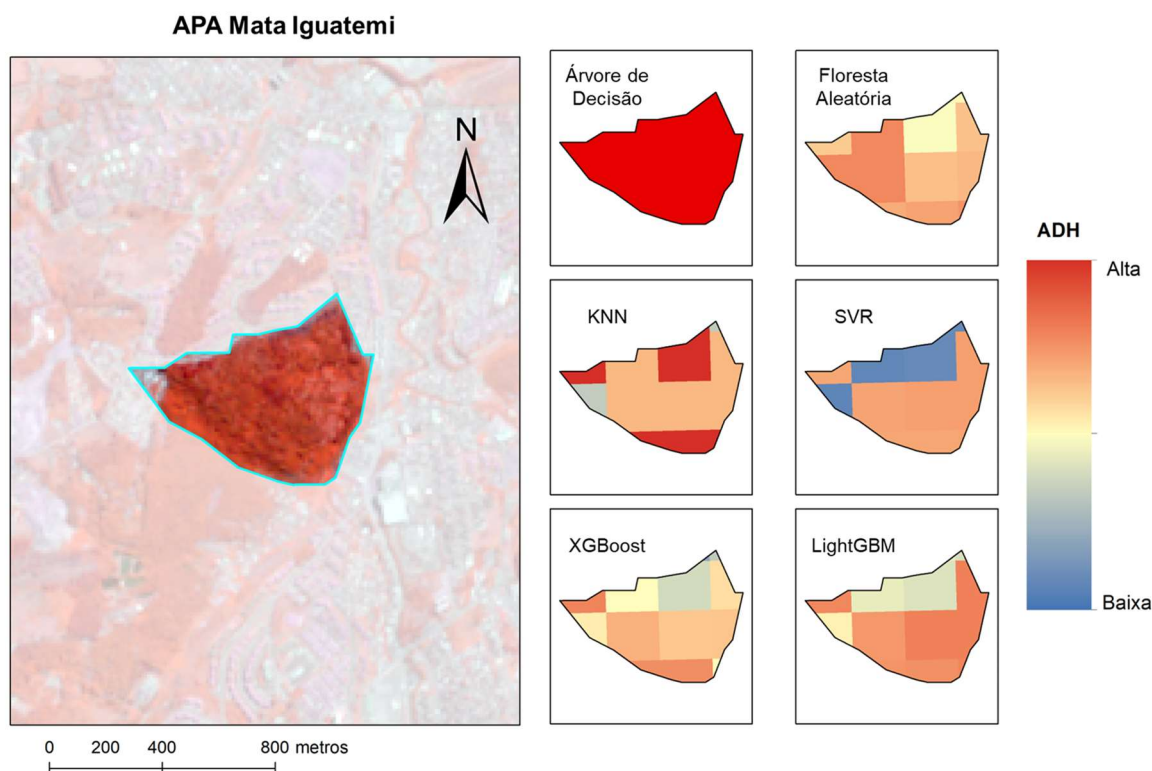
<sup>119</sup> Plano de Manejo do Monumento Natural Estadual Pedra Grande. Disponível em <[http://s.ambiente.sp.gov.br/fundacaoflorestal/planos-manejo/concluidos/mona\\_pedra\\_grande/Plano\\_Manejo\\_Mona\\_Pedra\\_Grande.pdf](http://s.ambiente.sp.gov.br/fundacaoflorestal/planos-manejo/concluidos/mona_pedra_grande/Plano_Manejo_Mona_Pedra_Grande.pdf)>. Acesso em 15 nov. 2022.

<sup>120</sup> Plano de Manejo do Parque Estadual de Campos do Jordão. Disponível em <<https://smastr16.blob.core.windows.net/fundacaoflorestal/2017/02/Diagn%C3%B3stico-e-Planejamento.pdf>>. Acesso em 15 nov. 2022.

As APAs, mais do que proteção à biodiversidade, devem garantir a ordenação da ocupação, de modo que as atividades antrópicas ali praticadas sejam condizentes com a conservação de recursos que justificaram sua criação. Por permitir certo grau de ocupação humana e ser estabelecida em áreas públicas e privadas, as APAs podem abranger grandes extensões de paisagens agrícolas ou urbanas. Então, é admissível que a AdH para mamíferos não seja alta em uma APA ou em grande parte do seu domínio. A APA Mata do Iguatemi tem alta AdH média, variando de 0,51 (SVR) a 1,00 (Árvore de Decisão); trata-se de uma UC de apenas 30 ha que, embora 76% da sua extensão seja coberta por Formação Florestal, está inserida em uma área altamente urbanizada no município de São Paulo (Figura 150), o que não é conveniente para a fauna silvestre nem para as populações humanas no entorno.

O caso da APA Mata Iguatemi remete à importância do tamanho do pixel (ou célula) considerado na modelagem. As áreas protegidas muito pequenas (menores que 4 pixels ou 25 km<sup>2</sup>) podem sofrer efeito da amostragem de pixel e da limitação da resolução espacial dos dados, pois a UC é representada por poucos pixels. Além disso, os pixels nas bordas das UCs podem ser mais representativos do entorno do que da própria unidade. No caso das variáveis numéricas, existe a autocorrelação espacial que atenua imprecisões relacionadas às bordas das UCs e a eventuais erros posicionais (das variáveis ambientais e dos pontos amostrais). Por outro lado, os valores dos pixels das variáveis categóricas não tem uma relação de ordem ou numérica e eventuais erros não podem ser atenuados pela autocorrelação espacial. Sendo assim, a mudança demarcada da cobertura da terra na borda das UCs e erros de classificação podem comprometer a acurácia da modelagem em escala local. Nesta pesquisa, a adequabilidade ambiental é analisada em escala regional (Estado de São Paulo) e entende-se que a AdH deve ser medida de forma padronizada para todas as UCs.

Figura 150 - AdH na APA Mata Iguatemi



As FLONAs, que são unidades de Uso Sustentável, tem como um de seus objetivos, conforme o Decreto 1.298/1994<sup>121</sup>, “promover o manejo dos recursos naturais, com ênfase na produção de madeira e outros produtos vegetais”. Portanto, é justificada a conversão de parte da cobertura nativa para extração e transporte do que é produzido na unidade, contanto que essas alterações estejam previstas no Plano de Manejo e limitem-se “ao estritamente necessário, com um mínimo impacto sobre a paisagem e os ecossistemas”<sup>81</sup>. Essa norma estende-se às FEs, conforme o Artigo 17 do SNUC. As FLONAs e FEs não são necessariamente dominadas por formações naturais florestais. Na FE Edmundo Navarro de Andrade (Rio Claro), a vegetação nativa é resultante do manejo florestal ou da ausência de ocupação de talhões antes florestados com eucalipto, o que permitiu a formação de sub-bosque por regeneração ou chuva de sementes provenientes do entorno<sup>122</sup>. A FLONA de Ipanema (Iperó) contempla em seu Zoneamento a Zona de Manejo Florestal Sustentável, que inclui área de floresta plantada com eucalipto para

<sup>121</sup> Decreto nº 1.298, de 27 de outubro de 1994. Aprova o Regulamento das Florestas Nacionais, e dá outras providências. Disponível em <[http://www.planalto.gov.br/ccivil\\_03/decreto/1990-1994/d1298.htm#:~:text=DECRETO%20N%C2%BA%201.298%2C%20DE%2027,vista%20o%20disposto%20nos%20arts.](http://www.planalto.gov.br/ccivil_03/decreto/1990-1994/d1298.htm#:~:text=DECRETO%20N%C2%BA%201.298%2C%20DE%2027,vista%20o%20disposto%20nos%20arts.)>. Acesso em 21 nov. 2022.

<sup>122</sup> Plano de Manejo - FE Edmundo Navarro de Andrade. Disponível em <<https://www.infraestruturameioambiente.sp.gov.br/fundacaoflorestal/planos-de-manejo/planos-de-manejo-planos-concluidos/plano-de-manejo-fe-edmundo-navarro-de-andrade/>>. Acesso em 15 nov. 2022.

o manejo sustentável de produtos madeireiros<sup>123</sup>. A criação da FLONA de Capão Bonito ocorreu no fim da década de 1960, quando o plantio de pinus e eucalipto fez parte de uma estratégia de desenvolvimento por meio de incentivos fiscais para plantio de florestas para garantir a cadeia produtiva da madeira; 58% da UC é coberta por Silvicultura<sup>124</sup>. Outros exemplos são a FES de Avaré I e II (Figura 151) e FE Piraju (Figura 152) que compreendem 75% e 61% da sua extensão em Silvicultura. Estes percentuais foram calculados a partir dos dados do MapBiomas - 2020 com resolução espacial de 30 metros. Ou seja, a composição florística nas FLONAS não é necessariamente de espécies nativas, e as espécies de porte arbóreo podem ser exóticas com finalidade econômica. Isto não significa que estas UCs sejam inadequadas para manter mamíferos não voadores. Os Planos de Manejo dessas unidades mencionam a ocorrência de mamíferos de médio e grande porte na sua composição faunística. É importante ressaltar que apenas 10,6% das amostras incidentes em áreas de Silvicultura são pontos de Presença e essa situação impacta no resultado dos modelos.

A proporção de pontos de Presença para a variável Silvicultura pode ter sido a causa da baixa AdH média na APA Corumbataí, Botucatu, Tejupá. No perímetro Botucatu, a cobertura da terra dominante é Silvicultura (24%) e a AdH ficou entre 0,34 (Árvore de Decisão) e 0,49 (SVR) (Figura 153).

O Parna da Serra da Bocaina (PNSB), único Parna avaliado, apresentou baixa ADH média (0,43 a 0,47), com exceção do SVR que resultou em AdH média de 0,72. O PNSB é interestadual e parte da sua extensão não foi modelada por estar no estado do Rio de Janeiro. O Parque está localizado em uma área onde a Temperatura Mínima, variável ambiental com maior pontuação SHAP, é muito baixa e atípica no estado de São Paulo (1,4 a 15°C). Além disso, há apenas uma amostra de Presença na parte paulista do PNSB.

As Figuras 154 a 157 mostram a ADH média para todas as UCs.

---

<sup>123</sup> Plano de Manejo - Floresta Nacional de Ipanema. Disponível em <[https://www.gov.br/icmbio/pt-br/assuntos/biodiversidade/unidade-de-conservacao/unidades-de-biomas/mata-atlantica/lista-de-ucs/flona-de-ipanema/arquivos/pm\\_flona\\_de\\_ipanema\\_vol\\_ii\\_planejamento.pdf](https://www.gov.br/icmbio/pt-br/assuntos/biodiversidade/unidade-de-conservacao/unidades-de-biomas/mata-atlantica/lista-de-ucs/flona-de-ipanema/arquivos/pm_flona_de_ipanema_vol_ii_planejamento.pdf)>. Acesso em 15 nov. 2022.

<sup>124</sup> Plano de Manejo - Floresta Nacional de Capão Bonito. Disponível em <[https://www.gov.br/icmbio/pt-br/assuntos/biodiversidade/unidade-de-conservacao/unidades-de-biomas/cerrado/lista-de-ucs/flona-de-capao-bonito/arquivos/plano\\_de\\_manejo\\_flona\\_de\\_capao\\_bonito\\_vol1\\_diagnostico.pdf](https://www.gov.br/icmbio/pt-br/assuntos/biodiversidade/unidade-de-conservacao/unidades-de-biomas/cerrado/lista-de-ucs/flona-de-capao-bonito/arquivos/plano_de_manejo_flona_de_capao_bonito_vol1_diagnostico.pdf)>. Acesso em 15 nov. 2022.

Figura 151 - AdH na FE Avaré I e II

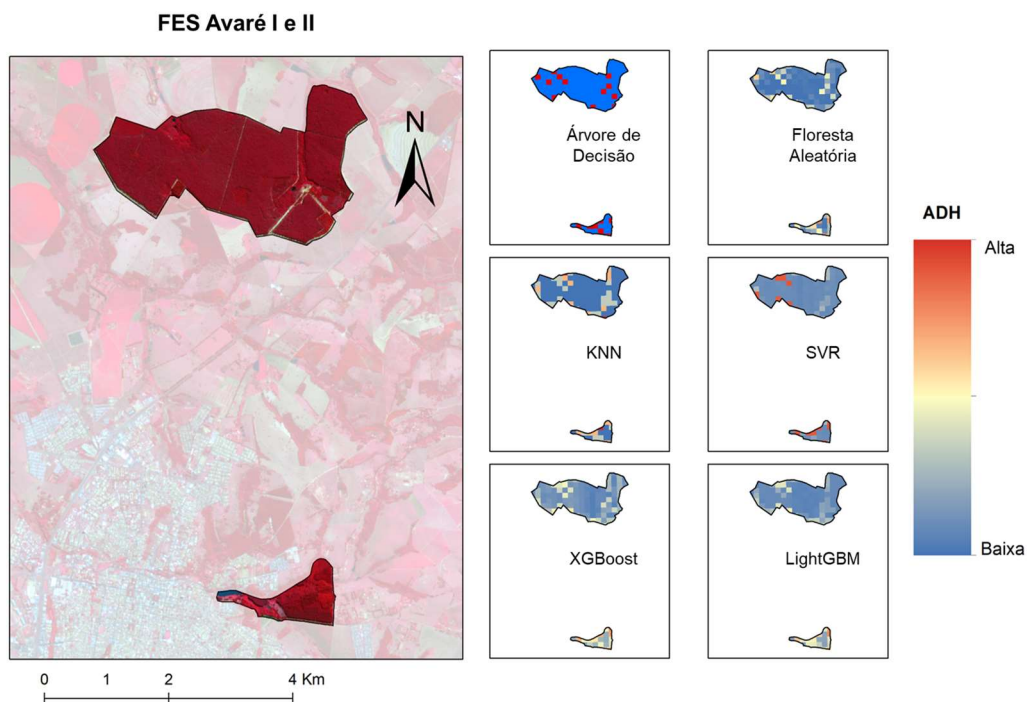


Figura 152 - AdH na FE Piraju

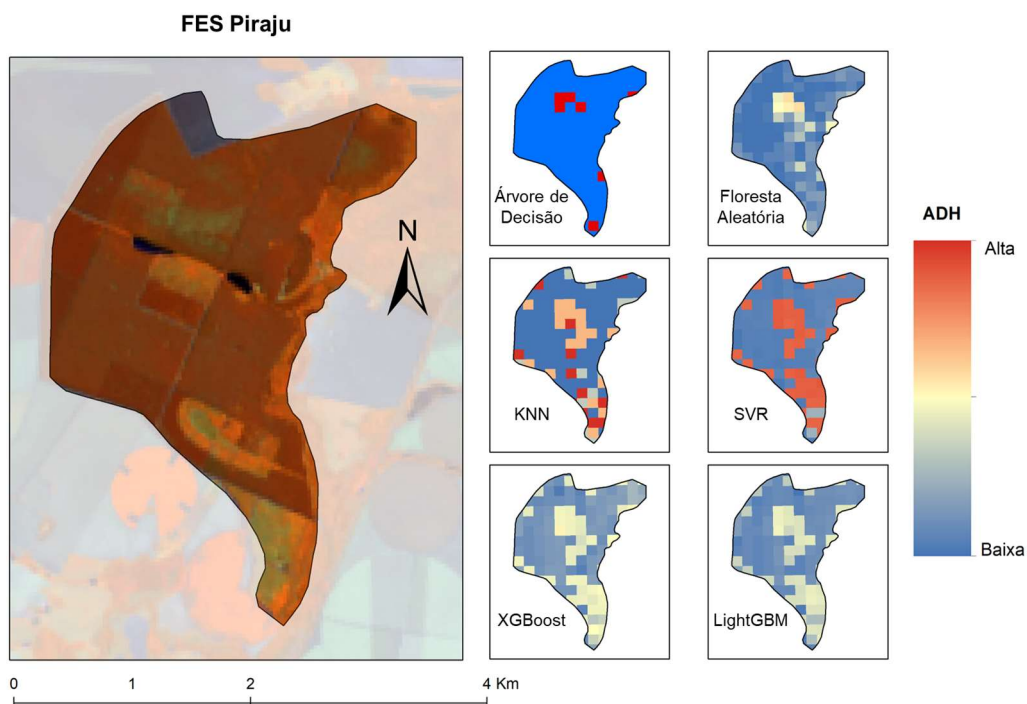




Figura 153 - AdH na APA Corumbataí, Botucatu, Tejuπά (Perímetro Corumbataí) e na APA dos Rios Piracicaba e Juqueri-Mirim

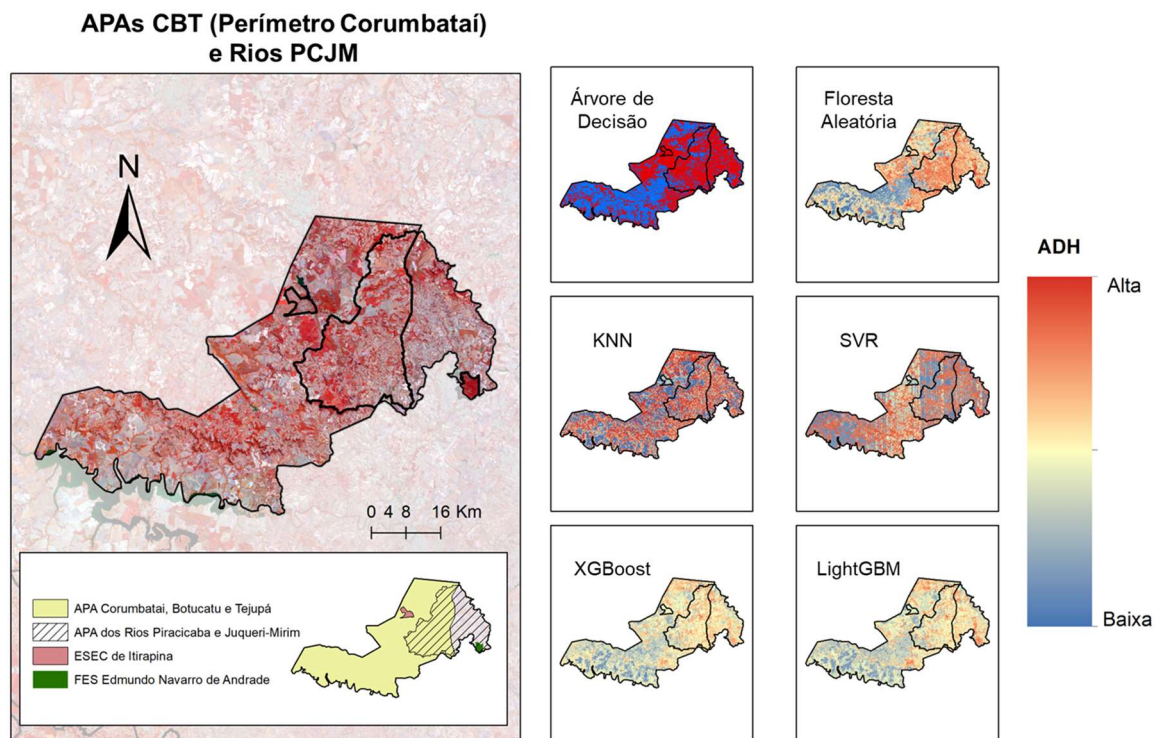


Figura 154 - AdH média nas UCs (I)

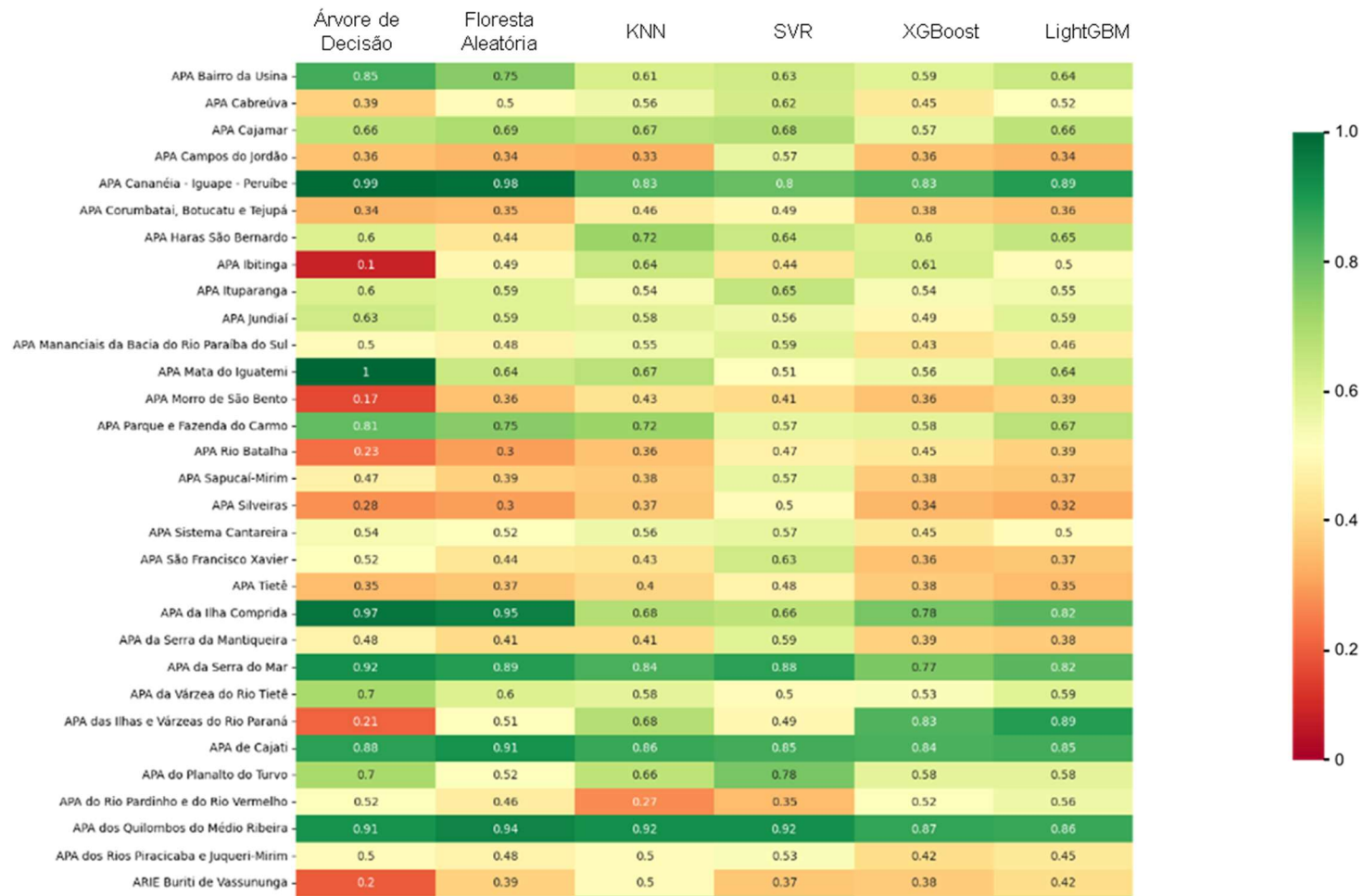


Figura 155 - ADH média nas UCs (II)

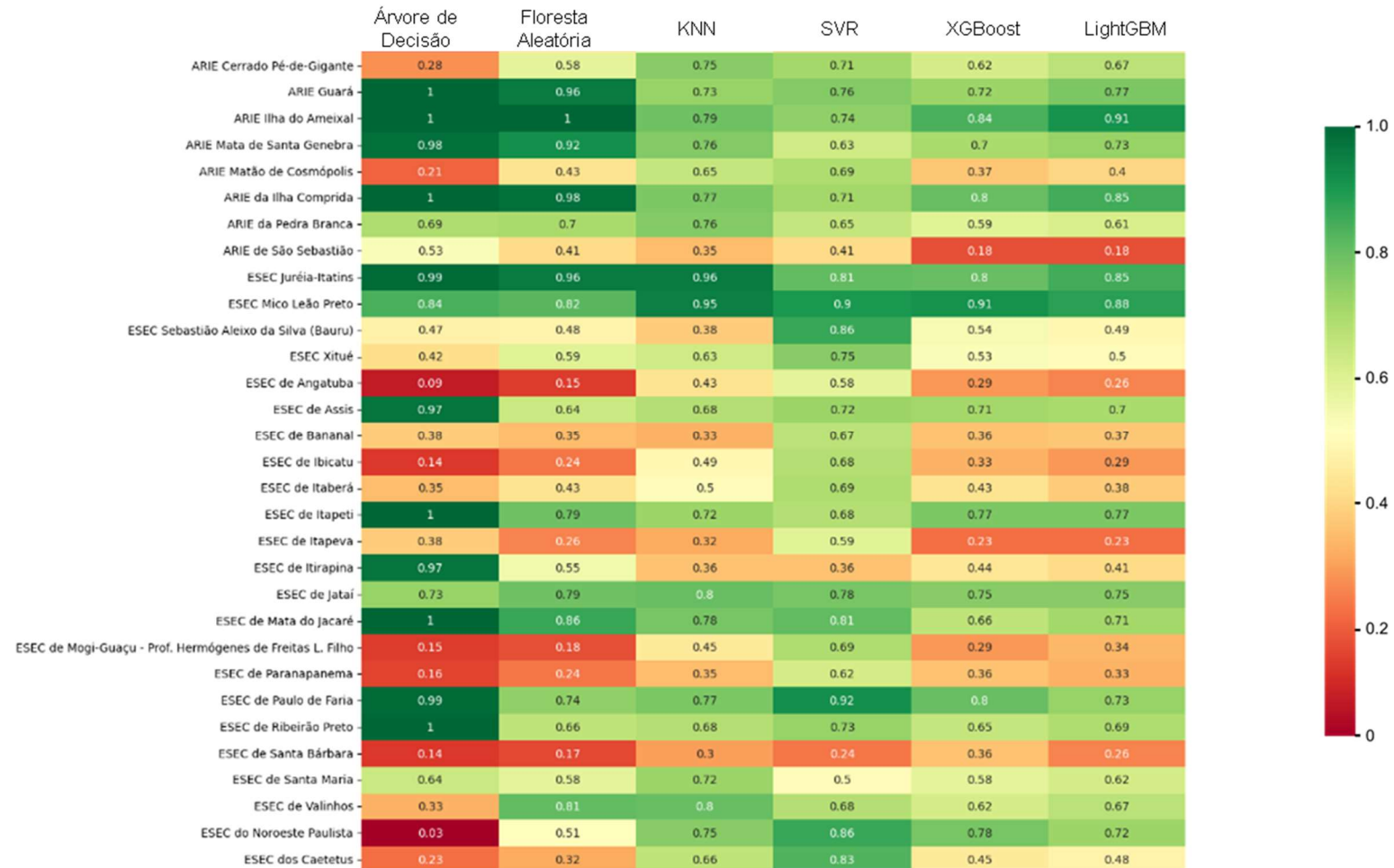


Figura 156 - AdH média nas UCs (III)

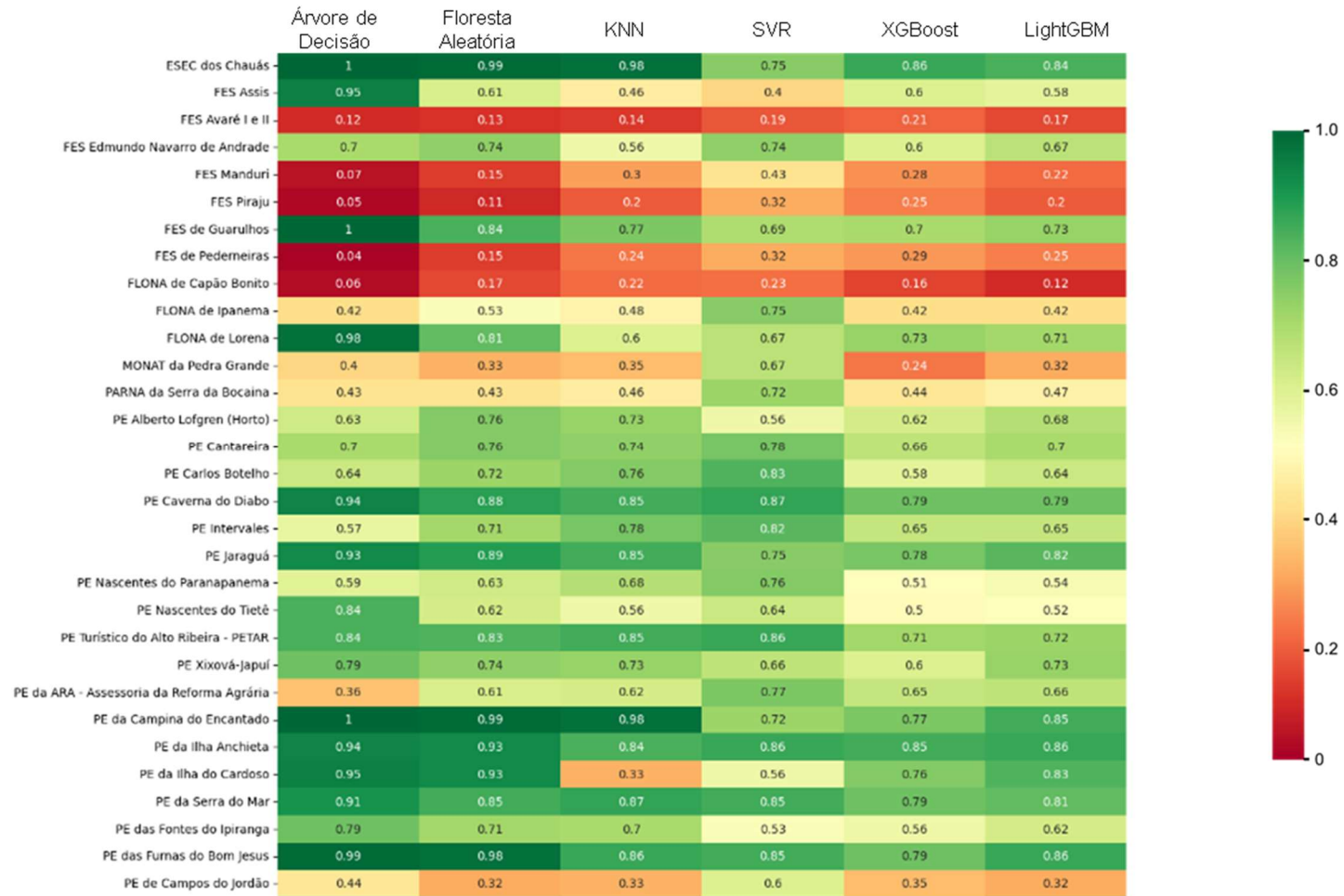
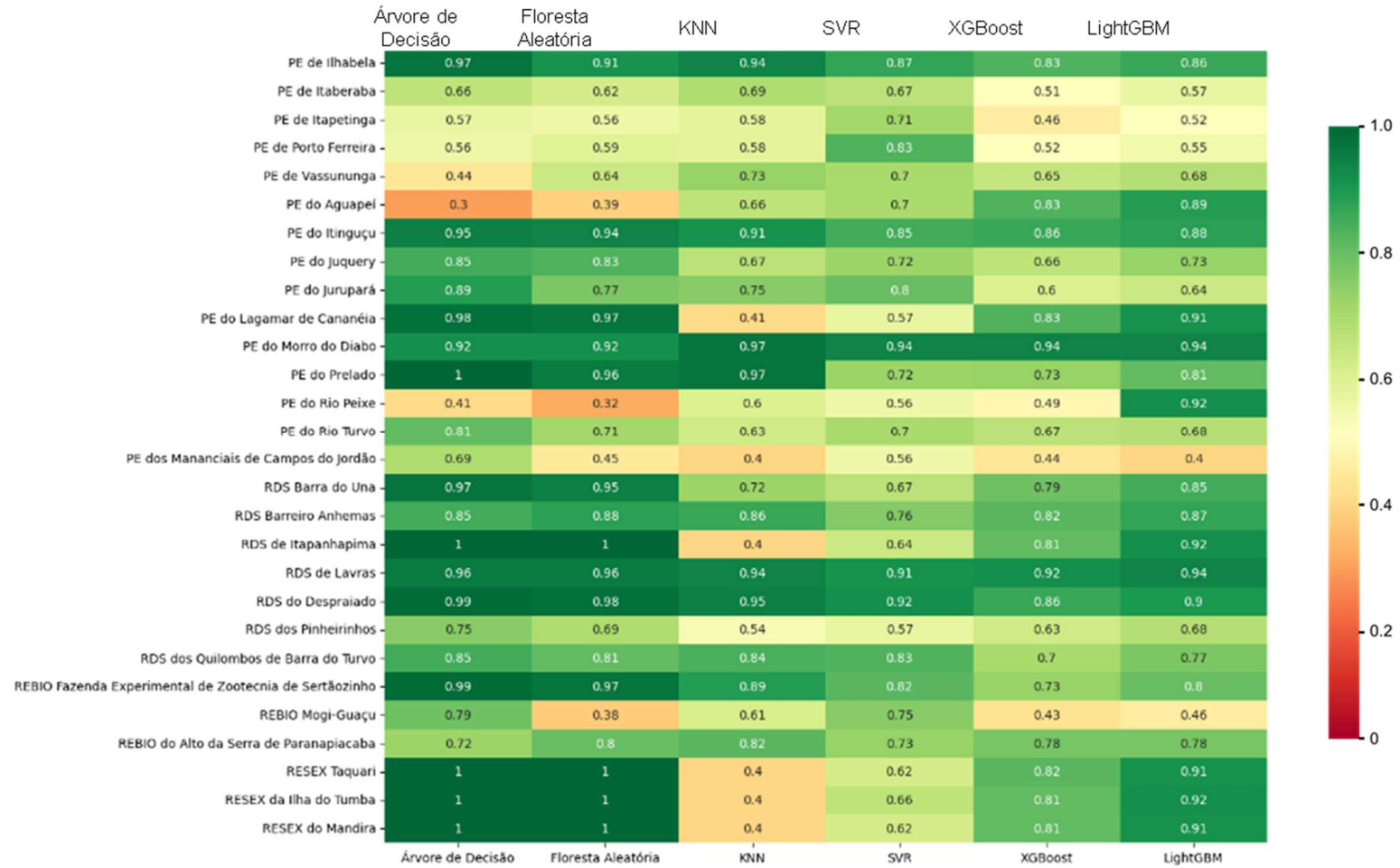


Figura 157 - ADH média nas UCs (IV)



#### 4.10 CONSIDERAÇÕES SOBRE A ABRANGÊNCIA ESPACIAL EM MAHs

Um aspecto importante na modelagem da distribuição espacial de espécies é a definição da área de estudo. Considerar toda a área de vida da espécie pode implicar no processamento de áreas muito extensas, principalmente no caso de espécies não endêmicas ou muito generalistas, o que pode demandar recursos computacionais de alto desempenho ou grande quantidade de tempo para sua conclusão, comprometendo a reprodutibilidade da pesquisa (ALSTON; RICK, 2021; KONKOL et al., 2019; PENG, 2018). Uma solução possível é aumentar a área de amostragem, aumentando o número de pontos amostrais e, assim, incrementar a variedade de ambientes. Todavia, seria necessário diminuir o tamanho do pixel dos planos de informação das variáveis ambientais, o que seria equivalente a diminuir a resolução espacial do dado, prejudicando a acurácia do modelo. A Tabela 38 exemplifica alguns estudos de MAHs realizados em diferentes escalas. Quase todos os trabalhos listados restringem a modelagem à área de interesse ou apenas incluem uma zona de buffer, dispensando a informação sobre a área de vida das espécies.

Tabela 38 - Exemplos na literatura sobre MAH

<b>Autor</b>	<b>Táxon</b>	<b>Modelo</b>	<b>Local</b>	<b>Tamanho do pixel</b>
Bramuth; Paprocki, 2022	Efemeroptera: <i>Campylocia burmeisteri</i>	MaxEnt	Treinado na Reserva da Biosfera da Serra do Espinhaço no Estado de Minas Gerais e aplicado para Região Sudeste do Brasil (300 km de buffer dos registros de ocorrência)	Não consta. Usou dados do WorldClim com 30' de resolução espacial.
Hurtado; Burton, 2022	Felídeo: <i>Puma concolor</i>	GLM e ENFA	Tumbes (Peru)	Não consta. Usou apenas dados com 30 metros de resolução espacial.
Kafash et al., 2022	Quirópteros: <i>Pipistrellus kuhlii</i> , <i>Miniopterus pallidus</i> , <i>Rhinopoma</i>	MaxEnt	Irã	Não consta. Usou dados ambientais de até 1 km de

Autor	Táxon	Modelo	Local	Tamanho do pixel
	<i>microphyllum</i> e <i>Rousettus aegyptiacus</i>			resolução espacial.
Rather et al., 2021	Ursídeo: <i>Melursus ursinus</i>	RF	Reserva de Tigres de Bandhavgarh, na Índia e sua Zona de Amortecimento (1.536 km <sup>2</sup> )	90 m
Su et al., 2021	Carnívoros: <i>Ursus thibetanus</i> e <i>Ailurus fulgens</i>	MaxEnt e GARP	Parque Nacional Makalu Barun, no Himalaia do Nepal (1.500 km <sup>2</sup> ) e Zona de Amortecimento (830 km <sup>2</sup> )	100 metros
Chahad-Ehlers et al., 2018	Dipteros: <i>Aedeomyia squamipennis</i> , <i>Aedes fluviatilis</i> , <i>Anopheles albitarsis</i> , <i>Culex coronator</i> , <i>Culex nigripalpus</i> , <i>Culex quinquefasciatus</i> , <i>Culex saltanensis</i> e <i>Mansonia titilans</i>	MaxEnt	Amostragem: não específica  Modelagem: América do Sul.	Não consta. Usou dados do WorldClim com 30' de resolução espacial.
Jarnevich et al., 2018	Squamata: <i>Salvator merianae</i> , <i>Salvator rufescens</i> e <i>Tupinambis teguixin</i>	GLM, MARS, BRT, RF e MaxEnt	Treinado na América do Sul e aplicado para América do Norte	30'
Lauria et al., 2015	Elasmobrânquios: <i>Raja clavata</i> , <i>Raja oxyrinchus</i> , <i>Raja melitensis</i> , <i>Squalus blainvillei</i> , <i>Dalatias licha</i> , <i>Chimaera monstrosa</i> , <i>Mustelus mustelus</i> , <i>Torpedo marmorata</i> e <i>Centrophorus granulosus</i>	GAM	Mar Mediterrâneo central	Não consta. Usou variáveis de diferentes resoluções espaciais.
Yannick et al.,	Cervídeo:	GLM	Quebec e Labrador	Não consta

Autor	Táxon	Modelo	Local	Tamanho do pixel
2014	<i>Rangifer tarandus</i>		(Canadá)	
Kalle et al., 2013	Carnívoros: <i>Felis chaus</i> , <i>Viverricula indica</i> , <i>Paradoxurus</i> <i>hermaphroditus</i> , <i>Herpestes vitticollis</i> , <i>Herpestes smithii</i> , <i>Herpestes edwardsii</i>	MaxEnt	Reserva de Mudulai Tiger, na Índia (321 km <sup>2</sup> ) com 2 km de buffer	1 km

É importante considerar áreas mais extensas do que a região de interesse no caso de áreas muito pequenas com características ambientais homogêneas que não refletem a área de vida da espécie. O Estado de São Paulo é bastante diverso quanto às variáveis ambientais consideradas nesta pesquisa. Problemas com a amostragem não são restritos à extensão da área de estudo, mas devem ponderar também a significância da amostragem em relação aos locais potencialmente adequados para as espécies. Por exemplo, algumas coberturas da terra em que as espécies ocorrem podem deixar de ser amostradas em decorrência de viés amostral. Áreas de difícil acesso, como áreas particulares, de floresta densa ou de alta inclinação do terreno podem ser sub-amostradas.



## 5 CONCLUSÕES

A Adequabilidade de Habitat em uma UC deve ser condizente com a categoria a qual pertence e como os propósitos da sua criação, isto é, a AdH média não precisa ser alta em todas as áreas protegidas. No Estado de São Paulo, as APAs, por ser uma categoria com menos restrições de uso da terra, têm AdH média similares às de áreas não protegidas por UCs. As demais UCUS, embora tenham algum grau de antropismo, têm boa AdH média para a mastofauna; as UCPI, como previsto na hipótese desta pesquisa, são as unidades com melhor AdH. Embora as classes de cobertura da terra Formação Florestal e Mosaico de Agricultura e Pastagem tenham tido alta importância nos modelos, dada a alta proporção de pontos de Presença nestas classes, outras variáveis ambientais (Temperatura Mínima, Luzes Noturnas Artificiais e Precipitação) também influenciaram os resultados da modelagem. Os modelos apontam três principais áreas no Estado de São Paulo com alta AdH: 1) Noroeste do estado, onde as paisagens são fortemente marcadas pela agricultura; 2) Nordeste do estado, onde a produção de cana-de-açúcar é dominante, mas onde ainda há remanescentes de Cerrado não protegidos por UCs; e 3) domínio geomorfológico da Província Costeira Sul do Estado, onde há justaposição de UCs de várias categorias.

Os métodos de Aprendizado de Máquina baseados em árvore (Floresta Aleatória e Árvore de Decisão) tiveram melhor desempenho do que os métodos baseados em distâncias (KNN) e métodos de maximização de margens (SVR). A modelagem de processos ecológicos parece ajustar-se melhor à separação de grupos com divisões sucessivas (árvores) do que à tentativa de discriminar estes grupos em um espaço de atributos. Embora esses métodos permitam trabalhar com dados complexos e com muitas variáveis relacionados às complexas relações entre as espécies e seu ambiente, parecem mais identificar padrões no comportamento dos dados do que detectar uma relação das variáveis ambientais com a ocorrência de espécies.

A ocorrência da espécie sinaliza mas não define a Adequabilidade do Habitat. Os conceitos relacionados à distribuição de espécies devem ser bem definidos para viabilizar o aprimoramento da modelagem. Esta compreensão e o entendimento das limitações metodológicas proporciona o aprimoramento da sequência de etapas na modelagem ecológica como o estabelecimento de critérios de filtragem de dados, definição da escala de estudo e, em um momento adiante, beneficia tomadas de decisão a partir dos seus resultados. Considerar diferentes modelos orienta sobre quanto é uma boa acurácia para uma mesma área de estudo e reforça a identificação de padrões de adequabilidade nas UCs.

Ou seja, o produto da modelagem não é irrefutável e determinante, mas deve ser ponderado nas tomadas de decisão em relação à Conservação da Biodiversidade e à Gestão da Paisagem. Por exemplo, a informação de que a ocorrência de espécies é frequente em ambientes antropogênicos justifica a demanda por atividades de Educação Ambiental junto aos produtores rurais e proprietários de terra. Os mapas de Adequabilidade de Habitat podem também orientar a reintrodução de espécies para conservação *in situ* e o estabelecimento de estratégias de restauração. Nessa premissa, vale ainda ressaltar que um ecossistema agrícola (e até mesmo urbano) pode ser usado para o trânsito de indivíduos ou populações de mamíferos que, eventualmente, podem ser observados nesses locais. Contudo, os ecossistemas agrícolas não são suficientes para suprir todos os recursos para manutenção da vida silvestre, em termos de alimentação, segurança e conforto térmico, principalmente no caso de espécies especialistas. Esse uso de paisagens agrícolas é decorrente da alta fragmentação florestal que suprimiu o habitat original da mastofauna no Estado de São Paulo.

Os modelos são orientadores que dão indicativos da situação da adequabilidade em uma UC. Entende-se que, se aquele local é apropriado como habitat de mamíferos, pode também estar em condições de abrigar outras espécies co-ocorrentes. Os modelos também contribuem para localizar áreas que, embora não sejam protegidas por UCs, possam ser consideradas na criação de novas áreas protegidas, especialmente de categorias que incluem propriedades privadas e que não requerem desapropriação de terras. Unidades em que é esperada alta Adequabilidade de Habitat, como as UCPIs, mas que não correspondem a essa condição, também podem ser identificadas e então, o gestor responsável pela Unidade, em parceria com o modelador, pode investigar sobre as razões da baixa AdH.

A sequência de etapas e as colocações a respeito da importância das variáveis, limitações dos dados disponíveis e desempenho dos modelos podem orientar futuros trabalhos sobre a modelagem espacial de espécies. Além do resultado em si, isto é, do reconhecimento de como está a adequabilidade de uma determinada área para as espécies estudadas, a modelagem ajuda a compreender aspectos sobre como os modelos se comportam diante da qualidade dos dados. Este ponto pode orientar sobre como refinar a coleta de dados.

Os MAHs são baseados em dados de campo e dados de sensores orbitais, ou seja, são empíricos e potencialmente refletem a realidade. Porém, existem desafios para aprimorar a modelagem, sendo um deles a disponibilidade de dados confiáveis em escala apropriada. Nem sempre os recursos de entrada são ideais e existem limitações quanto a resolução espacial das

variáveis ambientais, o desenho e tamanho amostral dos pontos de Presença, poder computacional para o processamento de dados de alta resolução que cobrem grandes áreas de interesse e discriminação estatística entre as amostras de Presença e de Pseudo-ausência. Este é um ponto inerente aos estudos de Síntese Ecológica. Os dados de ocorrência de espécies, por exemplo, embora sejam abundantes e com disponibilidade, não são integralmente úteis e precisam ser selecionados dependendo da abordagem do estudo. Além da confiabilidade, a insuficiência amostral é um gargalo metodológico. Seria apropriado um esforço de amostragem que contemplasse diferentes fitofisionomias e paisagens baseado em um desenho amostral com padrão estratificado, ou seja, na medida do possível aumentar o número de observações e, indiretamente, inferir pontos de ausência de espécies. A mesma proposta pode ser aplicada às variáveis numéricas, considerando a discretização em vários intervalos da variável, ou pelo menos em maiores amplitudes de valores das variáveis. Ainda que essa demanda possa ser inviável em grande escala, pode ser considerada para melhorar a amostragem de MAHs em escala regional.

O mapeamento da cobertura da terra pelo MapBiomas tem viabilizado diversos trabalhos nas mais variadas áreas do conhecimento (Meio Ambiente, Agricultura, Economia, Saúde e outros), por ser fundamentado em metodologia consolidada, ser produzido de forma sistemática e estar disponível em escala suficiente para estudos regionais. Seria conveniente a estruturação de um sistema em que os usuários apontassem eventuais inconsistências no mapeamento, no contexto da Ciência Cidadã.

Há pontos relevantes sobre o uso econômico de recursos naturais e uso do solo nas áreas protegidas, principalmente em relação às paisagens agrícolas. 1) A totalidade da área protegida não é coberta por florestas ou outras formações nativas. As APAs no estado de São Paulo são intensamente cobertas por atividades agrícolas, seja por Agricultura, Pecuária e/ou Silvicultura e, em alguns casos, até mineração. Mesmo as UCPI não são integralmente cobertas por formações nativas já que seu zoneamento prevê que parte da sua extensão seja destinada à infraestrutura da UC. 2) As UCs podem ser economicamente produtivas e seria incorreto considerar toda a sua extensão para contabilizar áreas indisponíveis para produção agrícola. As APAs são permissivas quanto ao uso do solo para agricultura, pastagem e silvicultura e as demais categorias de uso sustentável permitem a extração sustentável dos recursos naturais, seja para autoconsumo ou para fins econômicos. 3) Áreas com Mosaico de Agricultura e Pastagem podem ser convenientes para manter populações de mamíferos de médio e grande portes, visto que estas áreas sustentam populações de pequenos mamíferos que podem servir

como presas para espécies carnívoras. No entanto, não devem constituir extensas monoculturas e requerem a manutenção de formações nativas em APPs e RLs que, por sua vez, fornecem serviços ecossistêmicos para a própria produção agrícola e propiciam recursos alimentares, água, regulação da temperatura e local de descanso para a mastofauna, além de manter a conectividade entre fragmentos maiores e permitir o fluxo gênico em outros taxa. E, ainda, tais paisagens agrícolas devem limitar-se às APAs ou existir em pequena escala nas demais categorias de UCUS. Estes apontamentos são particularmente importantes dada a crescente demanda por biocombustíveis no Brasil e no mundo.

O Plano de Manejo é o elemento central na gestão de áreas protegidas e deve estabelecer os objetivos de gestão e integrar o plano diretor dos municípios que abrangem a UC no sentido de ordenar o uso do território na Zona de Amortecimento das unidades. Usualmente, os Planos estabelecem apenas recomendações gerais e não regras específicas sobre o uso dos recursos no entorno das UCs. Estas regras seriam particularmente importantes nos casos das UCPI que, mesmo sendo mais restritivas quanto ao uso, estão sujeitas à degradação. No Estado de São Paulo, onde a paisagem é predominantemente coberta por cana de açúcar, é importante definir diretrizes mais específicas sobre as Zonas de Amortecimento para que essas áreas sejam efetivamente um instrumento de proteção da biota que habita as áreas protegidas e permeia o seu entorno. Há casos, como a ARIE Buriti de Vassununga, em que a cobertura da terra já não era condizente com seus objetivos desde a época do decreto de sua criação. O documento deve mencionar atividades conflitantes no interior e entorno da UC e apresentar estratégias de minimização de seus eventuais impactos.

A metodologia aplicada nesta pesquisa segue os princípios da reprodutibilidade e pode ser adotada para outras áreas de interesse e incorporar novas variáveis. Os códigos, dados de entrada e análises passam a fazer parte de uma Ciência Aberta e podem ser verificados. A transparência da pesquisa proporcionada pela reprodutibilidade torna o conhecimento acessível e democrático e contribui para que a gestão do Meio Ambiente seja pautada na Ciência e em dados e análises coerentes, confiáveis e acessíveis.

Para pesquisas futuras recomenda-se considerar os seguintes apontamentos:

- 1) Considerar outras variáveis ambientais como distância de corpos d'água e da área urbana. O cálculo da distância não é meramente a distância euclidiana entre dois pontos, mas deve considerar a permeabilidade da paisagem, relevo, barreiras físicas (corpos

d'água, rodovias, áreas altamente urbanizadas). A definição de uma distância mínima pode ser um critério para filtragem espacial dos pontos de Presença de espécies.

- 2) Outras variáveis podem ser incorporadas ao conjunto de fatores ambientais que alimentam os modelos, como o tamanho dos fragmentos de vegetação nativa, conectividade e medidas de complexidade da paisagem. Produtos derivados de sensores multiespectrais, hiperespectrais, Radar (*Radio detection and ranging*) e LiDAR (*Light Detection and Ranging*) também podem aprimorar o conjunto de variáveis preditoras. Além dessas, vale considerar também a compreensão do entorno da Unidade de Conservação no contexto da paisagem e não apenas a informação isolada do pixel.
- 3) Os modelos baseados unicamente em variáveis climáticas devem ser repensados, visto que generalizam a predição com a possibilidade de ampliar a área de ocorrência ou adequabilidade a regiões para as quais a dispersão (e, portanto, a ocupação) é restrita. Podem ser úteis para a definição de áreas adequadas para a reintrodução de espécies, mas para esse objetivo devem ser usados com cautela. Neste caso, a modelagem de relações interespecíficas pode ser considerada.
- 4) Embora algumas variáveis possam não interferir diretamente na ocupação do habitat por uma espécie, podem estar relacionadas à existência de recursos alimentares. Por exemplo, a orientação de vertentes, tipo de solo, disponibilidade hídrica e a luminosidade podem determinar a existência em um determinado local de espécies vegetais que são recursos alimentares. Esta temática complementa a modelagem de relações interespecíficas citada no item anterior, especialmente no caso de espécies especialistas que consomem recursos que se desenvolvem em locais com características peculiares.
- 5) A modelagem multitemporal (multi-sazonal) de adequabilidade também pode ser considerada, visto que variáveis ambientais como temperatura, luminosidade e temperatura estão relacionadas à aspectos comportamentais e fisiológicos como migração e reprodução.
- 6) Neste trabalho, as variáveis categóricas com maior incidência de pontos amostrais tiveram maior importância nos modelos. Propõe-se avaliar se a amostragem estratificada, tentando manter o mesmo número de amostras por cobertura da terra, interfere no resultado dos modelos, principalmente KNN e SVR que deram maior

importância às variáveis categóricas. Ainda, pode ser avaliado se há influência das anomalias na importância dos modelos.

- 7) Em relação à filtragem espacial, os diferentes valores de  $\epsilon$  considerados no DBSCAN e, portanto, o tamanho amostral, podem ser considerados para verificar sua relação com a acurácia dos modelos.
- 8) O tamanho do pixel tem relação inversa à quantidade de pixels. Sendo assim, é considerável a possibilidade de que exista uma relação entre a proporção tamanho amostral / número total de pixels e acurácia do modelo. Não é apenas avaliar a acurácia como uma função do tamanho amostral, mas introduzir o tamanho do pixel no cálculo da acurácia e no poder de generalização do modelo.
- 9) Os métodos *Positive and Unlabeled Learning* e *Isolation Forest* podem ser considerados como uma alternativa para gerar o plano de fundo de Pseudo-ausência em comparação ao OC-SVM.
- 10) A criação de RPPNs, que não fizeram parte deste estudo, não requer uma área mínima, embora seja necessário que disponham de atributos para seu reconhecimento. Os modelos aqui produzidos podem ser usados para avaliar a AdH nesta categoria, já que podem sofrer efeitos de borda provenientes do entorno.

## REFERÊNCIAS

- ABRA, F. D.; HUIJSERM, P.; MAGIOLI, M.; BOVO, A. A. A.; FERRAZ, K. M. P. M. B. An estimate of wild mammal roadkill in São Paulo state, Brazil. **Heliyon**, v. 7, n. 1, e06015, 2021.
- AFONSO, C. M. **A paisagem da Baixada Santista: urbanização, transformação e conservação**. São Paulo: Editora da Universidade de São Paulo / FAPESP, 2006. 309 p.
- AI, Z.; ISHIHAMA, F.; HANASAKI, N. Mapping Current and Future Seawater Desalination Plants Globally Using Species Distribution Models. **Water Resources Research**, v. 58, n. 7, e2021WR031156, 2022.
- AIELLO-LAMMENS, M. E.; BORJA, R. A.; RADOSAVLJEVIC, A.; VILELA, B.; ANDERSON, R. P. spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. **Ecography**, v. 38, p. 541-545, 2015
- ALKATHIRI, M.; JHUMMARWALA, A.; POTDAR, M. B. Multi-dimensional geospatial data mining in a distributed environment using MapReduce. **Journal of Big Data**, v. 6, 82, 2019.
- ALMEIDA, F. F. M. Fundamentos geológicos do relevo paulista. **Boletim do Instituto Geográfico e Geológico**, v. 41, p. 167-263, 1964.
- ALMEIDA, E. **Econometria Espacial Aplicada**. Campinas: Ed. Alínea, 2012. 498 p.
- ALSTON, J. M.; RICK, J. A. A Beginner's Guide to Conducting Reproducible Research. **Bulletin of the Ecological Society of America**, v. 102, n. 2, e01801, 2021.
- ALTMAN, N. S. An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. **The American Statistician**, v. 46, n. 3, p. 175-185, 1992.
- ALVARENGA, G. C.; RAMALHO, E. E.; BACCARO, F. B.; ROCHA, D. G.; FERREIRA-FERREIRA, J.; BOBROWIEC, P. E. D. Spatial patterns of medium and large size mammal assemblages in várzea and terra firme forests, Central Amazonia, Brazil. **PLoS ONE**, v. 13, n. 5, e0198120, 2018.
- AMADO, T. F.; MOURA, T. A.; RIUL, P.; LIRA, A. F. A.; BADILLO-MONTAÑO, R.; MARTINEZ, P. A. Vulnerable areas to accidents with scorpions in Brazil. **Tropical Medicine and International Health**, v. 26, n. 5, 2021.
- AMIOT, C.; SANTOS, C. C.; ARVOR, D.; BELLÓN, B.; FRITZ, H.; HARMANGE, C.; HOLLAND, J. D.; MELO, I.; METZGER, J. P.; RENAUD, P. C.; ROQUE, F. O.; SOUZA, F. L.; PAYS, O. The scale of effect depends on operational definition of forest cover—evidence from terrestrial mammals of the Brazilian savanna. **Landscape Ecology**, v. 36, p. 973-987, 2021.
- AMIRI, M.; TARKESH, M.; JAFARI, R.; JETSCHKE, G. Bioclimatic variables from precipitation and temperature records vs. remote sensing-based bioclimatic variables: Which side can perform better in species distribution modeling? **Ecological Informatics**, v. 57, 101060, 2020.

ANDRADE, A. C.; MEDEIROS, S.; CHIARELLO, A. G. City sloths and marmosets in Atlantic forest fragments with contrasting levels of anthropogenic disturbance. **Mammal Research**, v. 65, p. 481–491, 2020.

ANSELIN, L. Local Indicators of Spatial Association — LISA. **Geographical Analysis**, v. 27, p. 93–115, 2010.

ANSELIN, L. The Moran Scatterplot as an ESDA Tool to Assess Local Instability in Spatial Association. In: FISCHER, M.; SCHOLTEN, H.; UNWIN, D. (eds). **Spatial Analytical Perspectives on GIS in Environmental and Socio-Economic Sciences**. London: Taylor; Francis, 1996. p. 111–125.

ARAÚJO, R. P. C.; RODRIGUES, A. L. A. M.; GALLIEZ, M. Sharing the nature: mammals and neighbor human population of a protected area in the Atlantic Forest, Brazil. **Oecologia Australis**, v. 24, n. 3, p. 721-735, 2020.

ARROYO, J.; DÍEZ, B.; KEMPES, C. P.; MARQUET, P. A. A general theory for temperature dependence in biology. **Proceedings of the National Academy of Sciences of United States of America**, v. 30, n.9, e2119872119, 2022.

ARTEAGA, M. C.; VENTICINQUE, E. M. Influence of topography on the location and density of armadillo burrows (Dasypodidae: Xenarthra) in the central Amazon, Brazil. **Mammalian Biology**, v. 73, p.262-266, 2008.

ATTIAS, N.; OLIVEIRA-SATOS, L. G. R.; FAGAN, W. F.; MOURÃO, G. Effects of air temperature on habitat selection and activity patterns of two tropical imperfect homeotherms. **Animal Behaviour**, v. 140, p. 129-140, 2018.

BABAR, M.; ARIF, F.; JAN, M. A.; TAN, Z.; KHAN, F. Urban data management system: Towards Big Data analytics for Internet of Things based smart urban environment using customized Hadoop. **Future Generation Computer Systems**, v. 96, p. 398-409, 2019.

BAITELLO, J. B.; AGUIAR, O. T.; PASTORE, J. A.; ARZOLLA, F. A. R. C. P. Parque Estadual do Juquery: Refúgio de Cerrado no Domínio Atlântico. **IF Série Registros**, n. 50, p. 1-46, 2013.

BAKER, D. J.; MACLEAN, I. M. D.; GOODALL, M.; GASTON, K. J. Species distribution modelling is needed to support ecological impact assessments. **Journal of Applied Ecology**, v. 58, n. 1, p. 21-26, 2020.

BALLARINI, Y.; CHAVES, F. G.; VECCHI, M. B.; ALVES, M. A. S. High Rates of Predation of the Nests of Two Endemic Antbirds of the Brazilian Atlantic Forest by Invasive Marmosets (*Callithrix* spp.). **Annales Zoologici Fennici**, v. 58, n.1-3, p. 31-40, 2021.

BARBET-MASSIN, M.; JIGUET, F.; ALBERT, C. H.; THUILLER, W. Selecting pseudo-absences for species distribution models: how, where and how many? **Methods in Ecology and Evolution**, v. 3, n. 2, p. 327-338, 2012.

BARBOSA, E. G.; PIVELLO, V. R.; MEIRELLES, S. T. Allelopathic evidence in *Brachiaria decumbens* and its potential to invade the Brazilian Cerrados. **Brazilian Archives of Biology and Technology**, v. 51, n. 4, p. 825-831, 2008.



BARROS, L. S. C.; LEUZINGER, M. D. Planos de Manejo: panorama, desafios e perspectivas. **Cadernos do Programa de Pós-Graduação**, v. 13, n. 2, p. 281-303, 2018.

BASTOS, B.; PRADHAN, N.; TARROSO, P.; BRITTO, J. C.; BORATYŃSKI, Z. Environmental determinants of minimum body temperature in mammals. **Journal of Vertebrate Biology**, v. 70, n. 2, 2021.

BEAN, W. T.; STAFFORD, R.; BRASHARES, J. S. 2012. The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. **Ecography**, v. 35, n. 3, p. 250-258, 2012.

BEAUDROT, L.; ACEVEDO, M. A.; LESSAR, J. P.; ZVOLEFF, A.; JANSEN, P. A.; SHEIL, D.; ROVERO, F.; O'BRIEN, T.; LARNEY, E.; FLETCHER, C.; ANDELMAN, S.; AHUMADA, J. Local temperature and ecological similarity drive distributional dynamics of tropical mammals worldwide. **Global Ecology and Biogeography**, v. 28, n.7, p. 976-991, 2019.

BEERY, S.; COLE, E.; PARKER, J.; PERONA, P.; WINNER, K. Species Distribution Modeling for Machine Learning Practitioners: A Review. *In: ACM SIGCAS CONFERENCE ON COMPUTING AND SUSTAINABLE SOCIETIES (COMPASS '21)*, 1. 2021, New York. **Proceedings** [...]. New York, NY, USA: Association for Computing Machinery, 2021. p. 329–348. Disponível em <<https://dl.acm.org/doi/pdf/10.1145/3460112.3471966>>. Acesso em 07 set 2022.

BELLIS, J.; LONGDEN, M.; STYLES, J.; DALRYMPLE, S. Using macroecological species distribution models to estimate changes in the suitability of sites for threatened species reintroduction. **Ecological Solutions and Evidence**, v. 2, n. 1, e12050, 2021.

BELLÓN, B.; BLANCO, J.; DE VOS, A.; ROQUE, F. O.; PAYS, O.; RENAUD, P. C. Integrated Landscape Change Analysis of Protected Areas and their Surrounding Landscapes: Application in the Brazilian Cerrado. **Remote Sensing**, v. 12, n. 9, 1413, 2020.

BENEMANN, V. R. F.; ARAUJO, L. D.; FABBRIS, A. Z.; MONTONE, R. C.; PETRY, M. V. Nesting distribution of masked booby *Sula dactylatra* at Trindade Island, Western South Atlantic Ocean. **Marine Ornithology**, v. 50, p. 189-195, 2022.

BENNETSEN, E.; GOBEYN, S.; GOETHALS, P. L. M. Species distribution models grounded in ecological theory for decision support in river management. **Ecological Modelling**, v. 325, p. 1-12, 2016.

BERGSTRA, J.; BENGIO, Y. Random Search for Hyper-Parameter Optimization. **Journal of Machine Learning Research**, v. 13, p. 281-305, 2012.

BERTASSONI, A.; COSTA, R. T.; GOUVEA, J. A.; BIANCHI, R. C.; RIBEIRO, J. W.; VANCINE, M. H.; RIBEIRO, M. C. Land-use changes and the expansion of biofuel crops threaten the giant anteater in southeastern Brazil. **Journal of Mammalogy**, v. 100, n. 2, p. 435-444, 2019.

BERTASSONI, A.; CAIXETA-OLIVEIRA, D.; SILVA, J. M. M.; CASTRO, V. G.; AMARAL, G. L.; YOGUI, D. R.; KLUYBER, D.; DESBIEZ, A. L. Do telemetry harnesses affect giant anteater (*Myrmecophaga tridactyla*) behavior and welfare? **European Journal of Wildlife Research**, v. 68, 40, 2022.

BEVEN, K. Towards a coherent philosophy for modelling the environment. **Proceedings of the Royal Society Publishing A: Mathematical, Physical and Engineering Sciences**, v. 458, n. 2026, p. 2465-2484, 2002.

BEVEN, K. **Environmental modelling: an uncertain future?** New York, USA: Routledge, 2009.

BEZJAK, S.; CLYBURN-SHERIN, A.; CONZETT, P.; FERNANDES, P.; GÖRÖGH, E.; HELBIG, K.; KRAMER, B.; LABASTIDA, I.; NIEMEYER, K.; PSOMOPOULOS, F.; ROSS-HELLAUER, T.; SCHNEIDER, R.; TENNANT, J.; VERBAKEL, E.; BRINKEN, H.; HELLER, L. **Open Science Training Handbook**. [s.l.]: Zenodo, 2018. doi: 10.5281/zenodo.1212496.

BICCA-MARQUES, J. C.; FREITAS, D. S. The Role of Monkeys, Mosquitoes, and Humans in the Occurrence of a Yellow Fever Outbreak in a Fragmented Landscape in South Brazil: Protecting Howler Monkeys is a Matter of Public Health. **Tropical Conservation Science**, v. 3, n. 1, p. 78-89, 2010.

BICKLEY, S. M.; LEMOS, F. G.; GILMORE, M. P.; AZEVEDO, F. C.; FREEMAN, E. W.; SONGSASEN, N. Human perceptions of and interactions with wild canids on cattle ranches in central Brazil. **Oryx**, v. 54, n. 4, p. 546-553, 2020.

BITTNER, R. E.; ROESLER, E. L.; BARNES, M. A. Using species distribution models to guide seagrass management. **Estuarine, Coastal and Shelf Science**, v. 240, 106790, 2020.

BLANCO, A.; LARRINAGA, A. R.; NETO, J. M.; TRONCOSO, J.; MÉNDEZ, G.; DOMÍNGUEZ-LAPIDO, P.; OVEJERO, A.; PEREIRA, L.; MOUGA, T. M.; GASPAR, R.; MARTÍNEZ, B.; LEMOS, M. F. L.; OLABARRIA, C. Spotting intruders: Species distribution models for managing invasive intertidal macroalgae. **Journal of Environmental Management**, v. 281, 111861, 2021.

BOEING, G. Clustering to Reduce Spatial Data Set Size, **SocArXiv**, pp. 1-7, 2018.

BOGONI, J. A.; PERES, C. A.; FERRAZ, K. M. P. M. B. Effects of mammal defaunation on natural ecosystem services and human well being throughout the entire Neotropical realm, **Ecosystem Services**, v. 45, 101173, 2020.

BOLAM, F. C.; GRAINGER, M. J.; MENGERSEN, K. L.; STEWART, G. B.; SUTHERLAND, W. J.; RUNGE, M. C.; MCGOWAN, P. J. K. Using the Value of Information to improve conservation decision making. **Biological Reviews**, v. 94, n. 2, p. 629-647, 2019.

BOOTH, T. H. Checking bioclimatic variables that combine temperature and precipitation data before their use in species distribution models. **Austral Ecology**, v. 47, n. 2, p. 1506-1514, 2022.

BORDONAL, R. O.; CARVALHO, J. L. N.; LAL, R.; FIGUEIREDO, E. B.; OLIVEIRA, B. G.; LA SCALA JR, N. Sustainability of sugarcane production in Brazil. A review. **Agronomy for Sustainable Development**, v. 38, 13, 2018.

BOROUSHAKI, S.; MALCZEWSKI, J. ParticipatoryGIS: A Web-based Collaborative GIS and Multicriteria Decision Analysis. **Journal of the Urban and Regional Information Systems Association**, v. 22, n. 1, p. 23-32, 2010.

- BOSCHETTI, A.; MASSARON, L. **Python Data Science Essentials: A Practitioner's Guide Covering Essential Data Science Principles, Tools, and Techniques**. 3. ed. Birmingham, UK: Packt Publishing, 2018. 472 p.
- BOU, N.; CUYCKENS, G. A. E.; GONZÁLEZ, E. M.; MENEGHEL, M. Conservation planning in Uruguay based on small felids (*Leopardus* spp.) as umbrella species. **Studies on Neotropical Fauna and Environment**, v. 54, n. 3, p. 169-180, 2019.
- BOURAMTANE, T.; HILAL, H.; REZENDE-FILHO, A. T.; BORAMTANE, K.; BARBIERO, L.; ABRAHAM, S.; VALLES, V.; KACIMI, I.; SANHAJI, H.; TORRES-RONDON, L.; CASTRO, D. D.; SANTOS, J. C. V.; OUARDI, J.; EL BEQQALI, O.; KASSOU, N.; MORARECH, M. Mapping Gully Erosion Variability and Susceptibility Using Remote Sensing, Multivariate Statistical Analysis, and Machine Learning in South Mato Grosso, Brazil. **Geosciences**, v. 12, n. 6, 235, 2022.
- BOX, G. E. P. Science and Statistics. **Journal of the American Statistical Association**, v. 71, n. 356, p. 791-799. doi: 10.1080/01621459.1976.10480949
- BRADLEY, B. A.; OLSSON, A. D.; WANG, O.; DICKSON, B. G.; PELECH, L.; SESNIE, S. E.; ZACHMANN, L. J. Species detection vs. habitat suitability: Are we biasing habitat suitability models with remotely sensed data? **Ecological Modelling**, v. 244, p. 57-64, 2012.
- BRAMUTH, A.; PAPROCKI, H. A habitat suitability modeling of *Campylocia burmeisteri* (EPHEMEROPTERA: Euthyplociidae) (Hagen, 1888) and its application on freshwater resources conservation. **Acta Limnologica Brasiliensis**, v. 34 e7, 2022.
- BRANCO, I. H. D. C.; DOMINGUES, E. N.; SÉRIO, F. C.; DEL CALI, I. H.; MATTOS, I. F. A.; BERTONI, J. E. A.; ROSSI, M.; ESTON, M. R.; PFEIFER, R. M.; ANDRADE, W. J. Plano Conceitual de Manejo - Parque Estadual das Furnas do Bom Jesus. **Revista do IF - Memórias do Instituto Florestal**, v. 3, n. 2, p. 137-155, 1991.
- BRASIL, S. B. M.; CARVALHO, R. G. Desafios para a sustentabilidade das unidades de conservação no Brasil: novo caminho trilhado para os planos de manejo. **Paper do NAEA**, v. 30, n. 1, Edição 525, 2021.
- BREIMAN, L. Random Forests. **Machine Learning**, v. 45, p. 5-32, 2001.
- BREIMAN, L.; FRIEDMAN, J.; OLSHEN, R.; STONE, C. **Classification and Regression Trees**. Wadsworth, Belmont, CA: Chapman & Hall, 1984. 368 p.
- BREUSCH, T. S.; PAGAN, A. R. A Simple Test for Heteroscedasticity and Random Coefficient Variation. **Econometrica**, v. 47, n. 5, p. 1287-1294. 1979.
- BRITO, D. V.; BOCCHIGLIERI, A. Bats community (Mammalia, Chiroptera) in Refúgio de Vida Silvestre Mata do Junco, Sergipe, northeastern Brazil. **Biota Neotropica**, v. 12, n. 3, p. 254-262, 2012.
- BROCKERHOFF, E. G.; JACTEL, H.; PARROTTA, J. A.; FERRAZ, S. F. B. Role of eucalypt and other planted forests in biodiversity conservation and the provision of biodiversity-related ecosystem services. **Forest Ecology and Management**, v. 301, p. 43-50, 2013.

- BRODIE, J. F.; REDFORD, K. H.; DOAK, D. F. Ecological Function Analysis: Incorporating Species Roles into Conservation. **Trends in Ecology and Evolution**, v. 33, n. 11, p. 840-850, 2018.
- BURIVALOVA, Z.; GAME, E. T.; BUTLER, R. A. The sound of a tropical forest. **Science**, v. 363, n. 6422, p. 28-29, 2019.
- CACERES, N. C. Biological characteristics influence mammal road kill in an Atlantic Forest–Cerrado interface in south-western Brazil. **Italian Journal of Zoology**, v. 78, n. 3, p. 379-389, 2011.
- CALDWELL, C. A.; WHITEN, A. Scrounging facilitates social learning in common marmosets, *Callithrix jacchus*. **Animal Behaviour**, v. 65, n. 6, p. 1085-1092, 2003.
- CAMPOS-CERQUEIRA, M.; TERANDO, A. J.; MURRAY, B. A.; COLLAZO, J. A.; AIDE, T. M. Climate change is creating a mismatch between protected areas and suitable habitats for frogs and birds in Puerto Rico. **Biodiversity and Conservation**, v. 30, p. 3509–3528, 2021.
- CARVALHO, A. L. C.; ARAÚJO, A. R.; MACHADO, T. M. M.; RIBON, R.; LOPES, L. E. Wildlife and damage to agriculture: an ethnobiological approach with rural producers in southeastern Brazil. **Revista Brasileira de Ornitologia**, v. 27, p. 17–26, 2019.
- CARVALHO, M.; DELGADO, D. B. M.; LIMA, K. M.; CANCELA, M. C.; SIQUEIRA, C. A.; SOUZA, D. L. B. Effects of the COVID-19 pandemic on the Brazilian electricity consumption patterns. **International Journal of Energy Research**, v. 45, n. 2, p. 3358-3364, 2021a.
- CARVALHO, A. F.; FERNANDES-FILHO, E. I.; DAHER, M.; GOMES, L. C.; CARDOSO, I. M.; FERNANDES, R. B. A.; SCHAEFER, C. E. G. R. Microclimate and soil and water loss in shaded and unshaded agroforestry coffee systems. **Agroforestry Systems**, v. 95, n. 119-134, 2021b.
- CARVALHO, C. S.; OLIVEIRA, M. E.; RODRIGUEZ-CASTRO, K. G.; SARANHOLI, B. H.; GALETTI JR., P. M. Efficiency of eDNA and iDNA in assessing vertebrate diversity and its abundance. **Molecular Ecology Resources**, v. 22, n. 4, p. 1262-1273, 2022.
- CARVALHO, F. R. **Articulação sistemática de folhas de cartas projeção TM** (conforme de Gauss). São José dos Campos: [s.n.], 1980. 41 p. Disponível em <<https://periodicos.unb.br/index.php/rbbsb/article/download/29045/26152/72476>>. Acesso em 24 abr. 2022.
- CASSEY, P.; BLACKBURN, T. M. Reproducibility and Repeatability in Ecology. **BioScience**, v. 56, n. 12, p. 958-959, 2006.
- CASTELLANOS, A. A.; HUNTLEY, J. W.; VOELKER, G.; LAWING, A. M. Environmental filtering improves ecological niche models across multiple scales. **Methods in Ecology and Evolution**, v. 10, p. 481-492, 2019.
- CASTRO, L. N.; FERRARI, D. G. **Introdução à Mineração de Dados: conceitos básicos, algoritmos e aplicações**. São Paulo: Saraiva, 2016. 351 p.

- CATOJO, A. M. Z.; JESUS, S. C. As unidades de conservação do estado de São Paulo: planos de manejo e representatividade. **Revista Brasileira de Geografia Física**, v. 15, n. 6, p. 2921-2943, 2022.
- CAUDILL, S. A.; DECLERCK, F. C. A.; HUSBAND, T. P. Connecting sustainable agriculture and wildlife conservation: Does shade coffee provide habitat for mammals? **Agriculture, Ecosystems & Environment**, v. 199, p. 85-93, 2015.
- CAYUELA, L.; GOLICHER, D. J.; NEWTON, A. C.; KOLB, M.; ALBURQUERQUE, F. S.; ARETS, E. J. M. M.; ALKEMADE, J. R. M.; PÉREZ, A. M. Species Distribution Modeling in the Tropics: Problems, Potentialities, and the Role of Biological Data for Effective Species Conservation. **Tropical Conservation Science**, v. 2, n. 3, p. 319-352, 2009.
- CELERE, B. S. **As queimadas na cana-de-açúcar: impactos na fauna silvestre**. São Carlos: [s.n.], 2012. 21 p. (Relatório de Iniciação Científica).
- CERRI, R. I.; REIS, F. A. G. V.; GRAMANI, M. F.; ROSOLEN, V.; LUVIZOTTO, G. L.; GIORDANO, L. C.; GABELINI, B. M. Assessment of landslide occurrences in Serra do Mar mountain range using kinematic analyses. **Environmental Earth Sciences**, v. 77, 325, 2018.
- CHAHAD-EHLERS, S.; FUSHITA, A. T.; LACORTE, G. A.; ASSIS, P. C. P.; DEL LAMA, S. N. Effects of habitat suitability for vectors, environmental factors and host characteristics on the spatial distribution of the diversity and prevalence of haemosporidians in waterbirds from three Brazilian wetlands. **Parasites & Vectors**, v. 11, 276, 2018.
- CHAPMAN, D. S.; PURSE, B. V. Community versus single-species distribution models for British plants. **Journal of Biogeography**, v. 38, p. 1524-1535, 2011.
- CHAPMAN, C. A.; STEINICHE, T.; BENAVIDEZ, K. M.; SARKAR, D.; AMATO, K.; SERIO-SILVA, J. C.; VENIER, M.; WASSERMAN, M. D. The chemical landscape of tropical mammals in the Anthropocene. **Biological Conservation**, v. 269, 109522, 2022.
- CHAPE, S.; HARRISON, J.; SPALDING, M.; LYSENKO, I. Measuring the extent and effectiveness of protected areas as an indicator for meeting global biodiversity targets. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 360, n. 1454, p. 443-455, 2005.
- CHAVES, O.M.; FORTES, V. B.; HASS, G. P.; AZEVEDO, R. B.; STONER, K. E.; BICCA-MARQUES, J. C. Flower consumption, ambient temperature and rainfall modulate drinking behavior in a folivorous-frugivorous arboreal mammal. **PLoS ONE**, v. 16, n. 2, e0236974, 2021.
- CHEN, S.; DING, S. Machine Learning and Its Applications in Studying the Geographical Distribution of Ants, **Diversity**, v. 14, n. 9, 706, 2022.
- CHEN, T.; GUESTRIN, C. Xgboost: A scalable tree boosting system. *In: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 22., 2016, California, San Francisco, USA. Proceedings [...].* New York, NY, United States: Association for Computing Machinery, 2016. p. 785-794. Disponível em <<https://dl.acm.org/doi/abs/10.1145/2939672.2939785>>. Acesso em 30 ago 2022
- CHEN, D.; XUE, Y.; FINK, D.; CHEN, S.; GOMES, C.P. Deep Multi-species Embedding. *In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 26.,*

2017, Melbourne, Australia. **Proceedings** [...]. Freiburg, Germany: International Joint Conferences on Artificial Intelligence (IJCAI), 2017. p. 3639-364. Disponível em <<https://www.ijcai.org/proceedings/2017/0509.pdf>>. Acesso em 30 ago. 2022.

CHIARELLO, A. G. Density and Population Size of Mammals in Remnants of Brazilian Atlantic Forest. **Conservation Biology**, v. 14, n. 6, p. 1649-1657, 2000.

CHINCHILLA, S.; VAN DEN BERGHE, E.; POLISAR, J.; ARÉVALO, C.; BONACIC, C. Livestock–carnivore coexistence: moving beyond preventive killing. **Animals**, v. 12, n. 4, 479, 2022.

CHO, S. H.; BOWKER, J. M.; PARK, W. M. Measuring the Contribution of Water and Green Space Amenities to Housing Values: An Application and Comparison of Spatially Weighted Hedonic Models. **Journal of Agricultural and Resource Economics**, v. 31, n. 3, p. 485-507, 2006.

CHOPIN, P.; BERGKVIST, G.; HOSSARD, L. Modelling biodiversity change in agricultural landscape scenarios - A review and prospects for future research. **Biological Conservation**, v. 235, p. 1-17, 2019.

CHUA, P. Y. S.; CRAMPTON-PLATT, A.; LAMMERS, Y.; ALSOS, I. G.; BOESSENKOOL, S.; BOHMANN, K. Metagenomics: A viable tool for reconstructing herbivore diet. **Molecular Ecology Resources**, v. 21, n. 7, p. 2249-2263, 2021.

CLAUZEL, C.; GIRARDET, X.; FOLTÊTE, J. C. Impact assessment of a high-speed railway line on species distribution: Application to the European tree frog (*Hyla arborea*) in Franche-Comté. **Journal of Environmental Management**, v. 127, p. 125-134, 2013.

CLIFF, A. C.; ORD, J. K. **Spatial autocorrelation**. London, UK: Pion Limited, 1973.

CNUC - CADASTRO NACIONAL DE UNIDADES DE CONSERVAÇÃO. **Painel Unidades de Conservação Brasileiras**. Disponível em <<https://cnucc.mma.gov.br/powerbi>>. Acesso em 14 jan. 2023.

COAD, L.; LEVERINGTON, F.; KNIGHTS, K.; GELDMANN, J.; EASSOM, A. KAPO, V.; KINGSTON, N.; LIMA, M.; ZAMORA, C.; CUARDOS, I.; NOLTE, C.; BURGESS, N.D.; HOCKINGS, M. Measuring impact of protected area management interventions: current and future use of the Global Database of Protected Area Management Effectiveness. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 30, 20140281, 2015.

COELHO, M. T. P.; DINIZ-FILHO, J. A.; RANGEL, T. F. A parsimonious view of the parsimony principle in ecology and evolution, **Ecography**, v. 42, n. 5, p. 968-976, 2019.

COELHO-JUNIOR, M. G.; OLIVEIRA, A. L.; SILVA-NETO, E. C.; CASTOR-NETO, T. C.; TAVARES, A. A. O.; BASSO, V. M.; TURETTA, A. P. D.; PERKINS, P. E.; CARVALHO, A. G. Exploring Plural Values of Ecosystem Services: Local Peoples' Perceptions and Implications for Protected Area Management in the Atlantic Forest of Brazil. **Sustainability**, v. 13, n. 3, 1019, 2021.

CONCEIÇÃO, R. S.; COSTA, N. M. C.; COSTA, V. C. A importância da evolução do uso do solo como geoindicador para o planejamento do Ecoturismo em Unidades de Conservação:

aplicação no Parque Estadual da Pedra Branca (RJ). **Revista Brasileira de Ecoturismo**, v.3, n.3, p.408-427, 2010.

CONCEIÇÃO, E. O.; GARCIA, J. M.; ALVES, G. H. Z.; DELANIRA-SANTOS, D.; CORBETTA, D. F.; BETIOL, T. C. C.; PACIFICO, R.; ROMAGNOLO, M. B.; BATISTA-SILVA, V.; BAILY, D.; FERREIRA, J. H. D.; COUTO, E. V. The impact of downsizing protected areas: How a misguided policy may enhance landscape fragmentation and biodiversity loss. **Land Use Policy**, v. 112, 105835, 2022.

CONDRO, A. A.; PRASETYO, L. B.; RUSHAYATI, S. B.; SANTIKAYASA, P.; ISKANDAR, E. Predicting Hotspots and Prioritizing Protected Areas for Endangered Primate Species in Indonesia under Changing Climate. **Biology (Basel)**, v. 10, 154, 2021.

CONSTANZA, R.; D'ARGE, R.; DE GROOT, R.; FARBER, S.; GRASSO, M.; HANNON, B.; LIMBURG, K.; NAEEM, S.; O'NEILL, R. V.; PARUELO, J.; RASKIN, R. G.; SUTTON, P.; VAN DEN BELT, M. The value of the world's ecosystem services and natural capital. **Nature**, v. 387, n. 6630, p. 253-260, 1997.

CORCORAN, E.; DENMAN, S.; HAMILTON, G. Evaluating new technology for biodiversity monitoring: Are drone surveys biased? **Ecology and Evolution**, v. 11, n. 11, p. 6649-6656, 2021.

CORLETT, R. T. Present-day drivers do not explain biodiversity patterns in mammals. **Proceedings of the National Academy of Sciences**, v. 117, n. 4, p. 1836-1838, 2020.

CORO, G.; MAGLIOZZI, C.; BERGHE, E. V.; BAILLY, N.; ELLENBROEK, A.; PAGANO, P. Estimating absence locations of marine species from data of scientific surveys in OBIS. **Ecological Modelling**, v. 323, p. 61-76, 2016.

CORSI, F.; LEEUW, J.; SKIDMORE, A. Modeling Species Distribution in GIS. *In*: BOITANI, L.; FULLER, T. (ed.). **Research Techniques in Animal Ecology: Controversies and Consequences**. New York: Columbia University Press, 2000. p. 389-421.

COSTA, M. C.; HAGOPIÁN, D.; SIMÓ, M.; GUERRERO, J. C.; LABORDA, A. First description of the male of *Mastophora extraordinaria* Holmberg, 1876 (Araneae: Araneidae), with notes on the natural history and potential distribution of the species. **Boletín de la Sociedad Zoológica del Uruguay**, v. 30, n. 1, p. 14-24, 2021.

COSTA, Y.; MARTINS, I.; CARVALHO, G. C.; BARROS, F. Sea-level rise effects on macrozoobenthos distribution within an estuarine gradient using Species Distribution Modeling. **Ecological Informatics**, v. 71, 101816, 2022.

COUTO, M. S. H. Ocupação irregular e criminalidade na região da Serra da Cantareira - SP. **Revista Laboratório de Estudos da Violência da UNESP / Marília**. v. 8, p. 107-126, 2011.

COVENEY, P. V.; GROEN, D.; HOEKSTRA, A. G. Reliability and reproducibility in computational science: implementing validation, verification and uncertainty quantification *in silico*. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v. 379, n. 2179, 20200409, 2021.

CROOKS, K. R.; BURDETT, C. L.; THEOBALD, D. M.; KING, S. R. B.; DI MARCO, M.; RONDININI, C.; BOITANI, L. Quantification of habitat fragmentation reveals extinction risk

in terrestrial mammals. **Proceedings of the National Academy of Sciences**, v. 114, n. 29, p. 7635-7640, 2017.

CUNHA, E. R.; SANTOS, C. A. G.; SILVA, R. M.; BACANI, V. M.; POTT, A. Future scenarios based on a CA-Markov land use and land cover simulation model for a tropical humid basin in the Cerrado/Atlantic forest ecotone of Brazil. **Land Use Policy**, v. 101, 105141, 2021.

D'AGOSTINO, R.; PEARSON, E. S. Tests for departure from normality. **Biometrika**, v. 60, p. 613-622, 1973.

DAGOUMAS, A. S.; KOLTSAKLIS, N. E. Review of models for integrating renewable energy in the generation expansion planning. **Applied Energy**, v. 242, p. 1573-1587, 2019.

DALLAPICOLLA, J.; ABREU, E. F.; PRADO, J. R.; CHIQUITO, E. A.; ROTH, P. R. O.; BRENNAND, P. G. G.; PAVAN, A. C. D.; PEREIRA, A.; MENDES, F. R.; ALVAREZ, M. R. V.; RIOS, E.O.; CASSANO, C.R.; MIRETZKI, M.; VÉLEZ, F.; SEVÁ, A.P.; PERCEQUILLO, A.R.; BOVENDORP, R.S. Areas of endemism of small mammals are underprotected in the Atlantic Forest. **Journal of Mammalogy**, v. 102, n. 5. p. 1390-1404, 2021.

DAMBROS, C.; CÁCERES, N.; BASELGA, A. The prevalence of temperature and dispersal limitation as drivers of diversity in Neotropical small mammals. **Austral Ecology**, v. 47, n. 3, p. 567-579, 2022.

DANCHEV, V. Reproducible Data Science with Python: An Open Learning Resource. **Journal of Open Source Education**, v. 5, n. 56, 156, 2022.

DARABI, H.; CHOUBIN, B.; RAHMATI, O.; HAGHIGHI, A. T.; PRADHAN, B.; KLØVE, B. Urban flood risk mapping using the GARP and QUEST models: A comparative study of machine learning techniques. **Journal of Hydrology**, v. 569, p. 142-154, 2019.

DARRAS, K.; BATÁRY, P.; FURNAS, B. J.; GRASS, I.; MULYANI, Y. A.; TSCHARNTKE, T. Autonomous sound recording outperforms human observation for sampling birds: a systematic map and user guide. **Ecological Applications**, v. 29, n. 6, e01954, 2019.

DASZAK, P.; CUNNINGHAM, A. A.; HYATT, A. D. Emerging infectious diseases of wildlife--threats to biodiversity and human health. **Science**, v. 287, n. 5459, p-443-449, 2000.

DAUNT, A. B. P.; SILVA, T. S. F.; BÜRGI, M.; HERSPERGER, A. M. Urban expansion and forest reserves: Drivers of change and persistence on the coast of Sao Paulo State (Brazil). **Land Use Policy**, v. 101, 105189, 2021.

DAUSMANN, K. H. Flexible patterns in energy savings: heterothermy in primates. **Journal of Zoology**, v. 292, n. 2, p. 101-111, 2014.

DEAN, W. **A ferro e fogo: a história e a devastação da Mata Atlântica brasileira**. 1.ed. São Paulo: Cia. das Letras, 2004. 484 p.

DE BONDI, N.; WHITE, J. G.; STEVENS, M.; COOKE, R. A comparison of the effectiveness of camera trapping and live trapping for sampling terrestrial small-mammal communities. **Wildlife Research**, v. 37, n.6, p. 456-465, 2010.



DE MARCO JR, P.; VILLÉN, S.; MENDES, P.; NÓBREGA, C.; CORTES, L.; CASTRO, T.; SOUZA, R. Vulnerability of Cerrado threatened mammals: an integrative landscape and climate modeling approach. **Biodiversity and Conservation**, v. 29, p. 1637-1658, 2020.

DE PINHO JÚNIOR, G. V.; NASCIMENTO, A. R. T.; VALVERDE, B. T.; CLEMENTE, L. H. Brazilian savanna re-establishment in a monoculture forest: diversity and environmental relations of native regenerating understory in *Pinus caribaea* Morelet. stands. **Journal of Forestry Research**, v. 26, p. 571-579, 2015.

DESBIEZ, A. L. J.; BODMET, R. E.; SANTOS, S. A. Wildlife habitat selection and sustainable resources management in a Neotropical wetland. **International Journal of Biodiversity and Conservation**, v. 1, n. 1, p. 11-20, 2009.

DESBIEZ, A. L. J.; KLUYBER, D.; MASSOCATO, G. F.; ATTIAS, N. Methods for the characterization of activity patterns in elusive species: the giant armadillo in the Brazilian Pantanal. **Journal of Zoology**, v. 315, n. 4, p. 301-312, 2021.

DESJONQUÈRES, C.; VILLÉN-PÉREZ, S.; DE MARCO, P.; MÁRQUEZ, R.; BELTRÁN, J. F.; LLUSIA, D. Acoustic species distribution models (aSDMs): A framework to forecast shifts in calling behaviour under climate change. **Methods in Ecology and Evolution**, v. 13, n. 10, p. 2275-2288, 2022.

DIAS, W. A. F.; TEZORI, R. F. F.; OLIVEIRA, A. K. Registro de mamíferos de médio e grande porte em dois fragmentos florestais no município de São Carlos, Estado de São Paulo, **Multiciência**, v. 11, p. 277-293, 2012.

DIAS, D. M.; FERREGUETTI, A. C.; RODRIGUES, F. H. G. Using an occupancy approach to identify poaching hotspots in protected areas in a seasonally dry tropical forest. **Biological Conservation**, v. 251, 108796, 2020.

DICK, M.; SILVA, M. A.; SILVA, R. R. F.; FERREIRA, O. G. L.; MAIA, M. S.; LIMA, S. F.; PAIVA NETO, V. B.; DEWES, H. Environmental impacts of Brazilian beef cattle production in the Amazon, Cerrado, Pampa, and Pantanal biomes. **Journal of Cleaner Production**, v. 311, 127750, 2021.

DIDONÉ, E. J.; MINELLA, J. P. G.; TIECHER, T.; ZANELLA, R.; PRESTES, O. D.; EVRARD, O. Mobilization and transport of pesticides with runoff and suspended sediment during flooding events in an agricultural catchment of Southern Brazil. **Environmental Science and Pollution Research**, v. 28, p. 39370–39386, 2021.

DINIZ, M. F.; MACHADO, R. B.; BISPO, A. A.; DE MARCO JR, P. Can we face different types of storms under the same umbrella? Efficiency and consistency of connectivity umbrellas across different patchy landscape patterns. **Landscape Ecology**, v. 33, p. 1911–1923. 2018.

DINIZ, M. F.; COELHO, M. T. P.; SÁNCHEZ-CUERVO, A. M.; LOYOLA, R. How 30 years of land-use changes have affected habitat suitability and connectivity for Atlantic Forest species. **Biological Conservation**, v. 274, 109737, 2022.

DOBSON, B.; BARRY S.; MAES-PRIOR, R.; MIJIC, A.; WOODWARD, G.; PEARSE, W. D. Predicting catchment suitability for biodiversity at national scales. **Water Research**, v. 221, 118764, 2022

DOLATSHAH, M.; HADIAN, A.; MINAEI-BODGOLI, B. Ball\*-tree: Efficient spatial indexing for constrained nearest-neighbor search in metric spaces. **SocArXiv**, pp. 1-15, 2015.

DOMINONI, D.; SMIT, J. A. H.; VISSER, M. E.; HALFWERK, W. Multisensory pollution: Artificial light at night and anthropogenic noise have interactive effects on activity patterns of great tits (*Parus major*). **Environmental Pollution**, v. 256, 113314, 2020.

DODONOV, P.; HARPER, K. A.; SILVA-MATOS, D. M. The role of edge contrast and forest structure in edge influence: vegetation and microclimate at edges in the Brazilian cerrado. **Plant Ecology**, v. 214, p. 1345-1359, 2013.

DONOHO, D. 50 Years of Data Science. **Journal of Computational and Graphical Statistics**, v. 26, n. 4, p. 745-766, 2017.

DOTTA, G.; VERDADE, L. M. Medium to large-sized mammals in agricultural landscapes of south-eastern Brazil. **Mammalia**, v. 75, n. 4, p. 345-352, 2011.

DOUROJEANNI, M. J.; PÁDUA, M. T. J. **Biodiversidade: a hora decisiva**. 2. ed. Curitiba: Editora UFPR, 2007. 284 p.

DRUCKER, H.; BURGESS, C. J. C.; KAUFMAN, L.; SMOLA, A.; VAPNIK, V. Support Vector Regression Machines. *In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS*, 9., 1996, Denver, CO, USA. **Proceedings** [...]. Cambridge, Massachusetts: MIT Press, 1997. p. 155-161. Disponível em <<https://dblp.org/db/conf/nips/nipsN1996.html>>. Acesso em 01 ago. 2022.

DUBOST, G.; HENRY, O. Seasonal Reproduction in Neotropical Rainforest Mammals. **Zoological Studies**, v. 56, n. 2, 2017.

DUFEK, A. S.; AMBRIZZI, T. Precipitation variability in São Paulo State, Brazil. **Theoretical and Applied Climatology**, v. 93, n. 3-4, p. 167-178, 2007.

DURY, J.; GARCIA, F.; REYNAUD, A.; THÉRON, O.; BERGEZ, J. E. Modelling the Complexity of the Cropping Plan Decision-making. *In: INTERNATIONAL CONGRESS ON ENVIRONMENTAL MODELLING AND SOFTWARE*, 5., 2010, Ottawa, Ontario, Canada. **Proceedings** [...]. Manno, Switzerland: International Environmental Modelling and Software Society, 2010. p. 569-576. Disponível em <<http://iemss.org/iemss2010/?n=Main.Proceedings>>. Acesso em 21 dez. 2022.

EBENSPERGER, L. A. Sociality in rodents: the New World fossorial hystricognaths as study models. **Revista Chilena de História Natural**, v. 71, p. 65-77, 1998.

ELDAWY, A.; MOKBEL, M. F. The ecosystem of SpatialHadoop. **SIGSPATIAL**, v. 6, n. 3, p. 3-10, 2014.

ELITH, J.; FRANKLIN, E. Species Distribution Modeling. **Encyclopedia of Biodiversity**, v. 6, p. 692-705, 2013.

ENGLER, R.; GUIGAN, A.; RECHSTEINER, L. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. **Journal of Applied Ecology**, v. 41, n. 2, p. 263-274, 2004.

ENGSTROM, K.; ROUNSEVELL, M. D. A.; MURRAY-RUST, D. HARDACRE, C.; ALEXANDER, P.; CUI, X.; PALMER, P. I.; ARNETH, A. Applying Occam's razor to global agricultural land use change. **Environmental Modelling & Software**, v. 75, p. 212-220, 2016.

ERTEL, W. **Introduction to Artificial Intelligence**. Cham, Switzerland: Springer, 2017. 356p.

ESPINDOLA, G. M.; FIGUEREDO, E. S.; PICANÇO JUNIOR, P.; REIS FILHO, A. A. Cropland expansion as a driver of land-use change: the case of Cerrado-Caatinga transition zone in Brazil. **Environment, Development and Sustainability**, v. 23, p. 17146-17160, 2021.

ESSAWY, B. T.; GOODALL, J. L.; VOCE, D.; MORSY, M. M.; SADLER, J. M.; DON CHOI, Y.; TARBOTON, D. G.; MALIK, T. A taxonomy for reproducible and replicable research in environmental modelling. **Environmental Modelling & Software**, v. 134:104753, 2020.

ESTER, M.; KRIEGEL, H. P.; SANDER, J.; XU, X. A density based-algorithm for discovering clusters in large spatial databases with noise. *In*: INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 2. 1996, Portland Oregon, USA. **Proceedings** [...]. Palo Alto, California: Association for the Advancement of Artificial Intelligence, 1996. p. 226-231.

ETHERINGTON, T. R. 2021. Mahalanobis distances for ecological niche modelling and outlier detection: implications of sample size, error, and bias for selecting and parameterising a multivariate location and scatter method. **PeerJ - Life and Environment**, v. 9, e11436, 2021.

EVANGELISTA-VALE, J. C.; WEIHS, M.; JOSÉ-SILVA, L.; ARRUDA, R.; SANDER, N. L.; GOMIDES, S. C.; MACHADO, T. M.; PIRES-OLIVEIRA, J. C.; BARROS-ROSA, L.; CASTUERA-OLIVEIRA, L.; MATIAS, R. A. A. M. MARTINS-OLIVEIRA, A. T.; SÃO BERNARDO, C. S.; SILVA-PEREIRA, I.; CARNICER, C.; CARPANEDO, R. S.; EISENLOHR, P. V. Climate change may affect the future of extractivism in the Brazilian Amazon. **Biological Conservation**, v. 257, 109093, 2021.

EVANS, N. M. Ecosystem Services: On Idealization and Understanding Complexity. **Ecological Economics**, v. 156, p. 427-430, 2019.

EVANS, M. V.; DRAKE, J. M. A Data-driven Horizon Scan of Bacterial Pathogens at the Wildlife–livestock Interface. **EcoHealth**, v. 19, p. 246-258, 2022.

VAN EUPEN, C.; MAES, D.; HERREMANS, M.; SWINNEN, K. R. R.; SOMERS, B.; LUCA, S. The impact of data quality filtering of opportunistic citizen science data on species distribution model performance. **Ecological Modelling**, v. 444, 109453. 2021.

FABIANE, D.; SPINELLI, J.; SOUZA, R. J. A natureza e a produção de valorização imobiliária nas cidades brasileiras. **Brazilian Geographical Journal**, v. 11, n. 1, p. 199-213, 2020.

- FACELI, K.; LORENA, A. C.; GAMA, J.; ALMEIDA, T. A.; CARVALHO, A. C. P. L. F. **Inteligência artificial: uma abordagem de Aprendizado de Máquina**. 2.ed. Rio de Janeiro: LTC, 2021. 400 p.
- FARIAS, F. O.; BARGOS, D. C.; MATIAS, L. F. Aplicação de geotecnologias no estudo da relação entre valorização da terra urbana e a presença de áreas verdes na cidade de Paulínia (SP). **Revista Brasileira de Cartografia**, v. 68, n. 2, p. 275-287, 2016.
- FARLEY, S. C.; DAWSON, A.; GORING, S. J.; WILLIAMS, J. W. Situating Ecology as a Big-Data Science: Current Advances, Challenges, and Solutions. **BioScience**, v. 68, n. 8, p. 563-576, 2018.
- FAWCETT, S.; AGOSTI, D.; COLE, S. R.; WRIGHT, D. F. Digital accessible knowledge: Mobilizing legacy data and the future of taxonomic publishing. **Bulletin of the Society of Systematic Biologists**, v. 1, n. 1, 8296, 2022.
- FENDRICH, A. N.; ROCHA, A. G.; RANIERI, V. E. L. Comparison between official priority studies guidelines and Protected Areas created in Brazil. **Land Use Policy**, v. 82, p. 240-246, 2019.
- FERNANDES, R. F.; HONRADO, J. P.; GUIBAN, A.; ROXO, A.; ALVES, P. MARTINS, J.; VICENTE, J. R. Species distribution models support the need of international cooperation towards successful management of plant invasions. **Journal for Nature Conservation**, v. 49, p. 85-94, 2019.
- FERNANDES-SANTOS, R. C.; MEDICI, E. P.; MICHELETTI, T. Health assessment of wild lowland tapirs (*Tapirus terrestris*) in the highly threatened Cerrado biome, Brazil. **Journal of Wildlife Diseases**, v. 56, n. 1, p. 34-46, 2020.
- FERREGUETTI, A. C.; GRACIANO, J. M.; LUPPI, A. P.; PEREIRA-RIBEIRO, J.; ROCHA, C. F. D.; BERGALLO, H. G. Roadkill of medium to large mammals along a Brazilian road (BR-262) in Southeastern Brazil: spatial distribution and seasonal variation. **Studies on Neotropical Fauna and Environment**, v. 55, n. 3, p. 216-225, 2020.
- FERREIRA, A. S.; PERES, C. A.; DODONOV, P.; CASSANO, C. R. Multi-scale mammal responses to agroforestry landscapes in the Brazilian Atlantic Forest: the conservation value of forest and traditional shade plantations. **Agroforestry Systems**, v. 94, p. 2331–2341, 2020.
- FERRIER, S.; DRIELSMA, M. Synthesis of pattern and process in biodiversity conservation assessment: a flexible whole-landscape modelling framework. **Diversity and Distributions**, v. 16, n. 3, p. 386-402.
- FLETCHER, R. J.; HEFLEY, T. J.; ROBERTSON, E. P.; ZUCKERBERG, B.; MCCLEERY, R. A.; DORAZIO, R. M. A practical guide for combining data to model species distributions. **Ecology**, e02710.2019, 2019.
- FICK, S. E.; HIJMANS, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. **International Journal of Climatology**, v. 37, n. 12, p. 4302-4315, 2017.
- FIGERIO, D.; RICHTER, A.; PER, E.; PRUSE, B.; VOHLAND, K. CITIZEN SCIENCE IN NATURAL SCIENCES. *In*: VOHLAND, K.; LAND-ZANDSTR, A.; CECCARONI, L.; LEMMENS, R.; PERELLÓ, J.; PONTI, M.; SAMSON, R.; WAGENKNECHT, K. (eds). **The Science of Citizen Science**. Cham, Switzerland: Springer, 2021. p. 79-96.

FITKAU, E. J. The fauna of South America. *In*: FITTKAU, E. J.; ILLIES, J.; KLINGE, H.; SCHWABE, G.H.; SIOLI, H. (ed.). **Biogeography and ecology in South America**. The Hague, Netherlands: W. Junk N.V. Publishers, 1969. p. 624-58.

FRANKLIN, J. **Mapping species distributions: spatial inference and prediction**. Cambridge, UK: Cambridge University Press, 2009. 320 p.

FREITAS, C. H.; SETZ, E. Z. F.; ARAÚJO, A. R. B.; GOBBI, N. Agricultural crops in the diet of bearded capuchin monkeys, *Cebus libidinosus* Spix (Primates: Cebidae), in forest fragments in southeast Brazil. **Revista Brasileira de Zoologia**, v. 25, n. 1, p. 32-39, 2008.

FREITAS, S. R.; LIGNANI, L. B.; CABRAL, D. C. Influence of Landscape Features on Forest Maturity: the Case of a Fragmented Landscape in the Serra do Mar Coastal Forest in Brazil. **Natureza & Conservação**, v. 9, n. 2, p. 194-199, 2011.

FREITAS, S. R.; OLIVEIRA, A. N.; CIOCHETI, G.; VIEIRA, M. V.; MATOS, D. M. S. How landscape features influence road-kill of three species of mammals in the Brazilian savanna? **Oecologia Australis**, v. 18, p. 35-45, 2015.

FREITAS-OLIVEIRA, R.; HANNIBAL, W.; LIMA-RIBEIRO, M.; TERRIBILE, L. C. Implications of climate change for the distribution of the water opossum (*Chironectes minimus*): habitat loss and conservation opportunities. **Mammalian Biology**, v. 101, 729-737, 2021.

FUNDAÇÃO FLORESTAL. **Plano de Manejo do Parque Estadual de Campos do Jordão**. Disponível em <<https://smastr16.blob.core.windows.net/fundacaoflorestal/2017/02/Diagn%C3%B3stico-e-Planejamento.pdf>> Acesso em 14 jan. 2023.

GARCÊS, A.; PIRES, I.; RODRIGUES, P. Teratological effects of pesticides in vertebrates: a review. **Journal of Environmental Science and Health, Part B: Pesticides, Food Contaminants, and Agricultural Wastes**, v. 55, n. 1, p. 75-89, 2020.

GARCIA, A. S.; SAWAKUCHI, H. O.; FERREIRA, M. E.; BALLESTER, M. V. R. Landscape changes in a neotropical forest-savanna ecotone zone in central Brazil: The role of protected areas in the maintenance of native vegetation. **Journal of Environmental Management**, v. 187, p. 16-23, 2017.

GASPARETTO, T. C. **Pressão Urbana e Conectividade da Paisagem no Entorno dos Parques Estaduais Itapetinga, Itaberaba, Cantareira, Juquery e Jaraguá na Região Metropolitana de São Paulo**. 2014. Dissertação (Mestrado em Geografia Física). Departamento de Geografia, Universidade de São Paulo, São Paulo, 2014. Disponível em <[https://www.teses.usp.br/teses/disponiveis/8/8135/tde-29042015-143428/publico/2014\\_TadeuDaCostaGaspareto\\_VCorr.pdf](https://www.teses.usp.br/teses/disponiveis/8/8135/tde-29042015-143428/publico/2014_TadeuDaCostaGaspareto_VCorr.pdf)>. Acesso em 01 set 2009.

GASTÓN, A.; GARCÍA-VIÑAS, J. I. Evaluating the predictive performance of stacked species distribution models applied to plant species selection in ecological restoration. **Ecological Modelling**, v. 263, p. 103-108, 2013.

GASTÓN, A.; GARCÍA-VIÑAS, J. I.; BRAVO-FERNÁNDEZ, A. J.; LÓPEZ-LEIVA, C.; OLIET, J. A.; ROIG, S.; SERRADA, R. Species distribution models applied to plant species selection in forest restoration: are model predictions comparable to expert opinion? **New Forests**, v. 45, p. 641-653, 2014.

- GEISE, L.; PEREIRA, L. G.; BOSSI, D. E. P.; BERGALLO, H. G. Pattern of elevational distribution and richness of non volant mammals in Itatiaia National Park and its surroundings, in Southeastern Brazil. **Brazilian Journal of Biology**, v. 64, n. 3B, p. 599-612, 2004.
- GEISER, F.; COOPER, C. E. Daily Torpor, Hibernation, and Heterothermy in Marsupials. American and Australasian Marsupials. *In*: CÁCERES, N.C.; DICKMAN, C.R. (eds). **American and Australasian Marsupials**. Edinburgh, Scotland: Springer, Cham., 2023. p. 1-28.
- GELDMANN, J.; BARNES, M.; COAD, L.; CRAIGIE, I. D.; HOCKINGS, M.; BURGESS, N. D. Effectiveness of terrestrial protected areas in reducing habitat loss and population declines. **Biological Conservation**, v. 161, p. 230-238, 2013.
- GENELETTI, D.; ESMAIL, B. A.; SCOLOZZI, R.; ASSANDRI, G.; BRAMBILLA, M.; PEDRINI, P. Modelling Biodiversity and Ecosystem Services Trade-Offs in Agricultural Landscapes to Support Planning and Policy-Making. *In*: MIRSCHEL, W.; TERLEEVE, V.; WENKEL, K. O. (eds). **Landscape Modelling and Decision Support: Innovations in Landscape Research**. Edinburgh, Scotland: Springer, Cham., 2020. p. 421-441.
- GHELER-COSTA, C.; VETTORAZZI, C. A.; PARDINI, R.; VERDADE, L. M. The distribution and abundance of small mammals in agroecosystems of southeastern Brazil. **Mammalia**, v. 76, n. 2, p. 185-192, 2012.
- GHELER-COSTA, C.; SABINO-SANTOS JR, G.; AMORIM, L. S.; ROSALINO, L. M.; FIGUEIREDO, L. T. M.; VERDADE, L. M. The effect of pre-harvest fire on the small mammal assemblage in sugarcane fields. **Agriculture, Ecosystem and Environment**, v. 171, p. 85-89, 2013.
- GIERE, R. N. How models are used to represent reality. **Philosophy of Science**, v. 71, n. 5, p. 742-752, 2004.
- GILBERT, N.; AHRWEILER, P.; BARBROOK-JOHNSON, P.; NARASIMHAN, K. P. WILKINSON, H. Computational modelling of public policy: reflections on practice. **Journal of Artificial Societies and Social Simulation**, v. 21, n.1, 14, 2018.
- GITTLEMAN, J. L. Carnivore Group Living: Comparative Trends. *In*: GITTLEMAN, J. L. (ed.). **Carnivore Behavior, Ecology, and Evolution**. Boston, MA, USA: Springer, 1989.
- GOBEYN, S.; MOUTON, A. M.; CORD, A. F.; KAIM, A.; VOLK, M.; GOETHALS, L. M. Evolutionary algorithms for species distribution modelling: A review in the context of machine learning. **Ecological Modelling**, v. 392, p. 179-195, 2019.
- GOEL, I.; GORADIA, S. R.; KAKELLI, A. K. Predicting the Presence of Amphibians Near Road Construction Sites Using Emerging Machine Learning Algorithms. *In*: INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND MACHINE VISION, 2., 2021, Settat, Morocco. **Proceedings [...]**. Edinburgh, Scotland: Springer, Cham, 2021. p. 1-6.
- GOMES, L. C.; BIANCHI, F. J. J. A.; CARDOSO, I. M.; FERNANDES, R. B. A.; FERNANDES-FILHO, E. I.; SCHULTE, R. P. O. Agroforestry systems can mitigate the

impacts of climate change on coffee production: A spatially explicit assessment in Brazil. **Agriculture, Ecosystems & Environment**, v. 294, 106858, 2020.

GOMEZ, M. D.; CODA, J. A.; SERAFINI, V. N.; STEINMANN, A. R.; PRIOTTO, J. W. Small mammals in agroecosystems: responses to land use intensity and farming management. **Mastozoologia Neotropical**, v. 24, n. 2, p. 289-300, 2017.

GONÇALVES, T. V.; PARREIRA, M. R.; NABOUT, J. C. Brazilian protected areas that are larger, older, and closer to urban areas are more studied by scientists. **Biological Conservation**, v. 257, 109123, 2021.

GONÇALVES-OLIVEIRA, J.; ROZENTAL, T.; GUTERRES, A.; TEIXEIRA, B. R.; ANDRADE-SILVA, B. E.; COSTA-NETO, S. F.; FURTADO, M. C.; MORATELLI, R.; D'ANDREA, P. S.; LEMOS, E. R. S. Investigation of *Bartonella* spp. in Brazilian mammals with emphasis on rodents and bats from the Atlantic Forest. **International Journal for Parasitology: Parasites and Wildlife**, v. 13, p. 80-89, 2020.

GONÇALVES-SOUZA, D.; VILELA, B.; PHALAN, B.; DOBROVOLSKI, R. The role of protected areas in maintaining natural vegetation in Brazil. **Science Advances**, v. 7, n. 38, eabh2932, 2021.

GONZÁLEZ, J. D. M.; FONSECA, E. L.; PEREZ, N. B. Invasion by *Eragrostis plana* Nees in areas of the Brazilian Pampa biome modelled with remotely sensed data and GARP species distribution model. **The International Journal of Engineering and Science (IJES)**, v. 9, n. 6, p. 14-20, 2020.

GONZÁLEZ-SOLÍS, J.; GUIX, J. C.; MATEOS, E.; LLORENS, L. Population density of primates in a large fragment of the Brazilian Atlantic rainforest. **Biodiversity and Conservation**, v. 10, p. 1267-1282, 2001.

GOODCHILD, M. F. Citizens as sensors: Web 2.0 and the volunteering of geographic information. **GeoFocus**, v. 7, p. 8-10, 2007.

GORDO, M.; CALLEIA, F. O.; VASCONCELOS, S. A.; LEITE, J. J. F.; FERRARI, S. F. The Challenges of Survival in a Concrete Jungle: Conservation of the Pied Tamarin (*Saguinus bicolor*) in the Urban Landscape of Manaus, Brazil. In: MARSH, L., CHAPMAN, C. (eds.). **Primates in Fragments**. Developments in Primatology: Progress and Prospects. New York, NY, USA: Springer, 2013. p. 357-370.

GORELICK, N.; HANCHER, M.; DIXON, M.; ILYUSHCHENKO, S.; THAU, D.; MOORE, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. **Remote Sensing of Environment**, v. 202, p. 18-27, 2017.

GRAEFE, S.; DUFOUR, D.; ZONNEVELD, M.; RODRIGUEZ, F.; GONZALEZ, A. Peach palm (*Bactris gasipaes*) in tropical Latin America: implications for biodiversity conservation, natural resource management and human nutrition. **Biodiversity and Conservation**, v. 22, p. 269-300, 2013.

GRECCHI, R. C.; GWYN, Q. H. J.; BÉNIÉ, G. B.; FORMAGGIO, A. R.; FAHL, F. C. Land use and land cover changes in the Brazilian Cerrado: A multidisciplinary approach to assess the impacts of agricultural expansion. **Applied Geography**, v. 55, p. 300-312, 2014.

DE GROOT, R. S.; ALKEMADE, R.; BRAAT, L.; HEIN, L.; WILLEMEN, L. Challenges in integrating the concept of ecosystem services and values in landscape planning, management and decision making. **Ecological Complexity**, v. 7, n. 3, p. 2660-272, 2010.

GROTTA-NETO, F.; DUARTE, J. M. B. Movements of Neotropical Forest Deer: What Do We Know? *In*: REYNA-HURTADO, R., CHAPMAN, C. (eds.). **Movement Ecology of Neotropical Forest Mammals**. Edinburg, Scotland: Springer, Cham, 2019. p. 95–109.

GUERRA, A.J.T.; COELHO, M.C.N. **Unidades de Conservação: abordagens e características geográficas**. 1. ed. Rio de Janeiro: Bertrand Brasil, 2012. 298 p.

GUETA, T.; CARMEL, Y. Quantifying the value of user-level data cleaning for big data: A case study using mammal distribution models. **Ecological Informatics**, v. 34, p. 139-145, 2016.

GUILLEM, E. E.; MURRAY-RUST, D.; ROBINSON, D. T.; BARNES, A.; ROUNSEVELL, M. D. A. Modelling farmer decision-making to anticipate tradeoffs between provisioning ecosystem services and biodiversity. **Agricultural Systems**, v. 137, p. 12-23, 2015.

GUO, J.; HUANG, C.; HOU, J. A Scalable Computing Resources System for Remote Sensing Big Data Processing Using GeoPySpark Based on Spark on K8s. **Remote Sensing**, v. 14, n. 3, 521, 2022.

GUISAN, A.; ZIMMERMANN, N. E. Predictive habitat distribution models in ecology. **Ecological Modelling**, v. 135, n. 2-3, p. 147-186, 2000.

GUISAN, A.; THUILLER, W. Zimmermann, N.E. **Habitat suitability and distribution models with applications in R**. Cambridge, UK: Cambridge University Press, 2017. 462 p.

GURGEL, H.; HARRAVE, J.; FRANÇA, F.; HOLMES, R. M. RICARTE, F. M.; DIAS, B. F. S.; RODRIGUES, C. G. O.; BRITO, M. C. W. Unidades de Conservação e o falso dilema entre conservação e desenvolvimento. **Boletim Regional, Urbano e Ambiental**, v. 3, p. 109-119, 2009. Disponível em <[http://repositorio.ipea.gov.br/bitstream/11058/4684/1/BRU\\_n03.pdf](http://repositorio.ipea.gov.br/bitstream/11058/4684/1/BRU_n03.pdf)>. Acesso em 30 nov. 2022.

GUZMÁN, S. D. M.; DRUMMOND, J. A.; BARRETO, C. G. Undesirable neighbours: Eucalyptus and protected areas. *In*: SURATMAN, M.N. **Protected Area Management**. London, UK: IntechOpen, 2022. p. 343-353.

HAKLAY, M. Geographic citizen science: an overview. *In*: SKARLATIDOU, A.; HAKLAY, M. eds.). **Geographic Citizen Science Design: No one left behind**. London, UK: UCL Press, 2021. p. 15-37.

HALPERN, B. S.; BERLOW, E.; WILLIAMS, R.; BORER, E. T.; DAVIS, F. W.; DOBSON, A.; ENQUIST, B. J.; FROEHLICH, H. E.; GERBER, L. R.; LORTIE, C. J.; O'CONNOR, M. I.; REGAN, H.; VÁZQUEZ, D. P.; WILLARD, G. Ecological Synthesis and its roles in advancing knowledge. **BioScience**, v. 70, n. 11, p. 1005-1014, 2020.

HAINES-YOUNG, R. Land use biodiversity relationships. **Land Use Policy**, v. 265, p. S178-S186, 2009.



HANSEN, P. J. Effects of heat stress on mammalian reproduction. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 364, p. 3341–3350, 2009.

HAPPI, C.; ADETIFA, I.; MBALA, P.; NJOUOM, R.; NAKOUNE, E.; HAPPI, A.; NDODO, N.; AYANSOLA, O.; MBOOWA, G.; BEDFORD, T.; NEHER, R.A.; ROEMER, C.; HODCROFT, E.; TEGALLY, H.; O'TOOLE, A.; RAMBAUT, A.; PYBUS, O.; KRAEMER, M.U.G.; WILKINSON, E.; ISIDRO, J.; BORGES, V.; PINTO, M.; GOMES, J.P.; FREITAS, L.; RESENDE, P. C.; LEE, R. T. C.; MAURER-STROH, S.; BAXTER, C.; LESSELS, R.; OGWELL, A. E.; KEBEDE, Y.; TESSEMA, S. K.; OLIVEIRA, T. Urgent need for a non-discriminatory and non-stigmatizing nomenclature for monkeypox virus. **PLoS Biology**, v. 20, n. 8, e3001769, 2022.

HARRISON, M. **Machine Learning: Guia de Referência Rápida - Trabalhando com dados estruturados em Python**. São Paulo: O'Reilly / Novatec, 2020. 272 p.

HARVEY, E.; HOLMES, E. C. Diversity and evolution of the animal virome. **Nature Reviews Microbiology**, v. 20, p. 321–334, 2022.

HEINRICH, M.; MAH, J.; AMIRKIA, V. Alkaloids Used as Medicines: Structural Phytochemistry Meets Biodiversity—An Update and Forward Look. **Molecules**, v. 26, n. 7, 1836, 2021.

HEIL, B. J.; HOFFMAN, M. M.; MARKOWETZ, F.; LEE, S. I.; GREENE, C. S.; HICKS, S.C. Reproducibility standards for machine learning in the life sciences. **Nature Methods**, v. 18, p. 1132-1135, 2021.

HENGL, H. Finding the right pixel size. **Computers & Geosciences**, v. 32, p. 1283-1398, 2006.

HENRIQUE, W. A cidade e a natureza: a apropriação, a valorização e a sofisticação da natureza nos empreendimentos imobiliários de alto padrão em São Paulo. **GEOUSP Espaço e Tempo (Online)**, v. 10, n. 1, p. 65-77, 2006. Disponível em <<https://www.revistas.usp.br/geousp/article/view/74008>>. Acesso em 09 set. 2022.

HERNANDEZ; P. A.; GRAHAM, C. H.; MASTER, L. L.; ALBERT, D. L. The effect of sample size and species characteristics on performance of different species distribution modeling methods. **Ecography**, v. 29, n. 5, p. 773-785, 2006.

HILÁRIO, R. R.; CARVALHO, W. D.; GHELER-COSTA, C.; ROSALINO, L. M. C.; MARQUES, T. A.; ADANIA, C. H.; PAULINO, J. S.; ALMEIDA, P. M.; MUSTIN, K. Drivers of human-wildlife impact events involving mammals in Southeastern Brazil. **Science of The Total Environment**, v. 794, 148600, 2021.

HILL, C. M. Primate Crop Feeding Behavior, Crop Protection, and Conservation. **International Journal of Primatology**, v. 38, p. 385–400, 2017.

HOCKINGS, M. Systems for Assessing the Effectiveness of Management in Protected Areas. **BioScience**, v. 53, n. 9, p. 823-832, 2003.

HOEFEFEL, J. L.; RIBEIRO, P. M.; RAMOS, J. R.; LAS-CASAS E NOVAES, M. C.; PINHEIRO, M.; SIMÕES, W. C. B.; BARBARICK, D.; NICKEL, D. Environmental Issues and Urban Expansion—A Study of the Atibaia River Basin in São Paulo, Brazil. *In*: LEAL FILHO, W.; AZEITEIRO, U.M.; SETTI, A.F.F. (eds.). **Sustainability in Natural Resources**

**Management and Land Planning.** World Sustainability Series. Edinburg, Scotland: Springer, Cham, 2021. p 327-341.

HOBERN, D.; BAPTISTE, B.; COPAS, K.; GURALNICK, R.; HAHN, A.; VAN HUIS, E.; KIM, E. S.; MCGEOCH, M.; NAICKER, I.; NAVARRO, L.; NOESGAARD, D.; PRICE, M.; RODRIGUES, A.; SCHIGEL, D.; SHEFFIELD, C. A.; WIECZOREK, J. Connecting data and expertise: a new alliance for biodiversity knowledge. **Biodiversity Data Journal**, v. 7, e33679, 2019.

HOFFMANN, J.; PALME, R.; ECCARD, J. A. Long-term dim light during nighttime changes activity patterns and space use in experimental small mammal populations. **Environmental Pollution**, v. 238, p. 844-851, 2018.

HONG, E. M.; PACHEPSKY, Y. A.; WHELAN, G.; NICHOLSON, T. Simpler models in environmental studies and predictions. **Critical Reviews in Environmental Science and Technology**, v. 47, n. 18, p. 1669-1712, 2017.

HSU, B. M. Comparison of Supervised Classification Models on Textual Data. **Mathematics**, v. 8, n. 5, 851, 2020.

HU, W.; WANG, Y.; ZHANG, D.; YU, W.; CHEN, G.; XIE, T.; LIU, Z.; MA, Z.; DU, J.; CHAO, B.; LEI, G.; CHEN, B. Mapping the potential of mangrove forest restoration based on species distribution models: A case study in China. **Science of The Total Environment**, v. 748, 142321, 2020.

HUNT, J. R.; BALDOCCHI, D. D.; VAN INGEN, C. Redefining ecological science using data. In: HEY, A.; TRANSLEY, S.; TOLLE, K. (orgs.). **The Fourth Paradigm: data-intensive scientific discovery**. Redmond, Washington, USA: Microsoft Research, 2009. p. 21-26.

HURST, Z. M.; MCCLEERY, R. A.; COLLIER, B. A.; SILVY, N. J.; TAYLOR, P. J.; MONADJEM, A. Linking changes in small mammal communities to ecosystem functions in an agricultural landscape. **Mammalian Biology**, v. 79, n. 1, p. 17-23, 2014.

HURTADO, C. M.; BURTON, A. C. Using habitat suitability information from across a species range to inform understudied ecosystems. **Journal of Mammalogy**, v. 103, n. 4, p. 814-925, 2022.

HUTCHINSON, G. Concluding remarks. In: COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY, 22., 1957, Huntington, NY, USA. **Proceedings** [...]. Huntington, NY: Cold Spring Harbor Laboratory Press, 1957. p. 415-427. Disponível em <<http://symposium.cshlp.org/>>. Acesso em 23 jan 2023.

INMAN, R.; FRANKLIN, J.; ESQUE, T. NUSSEAR, K. Comparing sample bias correction methods for species distribution modeling using virtual species. **Ecosphere**, v. 12, n. 3, e03422, 2021.

ISSAD, S. M.; BENHAFRI, N.; ALLALI, K. E.; FARSI, H.; OUALI-HASSENAOUI, S.; DEKAR-MADOUI, A. Effects of prolonged night-time light exposure and traffic noise on the behavior and body temperature rhythmicity of the wild desert rodent, *Gerbillus tarabuli*. **Chronobiology International**, v. 38, n. 3, p. 415-425, 2021.

IVIE, P.; THAIN, D. Reproducibility in scientific computing. **ACM Computing Surveys (CSUR)**, v. 51, n. 3, p. 1-36, 2018.

JAFARI, R. **Hands-On Data Preprocessing in Python: Learn how to Effectively Prepare Data for Successful Data Analytics**. Birmingham, UK: Packt Publishing, 2022. 602 p.

JARNEVICH, C. S.; HAYES, M. A.; FITZGERALD, L. A.; ADAMS, A. A. Y.; FALK, B. G.; COLLIER, M. A. M.; BONEWELL, L. R.; KLUG, P. E.; NARETTO, S.; REED, R. N. Modeling the distributions of tegu lizards in native and potential invasive ranges. **Scientific Reports**, v. 8, 10193, 2018.

JASINSKI, E.; MORTON, D.; DEFRIES, R.; SHIMABUKURO, Y.; ANDERSON, L.; HANSEN, M. Physical Landscape Correlates of the Expansion of Mechanized Agriculture in Mato Grosso, Brazil. **Earth Interactions**, v. 9, n. 16, p. 1-18, 2005.

JESUS, S. C.; CATOJO, A. M. Z.. Distribuição espacial dos fragmentos florestais de Cerrado no Estado de São Paulo. *In*: WORKSHOP DO CERRADO DA UFSCAR, 1., 2018, São Carlos. **Anais [...]** São Carlos: UFSCar, 2016. p. 15.

JETZ, W.; MCGEOCH, M. A.; GURALNICK, R.; FERRIER, S.; BECK, J.; COSTELLO, M. J.; FERNANDEZ, M.; GELLER, G. N.; KEIL, P.; MEROW, C.; MEYER, C.; MULLER-KARGER, F. E.; PEREIRA, H. M.; REGAN, E. C.; SCHMELLER, D. S.; TURAK, E. Essential biodiversity variables for mapping and monitoring species populations. **Nature Ecology & Evolution**, v. 3, p. 539-551, 2019.

KAFASH, A.; ASHRAFI, S.; YOUSEFI, M. Modeling habitat suitability of bats to identify high priority areas for field monitoring and conservation. **Environmental Science and Pollution Research**, v. 29, p. 25881-25891, 2022.

KAJIN, M.; CERQUEIRA, R.; VIEIRA, M. V.; GENTILE, R. Nine-year demography of the black-eared opossum *Didelphis aurita* (Didelphimorphia: Didelphidae) using life tables. **Revista Brasileira de Zoologia**, v. 25, n. 2, p. 206-213, 2008.

KALLE, R.; RAMESH, T.; QURESHI, Q.; SANKAR, K. Predicting the Distribution Pattern of Small Carnivores in Response to Environmental Factors in the Western Ghats. **PLoS ONE**, v. 8, n. 11, e79295, 2013.

KAMILARIS, A.; OSTERMANN, F. O. Geospatial Analysis and the Internet of Things. **ISPRS International Journal of Geo-Information**, v. 7, n. 7, 269, 2018.

KAY, R. F.; MADDEN, R. H. Mammals and rainfall: paleoecology of the middle Miocene at La Venta (Colombia, South America). **Journal of Human Evolution**, v. 32, n. 2-3, p. 161-199, 1997.

KAYS, R.; SHEPPARD, J.; MCLEAN, K.; WELCH, C.; PAUNESCU, C.; WANG, V.; KRAVIT, G.; CROFOOT, M. Hot monkey, cold reality: surveying rainforest canopy mammals using drone-mounted thermal infrared sensors. **International Journal of Remote Sensing**, v. 40, n. 2, p. 407-419, 2019.

KE, G.; MENG, Q.; FINLEY, T.; WANG, T.; CHEN, W.; MA, W.; YE, Q.; L, T.Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *In*: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 30., 2017, Long Beach, CA, USA.

**Proceedings** [...]. La Jolla, California, USA: Neural Information Processing Systems Foundation, 2017. p. 2659-2669.

KENNEDY, S.; CALAOR, J.; ZURÁPITI, Y.; HANS, J.; YOSHIMURA, M.; CHOO, J.; ANDERSEN, J. C.; CALLAGHAN, J.; RODERICK, G. K.; KREHENWINKEL, H.; ROGERS, H.; GILLESPIE, R. G.; ECONOMO, E. P. Richness and resilience in the Pacific: DNA metabarcoding enables parallelized evaluation of biogeographic patterns. **Molecular Ecology**, Early View, 2022. Disponível em < <https://doi.org/10.1111/mec.16575>>. Acesso em 07 set. 2022.

KHAN, Z. A.; YUMNAMCHA, T.; MONDAL, G.; DECI, S. D.; RAJIV, C.; LABALA, R. K.; DEVI, H. S.; CHATTORAJ, A. Artificial Light at Night (ALAN): A Potential Anthropogenic Component for the COVID-19 and HCoV's Outbreak. **Frontiers in Endocrinology**, v. 11, 622, 2020.

KIM, M.; HONG, S. H. Relationship between the reflected brightness of artificial lighting and land-use types: a case study of the University of Arizona campus. **Landscape and Ecological Engineering**, v. 11, 39-45, 2015.

KIMMINS, J. P.; BLANCO, J. A.; SEELY, B.; WELHAM, C.; SCOULLAR, K. Complexity in modelling forest ecosystems: How much is enough? **Forest Ecology and Management**, v. 256, n. 10, p. 1646-1658, 2008.

KIRSCHNER, J.; KAPLAN, Ž. Taxonomic monographs in relation to global Red Lists. **Biodiversity and Conservations**, v. 51, n. 1, p. 155-158, 2002.

KISSLING, W. D.; DORMANN, C. F.; GROENEVELD, J.; HICKLER, T.; KÜHN, I.; MCINERNEY, G. J.; MONTOYA, J. M.; RÖMERMANN, C.; SCHIFFERS, K.; SCHURR, F. M.; SINGER, A.; SVENNING, J. C.; ZIMMERMANN, N. E.; O'HARA, R. B. Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. **Journal of Biogeography**, v. 39, n. 12, p. 2163-2178, 2012.

KONKOL, M.; KRAY, C.; PFEIFFER, M. Computational reproducibility in geoscientific papers: Insights from a series of studies with geoscientists and a reproduction study. **International Journal of Geographical Information Science**, v. 22, n. 2, p. 408-429, 2019.

KORDON, A. K. **Applying Data Science: How to Create Value with Artificial Intelligence**. New York, USA: Springer International Publishing, 2021. 494 p.

KORSTANJE, J. **Advanced Forecasting with Python**. Berkeley, CA, USA: Apress, 2021. 296 p.

KRAUSMAN, P.; NAUGLE, D. E.; FRISINA, M. R.; NORTHRUP, R.; BLEICH, V. C.; BLOCK, W. M.; WALLACE, M. C.; WRIGHT, J. D. Livestock Grazing, Wildlife Habitat, and Rangeland Values. **Rangelands**, v. 31, n. 5, p. 15-19, 2009.

KREBS, J. R.; DAVIES, N. B. **Introdução à Ecologia Comportamental**. 3.ed. São Paulo: Atheneu Editora, 1996. 420 p.

LACERDA, A. C. R.; TOMAS, W. M.; MARINHO-FILHO, T. J. Domestic dogs as an edge effect in the Brasília National Park, Brazil: interactions with native mammals. **Animal Conservation**, v. 12, n. 5, p. 477-487, 2009.

LACHER JR, T. E.; DAVIDSON, A. D.; FLEMING, T. H.; GÓMEZ-RUIZ, E. P.; MCCRACKEN, G. F.; OWEN-SMITH, N.; PERES, C. A.; VANDER WALL, S. B. The functional roles of mammals in ecosystems. **Journal of Mammalogy**, v. 100, n. 3, p. 942–964, 2019.

LAHOZ-MONFORT, J. J.; MAGRATH, M. J. L. A Comprehensive Overview of Technologies for Species and Habitat Monitoring and Conservation. **BioScience**, v. 71, n. 10, p. 1038-1062, 2021.

LAMBECK, R. J. Focal species: a multi-species umbrella for nature conservation. **Conservation Biology**, v. 11, p.849–856, 1997.

LAMPRECHT, A. L.; GARCIA, L.; KUZAK, M.; MARTINEZ, C.; ARCILA, R.; MARTIN DEL PICO, E.; DOMINGUEZ DEL ANGEL, V.; VAN DE SANDT, S.; ISON, J.; MARTÍNEZ, P. A.; MCQUILTON, P.; VALENCIA, A.; HARROW, J.; PSOMOPOULOS, F.; GELPI, J. L.; CHUE HONG, N.; GOBLE, C.; CAPELLA-GUTIERREZ, S. Towards FAIR principles for research software. **Data Science**, v. 3, n. 1, p. 37-59, 2020.

LARANJEIRO, T.; MAY, R.; VERONES, F. Impacts of onshore wind energy production on birds and bats: recommendations for future life cycle impact assessment developments. **The International Journal of Life Cycle Assessment**, v. 23, p. 2007–2023, 2018.

LAURANCE, W. F. et al. Averting biodiversity collapse in tropical forest protected areas. **Nature**, v. 489, n. 7415, p. 290-294, 2012.

LAURIA, V.; GRISTINA, M.; ATTRILL, M. J.; FIORENTINO, F.; GAROFALO, G. Predictive habitat suitability models to aid conservation of elasmobranch diversity in the central Mediterranean Sea. **Scientific Reports**, v. 5, 13245, 2015.

LAVOR, C. S.; QUEIROZ, F. S.; BACELAR, P. S.; ROSA, A. T. M.; DINIZ, M. C. Prospecção tecnológica sobre estudos bioacústicos em morcegos. **Journal of Biotechnology and Biodiversity**, v. 10, n. 2, p. 137–143, 2022.

LEEMPOEL, K.; HEBERT, T.; HADLY, E. A. A comparison of eDNA to camera trapping for assessment of terrestrial mammal diversity. **Proceedings of the Royal Society B: Biological Sciences**, v. 287, n. 1918, 20192353, 2020.

LEME MACHADO, P. A. L. **Direito Ambiental Brasileiro**. 28.<sup>a</sup> ed. São Paulo: Malheiros Editores / Juspodivm, 2022. 1232 p.

LESSA, I.; GUIMARÃES, T. C. S.; BERGALLO, H. G.; CUNHA, A.; VIEIRA, E. M. Domestic dogs in protected areas: a threat to Brazilian mammals? **Natureza & Conservação**, v. 14, n. 2, p. 46-56, 2016.

LEVENE, H. Robust tests for equality of variance. *In*: OLKIN, I; HOTELLING, H. (eds.). **Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling**. Redwood City, Califórnia, USA: Stanford University Press, 1960. p. 278–292, 1960.

LEVERINGTON, F.; COSTA, K. L.; PAVESE, H.; LISLE, A.; HOCKINGS, M. A Global analysis of protected area management effectiveness. **environmental management**, v. 46, n. 5, p. 685-698, 2010.

- LI, R., POWERS, R., XU, M., ZHENG, Y., ZHAO, S. Proposed biodiversity conservation areas: gap analysis and spatial prioritization on the inadequately studied Qinghai Plateau, China. **Nature Conservation**, v. 24, p.1-20, 2018.
- LI, J. S.; HAMANN, A.; BEAUBIEN, E. Outlier detection methods to improve the quality of citizen science data. **International Journal of Biometeorology**, v. 64, p. 1825-1833, 2020.
- LI, M.; FU, Q.; SINGH, V. P.; JI, Y.; LIU, D.; ZHANG, C.; LI, T. An optimal modelling approach for managing agricultural water-energy-food nexus under uncertainty. **Science of The Total Environment**, v. 651, n. 1, p. 1416-1434, 2019.
- LI, S. L.; ACOSTA, A. L.; HILL, S. C.; BRADY, O. J.; ALMEIDA, M. A. B.; CARDOSO, J. C.; HAMLET, A.; MUCCI, L. F.; DEUS, J. T.; IANI, F. C. M.; ALEXANDER, N. S.; WINT, G. R. W.; PYBUS, O. G.; KRAEMER, M. U. G.; FARIA, N. R.; MESSINA, J. P. Mapping environmental suitability of *Haemagogus* and *Sabethes* spp. mosquitoes to understand sylvatic transmission risk of yellow fever virus in Brazil. **PLoS Neglected Tropical Diseases**, v. 16, n. 1, e0010019, 2022.
- LIMA, F. P. “**Sem morcego, não tem pequi**”: percepção e conhecimento sobre os serviços ecossistêmicos para a conservação do Cerrado. 2019. Tese (Doutorado em Recursos Naturais do Cerrado RENAC) - Câmpus Anápolis de Ciências Exatas e tecnológicas - Henrique Santillo, Universidade Estadual de Goiás, Anápolis, 2019.
- LIMA, E. A. C. F.; RANIERI, V. E. L. Land use planning around protected areas: Case studies in four state parks in the Atlantic forest region of southeastern Brazil. **Land Use Policy**, v. 71, p. 453-458, 2018.
- LIMA, R. A. F.; SOUZA, V. C.; SIQUEIRA, M. F.; STEEGE, H. Defining endemism levels for biodiversity conservation: Tree species in the Atlantic Forest hotspot. **Biological Conservation**, v. 252, 108825, 2020.
- LIMA, V. P.; LIMA, R. A. F.; JONER, F.; SIDDIQUE, I. RAES, N.; STEEGE, H. Climate change threatens native potential agroforestry plant species in Brazil. **Scientific Reports**, v. 12, 2267, 2022.
- LIN, C. C. Ecoinformatics: A Review of Approach and Applications in Ecological Research. **Proceedings of National Institute of Ecology**, v. 1, n. 1, p. 9-21, 2020.
- LINERO, D.; CUERVO-ROBAYO, A. P.; ETTER, A. Assessing the future conservation potential of the Amazon and Andes Protected Areas: Using the woolly monkey (*Lagothrix lagothricha*) as an umbrella species. **Journal for Nature Conservation**, v. 58, 125926, 2020.
- LIZCANO, D.; PEDRAZA, C., ROMERO, L., CAMPOS-CERQUEIRA, M. Acoustic monitoring in nature travel destinations of Colombia. v1.1. Awake Travel. Dataset/Occurrence. 2022. Disponível em <<https://doi.org/10.15472/kzo6hg>>. Acesso em 09 set. 2022.
- LOCKERS, R.; KNAPEN, R.; JANSSEN, S.; VAN RANDEN, Y.; JANSEM J. Analysis of Big Data technologies for use in agro-environmental science. **Environmental Modelling & Software**, v. 84, p. 494-504, 2016.
- LOURENÇO-DE-MORAES, R.; LANSAC-TOHA, F. M. SCHWIND, L. T. F. ARRIEIRA, R. L.; ROSA, R. R.; TERRIBILE, L. C.; LEMES, P.; RANGEL, T. F.; DINIZ-FILHO, A. F.;

- BASTOS, R. P.; BAILLY, D. Climate change will decrease the range size of snake species under negligible protection in the Brazilian Atlantic Forest hotspot. **Scientific Reports**, v. 9, 8523, 2019.
- LOVEGROVE, B. G. The evolution of endothermy in Cenozoic mammals: a plesiomorphic-apomorphic continuum. **Biological Reviews**, v. 8, n. 1, p. 128-162, 2012.
- LOPION, P.; MARKEWITZ, P.; ROBINIUS, M.; STOLTEN, D. A review of current challenges and trends in energy systems modeling. **Renewable and Sustainable Energy Reviews**, v. 96, p. 156-166, 2018.
- LYET, A.; PELLISSIER, L. VALENTINI, A.; DEJEAN, T.; HEHMEYER, A.; NAIDOO, R. eDNA sampled from stream networks correlates with camera trap detection rates of terrestrial mammals. **Scientific Report**, v. 11, 11362, 2021.
- MA, T. Quantitative Responses of Satellite-Derived Nighttime Lighting Signals to Anthropogenic Land-Use and Land-Cover Changes across China. **Remote Sensing**, v. 10, n. 9, 1447, 2018.
- MAASS, J. M.; BALVANERA, P.; CASTILLO, A.; DAILY, G. C.; MOONEY, H. A.; EHRLICH, P.; QUESADA, M.; MIRANDAM A.; JARAMILLO, V. J.; GARCÍA-OLIVA, F.; MARTÍNEZ-YRIZAR, A.; COTLER, H.; LÓPEZ-BLANCO, J.; PÉREZ-JÍMENEZ, A.; BÚRQUEZ, A.; TINOCO, C.; CEBALLOS, G.; BARRAZA, L.; AYALA, R.; SARUKHÁN, J. Ecosystem Services of Tropical Dry Forests: Insights from Long-term Ecological and Social Research on the Pacific Coast of Mexico. **Ecology and Society**, v. 10, n. 1, 17, 2005.
- MAESTRI, R.; MARINHO, J. R. Singing in the rain. Rainfall and moonlight affect daily activity patterns of rodents in a Neotropical forest. **Acta Theriologica**, v. 59, n. 3, p. 427-433, 2014.
- MAHER, C. R.; BURGER, J. R. Intraspecific variation in space use, group size, and mating systems of caviomorph rodents. **Journal of Mammalogy**, v. 92, n. 1, p. 54-64, 2011.
- MALAKOUTIKHAH, S.; FAKHERAN, S.; HEMAMI, M. R.; TARKESH, M.; SENN, J. Assessing future distribution, suitability of corridors and efficiency of protected areas to conserve vulnerable ungulates under climate change. **Diversity and Distributions**, v. 26, n. 10, p. 1383-1396, 2020.
- MALAWSKA, A.; TOPPING, C. J.; NIELSEN, H. Ø. Why do we need to integrate farmer decision making and wildlife models for policy evaluation? **Land Use Policy**, v. 38, p. 732-740, 2014.
- MALDONADO, C.; MOLINA, C. I.; ZIZKA, A.; PERSSON, C.; TAYLOR, C. M.; ALBÁN, J.; CHILQUILLO, E.; RØNSTED, N.; ANTONELLI, A. Estimating species diversity and distribution in the era of Big Data: to what extent can we trust public databases? **Global Ecology and Biogeography**, v. 24, p. 973-984, 2015.
- MANCUSO, M. A. C.; SORATTO, R. P.; PERDONÁ, M. J. Produção de café sombreado. **Colloquium Agrariae**, v. 9, n.1, p. 31-44, 2013.
- MAPBIOMAS. **MapBiomias General “Handbook”**: Algorithm Theoretical Basis Document (ATBD). Collection 6. 49 p. 2022. Disponível em <<https://mapbiomas.org/download-dos-atbds>>. Acesso em 01 out. 2021.

- DE MARCO JÚNIOR, P.; NÓBREGA, C. C. Evaluating collinearity effects on species distribution models: An approach based on virtual species simulation. **PLoS ONE**, v. 13, n.9, e0202403, 2018.
- MARES, M. A. The geobiological interface: Granitic outcrops as a selective force in mammalian evolution. **Journal of the Royal Society of Western Australia**, v. 80, n. 3, 1997.
- MARQUES NETO, R. **Zoogeografia do Brasil: a fauna, a paisagem e as organizações espaciais**. Curitiba, PR: CRV. 2018. 390 p.
- MARTIN, P. S.; GHELIER-COSTA, C.; LOPES, P. C.; ROSALINO, L. M.; VERDADE, L. M. Terrestrial non-volant small mammals in agro-silvicultural landscapes of Southeastern Brazil. **Forest Ecology and Management**, v. 282, p. 185-195, 2012.
- MATOS, T. P. V.; MATOS, V. P. V.; MELLO, K.; VALENTE, R. A. Protected areas and forest fragmentation: sustainability index for prioritizing fragments for landscape restoration. **Geology, Ecology, and Landscapes**, v. 5, n. 1, p. 19-31, 2021.
- MATUTINI, F.; BAUDRY, J.; PAIN, G.; SINEAU, M.; PITHON, J. How citizen science could improve species distribution models and their independent assessment. **Ecology and Evolution**, v. 11, p. 3028-3039, 2021.
- MAZZA, V.; DAMMHAHN, M. LÖSCHE, E.; ECCARD, J. A. Small mammals in the big city: Behavioural adjustments of non-commensal rodents to urban environments. **Global Change Biology**, v. 26, n. 11, p. 6326-6337, 2020.
- MCCLEERY, R. Urban Mammals. *In*: AITKENHEAD-PETERSON, J.; VOLDER, A. (eds.). **Urban Ecosystem Ecology**. Madison, Wisconsin, USA: American Society of Agronomy, 2010. p. 87-102.
- MEDEIROS, R.; YOUNG, C. E. F.; PAVESE, H. B.; ARAÚJO, F. F. S. **Contribuição das unidades de conservação brasileiras para a economia nacional: Sumário Executivo**. Brasília: UNEP-WCMC, 2011. 44p. Disponível em <[https://ava.icmbio.gov.br/pluginfile.php/4592/mod\\_data/content/21005/estudocontribuicao.pdf](https://ava.icmbio.gov.br/pluginfile.php/4592/mod_data/content/21005/estudocontribuicao.pdf)>. Acesso em 03 dez. 2022.
- MEDINA, F. M.; BONNAUD, E.; VIDAL, E.; NOGALES, M. Underlying impacts of invasive cats on islands: not only a question of predation. **Biodiversity and Conservation**, v. 23, p. 327–342, 2014.
- MEIRELLES FILHO, J.; BORGES, C. R. S. (orgs.). **Patrimônio natural privado no Brasil**. São Paulo: Metalivros, 2016. 270 p.
- MELLO, K.; TOPPA, R. H.; ABESSA, D. M. Dinâmica da paisagem do município de Cubatão: crescimento entre portos, indústrias e a Serra do Mar. **O Mundo da Saúde**, v. 35, n. 1, p. 42-46, 2011.
- MELLO-THÉRY, N. A.; CORREIA, B. O. Pressão urbana em áreas de florestas: história e conflitos políticos da proteção ambiental. **Mercator - Revista de Geografia da UFC**, v. 8, n. 16, p. 33-44, 2009.
- MELLO-THÉRY, N. A. Conservation of natural areas in São Paulo. **Estudos Avançados**, v. 25, n. 71, 2011.



- MENA, J. L.; YAGUI, H.; LA ROSA, F.; PASTOR, P.; RIVERO, J.; APPLETON, R. Influence of topography gradient and seasonality on primate habitat use in Central Amazonia. **Mammalian Biology**, v. 101, p. 251-259, 2021.
- MERCHANT, D.; LATHROP, R. G.; SANTOS, C. D.; PALUDO, D.; NILES, L. SMITH, J. A. M.; FEIGIN, S.; DEY, A. Distribution Modeling and Gap Analysis of Shorebird Conservation in Northern Brazil. **Remote Sensing**, v. 15, n. 2, 452, 2023.
- MICHENER, W. K.; JONES, M. B. Ecoinformatics: supporting ecology as a data-intensive science. **Trends in Ecology and Evolution**, v. 27, n. 2, p. 85-93, 2012
- MILLER, J. Species Distribution Modeling. **Geography Compass**, v. 4, n.6, p 490-509, 2010.
- MILLS, L. S.; SOULÉ, M. E.; DOAK, D. F. The keystone-species concept in Ecology and Conservation. **BioScience**, v. 43, n. 4, p. 210-224, 1993.
- MILLS, S.; WEISS, S.; LIANG, C. VIIRS day/night band (DNB) stray light characterization and correction. *In*: EARTH OBSERVING SYSTEMS, 18., 2013, Bellingham, Washington USA. **Proceedings** [...]. Bellingham, Washington: SPIE, 2013.
- MIRANDA, J. M. D.; PASSOS, F. C. Composição e dinâmica de grupos de *Alouatta guariba clamitans* Cabrera (Primates, Atelidae) em Floresta Ombrófila Mista no Estado do Paraná, Brasil. **Revista Brasileira de Zoologia**, v. 22, n. 1, p. 99-106, 2005.
- MIRANDA, E. B. P.; MENEZES, J. F. S.; FARIÁS, C. C. L.; MUNN, C.; PERES, C. A. Species distribution modeling reveals strongholds and potential reintroduction areas for the world's largest eagle. **PLoS ONE**, v. 14, n. 5, e0216323, 2019.
- MITCHELL, T. M. **Machine Learning**. New York, USA: McGraw-Hill Science, 1997. 414 p.
- MITCHELL, D.; SNELLING, E. P.; HETERM. R. S.; MALONEY, S. K.; STRAUSS, W. M.; FULLER, A. Revisiting concepts of thermal physiology: Predicting responses of mammals to climate change. **Journal of Animal Ecology**, v. 87, n. 4, p. 956-973, 2018.
- MOLIN, P. G.; GERGEL, S. E.; SOARES-FILHO, B.; FERRAZ, S. F. B. Spatial determinants of Atlantic Forest loss and recovery in Brazil. **Landscape Ecology**, v. 32, p. 857-870, 2017.
- MOOERS, B. H. M. Modernizing computing by structural biologists with Jupyter and Colab. *In*: PYTHON IN SCIENCE CONFERENCE (SCIPY 2021), 20., 2021, Austin, Texas, USA. **Proceedings** [...]. [s.l]: SciPy, 2021. Disponível em <<https://pdfs.semanticscholar.org/5426/cae152b08f4fb413a734890a9cc5ab91e413.pdf>>. Acesso em 14 jan. 2023.
- MORAES, R. F. **Agrotóxicos no Brasil**: padrões de uso, política da regulação e prevenção da captura regulatória. Rio de Janeiro: IPEA, 2019. 84 p.
- MORAES, M. C. P.; MELLO, K.; TOPPA, R. T. Protected areas and agricultural expansion: Biodiversity conservation versus economic growth in the Southeast of Brazil. **Journal of Environmental Management**, v. 188, p. 73-84, 2017.

MORAES, B.; RAZGOUR, O.; SOUZA-ALVES, J. P.; BOUBLI, J. P.; BEZERRA, B. Habitat suitability for primate conservation in north-east Brazil. **Oryx**, v. 54, n. 6, 2020.

MORAN, P. The Interpretation of Statistical Maps. **Journal of the Royal Statistical Society**, v. 10, p. 243-251, 1948.

MORETTIN, P. A.; SINGER, J. M. **Estatística e Ciência de Dados**. Rio de Janeiro: LTC, 2023. 464 p.

MOTA-ROJAS, D.; TITTO, C. G.; ORIHUELA, A.; MARTÍNEZ-BURNES, J.; GÓMEZ-PRADO, J.; TORRES-BERNAL, F.; FLORES-PADILLA, K.; CARVAJAL-DE LA FUENTE, V.; WANG, D. Physiological and Behavioral Mechanisms of Thermoregulation in Mammals. **Animals**, v. 116, n. 6, 1733, 2021.

MOURA JÚNIOR, E. G.; NASCIMENTO, F. A. O.; LEMOS FILHO, J. P.; SEVERI, W.; KAMINO, L. H. Y. Limnological layers improve species distribution modeling of aquatic macrophytes at fine-spatial resolution. **Acta Botanica Brasilica**, v. 35, n. 1, p. 9-16, 2021.

MUHAMMAD, K.; ANANDHI, A.; CHEN, G. Comparing Methods for Estimating Habitat Suitability. **Land**, v. 11, 1754, 2022.

MÜLLER, F.; FORHER, N.; CHICHARO, L. The basic ideas of the Ecosystem Service concept. *In*: CHICHARO, L.; MÜLLER, F.; FOHRER, N. (eds). **Ecosystem Services and River Basin Ecohydrology**. Berlin, Germany: Springer, Dordrecht, 2015. P. 7-33. doi: 10.1007/978-94-017-9846-4\_2

MUSSOI, E. M. Segurança Alimentar, agricultura familiar e extensão rural. **Agropecuária Catarinense**, v. 16, n. 3, p. 18-21, 2003.

NATIVI, S.; MAZZETTI, P.; SANTORO, M.; PAPESCHI, F.; CRAGLIA, M.; OCHIAI, O. Big Data challenges in building the Global Earth Observation System of Systems. **Environmental Modelling & Software**, v. 68, p. 1-26, 2015.

NIAZ, M. The role of idealization in science and its implications for science education. **Journal of Science Education and Technology**, v. 8, n. 145-150, 1999.

NICHOLSON, E.; FULTON, E. A.; BROOKS, T. M.; BLANCHARD, R.; LEADLEY, P. METZGER, J. P. MOKANY, K.; STEVENSON, S.; WINTLE, B. A.; WOOLEY, S. N. C.; BARNES, M.; WATSON, J. E. M.; FERRIER, S. Scenarios and models to support global conservation targets. **Trends in Ecology and Evolution**, v. 34, n. 1, p. 57-68, 2019.

NOGUERA-URBANO, E. A.; FERRO, I. Environmental factors related to biogeographical transition zones of areas of endemism of Neotropical mammals. **Australian Systematic Botany**, v. 30, n. 5-6, p. 485-494, 2017.

NORBERG, A.; ABREGO, N.; BLANCHET, G.; ADLER, F. R.; ANDERSON, B. J.; ANTTILA, J.; ARAÚJO, M. B.; DALLAS, T.; DUNSON, D.; ELITH, J.; FOSTER, S. D.; FOX, R.; FRANKLIN, D.; GODSOE, W.; GUISAN, A.; O'HARA, B.; HILL, N. A.; HOLT, R. D.; HUI, F. K. C.; HUSBY, M.; KALAS, J. A.; LEEHIKONEN, A.; LUOTO, M.; MOD, H. K.; NEWELL, G.; RENNER, I.; ROSLIN, T.; SOININEN, J.; THUILLER, W.; VANHATALO, J.; WARTON, D.; WHITE, M.; ZIMMERMANN, N. E.; GRAVEL, D.; OVASKAINEN, O. A comprehensive evaluation of predictive performance of 33 species

distribution models at species and community levels. **Ecological Monographs**, v. 89, n. 3, e01370, 2019.

NORD, A.; GIROUD, S. Lifelong Effects of Thermal Challenges During Development in Birds and Mammals. **Frontiers in Physiology**, v. 11, 419, 2020.

NUNES, H.; ROCHA, F. L.; CORDEIRO-ESTRELA, P. Bats in urban areas of Brazil: roosts, food resources and parasites in disturbed environments. **Urban Ecosystems**, v. 20, p. 953-969, 2017.

NUNES, A.; SCHMITZ, C.; MOURA, S.; MARASCHIN, M. The use of pesticides in Brazil and the risks linked to human health. **Brazilian Journal of Development**, v. 7, n. 4, p. 37885-37904, 2021.

OGDEN, L. E. The Emergence of eDNA: An interdisciplinary tool helps monitor biodiversity and health. **BioScience**, v. 72, n. 1, p. 5-12, 2022.

OKE, O. A.; THOMPSON, K. A. Distribution models for mountain plant species: The value of elevation. **Ecological Modelling**, v. 301, p. 72-77, 2015.

OLIVEIRA, V. B.; LINARES, A. M.; CORRÊA, G. L. C.; CHIARELLO, A. G. Predation on the black capuchin monkey *Cebus nigritus* (Primates: Cebidae) by domestic dogs *Canis lupus familiaris* (Carnivora: Canidae), in the Parque Estadual Serra do Brigadeiro, Minas Gerais, Brazil. **Revista Brasileira de Zoologia**, v. 25, n. 2, p. 376-378, 2008.

OLIVEIRA, W. C.; ARRAES, F. D. D.; OLIVEIRA, J. B.; MACÊDO, K. G.; LIMA JUNIOR, J. C. Energy balance in areas with different land uses in the Chapada do Araripe. **Revista Geama**, v. 3, n. 4, p. 252-256, 2017.

OLIVEIRA, K. D.; TOMASELLA, J.; CARAM R. O.; FALCK, A. S.; SANCHES, I. D. A. Impacts of sugarcane expansion on water availability in a river basin in southeastern Brazil. **CATENA**, v. 216, B, 106, 437, 2022.

OLIVEIRA-SILVA, A. E.; PIRATELLI, A. J.; ZURELL, D.; SILVA, F. R. Vegetation cover restricts habitat suitability predictions of endemic Brazilian Atlantic Forest birds. **Perspectives in Ecology and Conservation**, v. 20, n. 1, p. 1-8, 2022.

OMOHUNDRO, S. M. **Five balltree construction algorithms**. International Computer Science Institute Technical Report. 1989. Disponível em <[https://www.academia.edu/70430197/Five\\_Balltree\\_Construction\\_Algorithms](https://www.academia.edu/70430197/Five_Balltree_Construction_Algorithms)>. Acesso em 04 abr. 2022.

ORTMANN, C. R.; JOHNSON, S. D. How reliable are motion-triggered camera traps for detecting small mammals and birds in ecological studies? **Journal of Zoology**, v. 313, n. 3, p. 202-207, 2021.

OTEGUI, J.; ARIÑO, A. H.; CHAVAN, V. C.; GAIJI, S. On the dates of the GBIF mobilised primary biodiversity data records. **Biodiversity Informatics**, v.8, p. 173-184, 2013.

OTEGUI, J.; GURALNICK, R. P. The geospatial data quality REST API for primary biodiversity data. **Bioinformatics**, v. 32, n. 11, p. 1755-1757, 2016.

OWEN, L. **Hyperparameter Tuning with Python: Boost Your Machine Learning Model's Performance Via Hyperparameter Tuning**. Birmingham, UK: Packt Publishing. 2022. 306 p.

PALMEIRA, F. B. L.; TRINCA, C. T.; HADDAD, C. M. Livestock predation by Puma (*Puma concolor*) in the Highlands of a Southeastern Brazilian Atlantic Forest. **Environmental Management**, v. 56, p. 903–915, 2015.

PAINE, R. T. A note on trophic complexity and community stability. **American Naturalist**, v. 103, p. 91-93, 1966.

PALHARES, J. C. P. **Consumo de água na produção animal**. Comunicado Técnico 102. São Carlos: Embrapa, 2013. Disponível em <<https://ainfo.cnptia.embrapa.br/digital/bitstream/item/92401/1/PROCIJCPP2013.00165.pdf>>. Acesso em 03 dez. 2022.

PALHARES, J. C. P.; MORELLI, M.; NOVELLI, T. I. Water footprint of a tropical beef cattle production system: The impact of individual-animal and feed management. **Advances in Water Resources**, v. 149, 103853, 2021.

PAPADOPOULOS, T.; BALTA, M. E. Climate Change and big data analytics: Challenges and opportunities. **International Journal of Information Management**, v. 63, 102448, 2022.

PARDO, J. M.; CRUZ, P.; MOYA, S.; PIZZIO, E.; FOLETTTO, F.; ROBINO, F.; AQUINO, J.; COSTA, S.; BARROS, Y.; CLEO, F.; DI BITETTI, M. S.; IEZZI, M. E.; PAVIOLO, A.; DE ANGELO, C. Predicting poaching hotspots in the largest remnant of the Atlantic Forest by combining passive acoustic monitoring and occupancy models. **Biological Conservation**, v. 272, 109600, 2022.

PARK, Y. S.; CHON, T. S. Biologically-inspired machine learning implemented to ecological informatics. **Ecological Modelling**, v. 203, n. 1-2, p. 1-7, 2007.

PAS, J. J. W.; PAS, T. E. E. Invited Review: Water consumption, and drinking behavior of beef cattle, and effects of water quality. **Applied Animal Science**, v. 37, n. 4, p. 418-435, 2021.

PATARROYO CAMARGO, G. D; OCHOA SANABRIA, L. P. Foraminíferos presentes en una laguna arrecifal, registrados durante la Expedición Seaflower 2019 - Proyecto Colombia BIO. v1.0. Stratos Consultoría Geológica. Dataset/Occurrence. 2021. Disponível em <<https://doi.org/10.15472/r06dix>>. Acesso em 09 set. 2022.

PAULA, R. R.; OLIVEIRA, I. R.; GONÇALVES, J. L. M.; FERAZ, A. V. Why mixed forest plantation? *In*: CARDOSO, E. J. B. N.; GONÇALVES, J. L. M.; BALIEIRO, F. C.; FRANCO, A. A. **Mixed Plantations of Eucalyptus and Leguminous Trees**. Edinburgh, Scotland: Springer, Cham., 2020. p. 1-13.

PAYÉS, A.C.L.M.; PAVÃO, T.; SANTOS, R.F. The conservation success over time: Evaluating the land use and cover change in a protected area under a long re-categorization process. **Land Use Policy**, v. 30, p. 177-185, 2013.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A. COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine Learning in Python. **Journal of Machine Learning Research**, v. 12, n. 85, p. 2825-2830, 2011.

PENG, R. D. Reproducible Research in Computational Science. **Science**, v. 334, n. 6060, p. 1226-1227, 2011.

PERCEQUILLO, A. R.; DALAPICOLLA, J.; ABREU-JÚNIOR, E. F.; ROTH, P. R. O.; FERRAZ, K. M. P. M. B.; CHIQUITO, E. A. How many species of mammals are there in Brazil? New records of rare rodents (Rodentia: Cricetidae: Sigmodontinae) from Amazonia raise the current known diversity. **PeerJ**, v. 5, e4071, 2017.

PEREIRA, C. H.; PATINO, H. O.; HOSHIDE, A. K.; ABREU, D. C.; ROTZ, A. Grazing supplementation and crop diversification benefits for southern Brazil beef: A case study. **Agricultural Systems**, v. 162, p. 1-9, 2018.

PEREIRA, A. A. B. G.; DIAS, B.; CASTRO, S. I.; LANDI, M. F. A.; MELO, C. B.; WILSON, T. M.; COSTA, G. R. T.; PASSOS, P. H. O.; ROMANO, A. P.; SZABÓ, M. P. J.; CASTRO, M. B. Electrocutions in free-living black-tufted marmosets (*Callithrix penicillata*) in anthropogenic environments in the Federal District and surrounding areas, Brazil. **Primates**, v. 61, p. 321-329, 2020.

PERINO, A.; PEREIRA, H. M.; FELIPE-LUCIA, M.; KIM, H.; KÜHL, H. S.; MARSELLE, M. R.; MEYA, J. N.; MEYER, C.; NAVARRO, L. M.; VAN KLINK, R.; ALBERT, G.; BARRATT, C. D.; BRUELHEIDE, H.; CAO, Y.; CHAMOIN, A.; DARBI, M.; DORNELAS, M.; EISENHAEUER, N.; ESSL, F.; FARWIG, N.; FÖRSTER, J.; FREYHOF, J.; GESCHKE, J.; GOTTSCHALL, F.; GUERRA, C.; HAASE, P.; HICKLER, T.; JACOB, U.; KASTNER, T.; KORELL, L.; KÜHN, I.; LEHMANN, G. U. C.; LENZNER, B.; MARQUES, A.; ŠVARA, E. M.; QUINTERO, L. C.; PACHECO, A.; POPP, A.; ROUET-LEDUC, J.; SCHNABEL, F.; SIEBERT, J.; STAUDE, I. R.; TROGISCH, S.; ŠVARA, V.; SVENNING, J. C.; PE'ER, G.; RAAB, K.; RAKOSY, D.; VANDEWALLE, M.; WERNER, A. S.; WIRTH, C.; XU, H.; YU, D.; ZINNGREBE, Y.; BONN, A. Biodiversity post-2020: Closing the gap between global targets and national-level implementation. **Conservation Letters**, v. 15, n. 2, e12848, 2022.

PETRONI, M. L.; SIQUEIRA-GAY, J.; GALLARDO, A. L. C. F. Understanding land use change impacts on ecosystem services within urban protected areas. **Landscape and Urban Planning**, v. 223, 104404, 2022.

PHILOGENE, S.; NI-MEISTER, W. Relationship between fire events and land use changes in the State of São Paulo, Brazil. **Remote Sensing**, v. 13, n. 15, 2853, 2021.

PIANKA, E. R. On r and K selection. **American Naturalist**, v. 104, p. 592-597, 1970.

PINTO, F. A. S.; CIRINO, D. W.; CERQUEIRA, R. C.; ROSA, C.; FREITAS, S. R. How many mammals are killed on Brazilian roads? Assessing impacts and conservation implications. **Diversity**, v. 14, n. 10, 835, 2022.

PLÁCIDO, R. A. A.; BORGES, S. H.; SILVA, E. G. A protocol to evaluate the potential of protected areas for birdwatching tourism: a study case in the Brazilian Amazon. **Revista Brasileira de Ecoturismo**, v. 14, n. 4, p. 521-553, 2021.

POGGIATO, G.; MÜNKEMÜLLER, T.; BYSTROVA, D.; ARBEL, J.; CLARK, J. THUILLER, W. Community confounding in joint species distribution models. **Trends in Ecology and Evolution**, v. 36, n. 5, p. 391-401, 2021.

POISOT, T.; LABRIE, R.; LARSON, E.; RAHLIN, A.; SIMMONS, B. I. Data-based, synthesis driven: Setting the agenda for computational ecology. **Ideas in Ecology and Evolution**, v. 12, p. 9-21, 2019.

POLLOCK, L. J.; TINGLEY, R.; MORRIS, W. K.; GOLDING, N.; O'HARA, R. B.; PARRIS, K. M.; VESK, P. A.; MCCARTHY, M. A. Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). **Methods in Ecology and Evolution**, v. 5, n. 5, p. 397-406, 2014.

POLLOCK, L. J.; O'CONNOR, L. M. J.; MOKANY, K.; ROSAUER, D. F.; TALLUTO, M. V.; THUILLER, W. Protecting Biodiversity (in All Its Complexity): New Models and Methods. **Trends in Ecology & Evolution**, v. 35, n. 12, p. 1119-1128, 2020.

POPPER, K. **The Logic of Scientific Discovery**. New York, USA: Hutchinson & Co, 1959. 513 p.

PRASETYA, D. A.; NGUYEN, P. T.; FAIZULLIN, R.; ISWANTO, I. ARMA, E. F. Resolving the shortest path problem using the Haversine algorithm. **Journal of Critical Reviews**, v. 7, n. 1, 2020.

PRASNIEWSKI, V. M.; SZINWELSKI, N.; BERTRAND, A. S.; MARTELLO, F.; BROCARD, C. R.; CUHNA, J.; SPERBER, C. F.; VIANA, R.; SANTOS, B. G.; FEARNESIDE, P. M.; SOBRAL-SOUZA, T. Brazil's Iguazu National Park threatened by illegal activities: predicting consequences of proposed downgrading and road construction. **Environmental Research Letters**, v. 17, n. 2, 024024, 2022.

PRESSEY, R. L.; VISCONTI, P.; MCKINNON, M. C.; GURNEY, G. G.; BARNES, M. D.; GLEW, L.; MARON, M. The mismeasure of conservation. **Trends in Ecology and Evolution**, v. 36, n. 9, p 808-821, 2021.

PRIMACK, R. B.; RODRIGUES, E. **Biologia da Conservação**. Londrina: Editora Planta, 2001. 328 p.

PRISTM P. R.; PRADO, A.; TAMBOSI, L. R.; UMETSU, F.; BUENO, A. A.; PARDINI, R.; METZGER, J. P. Moving to healthier landscapes: Forest restoration decreases the abundance of Hantavirus reservoir rodents in tropical forests. **Science of The Total Environment**, v. 752, 141967, 2021.

PUREZA, F.; PELLIN, A.; PADUA, C. **Unidades de Conservação: Fatos e personagens que fizeram a história das categorias de manejo**. 1. ed. São Paulo: Matrix, 2015. 240 p.

QUARANTA, L.; CALEFATO, F.; LANUBILE, F. A Taxonomy of Tools for Reproducible Machine Learning Experiments. *In: INTERNATIONAL CONFERENCE OF THE ITALIAN ASSOCIATION FOR ARTIFICIAL INTELLIGENCE*, 20., 2021, Milão, Itália, 2021. p. 65-76. Disponível em <<https://ceur-ws.org/Vol-3078/>>. Acesso em 14 jan. 2023.

QUEIROZ, J. A.; QUIRINI, Z. G. M.; LOPES, A. V.; MACHADO, I. C. Vertebrate mixed pollination system in *Encholirium spectabile*: A bromeliad pollinated by bats, opossum and hummingbirds in a tropical dry forest. **Journal of Arid Environments**, v. 125, p. 21-30, 2016.

QUINTERO, S.; ABRAHAMS, M. I.; BEIRNE, C.; BLAKE, J.; CARVALHO JR, E.; COSTA, H. C. M.; PAULA, M. J.; ENDO, W.; HAUGAASEN, T.; LIMA, M. G. M.; MICHALSKI, F.; MOSQUERA, D.; NORRIS, D.; OLIVEIRA, T.; PAEMELAERE, E.; PERES, C. A.; PEZZUTI, J.; ROMERO, S.; SANTOS, F.; SILLERO-ZUBIRI, C.; TAN, C. K. W. Effects of human-induced habitat changes on site-use patterns in large Amazonian Forest mammals. **Biological Conservation**, v. 279, 109904, 2023.

RAHMANIAN, S.; POUYAN, S.; KARAMI, S.; POURGHASEMI, H. R. Predictive habitat suitability models for *Teucrium polium* L. using boosted regression trees. In: POURGHASEMI, H. R. (ed.). **Computers in Earth and Environmental Sciences**. Amsterdã, Países Baixos: Elsevier, 2022. p. 245-254.

RAJ, A.; JHARIYA, M. K.; BARGALI, S. S. Bund based agroforestry using Eucalyptus species: a review. **Current Agriculture Research Journal**, v. 4, n. 2, p. 148-158, 2016.

RAMOS, F. N.; SANTOS, F. A. M. Microclimate of Atlantic Forest Fragments: Regional and Local Scale Heterogeneity. **Brazilian Archives of Biology and Technology**, v. 49, n. 6, p. 935-944, 2006.

RAMOS, L. M.; BAZZAN, T.; MOTTA, M. F. B.; BERNARDES, G. P.; GIACHETI, H. L. Landslide susceptibility mapping based on rainfall scenarios: a case study from Sao Paulo in Brazil. **AMS Geosciences**, v. 8, n. 3, p. 438-451, 2022.

RAMOS-RIBEIRO, R. R. Green Belt Biosphere Reserve in the Brazilian City of São Paulo. **Ecological Questions**, v. 20, p. 93-97, 2014.

RATHER, T. A.; KUMAR, S.; KHAN, J. A. Multi-scale habitat modelling and predicting change in the distribution of tiger and leopard using random forest algorithm. **Scientific Reports**, v. 10, 11473, 2019.

RATHER, T. A.; KUMAR, S.; KHAN, J. A. Using machine learning to predict habitat suitability of sloth bears at multiple spatial scales. **Ecological Processes**, v. 10, 48, 2021.

RAWSON, A.; SABEUR, Z.; BRITO, M. Intelligent geospatial maritime risk analytics using the Discrete Global Grid System. **Big Earth Data**, v. 6, n. 3, p. 294-322, 2022.

RESENDE, F. M.; CIMON-MORIN, J.; POULIN, M.; MEYER, L.; JONER, D. C.; LOYOLA, R. The importance of protected areas and Indigenous lands in securing ecosystem services and biodiversity in the Cerrado. **Ecosystem Services**, v. 49, 101282, 2021.

RIBEIRO, B. R.; MARTINS, E.; MARTINELLI, G.; LOYOLA, R. The effectiveness of protected areas and indigenous lands in representing threatened plant species in Brazil. **Rodriguésia**, v. 69, n. 4, p. 1539-1546, 2018.

RICKETTS, T. H.; SOARES-FILHO, B.; FONSECA, G. A. B.; NEPSTAD, D.; PFAFF, A.; PETSONK, A.; ANDERSON, A.; BOUCHER, D.; CATTANEO, A.; CONTE, M.; CREIGHTON, K.; LINDEN, L.; MARETTI, C.; MOUTINHO, P.; ULLMAN, R.; VICTURINE, R. A protocol to evaluate the potential of protected areas for birdwatching tourism: a study case in the Brazilian Amazon. **PLoS Biology**, v. 8, n. 3, e1000331, 2010.

REGINATO, M. monographaR: An R package to facilitate the production of plant taxonomic monographs. **Brittonia**, v. 68, n. 2, p. 212-216, 2016.

- REIS, F. C.; VICTÓRIA FILHO, R.; ANDRADE, M. T.; BARROSO, A. A. M. Use of Herbicides in Sugarcane in the São Paulo State. **Planta Daninha**, v. 37, e019184227, 2019.
- RICCI, M. S. F.; COSTA, J. R.; PINTO, A. N.; SANTOS, V. L. S. Cultivo orgânico de cultivares de café a pleno sol e sombreado. **Pesquisa Agropecuária Brasileira**, v. 41, n. 4, p. 569-575, 2006.
- RICHARD, U.; BYAMUNGU, R. M.; MAGIGE, F.; MAKONDA, F. B. S. Microhabitat, altitude and seasonal influence on the abundance of non-volant small mammals in Mount Rungwe forest nature reserve. **Global Ecology and Conservation**, v. 35, e02069, 2022.
- ROBERGE, J. M.; ALGESTAM, P. Usefulness of the Umbrella Species Concept as a Conservation Tool. **Conservation Biology**, v. 18, n. 1, p. 76-85, 2004.
- ROCHA, C. M.; LASTRE-ACOSTA, A. M.; PARIZI, M. P. S.; TEIXEIRA, A. C. S. C. Environmental photochemical fate of pesticides ametryn and imidacloprid in surface water (Parapanema River, São Paulo, Brazil). **Environmental Science and Pollution Research**, v. 29, p. 42290–42304, 2022.
- RODGERS, M. S. M.; BAVIA, M. E.; FONSECA, E. O. L.; COVA, B. O.; SILVA, M. M. N.; CARNEIRO, D. D. M. T.; CARDIM, L. L.; MALONE, J. B. Ecological niche models for sand fly species and predicted distribution of *Lutzomyia longipalpis* (Diptera: Psychodidae) and visceral leishmaniasis in Bahia state, Brazil. **Environmental Monitoring and Assessment**, v. 191, 331, 2019.
- RODRIGUES, A. S. L.; CAZALIS, V. The multifaceted challenge of evaluating protected area effectiveness. **Nature Communications**, v. 11, 5147, 2020.
- RODRIGUES JUNIOR, U. J.; DZIEDZIC, M. The water footprint of beef cattle in the amazon region, Brazil. **Ciência Rural**, v. 51, n. 8, 2021.
- RODRIGUEZ-REY, M.; CONSUEGRA, S.; BÖRGER, L.; LEANIZ, C. G. Improving Species Distribution Modelling of freshwater invasive species for management applications. **PLoS ONE**, v. 1, n. 6, e0217896, 2019.
- ROGER, F.; GHANAVI, H. R.; DANIELSSON, N.; WAHLBERG, N.; LÖNDAHL, J.; PETTERSSON, L. B.; ANDERSSON, G. K. S.; OLÉN, N. B.; CLOUGH, Y. Airborne environmental DNA metabarcoding for the monitoring of terrestrial insects—A proof of concept from the field. **Environmental DNA**, v. 4, n. 4, p. 790-807, 2022.
- ROL, M. Idealization, abstraction, and the policy relevance of economic theories. **Journal of Economic Methodology**, v. 15, n. 1, p. 69-97, 2008.
- ROLIM, E. C. A. **A promoção imobiliária e sua relação com as áreas verdes em Natal/RN**. 2019. Dissertação (Mestrado em Engenharia Civil) - Centro de Tecnologia, Universidade Federal do Rio Grande do Norte, Natal, 2019.
- RONQUIM, C. C. **Avanço das áreas de cana-de-açúcar e alterações em áreas de agropecuária no interior paulista**. Campinas: Embrapa Territorial, 2018. 48 p.
- ROSENHEIM, J. A.; GRATTON, C. Ecoinformatics (Big Data) for Agricultural Entomology: pitfalls, progress, and promise. **Annual Review of Entomology**, v. 62, p. 399-417, 2017.



ROTH, A. E. **The Shapley value:** Essays in honor of Lloyd S. Shapley. Melbourne, Australia: Press Syndicate of the University of Cambridge, 1988. 338 p.

RUNQUIST, R. D. B.; LAKE, T. A.; MOELLER, D. A. Improving predictions of range expansion for invasive species using joint species distribution models and surrogate co-occurring species. **Journal of Biogeography**, v. 48, n. 7, p. 1693-1705, 2021.

RUNTING, R. K.; PHINN, S.; XIE, Z.; VENTER, O.; WATSON, J. E. M. Opportunities for big data in conservation and sustainability. **Nature Communications**, v. 11, 2003, 2020.

RUSSELL, J. C.; ABRAHÃO, C. R.; SILVA, J. C. R.; DIAS, R. A. Management of cats and rodents on inhabited islands: An overview and case study of Fernando de Noronha, Brazil. **Perspectives in Ecology and Conservation**, v. 16, n. 4, p. 193-200, 2018.

SALES, N.; G.; KAIZER, M. C.; COSCIA, I.; PERKINS, J. C.; HIGHLANDS, A.; BOUBLI, J. P.; MAGNUSSON, W. E.; SILVA, M. N. F.; BENVENUTO, C.; MCDEVITT, A. D. Assessing the potential of environmental DNA metabarcoding for monitoring Neotropical mammals: a case study in the Amazon and Atlantic Forest, Brazil. **Mammal Review**, v. 50, n. 3, p. 221-225, 2020.

SÁNCHEZ, A. C.; BRICEÑO, N. B. R.; BANDOPADHYAY, S.; GHOSH, S.; GUZMAN, C. T.; OLIVA, M.; GUZMAN, B. K.; LÓPEZ, R. S. Biogeographic Distribution of *Cedrela* spp. Genus in Peru Using MaxEnt Modeling: A Conservation and Restoration Approach. **Diversity**, v. 13, n. 6, 261, 2021.

SANI, N. A. A survey on current distribution and habitat suitability of the Great Bustard in West Azerbaijan, Iran. **Journal of Wildlife and Biodiversity**, v.1, n. 2, p. 88–93, 2017.

SANTANA, V. V.; SANTOS, P. R.; BARBOSA, M. V. Contribuições do Plano de Manejo e do Conselho Gestor em Unidades de Conservação. **Meio Ambiente (Brasil)**, v. 2, n. 2, p. 18-29, 2020.

SANTINI, L.; GONZÁLEZ-SUÁREZ, M.; RUSSO, D.; GONZALEZ-VOYER, A.; VON HARDENBERG, A.; ANCILOTTO, L. One strategy does not fit all: determinants of urban adaptation in mammals. **Ecology Letters**, v. 22, n. 2, p. 365-376, 2019.

SANTOS, K. M. P.; GARAVELLO, M. E. P. E. Uma análise agroalimentar: o caso dos agricultores quilombolas da Reserva de Desenvolvimento Sustentável Quilombos Barra do Turvo - SP. **Redes (St. Cruz Sul, Online)**, v. 21, n. 3, p. 196-216, 2016. Disponível em <<https://online.unisc.br/seer/index.php/redes/article/view/7429>>. Acesso em 03 nov. 2022.

SARKAR, R.; BHADRA, A. How do animals navigate the urban jungle? A review of cognition in urban-adapted animals. **Current Opinion in Behavioral Sciences**, v. 46, 202277, 2022.

SCHMIDT, G. M.; LEWISON, R. L.; SWARTS, H. M. Identifying landscape predictors of ocelot road mortality. **Landscape Ecology**, v. 35, p. 1651-1666, 2020.

SCHMITT, S.; POUTEAU, R.; JUSTEAU, D.; BOISSIEU, F.; BIRNBAUM, P. SSDM: An r package to predict distribution of species richness and composition based on stacked species distribution models. **Methods in Ecology and Evolution**, v. 8, n. 12, p. 1795-1803, 2017.

SCHMITT, C. J.; COOK, J. A.; ZAMUDIO, K. R.; EDWARDS, S. V. Museum specimens of terrestrial vertebrates are sensitive indicators of environmental change in the Anthropocene. **Philosophical Transactions of the Royal Society B: Biological Science**, v. 374, n. 1763, 20170387. 2018.

SCHÖLKOPF, B.; PLATT, J. C.; SHAW-TAYLOR, J.; SMOLA, A. J.; WILLIAMSON, R. C. Estimating Support of a High-Dimensional Distribution. **Neural Computation**, v. 13, n. 7, p. 1443-1471, 2001.

SCHUTTLER, S.; RAMCHARAN, S.; BOONE, H.; STONE, S. O'SHEA, B. J.; GAJAPERSAD, K.; KAYS, R. Can mammals thrive near urban areas in the Neotropics? Characterizing the community of a reclaimed tropical forest. **Tropical Ecology**, v. 62, p. 174-185, 2021.

SCRIVEN, S.; WILLIAMNS, S. H.; GHANI, M. A.; AGAMA, A. L.; BENEDICK, S.; BRODIE, J. F.; HAMER, K. C.; MCCLEAN, C. J.; REYNOLDS, G.; HILL, J. K. D. Assessing the effectiveness of protected areas for conserving range-restricted rainforest butterflies in Sabah, Borneo. **Biotropica**, v. 52, n.2, p. 380-391, 2020.

SENAY, S. D. WORNER, S. P.; IKEDA, T. Novel Three-Step Pseudo-Absence Selection Technique for Improved Species Distribution Modelling. **PLoS ONE**, v. 8, n. 8, e71218, 2013.

SEQUEIRA, A. M. M.; BOUCHET, P. J.; YATES, K. L.; MENGERSEN, K.; CALEY, M. J. Transferring biodiversity models for conservation: Opportunities and challenges. **Methods in Ecology and Evolution**, v. 9, n. 5, p. 1250-1264.

SERRA-DIAZ, J. M.; ENQUIST, B. J.; MAITNER, B.; MEROW, C.; SVENNING, J. C. Big data of tree species distributions: how big and how good? **Forest Ecosystems**, v. 4, 30, 2017.

SHAPLEY, L. A value for N-person games. *In*: KUHN, H.W.; TUCKER, A. (eds.). **Contributions to the Theory of Games II**. Princeton, New Jersey, USA: Princeton University Press, 1953. pp. 251-263.

SHEN, X.; LI, S.; MCSHEA, W. J.; WANG, D.; YU, J.; SHI, X.; DONG, W.; MI, X.; MA, K. Effectiveness of management zoning designed for flagship species in protecting sympatric species. **Conservation Biology**, v. 34, n. 1, p. 158-167, 2020.

SHOU, W.; BERGSTROM, C. T.; CHAKRABORTY, A. K.; SKINNER, F. K. Research: Theory, models and biology. **eLife**, v. 4, e07158.

SILVA, F. C.; BERGAMASCO, A. F. Levantamento de modelos matemáticos descritos para a cultura da cana-de-açúcar. **Revista Biociências**, v. 7, n. 1, p. 7-14, 2001.

SILVA, A. R.; MELLO, J. S. A. Viabilidade de políticas públicas no Sistema Nacional das Unidades de Conservação da Natureza – SNUC (Lei nº 9.985/2000). **Revista Processus de Políticas Públicas e Desenvolvimento Social**, v. 1, n. 2, p. 71-107, 2019.

SILVA, D. F.; SOUZA, V. M. A.; BATISTA, G. E. A. P. A.; KEOGH, E.; ELLIS, D. P. Applying Machine Learning and Audio Analysis Techniques to Insect Recognition in Intelligent Traps. *In*: INTERNATIONAL CONFERENCE ON MACHINE LEARNING

AND APPLICATIONS, 12., 2013, Miami, Florida, USA. **Proceedings** [...]. Piscataway, New Jersey: IEEE Xplore, 2013. p. 99-104.

SILVA, A. R.; FORNECK, E. D.; BORDINGNON, A. L.; CADEMARTORI, C. V. Diet of *Didelphis albiventris* Lund, 1840 (Didelphimorphia, Didelphidae) in two periurban areas in southern Brazil. **Acta Scientiarum**, Biological Sciences, v. 36, n. 2, p. 241-247, 2014.

SILVA, J. S.; SILVA, R. M.; SANTOS, C. A. G. Spatiotemporal impact of land use/land cover changes on urban heat islands: A case study of Paço do Lumiar, Brazil. **Building and Environment**, v. 136, p. 279-292, 2018.a

SILVA, C. O. F.; MANZIONE, R. L.; ALBUQUERQUE FILHO, J. L. Large-Scale Spatial Modeling of Crop Coefficient and Biomass Production in Agroecosystems in Southeast Brazil. **Horticulturae**, v. 4, n. 4, 44, 2018.b

SILVA, D. L.; CORRÊA, P. L. P.; JUAREZ, K. M.; FONSECA, R. L. **Diretrizes para a Integração de Dados de Biodiversidade**. Brasília: MMA, 2015. 100 p.

SILVA, M.; CHAME, M.; MORATELLI, R. Fiocruz Biological Collections: strengthening Brazil's biodiversity knowledge and scientific applications opportunities. **Biodiversity Data Journal**, v. 8, e53607, 2020.

SILVA, I. A.; DUPAS, F. A.; COSTA, C. W.; MEDEIROS, G. O. R.; SOUZA, A. R. Spatiotemporal changes in land cover land use in protected areas of Serra da Mantiqueira, Southeastern Brazil. **Environmental Challenges**, v. 4, 100195, 2021.a

SILVA, G. O.; OLIVEIRA, F. H. ; MELLO, R. Os conflitos socioambientais e sua relação com a reclassificação e recategorização de unidades de conservação. **Brazilian Journal of Development**, v.7, n.7, p. 67538-67558, 2021.b

SILVA, L. O.; RESENDE, M.; GALHARDAS, H.; MANQUINHO, V.; LYNCEM I. DeepData: Machine learning in the marine ecosystems. **Expert Systems with Applications**, v. 206, 117841, 2022.

SILVEIRA, L.; BOULHOSA, R.; ASTETE, S.; JÁCOMO, A. T. A. Management of domestic livestock predation by jaguars in Brazil. **CAT News, Special Issue**, v. 4, p. 21-25, 2008.

SIMÕES, S. S.; ZAPPI, D.; COSTA, G. M.; OLIVEIRA, G.; AONA, L. Y. S. Spatial niche modelling of five endemic cacti from the Brazilian Caatinga: Past, present and future. **Austral Ecology**, v. 45, n. 1, p. 35-47, 2020.

SIMONETTI, J. A.; GREZ, A. A.; ESTADES, C. F. Providing Habitat for Native Mammals through Understory Enhancement in Forestry Plantations. **Conservation Biology**, v. 27, n. 5, p. 1117-1121, 2013.

SINGHAL, R. K.; CHAUHAN, J.; JATAV, H. S.; RAJPUT, V. D.; SINGH, G. S.; BOSE, B. Artificial night light alters ecosystem services provided by biotic components. **Biologia Futura**, v. 72, n. 2, p. 169-185, 2021.

SMALL, C.; SOUSA, D. The role of edge contrast and forest structure in edge influence: vegetation and microclimate at edges in the Brazilian cerrado. **Anthropocene**, v. 14, p. 1-33, 2016.

SMITH, F. A.; SMITH, R. E. E.; LYONS, S. K.; PAYNE, J. L. Body size downgrading of mammals over the late Quaternary. **Science**, v. 360, n. 6386, p. 310-313, 2018.

SMITH, I. T.; KNETTER, S. J.; SVANCARA, L. K.; KARL, J. W.; JOHNSON, T. R.; RACHLOW, J. L. Overlap Between Sagebrush Habitat Specialists Differs Among Seasons: Implications for Umbrella Species Conservation. **Rangeland Ecology & Management**, v. 78, p. 142-154, 2021.

SOARES-FILHO, B.; MOUTINHO, P.; NEPSTAD, D.; MARETTI, C. Role of Brazilian Amazon protected areas in climate change mitigation. **Proceedings of the National Academy of Sciences**, v. 107, n. 24, p. 10821-10286, 2010.

SÓBERON, J.; NAKAMURA, M. Niches and distributional areas: Concepts, methods, and assumptions. **Proceedings of the National Academy of Sciences**, v. 106, n. 2, 19644-50, 2009.

SOBRAL-SOUZA, T.; VANCINE, M. H.; RIBEIRO, M. C.; LIMA-RIBEIRO, M. Efficiency of protected areas in Amazon and Atlantic Forest conservation: A spatio-temporal view. **Acta Oecologica**, v. 87, p. 1-7, 2018.

SOHL, T. L.; CLAGGETT, P. R. Clarity versus complexity: Land-use modeling as a practical tool for decision-makers. **Journal of Environmental Management**, v. 129, p. 235-243, 2013.

SOLOMON, N. G.; KEANE, B. Rodent sociality: a comparison between caviomorphs and other rodent model systems. *In*: EBENSPERGER, L.A.; HAYES, L.D. (eds.). **Sociobiology of Caviomorph Rodents: An Integrative Approach**. Hoboken, New Jersey, USA: John Wiley & Sons, 2016. p. 201-227.

SOUZA, A. M.; JACOBI, L.; PEREIRA, J. E. **Gráficos de controle de regressão usando o Statistica**. Florianópolis: Visual Books, 2005. 112 p.

SOUZA, D. O.; ALVALÁ, R. C. S.; NASCIMENTO, M. G. Urbanization effects on the microclimate of Manaus: A modeling study. **Atmospheric Research**, v. 167, p. 237-248, 2016.

SOUZA, A. C.; PREVEDELLO, J. A. Geographic distribution of the threatened palm *Euterpe edulis* Mart. in the Atlantic Forest: implications for conservation. **Oecologia Australis**, v. 23, n. 3, p. 636-643, 2019.

SOUZA JR, C. M.; SHIMBO, J. Z.; ROSA, M. R.; PARENTE, L. L.; ALENCAR, A. A.; RUDORFF, B. F. T.; HASENACK, H.; MATSUMOTO, M.; FERREIRA, L. G.; SOUZA-FILHO, P. W. M.; OLIVEIRA, S. W.; ROCHA, W. F.; FONSECA, A. V.; MARQUES, C. B.; DINIZ, C. G.; COSTA, D.; MONTEIRO, D.; ROSA, E. R.; VÉLEZ-MARTIN, E.; WEBER, E. J.; LENTI, F. E. B.; PATERNOST, F. F.; PAREYN, F. G. C.; SIQUEIRA, J. V.; VIERA, J. L.; FERREIRA NETO, L. C.; SARAIVA, M. M.; SALES, M. H.; SALGADO, M. P. G.; VASCONCELOS, R.; GALANO, S.; MESQUITA, V. V.; AZEVEDO, T. Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine. **Remote Sensing**, v. 12, n. 17, 2735, 2020.

SPAAN, D.; DI FIORE, A.; RANGEL-RIVERA, C. E.; HUTSCHENREITER, A.; WICH, S.; AURELI, F. Detecting spider monkeys from the sky using a high-definition RGB camera: a rapid-assessment survey method? **Biodiversity and Conservation**, v. 31, p. 479–496, 2022.

SPERA, S. A.; COHN, A. S.; VANWEY, L. K.; MUSTARD, J. F.; RUDORFF, B. F.; RISSO, J.; ADAMI, M. Recent cropping frequency, expansion, and abandonment in Mato Grosso, Brazil had selective land characteristics. **Environmental Research Letters**, v. 9, n. 6, 064010, 2014.

SPIERENBURG, M. Getting the message across biodiversity science and policy interfaces – a review. **GAIA – Ecological Perspectives for Science and Society**, v. 21, n. 2, p. 125-134, 2012.

SRIVASTAVA, V.; LAFOND, V.; GRIESS, V. C. Species distribution models (SDM): applications, benefits and challenges in invasive species management. **CABI Reviews**, v. 2019, p. 1-13, 2019.

STARLINGS, J. R. The importance of understorey on wildlife in a Brazilian eucalypt plantation. **Revista Brasileira de Zoologia**, v. 7, n. 3, p. 267-276, 1990.

STEEN, V. A.; ELPHICK, C. S.; TINGLEY, M. W. An evaluation of stringent filtering to improve species distribution models from citizen science data. **Biodiversity Research**, v. 25, n. 12, p. 1857-1869, 2019.

STOCKWELL, D. R.; PETERSON, A. T. Effects of sample size on accuracy of species distribution models. **Ecological Modelling**, v. 148, n. 1, p. 1-13, 2002.

STOWELL D. Computational bioacoustics with deep learning: a review and roadmap. **PeerJ** 10:e13152, 2022.

STRAND, J.; SOARES-FILHO, B.; COSTA, M. H.; OLIVEIRA, U.; RIBEIRO, S. C.; PIRES, G. F.; OLIVEIRA, A.; RAJÃO, R.; MAY, P.; VAN DER HOFF, R.; SIIKAMÄKI, J.; MOTTA, R. S.; TOMAN, M. Spatially explicit valuation of the Brazilian Amazon Forest's Ecosystem Services. **Nature Sustainability**, v. 1, p. 657-664, 2018.

SU, H.; BISTA, M.; LI, M. Mapping habitat suitability for Asiatic black bear and red panda in Makalu Barun National Park of Nepal from Maxent and GARP models. **Scientific Reports**, v. 11, 14136, 2021.

SUMBH, O.; HOF, A. R. Can pikas hold the umbrella? Understanding the current and future umbrella potential of keystone species Pika (*Ochotona* spp.). **Global Ecology and Conservation**, v. 38, e02247, 2022.

SUSSMAN, R. W.; GARBER, P. A.; CHEVERUD, J. M. Importance of Cooperation and Affiliation in the Evolution of Primate Sociality. **American Journal of Physical Anthropology**, v. 128, n. 1, p. 84-97, 2005.

TAGLIARI, M. M.; VIEILLEDENT, G.; ALVES, J.; SILVEIRA, T. C. L. PERONI, N. Relict populations of *Araucaria angustifolia* will be isolated, poorly protected, and unconnected under climate and land-use change in Brazil. **Biodiversity and Conservation**, v. 30, p. 3665–3684, 2021.

TARABOM, S.; BERGÈS, L.; DUTOIT, T.; ISSELIN-NONDEDEU, F. Environmental impact assessment of development projects improved by merging species distribution and habitat connectivity modelling. **Journal of Environmental Management**, v. 241, p. 439-449, 2019.

TAVARES, P. D.; UZÊDA, M. C.; PIRES, A. S. Biodiversity Conservation in Agricultural Landscapes: the Importance of the Matrix. **Floresta e Ambiente**, v. 26, n. 4, e20170664, 2019.

TAYLOR, L. Monkeypox: WHO to rename disease to prevent stigma. **BMJ**, v. 377, o1489, 2022.

TEIXEIRA, F. G. **Os investimentos públicos em áreas verdes e a produção do espaço urbano em Campos dos Goytacazes (RJ)**. 2019. Dissertação (Mestrado em Desenvolvimento Regional, Ambiente e Políticas Públicas) - Instituto de Ciências da Sociedade e Desenvolvimento Regional, Universidade Federal Fluminense, Campos dos Goytacazes, 2019. Disponível em <<https://app.uff.br/riuff/handle/1/13366>>. Acesso em 23 out. 2022.

TERBORGH, J.; VAN SCHAİK, C. P. Minimizing Species Loss: The Imperative of Protection. *In*: KRAMER, R.; VAN SCHAİK, C.; JOHNSON, J. (eds.), **Protected Areas and the Defense of Tropical Biodiversity**, New York USA: Oxford University Press, 1997. p.15-35.

THONIS, A.; CEBALLOS, R. M.; TUEN, A. A.; LOVEGROVE, B. G.; LEVESQUE, D. L. Small tropical mammals can take the heat: high upper limits of thermoneutrality in a Bornean treeshrew. **Physiological and Biochemical Zoology**, v. 93, n. 3, p. 199-209, 2020.

TILMAN, D.; ISABELL, F.; COWLES, J. M. Biodiversity and Ecosystem Functioning. **Annual Review of Ecology, Evolution and Systematics**, v. 45, n. 1, p. 471-493, 2014.

TIMO, T. P. C.; LYRA-JORGE, M. C.; GHELIER-COSTA, C. Effect of the plantation age on the use of Eucalyptus stands by medium to large-sized wild mammals in south-eastern Brazil. **iForest - Biogeosciences and Forestry**, v. 8, n. 2, p. 108-113. 2015, 2014.

TOILE, K.; TANSLEY, S.; HEY, T. Jim Gray on eScience: A Transformed Scientific Method. *In*: HEY, T.; TANSLEY, S.; TOILE, K. (eds.). **The fourth paradigm: data-intensive scientific discovery**. Redmond, Washington, USA: Microsoft Research, 2009. p. xvii-xxxi.

TORRES, P. H. C.; RAMOS, R. F.; GONÇALVES, L. R. Conflitos ambientais na macrometrópole paulista: Paranapiacaba e São Sebastião. **Ambiente & Sociedade**, v. 22, e0101, 2019.

TORRES-ROMERO, E. J.; GIORDANO A. J. Impact of the Anthropocene on the status of the world's small carnivores: A global macroecological perspective. **Journal of Biogeography**, v. 49, n. 5, p. 916-929, 2022.

TOURINHO, L.; PREVEDELLO, J. A.; CARVALHO, B. M. ROCHA, D. S. B.; VALE, M. M. Macroscale climate change predictions have little influence on landscape-scale habitat suitability. **Perspectives in Ecology and Conservation**, v. 20, n. 1, p. 29-37, 2022.

TUCKER, M. A.; BÖHNING-GAESE, K.; FAGAN, W. F.; FRYXELL, J. M.; VAN MOORTER, B.; ALBERTS, S. C.; ALI, A. H.; ALLEN, A. M.; ATTIAS, N.; AVGAR, T.; BARTLAM-BROOKS, H.; BAYARBAATAR, B.; BELANT, J. L.; BERTASSONI, A.; BEYER, D.; BIDNER, L.; VAN BEEST, F. M.; BLAKE, S.; BLAUM, N.; BRACIS, C.; BROWN, D.; DE BRUYN, P. J. N.; CAGNACCI, F.; CALABRESE, J. M.; CAMILO-ALVES, C.; CHAMAILLÉ-JAMMES, S.; CHIARADIA, A.; DAVIDSON, S. C.; DENNIS, T.; DESTEFANO, S.; DIEFENBACH, D.; DOUGLAS-HAMILTON, I.; FENNESSY, J.; FICHTEL, C.; FIEDLER, W.; FISCHER, C.; FISCHHOFF, I.; FLEMING, C. H.; FORD, A. T.; FRITZ, S. A.; GEHR, B.; GOHEEN, J. R.; GURARIE, E.; HEBBLEWHITE, M.; HEURICH, M.; HEWISON, A. J. M.; HOF, C.; HURME, E.; ISBELL, L. A.; JANSSEN, R.; JELTSCH, F.; KACZENSKY, P.; KANE, A.; KAPPELER, P. M.; KAUFFMAN, M.; KAYS, R.; KIMUYU, D.; KOCH, F.; KRANSTAUBER, B.; LAPOINT, S.; LEIMGRUBER, P.; LINNELL, J. D. C.; LÓPEZ-LÓPEZ, P.; MARKHAM, A. C.; MATTISSON, J.; MEDICI, E. P.; MELLONE, U.; MERRILL, E.; DE MIRANDA MOURÃO, G.; MORATO, R. G.; MORELLET, N.; MORRISON, T. A.; DÍAZ-MUÑOZ, S. L.; MYSTERUD, A.; NANDINTSETSEG, D.; NATHAN, R.; NIAMIR, A.; ODDEN, J.; O'HARA, R. B.; OLIVEIRA-SANTOS, L. G. R.; OLSON, K. A.; PATTERSON, B. D.; CUNHA DE PAULA, R.; PEDROTTI, L.; REINEKING, B.; RIMMLER, M.; ROGERS, T. L.; ROLANDSEN, C. M.; ROSENBERRY, C. S.; RUBENSTEIN, D. I.; SAFI, K.; SAÏD, S.; SAPIR, N.; SAWYER, H.; SCHMIDT, N. M.; SELVA, N.; SERGIEL, A.; SHIILEGDAMBA, E.; SILVA, J. P.; SINGH, N.; SOLBERG, E. J.; SPIEGEL, O.; STRAND, O.; SUNDARESAN, S.; ULLMANN, W.; VOIGT, U.; WALL, J.; WATTLES, D.; WIKELSKI, M.; WILMERS, C. C.; WILSON, J. W.; WITTEMYER, G.; ZIĘBA, F.; ZWIJACZ-KOZICA, T.; MUELLER, T. Moving in the Anthropocene: Global reductions in terrestrial mammalian movements. *Science*, v. 359, p. 466–469, 2018.

TURNER, W. Sensing biodiversity. *Science*, v. 346, n. 6207, p. 301-302, 2014.

TWISS, S. D.; BRANNAN, N.; SHUERT, C. R.; BISHOP, A. M.; POMEROY, P. P.; MOSS, S. An external telemetry system for recording resting heart rate variability and heart rate in free-ranging large wild mammals. *PLoS ONE*, v. 16, n. 6, e0252013, 2021.

UDY, K.; FRITSCH, M.; MEYER, K. M.; GRASS, I.; HANSS, S.; KNEIB, T.; KREFT, H.; KUKUNDA, C. B.; PE'ER, G.; REININGHAUS, H.; TIETEJEN, B.; TSCHARNTKE, T.; VAN WAVEREN, C. S.; WIEGAND, K. Environmental heterogeneity predicts global species richness patterns better than area. *Global Ecology and Biogeography*, v. 30, n. 4, p. 842-851, 2021.

URBINA-CARDONA, J. N.; LOYOLA, R. D. Applying niche-based models to predict endangered-hylid potential distributions: are neotropical protected areas effective enough? *Tropical Conservation Science*, v.1, n. 4, p. 417-445, 2008.

VAN EE, J. J.; IVAN, J. S.; HOOTEN, M. B. Community confounding in joint species distribution models. *Scientific Reports*, v. 12, 12235, 2022.

VAN ZYL, J. J. The Shuttle Radar Topography Mission (SRTM): a breakthrough in remote sensing of topography. *Acta Astronautica*, v. 48, n.5-12, p. 559-565, 2001.

VAPNIK, V. 1995. *The nature of statistical learning theory*. New York, USA: Springer Verlag, 1995. 188 p.

- VARELA, S.; ANDERSON, R. P.; GARCÍA-VALDÉS, R.; FERNÁNDEZ-GONZÁLEZ, F. Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. **Ecography**, v. 37, n. 11, p. 1084-1091, 2014.
- VASUDEVA, V.; UPGUPTA, S.; SINGH, A.; SHERWANI, N.; DUTTA, S.; RAJAMARAN, R.; CHAUDHURI, S.; VERMA, S.; JOHNSON, J. A.; KRISHNAMURTHY, R. Conservation Prioritization in a Tiger Landscape: Is Umbrella Species Enough? **Land**, v. 11, n. 3, 371, 2022.
- VERMEIREN, P.; REICHERT, P.; SCHUWIRTH, N. Integrating uncertain prior knowledge regarding ecological preferences into multi-species distribution models: Effects of model complexity on predictive performance. **Ecological Modelling**, v. 420, 108956, 2020.
- VERÍSSIMO, I.; CUPOLILLO, G.; JORGE, B. M. S.; NOVAES, R. L. M.; TAVARES, J. A.; GABRIEL, M. M.; COSTA-NETO, S. F.; COUTO, A. L. G.; SCHMIDT, E.; MIRANDA, A.; ANDREAZZI, C. S.; MORATELLI, R. Medium- and large-sized mammals from Estação Biológica Fiocruz Mata Atlântica, Rio de Janeiro, south-eastern Brazil. **Biodiversity Data Journal**, v. 10, e86756, 2022.
- VETTER, D.; HANSBAUER, M. M.; VÉGVÁRI, Z.; STORCH, I. Predictors of forest fragmentation sensitivity in Neotropical vertebrates: a quantitative review. **Ecography**, v. 34, p. 1-8, 2011.
- VIANNA, M. S.; NASSIF, D. S. P.; CARVALHO, K. S.; MARIN, R. R. Modelling the trash blanket effect on sugarcane growth and water use. **Computers and Electronics in Agriculture**, v. 172, 105361, 2020.
- VICTOR, R. A. B. M.; COSTA NETO, J. B.; AB'SABER, A. N.; SERRANO, O.; DOMINGOS, M.; PIRES, B. C. C.; AMAZONAS, M.; VICTOR, M. A. Application of the Biosphere Reserve Concept to Urban Areas: The Case of São Paulo City Green Belt Biosphere Reserve, Brazil—São Paulo Forest Institute: A Case Study for UNESCO. **Annals of the New York Academy of Sciences**, v. 1023, n. 1, p 237-281, 2004.
- VIJAY, V.; ARMSWORTH, P. R. Pervasive cropland in protected areas highlight trade-offs between conservation and food security. **Proceedings of the National Academy of Sciences**, v. 118, n. 4, e2010121118, 2021.
- VILAS BOAS, A. H.; VERÍSSIMO, I.; NOVAES, R. L. M.; CUPOLILLO, G.; ANDREAZZI, C. S.; COSTA-NETO, S. F.; MORATELLI, R. Survey of medium- and large-sized mammals in Atlantic Forest remnants of Conceição dos Ouros, Minas Gerais, Brazil. **Biodiversity Data Journal**, v. 10:e82139, 2022.
- VILELA, G. F.; BENTES, M. P. M.; OLIVEIRA, Y. M. M.; MARQUES, D. K. S.; SILVA, J. C. B. **Life on land: Contributions of Embrapa**. Brasília: Embrapa, 2019. 15 p. Disponível em <<https://www.alice.cnptia.embrapa.br/alice/bitstream/doc/1113495/1/Alienspecieseconomicaluse2019.pdf>>. Acesso em 22 out. 2022.
- VILLA-MACHÍO, I.; CASTRO, A. G. F.; FUERTES-AGUILAR, J.; FELINER, G. N. Colonization history of the Canary Islands endemic *Lavatera acerifolia*, (Malvaceae) unveiled with genotyping-by-sequencing data and niche modelling. **Journal of Biogeography**, v. 47, n. 4, p. 993-1005, 2020.



VILLAR, N.; PAZ, C.; ZIPPARO, V.; NAZARETH, S.; BULASCOSCHI, L.; BAKKER, E.S.; GALETTI, M. Frugivory underpins the nitrogen cycle. **Functional Ecology**, v. 35, n. 2, p. 357-368, 2021.

VILLAR, N.; ROCHA-MENDES, F.; GUEVARA, R.; GALETTI, M. Large herbivore-palm interactions modulate the spatial structure of seedling communities and productivity in Neotropical forests. **Perspectives in Ecology and Conservation**, v. 20, n. 1, p. 45-59, 2022.

VILLERO, D.; PLA, M.; CAMPS, D.; RUIZ-OLMO, J.; BROTONS, L. Integrating species distribution modelling into decision-making to inform conservation actions. **Biodiversity and Conservation**, v. 26, p. 251-271, 2017.

VINCENT, H.; WIERSEMA, J.; KELL, S.; FIELDER, H.; DOBBIE, S.; CASTAÑEDA-ÁLVAREZ, N. P.; GUARINO, L.; EASTWOOD, R.; LEÓN, B.; MAXTED, N. A prioritized crop wild relative inventory to help underpin global food security. **Biological Conservation**, v. 167, p. 265-275, 2013.

VINCENT, H.; AMRI, A.; CASTAÑEDA-ÁLVAREZ, N. P.; DEMPEWOLF, H.; DULLOO, E.; GUARINO, L.; HOLE, D.; MBA, C.; TOLEDO, A.; MAXTED, N. Modeling of crop wild relative species identifies areas globally for in situ conservation. **Communications Biology**, v. 2, 136, 2019.

VIVO, M.; CARMIGNOTTO, A. P.; GREGORIN, R.; HINGST-ZAHER, E.; IACK-XIMENES, G. E.; MIRETZKI, M.; PERCEQUILLO, A.R.; ROLLO JUNIOR, M. M.; ROSSI, R. V.; TADDEI, V. A. Checklist dos mamíferos do Estado de São Paulo, Brasil. **Biota Neotropica**, v. 11, n. 1, 2011.

VOLTOLINI, T. C. **Desenvolvimento Local no município de Altinópolis - SP com destaque para o turismo**: diagnóstico de uma proposta. 2011. Trabalho de Conclusão de Curso (Bacharelado em Geografia) - Instituto de Geociências, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Rio Claro, 2011.

WANG, L., REN, Y. Python Teaching Research and Practice from the Perspective of Computational Ecology. *In*: KOUNTCHEV, R., MAHANTI, A., CHONG, S., PATNAIK, S., FAVORSKAYA, M. (eds.). **Advances in Wireless Communications and Applications**. Smart Innovation, Systems and Technologies, v. 191. Singapore: Springer, 2021. p. 177-182.

WANG, Y.; LIU, Y.; JING, W. Hadoop-based Parallel Algorithm for Data Mining in Remote Sensing Images. **International Journal of Performability Engineering**, v. 15, n. 11, p. 2860-2870, 2019.

WEERSINK, A; FRASERM E.; PANNEL, D.; DUNCANM E.; ROTZ, S. Opportunities and Challenges for Big Data in Agricultural and Environmental Analysis. **Annual Review of Resource Economics**, v. 10, p. 19-37, 2018.

WEGE, M.; SALAS, L.; LARUE, M. Citizen science and habitat modelling facilitates conservation planning for crabeater seals in the Weddell Sea. **Biodiversity Research**, v. 26, n. 10, p. 1291-1304, 2020.

WEIJS, S. V.; RUDELLE, B. L. Debates: Does Information Theory Provide a New Paradigm for Earth Science? Sharper Predictions Using Occam's Digital Razor. **Water Resources Research**, v. 56, n. 2, e2019WR026471, 2020.

- WEISS, D. J.; ATKINSON, P. M.; BHATT, S.; MAPPIN, B.; HAY, S. I.; GETHING, P. W. An effective approach for gap-filling continental scale remotely sensed time-series. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 98, p. 106-118, 2014.
- WELCH, B. L. The generalization of "Student's" problem when several different population variances are involved. **Biometrika**, v. 34, n. 1-2, p. 28-35, 1947.
- WENCESLAU, F. F.; BARDEN, J. E.; TURATTI, L. O Brasil e as Metas de Aichi: uma análise sobre o cumprimento da Meta 11. **Revista Internacional de Direito Ambiental**, v. 9, n. 25, p. 113-132, 2020.
- WICH, S. A.; KOH, L. P. **Conservation Drones: Mapping and Monitoring Biodiversity**. New York, USA: Oxford University Press, 2018. 144 p.
- WIECZOREK, J.; BLOOM, D.; GURALNICK, R.; BLUM, S.; DÖRING, M.; GIOVANNI, R.; ROBERTSON, T.; VIEGLAIS, D. Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. **PLoS ONE**, v. 7, n. 1, e29715, 2012.
- WILLCOCK, S.; MARTÍNEZ-LÓPEZ, J.; HOOFTMAN, D. A. P.; BAGSTAD, K. J.; BALBI, S.; MARZO, A.; PRATO, C.; SCIANDRELLO, S.; SIGNORELLO, G.; VOIGHT, B.; VILLA, F.; BULLOCK, J. M.; ATHANASIADIS I. N. Machine learning for ecosystem services. **Ecosystem Services**, v. 33, p. 165-174, 2018.
- WILSON, J. W.; SEXTON, J. O.; JOBE, R. T.; HADDAD, N. M. The relative contribution of terrain, land cover, and vegetation structure indices to species distribution models. **Biological Conservation**, v. 164, p. 170-176, 2013.
- WOOD, J. R.; HOLDAWAY, R. J.; ORWIN, K. H.; MORSE, C.; BONNER, K. I.; DAVIS, C.; BOLSTRIDGE, N.; DICKIE, I. A. No single driver of biodiversity: divergent responses of multiple taxa across land use type. **Ecosphere**, v. 9, n. 11, e01997, 2017.
- WU, J.; GUO, S.; LI, J.; ZENG, D. Big Data Meet Green Challenges: Big Data Toward Green Applications. **IEEE Systems Journal**, v. 10, n. 3, p. 888-900, 2016.
- YANG, C.; YU, M.; HU, M.; JIANG, Y.; LI, Y. Utilizing Cloud Computing to address big geospatial data challenges. **Computers, Environment and Urban Systems**, v. 61, p. 120-128, 2017.
- YANNIC, G.; PELLISSIER, L.; LE CORRE, M.; DUSSAULT, C.; BERNATCHEZ, L.; CÔTE, S. D. Temporally dynamic habitat suitability predicts genetic relatedness among caribou. **Proceedings of the Royal Society B**, v. 281, n. 1792, 20140502, 2014.
- YI, H.; MA, H. Monitoring and control system for logistics protection and transportation of urban solid waste based on big data. *Global NEST Journal*, v. 22, n. 4, p. 455-462, 2020.
- YING, X. An overview of overfitting and its solutions. **Journal of Physics: Conference Series**, v. 1168, 022022, 2019.
- YOUNG, A. F. Urban expansion and environmental risk in the São Paulo Metropolitan Area. **Climate Research**, v. 57, p. 73-80, 2013.

YOUNG, P.; PARKINSON, S.; LEES, M. Simplicity out of complexity in environmental modelling: Occam's razor revisited. **Journal of Applied Statistics**, v. 23, n. 2-3, p. 165-210, 1996.

YOUNG, C. E. F.; CASTRO, B. S. Financing mechanisms to bridge the resource gap to conserve biodiversity and ecosystem services in Brazil. **Ecosystem Services**, v. 50, 101321, 2021.

YU, J.; ZHANG, Z.; SARWAT, M. Spatial data management in apache spark: the GeoSpark perspective and beyond. **GeoInformatica**, v. 23, p. 37-78, 2019.

YUE, P.; JIANG, L. BigGIS: How big data can shape next-generation GIS. *In*: INTERNATIONAL CONFERENCE ON AGRO-GEOINFORMATICS, 3., 2014, Beijing, China. **Proceedings** [...] [s.l.]: IEEE, 2014. p. 1-6. Disponível em <<https://ieeexplore.ieee.org/document/6910649>>. Acesso em 28 jan. 2023.

ZANIEWSKI, A. E.; LEHMANN, A.; OVERTON, J. M. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. **Ecological Modelling**, v. 158, n. 2-3, p. 261-280, 2002.

ZARAGOZI, B. M.; TRILLES, S.; NAVARRO-CARRIÓN, J. T. Leveraging Container Technologies in a GIScience Project: A Perspective from Open Reproducible Research. **ISPRS International Journal of Geo-Information**, v. 9, n. 3, 138, 2020.

ZELLMER, A. J.; CLAISSE, J. T.; WILLIAMS, C. M.; SCHWAB, S.; PONDELLA II, D. J. Predicting Optimal Sites for Ecosystem Restoration Using Stacked-Species Distribution Modeling. **Frontiers in Marine Science**, v. 6, 3, 2019.

ZARO, G. C.; CARAMORI, P. H.; YADA JUNIOR, G. M.; SANQUETTA, C. R.; ANDROCIOLI FILHO, A.; NUNES, A. L. P.; PRETE, C. E. C.; VORONEY, P. Carbon sequestration in an agroforestry system of coffee with rubber trees compared to open-grown coffee in southern Brazil. **Agroforestry Systems**, v. 94, p. 799-809, 2020.

ZARO, G. C.; CARAMORI, P. H.; WREGGE, M. S.; CALDANA, N. F. S.; VIRGENS FILHO, J. S.; MORAIS, H.; YADA JUNIOR, G. M.; CARAMORI, D. C. Coffee crops adaptation to climate change in agroforestry systems with rubber trees in southern Brazil. *Scientia Agricola*, v. 8, e20210142, 2023.

ZHANG, J. Biodiversity science and macroecology in the era of big data. **Biodiversity Science**, v. 25, n. 4, p. 355-363, 2017.

ZHANG, J.; LI, S. A Review of Machine Learning Based Species' Distribution Modelling. *In*: INTERNATIONAL CONFERENCE ON INDUSTRIAL INFORMATICS - COMPUTING TECHNOLOGY, INTELLIGENT TECHNOLOGY, INDUSTRIAL INFORMATION INTEGRATION, 3., 2017, Wuhan, China. **Proceedings** [...] [s.l.]: IEEE, 2017. p. 199-206.

ZHANG, C.; LI, X. Land Use and Land Cover Mapping in the Era of Big Data. **Land**, v. 11, 1692, 2022.

ZIZKA, A.; CARVALHO, F. A.; CALVENTE, A.; BAEZ-LIZARAZO, M. R.; CABRAL, A.; COELHO, J. F. R.; COLLI-SILVA, M.; FANTINATI, M. R.; FERNANDES, M. F.;

FERREIRA-ARAÚJO, T.; MOREIRA, F. G. L.; SANTOS, N. M. C.; SANTOS, T. A. B.; SANTOS-COSTA, R. C.; SERRANO, F. C.; SILVA, A. P. A.; SOARES, A. S.; SOUZA, P. G. C.; TOMAZ, E. C.; VALE, V. F.; VIEIRA, T. L.; ANTONELLI, A. No one-size-fits-all solution to clean GBIF. **PeerJ**, v. 8, e9916, 2020.

ZWEIG, M. H.; CAMPBELL, G. Receiver-Operating Characteristic (ROC) Plots: a fundamental evaluation tool in Clinical Medicine. **Clinical Chemistry**, v. 39, n. 4, p. 561-577, 1993.

### ANEXO I: Bibliotecas Python usadas nesta pesquisa

Biblioteca e Módulos	Função	Fonte
csv	Implementa classes para ler e gravar dados tabulares no formato CSV.	<a href="https://docs.python.org/3/library/csv.html">https://docs.python.org/3/library/csv.html</a>
dtreeviz	Visualização de árvores de decisão e interpretação de modelos.	<a href="https://pypi.org/project/dtreeviz/">https://pypi.org/project/dtreeviz/</a>
earthengine-api	Interação com o Google Earth Engine usando a linguagem de programação Python.	<a href="https://pypi.org/project/earthengine-api/">https://pypi.org/project/earthengine-api/</a>
elevation	Acesso e processamento de modelos digitais de elevação de terreno global, SRTM 30m DEM e SRTM 90m DEM.	<a href="https://pypi.org/project/elevation/">https://pypi.org/project/elevation/</a>
esda	<p style="text-align: center;">Moran</p> <hr/> <p style="text-align: center;">Moran_BV</p> <hr/> <p style="text-align: center;">Moran_Local</p>	<p style="text-align: center;">Análise Exploratória de Dados Espaciais. Métodos para testar a autocorrelação global e local em dados unitários de área.</p> <p style="text-align: right;"><a href="https://pypi.org/project/esda/">https://pypi.org/project/esda/</a></p>

## Moran\_Local\_BV

fiona	Leitura e gravação de dados em formatos GIS.	<a href="https://pypi.org/project/Fiona/">https://pypi.org/project/Fiona/</a>
folium	Visualização interativa de dados espaciais	<a href="https://pypi.org/project/folium/">https://pypi.org/project/folium/</a>
geobr	Download de conjuntos de dados espaciais oficiais do Brasil.	<a href="https://pypi.org/project/geobr/">https://pypi.org/project/geobr/</a>
geopandas	Estende os tipos de dados usados pelos pandas para permitir operações espaciais em tipos geométricos.	<a href="https://geopandas.org/">https://geopandas.org/</a>
glob	Encontra todos os nomes de caminho que correspondem a um padrão especificado.	<a href="https://docs.python.org/3/library/glob.html">https://docs.python.org/3/library/glob.html</a>
google.colab drive	Bibliotecas específicas do Python do Google Colaboratory.	<a href="https://pypi.org/project/google-colab/">https://pypi.org/project/google-colab/</a>
io StringIO	Lida com vários tipos de I/O.	<a href="https://docs.python.org/3/library/io.html">https://docs.python.org/3/library/io.html</a>

<p>IPython.display</p> <p>Image</p>	<p>API pública para ferramentas de exibição em IPython.</p> <p><a href="https://ipython.readthedocs.io/en/stable/api/generated/IPython.display.html">https://ipython.readthedocs.io/en/stable/api/generated/IPython.display.html</a></p>
<p>kneed</p> <p>KneeLocator</p>	<p>Retorna o ponto de curvatura máxima de uma função.</p> <p><a href="https://pypi.org/project/kneed/">https://pypi.org/project/kneed/</a></p>
<p>libpysal</p> <hr/> <p>libpysal.weights</p> <p>Kernel</p> <hr/> <p>libpysal.weights.contiguity</p> <p>Queen</p>	<p>Oferece quatro módulos que formam os blocos de construção em muitos pacotes upstream na família PySAL: 1) Pesos Espaciais: libpysal.weights; 2) Entrada e saída: libpysal.io; 3) Geometria computacional: libpysal.cg; 4) Conjuntos de dados de exemplo integrados libpysal.examples.</p> <p><a href="https://pysal.org/libpysal/">https://pysal.org/libpysal/</a></p>
<p>matplotlib</p> <hr/> <p>cm</p> <p>get_cmap</p> <hr/> <p>matplotlib.pyplot</p> <hr/> <p>matplotlib.ticker</p>	<p>Cria visualizações estáticas, animadas e interativas em Python.</p> <p><a href="https://matplotlib.org/">https://matplotlib.org/</a></p>

numpy	Biblioteca Python usada para trabalhar com arrays, com funções para trabalhar no domínio da álgebra linear, transformada de Fourier e matrizes.	<a href="https://numpy.org/">https://numpy.org/</a>
openpyxl	Leitura e gravação de arquivos do Excel 2010 xls/xlsm/xlsx/xltx/xltn.	<a href="https://openpyxl.readthedocs.io/en/stable/">https://openpyxl.readthedocs.io/en/stable/</a>
osgeo gdal	Importação / exportação, processamento e visualização de dados geoespaciais.	<a href="https://www.osgeo.org/">https://www.osgeo.org/</a>
gdal_array		
pandas	Processamento e análise de dados estruturados.	<a href="https://pandas.pydata.org/">https://pandas.pydata.org/</a>
plotly.express	API de alto nível para criação de figuras.	<a href="https://plotly.com/python/plotly-express/">https://plotly.com/python/plotly-express/</a>
pydotplus	Fornece uma interface Python para a linguagem Dot do Graphviz.	<a href="https://pydotplus.readthedocs.io/">https://pydotplus.readthedocs.io/</a>
pygeos	Biblioteca C/Python com funções geométricas vetorizadas.	<a href="https://pygeos.readthedocs.io/en/stable/">https://pygeos.readthedocs.io/en/stable/</a>



	mask		
	features		
	rasterio.features	shapes	
rasterio	rasterio.merge		Processa dados raster geoespaciais baseados em geometrias vetoriais. Inclui funções para estatísticas zonais e consultas de pontos interpolados.
	rasterio.plot		
	rasterio.warp		
		point_query	
	rasterstats	zonal_stats	
	requests		Permite o envio de solicitações HTTP/1.1.
			<a href="https://pypi.org/project/requests/">https://pypi.org/project/requests/</a>
	richdem		Conjunto de ferramentas de análise hidrológica do modelo digital de elevação (DEM).
			<a href="https://richdem.readthedocs.io/en/latest/">https://richdem.readthedocs.io/en/latest/</a>

rioxarray			Extensão xarray geoespacial alimentada pela biblioteca rasterio	<a href="https://pypi.org/project/rioxarray/">https://pypi.org/project/rioxarray/</a>
rpy2.ipython			Interface de comando mágico para trabalho interativo com R em ipython.	<a href="https://rpy2.github.io/doc/latest/html/interactive.html">https://rpy2.github.io/doc/latest/html/interactive.html</a>
rtree			Índice espacial R-Tree para Python GIS.	<a href="https://pypi.org/project/Rtree/">https://pypi.org/project/Rtree/</a>
	stats			
scipy	scipy.spatial	distance_matrix	Fornecer algoritmos para otimização, integração, interpolação, problemas de autovalor, equações algébricas, equações diferenciais, estatísticas e muitas outras classes de problemas. Contém principalmente funções que são realmente funções numpy	<a href="https://scipy.org/">https://scipy.org/</a>
	scipy.stats	mannwhitneyu		
seaborn			Biblioteca de visualização de dados baseada em matplotlib.	<a href="https://seaborn.pydata.org/">https://seaborn.pydata.org/</a>
shap			Uma abordagem unificada para explicar a saída de qualquer modelo de aprendizado de máquina.	<a href="https://shap.readthedocs.io/">https://shap.readthedocs.io/</a>

shapely	shapely.geometry	shape	Processamento e análise de objetos geométricos planares.	<a href="https://pypi.org/project/shapely/">https://pypi.org/project/shapely/</a>
	shapely.speedups			
	shutil		Operações de alto nível em arquivos e coleções de arquivos. Em particular, são fornecidas funções que suportam a cópia e remoção de arquivos.	<a href="https://docs.python.org/3/library/shutil.html">https://docs.python.org/3/library/shutil.html</a>
sklearn	metrics		Biblioteca para Aprendizado de Máquina que fornece recursos para modelagem estatística, análise e mineração de dados, além de suporte ao aprendizado supervisionado e não supervisionado	<a href="https://scikit-learn.org/">https://scikit-learn.org/</a>
	svm			
	tree			
	sklearn.cluster	DBSCAN		
	sklearn.dummy	DummyRegressor		
	sklearn.decomposition	PCA		
	sklearn.ensemble	RandomForestRegressor		

---

sklearn.linear_model	LinearRegression
	accuracy_score
	confusion_matrix
sklearn.metrics	recall_score
	precision_score
	f1_score
sklearn.model_selection	train_test_split
	GridSearchCV
sklearn.neighbors	DistanceMetric
	KNeighborsRegressor
sklearn.preprocessing	LabelEncoder
	MinMaxScaler

---

		OneClassSVM	
	sklearn.svm	SVR	
		DecisionTreeRegressor	
	sklearn.tree	export_graphviz	
	splot.mapping	mapclassify_bin	
splot	vba_choropleth		Análise visual para análise espacial com PySAL. <a href="https://readthedocs.org/projects/splot/downloads/pdf/stable/">https://readthedocs.org/projects/splot/downloads/pdf/stable/</a>
	vba_legend		
	splot.esda	lisa_cluster	
	moran_scatterplot		
		plot_moran_bv	
statsmodels	statsmodels.stats.api		Fornece classes e funções para a estimativa de modelos estatísticos, e também para a realização de testes estatísticos e exploração de dados. <a href="https://www.statsmodels.org/">https://www.statsmodels.org/</a>
	statsmodels.stats.diagnostic		

statsmodels.regression.linear_model		
statsmodels.tools.tools		
time		Fornecer várias funções relacionadas ao tempo. <a href="https://docs.python.org/3/library/time.html">https://docs.python.org/3/library/time.html</a>
U kaleido		Biblioteca multiplataforma para gerar imagens estáticas (por exemplo, png, svg, pdf, etc.) para bibliotecas de visualização baseadas na web, com foco particular na eliminação de dependências externas. <a href="https://pypi.org/project/kaleido/">https://pypi.org/project/kaleido/</a>
xgboost		Implementa algoritmos de aprendizado de máquina sob a estrutura Gradient Boosting. <a href="https://xgboost.readthedocs.io/en/stable/">https://xgboost.readthedocs.io/en/stable/</a>
yellowbrick	yellowbrick.features <hr/> yellowbrick.regressor	FeatureImportances <hr/> PredictionError <hr/> ResidualsPlot
		Estende a API Scikit-Learn para facilitar a seleção de modelos e o ajuste de hiperparâmetros. <a href="https://www.scikit-yb.org/en/latest/">https://www.scikit-yb.org/en/latest/</a>

## ANEXO II: Variáveis dos registros de ocorrência do GBIF

Nome da variável	Tipo da variável	Definição <sup>125</sup>
acceptedScientificName	object	O nome atualmente válido (zoológico) ou aceito (botânico) para o nome científico.
acceptedTaxonKey	int64	*
accessRights	object	Uma descrição das restrições sobre o uso dos dados como compartilhados ou acesso a outros dados que não são compartilhados.
associatedReferences	object	Uma lista (concatenada e separada) de identificadores (publicação, referência bibliográfica, identificador único global, URI) da literatura associada à Ocorrência.
associatedSequences	object	Uma lista (concatenada e separada) de identificadores (publicação, identificador único global, URI) de informações de sequência genética associadas à Ocorrência.
basisOfRecord	object	A natureza específica do registro de dados. Exemplos: PreservedSpecimen, FossilSpecimen, LivingSpecimen, MaterialSample, Event, HumanObservation, MachineObservation, Taxon, Occurrence, MaterialCitation
behavior	object	Uma descrição do comportamento apresentado pelo sujeito no momento em que a Ocorrência foi registrada.
bibliographicCitation	object	Uma referência bibliográfica para o recurso.
catalogNumber	object	Um identificador (de preferência único) para o registro dentro do conjunto de dados ou coleção.

<sup>125</sup> TDWG - List of Darwin Core terms. Disponível em <<https://dwc.tdwg.org/list/>>. Acesso em 27 nov 2022.

<b>Nome da variável</b>	<b>Tipo da variável</b>	<b>Definição<sup>125</sup></b>
class	object	O nome científico completo da classe na qual o táxon está classificado.
classKey	int64	*
collectionCode	object	O nome, acrônimo, código ou inicialismo que identifica a coleção ou conjunto de dados do qual o registro foi derivado.
collectionID	object	*
collectionKey	object	*
continent	object	O nome do continente em que o Local ocorre.
coordinatePrecision	float64	Uma representação decimal da precisão das coordenadas fornecidas na Latitude decimal e na Longitude decimal.
coordinateUncertaintyInMeters	float64	A distância horizontal (em metros) da dada decimalLatitude e decimalLongitude que descreve o menor círculo que contém a localização inteira. Deixe o valor em branco se a incerteza for desconhecida, não puder ser estimada ou não for aplicável (porque não há coordenadas). Zero não é um valor válido para este termo.
country	object	O nome do país ou unidade administrativa principal em que o Local ocorre.
countryCode	object	O código padrão para o país em que o Local ocorre.



Nome da variável	Tipo da variável	Definição <sup>125</sup>
county	object	O nome completo e não abreviado da próxima região administrativa menor que stateProvince (condado, condado, departamento etc.) em que o Local ocorre.
crawlId	int64	*
datasetID	object	Um identificador para o conjunto de dados. Pode ser um identificador global exclusivo ou um identificador específico de uma coleção ou instituição.
datasetKey	object	*
datasetName	object	O nome que identifica o conjunto de dados do qual o registro foi derivado.
dateIdentified	object	A data em que o espécime foi determinado como representante do Táxon.
day	float64	O dia inteiro do mês em que o Evento ocorreu.
decimalLatitude	float64	A latitude geográfica (em graus decimais, usando o sistema de referência espacial dado em geodeticDatum) do centro geográfico de um Local. Os valores positivos estão ao norte do Equador, os valores negativos estão ao sul dele. Os valores legais situam-se entre -90 e 90, inclusive.
decimalLongitude	float64	A longitude geográfica (em graus decimais, usando o sistema de referência espacial dado em geodeticDatum) do centro geográfico de um Local. Os valores positivos estão a leste do Meridiano de Greenwich, os valores negativos estão a oeste dele. Os valores legais estão entre -180 e 180, inclusive.

<b>Nome da variável</b>	<b>Tipo da variável</b>	<b>Definição<sup>125</sup></b>
disposition	object	O estado atual de uma amostra em relação à coleção identificada em collectionCode ou collectionID.
dynamicProperties	object	Uma lista de medições adicionais, fatos, características ou afirmações sobre o registro. Destina-se a fornecer um mecanismo para conteúdo estruturado.
earliestAgeOrLowestStage	object	O nome completo da idade geocronológica mais precoce possível ou estágio cronoestratigráfico mais baixo atribuível ao horizonte stratigráfico do qual o item catalogado foi coletado. Exemplos: Atlantic, Boreal, Skullrockian
elevation	float64	*
elevationAccuracy	float64	*
endDayOfYear	object	O último dia inteiro do ano em que o Evento ocorreu (1 para 1º de janeiro, 365 para 31 de dezembro, exceto em um ano bissexto, caso em que é 366).
establishmentMeans	object	O processo pelo qual o(s) indivíduo(s) biológico(s) representado(s) na Ocorrência se estabeleceram no local.
eventDate	object	A data-hora ou intervalo durante o qual ocorreu um Evento. Para ocorrências, esta é a data e hora em que o evento foi registrado. Não é adequado para um tempo em um contexto geológico.
eventRemarks	object	Comentários ou notas sobre o Evento.
eventTime	object	O tempo ou intervalo durante o qual um Evento ocorreu.

Nome da variável	Tipo da variável	Definição <sup>125</sup>
extensions	object	*
facts	object	*
family	object	O nome científico completo da família na qual o táxon é classificado.
familyKey	float64	*
fieldNumber	object	Um identificador fornecido ao evento no campo. Muitas vezes serve como um elo entre as notas de campo e o Evento.
footprintWKT	object	Uma representação de texto conhecido ( <i>Well-Known Text</i> - WKT) da forma (pegada, geometria) que define o local. Um local pode ter uma representação de ponto-raio (consulte decimalLatitude) e uma representação de pegada, e elas podem diferir uma da outra.
formation	object	O nome completo da formação litoestratigráfica da qual o item catalogado foi coletado.
gadm	object	*
gbifID	object	*
genericName	object	A parte do gênero do nome científico sem autoria.
genus	object	O nome científico completo do gênero no qual o táxon é classificado.

<b>Nome da variável</b>	<b>Tipo da variável</b>	<b>Definição<sup>125</sup></b>
genusKey	float64	*
geodeticDatum	object	O elipsóide, datum geodésico ou sistema de referência espacial (SRS) no qual se baseiam as coordenadas geográficas fornecidas em decimalLatitude e decimalLongitude.
georeferenceProtocol	object	Uma descrição ou referência aos métodos usados para determinar a pegada espacial, coordenadas e incertezas.
georeferenceRemarks	object	Notas ou comentários sobre a determinação da descrição espacial, explicando suposições adicionais ou contrapostas às formalizadas no método referido no Protocolo de georreferenciamento.
georeferenceSources	object	Um mapa, dicionário geográfico ou outro recurso usado para georreferenciar o local.
georeferenceVerificationStatus	object	Uma descrição categórica da extensão em que a georeferência foi verificada para representar a melhor descrição espacial possível para a Localização da Ocorrência.
georeferencedBy	object	Uma lista (concatenada e separada) de nomes de pessoas, grupos ou organizações que determinaram a georeferência (representação espacial) para o Local.
georeferencedDate	object	A data em que o Local foi georreferenciado.
group	object	O nome completo do grupo litoestratigráfico do qual o item catalogado foi coletado.
habitat	object	Uma categoria ou descrição do habitat em que o Evento ocorreu.

<b>Nome da variável</b>	<b>Tipo da variável</b>	<b>Definição<sup>125</sup></b>
higherClassification	object	Uma lista (concatenada e separada) de nomes de táxons terminando na classificação imediatamente superior ao táxon referenciado no registro do táxon.
higherGeography	object	Uma lista (concatenada e separada) de nomes geográficos menos específicos do que as informações capturadas no termo de localidade.
higherGeographyID	object	Um identificador para a região geográfica na qual o local ocorreu.
hostingOrganizationKey	object	*
<a href="http://unknown.org/canonicalName">http://unknown.org/canonicalName</a>	object	*
<a href="http://unknown.org/captive">http://unknown.org/captive</a>	object	*
<a href="http://unknown.org/combinationAuthors">http://unknown.org/combinationAuthors</a>	object	*
<a href="http://unknown.org/combinationYear">http://unknown.org/combinationYear</a>	object	*
<a href="http://unknown.org/language">http://unknown.org/language</a>	object	*
<a href="http://unknown.org/nick">http://unknown.org/nick</a>	object	*
<a href="http://unknown.org/rights">http://unknown.org/rights</a>	object	*
<a href="http://unknown.org/rightsHolder">http://unknown.org/rightsHolder</a>	object	*

Nome da variável	Tipo da variável	Definição <sup>125</sup>
<a href="http://unknown.org/verbatimScientificName">http://unknown.org/verbatimScientificName</a>	object	*
identificationID	object	Um identificador para a Identificação (o conjunto de informações associado à atribuição de um nome científico). Pode ser um identificador global exclusivo ou um identificador específico para o conjunto de dados.
identificationQualifier	object	Um valor controlado para expressar as dúvidas do determinador sobre a Identificação.
identificationReferences	object	Uma lista (concatenada e separada) de referências (publicação, identificador único global, URI) usadas na Identificação.
identificationRemarks	object	Comentários ou notas sobre a Identificação.
identificationVerificationStatus	object	Um indicador categórico da extensão em que a identificação taxonômica foi verificada como correta.
identifiedBy	object	Uma lista (concatenada e separada) de nomes de pessoas, grupos ou organizações que atribuíram o Táxon ao sujeito.
identifiedByIDs	object	*
identifier	object	*
identifiers	object	*
inCluster	bool	*

<b>Nome da variável</b>	<b>Tipo da variável</b>	<b>Definição<sup>125</sup></b>
individualCount	float64	O número de indivíduos presentes no momento da Ocorrência.
informationWithheld	object	Informações adicionais que existem, mas que não foram compartilhadas no registro fornecido.
infraspecificEpithet	object	O nome do epíteto infraespecífico mais baixo ou terminal do nome científico, excluindo qualquer designação de classificação.
installationKey	object	*
institutionCode	object	O nome (ou sigla) em uso pela instituição detentora do(s) objeto(s) ou informação referida no registro.
institutionID	object	Um identificador para a instituição que detém a custódia do(s) objeto(s) ou informação referida no registro.
institutionKey	object	*
isInCluster	bool	*
island	object	O nome da ilha em ou perto da qual o Local ocorre.
issues	object	*
iucnRedListCategory	object	*

<b>Nome da variável</b>	<b>Tipo da variável</b>	<b>Definição<sup>125</sup></b>
key	int64	*
kingdom	object	O nome científico completo do reino em que o táxon é classificado.
kingdomKey	int64	*
language	object	O idioma do recurso.
lastCrawled	object	*
lastInterpreted	object	*
lastParsed	object	*
latestAgeOrHighestStage	object	O nome completo da última idade geocronológica possível ou estágio cronoestratigráfico mais alto atribuível ao horizonte estratigráfico do qual o item catalogado foi coletado.
license	object	Um documento legal que dá permissão oficial para fazer algo com o recurso.
lifeStage	object	A classe de idade ou estágio de vida do(s) Organismo(s) no momento em que a Ocorrência foi registrada.
lithostratigraphicTerms	object	A combinação de todos os nomes litoestratigráficos para a rocha da qual o item catalogado foi coletado.
locality	object	A descrição específica do local.



<b>Nome da variável</b>	<b>Tipo da variável</b>	<b>Definição<sup>125</sup></b>
locationAccordingTo	object	Informações sobre a fonte dessas informações de localização. Pode ser uma publicação (diário), instituição ou equipe de indivíduos.
locationID	object	Um identificador para o conjunto de informações de localização (dados associados a dterms:Location). Pode ser um identificador global exclusivo ou um identificador específico para o conjunto de dados.
locationRemarks	object	Observações de localização
media	object	*
modified	object	A data e hora mais recente em que o recurso foi alterado.
month	float64	O mês inteiro em que o Evento ocorreu.
municipality	object	O nome completo e não abreviado da próxima região administrativa menor que o condado (cidade, município etc.) em que o Local ocorre. Não use este termo para um local nomeado próximo que não contenha a localização real.
namePublishedInYear	object	O ano de quatro dígitos em que o nome científico foi publicado.
networkKeys	object	*
nomenclaturalCode	object	O código nomenclatural (ou códigos no caso de um nome ambirenal) sob o qual o nome científico é construído.

<b>Nome da variável</b>	<b>Tipo da variável</b>	<b>Definição<sup>125</sup></b>
occurrenceID	object	Um identificador para a Ocorrência (em oposição a um registro digital específico da ocorrência). Na ausência de um identificador exclusivo global persistente, construa um a partir de uma combinação de identificadores no registro que tornará mais próximo o código de ocorrência globalmente exclusivo.
occurrenceRemarks	object	Comentários ou notas sobre a Ocorrência.
occurrenceStatus	object	Uma declaração sobre a presença ou ausência de um Táxon em um Local.
order	object	O nome científico completo da ordem em que o táxon é classificado.
orderKey	float64	*
organismID	object	Um identificador para a instância do Organismo (em oposição a um registro digital específico do Organismo). Pode ser um identificador globalmente exclusivo ou um identificador específico para o conjunto de dados.
ownerInstitutionCode	object	O nome (ou sigla) em uso pela instituição detentora do(s) objeto(s) ou informação referida no registro.
phylum	object	O nome científico completo do filo ou divisão em que o táxon é classificado.
phylumKey	int64	*
preparations	object	Um método de preparação ou preservação para uma amostra.

<b>Nome da variável</b>	<b>Tipo da variável</b>	<b>Definição<sup>125</sup></b>
previousIdentifications	object	Uma lista (concatenada e separada) de ScientificNames anteriores aos quais a amostra foi identificada.
projectId	object	*
protocol	object	*
publishingCountry	object	*
publishingOrgKey	object	*
recordNumber	object	Um identificador dado à Ocorrência no momento em que foi registrada. Muitas vezes serve como um elo entre as notas de campo e um registro de Ocorrência, como o número de um coletor de espécimes.
recordedBy	object	Uma lista (concatenada e separada) de nomes de pessoas, grupos ou organizações responsáveis por registrar a Ocorrência original. O coletor ou observador primário, especialmente aquele que aplica um identificador pessoal (recordNumber), deve ser listado primeiro.
recordedByIDs	object	*
references	object	Um recurso relacionado que é referenciado, citado ou apontado de outra forma pelo recurso descrito.
relations	object	Um identificador para um recurso relacionado (o objeto, em vez do assunto do relacionamento).

Nome da variável	Tipo da variável	Definição <sup>125</sup>
reproductiveCondition	object	A condição reprodutiva do(s) indivíduo(s) biológico(s) representado(s) na Ocorrência.
rights	object	Informações sobre os direitos detidos no e sobre o recurso.
rightsHolder	object	Uma pessoa ou organização que possui ou gerencia direitos sobre o recurso.
samplingEffort	object	A quantidade de esforço gasto durante um Evento.
samplingProtocol	object	Os métodos ou protocolos usados durante um Evento, indicados por um IRI ( <i>Internationalized Resource Identifier</i> ).
scientificName	object	O nome científico completo, com informações de autoria e data, se conhecido. Ao fazer parte de uma Identificação, este deve ser o nome na classificação taxonômica de nível mais baixo que pode ser determinado. Este termo não deve conter qualificações de identificação, que devem ser fornecidas no termo IdentificationQualifier.
scientificNameID	object	Um identificador para os detalhes nomenclaturais (não taxonômicos) de um nome científico.
sex	object	O sexo biológico do(s) indivíduo(s) representado(s) na Ocorrência.
species	object	*
speciesKey	float64	*
specificEpithet	object	O nome do primeiro ou epíteto da espécie do Nome científico.

Nome da variável	Tipo da variável	Definição <sup>125</sup>
startDayOfYear	object	O primeiro dia inteiro do ano em que o Evento ocorreu (1 para 1º de janeiro, 365 para 31 de dezembro, exceto em um ano bissexto, caso em que é 366).
stateProvince	object	O nome da próxima região administrativa menor que o país (estado, província, cantão, departamento, região etc.) em que o Local ocorre.
taxonConceptID	object	Um identificador para o conceito taxonômico ao qual o registro se refere - não para os detalhes nomenclaturais de um táxon.
taxonID	object	Um identificador global exclusivo para o táxon (nome em uma classificação).
taxonKey	int64	*
taxonRank	object	A classificação taxonômica do nome mais específico no nome científico.
taxonRemarks	object	Comentários ou notas sobre o táxon ou nome.
taxonomicStatus	object	O status do uso do nome científico como um rótulo para um táxon. Requer opinião taxonômica para definir o escopo de um táxon. Regras de prioridade são então utilizadas para definir o status taxonômico contido da nomenclatura escopo, combinados com a opinião dos especialistas. Deve estar vinculado a uma referência taxonômica específica que define o conceito.
type	object	A natureza ou gênero do recurso.
typeStatus	object	Uma lista (concatenada e separada) de tipos nomenclaturais (status do tipo, nome científico tipificado, publicação) aplicados ao assunto.

<b>Nome da variável</b>	<b>Tipo da variável</b>	<b>Definição<sup>125</sup></b>
typifiedName	object	*
verbatimCoordinateSystem	object	O formato de coordenadas para a Latitude literal e Longitude literal ou as Coordenadas literais do Local.
verbatimEventDate	object	A representação original literal das informações de data e hora de um Evento.
verbatimLabel	object	*
verbatimLocality	object	A descrição textual original do local.
verbatimTaxonRank	object	A classificação taxonômica do nome mais específico no nome científico como aparece no registro original.
vernacularName	object	Um nome comum ou vernáculo.
waterBody	object	O nome do corpo d'água em que o Local ocorre.
year	float64	O ano de quatro dígitos em que o Evento ocorreu, de acordo com o Calendário da Era Comum.

**ANEXO III: Lista das Unidades de Conservação que estão na base de dados do Instituto Socioambiental e foram consideradas neste estudo**

<b>Nome</b>	<b>Uso</b>	<b>Categoria</b>	<b>Esfera</b>	<b>Área (ha)</b>	<b>% na área de estudo*</b>
<b>ESEC de Angatuba</b>	Proteção Integral	ESEC	Estadual	1.366,1	100%
ESEC de Assis	Proteção Integral	ESEC	Estadual	1.797,8	100%
ESEC de Bananal	Proteção Integral	ESEC	Estadual	878,6	100%
ESEC de Ibicatu	Proteção Integral	ESEC	Estadual	85,8	100%
ESEC de Itaberá	Proteção Integral	ESEC	Estadual	191,1	100%
ESEC de Itapeti	Proteção Integral	ESEC	Estadual	84,2	100%
ESEC de Itapeva	Proteção Integral	ESEC	Estadual	92,5	100%
ESEC de Itirapina	Proteção Integral	ESEC	Estadual	2.292,0	100%
ESEC de Jataí	Proteção Integral	ESEC	Estadual	8.852,1	100%
ESEC de Mata do Jacaré	Proteção Integral	ESEC	Estadual	67,4	100%
ESEC de Mogi-Guaçu - Prof. Hermógenes de Freitas L. Filho	Proteção Integral	ESEC	Estadual	975,7	100%
ESEC de Paranapanema	Proteção Integral	ESEC	Estadual	631,6	100%
ESEC de Paulo de Faria	Proteção Integral	ESEC	Estadual	440,3	100%
ESEC de Ribeirão Preto	Proteção Integral	ESEC	Estadual	147,5	100%
ESEC de Santa Bárbara	Proteção Integral	ESEC	Estadual	3.172,3	100%
ESEC de Santa Maria	Proteção Integral	ESEC	Estadual	1.315,7	100%
ESEC de Valinhos	Proteção Integral	ESEC	Estadual	18,1	100%
ESEC do Noroeste Paulista	Proteção Integral	ESEC	Estadual	188,2	100%
ESEC dos Caetetus	Proteção Integral	ESEC	Estadual	2.320,4	100%
ESEC dos Chauás	Proteção Integral	ESEC	Estadual	2.648,0	100%
ESEC Juréia-Itatins	Proteção Integral	ESEC	Estadual	84.611,9	100%
ESEC Mico Leão Preto	Proteção Integral	ESEC	Federal	6.697,6	100%
ESEC Sebastião Aleixo da Silva (Bauru)	Proteção Integral	ESEC	Estadual	284,1	100%
ESEC Xitué	Proteção Integral	ESEC	Estadual	2.847,0	100%
MONAT da Pedra Grande	Proteção Integral	MONAT	Estadual	3.305,6	100%
PARNA da Serra da Bocaina	Proteção Integral	PARNA	Federal	104.940,4	35%
PE Alberto Lofgren (Horto)	Proteção Integral	PE	Estadual	179,1	100%
PE ARA	Proteção Integral	PE	Estadual	65,4	100%

<b>Nome</b>	<b>Uso</b>	<b>Categoria</b>	<b>Esfera</b>	<b>Área (ha)</b>	<b>% na área de estudo*</b>
PE Cantareira	Proteção Integral	PE	Estadual	6.891,3	100%
PE Carlos Botelho	Proteção Integral	PE	Estadual	39.837,3	100%
PE Caverna do Diabo	Proteção Integral	PE	Estadual	40.027,0	100%
PE da Campina do Encantado	Proteção Integral	PE	Estadual	3.133,2	100%
PE da Ilha Anchieta	Proteção Integral	PE	Estadual	823,8	90%
PE da Ilha do Cardoso	Proteção Integral	PE	Estadual	13.576,6	99%
PE da Serra do Mar	Proteção Integral	PE	Estadual	305.222,9	100%
PE das Fontes do Ipiranga	Proteção Integral	PE	Estadual	477,5	100%
PE das Furnas do Bom Jesus	Proteção Integral	PE	Estadual	2.159,0	100%
PE de Campos do Jordão	Proteção Integral	PE	Estadual	7.921,8	100%
PE de Ilhabela	Proteção Integral	PE	Estadual	28.903,9	99%
PE de Itaberaba	Proteção Integral	PE	Estadual	13.288,7	100%
PE de Itapetinga	Proteção Integral	PE	Estadual	10.218,3	100%
PE de Porto Ferreira	Proteção Integral	PE	Estadual	627,5	100%
PE de Vassununga	Proteção Integral	PE	Estadual	2.045,7	100%
PE do Aguapeí	Proteção Integral	PE	Estadual	9.074,0	100%
PE do Itinguçu	Proteção Integral	PE	Estadual	4.983,4	99%
PE do Juquery	Proteção Integral	PE	Estadual	1.930,2	100%
PE do Jurupará	Proteção Integral	PE	Estadual	24.137,0	100%
PE do Lagamar de Cananéia	Proteção Integral	PE	Estadual	40.836,3	99%
PE do Morro do Diabo	Proteção Integral	PE	Estadual	36.641,5	100%
PE do Prelado	Proteção Integral	PE	Estadual	1.837,8	100%
PE do Rio Peixe	Proteção Integral	PE	Estadual	7.733,2	100%
PE do Rio Turvo	Proteção Integral	PE	Estadual	74.032,5	99%
PE dos Mananciais de Campos do Jordão	Proteção Integral	PE	Estadual	487,4	100%
PE Intervalos	Proteção Integral	PE	Estadual	41.977,0	100%
PE Jaraguá	Proteção Integral	PE	Estadual	485,5	100%
PE Nascentes do Paranapanema	Proteção Integral	PE	Estadual	22.113,5	100%
PE Nascentes do Tietê	Proteção Integral	PE	Estadual	122,7	100%
PE Turístico do Alto Ribeira - PETAR	Proteção Integral	PE	Estadual	35.851,4	100%



<b>Nome</b>	<b>Uso</b>	<b>Categoria</b>	<b>Esfera</b>	<b>Área (ha)</b>	<b>% na área de estudo*</b>
PE Xixová-Japuí	Proteção Integral	PE	Estadual	897,2	70%
REBIO do Alto da Serra de Paranapiacaba	Proteção Integral	REBIO	Estadual	336,8	100%
REBIO Fazenda Experimental de Zootecnia de Sertãozinho	Proteção Integral	REBIO	Estadual	2.373,0	100%
REBIO Mogi-Guaçu	Proteção Integral	REBIO	Estadual	353,0	100%
APA Bairro da Usina	Uso Sustentável	APA	Estadual	998,4	100%
APA Cabreúva	Uso Sustentável	APA	Estadual	25.801,5	100%
APA Cajamar	Uso Sustentável	APA	Estadual	12.830,3	100%
APA Campos do Jordão	Uso Sustentável	APA	Estadual	29.015,0	99%
APA Cananéia - Iguape - Peruíbe	Uso Sustentável	APA	Federal	192.241,4	100%
APA Corumbatai, Botucatu e Tejuπά	Uso Sustentável	APA	Estadual	615.148,5	100%
APA da Ilha Comprida	Uso Sustentável	APA	Estadual	18.184,4	100%
APA da Serra da Mantiqueira	Uso Sustentável	APA	Federal	432.605,6	27%
APA da Serra do Mar	Uso Sustentável	APA	Estadual	378.890,3	100%
APA da Várzea do Rio Tietê	Uso Sustentável	APA	Estadual	10.198,2	100%
APA das Ilhas e Várzeas do Rio Paraná	Uso Sustentável	APA	Federal	1.012.290,1	1%
APA de Cajati	Uso Sustentável	APA	Estadual	2.976,9	100%
APA do Planalto do Turvo	Uso Sustentável	APA	Estadual	2.724,7	100%
APA do Rio Pardinho e do Rio Vermelho	Uso Sustentável	APA	Estadual	3.236,8	99%
APA dos Quilombos do Médio Ribeira	Uso Sustentável	APA	Estadual	65.671,4	100%
APA dos Rios Piracicaba e Juqueri-Mirim	Uso Sustentável	APA	Estadual	399.282,0	100%
APA Haras São Bernardo	Uso Sustentável	APA	Estadual	35,9	100%
APA Ibitinga	Uso Sustentável	APA	Estadual	65.321,8	100%
APA Ituparanga	Uso Sustentável	APA	Estadual	95.393,5	100%
APA Jundiá	Uso Sustentável	APA	Estadual	43.298,8	100%
APA Mananciais da Bacia do Rio Paraíba do Sul	Uso Sustentável	APA	Federal	292.839,2	100%
APA Mata do Iguatemi	Uso Sustentável	APA	Estadual	32,9	100%
APA Morro de São Bento	Uso Sustentável	APA	Estadual	33,9	100%
APA Parque e Fazenda do Carmo	Uso Sustentável	APA	Estadual	996,8	100%
APA Rio Batalha	Uso Sustentável	APA	Estadual	232.654,9	100%

<b>Nome</b>	<b>Uso</b>	<b>Categoria</b>	<b>Esfera</b>	<b>Área (ha)</b>	<b>% na área de estudo*</b>
APA São Francisco Xavier	Uso Sustentável	APA	Estadual	11.161,7	99%
APA Sapucaí-Mirim	Uso Sustentável	APA	Estadual	38.178,8	99%
APA Silveiras	Uso Sustentável	APA	Estadual	41.571,7	100%
APA Sistema Cantareira	Uso Sustentável	APA	Estadual	254.955,0	100%
APA Tietê	Uso Sustentável	APA	Estadual	46.315,2	100%
ARIE Buriti de Vassununga	Uso Sustentável	ARIE	Federal	151,4	100%
ARIE Cerrado Pé-de-Gigante	Uso Sustentável	ARIE	Federal	1.202,2	100%
ARIE da Ilha Comprida	Uso Sustentável	ARIE	Estadual	11.749,7	100%
ARIE da Pedra Branca	Uso Sustentável	ARIE	Estadual	639,1	100%
ARIE de São Sebastião	Uso Sustentável	ARIE	Estadual	640,8	56%
ARIE Guará	Uso Sustentável	ARIE	Estadual	445,8	93%
ARIE Ilha do Ameixal	Uso Sustentável	ARIE	Federal	434,2	100%
ARIE Mata de Santa Genebra	Uso Sustentável	ARIE	Federal	252,6	100%
ARIE Matão de Cosmópolis	Uso Sustentável	ARIE	Federal	169,5	100%
FES Assis	Uso Sustentável	FES	Estadual	3.152,2	100%
FES Avaré I e II	Uso Sustentável	FES	Estadual	721,8	100%
FES de Guarulhos	Uso Sustentável	FES	Estadual	98,5	100%
FES de Pederneiras	Uso Sustentável	FES	Estadual	1.948,0	100%
FES Edmundo Navarro de Andrade	Uso Sustentável	FES	Estadual	2.227,3	100%
FES Manduri	Uso Sustentável	FES	Estadual	1.497,7	100%
FES Piraju	Uso Sustentável	FES	Estadual	684,0	100%
FLONA de Capão Bonito	Uso Sustentável	FLONA	Federal	4.246,5	100%
FLONA de Ipanema	Uso Sustentável	FLONA	Federal	5.330,1	100%
FLONA de Lorena	Uso Sustentável	FLONA	Federal	282,1	100%
RDS Barra do Una	Uso Sustentável	RDS	Estadual	1.463,6	98%
RDS Barreiro Anhemas	Uso Sustentável	RDS	Estadual	3.272,9	100%
RDS de Itapanhapima	Uso Sustentável	RDS	Estadual	1.243,3	100%
RDS de Lavras	Uso Sustentável	RDS	Estadual	889,0	100%
RDS do Despraiado	Uso Sustentável	RDS	Estadual	3.969,3	100%
RDS dos Pinheirinhos	Uso Sustentável	RDS	Estadual	1.483,0	99%

<b>Nome</b>	<b>Uso</b>	<b>Categoria</b>	<b>Esfera</b>	<b>Área (ha)</b>	<b>% na área de estudo*</b>
RDS dos Quilombos de Barra do Turvo	Uso Sustentável	RDS	Estadual	5.835,5	100%
RESEX da Ilha do Tumba	Uso Sustentável	RESEX	Estadual	1.113,0	100%
RESEX do Mandira	Uso Sustentável	RESEX	Federal	1.091,3	100%
RESEX Taquari	Uso Sustentável	RESEX	Estadual	1.651,0	100%

\*Os percentuais referem-se à intersecção entre os planos de informação de Unidades de Conservação e da área de estudo.