

UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
INSTITUTO DE GEOGRAFIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM GEOGRAFIA

Laboratório de Pedologia e Estudos da Paisagem - LAPEP

**A DISTRIBUIÇÃO DAS FORMAÇÕES VEGETACIONAIS DO  
SEMIÁRIDO BRASILEIRO NOS ÚLTIMOS 21.000 ANOS**

*Rafael César Silva*

**Orientador:**

*Prof. Dr. Guilherme Resende Corrêa*

**Uberlândia, março de 2021.**

UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
INSTITUTO DE GEOGRAFIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM GEOGRAFIA

Laboratório de Pedologia e Estudos da Paisagem - LAPEP

**A DISTRIBUIÇÃO DAS FORMAÇÕES VEGETACIONAIS DO  
SEMIÁRIDO BRASILEIRO NOS ÚLTIMOS 21.000 ANOS**

Rafael César Silva

Dissertação realizada  
como requisito  
obrigatório para obtenção  
do título de Mestre em  
Geografia na  
Universidade Federal de  
Uberlândia

Banca Examinadora:

Prof. Dr. Guilherme Resende Corrêa (Instituto de Geografia/UFU) - Orientador

Prof.<sup>a</sup> Dr<sup>a</sup> Prímula Viana Campos (Universidade Federal de Viçosa)

Prof. Dr. Daniel Meira Arruda (Universidade Estadual de Montes Claros)

Prof. Dr. Marcelo Leandro Bueno (Universidade Estadual do Mato Grosso do Sul)

**Uberlândia, março de 2021.**

Dados Internacionais de Catalogação na Publicação (CIP) Sistema de Bibliotecas da  
UFU, MG, Brasil.

---

S586d Silva, Rafael César, 1996-  
2021 A distribuição das formações vegetacionais do semiárido brasileiro nos últimos 21.000 anos  
[recurso eletrônico] / Rafael César Silva. - 2021.

Orientador: Guilherme Resende Corrêa.

Dissertação (mestrado) - Universidade Federal de Uberlândia.

Programa de Pós-Graduação em Geografia.

Modo de acesso: Internet.

Disponível em: <http://doi.org/10.14393/ufu.di.2021.5513> Inclui bibliografia.

Inclui ilustrações.

1. Geografia. I. Corrêa, Guilherme Resende, 1981-, (Orient.). II. Universidade Federal  
de Uberlândia. Programa de Pós-Graduação em Geografia. III. Título.

CDU:910.1

---

Glória Aparecida – CRB-6/2047

# A DISTRIBUIÇÃO DAS FORMAÇÕES VEGETACIONAIS DO SEMIÁRIDO BRASILEIRO NOS ÚLTIMOS 21.000 ANOS

Dissertação realizada como requisito obrigatório para obtenção do título de Mestre em Geografia na Universidade Federal de Uberlândia pela banca examinadora formada por:

Uberlândia, 31 de março de 2021.



---

Prof.<sup>a</sup> Dr<sup>a</sup> Prímula Viana Campos, UFV/MG



---

Prof. Dr. Daniel Meira Arruda, UNIMONTES/MG



---

Prof. Dr. Marcelo Leandro Bueno, UEMS/MS

Dedico este trabalho à minha família,  
amigos e professores.

## AGRADECIMENTOS

Primeiramente gostaria de agradecer aos meus pais Carlos Cezar e Maria Madalena e meu irmão Samuel, por me proporcionarem todas as ferramentas e suporte possível para desenvolvimento deste mestrado. Sem o apoio familiar nenhuma das minhas conquistas seriam alcançadas. Deixo também meu agradecimento à minha companheira Jéssica, que me forneceu apoio sem fim durante todos estes anos.

Gostaria de agradecer ao meu orientador, professor Guilherme, por toda liderança, disposição e paciência para que este trabalho fosse desenvolvido apesar de todas as dificuldades nos últimos anos. Gostaria de agradecer também aos professores Daniel Arruda, Guilherme Oliveira, Elpídio Inácio, Gustavo Veloso e Prímula Viana pelas inúmeras contribuições ao longo do trabalho.

Estendo aqui meu enorme obrigado a todos meus colegas de laboratório, pois a convivência diária foi minha energia durante diversos momentos para conclusão. Em especial deixo meu agradecimento ao Gabriel Palucci, Márcio Vêna e Viviane Koch pelo auxílio prestado no trabalho de campo e posterior. Agradeço também a contribuição de todos meus amigos em todos estes anos de mestrado, a citação de cada um deles aqui tornaria a lista extensa e não faria justiça a enorme contribuição emocional, intelectual e muitas vezes metafísica de cada um.

Agradeço aos meus professores ao longo desta caminhada, pois cada aula, dica e momento foi de grande valia. Sou grato pela estrutura física fornecida pelas universidades, em especial a UFU, UFV e UFMG que participaram deste trabalho de forma direta. Também deixo registrado o meu agradecimento a cada um dos profissionais envolvidos na confecção das muitas bases de dados utilizadas neste trabalho, disponibilizadas de forma gratuita e facilitada, sem as quais não seria possível a realização dele.

Agradeço a CNPQ, CAPES, UFU e IG pelo apoio financeiro e estrutural que me foi concedido.

## RESUMO

SILVA, R. C. **A distribuição das formações vegetacionais do Semiárido Brasileiro nos últimos 21.000 anos.** 2021. 61 f. Dissertação (Mestrado em Geografia) – Programa de Pós-Graduação em Geografia, Instituto de Geografia, Universidade Federal de Uberlândia. Uberlândia, 2021.

No contexto da América do Sul umas das principais regiões que necessitam de abordagem nos estudos paleoecológicos é o Semiárido Brasileiro, uma área marcada pela aridez no presente e está localizada em um encontro de vegetações de diferentes biomas. O passado da região reconstruído pelos modelos paleoclimáticos nos períodos do Pleistoceno e Holoceno está relacionado com mudanças climáticas intensas em curtas escalas de tempo. Neste trabalho nosso objetivo foi utilizar a combinação entre preditores bioclimáticos e edáficos para construir modelos com algoritmos de *machine learning* capazes de prever a vegetação do presente e dos últimos 21.000 anos. Nossos resultados evidenciam que o ganho de acurácia na classificação das formações vegetacionais é constante na medida que se acrescentam preditores edáficos. A opção por utilizar horizontes subsuperficiais para gerar atributos edáficos também resultou em um maior detalhamento local na distribuição da vegetação. Nas predições realizadas em períodos passados, a maior alteração é promovida durante o período Heinrich Stadial 1, com a retração das áreas de formações vegetacionais abertas de Caatinga Hiperxerófila, avanço das áreas de Caatinga Arbórea, Caatinga Florestada e Floresta Estacional Decidual. O padrão de vegetação presente teve seu estabelecimento observado apenas no Holoceno Superior, o que indica que as condições prévias de umidade no Semiárido brasileiro permitiram uma ocupação maior de formações florestais durante o final do Pleistoceno.

**Palavras-chave:** Semiárido Brasileiro; Formações Vegetacionais; Pleistoceno; Holoceno; Modelagem Paleoclimática; Solos;

## ABSTRACT

SILVA, R. C. **The distribution of vegetation formations in the Brazilian Semiarid region over the past 21,000 years.** 2021. 61 p. Dissertation (Master in Geography) - Postgraduate Program in Geography, Institute of Geography, Federal University of Uberlândia. Uberlândia, 2021.

In the context of South America, one of the main regions that need to be approached by paleoecological studies is the Brazilian Semiarid, an area marked by aridity in the present and located in a meeting of vegetation from different biomes. The region's past reconstructed by paleoclimatic models in the Pleistocene and Holocene periods is related to intense climatic changes in short time scales. In this work, our objective was to use the combination of bioclimatic and edaphic predictors to build models with machine learning algorithms capable of predicting the vegetation of the present and the last 21,000 years. Our results show that the gain in accuracy in the classification of vegetation formations is constant as edaphic predictors are added. The option to use subsurface horizons to generate edaphic attributes also resulted in greater local detail in the distribution of vegetation. In predictions made in past periods, the greatest change is promoted during the Heinrich Stadial 1 period, with the retraction of the areas of open vegetation formations of the Hyperxerophilous Caatinga, advance of the Arboreal Caatinga, Forested Caatinga and Deciduous Forest areas. The vegetation pattern present had its establishment observed only in the Late Holocene, which indicates that the previous humidity conditions in the Brazilian Semiarid allowed a greater occupation of forest formations during the end of the Pleistocene.

**Keywords:** Brazilian Semiarid region; Vegetation Formations; Pleistocene; Holocene; Paleoclimatic Modeling; Soils;



## LISTA DE FIGURAS

Figura 1. Mapa de relevo e áreas de interesse do SAB. ....	16
Figura 2. Localização e vegetação (Formações vegetacionais e Biomas) do Semiárido Brasileiro. ....	17
Figura 3. Solos do Semiárido Brasileiro.. ....	18
Figura 4. Fotografias que exemplificam as formações vegetacionais adotadas no estudo. ....	19
Figura 5. Fluxograma metodológico do estudo. ....	21
Figura 6: Reconstruções paleoclimáticas realizadas por modelos de circulação atmosférica global representando os sete períodos analisados no estudo.. ....	23
Figura 7. Formações vegetacionais potenciais referentes ao período atual (1970- 2000) no Semiárido Brasileiro predita pelos modelos edafoclimáticos.....	26
Figura 8. Reconstrução das formações vegetacionais no SAB durante o final do Pleistoceno e início do Holoceno. ....	28
Figura 9. Variação em % das áreas de vegetação preditas para o passado (21 a 0,4 ky AP) em comparação com a predição presente (1970-2000). ....	29

## LISTA DE ABREVIATURAS E SIGLAS

AOGCM - *Atmospheric-ocean Global Circulation Model* - Modelos de Circulação Global Atmosférica e Oceânica

AP – Antes do Presente (1950)

AvgNnet – Averaged Neural Net

BA - *Bølling-Allerød Stadial* – Interstadial Bølling-Allerød

EH – *Early Holocene* - Holoceno Inferior

Embrapa – Empresa Brasileira de Pesquisa Agropecuária

ESM - *Earth System Models* - Modelos do Sistema Terrestre

HS1 – *Heinrich Stadial 1*

IBGE – Instituto Brasileiro de Geografia e Estatística

KKNN – K-Nearest Neighbor

LGM – *Last Glacial Maximum* – UMG ou Último Máximo Glacial

LH – *Late Holocene* – Holoceno Superior

MH – *Middle Holocene* – Holoceno Médio

RF- randomForest

SAB – Semiárido Brasileiro

SVM – Supported Vector Machine

YD – *Younger Dryas Stadial* – Younger Dryas

ZCIT – Zona de Convergência Intertropical

## SUMÁRIO

<b>1.1 INTRODUÇÃO .....</b>	<b>12</b>
<b>1.2 MATERIAL E MÉTODOS .....</b>	<b>15</b>
<b>1.2.1 – Área de Estudo.....</b>	<b>15</b>
<b>1.2.2 – Amostragem e processamento dos dados.....</b>	<b>19</b>
<b>1.2.3 –Modelagem e espacialização dos modelos de vegetação .....</b>	<b>20</b>
<b>1.2.3.1. – Extração/Pré-processamento dos dados .....</b>	<b>21</b>
<b>1.2.3.2. – Seleção de covariáveis.....</b>	<b>22</b>
<b>1.2.3.3. – Treinamento e espacialização da vegetação atual e dos períodos paleoclimáticos .....</b>	<b>23</b>
<b>1.2.3.3. – Mapas finais e Resultados finais.....</b>	<b>23</b>
<b>1.2.4. – Softwares utilizados .....</b>	<b>24</b>
<b>1.3 RESULTADOS.....</b>	<b>25</b>
<b>1.3.1 – Performance dos Algoritmos.....</b>	<b>25</b>
<b>1.3.2 – Análise da vegetação potencial do presente .....</b>	<b>27</b>
<b>1.3.3 – Reconstrução da vegetação do passado .....</b>	<b>28</b>
<b>1.4 DISCUSSÃO .....</b>	<b>29</b>
<b>1.5 CONCLUSÕES .....</b>	<b>33</b>
<b>1.6 REFERÊNCIAS.....</b>	<b>34</b>
<b>1.7 ANEXOS .....</b>	<b>42</b>

## 1.1 INTRODUÇÃO

A cobertura vegetal da América do Sul durante o Quaternário é objeto frequente de estudos paleoecológicos, no qual um contexto de flutuações climáticas durante o final do Pleistoceno indica a distribuição e riqueza de espécies que encontramos hoje. Duas grandes teorias se apresentam sobre a região como um todo. A Teoria dos Refúgios (HAFFER, J. 1969; AB' SABER, A. 2002) que indica a retração das áreas florestais da Amazônia durante os períodos mais frios e secos do Último Máximo Glacial (21 a 18 ky AP); as áreas de savana com vegetação aberta apresentariam uma maior expansão neste período. Já a teoria do Arco Pleistocênico (PRADO E GIBBS, 1993; PENNINGTON, PRADO E PENDRY, 2000) indica a formação de um arco contínuo de vegetações sazonais durante os períodos glaciais, envolvendo a área da Caatinga, o Cerrado e o Chaco.

Para confirmar ou rejeitar estas teorias, diversos pesquisadores se valem de diferentes abordagens para reconstruir a cobertura vegetal. Uma delas é a utilização de simulações paleoclimáticas geradas por Modelos de Circulação Global Atmosférica e Oceânica (AOGCM - *Atmospheric-ocean Global Circulation Model*) e Modelos do Sistema Terrestre (ESM - *Earth System Models*). A maioria dos trabalhos que utilizaram estes métodos (MAYLE, F.E. 2004; COLLEVATTI et al., 2013; BUENO et al., 2017; ARRUDA et al., 2018; COSTA et al., 2018; MAKSIC et al., 2019; PLUMPTON, MAYLE E WHITNEY, 2020) encontraram resultados satisfatórios que podem ser validados através de registros fósseis. Duas lacunas se apresentam para novos estudos: o aumento da resolução espacial dos resultados e também da resolução temporal dos períodos estudados.

Uma das regiões de interesse para a modelagem ambiental é o Semiárido Brasileiro (SAB), área municipal definida de acordo com o baixo índice pluviométrico. Apesar da atual condição árida do SAB, estudos paleoclimáticos do Quaternário indicam a possibilidade de a região ter sido dominada por condições úmidas até períodos recentes do Holoceno. Reconstruções através de pólen, carvão e sedimentos terrestres (DE OLIVEIRA, BARRETO E SUGUIO, 1999; SANTOS, JC 2007; PESSENDA et al., 2010; MEDEIROS et al., 2018; DE MORAES, OLIVEIRA E BEHLING 2020), espeleotemas (AULER E SMART, 2001; WANG et al., 2004; CRUZ et al., 2009; NOVELLO et al., 2012; STRIKIS et al., 2015, 2018; UTIDA et al., 2020) e sedimentos marinhos (BEHLING et al., 2000; MARTINS et al., 2020; VENANCIO et al., 2020) indicam períodos úmidos na transição entre o Último Máximo Glacial (LGM) e o Holoceno.

O período entre 21 e 17 ky AP é conhecido como o Último Máximo Glacial (LGM), que representa o fim da última glaciação com redução de temperatura de  $-6,1^{\circ}\text{C}$  na média de todo globo (TIERNEY et al., 2020) em comparação com a média do presente. Para a América do Sul, isso significou o avanço de geleiras no sul próximo a Cordilheira dos Andes (HULTON et al., 2002), diminuição da umidade e áreas florestadas (CLAPPERTON, C.M. 1993) e uma redução

generalizada na temperatura do continente. Apesar disso, alguns autores indicam a possibilidade de mudanças climáticas mais regionais dentro do contexto do LGM (Sylvestre, F. 2009) com as áreas mais ao norte da América do Sul recebendo um maior aporte de umidade devido ao posicionamento da Zona de Convergência Intertropical (ZCIT) e um fortalecimento do Sistema de Monção da América do Sul.

Este posicionamento da ZCIT para sul é ainda mais bem documentado durante o evento conhecido como Heinrich Stadial 1 (HS1). Se estendendo entre 17 e 14,7 ky AP, este evento é marcado por um maior desprendimento de icebergs no Atlântico Norte, o que alterou a circulação oceânica e proporcionou mudanças na circulação atmosférica (STRIKIS et al., 2018). Este período possui grandes probabilidades de apresentar a maior flutuação climática, tendo em vista que representa o início do interglacial e as alterações verificadas em registros fósseis (CRUZ et al., 2009; STRIKIS et al., 2015; UTIDA et al., 2020) indicam que a circulação no SAB não teria seguido o padrão frio/seco indicado pela maior parte dos estudos que avaliam o final do período glacial.

Já o período conhecido como Bølling-Allerød (14,9 a 12,9 ky AP) é caracterizado por um aquecimento abrupto das condições atmosféricas, no qual a temperatura apresentou um aumento de até 3°-5°C em algumas regiões do globo (RASMUSSEN et al., 2006). As causas deste aquecimento são elencadas como alterações na circulação profunda do Atlântico Norte (SU, INGERSOLL E HE, 2016), sincronizado com um aumento abrupto na quantidade de CO<sub>2</sub> atmosférico e no nível do mar (KÖHLER et al., 2011). Para o SAB, esta mudança abrupta em um curto período de tempo (2 ky) pode ter representado uma grande alteração na cobertura vegetal que estava condicionada a maiores volumes de precipitação do HS1.

O breve aquecimento durante o período Bølling-Allerød foi seguido pelo Younger Dryas (12,9 a 11,7 ky AP), um período de redução das temperaturas que é contabilizado em registros em ambos os polos (RASMUSSEN et al., 2006; MASSON-DELMOTTE et al., 2010). Conforme é indicado por estudos prévios (MAYEWSKI et al., 1993), o estabelecimento e fim do Younger Dryas ocorre em um curto período de tempo, estando localizado entre dois períodos de aquecimento da atmosfera. Na América do Sul, esse resfriamento é registrado por dados palinológicos (MARKGRAF, V. 1993; MENDES, V.R. 2016, UTIDA et al., 2020), o que indica também um último período de maior umidade antes do final do Pleistoceno, com o resfriamento abrupto atuando como mais um filtro da vegetação do SAB.

O Holoceno (11,7 ky até o presente) é um período de constante aquecimento da atmosfera, caminhando para um período interglacial que registrará as maiores temperaturas em milhões de anos (MASSON-DELMOTTE et al., 2010). Para o SAB, as mudanças climáticas observadas durante este período contribuem para o estabelecimento de vegetações mais abertas e adaptadas as condições de aridez (MORAES, OLIVEIRA E BEHLING, 2020). Apesar

disso, a divisão entre Holoceno Inferior, Médio e Superior pode demonstrar mudanças climáticas regionais não observadas em estudos anteriores.

O contexto de flutuações climáticas intensas durante o Quaternário é refletido em um domínio importante para a vegetação, os solos do SAB. Em toda a região são encontradas áreas com solos extremamente profundos e com seus nutrientes removidos, herança de um clima mais úmido que acelerou os processos de formação do solo (SCHAEFER, C. E. G. R. 2013). A existência de solos com teores de carbonato menores que o esperado também indica que algumas regiões do SAB possuem estabelecimento de clima seco apenas nos últimos milhares de anos. Tal fator também privilegia a utilização das informações dos horizontes subsuperficiais, tendo em vista que eles preservam as informações pedogenéticas (LEPSCH, I.F. 2016) com maior fidelidade e são mais estáveis a alteração climática.

A relação entre solo e vegetação já está amplamente documentada por pesquisas realizadas no Brasil (ARRUDA et al., 2015, SILVA et al., 2015; FEITOSA et al., 2016, ARRUDA et al., 2020). O fornecimento de nutrientes, a disponibilidade hídrica e a profundidade ocupada pelas raízes são alguns fatores que privilegiam a ocupação de certas espécies nas diferentes classes de solo. No Sistema Brasileiro de Classificação de Solos (DOS SANTOS et al., 2018), horizontes de subsuperfície diagnósticos são usados para categorizar os níveis mais altos entre as classes de solo, refletindo os processos de formação e intemperismo. Embora o horizonte superficial esteja em contato direto com a vegetação, seu posicionamento em contato com fatores abióticos e bióticos pode resultar em sua maior degradação. Assim, o horizonte B pode refletir um registro mais confiável da formação da paisagem.

Ao inserir as informações edáficas nos estudos realizados, podemos acrescentar maior poder preditivo aos modelos e aumentar a escala do trabalho, tendo em vista que a variação do solo ocorre em escala local, enquanto a variação climática é regional. Alguns resultados (DE OLIVEIRA, A. R. 2016; VASCONCELOS, B. N. F. 2016; ARRUDA et al., 2017, OLIVEIRA et al., 2021) já mostram que a pedometria incorpora o poder descritivo da análise do solo e seus reflexos na paisagem. Sendo assim, um estudo que utilize covariáveis edáficas pode encontrar resultados superiores ao construir um nicho edafoclimático para cada formação vegetal, evitando extrapolações de ocorrência como ao modelar espécies (COLLEVATTI et al., 2013). Como a região do SAB apresenta fitofisionomias que podem ser semelhantes (IBGE, 2012), estabelecer a modelagem em nível regional com informações derivadas de solo, clima e vegetação certamente indicará novos resultados.

As reconstruções da paleovegetação realizadas por meio de modelos (ARRUDA et al., 2018, COSTA et al., 2018, MAKŠIĆ et al., 2019, ZULAR et al., 2019) apresentam divergências na distribuição da vegetação nos últimos 21 ka AP. As incertezas variam entre uma expansão da vegetação sazonal (Cerrados, Florestas Estacionais) durante o LGM e uma estabilidade nas áreas centrais de cada bioma, ambas proporcionadas pela alternância entre climas seco / quente

e frio / úmido. Um trabalho que privilegie uma escala temporal maior pode encontrar mudanças na cobertura vegetal ao analisar períodos entre o LGM e o presente que não foram analisados pelos estudos citados.

O objetivo deste estudo é: i) treinar modelos de aprendizado de máquina para prever a vegetação atual com base em dois conjuntos de dados, um usando as variáveis bioclimáticas e outro com variáveis bioclimáticas e de solo; ii) usar os modelos para prever a vegetação passada, testando a hipótese de que as mudanças climáticas do final do Pleistoceno – início do Holoceno (de 21 a 0,4 ka AP) e os solos da região explicam o estabelecimento da vegetação presente no SAB apenas em períodos recentes do Holoceno. Além disso, a ocorrência de períodos super-úmidos pode ter auxiliado no intemperismo dos solos da região, proporcionando a ocorrência de outras formações vegetacionais na área que hoje consideramos núcleo da Caatinga; iii) comparar os resultados da modelagem com registros fósseis da região que abarquem o mesmo período de estudo.

## **1.2 MATERIAL E MÉTODOS**

### **1.2.1 – Área de Estudo**

O Semiárido Brasileiro (SAB) é uma região delimitada por suas condições climáticas áridas (DA SILVA et al., 2010; MARENGO et al., 2011). São municípios com precipitação média anual de 800 mm ou menos; Índice de Aridez de Thornthwaite igual ou inferior a 0,50; porcentagem diária de déficit hídrico igual ou superior a 60% em todos os dias do ano. A delimitação de 2017 abarcava 1.127.952,744 km<sup>2</sup>, abrangendo municípios de todos os estados do Nordeste e do Norte de Minas Gerais (Figura 2).

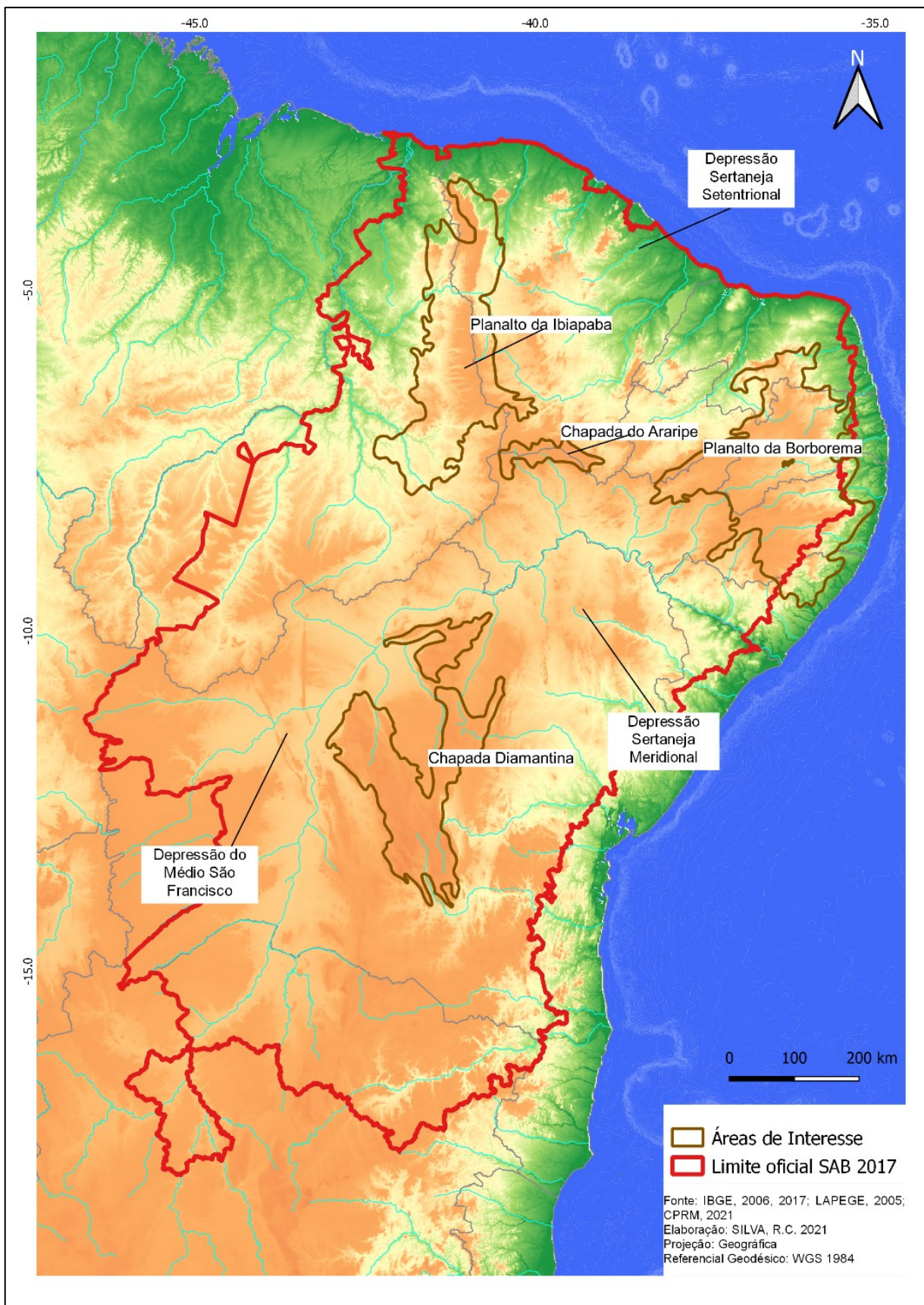


Figura 1. Mapa de relevo e áreas de interesse do SAB. Fonte: IBGE, 2006, 2017; LAPEGE, 2006; CPRM, 2021.

A litologia da região é geralmente do Pré-Cambriano, com relevo plano (figura 1) na área conhecida como Depressão Sertaneja, áreas elevadas como a Chapada Diamantina, Chapada do Araripe e Planalto da Borborema (AB'SÁBER, A. N. 2003). A maior parte do SAB é ocupada



por vegetação de Caatinga, com as porções sul e oeste apresentando contato ecotonal com o bioma Cerrado, enquanto as áreas a leste possuem contato com a Mata Atlântica.

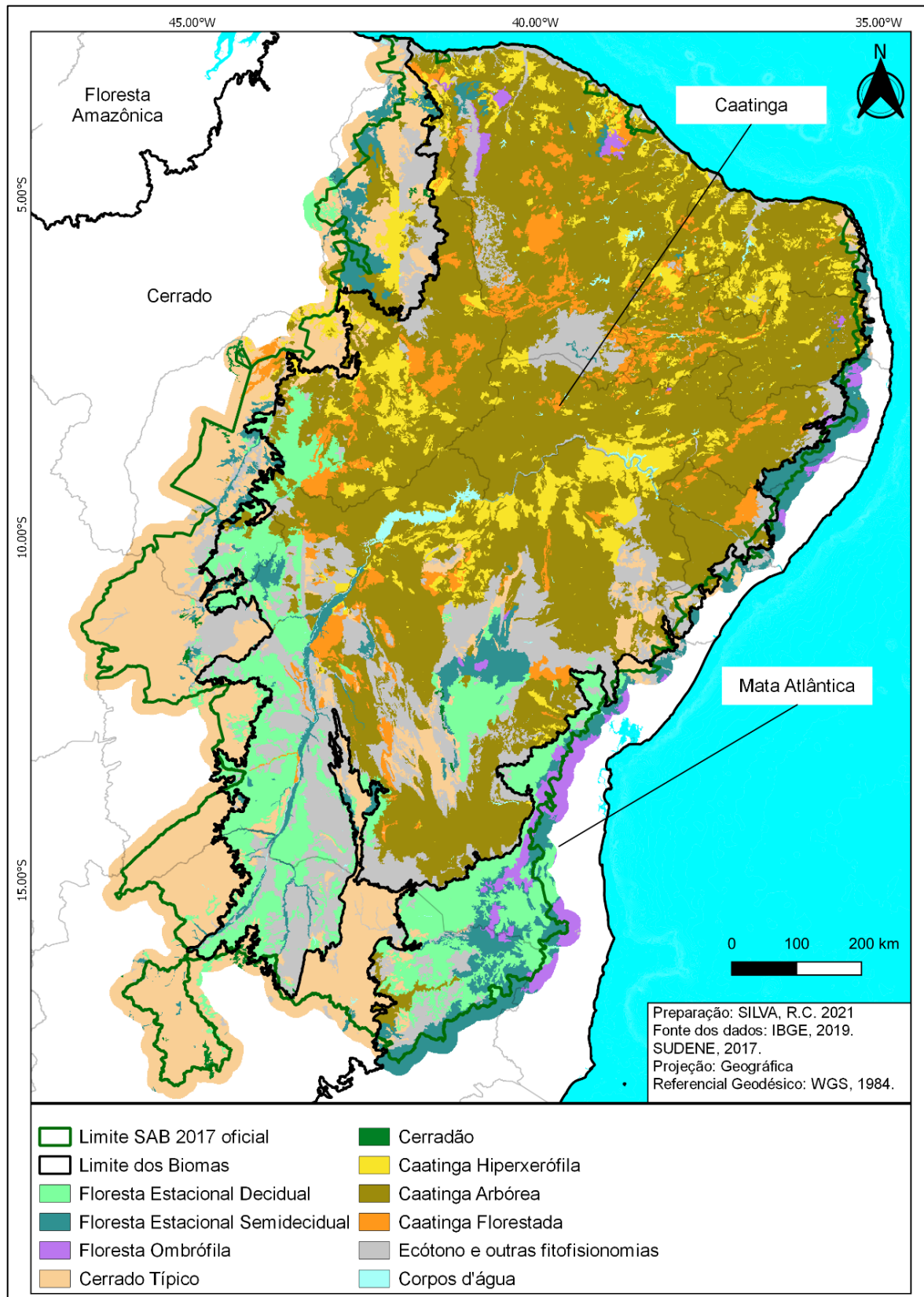


Figura 2. Localização e vegetação (Formações vegetacionais e Biomas) do Semiárido Brasileiro. Vegetação adaptada do tema “Vegetação” do BDiA- IBGE, 2019.

Os solos presentes no SAB (Figura 3) refletem seus fatores de formação. Na porção sul, ocorrem principalmente Latossolos e Argissolos, com maior profundidade, melhor drenagem e redução da fertilidade natural. Na porção norte, os solos apresentam maior diversidade, com grande presença de Luvisolos, Planossolos, Vertissolos e Neossolos Quartzarênicos. Essa região possui solos mais rasos, com maior fertilidade natural devido ao menor desgaste em sua formação. Nas áreas mais altas do SAB, podemos encontrar classes de solos menos frequentes, como Neossolos Regolíticos, Cambissolos e afloramentos rochosos.

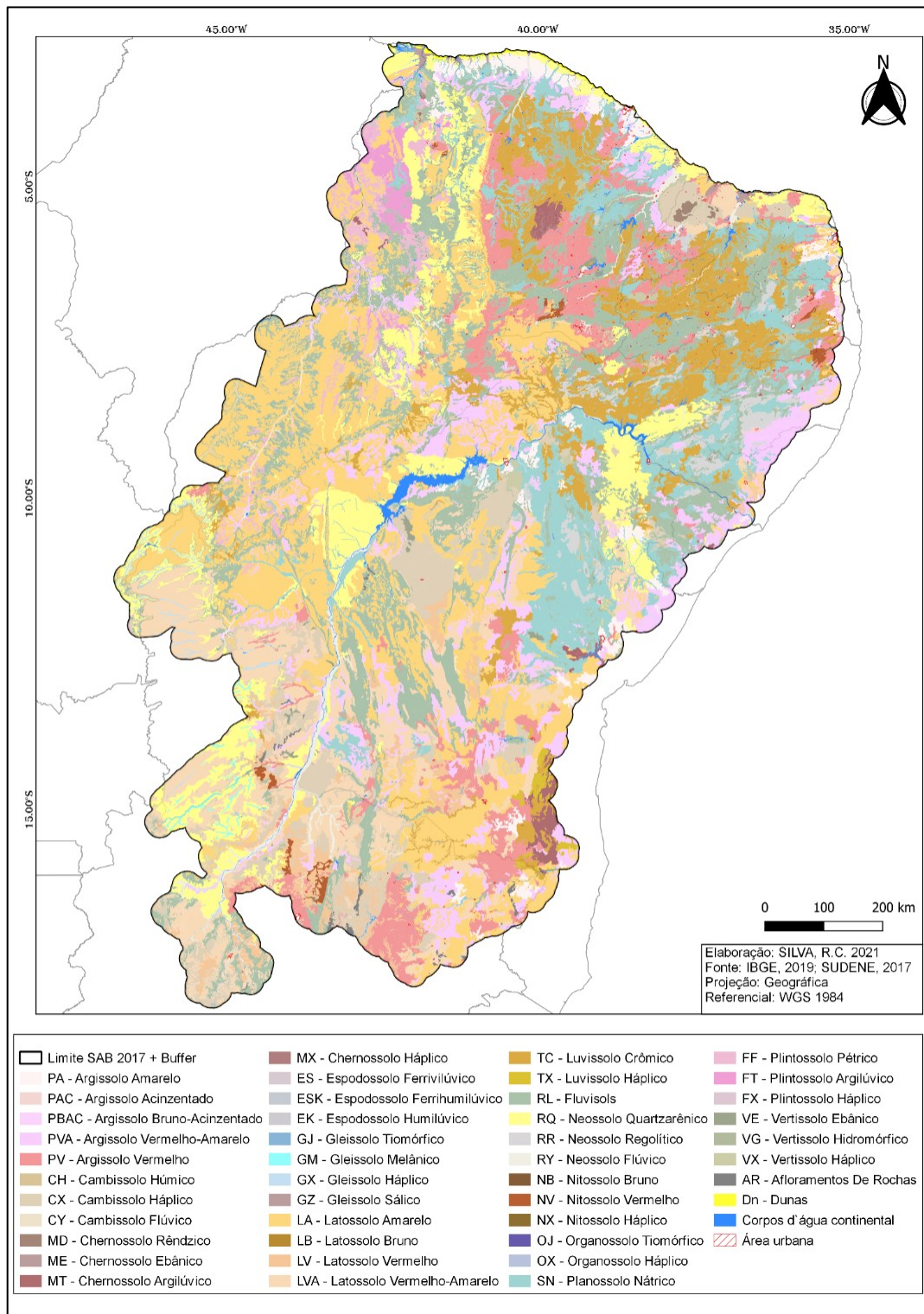


Figura 3. Solos do Semiárido Brasileiro. Adaptado do mapa do IBGE e RADAMBRASIL.

## 1.2.2 – Amostragem e processamento dos dados

As unidades de vegetação adotadas neste trabalho foram selecionadas por meio do tema “Vegetação” do Banco de Dados de Informações Ambientais (BDiA) do Instituto Brasileiro de Geografia e Estatística (IBGE, 2018). A escala dos polígonos de vegetação selecionados foi de 1: 250.000, adotando o conceito de Região Fitoecológica (IBGE, 2012). Posteriormente, foi criado um arquivo *raster* de classes de vegetação para o SAB com buffer de 20 km (BRASIL, 2017), sendo reamostrado para resolução de 1 km<sup>2</sup>. Este buffer é necessário para aumentar a área de amostragem e minimizar os efeitos de subamostragem de algumas classes de vegetação (Anexo 1.A). Foram consideradas apenas as regiões fitoecológicas (Anexo 2.A) mais representativas do espaço SAB, sendo agrupadas em formações vegetacionais para fins de classificação utilizando as definições de Floresta Estacional Decidual (FED), Floresta Estacional Semidecidual (FES), Floresta Ombrófila (FO), Cerrado Típico (ST), Cerradão (SF), Caatinga Hiperxerófila (CH), Caatinga Arbórea (CA) e Caatinga Florestada (CF), conforme a figura 4 ilustra.

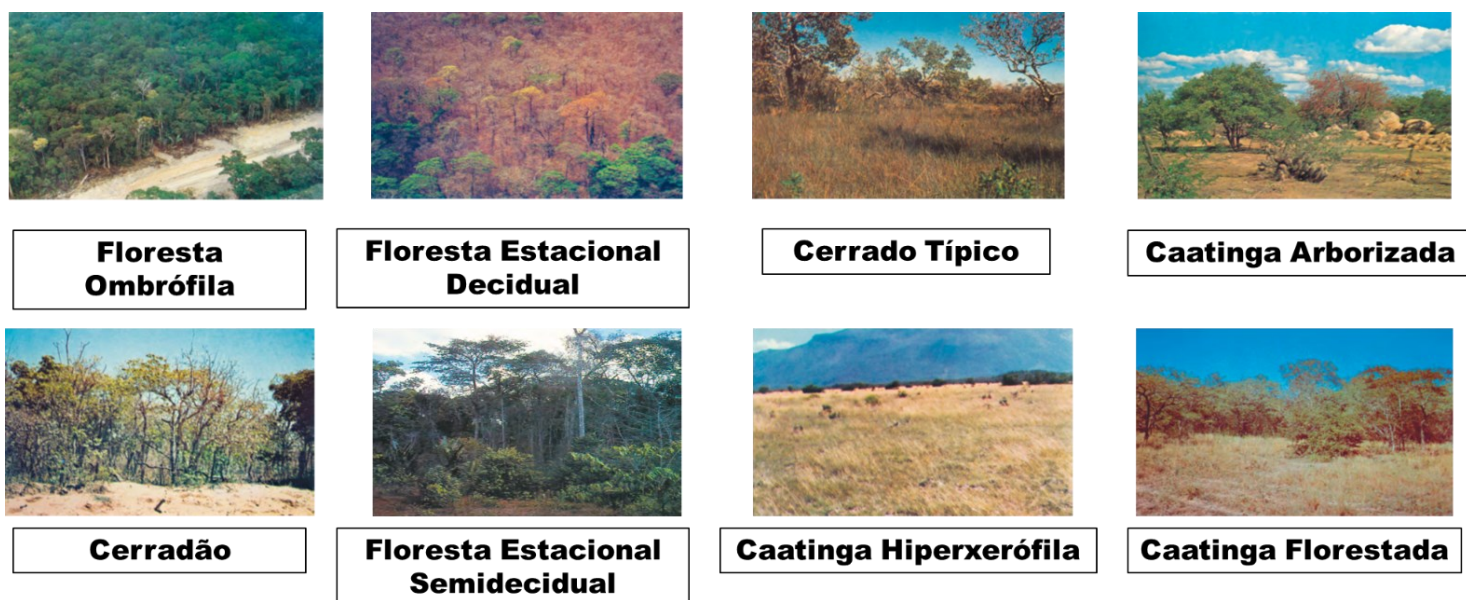


Figura 4. Fotografias que exemplificam as formações vegetacionais adotadas no estudo. Fonte: IBGE, 2012.

Os dados do solo foram obtidos de duas fontes diferentes. O mapeamento de solos também foi fornecido pelo BDiA (IBGE, 2018), na escala de 1: 250.000. Os polígonos (unidades de mapeamento) foram exportados com a classificação até o segundo nível categórico do SiBCS. Os perfis amostrados com análises químicas, físicas e mineralógicas foram obtidos no Banco de Dados de Solos da Embrapa (BDSolos) (SIMÕES, M.G. et al., 2017; EMBRAPA, 2019). Esses perfis foram selecionados de acordo com a disponibilidade de informações sobre horizontes subsuperficiais (qualquer horizonte que não fosse o superior) e proximidade espacial com o Semiárido Brasileiro.

A base de dados de perfis recebeu avaliação e tratamento estatístico. A média dos atributos dos horizontes subsuperficiais foi calculada para homogeneizar os perfis, criando assim um horizonte subsuperficial médio para representar cada perfil individualmente. Posteriormente, os dados do perfil receberam o preenchimento de ausências ajustado com um modelo de regressão com o algoritmo *randomForest*, do pacote *missForest* (STEKHOVEN, D. J. 2011) em R. Foram selecionadas para o estudo as variáveis que anteriormente possuíam valor máximo de 10% de ausência.

As variáveis de solo selecionadas foram areia grossa, areia fina, areia total, argila, silte, relação silte / argila, cálcio e magnésio ( $\text{Ca}^{2+} + \text{Mg}^{2+}$  cmolc/kg), alumínio trocável ( $\text{Al}^{3+}$  cmolc.kg<sup>-1</sup>), hidrogênio ( $\text{H}^+$  cmolc.kg<sup>-1</sup>), potássio ( $\text{K}^+$  cmolc.kg<sup>-1</sup>), saturação de alumínio ( $(100 \cdot \text{Al}^{3+} / \text{S} + \text{Al}^{3+})$  %, cmolc.kg<sup>-1</sup>), nitrogênio (N g/kg), sódio ( $\text{Na}^+$  g/kg), pH (H<sub>2</sub>O), pH (KCl), soma de bases ( $\text{Ca}^{2+} + \text{Mg}^{2+} + \text{K}^+ + \text{Na}^+$  - cmolc / kgS), valor T ( $\text{S} + \text{H}^+ + \text{Al}^{3+}$  cmolc / kg), saturação de bases ( $100 \cdot \text{S} / \text{T}$ %), carbono orgânico (CO - g / kg), razão carbono/nitrogênio (C/N %). Esta base de dados foi relacionada aos polígonos do mapa de solos, tomando o 1º nível categórico do Sistema Brasileiro de Classificação de Solos (Dos Santos et al., 2018) e a proximidade espacial como referência para junção. Finalmente, cada variável foi convertida para um *raster* de resolução de 1 km<sup>2</sup>. As variáveis topográficas vêm do modelo digital de elevação SRTM (*Shuttle Radar Topographic Mission*). As variáveis altitude e declive foram obtidas no site da AMBDATA (AMARAL et al., 2013, INPE, 2019), com resolução espacial de 1 km<sup>2</sup> e foram incluídas na base de dados de solos.

As variáveis climáticas foram obtidas no site WorldClim, em sua versão 2.1 referente aos anos 1970-2000 (FICK E HIJMANS, 2017). Os dados estão organizados em 19 variáveis bioclimáticas, oriundas da média, máxima, mínima temperatura e precipitação. Os dados têm resolução de 1 km<sup>2</sup> em formato matricial (*raster*).

### **1.2.3 – Modelagem e espacialização dos modelos de vegetação**

O arcabouço metodológico do processo de modelagem está contido no fluxograma (Figura 5). Os metodologia pode ser dividida em quatro fases: Extração/Pré-processamento dos dados, seleção de covariáveis, treinamento/espacialização dos modelos e Mapas e Resultados Finais.

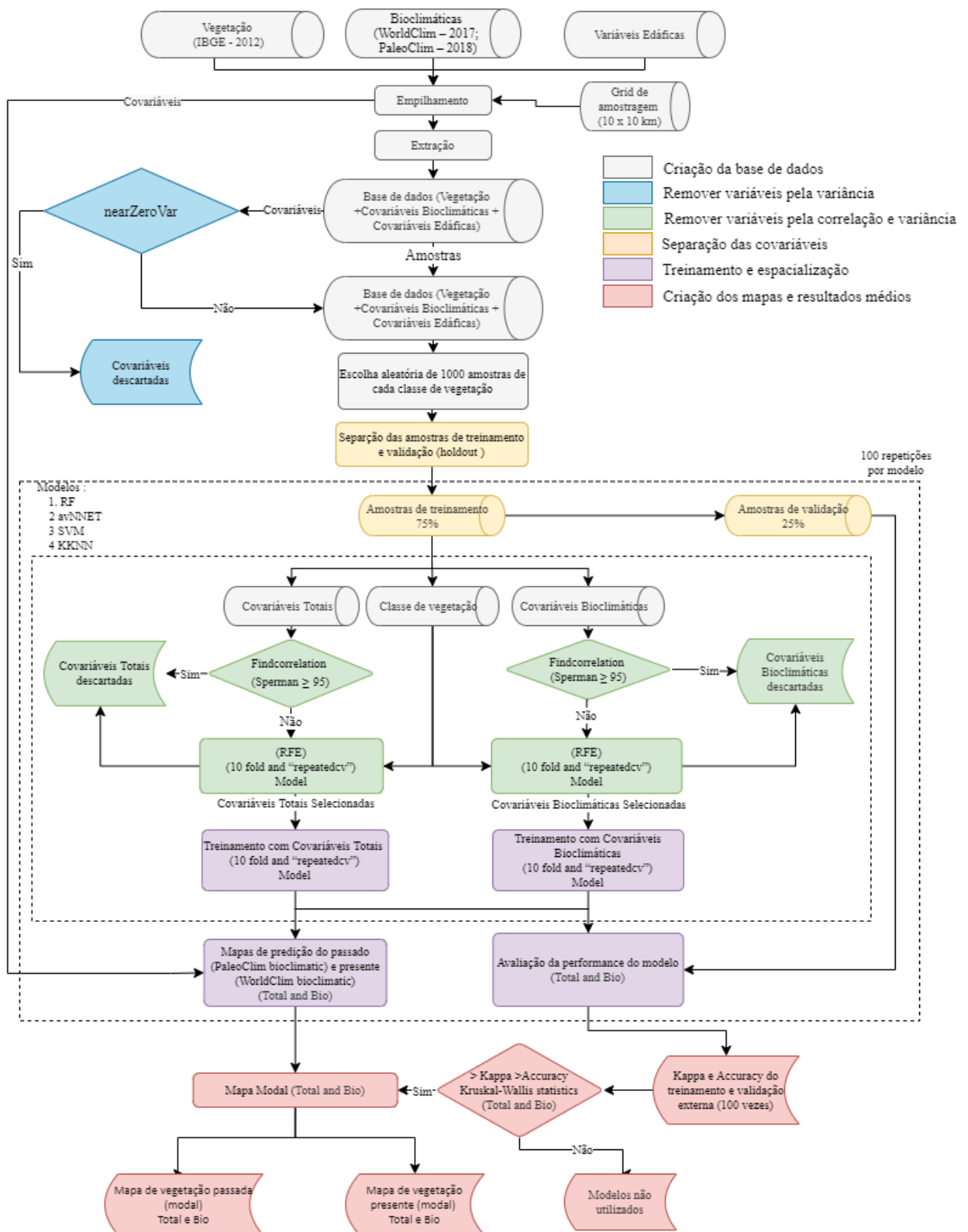


Figura 5. Fluxograma metodológico do estudo.

### 1.2.3.1. – Extração/Pré-processamento dos dados

Os conjuntos de covariáveis foram empilhados com o mapa de vegetação, sendo feita a extração dos dados usando um grid regular de distância 10 x 10 km. As covariáveis deste conjunto foram avaliadas em relação a variância, sendo retiradas as que apresentassem

variância próxima de zero; este processo é necessário para reduzir o custo computacional nos processos posteriores (seleção de covariáveis e treinamento do modelo), visto que as covariáveis que apresentam variância próxima a zero ou nula não apresentam importância na modelagem (KUHN, M. 2008). Após este processo foram escolhidas 1000 amostras de cada classe de vegetação, o qual compõem o conjunto base para processo de modelagem. O conjunto base foi separado o conjunto em treinamento (75%) e validação separação amostras (25%) (hold-out).

### **1.2.3.2. – Seleção de covariáveis**

A seleção de covariáveis é necessária visando a redução do custo computacional e aumento da parcimônia do sistema. O processo de seleção feito neste trabalho segue os princípios usados por Gomes et al. (2018), Souza et al. (2020) e Ferreira et al. (2021), sendo dividido em duas fases: seleção por correlação e importância. Estes processos tendem a serem utilizados em conjunto visto que aplicação somente da importância reduz o efeito de covariáveis altamente correlacionadas, mas o não elimina e impede o aparecimento de covariáveis causais (raras) impedido de detectá-las (DARST, MALECKI E ENGELMAN 2018).

Na primeira parte foi feita avaliação a correlação das covariáveis, usando a correlação de Spearman para avaliar as duplas que obtiveram correlação superior ou igual a 95%. A duplas acima desse valor foram avaliadas com relação a correlação com as outras covariáveis do trabalho, sendo eliminadas as que apresentaram maior valor da soma da correlação absoluta das outras covariáveis da entrada dessa fase. Este passo foi feito pela função `findCorrelation` do pacote `Caret` (KUHN, et al., 2008). As covariáveis restantes passaram pela segunda fase.

A remoção pela importância da covariável é feita através do algoritmo Recursive Feature Elimination – RFE contido no pacote `Caret` (KUHN, et al., 2008; KUHN et al., 2013). O RFE reduz o número dos preditores de forma iterativa, através da verificação do conjunto ideal de preditores que geram o melhor desempenho dos algoritmos testados com relação a sua performance (Cohen's Kappa – Kappa). O RFE foi rodado com conjunto total de variáveis (Anexo 3.A) e 16 subconjuntos de preditores: 5,6,7, 20 e 25 preditores. A seleção do subconjunto de preditores ótimo foi baseado na validação cruzada com 10 folds, uma repetição e 5 valores de cada um dos hiperparâmetros internos de cada algoritmo testados (`tuneLength`). Os hiperparâmetros de cada algoritmo estão descritos em manual do pacote `Caret` no capítulo 6. Os algoritmos que foram testados foram: Model Averaged Neural Network (`avNNet`), Random Forest (RF), Support Vector Machine Radial Sigma (SVM) e k-Nearest Neighbors (KNN). O subconjunto de preditores ótimo será usado para o treinamento de cada algoritmo testados.

### 1.2.3.3. – Treinamento e espacialização da vegetação atual e dos períodos paleoclimáticos

O treinamento dos modelos foi feito usando o subconjunto ideal gerado na fase de seleção por importância (RFE). A otimização dos parâmetros internos dos modelos no treinamento foi feita usando validação cruzada repetida (*repeatedcv*) com 10 folds, 10 repetições e 5 valores de cada uma dos hiperparâmetros internos de cada algoritmo testados (*tuneLength*). O resultado do treinamento foi usado para predição do mapa final da vegetação atual. Os mapas paleoclimáticos foram criados usando modelos paleoclimáticos Community Climate System Model 3.0 (COLLINS et al., 2006), sendo feito um *downscaling* para resolução igual a 1 km<sup>2</sup> (BROWN, J. L. et al., 2018). Os dados dos modelos paleoclimáticos substituíram os dados climáticos usados no modelo atual gerando mapas para 7 modelos.

Os mapas paleoclimáticos foram criados usando dados paleoclimáticos (Figura 6) obtidos no site PaleoClim, (BROWN et al., 2018). Eles são o resultado de dois experimentos: CHELSA e TRACE21ka (HE, FENG 2011; MEINSHAUSEN et al., 2011). Esses dados também estão na forma de variáveis bioclimáticas e foram reamostrados para a resolução de 1 km<sup>2</sup> neste estudo.

Período	Sigla	Nome do Período	Autor
4.2-0.3 ka	LH	Holoceno Inferior - <i>Late Holocene</i>	Fordham et al., 2017
8.32-4.2 ka	MH	Holoceno Médio - <i>Mid Holocene</i>	Fordham et al., 2017
11.7-8.32 ka	EH	Holoceno Superior - <i>Early Holocene</i>	Fordham et al., 2017
12.9-11.7 ka	YD	Younger Dryas - <i>Younger Dryas Stadial</i>	Fordham et al., 2017
14.7-12.9 ka	BA	Interstadial Bolling Allerod - <i>Bølling-Allerød</i>	Fordham et al., 2017
17.0-14.7 ka	HS1	Heinrich Stadial 1 - <i>Heinrich Stadial 1</i>	Fordham et al., 2017
21-17.0 ka	LGM	Último Máximo Glacial - <i>Last Glacial Maximum</i>	Karger et al., 2017

Figura 6: Reconstruções paleoclimáticas realizadas por modelos de circulação atmosférica global representando os sete períodos analisados no estudo. Fonte: Brown et al., 2018.

A predição da vegetação potencial nestes sete períodos foi realizada nos mesmos modelos usados pela vegetação atual, substituindo as variáveis bioclimáticas do WorldClim pelos experimentos do PaleoClim. Os resultados são comparados com levantamento bibliográfico de registros fósseis da região.

### 1.2.3.3. – Mapas finais e Resultados finais

O processo de seleção de variáveis, treinamento e predição dos mapas foi repetido 100 vezes com amostras de treinamento e validação diferentes. Este processo é importante para avaliar a variabilidade da predição, visto que estes subconjuntos devem gerar resultados diferentes e por consequência performance para cada modelo (KUHN E JOHNSON, 2013). Os

resultados finais performance foram calculadas pelas médias das 100 rodadas. Os mapas finais foram confeccionados calculando a modais (maior frequência entre as classes de vegetação para cada célula predita do grid) e mapas de frequência da moda. das 100 rodadas para cada pixel para os mapas presentes e mapas paleoclimáticos em cada algoritmo.

A escolha do melhor algoritmo foi feita usando parâmetros de performance como os índices Kappa (k) Eq. (1) e Acurácia Eq. (2) (CONGALTON, 1991). Os resultados das 100 rodadas foram testados usando Kruskal-Wallis e teste post-hoc de Dunn, havendo diferença estatísticas entre os algoritmos testados, os resultados dos dois melhores algoritmos foram avaliados comparando os mapas de vegetação presente e passada para suas características e com isso escolher o algoritmo mais adequado a cada situação.

$$kappa = \frac{n \sum_{i=1}^c n_{ii} - \sum_{i=1}^c n_{i+} + n_{+i}}{n^2 - \sum_{i=1}^c n_{i+} + n_{+i}} \quad (1)$$

Onde: K é a estimativa do Kappa;  $n_{ii}$  é o valor na linha  $i$  e coluna  $i$  (concordância observada);  $n_{i+}$  é a soma da linha  $i$ , e  $n_{+i}$  é a soma da coluna  $i$  da matriz de confusão (produto das marginais, sendo a concordância esperada);  $n$  é o número total de amostras; e  $C$  é o número total de classes.

$$Accuracy = \frac{\sum xi}{n} \times 100 \quad (2)$$

Onde: Accuracy é a exatidão global,  $xi$  representa o somatório de todos os elementos na diagonal da matriz de confusão e  $n$  é igual ao número total de amostras.

Todas as predições foram realizadas utilizando o conjunto de covariáveis edafoclimáticas e o conjunto com apenas covariáveis bioclimáticas, para verificar o efeito da adição das variáveis edáficas na performance dos modelos, qualidade da representação da vegetação no presente e possibilidade de uso na predição da vegetação do passado.

#### 1.2.4. – Softwares utilizados

As análises e gráficos gerados neste trabalho foram desenvolvidos com o software R versão 4.0.0 e suas bibliotecas. O processamento e modelagem de dos dados foram feitos em parceria com o Laboratório de Geoprocessamento da UFV (LABGEO-UFV) e auxílio do cluster da UFV (Universidade Federal de Viçosa).



## 1.3 RESULTADOS

### 1.3.1 – Performance dos Algoritmos

Dos quatro algoritmos selecionados, todos apresentaram Kappa > 0,6 e foram utilizados para as análises (Anexo 4.A.). Todos os modelos, exceto KNN, apresentaram ganho de Kappa e acurácia ao receber as variáveis edáficas (Tabela 1), evidenciando a necessidade de incluí-los na análise. Um comparativo entre as predições com covariáveis bioclimáticas e as predições com covariáveis bioclimáticas e edáficas foi realizado para verificação do efeito em larga escala (Anexo 5.A). A partir destes resultados, optamos por utilizar os modelos edafoclimáticos para a posterior análise.

Tabela 1. Performance dos modelos na predição das formações vegetacionais do SAB no presente (1970-2000) nas 100 repetições e validações.

Dataset	Modelo	Kappa	Kappa CV	Acurácia	Acurácia CV
Bioclim	avNNet	0.59	2.41	0.64	1.94
Bioclim+soil	avNNet	0.63	1.86	0.67	1.51
Bioclim	kkn	0.77	1.18	0.80	0.99
Bioclim+soil	kkn	0.76	1.21	0.79	1.02
Bioclim	rf	0.79	1.11	0.81	0.94
Bioclim+soil	rf	0.80	0.99	0.83	0.84
Bioclim	svmRadialSigma	0.72	1.32	0.75	1.10
Bioclim+soil	svmRadialSigma	0.73	1.29	0.76	1.08

Ao comparar os resultados dos modelos, fica evidente que o modelo RF tende a favorecer as variáveis bioclimáticas (Anexo 6.A). Das dez variáveis mais importantes, apenas uma não é bioclimática (declividade). Esse fator ajuda a prever distribuições em grande escala (Figura 7) em comparação com o mapa do IBGE.

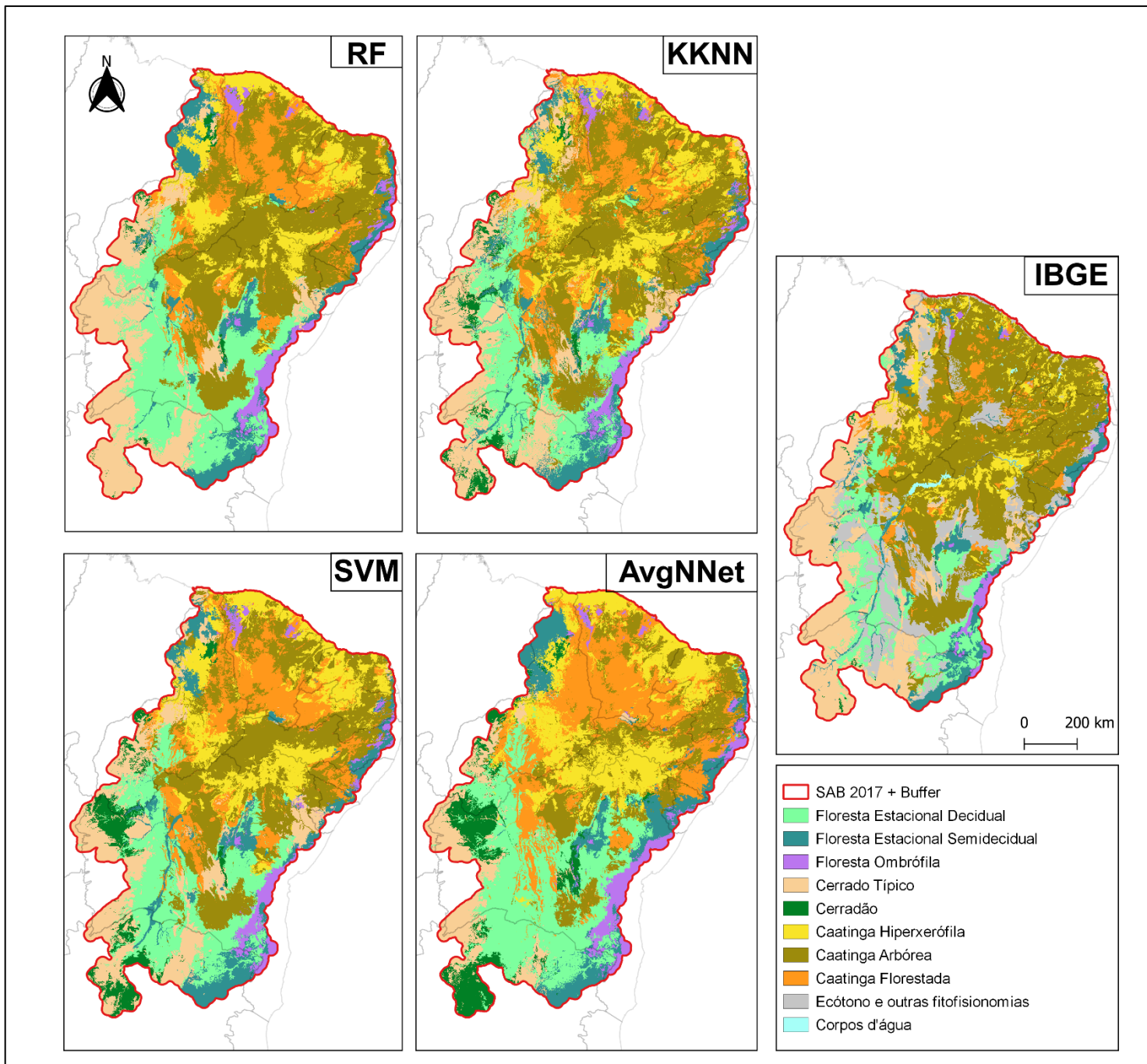


Figura 7. Formações vegetacionais potenciais referentes ao período atual (1970-2000) no Semiárido Brasileiro predita pelos modelos edafoclimáticos.

O modelo KKNN é o modelo que atua melhor na previsão das variações locais da formação vegetacional, privilegiando também as variáveis edáficas em importância (Anexo 6.A). Cinco variáveis edáficas estão entre as mais importantes para o modelo (declividade, altitude, pH em água, alumínio trocável e pH-KCl). Os modelos SVM e AvgNNet também apresentaram comportamento semelhante na importância das variáveis (Anexo 6.A).

Nossos resultados indicam que o modelo RF tende a favorecer as variáveis bioclimáticas, enquanto o modelo KKNN permite uma maior representação da variação edáfica em suas previsões. Essa diferença gera uma maior diversidade de formações vegetacionais na escala local no modelo KKNN, evidenciando que seu desempenho pode ser mais bem aproveitado para um trabalho desta escala. Todos os modelos foram usados para prever a vegetação passada e

presente. Através da análise do desempenho pelas métricas Kappa e Acurácia (Tabela 2), importância de variáveis (Anexo 6.A) e também pela comparação com o mapeamento do IBGE, optamos por utilizar o modelo KNN para análise da vegetação. As demais predições estão registradas nos anexos (Anexos 7.A a 13.A)

### **1.3.2 – Análise da vegetação potencial do presente**

As formações vegetacionais selecionadas para este estudo representam gradientes climáticos e edáficos de interesse. As três formações da Caatinga apresentam as maiores temperaturas médias do estudo (Anexo 14.A), além de apresentar os menores volumes anuais de precipitação. É interessante destacar que a CF apresenta menor sazonalidade e um volume de precipitação um pouco maior. Essas vegetações também estão posicionadas em solos mais arenosos (Anexo 14.A) e com menos carbono orgânico. Podemos concluir que a distinção entre as três fisionomias se deve ao maior volume de precipitação no período chuvoso, embora outros fatores podem estar presentes.

As duas formações de Cerrado selecionadas apresentam diferenças por fatores edáficos. No contexto do SAB, o SF possui área de cobertura menor, mas com menor variação na precipitação (Anexo 14.A). Apresenta também uma variação maior de argila e carbono orgânico, em linha com as observações de outros estudos. Outro gradiente relevante é aquele estabelecido entre a CF, FED e FES. As condições de precipitação entre os dois primeiros são evidentes (Anexo 14.A), enquanto o FES ocupa um nicho de maior disponibilidade hídrica. Os solos ocupados por FED e CF tendem a ser mais férteis e menos argilosos.

O modelo KNN que foi escolhido para análise reflete bem a variação regional das formações vegetacionais dentro de um mesmo bioma, atribuindo mais peso as covariáveis edáficas da região. Isso permite uma maior resolução nos resultados obtidos, preenchendo as áreas de ecótono indicadas no mapeamento do IBGE com a formação vegetacional que estaria mais adaptada a aquele nicho edafoclimático.

Dentro do contexto do SAB, o modelo apresenta uma área menor ocupada por CA (Anexo 15.A) em comparação com o mapeamento do IBGE, cerca de 10,7% a menos que o original. Parte dessa substituição é representada pelo aumento das áreas de CH e CF (+7,2% e +7,8% respectivamente, que também preenchem as áreas de ecótono no norte do SAB (Figura 7). Nas formações vegetacionais do Cerrado, as áreas de ST (-1,4% em comparação com IBGE) foram substituídas por SF (+2,5%) na porção oeste do SAB.

Nas áreas mais elevadas do SAB (Figura 1), a Chapada do Araripe é repleta de maior diversidade de formações (Figura 7), enquanto a Chapada Diamantina e o rio São Francisco apresentam áreas maiores com FES e FED; no Planalto da Ibiapaba, grande parte da área é preenchida com CF, com áreas de FO ao norte. De forma geral, os resultados indicaram que

todas as formações vegetacionais florestais (CF, SF, FED, FES e FO) apresentaram aumento na cobertura predita pelo modelo, ocupando as áreas de ecótono ou áreas de CA e ST.

### 1.3.3 – Reconstrução da vegetação do passado

Um resumo das variações de áreas comparadas com o período atual pode ser observado na Figura 9. Todos os períodos analisados por ambos os modelos estão representados na Figura 8. Uma reconstrução das anomalias climáticas da região analisada está registrada no anexo 16.A.

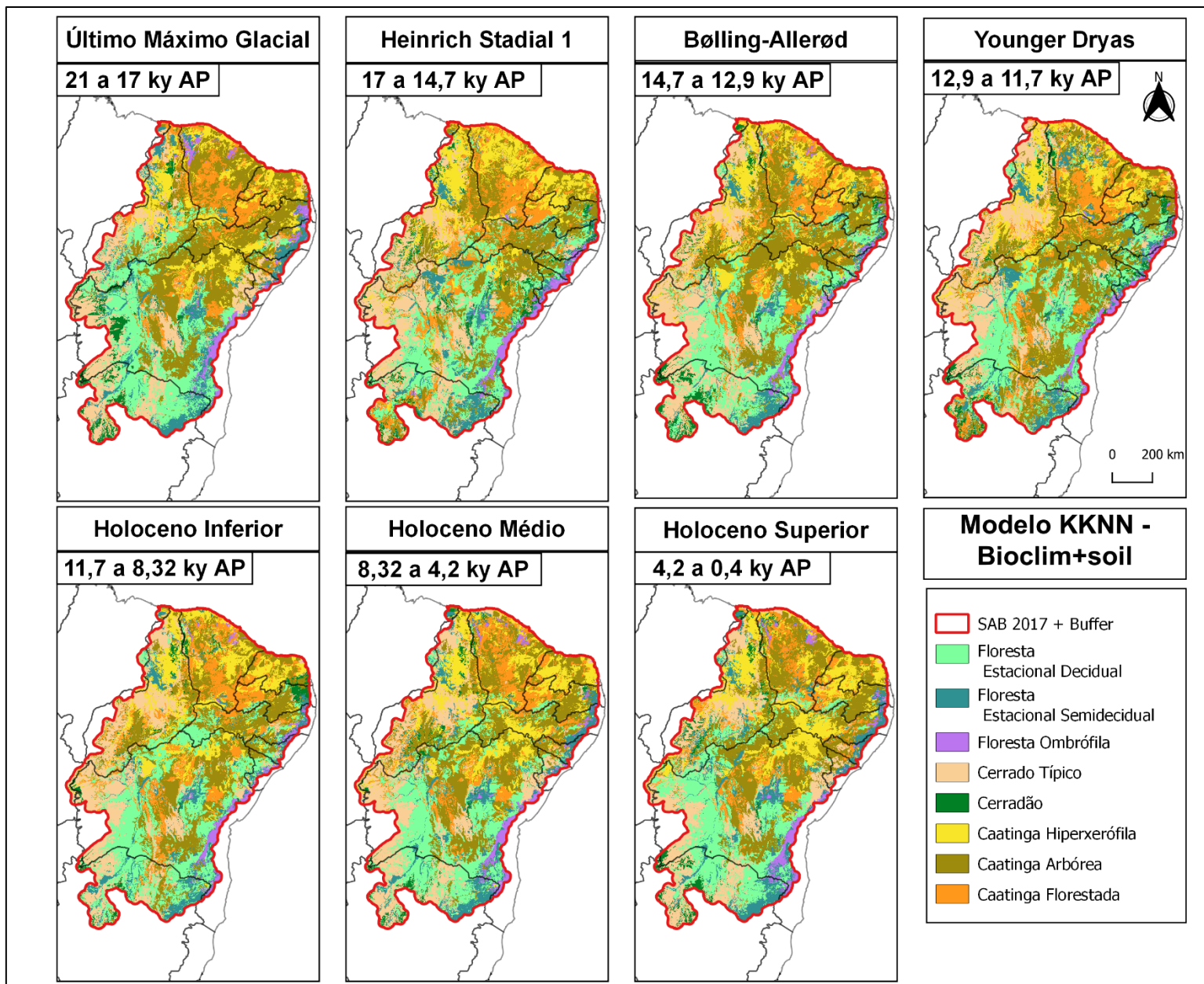


Figura 8. Reconstrução das formações vegetacionais no SAB durante o final do Pleistoceno e início do Holoceno.

Durante o UMG, as formações vegetacionais de Caatinga apresentaram expressiva redução no modelo KKNN, com a CH apresentando redução de -2,1%. No período seguinte, o HS1, as formações vegetacionais de CA e CF apresentam um aumento na área de +0,2 e +1%.

Nos períodos seguintes, é observado um aumento na cobertura de CH após o EH, seguindo a transição para um clima mais quente. As áreas de CF apresentam cobertura maior que a presente durante os períodos de maior umidade como o HS1 e o Younger Dryas, enquanto as áreas de CA apresentam cobertura inferior ao presente em todos os períodos.

As formações vegetacionais de Cerrado apresentaram um comportamento distinto. A formação vegetal ST está sempre representada por uma cobertura maior que a observada no presente, entre 0,2 e 2%. Já o SF apresenta cobertura maior que o atual em todos os períodos, exceto no EH e MH.

As áreas de FED apresentaram expansão em todos os períodos; o pico de expansão dessa formação vegetal é durante o EH, com um aumento de 5,9%. Já as áreas de FES foram subrepresentadas em quase todos os períodos. A maior retração verificada foi durante o período Bølling-Allerød (-2%), enquanto sua cobertura apresentou um aumento durante o LGM (+1,2%). As áreas de FO apresentaram retração em todos os períodos analisados, com seu pico durante o EH (-0,8%). Apesar disso, a magnitude das alterações dessa formação vegetal e da FES foi inferior as outras formações vegetacionais, principalmente se devendo a sua menor cobertura na área analisada.

Os dois maiores padrões observados nos resultados de predição são a cobertura inferior de CA em todos os períodos, enquanto as áreas de FED apresentam aumento. Os períodos em que estas mudanças são mais intensas são os períodos do HS1, EH e MH, indicando que períodos com maior alteração na precipitação e temperatura respectivamente são chave para promover a alteração na cobertura vegetal.

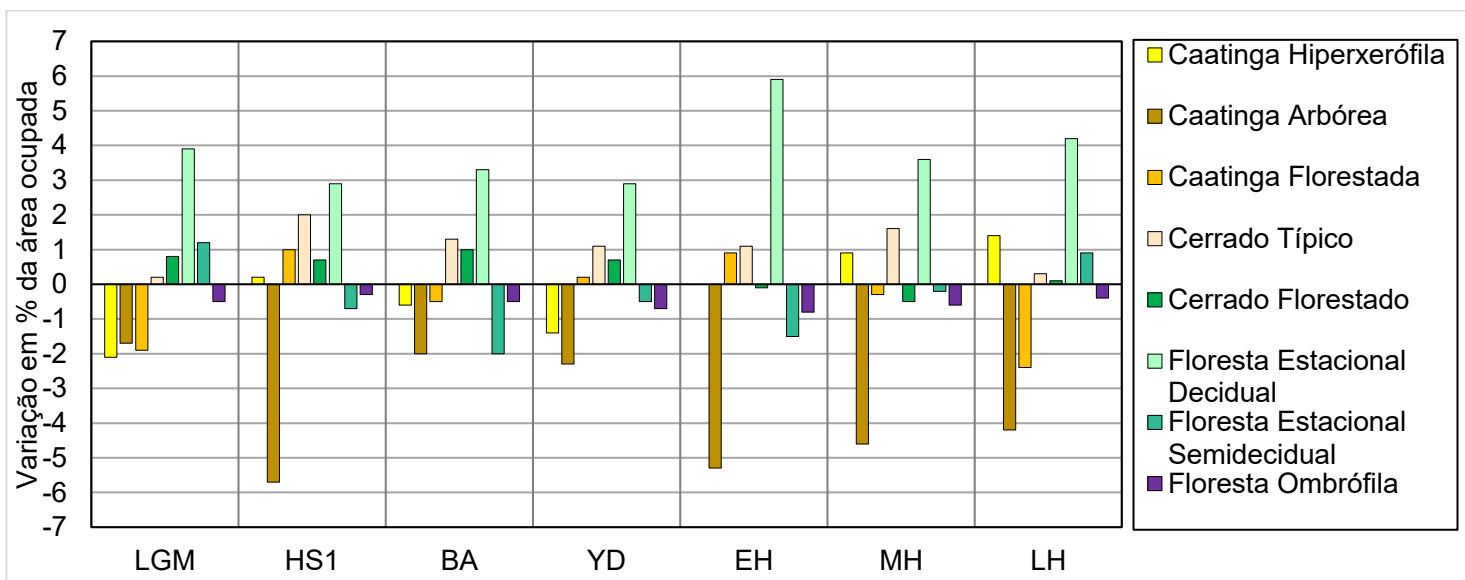


Figura 9. Variação em % das áreas de vegetação preditas para o passado (21 a 0,4 ky AP) em comparação com a predição presente (1970-2000).

## 1.4 DISCUSSÃO

De acordo com o objetivo i), os modelos apresentaram um desempenho superior quando adicionadas as variáveis edáficas. O ganho de acurácia foi observado em quase todos os

modelos, com a ressalva de que o modelo RF tende a privilegiar as variáveis bioclimáticas. O desempenho superior dos modelos de árvore de decisão já é documentado em outros trabalhos (ARRUDA et al., 2018), porém a importância das variáveis edáficas não é evidenciada nos resultados, o que é prejudicial em uma área tão diversa como o SAB. No contexto de aridez da região, espera-se que o material de origem tenha um efeito maior que os outros fatores de formação do solo (ARAÚJO FILHO, J. C. 2011), o que também nos levou a selecionar o modelo KNN para discussão dos resultados.

Um dos resultados observados foi a conversão das áreas de CA (Savana-estépica Arborizada no mapeamento original do IBGE) em CH (Savana-estépica, Savana-estépica Parque e Savana-estépica Gramíneo-lenhosa) e em CF (Savana-estépica Florestada). Uma das possíveis explicações está no fato das formações vegetacionais ocuparem nichos semelhantes, com sua diferenciação sendo estabelecida pela presença de estrato arbóreo esparso. Analisando em conjunto as três formações de Caatinga, não observamos a divisão característica entre as áreas cristalinas e sedimentares (QUEIROZ, L. P. 2006; MORO et al., 2016). Nossos resultados indicam que a melhor divisão entre as formações vegetacionais pode se dar pela distinção das áreas de CF como uma unidade separada, dada a semelhança observada entre CA e CH.

Ao analisar o modelo KNN, que dá mais importância para as variáveis edáficas, podemos obter uma maior diversidade local de vegetação. Este modelo indica áreas maiores de SF (Savana Florestada ou Cerradão), indicando que sua distinção das áreas de Cerrado pode se dar pelas variáveis edáficas. Por ocuparem nichos semelhantes em diferentes ambientes do SAB, este modelo também nos indicou que as formações de Caatinga e Florestas Estacionais estão relacionadas no contexto abiótico da região (ROSSATTO, CARVALHO E HARIDASAN, 2015). Isso vai de acordo com a teoria de que a Caatinga estaria mais relacionada com as florestas secas do que com outros biomas próximos como o Cerrado (FERNANDES E QUEIROZ, 2018).

Em relação ao objetivo ii) os modelos apresentaram que a diversidade estabelecida entre CA, CF e FED é mantida até o Holoceno Superior, indicando que o padrão atual de vegetação pode ter sido desenvolvido apenas nos últimos 0,4 ka. Estes resultados expandem os encontrados em estudos prévios (ARRUDA et al., 2018; COSTA et al., 2018; MAKSIĆ et al., 2019) ao aumentar a resolução espacial e temporal das reconstruções. Observamos também uma grande diminuição na área de FED no último período reconstruído, indicando que esta é uma formação vegetacional que pode ter tido sua área ocupada por outras formações mais adaptadas as mudanças climáticas como CA, CF e ST.

Podemos observar que o aumento de umidade privilegiou o avanço de FED ao sul da área atual do bioma da Caatinga que está inserida no SAB; Na região norte do SAB os modelos parecem privilegiar áreas maiores de CF e CA até mesmo em períodos mais secos. Estes resultados estão de acordo com registros palinológicos da região Nordeste, que indicam áreas

florestadas em períodos do final do Pleistoceno ao início do Holoceno (DE OLIVEIRA, BARRETO & SUGUIO, 1999; BEHLING et al., 2000; DE MEDEIROS et al., 2018).

Um dos resultados mais expressivos é o período HS1, no qual as anomalias de precipitação são intensas pelos resultados da reconstrução do modelo paleoclimático; essas alterações no regime de chuvas são observadas em registros fósseis da região (WANG et al., 2004; CRUZ et al., 2009; PESSENDA et al., 2010; STRIKIS et al., 2015; STRIKIS et al., 2018). A predição da vegetação mostra que grande parte da porção central do SAB apresentou uma mudança de vegetação de CA para CF, FED e FES; no norte do SAB as áreas de formações mais florestais são substituídas por CH, uma resposta ao deslocamento da umidade da ZCIT para sul. Como o registro parece indicar, este período parece ser o de maior alteração na circulação atmosférica, com alterações no sistema de monções que podem ter realmente promovido um grande favorecimento das formações florestais.

Além da influência na vegetação, eventos de grande alteração climática cumpriram um papel importante na formação dos solos do SAB. Enquanto os solos das áreas mais áridas da Depressão Sertaneja são intimamente relacionados com o material de origem e condições de drenagem (CÂMARA et al., 2021; PARAHYBA et al., 2021; DE SOUZA et al., 2020), os solos do sul do SAB estão em contato com as áreas de ecótono, com profundidades maiores sendo observadas (ARRUDA et al., 2015). Nestas áreas os solos mais férteis apresentam Florestas Estacionais, enquanto as áreas de Cerrado permanecem nos solos distróficos e profundos a oeste que foram originados de coberturas Cretáceas e Terciárias, conforme evidenciado pelos modelos.

No comparativo (Anexos 17.A e 18.A) realizado pelo objetivo iii), nossos resultados sobre o período do LGM são semelhantes aos encontrados na Toca da Boa Vista e no estudo comparativo envolvendo a Flona do Araripe, Parque Nacional das Sete Cidades e Reserva Biológica de Guaribas (AULER E SMART, 2001; PESSENDA et al., 2010), os quais indicam este período com maior umidade. Como são estudos de espeleotemas e datação de isótopos de carbono, não há um comparativo com a vegetação encontrada na região dos registros fósseis. O período do HS1 é representado na grande maioria dos registros fósseis da região Nordeste como um período de grande umidade, nos pontos 2, 3, 4, 5, 8, 9, 10, 12, 13, 14, 18, 19, 20 e 22 (Anexo. 17.A); o testemunho marinho GeoB 3104-1 (BEHLING et al., 2000) corrobora os resultados da modelagem ao indicar uma maior presença de vegetação florestal durante o HS1, indicando a sincronicidade entre o aumento da precipitação no SAB e as alterações na cobertura que permitiram a ocupação de formações vegetais como CF, FED e FES.

O período de Bølling-Allerød é marcado como um período seco nos espeleotemas de Rainha, Furna Nova, Abissal, Trapiá e Urubu (CRUZ et al., 2009; UTIDA et al., 2020); os demais registros analisados indicam um período úmido, enquanto o testemunho GeoB 3104-1 também indica uma maior cobertura florestal neste período. Apesar de uma retração nas áreas de CF e

FED nos resultados da modelagem referente a este período, a porção sul do SAB apresenta um clima mais úmido neste período, o que pode ter mantido a presença das florestas estacionais. A diminuição abrupta na precipitação e aumento na temperatura do Bølling-Allerød pode ter favorecido as vegetações abertas de CH e CA no norte do SAB conforme indicam nossos resultados.

O Younger Dryas é indicado por todos os registros fósseis analisados como um período úmido (Anexo 17.A.) Os resultados da modelagem são condizentes com a diminuição da temperatura e maior aporte de precipitação para o norte do SAB, aumentando a cobertura de CF e FES nessa região. O único registro de vegetação analisado, GeoB 3104-1 (BEHLING et al., 2000), indica a manutenção da vegetação florestal na porção norte do SAB desde o HS1 até o YD. Marcado como o final do Pleistoceno, esse último pulso de resfriamento é seguido pelo aumento da temperatura gradual, o que favorece a expansão da vegetação típica de CH e CA.

O Holoceno é um período com maior número de registros fósseis na região do SAB, marcadamente como um período mais quente e com diminuição da precipitação até as condições atuais. No período do Holoceno Inferior (EH) os registros fósseis do Vale do Rio Icatu (DE OLIVEIRA, BARRETO E SUGUIO, 1999) indicam uma maior presença de formações florestais, corroborando com os registros de Bom Jesus e Vale do Catimbau (DE MEDEIROS et al., 2018); O registro fóssil da foz do Rio Formoso (MARTINS et al., 2020) também marca este período como o início da ocupação por vegetação de restinga proveniente do aumento do nível do mar. Os resultados da modelagem são condizentes, com o aumento das formações de CH e redução das formações florestais de FES.

O período do Holoceno Médio é apresentado como o período mais quente da história recente anterior ao período Industrial. Os registros fósseis do Vale do Rio Icatu, GeoB 3104-1, Bom Jesus e Vale do Catimbau indicam um aumento na vegetação relacionada a Caatinga, corroborando nossos resultados de modelagem que indicam aumento na cobertura de CA e CH e uma redução na área de CF e FED. O período subsequente, o Holoceno Superior, é marcado pelo estabelecimento de vegetações adaptadas à aridez no registro fóssil da vegetação (pontos 1, 2, 5, 6, 16, 17, 21 no Anexo 18.A); os trabalhos referentes à Serra da Capivara e Vale do Catimbau são especialmente indicativos do domínio das espécies de CH e CA em relação as formações que possuem mais espécies arbóreas, o que reflete os resultados encontrados pela modelagem deste estudo.

A sequência de mudanças climáticas observadas por este estudo indica que a variação regional da precipitação e temperatura pode ter contribuído para o aumento da diversidade de espécies e formações vegetacionais encontradas na região do SAB. A adição de covariáveis edáficas permite observar novos resultados (OLIVEIRA et al., 2021) que não foram observados por estudos prévios que envolveram a modelagem da vegetação brasileira, indicando que as formações de CF e FED podem ocupar nichos muito próximos. O estabelecimento das condições



necessárias para o aumento da CH corrobora a hipótese de que essa formação da Caatinga pode ser originada em períodos recentes, sendo, portanto, distinta da CF (Oliveira et al., 2019).

Resultados recentes de estudos indicam que a influência antrópica nos solos pode ter se desenvolvido desde 10 kyr AP (SOUZA et al., 2020) na região da Caatinga; os primeiros registros da presença humana na região ainda são objetivo de debate entre pesquisadores (ALVIM, M. C. M 2008) que indicam datas anteriores a 12 kyr. De qualquer forma, estudos sobre o período mais recente indicam que cerca de 63% dos ambientes da Caatinga são ecossistemas antropogênicos (SILVA E BARBOSA, 2017). Como o Nordeste é a região de mais antiga ocupação pela exploração dos europeus, este estudo não descarta a possibilidade de uma interferência antrópica recente (0,4 ky AP até o período atual) capaz de alterar as formações mais florestais como a CF e CA para CH. Além disso, as áreas mais úmidas como FED, FES e FO foram as que sofreram maior alteração, também já documentada. Essas condições mantêm o debate sobre a conservação de grandes áreas incluídas no SAB que podem ser protegidas pela Lei da Mata Atlântica (BRASIL, 2006).

## **1.5 CONCLUSÕES**

O aumento da disponibilidade de dados ambientais permite a atualização de diversos estudos e teorias desenvolvidos sobre a distribuição da paleovegetação na América do Sul. Ao utilizar preditores climáticos e edáficos, conseguimos desenvolver modelos de predição com maior acurácia e melhor representação local da vegetação do presente na região do Semiárido brasileiro (SAB). Nossos resultados indicam uma divisão entre as fitofisionomias de Caatinga Hiperxerófila, Caatinga Arbórea e Caatinga Florestada, com grande semelhança de nicho ambiental com as Florestas Estacionais Deciduais. Com estes achados indicamos que as formações com maior estrato florestal na Caatinga são áreas mais semelhantes as Florestas Estacionais Deciduais de outros biomas.

As mudanças climáticas observadas no estudo podem ser uma das causas da grande diversidade de formações e espécies observada nessa região que é o encontro de três biomas. Ao reconstruir a vegetação do passado entre o final do Pleistoceno e o Holoceno, constatamos que o padrão de vegetação atual do SAB se estabeleceu entre 8,3 e 0,3 ka AP; durante os períodos de maior umidade ocorridos entre o Heinrich Stadial 1 e o Holoceno Inferior houve aumento na cobertura de Florestas Estacionais e Caatinga Florestada. Com a maior importância proporcionada para as variáveis edáficas no treinamento do modelo, nossos resultados também indicam que o estabelecimento da vegetação mais aberta da Caatinga pode ter sido extremamente recente, com a substituição de áreas de Caatinga Arbórea, Caatinga Florestada e Florestas Estacionais através da influência antrópica, não apenas por mudanças climáticas.

## 1.6 REFERÊNCIAS

AB'SÁBER, Aziz Nacib. **Os domínios de natureza no Brasil: potencialidades paisagísticas**. Ateliê editorial, 2003.

AMARAL, Silvana et al., AMBDATA: variáveis ambientais para Modelos de Distribuição de Espécies (MDEs). **Anais do XVI Simpósio Brasileiro de Sensoriamento Remoto**, v. 16, p. 6930-6937, 2013. Disponível em:  
<http://marte2.sid.inpe.br/attachment.cgi/dpi.inpe.br/marte2/2013/05.29.00.11.46/doc/p0944.pdf>

DE ARAUJO FILHO, J. C. Relação solo e paisagem no Bioma Caatinga. In: **Embrapa Solos- Artigo em anais de congresso (ALICE)**. In: SIMPÓSIO BRASILEIRO DE GEOGRAFIA FÍSICA APLICADA, 14., 2011, Dourados." Dinâmicas socioambientais das inter-relações às interdependências". Dourados: UFGD, 2011. Disponível em:  
<https://www.alice.cnptia.embrapa.br/bitstream/doc/896995/1/CoelhoXIVSBGFA1.pdf>

ARRUDA, Daniel M. et al., Landforms and soil attributes determine the vegetation structure in the Brazilian semiarid. **Folia Geobotanica**, v. 50, n. 3, p. 175-184, 2015.  
<https://doi.org/10.1007/s12224-015-9221-0>

ARRUDA, Daniel M. et al., Combining climatic and soil properties better predicts covers of Brazilian biomes. **The Science of Nature**, v. 104, n. 3-4, p. 32, 2017.  
<https://doi.org/10.1007/s00114-017-1456-6>

ARRUDA, Daniel M. et al., Vegetation cover of Brazil in the last 21 ka: New insights into the Amazonian refugia and Pleistocenic arc hypotheses. **Global Ecology and Biogeography**, v. 27, n. 1, p. 47-56, 2018. <https://doi.org/10.1111/geb.12646>

ARRUDA, Daniel M. et al., Soil and climate equally contribute to changes in the species compositions of Brazilian dry forests across 300 km. **Journal of Plant Ecology**, v. 13, n. 2, p. 171-176, 2020. <https://doi.org/10.1093/jpe/rtz059>

AULER, Augusto S.; SMART, Peter L. Late Quaternary paleoclimate in semiarid northeastern Brazil from U-series dating of travertine and water-table speleothems. **Quaternary Research**, v. 55, n. 2, p. 159-167, 2001. <https://doi.org/10.1006/qres.2000.2213>

BEHLING, Hermann et al., Late Quaternary vegetational and climate dynamics in northeastern Brazil, inferences from marine core GeoB 3104-1. **Quaternary Science Reviews**, v. 19, n. 10, p. 981-994, 2000. [https://doi.org/10.1016/S0277-3791\(99\)00046-3](https://doi.org/10.1016/S0277-3791(99)00046-3)

BRASIL, 2006. LEI Nº. 11.428, DE 22 DE DEZEMBRO DE 2006. **Dispõe sobre a utilização e proteção da vegetação nativa do Bioma Mata Atlântica, e dá outras providências.** Disponível em< [http://www.planalto.gov.br/ccivil\\_03/\\_ato2004-2006/2006/lei/l11428.htm](http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2006/lei/l11428.htm), v. 20, 2017.

BRASIL, 2017. Semiárido Brasileiro. Resolução nº 115, de 23 de novembro de 2017 do Conselho Deliberativo da Superintendência do Desenvolvimento do Nordeste. Disponível em: [https://www.in.gov.br/materia/-/asset\\_publisher/Kujrw0TZC2Mb/content/id/739568/do1-2017-12-05-resolucao-n-115-de-23-de-novembro-de-2017-739564](https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/739568/do1-2017-12-05-resolucao-n-115-de-23-de-novembro-de-2017-739564)

BROWN, Jason L. et al., PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. **Scientific data**, v. 5, n. 1, p. 1-9, 2018. <https://doi.org/10.1038/sdata.2018.254>

CÂMARA, Elis Regina Guimarães et al., Parent rock–pedogenesis relationship: How the weathering of metamorphic rocks influences the genesis of Planosols and Luvisols under a semiarid climate in NE Brazil. **Geoderma**, v. 385, p. 114878. <https://doi.org/10.1016/j.geoderma.2020.114878>

CLAPPERTON, Chalmers M. Nature of environmental changes in South America at the Last Glacial Maximum. **Palaeogeography, Palaeoclimatology, Palaeoecology**, v. 101, n. 3-4, p. 189-208, 1993. [https://doi.org/10.1016/0031-0182\(93\)90012-8](https://doi.org/10.1016/0031-0182(93)90012-8)

COLLEVATTI, Rosane G. et al., Drawbacks to palaeodistribution modelling: the case of South American seasonally dry forests. **Journal of Biogeography**, v. 40, n. 2, p. 345-358, 2013. <https://doi.org/10.1111/jbi.12005>

COLLINS, William D. et al., The community climate system model version 3 (CCSM3). **Journal of Climate**, v. 19, n. 11, p. 2122-2143, 2006. <https://doi.org/10.1175/JCLI3761.1>

CONGALTON, R. G. A review of assessing the accuracy of classifications of remotely sensed data. **Remote Sensing of Environment**, v. 37, n. 1, p. 35–46, 1991. [https://doi.org/10.1016/0034-4257\(91\)90048-B](https://doi.org/10.1016/0034-4257(91)90048-B)

COSTA, Gabriel C. et al., Biome stability in South America over the last 30 kyr: Inferences from long-term vegetation dynamics and habitat modelling. **Global Ecology and Biogeography**, v. 27, n. 3, p. 285-297, 2018. <https://doi.org/10.1111/geb.12694>

CRUZ, Francisco W. et al., Orbitally driven east–west antiphasing of South American precipitation. **Nature Geoscience**, v. 2, n. 3, p. 210-214, 2009. <https://doi.org/10.1038/ngeo444>

- DA SILVA et al., Caracterização do Semiárido brasileiro: fatores naturais e humanos. **Embrapa Semiárido-Capítulo em livro científico (ALICE)**, 2010. Disponível em:  
<http://www.alice.cnptia.embrapa.br/alice/handle/doc/861906>
- DARST, Burcu F.; MALECKI, Kristen C.; ENGELMAN, Corinne D. Using recursive feature elimination in random forest to account for correlated variables in high dimensional data. **BMC genetics**, v. 19, n. 1, p. 65, 2018. <https://doi.org/10.1186/s12863-018-0633-8>
- DE MORAES, Caio Alves; DE OLIVEIRA, Marcelo AT; BEHLING, Hermann. Late Holocene climate dynamics and human impact inferred from vegetation and fire history of the Caatinga, in Northeast Brazil. **Review of Palaeobotany and Palynology**, v. 282, p. 104299, 2020. <https://doi.org/10.1016/j.revpalbo.2020.104299>
- DE OLIVEIRA, Amanda Ribeiro. Relações solo-paisagem e modelo de predição de solos da bacia do córrego Marinheiro, Sete Lagoas (MG). 2016. Disponível em:  
<http://hdl.handle.net/1843/IGCC-AYEHAQ>
- DE OLIVEIRA, Paulo E.; BARRETO, Alcina Magnólia Franca; SUGUIO, Kenitiro. Late Pleistocene/Holocene climatic and vegetational history of the Brazilian caatinga: the fossil dunes of the middle São Francisco River. **Palaeogeography, palaeoclimatology, palaeoecology**, v. 152, n. 3-4, p. 319-337, 1999. [https://doi.org/10.1016/S0031-0182\(99\)00061-9](https://doi.org/10.1016/S0031-0182(99)00061-9)
- DE SOUZA, José João Lélis Leal et al., Archaeoanthrosol formation in the Brazilian semiarid. **CATENA**, v. 193, p. 104603, 2020. <https://doi.org/10.1016/j.catena.2020.104603>
- DOS SANTOS, Humberto Gonçalves et al., **Sistema brasileiro de classificação de solos**. Brasília, DF: Embrapa, 2018.
- EMBRAPA, 2019. BD – Solos, Banco de dados de solos brasileiros. Disponível em:  
[https://www.bdsolos.cnptia.embrapa.br/consulta\\_publica.html](https://www.bdsolos.cnptia.embrapa.br/consulta_publica.html). Acesso em: 20 de novembro de 2020.
- FEITOSA, Keily Katiany Almeida et al., Relações solo-vegetação em "ilhas" florestais e savanas adjacentes, no nordeste de Roraima. **Ciência Florestal**, v. 26, n. 1, p. 135-146, 2016. <https://doi.org/10.5902/1980509821098>
- FERREIRA, Renan Gon et al., Machine learning models for streamflow regionalization in a tropical watershed. **Journal of Environmental Management**, v. 280, p. 111713, 2021. <https://doi.org/10.1016/j.jenvman.2020.111713>

- FICK, Stephen E.; HIJMANS, Robert 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. <https://doi.org/10.1002/joc.5086>
- FORDHAM, Damien A. et al., PaleoView: a tool for generating continuous climate projections spanning the last 21 000 years at regional and global scales. **Ecography**, v. 40, n. 11, p. 1348-1358, 2017. <https://doi.org/10.1111/ecog.03031>
- GOMES, Lucas Carvalho et al., Modelling and mapping soil organic carbon stocks in Brazil. **Geoderma**, v. 340, p. 337-350, 2019. <https://doi.org/10.1016/j.geoderma.2019.01.007>
- GRUNWALD, Sabine (Ed.). **Environmental soil-landscape modeling: Geographic information technologies and pedometrics**. CRC Press, 2016.
- HE, Feng. **Simulating transient climate evolution of the last deglaciation with CCSM 3**. 2011. Disponível em: [https://www.researchgate.net/profile/Feng-He-27/publication/263618839\\_Simulating\\_transient\\_climate\\_evolution\\_of\\_the\\_last\\_deglaciation\\_with\\_CCSM3/links/0c96053b5fc619188f000000/Simulating-transient-climate-evolution-of-the-last-deglaciation-with-CCSM3.pdf](https://www.researchgate.net/profile/Feng-He-27/publication/263618839_Simulating_transient_climate_evolution_of_the_last_deglaciation_with_CCSM3/links/0c96053b5fc619188f000000/Simulating-transient-climate-evolution-of-the-last-deglaciation-with-CCSM3.pdf)
- HULTON, Nicholas RJ et al., The last glacial maximum and deglaciation in southern South America. **Quaternary Science Reviews**, v. 21, n. 1-3, p. 233-241, 2002. [https://doi.org/10.1016/S0277-3791\(01\)00103-2](https://doi.org/10.1016/S0277-3791(01)00103-2)
- IBGE. Manual técnico da vegetação brasileira. 2012.
- IBGE, 2018. Banco de Dados de Informações Ambientais - BDIA. Disponível em: <https://bdiaweb.ibge.gov.br/#/home>. Acesso em: nov. 2020.
- INPE, 2019. AMBDATA. Disponível em: <http://www.dpi.inpe.br/Ambdata/referencias.php>. Acesso em: nov. 2020
- KARGER, Dirk Nikolaus et al., Climatologies at high resolution for the earth's land surface areas. **Scientific data**, v. 4, n. 1, p. 1-20, 2017. <https://doi.org/10.1038/sdata.2017.122>
- KÖHLER, Peter et al., Abrupt rise in atmospheric CO<sub>2</sub> at the onset of the Bølling/Allerød: in-situ ice core data versus true atmospheric signals. **Climate of the Past**, v. 7, n. 2, p. 473-486, 2011. <https://doi.org/10.5194/cp-7-473-2011>
- KUHN, Max et al., Building predictive models in R using the caret package. **J Stat Softw**, v. 28, n. 5, p. 1-26, 2008. Disponível em: <http://www.math.chalmers.se/Stat/Grundutb/GU/MSA220/S18/caret-JSS.pdf>

KUHN, Max et al., **Applied predictive modeling**. New York: Springer, 2013.

KUHN, Max. Caret: classification and regression training. **ascl**, p. ascl: 1505.003, 2015.

Disponível em: <https://ui.adsabs.harvard.edu/abs/2015ascl.soft05003K/abstract>

MAKSIC, Jelena et al., Simulation of the Holocene climate over South America and impacts on the vegetation. **The Holocene**, v. 29, n. 2, p. 287-299, 2019.

<https://doi.org/10.1177%2F0959683618810406>

MARENGO et al., Variabilidade e mudanças climáticas no semiárido brasileiro. **Recursos hídricos em regiões áridas e semiáridas**, v. 1, p. 385-422, 2011.

<https://doi.org/10.1177%2F0959683618810406>

MARKGRAF, Vera. Younger Dryas in southernmost South America—an update. **Quaternary Science Reviews**, v. 12, n. 5, p. 351-355, 1993. [https://doi.org/10.1016/0277-3791\(93\)90043-L](https://doi.org/10.1016/0277-3791(93)90043-L)

MARTINS, Silvio Eduardo Matos et al., Holocene vegetation changes according to sea-level and climate dynamics on tidal flats of the Formoso River estuary, northeastern Brazil. **Quaternary International**, 2020. <https://doi.org/10.1016/j.quaint.2020.07.032>

MASSON-DELMOTTE, Valérie et al., EPICA Dome C record of glacial and interglacial intensities. **Quaternary Science Reviews**, v. 29, n. 1-2, p. 113-128, 2010.

<https://doi.org/10.1016/j.quascirev.2009.09.030>

MAYEWSKI, Paul A. et al., The atmosphere during the Younger Dryas. **Science**, v. 261, n. 5118, p. 195-197, 1993. <https://doi.org/10.1126/science.261.5118.195>

MAYLE, Francis E. Assessment of the Neotropical dry forest refugia hypothesis in the light of palaeoecological data and vegetation model simulations. **Journal of Quaternary Science**, v. 19, n. 7, p. 713-720, 2004. <https://doi.org/10.1002/jqs.887>

MEDEIROS, VANDA B. et al., New holocene pollen records from the Brazilian Caatinga. **Anais da Academia Brasileira de Ciências**, v. 90, n. 2, p. 2011-2023, 2018.

<https://doi.org/10.1590/0001-3765201820170161>

MEINSHAUSEN, M. et al., The paleoclimate modeling intercomparison project contribution to CMIP5. **WCRP Coupled Model Intercomparison Project-Phase 5-CMIP5**, p. 15, 2011.

Disponível: [https://ensembles-eu.metoffice.gov.uk/cmug/CLIVAR\\_Exchange.pdf#page=15](https://ensembles-eu.metoffice.gov.uk/cmug/CLIVAR_Exchange.pdf#page=15)

MORO, Marcelo Freire et al., A phytogeographical metaanalysis of the semiarid Caatinga domain in Brazil. **The Botanical Review**, v. 82, n. 2, p. 91-148, 2016.

<https://doi.org/10.1007/s12229-016-9164-z>

NOVELLO, Valdir F. et al., Multidecadal climate variability in Brazil's Nordeste during the last 3000 years based on speleothem isotope records. **Geophysical Research Letters**, v. 39, n. 23, 2012. <https://doi.org/10.1029/2012GL053936>

OBBINK, Elizabeth A.; CARLSON, Anders E.; KLINKHAMMER, Gary P. Eastern North American freshwater discharge during the Bølling-Allerød warm periods. **Geology**, v. 38, n. 2, p. 171-174, 2010. <https://doi.org/10.1130/G30389.1>

OLIVEIRA, Guilherme de Castro et al., Climate and soils at the Brazilian semiarid and the forest-Caatinga problem: new insights and implications for conservation. **Environmental Research Letters**, v. 14, n. 10, p. 104007, 2019. <https://doi.org/10.1088/1748-9326/ab3d7b>

OLIVEIRA, Guilherme de Castro et al., Soil predictors are crucial for modelling vegetation distribution and its responses to climate change. **Science of The Total Environment**, p. 146680, 2021. <https://doi.org/10.1016/j.scitotenv.2021.146680>

PARAHYBA, Roberto da Boa Viagem et al., Pedogenesis of planosols in a toposequence of the agreste region of Pernambuco, Brazil. **Revista Brasileira de Ciência do Solo**, v. 34, n. 6, p. 1991-2000, 2010. <https://doi.org/10.1590/S0100-06832010000600023>

PESSENDA, Luiz Carlos Ruiz et al., Late Pleistocene and Holocene vegetation changes in northeastern Brazil determined from carbon isotopes and charcoal records in soils. **Palaeogeography, Palaeoclimatology, Palaeoecology**, v. 297, n. 3-4, p. 597-608, 2010.

PLUMPTON, Heather J.; MAYLE, Francis E.; WHITNEY, Bronwen S. Long-term impacts of mid-Holocene drier climatic conditions on Bolivian tropical dry forests. **Quaternary Research**, v. 93, n. 1, p. 204-224, 2020. <https://doi.org/10.1017/qua.2019.55>

QUEIROZ, L. P. de. The Brazilian Caatinga: phytogeographical patterns inferred from distribution data of the Leguminosae. **Neotropical savannas and seasonally dry forests: plant diversity, biogeography and conservation**, p. 121-157, 2006.

R Core Team. R: A language and environment for statistical computing. 2020.

RASMUSSEN, Sune O. et al., A new Greenland ice core chronology for the last glacial termination. **Journal of Geophysical Research: Atmospheres**, v. 111, n. D6, 2006. <https://doi.org/10.1029/2005JD006079>

ROSSATTO, Davi Rodrigo; CARVALHO, Fabrício Alvim; HARIDASAN, Mundayatan. Soil and leaf nutrient content of tree species support deciduous forests on limestone outcrops as a

eutrophic ecosystem. **Acta Botanica Brasilica**, v. 29, n. 2, p. 231-238, 2015.

<https://doi.org/10.1590/0102-33062014abb0039>

SANTOS, Janaina Carla dos. O Quaternário do Parque Nacional Serra da Capivara, Piauí, Brasil: morfoestratigrafia, sedimentologia, geocronologia e paleoambientes. 2007. Disponível em: <https://repositorio.ufpe.br/handle/123456789/6387>

SILVA, José Maria Cardoso da; BARBOSA, Luís Cláudio Fernandes. Impact of human activities on the Caatinga. In: **Caatinga**. Springer, Cham, 2017. p. 359-368. [https://doi.org/10.1007/978-3-319-68339-3\\_13](https://doi.org/10.1007/978-3-319-68339-3_13)

SILVA, Renata Botelho Machado da et al., Relação solo/vegetação em ambiente de cerrado sobre influência do grupo Urucua. **Ciência Florestal**, v. 25, n. 2, p. 363-373, 2015. <https://doi.org/10.5902/1980509818455>

SYLVESTRE, Florence. Moisture pattern during the last glacial maximum in South America. **Past climate variability in South America and surrounding regions**, p. 3-27, 2009. [https://doi.org/10.1007/978-90-481-2672-9\\_1](https://doi.org/10.1007/978-90-481-2672-9_1)

SIMÕES, Margareth Gonçalves et al., Democratização da informação de solos do Brasil: Geoportal e banco de dados de solos com acesso via web. **Cadernos de Ciência & Tecnologia**, v. 32, n. 1/2, p. 55-69, 2017. <http://dx.doi.org/10.35977/0104-1096.cct2015.v32.23295>

SOUZA, Cristiano Marcelo Pereira et al., MACHINE LEARNING ALGORITHM IN THE PREDICTION OF GEOMORPHIC INDICES FOR APPRAISAL THE INFLUENCE OF LANDSCAPE STRUCTURE ON FLUVIAL SYSTEMS, SOUTHEASTERN-BRAZIL. **Revista Brasileira de Geomorfologia**, v. 21, n. 2, 2020. <http://dx.doi.org/10.20502/rbg.v21i2.1671>

STEKHOVEN, Daniel J. Using the missForest package. **R package**, p. 1-11, 2011.

STRÍKIS, Nicolás M. et al., Timing and structure of Mega-SACZ events during Heinrich Stadial 1. **Geophysical Research Letters**, v. 42, n. 13, p. 5477-5484A, 2015. <https://doi.org/10.1002/2015GL064048>

STRÍKIS, Nicolás M. et al., South American monsoon response to iceberg discharge in the North Atlantic. **Proceedings of the National Academy of Sciences**, v. 115, n. 15, p. 3788-3793, 2018. <https://doi.org/10.1073/pnas.1717784115>



SU, Zhan; INGERSOLL, Andrew P.; HE, Feng. On the abruptness of Bølling–Allerød warming. **Journal of Climate**, v. 29, n. 13, p. 4965-4975, 2016. <https://doi.org/10.1175/JCLI-D-15-0675.1>

TIERNEY, Jessica E. et al., Glacial cooling and climate sensitivity revisited. **Nature**, v. 584, n. 7822, p. 569-573, 2020. <https://doi.org/10.1038/s41586-020-2617-x>

UTIDA, Giselle et al., Climate changes in Northeastern Brazil from deglacial to Meghalayan periods and related environmental impacts. **Quaternary Science Reviews**, v. 250, p. 106655, 2020. <https://doi.org/10.1016/j.quascirev.2020.106655>

VASCONCELOS, B. N. F. **MAPEAMENTO DIGITAL DE SOLOS EM DIFERENTES ESCALAS: ABORDAGEM METODOLÓGICA**. 2016. Tese de Doutorado. Universidade Federal de Viçosa. <https://locus.ufv.br/handle/123456789/10432>

VENANCIO, I. M. et al., Ocean-atmosphere interactions over the western South Atlantic during Heinrich stadials. **Global and Planetary Change**, p. 103352, 2020. <https://doi.org/10.1016/j.gloplacha.2020.103352>

WANG, Xianfeng et al., Wet periods in northeastern Brazil over the past 210 kyr linked to distant climate anomalies. **Nature**, v. 432, n. 7018, p. 740-743, 2004. <https://doi.org/10.1038/nature03067>

ZULAR, André et al., The role of abrupt climate change in the formation of an open vegetation enclave in northern Amazonia during the late Quaternary. **Global and planetary change**, v. 172, p. 140-149, 2019. <https://doi.org/10.1016/j.gloplacha.2018.09.006>

# 1.7 ANEXOS

## Anexo 1.A

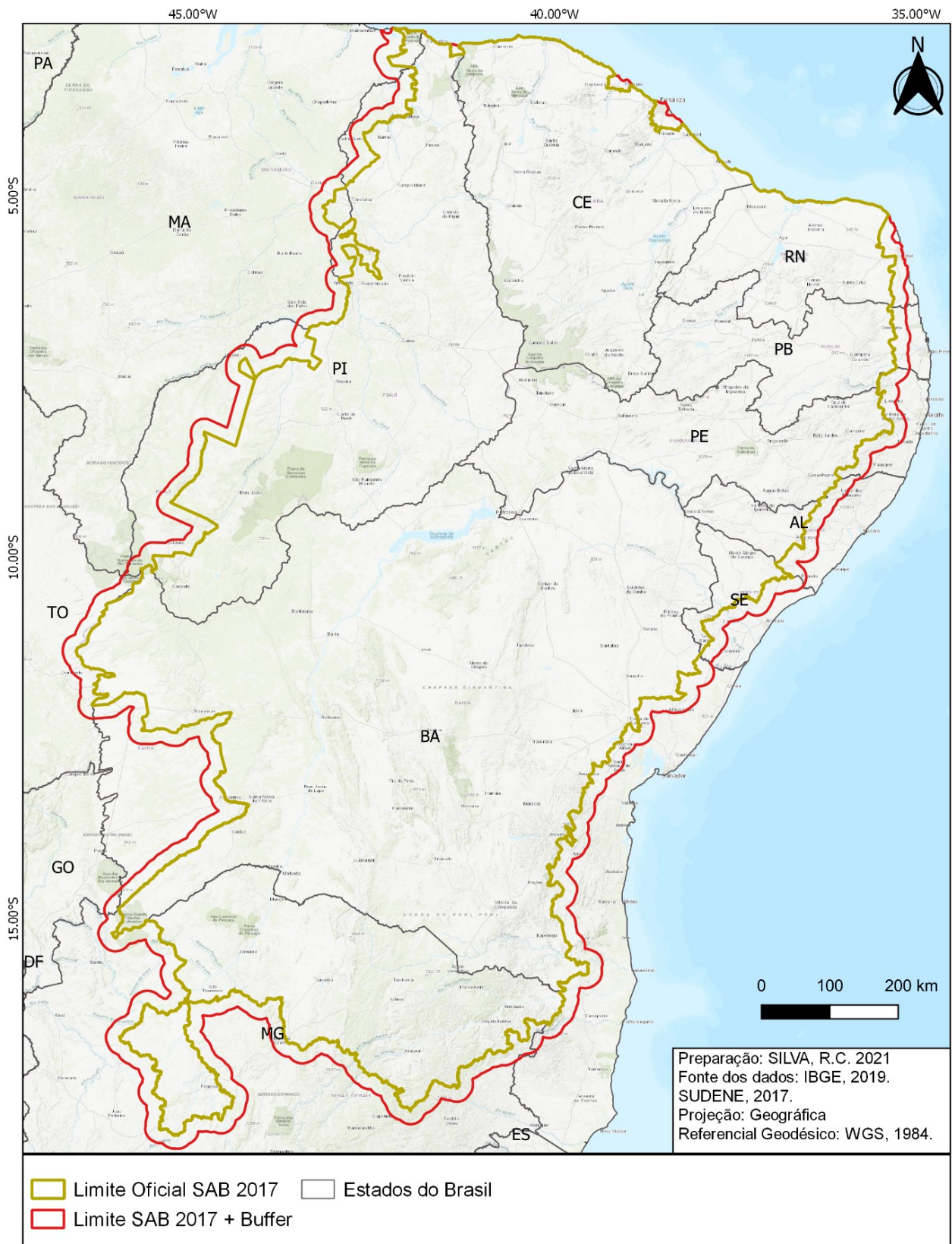


Figura 1.A: Localização do SAB e o buffer aplicado no estudo.

## Anexo 2.A.

Região Fitoecológica	Descrição	Sistema Vegetação IBGE	Tradução
Caatinga Hiperxerófila	Fisionomia arbustiva e florestal. Vegetação xeromórfica e predominante no SAB, adaptada ao clima mais quente e árido da região com temperaturas médias anuais atingindo os 26°C. A precipitação é extremamente sazonal, com um período seco superior a 9 meses e volumes que podem atingir os 400 mm anuais. Posicionada sobre solos relativamente férteis naturalmente.	Savana-Estépica, Savana-Estépica Parque, Savana-Estépica Gramíneo-Lenhosa	Hyperxerophilous Caatinga
Caatinga Arbustiva	Foi adotada a fitofisionomia "Savana-Estépica Arborizada", semelhante a Caatinga descrita anteriormente, se diferenciando por maior concentração de estrato arbustivo. Sua presença pode estar associada a áreas com solos mais profundos que as fitofisionomias Parque e Gramíneo-Lenhosa.	Savana-Estépica Arborizada	Arboreal Caatinga
Caatinga Florestada	Adaptada da fitofisionomia "Caatinga Florestada", possuindo grande semelhança com as Florestas Estacionais. Possui estrato arbóreo e arbustivo, com grande parte das árvores apresentando deciduidade. Marcada por um clima mais ameno que as áreas de Caatinga, também é conhecida como Caatinga hipoxerófila, estando posicionada sobre solos mais profundos e desenvolvidos que as outras duas fitofisionomias.	Savana-Estépica Florestada	Forested Caatinga
Cerrado Típico	Corresponde as fitofisionomias "Savana", "Savana Parque" e "Savana Arborizada". Fisionomia savânica, com transições para florestal e campestre. Condicionada ao clima tropical e sazonal, ocupa áreas na qual o solo apresenta baixa fertilidade natural. Está adaptada à falta de água e incêndios recorrentes, com sua formação florestal ocupando áreas mais úmidas e as formações campestres as áreas de solo mais pobre. No contexto do Semiárido Brasileiro, ocupa a porção oeste e sul da região, compondo uma área de grande contraste e contato ecológico.	Savana, Savana Parque, Savana Arborizada, Savana Gramíneo-Lenhosa	Typical Cerrado
Cerradão	Corresponde à fitofisionomia "Savana Florestada", conhecido pelas áreas de fisionomia florestal do bioma Cerrado. Seu estrato arbóreo varia entre 6 e 8 metros. Marcada por um clima estacional e com inverno seco, está posicionada sobre solos mais profundos e pode estar associada com a disponibilidade hídrica.	Savana Florestada	Forested Cerrado
Floresta Estacional Decidual	Uma formação florestal, de ocorrência em climas tropicais estacionais. Essa vegetação tolera um período seco de até 7 meses, na qual tende a perder 50% das suas folhas. No estudo, essa formação tende a ocupar áreas de transição entre um clima árido e um clima mais úmido, acompanhando áreas com solos mais férteis e relevo um pouco mais acidentado.	Floresta Estacional Decidual	Seasonal Deciduous Forest
Floresta Estacional Semidecidual	Formação florestal. Ocorre em áreas na qual o clima tropical está condicionado a estacionalidade (até quatro meses de período seco), com verões chuvosos e invernos secos. A vegetação apresenta queda de folhas entre 20 e 50%. No Semiárido Brasileiro, essa formação tende a estar em áreas de transição e ecótonos, aonde as condições climáticas e edáficas não favorecem o desenvolvimento de outras formações.	Floresta Estacional Semidecidual	Seasonal Semideciduous Forest
Floresta Ombrófila	Adaptadas as fitofisionomias "Floresta Ombrófila Aberta" e "Floresta Ombrófila Densa". Apresenta uma fisionomia florestal, condicionada a um clima tropical com temperaturas médias elevadas durante todo ano. Está adaptada a um período seco que beira a inexistência, com volumes de precipitação que podem ultrapassar os 2000 mm anuais. Neste estudo sua área principal é a Floresta Atlântica ao leste.	Floresta Ombrófila Aberta, Floresta Ombrófila Densa	Ombrophillous Forest

Quadro 2.A. Organização das formações vegetacionais utilizadas no estudo.

### Anexo 3.A.

Covariável	Sigla	Unidade	Tipo
Temperatura média anual	bio_01	°C	Bioclimática
Faixa média da temperatura diurna	bio_02	°C	Bioclimática
Isotermalidade (bio02/bio07)	bio_03	-	Bioclimática
Sazonalidade da temperatura	bio_04	-	Bioclimática
Temperatura máxima do mês mais quente	bio_05	°C	Bioclimática
Temperatura mínima do mês mais frio	bio_06	°C	Bioclimática
Faixa anual de temperatura (bio05-bio06)	bio_07	°C	Bioclimática
Temperatura média do trimestre mais úmido	bio_08	°C	Bioclimática
Temperatura média do trimestre mais seco	bio_09	°C	Bioclimática
Temperatura média do trimestre mais quente	bio_10	°C	Bioclimática
Temperatura média do trimestre mais frio	bio_11	°C	Bioclimática
Precipitação Anual	bio_12	mm/ano	Bioclimática
Precipitação do mês mais úmido	bio_13	mm/mês	Bioclimática
Precipitação do mês mais seco	bio_14	mm/mês	Bioclimática
Sazonalidade da precipitação	bio_15	-	Bioclimática
Precipitação do trimestre mais úmido	bio_16	mm/trimestre	Bioclimática
Precipitação do trimestre mais seco	bio_17	mm/trimestre	Bioclimática
Precipitação do trimestre mais quente	bio_18	mm/trimestre	Bioclimática
Precipitação do trimestre mais frio	bio_19	mm/trimestre	Bioclimática
Areia Grossa	areia_grossa	(g/kg)	Solo
Areia Fina	areia_fina	(g/kg)	Solo
Areia Total	areia_total	(g/kg)	Solo
Silte	silte	(g/kg)	Solo
Argila	argila	(g/kg)	Solo
Relação Silte/Argila	relacao_silte_argila	(g/kg)	Solo
Ca Mg	ca_mg	(cmolc.kg-1)	Solo
Alumínio Trocável (Al <sup>3+</sup> )	al	(cmolc.kg-1)	Solo
Hidrogênio (H <sup>+</sup> )	h	(cmolc.kg-1)	Solo
Potássio (K <sup>+</sup> )	k	(cmolc.kg-1)	Solo
Saturação por Alumínio (100. Al <sup>3+</sup> /S+Al <sup>3+</sup> )	m	(cmolc.kg-1)	Solo
Nitrogênio (N)	n	(g/kg)	Solo
Sódio (Na <sup>+</sup> )	na	(g/kg)	Solo
pH (H <sub>2</sub> O)	p_h_h2o	-	Solo
pH (KCl)	p_h_k_cl	-	Solo
Soma de Bases (Ca <sup>2+</sup> +Mg <sup>2+</sup> +K <sup>+</sup> +Na <sup>+</sup> )	sb	%	Solo
Valor T (S+H <sup>+</sup> +Al <sup>3+</sup> )	t	(cmolc.kg-1)	Solo
Valor V – Saturação por Bases (100. S/T)	v	%	Solo
Carbono orgânico	co	(g/kg)	Solo
Relação C/N	c_n	%	Solo
Altitude	altitude	metros	Solo
Declividade	declividade	%	Solo

Tabela 3.A. Conjunto total de preditores utilizados no estudo

## Anexo 4.A.

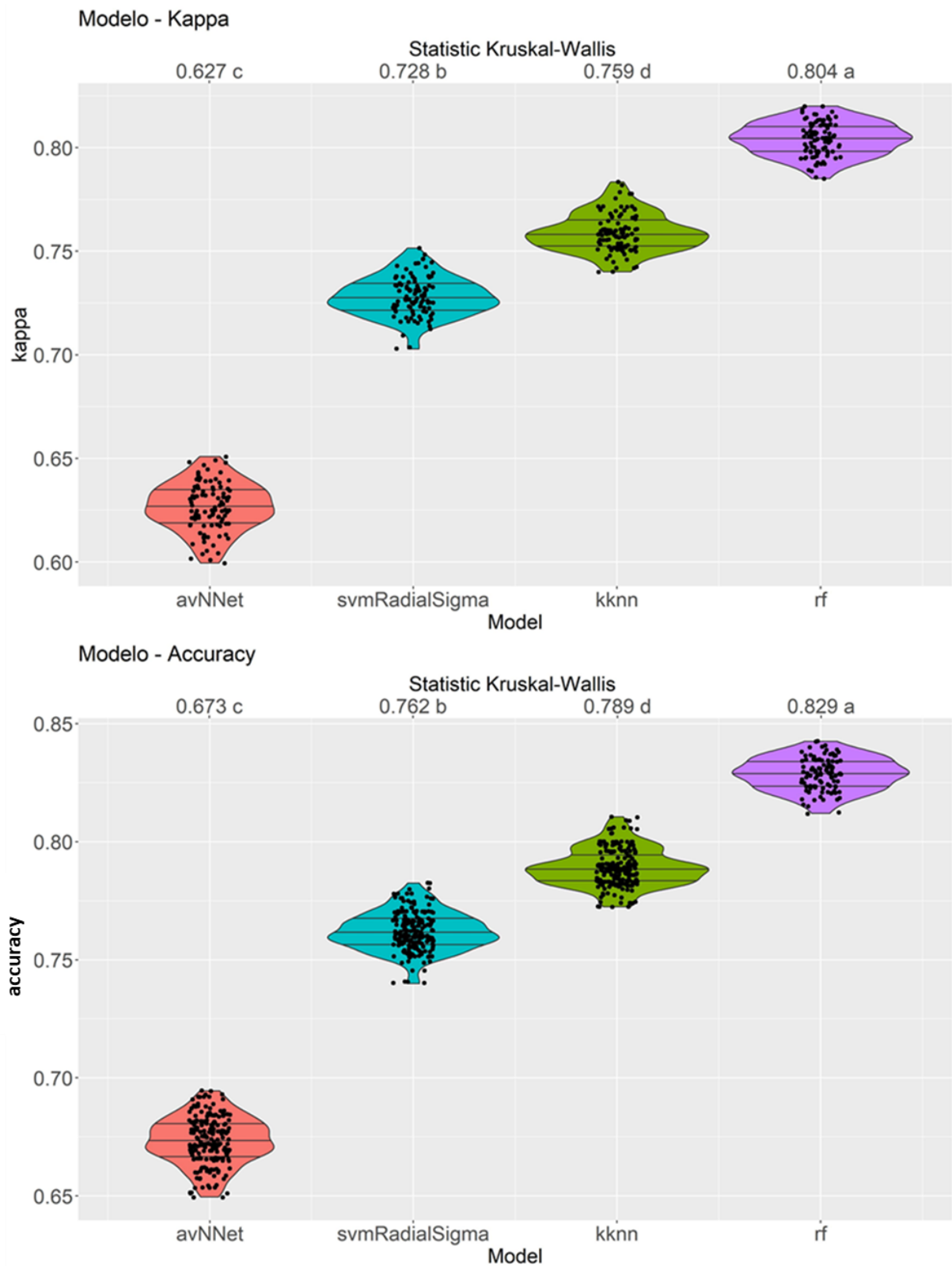


Figura 4.A. Desempenho dos modelos bioclim+soil nas 100 rodadas realizadas pelo estudo. Diferença nas médias obtidas por teste de Kruskal-Wallis e diferença entre grupos pelo teste de Dunn.

Anexo 5.A.

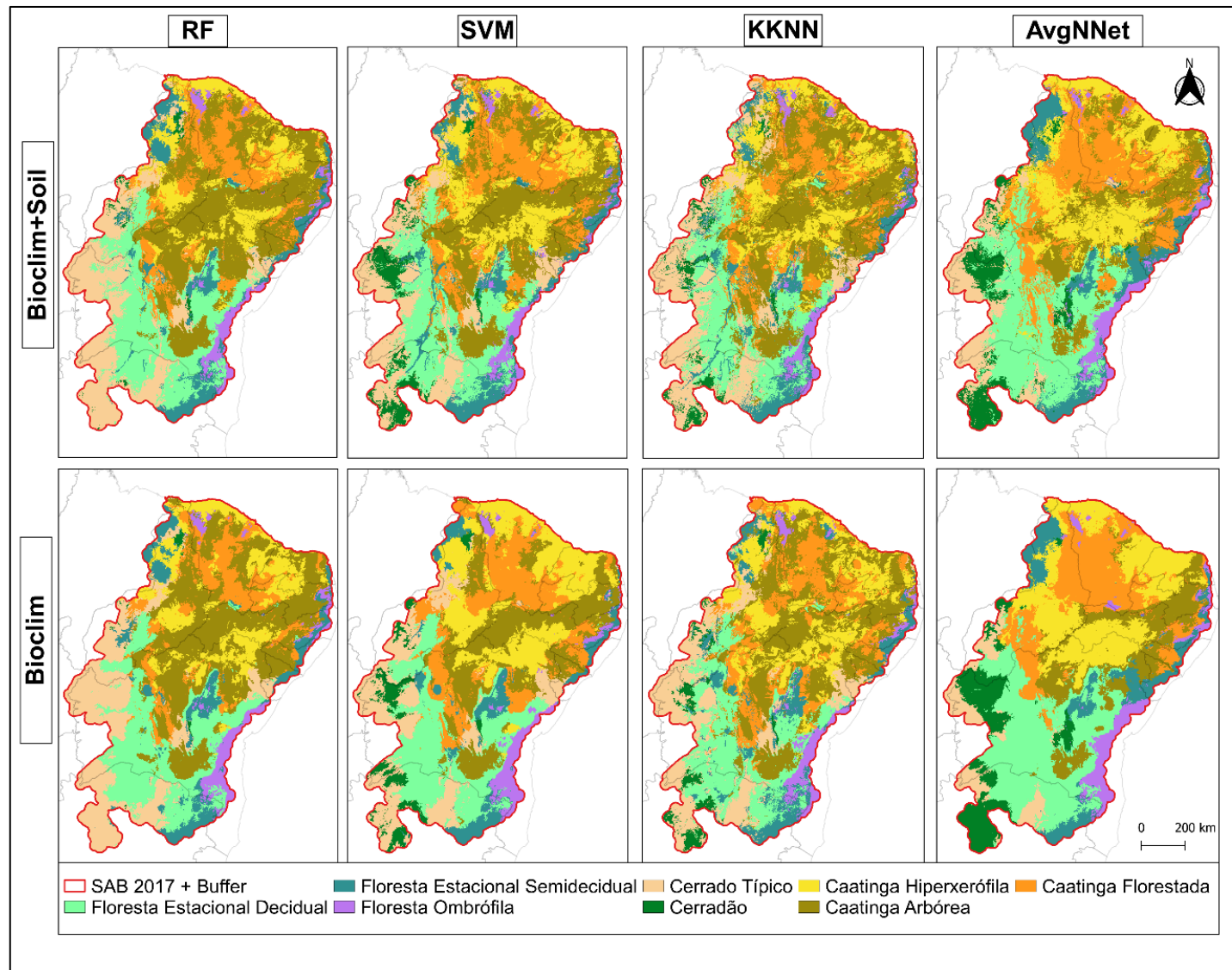


Figura 5.A: Comparativo de vegetação potencial entre os modelos que utilizaram covariáveis bioclimáticas e covariáveis bioclimáticas + edáficas.

## Anexo 6.A.

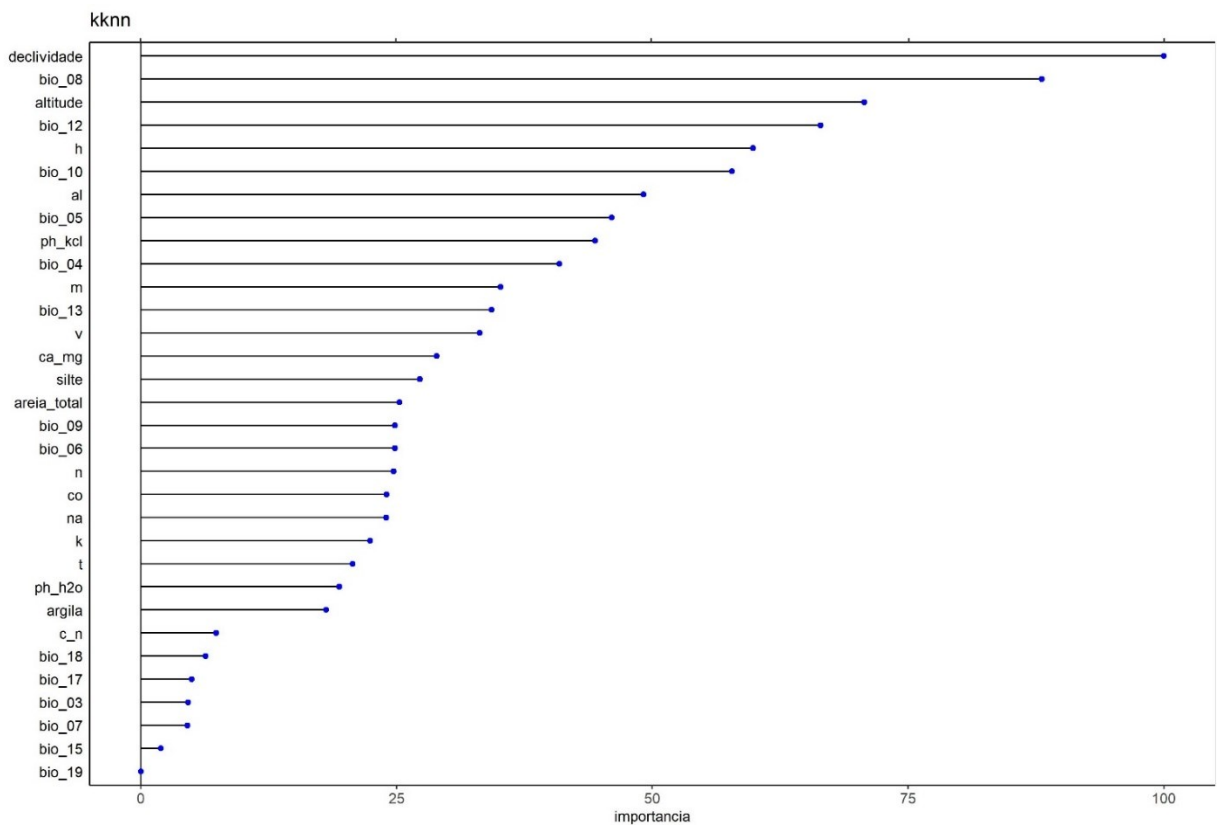
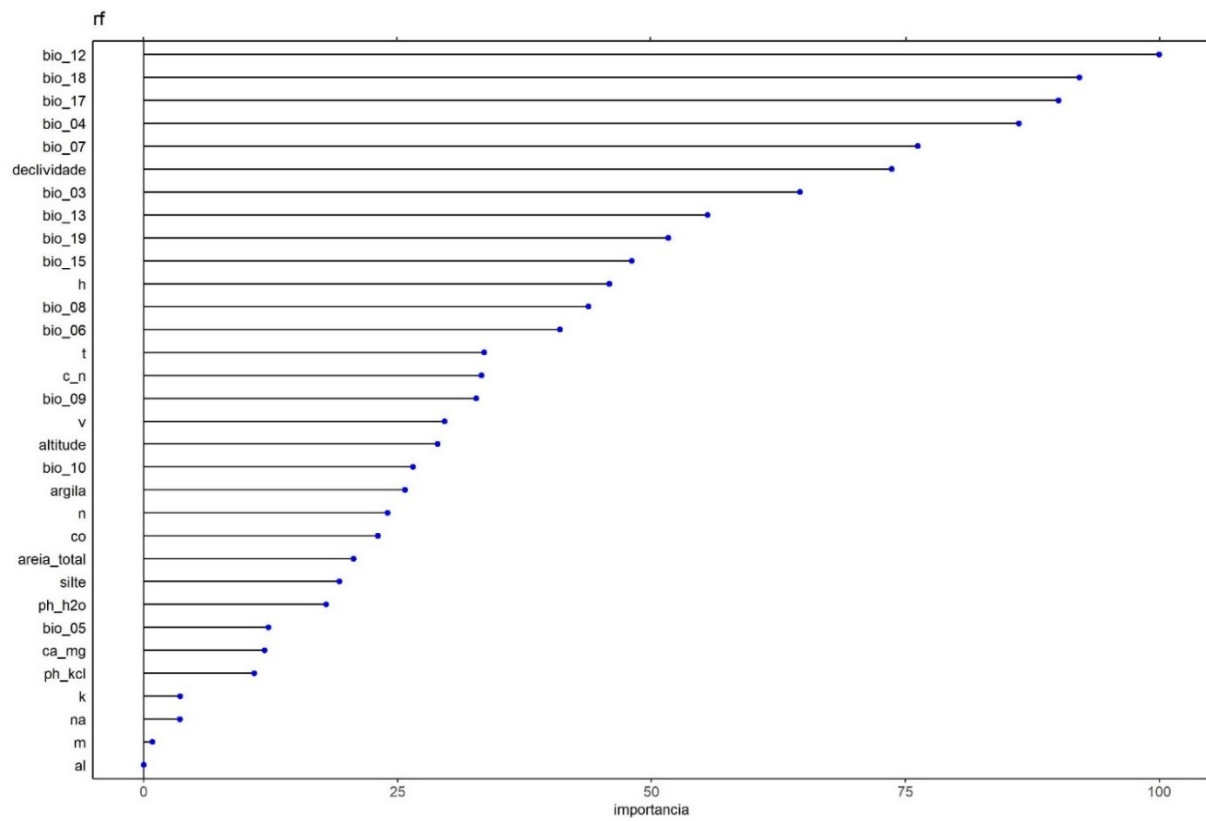
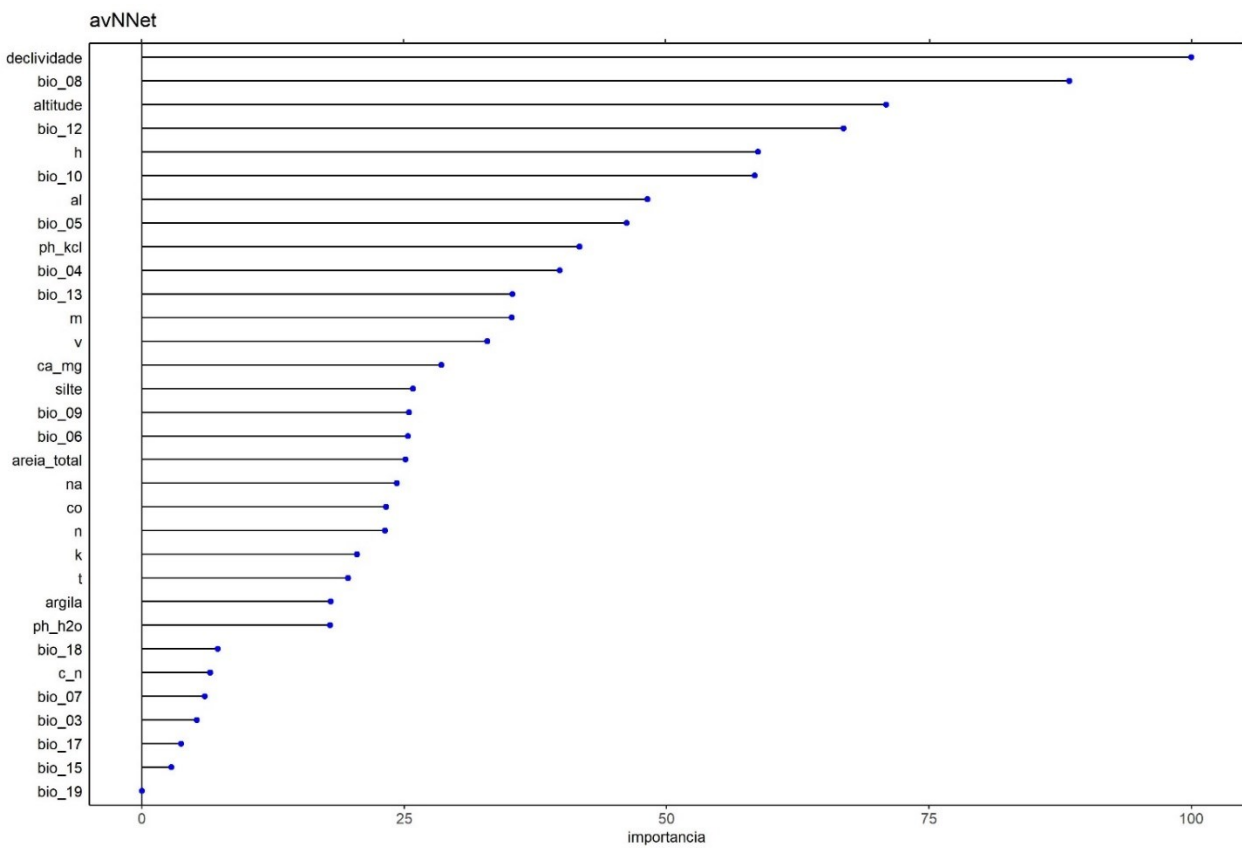
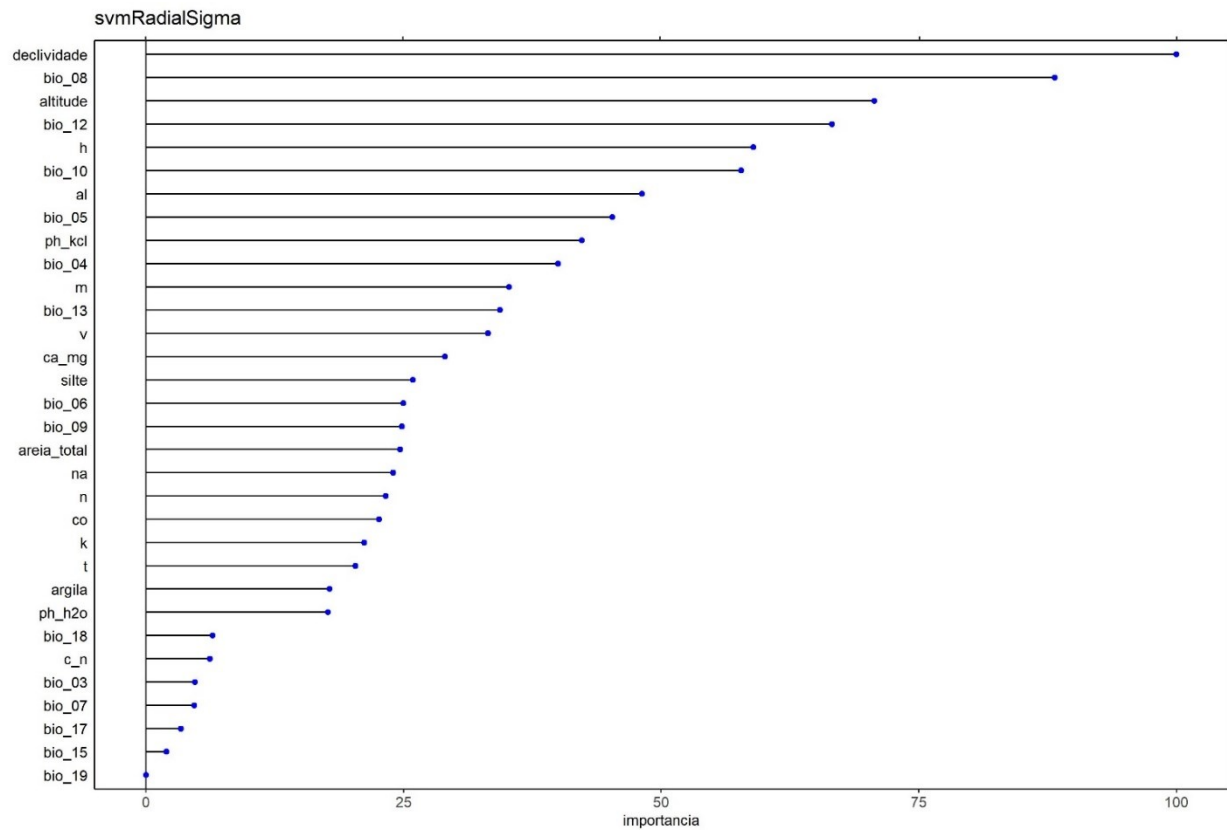


Figura 6.A: Importância das variáveis nas 100 rodadas dos modelos bioclim+soil na predição do presente, sendo na imagem os modelos RF e KKNN.

## Anexo 6.A Continuação



Continuação Figura 6.A: Importância das variáveis nas 100 rodadas dos modelos bioclim+soil na predição do presente, sendo na imagem os modelos SVM e AVGNNET.



Anexo 7.A.

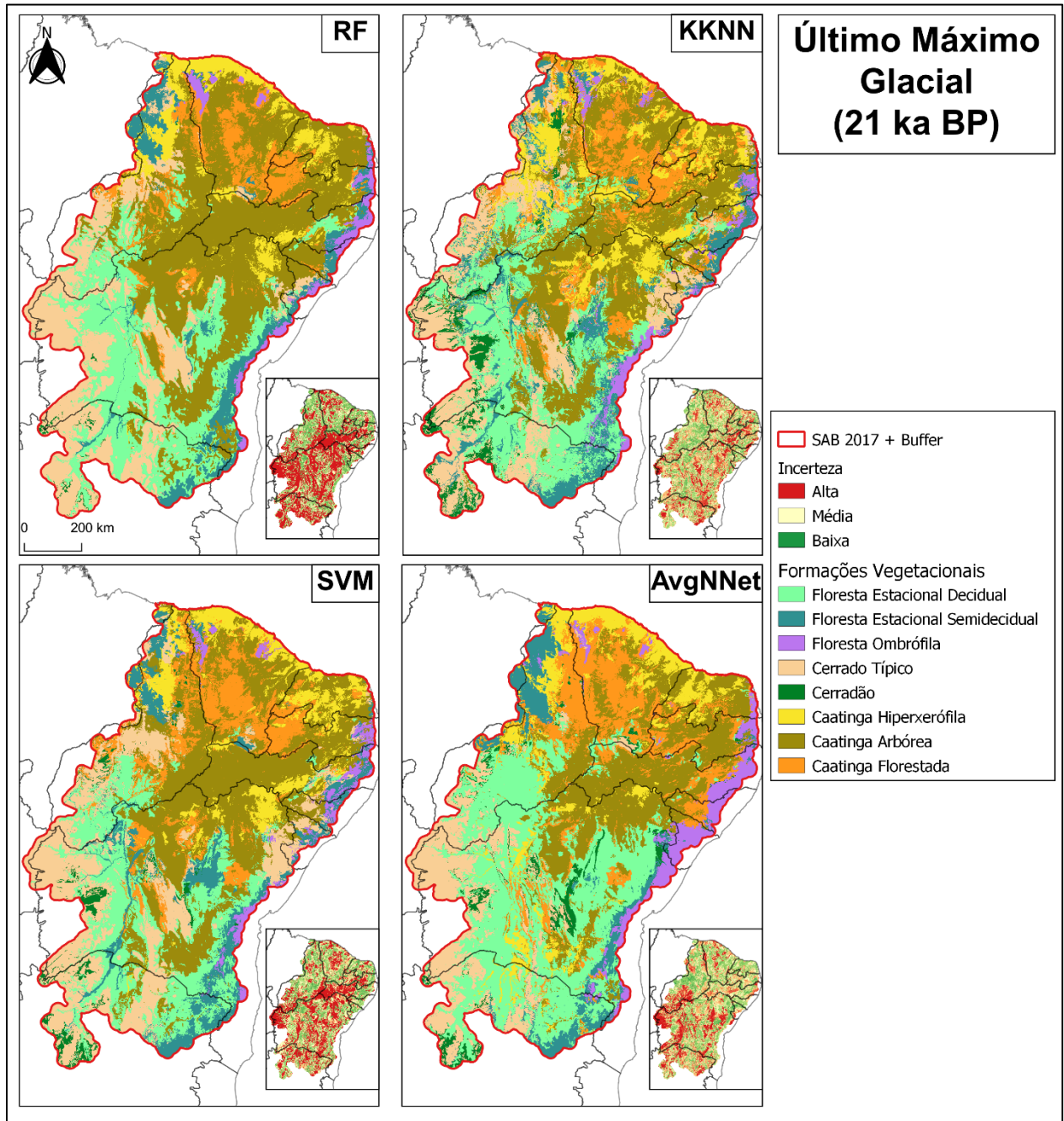


Figura 7.A. Resultados da predição dos modelos bioclim+soil no período do Último Máximo Glacial.

Anexo 8.A.

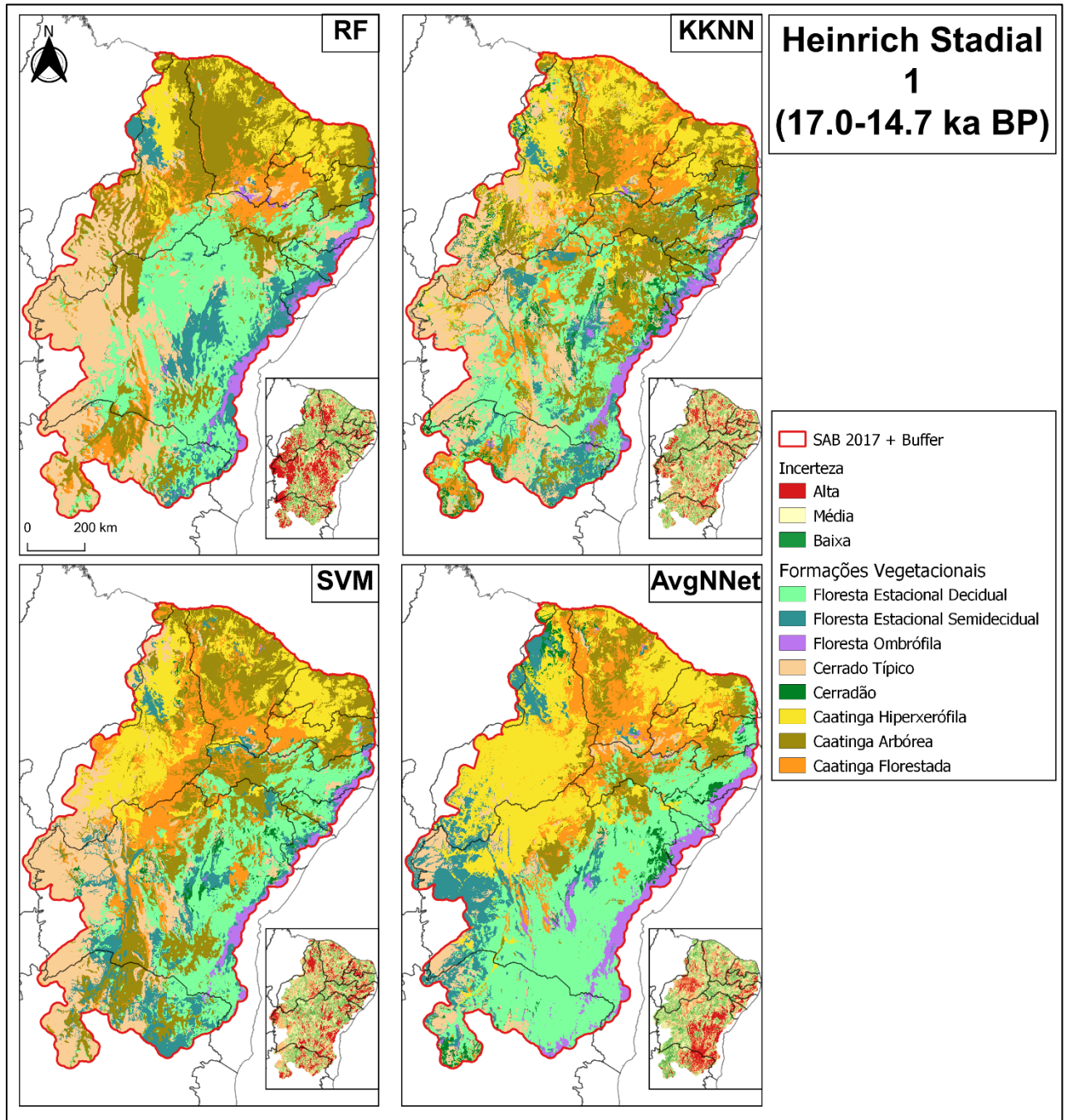


Figura 8.A. Resultados da predição dos modelos bioclim+soil no período de Heinrich Stadial 1.

Anexo 9.A.

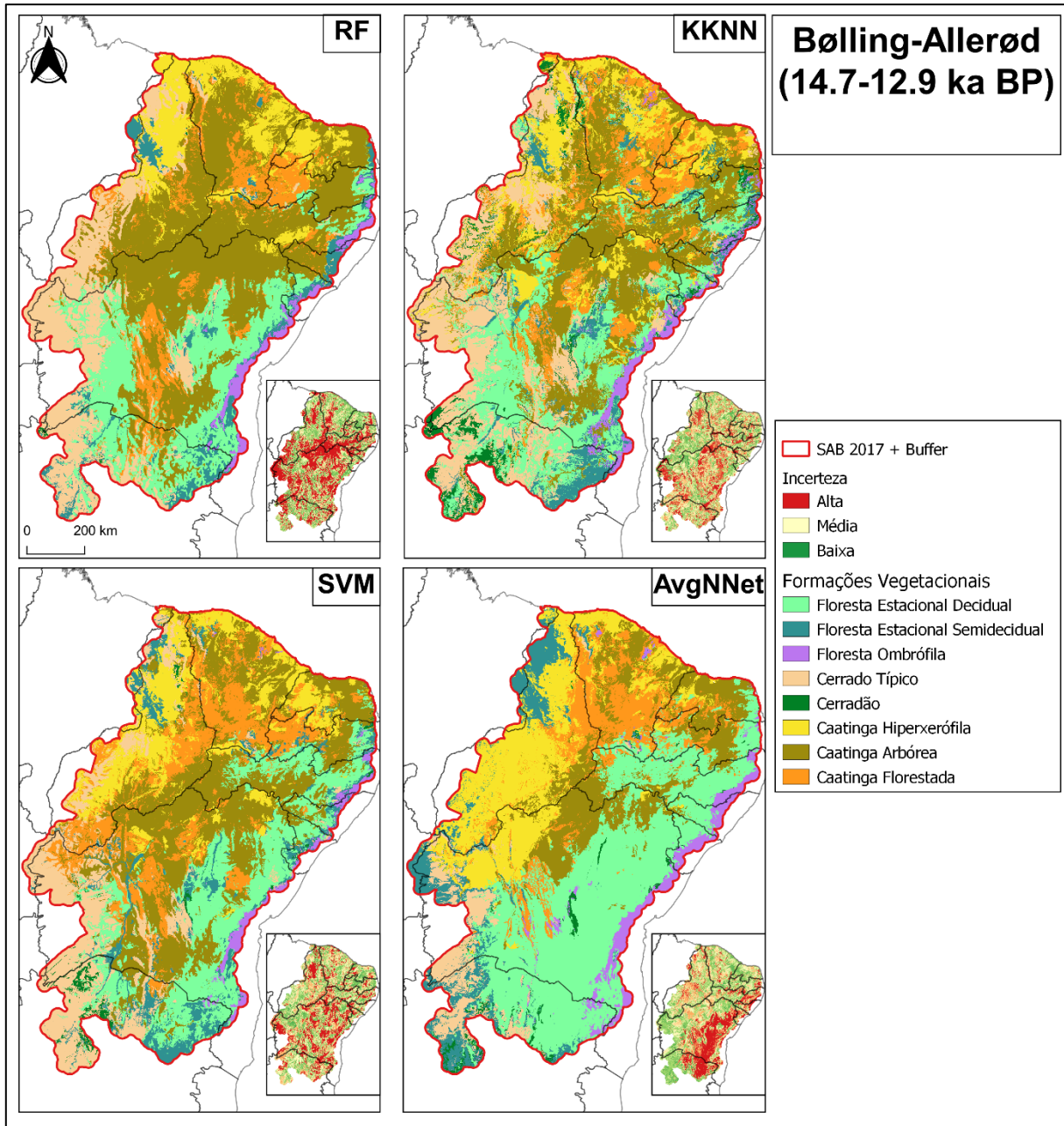


Figura 9.A. Resultados da predição dos modelos bioclim+soil no período de Bølling-Allerød.

Anexo 10.A.

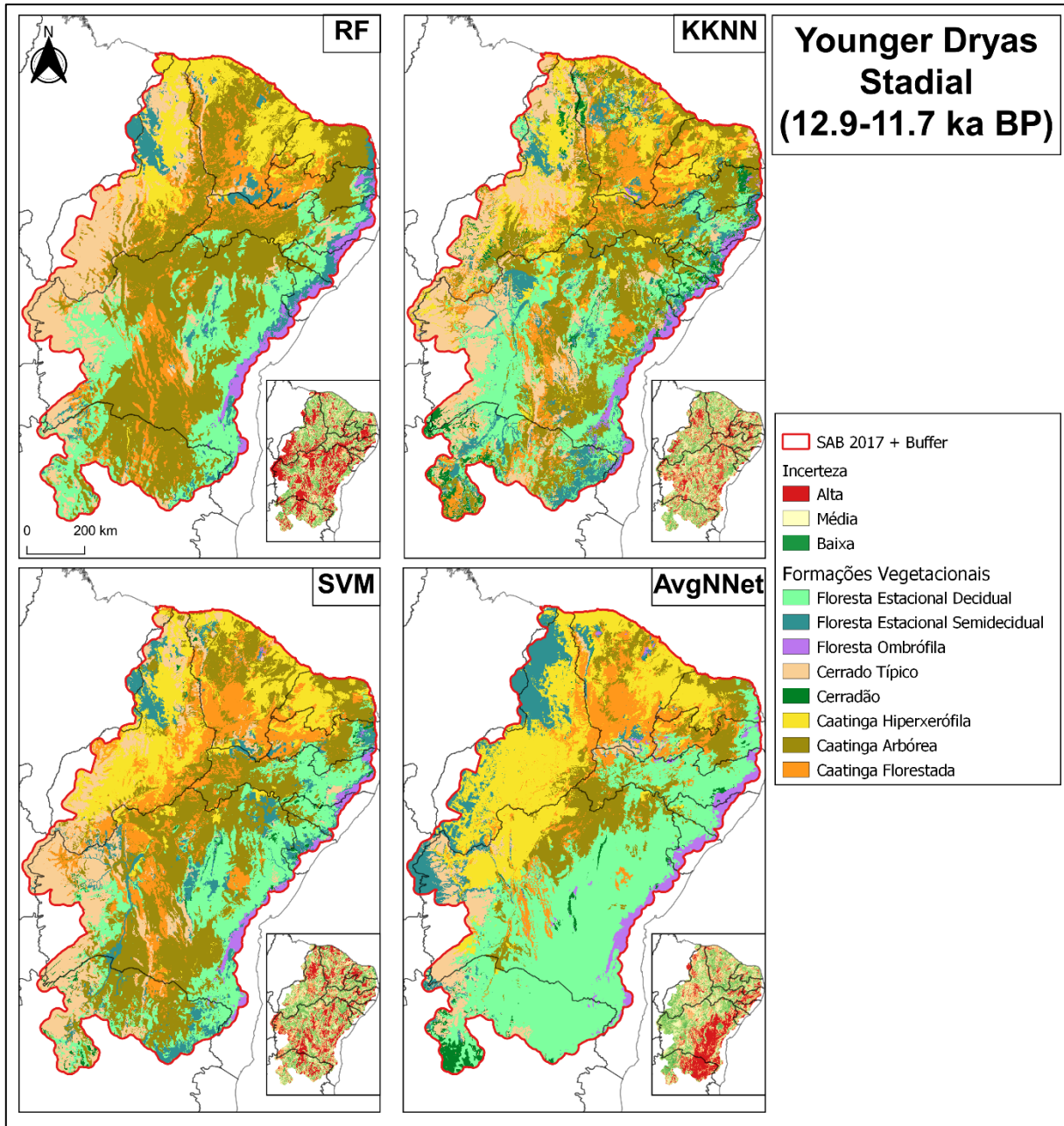


Figura 10.A. Resultados da predição dos modelos bioclim+soil no período de Younger Dryas.

Anexo 11.A.

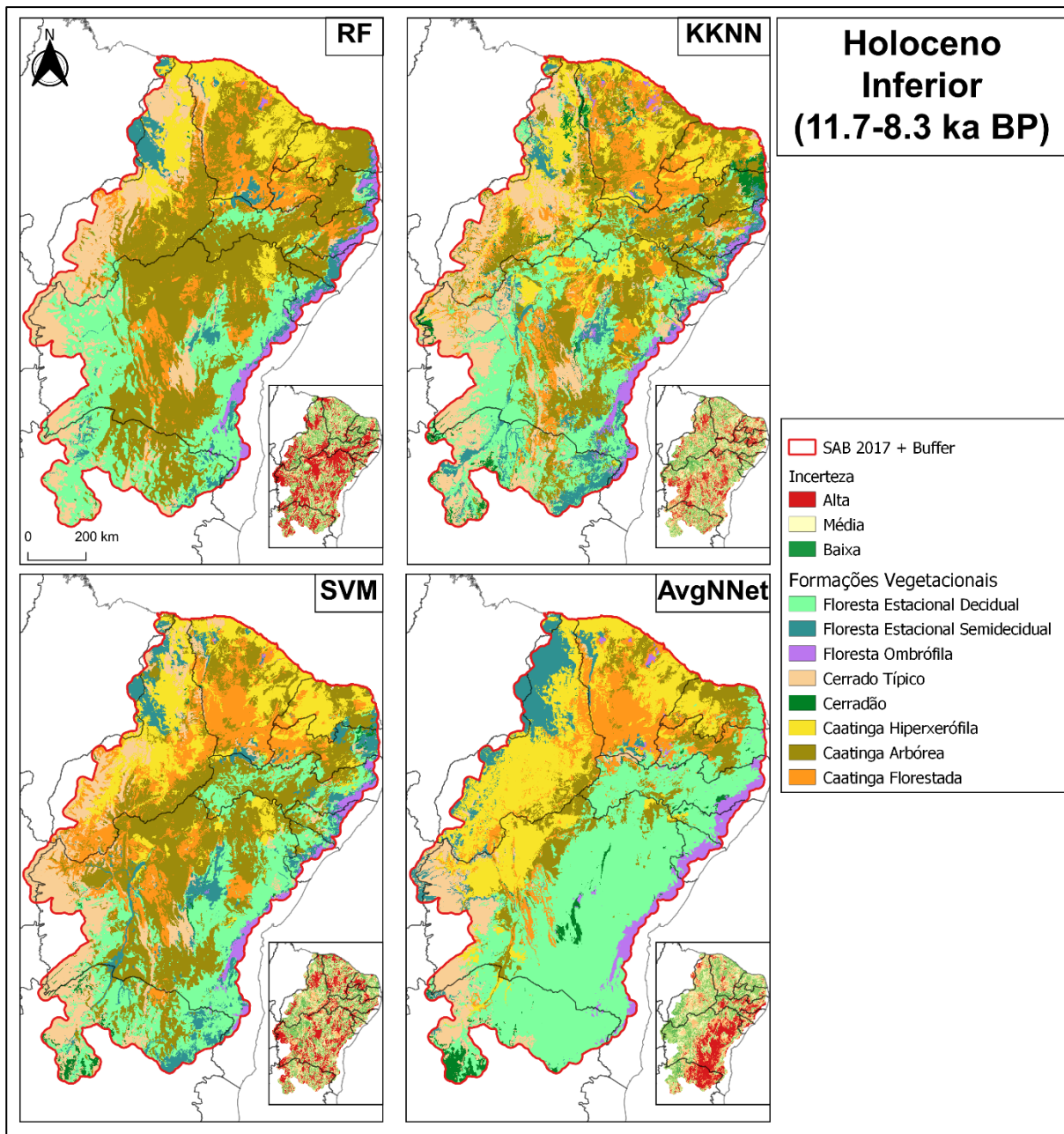


Figura 11.A. Resultados da predição dos modelos bioclim+soil no período do Holoceno Inferior.

Anexo 12.A.

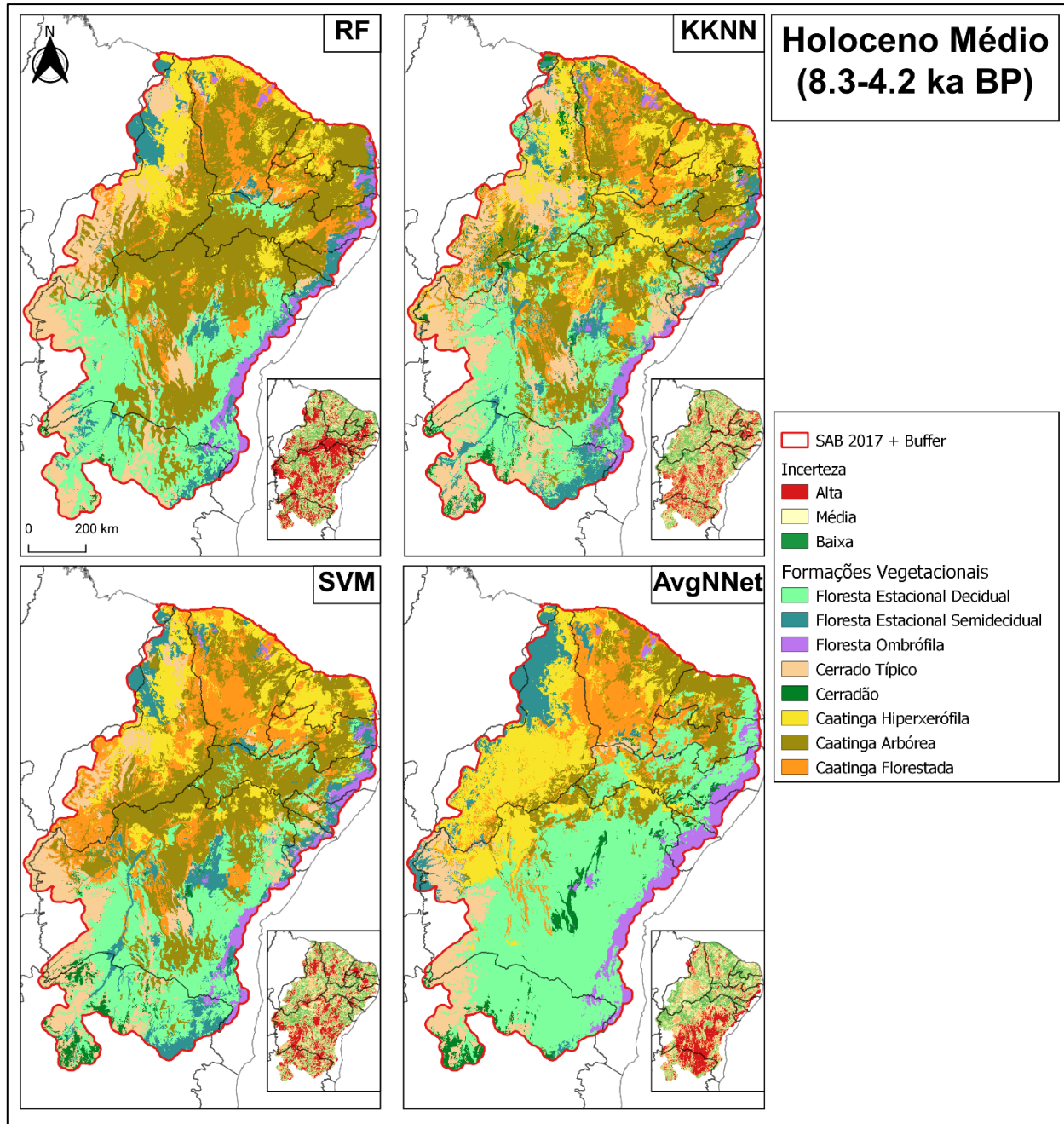


Figura 12.A. Resultados da predição dos modelos bioclim+soil no período do Holoceno Médio.

Anexo 13.A.

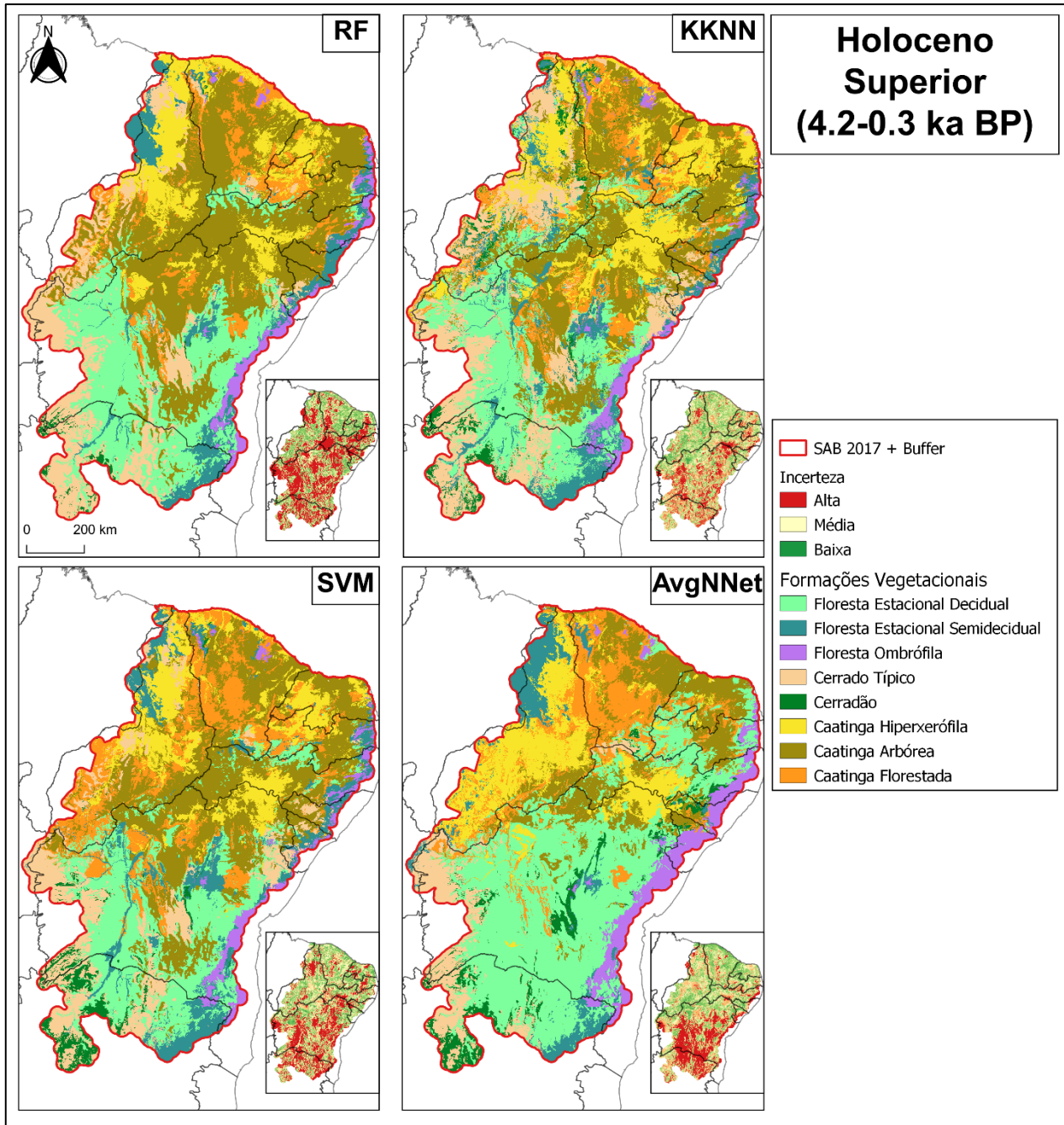


Figura 13.A. Resultados da predição dos modelos bioclim+soil no período de Holoceno Superior.

## Anexo 14.A.

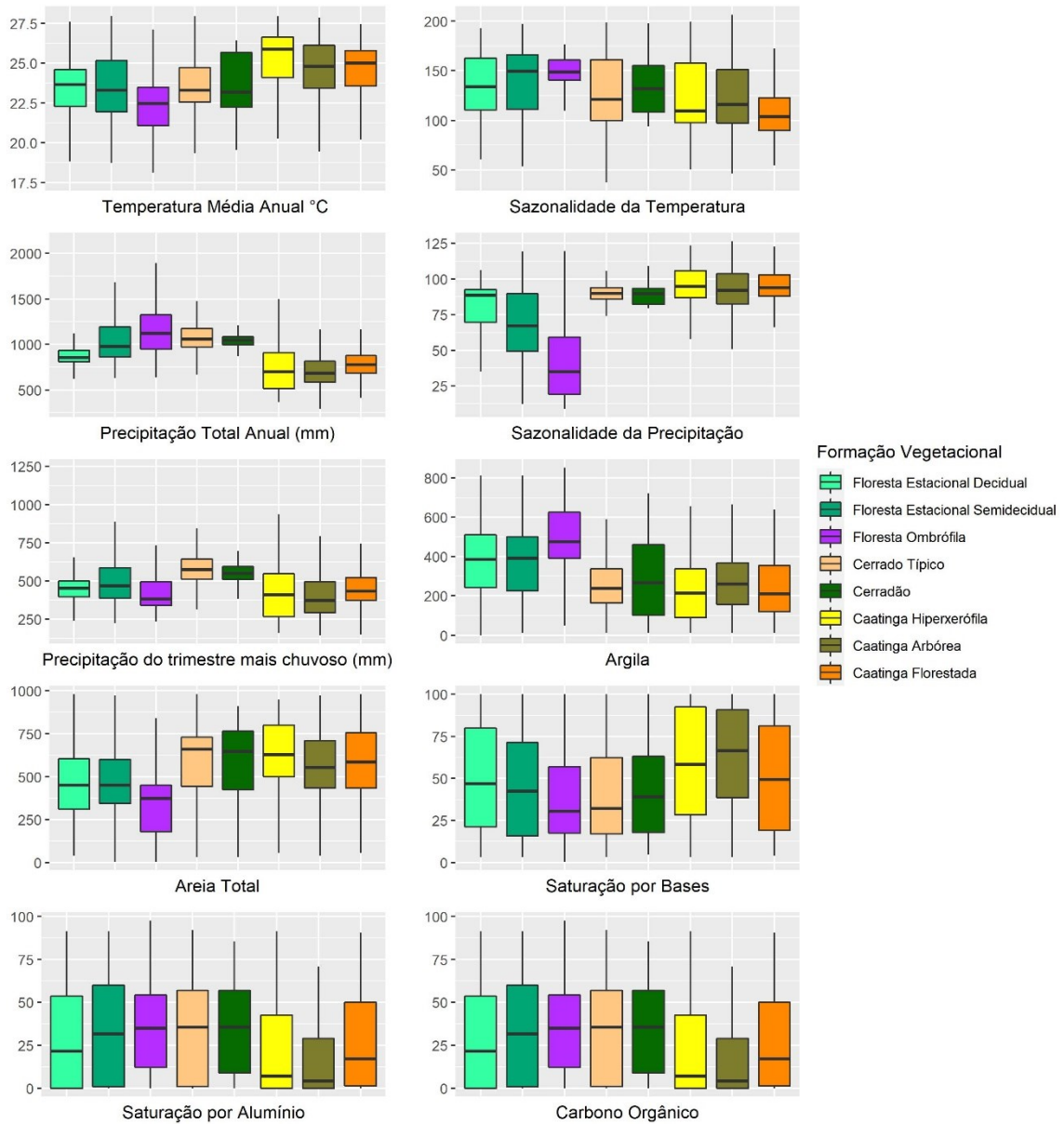


Figura 14.A. Variação de covariáveis nas formações vegetacionais adotadas no estudo do SAB. Fonte: Fick and Hijmans, 2017; Embrapa, 2019.



## Anexo 15.A.

	KKNN							
	LGM	HS1	BA	YD	EH	MH	LH	PRE
Caatinga Hiperxerófila	11.4	13.7	12.9	12.1	13.5	14.4	14.9	13.5
Caatinga Arbórea	26	22	25.7	25.4	22.4	23.1	23.5	27.7
Caatinga Florestada	10	12.9	11.4	12.1	12.8	11.6	9.5	11.9
Cerrado Típico	17.5	19.3	18.6	18.4	18.4	18.9	17.6	17.3
Cerradão	3.6	3.5	3.8	3.5	2.7	2.3	2.9	2.8
Floresta Estacional Decidual	18.7	17.7	18.1	17.7	20.7	18.4	19	14.8
Floresta Estacional Semidecidual	9.9	8	6.7	8.2	7.2	8.5	9.6	8.7
Floresta Ombrófila	2.8	3	2.8	2.6	2.5	2.7	2.9	3.3

Figura 15.A. Porcentagem da área ocupada por formação vegetacional nos diferentes períodos reconstruídos pelo modelo bioclim+soil KKNN.

## Anexo 16.A.

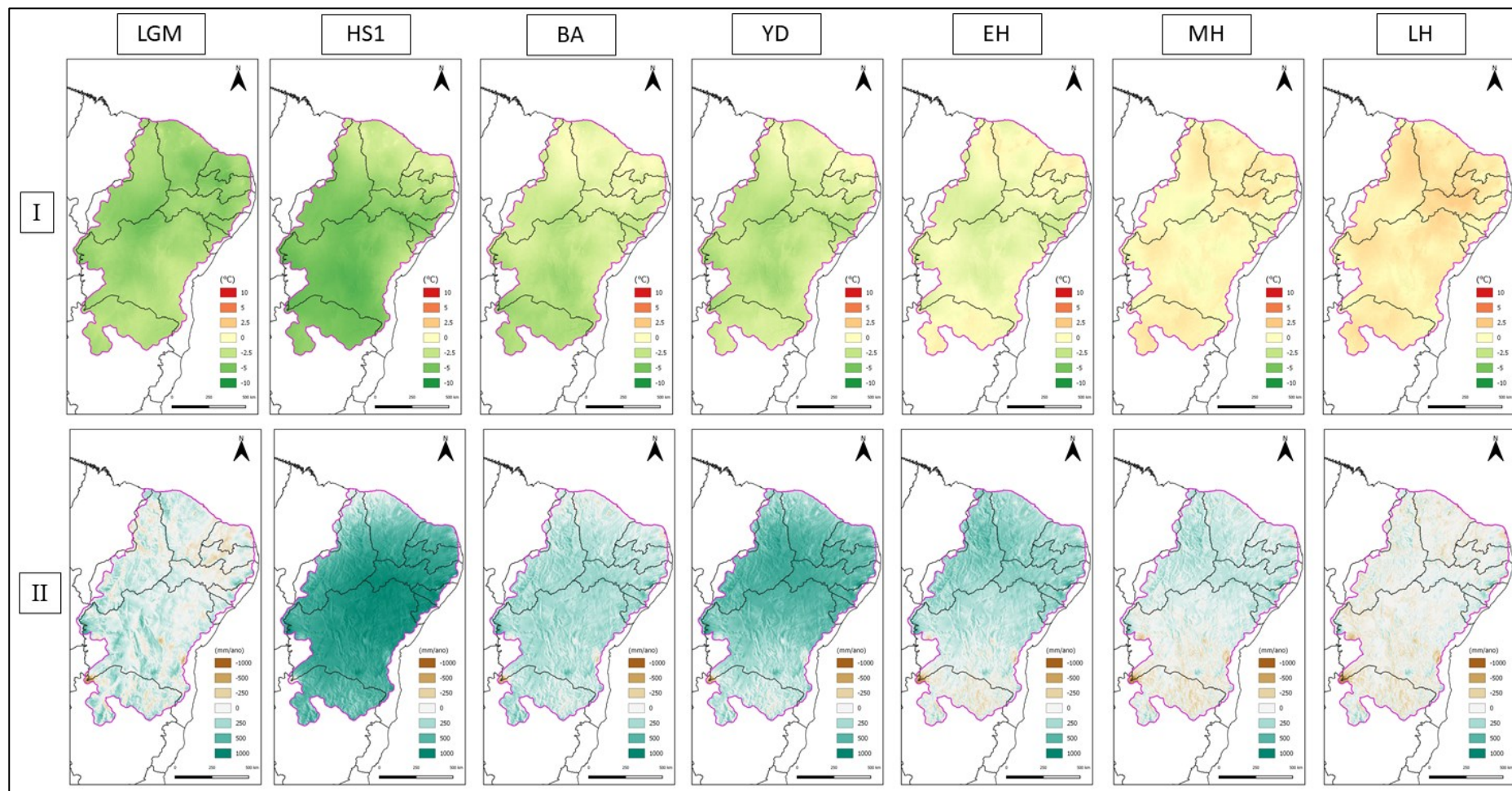


Figura 16.A: Anomalias paleoclimáticas entre o final do Pleistoceno e início do Holoceno. I) Anomalia de temperatura média anual °C (Climatologia: 1970-2000); II) Anomalia de precipitação total anual (Climatologia: 1970-2000). Clima presente: Fick e Hijmans, 2017. Clima passado: Brown et al., 2018.

## Anexo 17.A

Tipo	Autores	Área de Estudo	LGM	HS1	BA	YD	EH	MH	LH
Sequência de turfeiras	De Oliveira, Barreto & Suguio. 1999	Vale do Rio Icatu	x	x	x	x	Vegetação de floresta ombrófila e floresta montana	Vegetação de vereda, savana e savana estépica	Vegetação de savana-estépica
Testemunho de sedimentos marinhos	Behling et al., 2000	GeoB 3104-1	x	Vegetação de veredas, florestas de galeria e florestas úmidas			Vegetação de savana-estépica		
Espeleotema	Auler e Smart, 2001	Toca da Boa Vista	Indicação de período úmido				Indicação de período seco		
Espeleotema	Wang et al., 2004	Toca da Boa Vista, Toca da Barriguda e Lapa dos Brejões	x	Indicação de período úmido	x	x	x	x	x
Análise de sedimentos	Santos, J.C. 2007	Serra da Capivara	x	x	x	x	x	Período úmido, com espécies de áreas pantanosas	Transição para um período mais seco, com declínio das espécies arbóreas
Palinologia de argila rica em material orgânico	Nascimento, L. R. S. L. 2008	Parna Catimbau	x	x	x	x	x	Período úmido com alta taxa de sedimentação	Período seco (1,6 ky AP) com ocupação promovida por espécies típicas da caatinga
Espeleotema	Cruz et al., 2009	Rainha, Furna Nova e Abissal	Indicação de período seco	Indicação de período úmido	Indicação de período seco	x	Indicação de período úmido		Indicação de período seco
Datação de isótopos de carbono e registros de carvão	Pessenda et al., 2010	Flona Araripe, Parna Sete Cidades e Rebio Guaribas	Indicação de período úmido				Transição úmido para seco	Indicação de período seco	Breve período de umidade seguido por período seco
Espeleotema	Novello et al., 2012	Diva de Maura, Torrinha e Lapa Doce	x	x	x	x	x	x	Período seco
Espeleotema	Strikis et al., 2015	Lapa Sem Fim e Paixão	x	Período úmido com grande aumento na precipitação	x	x	x	x	x

Quadro 17.A: Resumo dos registros paleoecológicos analisados para este estudo.

### Continuação Anexo 17.A

Tipo	Autores	Área de Estudo	LGM	HS1	BA	YD	EH	MH	LH
Testemunho de sedimentos marinhos e lacustres	Mendes, V.R. 2016	Parna Serra da Capivara, Médio e Baixo Parnaíba, GeoB 16206	x	Indicação de período úmido	X	Indicação de período úmido	X	X	X
Sequência de turfeiras	De Medeiros et al., 2018	Bom Jesus e Vale do Catimbau	x	x	x	x	Período úmido com vegetação arbórea	Transição entre período úmido com vegetação próxima à Caatinga	Estabelecimento das condições secas predominantes
Espeleotema	Strikis et al., 2018	Lapa Grande, Lapa Sem Fim, Paixão e Marota	x	Período úmido com grande aumento na precipitação	x	x	x	x	x
Testemunho de sedimentos marinhos	Venancio et al., 2020	GL-1248, GeoB 16206	x	Indicação de período úmido	x	x	x	x	
Testemunho de sedimento em estuário	Martins et al., 2020	Rio Formoso	x	x	x	x	Início da ocupação de vegetação de restinga	Desenvolvimento completo da vegetação de restinga	
Testemunho de sedimento terrestres	Moraes, Oliveira e Behling 2020	Parna Catimbau	x	x	x	x	x	x	Período úmido entre 2,1 e 0,4 ky AP. Estabelecimento de condições secas nos últimos 450 anos.
Espeleotema	Utida et al., 2020	Rainha, Abissal, Trapiá e Urubu	x	Indicação de período úmido	Indicação de período seco	Indicação de período úmido	Transição para período seco	Indicação de período seco	
Testemunho de sedimentos terrestres	De Oliveira, Santos e Lemos, 2020	Serra da Capivara	Indicação de período seco	Indicação de período úmido	x	x	Indicação de período úmido	x	x

Continuação Quadro 17.A: Resumo dos registros paleoecológicos analisados para este estudo.

## Anexo 18.A

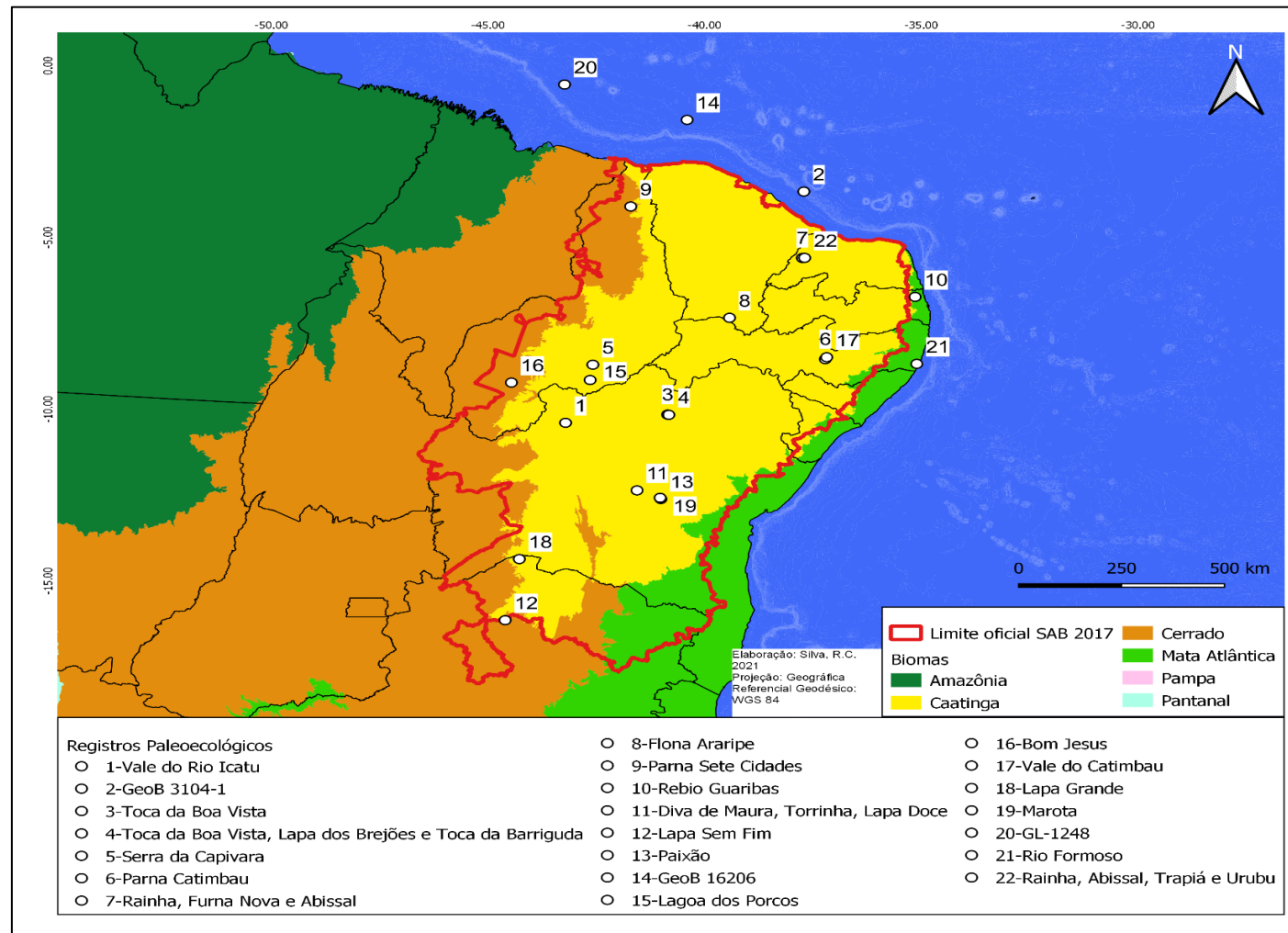


Figura 18.A. Registros paleoecológicos analisados neste estudo. Fonte: De Oliveira, Barreto & Suguio. 1999; Behling et al., 2000; Auler e Smart, 2001; Wang et al., 2004; Santos, J.C. 2007; Nascimento, L. R. S. L. 2008; Cruz et al., 2009; Pessenda et al., 2010; Novello et al., 2012; Strikis et al., 2015; Mendes, V.R. 2016; De Medeiros et al., 2018; Strikis et al., 2018; Venancio et al., 2020; Martins et al., 2020; Moraes, Oliveira e Behling 2020; Utida et al., 2020; De Oliveira, Santos e Lemos, 2020.