



UNIVERSIDADE FEDERAL DO PARAÍBA
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA (CCEN)
DEPARTAMENTO DE GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GEOGRAFIA
DOUTORADO EM GEOGRAFIA

RAMON SANTOS SOUZA

**ANÁLISE DAS PROPRIEDADES NUTRICIONAIS DO SOLO E DA ESTRUTURA
FLORESTAL NO BIOMA CAATINGA**

JOÃO PESSOA – PB,
2023

RAMON SANTOS SOUZA

ANÁLISE DAS PROPRIEDADES NUTRICIONAIS DO SOLO E DA ESTRUTURA
FLORESTAL NO BIOMA CAATINGA

Tese apresentada ao Programa de Pós-Graduação em Geografia do Centro de Ciências Exatas e da Natureza (CCEN) da Universidade Federal do Paraíba, como requisito parcial à obtenção do título de doutor em Geografia. Área de Concentração: Gestão do Território e Análise Geoambiental

Orientador: Prof. Dr. Bartolomeu Israel de Souza

Coorientador: Prof. Dr. José João Lelis Leal de Souza

JOÃO PESSOA – PB,

2023

Catálogo na publicação
Seção de Catalogação e Classificação

S729a Souza, Ramon Santos.

Análise das propriedades nutricionais do solo e da estrutura florestal no bioma Caatinga / Ramon Santos Souza. - João Pessoa, 2023.

134 f. : il.

Orientação: Bartolomeu Israel de Souza.

Coorientação: José João Lelis Leal de Souza.

Tese (Doutorado) - UFPB/CCEN.

1. Geografia dos vegetais. 2. Estrutura da vegetação. 3. Análise geoambiental. 4. Google Earth Engine. 5. Covariáveis ambientais. I. Souza, Bartolomeu Israel de. II. Souza, José João Lelis Leal de. III. Título.

UFPB/BC

CDU 581.9(043)



Universidade Federal da Paraíba
Centro de Ciências Exatas e da Natureza
Programa de Pós-Graduação em Geografia

ATA DA SESSÃO PÚBLICA DA DEFESA DE TESE DE
DOUTORADO DO ALUNO RAMON SANTOS
SOUZA, CANDIDATO AO TÍTULO DE DOUTOR EM
GEOGRAFIA, NA ÁREA DE TERRITÓRIO,
TRABALHO E AMBIENTE.

Aos 28 (vinte e oito) dias do mês de Fevereiro de dois mil e vinte e três, às 14:00 (quatorze) horas, na plataforma no Departamento de Geociências, reuniram-se, em caráter de solenidade pública, os membros da comissão designada para avaliar **Ramon Santos Souza**, candidato ao grau de doutor em Geografia, na área de **Território, Trabalho e Ambiente**. Foram componentes da banca examinadora os professores Bartolomeu Israel de Souza (orientador) – Doutor em Geografia; Jonas Otaviano Praca de Souza (examinador interno) – Doutor em Geografia; Rafael Albuquerque Xavier (examinador interno) - Doutor em Engenharia Civil; Valeria Raquel Porto de Lima (examinadora externa) – Doutora em Geografia – e Elpídio Inácio Fernandes Filho (examinador externo) - Doutor em Agronomia, sendo o primeiro, o segundo e o terceiro integrantes do corpo docente da UFPB, a quarta, do corpo docente do IFPB, e o quinto, do corpo docente da UFCG. Dando início aos trabalhos, o Presidente da banca, Prof. Bartolomeu Israel de Souza, após declarar os objetivos da reunião, apresentou o candidato Ramon Santos Souza, a quem concedeu a palavra para que defendesse, oral e sucintamente, sobre o tema apresentado, intitulado: " **ANÁLISE DAS PROPRIEDADES NUTRICIONAIS DO SOLO E DA ESTRUTURA FLORESTAL NO BIOMA CAATINGA** ". Após discorrer sobre o referido tema, o candidato foi arguido pelos examinadores, na forma regimental. Ato contínuo, passou a comissão em caráter secreto, a proceder à avaliação e julgamento do trabalho, concluindo por atribuir-lhe o conceito **Aprovado**. Face à aprovação, declarou o Presidente achar-se o avaliado legalmente habilitado a receber o Grau de doutor em Geografia, cabendo à Universidade Federal da Paraíba, providências, como de direito, à expedição do Diploma a que o mesmo fez jus. Nada mais havendo a tratar, foi lavrada a Ata, que segue assinada e aprovada pela Comissão Examinadora.

Documento assinado digitalmente



BARTOLOMEU ISRAEL DE SOUZA

Data: 27/11/2022 10:24:57-0300

Verifique em <https://verificador.iti.br>

Prof. Dr. Bartolomeu Israel de Souza
Orientador

Jonas Otaviano P. de Souza
Prof. Dr. Jonas Otaviano Praca de Souza

Examinador interno



Prof. Dr. Rafael Albuquerque Xavier
Examinador interno



Profª Drª. Valeria Raquel Porto de Lima
Examinadora Externa



Prof. Dr. Elpídio Inácio Fernandes Filho
Examinador externo

RAMON SANTOS SOUZA

ANÁLISE DAS PROPRIEDADES NUTRICIONAIS DO SOLO E DA ESTRUTURA
FLORESTAL NO BIOMA CAATINGA

Tese apresentada ao Programa de Pós-Graduação em Geografia do Centro de Ciências Exatas e da Natureza (CCEN) da Universidade Federal do Paraíba, como requisito parcial à obtenção do título de doutor em Geografia. Área de Concentração: Gestão do Território e Análise Geoambiental

Aprovada em:

BANCA EXAMINADORA

Prof. Dr. Bartolomeu Israel de
Souza (Orientador)
Universidade Federal do Paraíba (UFPB)

Prof. Dr. José João Lelis Leal de
Souza (Coorientador)
Universidade Federal do Paraíba (UFPB)

Prof. Dr. Eduardo Rodrigues Viana de Lima
Universidade Federal do Paraíba – UFPB

Prof. Dra. Valéria Raquel Porto de Lima
Universidade Estadual do Paraíba - UEPB

AGRADECIMENTOS

A Deus, a inteligência suprema, causa primária de todas as coisas. A Jesus, que é o caminho, a verdade e a vida. Agradeço a toda criação.

Ao amor de meus pais Maria da Penha dos Santos Souza e João de Souza Latão, por assumir a responsabilidade de minha tutela na Terra e acreditarem no meu potencial, por estarem ao meu lado em todos os momentos, auxiliando minha evolução, formação moral e intelectual. Aos meus irmãos Ricardo dos Santos Souza e Rafael dos Santos Souza.

À minha avó biológica Helena (“in memória”), avô Antônio Firmino (“in memórian”), avô Antônio Praxedes (“in memórian”), avô Samuel (“in memórian”), avó Helena Pereira, avó Maria do Socorro (Desaparecida!), tia-avó Maria Tibuço (“in memórian”), tia Maria Firmino (“in memórian”). E demais queridos familiares.

Ao meu querido amigo e Prof. Dr. Carlos Antonio Belarmino Alves (“in memória”), que ficou marcado em minha vida.

A Auricélia Alustal, esposa de Prof. Carlos Belarmino, que gentilmente me doou uma parte da biblioteca de professor Carlos, que serviu para a construção dessa tese.

Ao Programa de Pós-Graduação em Geografia da Universidade Federal da Paraíba - UFPB - Campus I.

A FAPESQ - Fundação de Apoio à Pesquisa do Estado da Paraíba, pelo financiamento da pesquisa e bolsa concedida ao estágio na Superintendência de Administração do Meio Ambiente (SUDEMA).

Aos meus colegas e amigos do trabalho na Universidade Estadual da Paraíba e SUDEMA.

Ao meu amigo e orientador Prof. Dr. Bartolomeu Israel de Souza, pela confiança em meu potencial. Sua sabedoria, sugestões e críticas foram essenciais para meu aperfeiçoamento profissional e humano.

Ao meu coorientador Prof. Dr. José João Lelis Leal de Souza pelo auxílio dado a toda construção da pesquisa.

A Profa. Dra. Valéria Raquel Porto de Lima (UEPB) por aceitar a fazer parte da banca examinadora como membro externo e suas contribuições para a concretização desse trabalho.

Ao Prof. Dr. Rafael Albuquerque Xavier (UFPB) por aceitar a fazer parte da banca examinadora como membro interno e suas considerações para a construção desse trabalho.

Ao Prof. Dr. Jonas Otaviano Praça de Souza por aceitar avaliar essa pesquisa.

Ao Prof. Dr. Elpídio Inácio Fernandes Filho pelas considerações para melhoria da pesquisa.

Ao Laboratório de Estudos do Semiárido - LAESA, meu sincero obrigado a todos os seus integrantes, que me ajudaram nesse período.

Ao meu amor e companheira de todas as horas, Simone da Silva.

A todos que fizeram de Turma de doutorado 2019.1 da UFPB, pelos momentos de descontração, alegria e dificuldades que passamos juntos. Aos professores do PPGG, pela dedicação em compartilhar seus conhecimentos.

Aos meus grandes amigos da UEPB Estevão Ferreira, Wellington Dantas, Joseline Silva, Amanda Azevedo, Suziane, Clemilson, entre outros, que fizeram parte da minha vida, que permaneceram ao meu lado sempre nessa caminhada.

A União Espírita Fabiano de Cristo do município de Cuitegi-PB e seus trabalhadores. Ao Núcleo Espírita Amigo da Paz (Guarabira-PB), Instituição Espírita Beneficente Allan Kardec (Guarabira-PB) e o Centro Espírita Vinha de Luz (Caiçara-PB), pelo trabalho caridoso e amor dedicado aos irmãos necessitados.

Por todas as pessoas que contribuíram direta e indiretamente na concretização desse sonho, os meus mais sinceros agradecimentos!

[...]

Se avexe não

Amanhã pode acontecer tudo, inclusive nada

Se avexe não

A lagarta rasteja até o dia em que cria asas

Se avexe não

Que a burrinha da felicidade nunca se atrasa

Se avexe não

Amanhã ela para na porta da sua casa

Se avexe não

Toda caminhada começa no primeiro passo

A natureza não tem pressa, segue seu compasso

Inexoravelmente chega lá

Se avexe não

Observe quem vai subindo a ladeira

Seja princesa ou seja lavadeira

Pra ir mais alto, vai ter que suar

[...]

(Flávio José, Composição: Acioli Neto. **A**

Natureza Das Coisas.)

RESUMO

Os nutrientes do solo e a estrutura das florestas desempenham papel fundamental nos ecossistemas terrestres e são primordiais para compreender os efeitos das mudanças globais. Na Caatinga ainda são escassos estudos que buscaram avaliar os controladores da variabilidade espacial dos C, N, P no solo e, menos ainda os que exploraram o uso da modelagem para estimar parâmetros estruturais da vegetação no espaço. Este trabalho tem como objetivo analisar teores de C, N, C:N e P em solos e parâmetros estruturais de altura e densidade da vegetação no bioma caatinga. Primeiramente, foi realizada uma análise descritiva dos dados. Após isso, utilizou-se de variáveis ambientais como entradas no modelo Random Forest. A validação foi realizada com as métricas e validação cruzada de 10 repetições (divididos em dados treino (75%) e teste (25%)) para os nutrientes e para os dados GEDI (rh98, cover e pai) dividiu as amostras em conjuntos de treino e teste para 353 grids ($0,5^\circ \times 0,5^\circ$) criados na área de estudo. O resultado descritivo apresenta os seguintes valores de média: $5,68 \pm 8,76$ g/kg de teor de carbono, $0,62 \pm 0,47$ g/kg para o teor de nitrogênio, $7,51 \pm 3,19$ de relação C:N e $13,79 \pm 46,95$ mg/kg para o fósforo. Os estratos Arbustivo Alto e Arbóreo Intermediário tiveram os melhores valores para C, N, C:N e P. Esses estratos são bons representativos da condição de caatinga preservada. Não foi possível identificar um padrão em relação aos nutrientes do solo e cobertura do dossel. No geral, as métricas de validação demonstraram uma moderada a baixa precisão dos modelos preditivos do solo. A altura do dossel modelado varia de 2,8 a 17,7 m e valor médio de $5,6 \pm 1,7$ m. A cobertura de dossel total modelada foi de 0% (sem cobertura de dossel presente) a 87% (cobertura de dossel densa a fechada) e média de $13 \pm 0,08\%$, o que destaca os níveis bastante esparsos das condições de cobertura do dossel na Caatinga. Já para o índice de área total da planta apresenta valor médio de $0,36 \pm 0,33$ m²/m². Portanto, esses resultados são essenciais para criação de políticas ambientais e desenvolvimento de novos estudos e, principalmente contribuir com entendimento de um dos biomas mais ameaçados e menos estudado do Brasil.

Palavras-chave: Ciclos biogeoquímicos. Estrutura da Vegetação. Covariáveis ambientais. Google Earth Engine.

ABSTRACT

Soil nutrients and forest structure play a fundamental role in terrestrial ecosystems and are essential for understanding the effects of global changes. In the Caatinga, there are still few studies that seek to evaluate the controllers of the spatial variability of C, N, P in the soil, and even less those that explore the use of modeling to estimate structural parameters of vegetation in space. This work aims to analyze C, N, C:N and P contents in soils and structural parameters of vegetation height and density in the caatinga biome. First, a descriptive analysis of the data was carried out. After that, environmental variables were used as entries in the Random Forest model. Validation was performed with metrics and cross-validation of 10 repetitions (divided into training data (75%) and test (25%)) for nutrients and for GEDI data (rh98, cover and f040) divided the samples into training and test sets for 353 grids ($0.5^\circ \times 0.5^\circ$) created in the study area. The descriptive result shows the following mean values: 5.68 ± 8.76 g/kg of carbon content, 0.62 ± 0.47 g/kg for nitrogen content, 7.51 of C:N ratio and 13.79 ± 46.95 mg/kg for phosphorus. The Tall Shrub and Intermediate Tree strata had the best values for C, N, C:N and P. These strata are representative assets of the preserved caatinga condition. It was not possible to identify a pattern in relation to soil nutrients and canopy cover. Overall, validation metrics demonstrated moderate to low accuracy of predictive soil models. The height of the modeled canopy ranges from 2.8 to 17.7 m and an average value of 5.6 ± 1.7 m. Modeled total canopy cover ranged from 0% (no canopy cover present) to 87% (dense to closed canopy cover) and averaged $13 \pm 0.08\%$, which highlights the rather sparse levels canopy cover conditions in the Caatinga. As for the total plant area index, it presents an average value of 0.36 ± 0.33 m²/m². Therefore, these results are essential for the creation of environmental policies and the development of new studies and, mainly, contribute to the understanding of one of the most threatened and least studied biomes in Brazil.

Keywords: Biogeochemical cycles. Structure of Vegetation. Environmental covariates. Google Earth Engine.

LISTA DE FIGURAS

Figura 1 – Fluxograma da estrutura resumida da tese.	21
Figura 2 – O solo e a paisagem. As conexões do solo com os componentes da paisagem são ilustradas por fluxos de massa (setas) via planos horizontais conceituados na interface entre o dossel da vegetação e a atmosfera; interface entre atmosfera, superfície terrestre, vegetação e perfil do solo; e interface entre o perfil inferior do solo e a zona vadosa/lençol freático e o limite inferior do fluxo representando a base rochosa inalterada da zona crítica. Abreviaturas: Ca, cálcio; CH ₄ , metano; CO ₂ , dióxido de carbono; H ₂ O, dióxido de hidrogênio; K, potássio; Mg, magnésio; N, nitrogênio; NO _x , óxidos de azoto; O ₂ , alótropo de oxigênio; P, fósforo.	24
Figura 3 – Províncias estruturais da Caatinga.	34
Figura 4 – Domínios morfoestruturais da Caatinga.	35
Figura 5 – Compartimentos de relevo da Caatinga.	36
Figura 6 – Ecorregiões da Caatinga.	37
Figura 7 – Solos do bioma Caatinga.	40
Figura 8 – Regiões fitoecológicas da Caatinga.	41
Figura 9 – Estrutura em grids em folhas de 0,5° × 0,5°.	53
Figura 10 – Importância relativa das covariáveis para a previsão da altura superior do dossel (rh98) na Caatinga.	56
Figura 11 – Importância relativa das covariáveis para a previsão da cobertura total do dossel na Caatinga.	56
Figura 12 – Importância relativa das covariáveis para a previsão da índice de área total da planta na Caatinga.	57
Figura 13 – Áreas de caatinga (a) preservada [8°7'23,95"S, 36°27'1,18"O] e (b) degradada [7°39'8,60"S, 36°41'26,33"O] na bacia do Alto Curso rio Paraíba no estado da Paraíba.	58
Figura 14 – Altura superior do dossel (rh98) em m modelada para o bioma Caatinga. . .	59
Figura 15 – Cobertura total do dossel em % modelada para o bioma Caatinga.	60
Figura 16 – Índice de área total da planta em m ² /m ² modelada para o bioma Caatinga. .	61
Figura 17 – Atributos derivados do modelo digital de elevação: (a) hipsometria (m), (b) índice de posição topográfica (TPI), (c) declividade, (d) aspecto.	67

Figura 17 – (Cont.) Atributos derivados do modelo digital de elevação: (e) índice de exposição ao norte, (f) índice de exposição ao leste, (g) curvatura horizontal, (h) curvatura vertical.	68
Figura 17 – (Cont.) Atributos derivados do modelo digital de elevação: (i) índice de forma (<i>Shape Index</i>), (j) índice de umidade topográfica (TWI), (k) índice de potência de escoamento (SPI).	69
Figura 18 – Mediana histórica (1985-2020) dos índices espectrais derivados dos satélites Landsat 5, 7 e 8: (a) NDVI, (b) MSAVI2, (c) EVI, (d) CFLUX.	73
Figura 18 – (Cont.) Mediana histórica (1985-2020) dos índices espectrais derivados dos satélites Landsat 5, 7 e 8: (e) GNDVI, (f) Albedo, (g) NDWI, (h) TCW.	74
Figura 19 – Média histórica da: (a) temperatura da superfície (Ts) diária e (b) Ts da noturna (2000-2021), (c) GPP e (d) NPP (2001-2021).	75
Figura 20 – A magnitude das contribuições, expressa pela média dos valores absolutos de Shapley, foi calculada sobre o conjunto de dados de calibração para prever os atributos do solo no bioma Caatinga. Essas contribuições são uma medida da importância relativa de cada atributo na previsão dos valores dessas variáveis. Ao analisar os valores absolutos de Shapley, podemos identificar quais atributos têm um impacto mais significativo na determinação dos teores de carbono (a), nitrogênio (b), razão C:N (c) e fósforo no solo (d).	110
Figura 21 – Os valores de Shapley para cada ponto individual no conjunto de dados de calibração representam a contribuição da covariável para a previsão dos teores de carbono (a), nitrogênio (b), razão C:N (c) e fósforo no solo (d) em um determinado local. A escala de cores utilizada reflete o valor da covariável normalizada no intervalo de 0 a 1.	111
Figura 22 – Comparação entre os valores previstos e observados para os teores de carbono (a), nitrogênio (b), razão C:N (c) e fósforo no solo (d) na faixa de profundidade de 0-20 cm, utilizando validação cruzada de 10 vezes a partir de modelos de regressão de floresta aleatória.	113
Figura 23 – Teor de carbono orgânico (g/kg) para o intervalo de profundidade de 0-20 cm estimado usando a abordagem de mapeamento digital do solo a partir de dados legados de solo.	117

Figura 24 – Teor de carbono nitrogênio (g/kg) para o intervalo de profundida de 0–20 cm estimado usando a abordagem de mapeamento digital do solo a partir de dados legados de solo.	118
Figura 25 – Razão C:N para o intervalo de profundida de 0–20 cm estimado usando a abordagem de mapeamento digital do solo a partir de dados legados de solo.	119
Figura 26 – Fósforo disponível no solo para o intervalo de profundida de 0–20 cm estimado usando a abordagem de mapeamento digital do solo a partir de dados legados de solo.	120

LISTA DE TABELAS

Tabela 1 – Erros do modelo do modelo <i>Random Forest</i> -RF para cada propriedade estrutural da vegetação Caatinga. RMSE (erro quadrático médio) e R^2 (coeficiente de determinação). Valores de média e mediana por critério de erro calculados para todos os <i>grids</i> modelados.	55
Tabela 2 – Estatística descritiva das propriedades nutricionais do solo agregados em diferentes intervalos de profundidade na Caatinga.	79
Tabela 3 – Estatística descritiva do teor de carbono orgânico (g/kg) da camada superficial de solos por subordem de solo.	81
Tabela 4 – Estatística descritiva do teor de nitrogênio (g/kg) da camada superficial de solos por subordem de solo.	82
Tabela 5 – Estatística descritiva da razão C:N da camada superficial de solos por subordem de solo.	83
Tabela 6 – Estatística descritiva do fósforo disponível (mg/kg) da camada superficial de solos por subordem de solo.	85
Tabela 7 – Estatística descritiva do teor de carbono orgânico (g/kg) da camada superficial, segundo classificação de solo e província geológica.	87
Tabela 7 – Cont. Estatística descritiva do teor de carbono orgânico (g/kg) da camada superficial, segundo classificação de solo e província geológica.	88
Tabela 7 – Cont. Estatística descritiva do teor de carbono orgânico (g/kg) da camada superficial, segundo classificação de solo e província geológica.	89
Tabela 7 – Cont. Estatística descritiva do teor de carbono orgânico (g/kg) da camada superficial, segundo classificação de solo e província geológica.	90
Tabela 8 – Estatística descritiva do teor de nitrogênio (g/kg) da camada superficial, segundo classificação de solo e província geológica.	91
Tabela 8 – Cont. Estatística descritiva do teor de nitrogênio (g/kg) da camada superficial, segundo classificação de solo e província geológica.	92
Tabela 8 – Cont. Estatística descritiva do teor de nitrogênio (g/kg) da camada superficial, segundo classificação de solo e província geológica.	93
Tabela 8 – Cont. Estatística descritiva do teor de nitrogênio (g/kg) da camada superficial, segundo classificação de solo e província geológica.	94

Tabela 9 – Estatística descritiva da razão C:N da camada superficial, segundo classificação de solo e província geológica.	95
Tabela 9 – Cont. Estatística descritiva da razão C:N da camada superficial, segundo classificação de solo e província geológica.	96
Tabela 9 – Cont. Estatística descritiva da razão C:N da camada superficial, segundo classificação de solo e província geológica.	97
Tabela 9 – Cont. Estatística descritiva da razão C:N da camada superficial, segundo classificação de solo e província geológica.	98
Tabela 10 – Estatística descritiva do fósforo disponível (mg/kg) da camada superficial, segundo classificação de solo e província geológica.	99
Tabela 10 – Cont. Estatística descritiva do fósforo disponível (mg/kg) da camada superficial, segundo classificação de solo e província geológica.	100
Tabela 10 – Cont. Estatística descritiva do fósforo disponível (mg/kg) da camada superficial, segundo classificação de solo e província geológica.	101
Tabela 10 – Cont. Estatística descritiva do fósforo disponível (mg/kg) da camada superficial, segundo classificação de solo e província geológica.	102
Tabela 11 – Estatística descritiva do carbono orgânico (g/kg) da camada superficial de solos por estratos da vegetação.	103
Tabela 12 – Estatística descritiva do teor de nitrogênio (g/kg) da camada superficial de solos por estratos da vegetação.	104
Tabela 13 – Estatística descritiva da razão C:N da camada superficial de solos por estratos da vegetação.	105
Tabela 14 – Estatística descritiva do fósforo disponível (mg/kg) da camada superficial de solos por estratos da vegetação.	106
Tabela 15 – Estatística descritiva do carbono orgânico (g/kg) da camada superficial de solos por cobertura da vegetação.	107
Tabela 16 – Estatística descritiva do teor de nitrogênio (g/kg) da camada superficial de solos por cobertura da vegetação.	107
Tabela 17 – Estatística descritiva da razão C:N da camada superficial de solos por cobertura da vegetação.	108
Tabela 18 – Estatística descritiva do fósforo disponível (mg/kg) da camada superficial de solos por cobertura da vegetação.	108

LISTA DE QUADROS

Quadro 1 – Tipos de variabilidade do solo.	25
--	----

SUMÁRIO

1	INTRODUÇÃO GERAL	19
1.1	HIPÓTESE	20
1.2	OBJETIVOS	20
1.2.1	Geral	20
1.2.2	Específicos	20
1.3	ESTRUTURA DA TESE	20
2	FUDAMENTAÇÃO TEÓRICA	22
2.1	PEDOMETRIA	26
2.1.1	Breve Histórico	26
2.1.2	Mapeamento Digital dos Solos	28
2.2	BIOMA CAATINGA	30
3	INTEGRANDO IMAGENS GEDI E SENTINEL PARA ESTIMAR PA- RÂMETROS ESTRUTURAIS DE UMA FLORESTA SAZONALMENTE SECA DO BRASIL	43
3.1	INTRODUÇÃO	43
3.2	MATERIAL E MÉTODOS	44
3.2.1	Área de Estudo	44
3.2.2	Levantamento de Dados	45
3.2.3	Covariáveis Ambientais	45
3.2.4	Treinamento, Avaliação e Predição Espacial	52
3.3	RESULTADOS E DISCUSSÃO	54
3.3.1	Avaliação de Erro	54
3.3.2	Variáveis Importantes	55
3.3.3	Atributos da Estrutura da Vegetação Modelada	57
4	MAPEAMENTO DIGITAL DAS PROPRIEDADES NUTRICIONAIS DO SOLO NO BIOMA CAATINGA	62
4.1	INTRODUÇÃO	62
4.2	MATERIAL E MÉTODOS	64
4.2.1	Área de Estudo	64
4.2.2	Procedimentos Metodológico	65

4.2.2.1	Levantamento de Dados	65
4.2.2.2	Análise Laboratorial	66
4.2.2.3	Covariáveis Preditoras	66
4.2.2.4	Treinamento, Avaliação e Predição Espacial	76
4.2.3	Análise Estatística	77
4.3	RESULTADOS E DISCUSSÃO	78
4.3.1	Estatística Descritiva dos Dados Amostrais	78
4.3.2	Caracterização Geoquímica dos Solos	79
4.3.3	Relação entre Propriedades Biogeoquímicas do Solo e Estrutura da Ve- getação	103
4.3.4	Variáveis importantes calculadas com valores SHAP	109
4.3.5	Avaliação e Precisão do Modelo	112
4.3.6	Previsão das Propriedades Nutricionais do Solo	114
5	CONSIDERAÇÕES FINAIS	121
	REFERÊNCIAS	123

1 INTRODUÇÃO GERAL

A pedosfera é a geomembrana que controla a maioria das interações biogeofísicas e químicas entre a terra, os corpos hídricos (superficiais e subterrâneas) e a atmosfera (WILDING; LIN, 2006). Visto que é através do solo que a água e solutos, bem como energia, gases, sólidos e organismos são trocados ativamente com a atmosfera, biosfera, hidrosfera e litosfera, criando assim um ambiente de sustentação da biosfera (LIN, 2010). Assim, o solo desempenha um papel importante na reciclagem de carbono, nitrogênio e outros produtos químicos para a atmosfera, hidrosfera e litosfera, além de armazenar água e descartar resíduos sólidos e líquidos (WILDING; LIN, 2006).

O solo é um componente essencial para os principais serviços ecossistêmicos que preservam e sustentam a biosfera (DINTWE; OKIN, 2018). Alguns grupos de bens e serviços ambientais dependem de processos dinâmicos que ocorrem no solo como ativos naturais (BANWART *et al.*, 2019; ROBINSON *et al.*, 2014). É o substrato da produção de alimentos e fibras obrigatórios na alimentação humana e fornece os principais recursos renováveis (TOTSCHKE *et al.*, 2010). Entre diversas funções desempenhadas pelos solos, destaca-se o sequestro de carbono que pode ocorrer naturalmente ou através do manejo, que ajudar a mitigar as mudanças climáticas e aumentar o acesso de nutrientes pelas plantas (ELBASIOUNY *et al.*, 2022).

Como observado, os solos atuam como fontes e sumidouros da maioria dos nutrientes essenciais requeridos pelos organismos. As plantas obtêm quase todos os seus nutrientes essenciais do solo e depende de sua capacidade de tamponar esses nutrientes essenciais e, de outra forma, retardar suas perdas, como a lixiviação da zona planta-raiz (STRAWN *et al.*, 2020).

Além da capacidade dos solos de sequestrar o carbono por períodos mais longos, as suas propriedades químicas e físicas determinarão a capacidade produtiva ou a taxa de crescimento das florestas e culturas agrícolas. A química do solo também determina quais atividades de manejo podem ser aplicadas para aumentar a taxa de crescimento das plantas ou se um solo contém elementos tóxicos que podem inibir o crescimento das plantas. As restrições do solo determinarão, em última análise, se a sociedade é capaz de continuar a coletar serviços ecossistêmicos e onde as culturas agrícolas podem ser cultivadas (VOGT *et al.*, 2015).

Em função disso, esta pesquisa teve como foco principal a utilização de modelagens computacionais que possa estimar espacialmente os nutrientes do solo e parâmetros estruturais da vegetação e sua relação no Caatinga. Esse é um dos biomas mais degradados do país, principalmente pelo uso inadequado da vegetação e do solo. Acredita-se que a Caatinga perdeu

metade de sua cobertura original, ocasionado pela ação antrópica para uso agropecuários, uso madeireira e consumo de lenha (SILVA *et al.*, 2017; ANTONGIOVANNI *et al.*, 2020).

1.1 HIPÓTESE

A Caatinga apresenta diversas características físicas e condições ecológicas que influenciam as propriedades do solo e da vegetação. A modelagem utilizando *machine learning* pode ser uma ferramenta poderosa para uma previsão precisa do teor de C, N, C:N e P com base em abordagens robustas de baixo custo obter resultados razoáveis. Além disso, o uso de dados ópticos *LiDAR GEDI* unidos aos satélites alta resolução (*Sentinel 1* e *2*) pode auxiliar em estudos da floresta e seus resultados nos ajudará a relacionar com os nutrientes do solo no bioma Caatinga.

1.2 OBJETIVOS

1.2.1 Geral

Analisar espacialmente e desenvolver uma proposta metodológica para estimar os teores de C, N, C:N e P em solos e sua relação com os parâmetros estruturais de altura e densidades no bioma Caatinga.

1.2.2 Específicos

- Estimar os parâmetros estruturais altura e densidade da vegetação;
- Avaliar a precisão e o desempenho da modelagem usando *Random Forest* para estimar a estrutura da vegetação e os nutrientes do solo;
- Relacionar os parâmetros estruturais da vegetação com os nutrientes do solo;
- Entender quais são os fatores mais importantes que afetam a distribuição espacial dos nutrientes do solo.

1.3 ESTRUTURA DA TESE

A estrutura da tese está subdividida em cinco capítulos, dois deles são artigos resultados desta pesquisa. Sendo o Capítulo 1 formado pela Introdução, Hipótese e Objetivos.

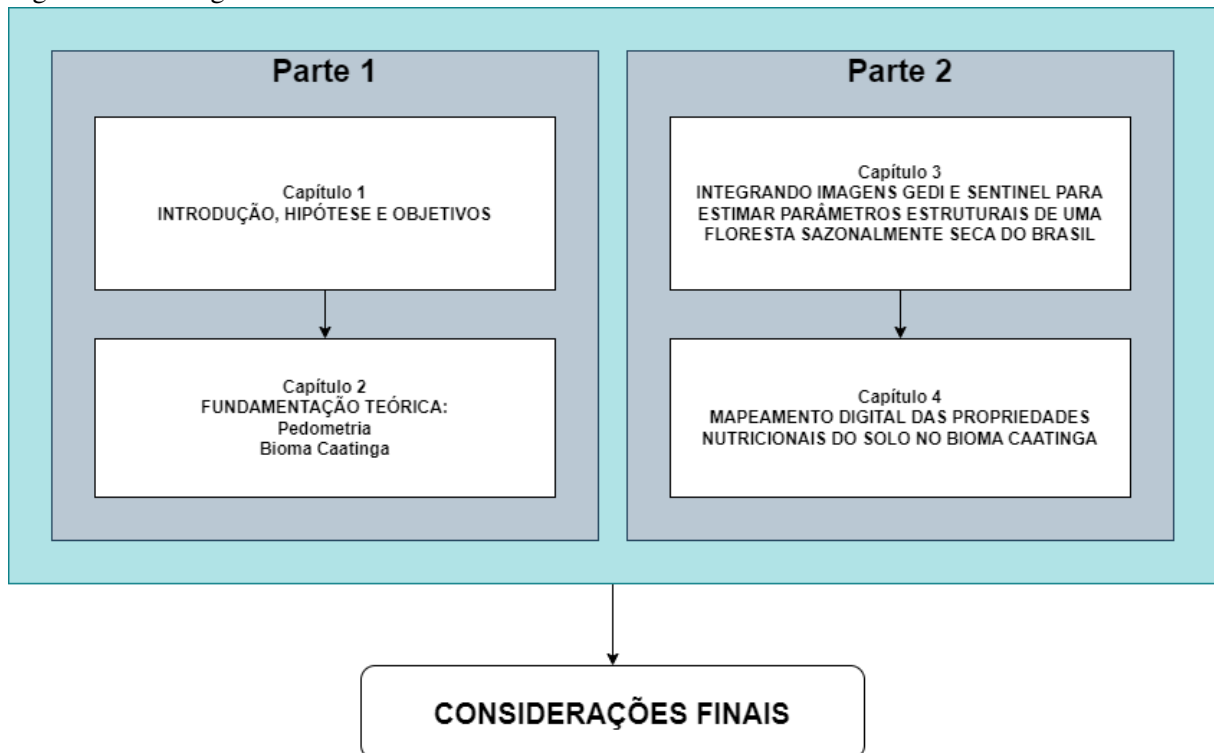
O Capítulo 2 denominado de Fundamentação Teórica, tem como foco inicial o con-

ceito e mapeamento digital de solos. Além disso, serão apresentadas considerações iniciais sobre pedometria e a caracterização do bioma Caatinga.

O Capítulo 3 apresenta o primeiro artigo denominado de ***INTEGRANDO IMAGENS GEDI E SENTINEL PARA ESTIMAR PARÂMETROS ESTRUTURAIS DE UMA FLORESTA SAZONALMENTE SECA DO BRASIL***, nos apoiamos de dados LiDAR GEDI para a estimativa de parâmetros estruturais da Caatinga, que será utilizado no Capítulo 4.

O Capítulo 4 encontra-se o segundo artigo chamado de ***MAPEAMENTO DIGITAL DAS PROPRIEDADES NUTRICIONAIS DO SOLO NO BIOMA CAATINGA*** que tem os principais resultados desta pesquisa que é a distribuição espacial dos nutrientes do solo e sua relação com os parâmetros estruturais da vegetação. E por fim, as considerações finais da Tese.

Figura 1 – Fluxograma da estrutura resumida da tese.



Fonte: Elaborada pelo autor.

2 FUDAMENTAÇÃO TEÓRICA

A humanidade sempre buscou se relacionar com os solos (BREVIK; HARTEMINK, 2010; MILLER; SCHAETZL, 2014). Desde as primeiras ocupações localizados nas bordas dos planaltos íngremes até as migrações posteriores para as planícies fluviais e lacustres mais baixas, surgiram sistemas complexos de irrigação para agricultura que permitiram o florescimento das primeiras civilizações (ARNOLD, 2006). Hoje observa-se que mais da metade dos solos da Terra já são utilizados para produção de culturas alimentícias e pastagem. Parte dos solos se encontra cobertos por infraestruturas residenciais, industriais, de transporte e usos recreativos. Além de serem alterados hidrologicamente, contaminados quimicamente e utilizados para disposição de resíduos. Assim, uma história complexa de impacto humano está se acumulando nos solos, à medida que novas mudanças se sobrepõem às do passado (RICHTER, 2007).

Neste contexto, os solos são entendidos como corpos naturais tridimensionais e dinâmicos, constituídos de materiais nos estados físicos: sólido, líquido e gasoso (CHESWORTH, 2008; EMBRAPA, 2018; STAFF, 2014). Eles são resultados da agregação de partículas orgânicas e/ou inorgânicas, que forma uma cobertura efêmera na superfície terrestre (CHESWORTH, 2008). Caracterizam-se pela presença de horizontes e/ou camadas contratantes formados das adições, perdas, transferências e transformações de energia e matéria (STAFF, 2014).

Os solos estão na interface entre a crosta e a atmosfera. Não são de interesse exclusivamente dos geólogos, cujo domínio é a terra sólida, nem pelos meteorologistas, cujo domínio é a atmosfera, nem pelos biólogos, cujo domínio são os habitantes vivos da Terra. No entanto, as preocupações de todas essas disciplinas se sobrepõem onde há solos. As rochas se desintegram para se tornar solo, e o ar e a água meteórica (definidos no glossário) ocupam grande parte dos solos, e as raízes das plantas e um grande zoológico de microorganismos e pequenos animais ocupam os solos. O solo e seus organismos estão constantemente trocando gases com a atmosfera. Assim, os solos estão nos domínios da geologia, meteorologia, hidrologia e biologia, mas nenhuma dessas disciplinas inclui todos os aspectos dos solos. A ciência dos solos é uma disciplina separada, mas não independente das outras disciplinas. (ALEXANDER, 2013).

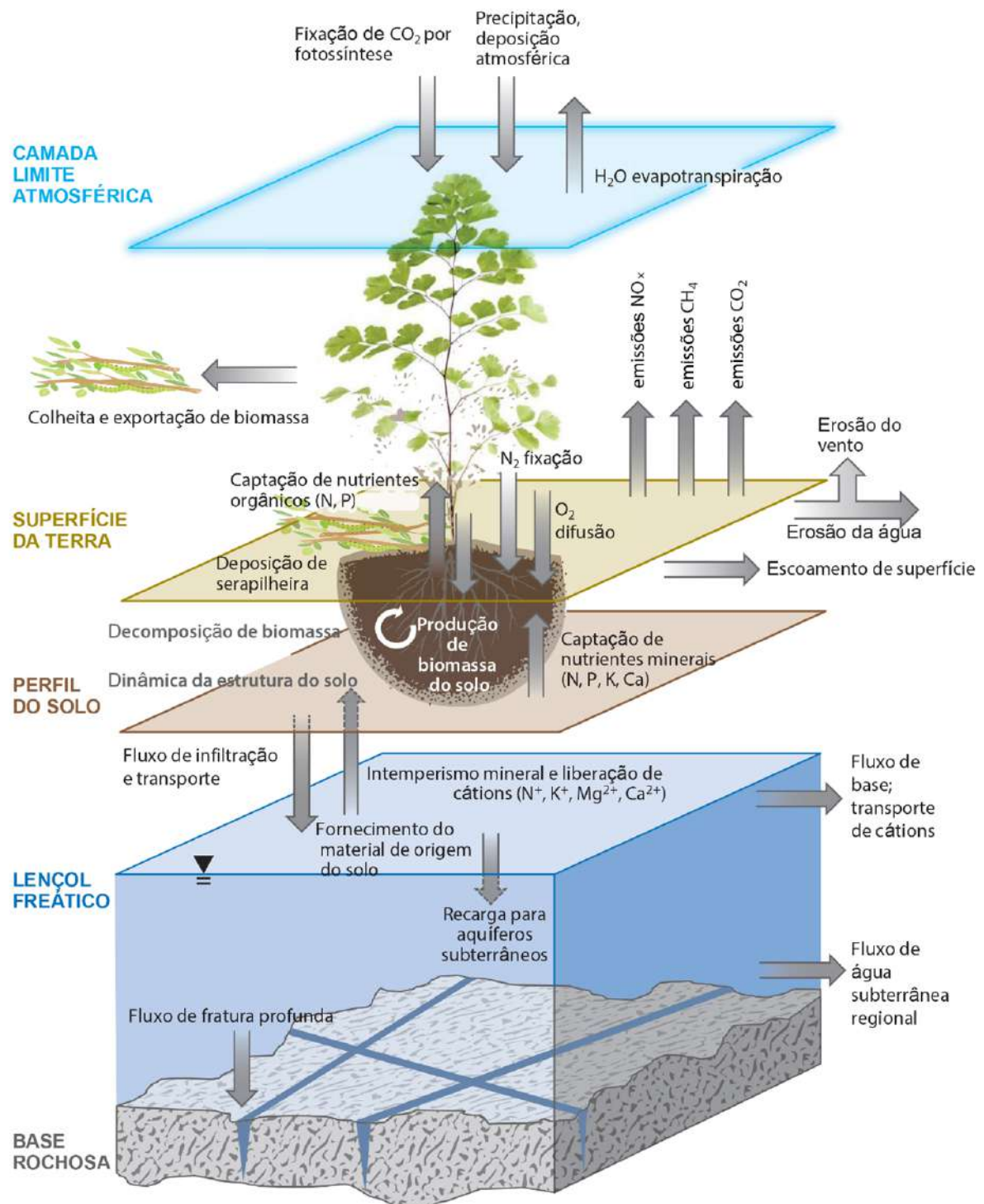
Por exemplo, entre as partículas de materiais sólidos do solo existem espaços vazios responsáveis pela passagem e preenchimento de: (a) ar – os mesmos gases da atmosfera terrestre, com suas proporções modificadas pelas reações químicas e bioquímicas decorrentes da interação água-mineral, respiração e decomposição orgânica; (b) água – apresenta-se em forma de solução aquosa de complexos inorgânicos e orgânicos, moléculas e íons; (c) micror-

ganismos – vivos e/ou mortos, dos quais a biomassa microbiana é de fundamental importância (BRADY; WEIL, 2013; CHESWORTH, 2008; OSMAN, 2013).

Neste contexto, o solo (pedosfera) é a interface central da Zona Crítica da Terra (ZCT) (BANWART *et al.*, 2019) – uma fina camada superficial do planeta que se estende desde o substrato rochoso inalterado até o dossel da vegetação e a camada limite atmosférica, faixa onde a litosfera, hidrosfera, biosfera e atmosfera se cruzam e interagem mutuamente (NRC, 2001). Tais interrelações determinam a disponibilidade de quase todos os recursos que sustentam a vida no planeta (WILDING; LIN, 2006).

Outra característica fundamental do solo é sua capacidade de atuar como uma memória da evolução da paisagem. O solo, como um corpo natural relativamente imóvel e formado *in situ*, faz que ele incessantemente reaja, suporte e registre as mudanças ambientais ocorridas no espaço, agindo como um “arquetipo” que guarda as heranças da dinâmica pretérita e atual da biosfera-geosfera. Essa memória assume múltiplas formas, como micromorfologia, características de perfil e várias propriedades físicas, químicas e biológicas do solo (LIN, 2010). Portanto, o solo dever ser compreendido como um componente, síntese da paisagem.

Figura 2 – O solo e a paisagem. As conexões do solo com os componentes da paisagem são ilustradas por fluxos de massa (setas) via planos horizontais conceituados na interface entre o dossel da vegetação e a atmosfera; interface entre atmosfera, superfície terrestre, vegetação e perfil do solo; e interface entre o perfil inferior do solo e a zona vadosa/lençol freático e o limite inferior do fluxo representando a base rochosa inalterada da zona crítica. Abreviaturas: Ca, cálcio; CH₄, metano; CO₂, dióxido de carbono; H₂O, dióxido de hidrogênio; K, potássio; Mg, magnésio; N, nitrogênio; NO_x, óxidos de azoto; O₂, alótropo de oxigênio; P, fósforo.



Fonte: Adaptada de Banwart *et al.* (2019).

Os atributos do solo apresentam variabilidade na dimensão horizontal e vertical, mudam ao longo do tempo devido à atuação de processos pedológicos, que podem ser acelerados ou não pelas ações humanas (GRUNWALD *et al.*, 2006; OSMAN, 2013; PEREIRA, 2020). Portanto, os solos são susceptíveis às transformações estabelecidas pelos ambientes em que estão inseridos (PEREIRA, 2020). Tais variações dos solos podem ocorrer de forma gradual ou abrupta em todos os níveis de observação da paisagem, uma vez que os fatores formadores do solo (clima, organismos, relevo, material de origem e tempo) e os processos pedogenéticos também variam nesta mesma paisagem (JOHNSON, 1963; WARRICK, 1998). Mais frequentemente, as variações pedológicas são graduais (SCHAETZL; ANDERSON, 2005). O Quadro 2 apresenta uma síntese conceitual dos tipos de variabilidades encontradas nas disposições espaço-temporal do solo, que podem ser de curto a longo alcance, de natureza multivariada e ter variação sistemática ou aleatória (SHUKLA, 2014).

Quadro 1 – Tipos de variabilidade do solo.

Tipo	Conceito
Sistemática	Indica uma mudança gradual nas propriedades do solo nas paisagens e pode ser devido aos fatores intrínsecos da formação do solo, elementos geomórficos e formas de relevo; além de fatores extrínsecos associados ao uso e manejo da terra.
Aleatória	Baseadas em fatores intrínsecos de formação do solo que são difíceis de explicar.
Intrínseca	Causada pela pedogênese e geralmente ocorre em grandes escalas de tempo.
Regionalizada	Causada pelos processos pedogenéticos, com áreas próximas consideradas mais semelhantes do que áreas mais distantes.
Extrínseca	Pode ter efeito relativamente rápido e não pode ser tratada como regionalizada.

Fonte: Adaptada de Shukla (2014).

2.1 PEDOMETRIA

O termo “pedometria” é um neologismo cunhado por McBratney (1986) com base na derivação das raízes gregas:

- πεδου “**PEDON**”: o chão, terra, solo;
- μετρου “**METRON**”: medição.

Essa definição abrange o caráter quantitativo incorporado à ciência do solo por meio dos métodos matemáticos, estatísticos e numéricos. Dessa maneira, a pedometria trata de pedologia quantitativa ou de métodos quantitativos aplicados à pedologia no estudo da distribuição e gênese do solo, bem como suas propriedades e processos variam no espaço-tempo (SANTOS; MANZATTO, 2007; MINASNY *et al.*, 2014; MCBRATNEY *et al.*, 2018).

Neste contexto, a pedometria está intimamente relacionada com a pedologia, buscando estudar o solo como parte do ambiente natural e como um recurso sustentável. Mas se concentra em problemas que podem ser formulados quantitativamente e que podem ser resolvidos com técnicas matemáticas e estatísticas (MINASNY *et al.*, 2014). Portanto, é essencialmente a aplicação de probabilidade e estatística ao solo, lidando com a incerteza e complexidade, devido à variação determinística ou estocástica, imprecisão e falta de conhecimento das propriedades e processos do solo (MCBRATNEY *et al.*, 2018). A seguir na subseção 2.1.1 faremos um breve histórico da pedometria.

2.1.1 Breve Histórico

A quantificação tem sido aplicada a estudos de solo desde, pelo menos, a década de 1960 com a disponibilidade de computadores e softwares (MCBRATNEY *et al.*, 2018). A pedometria auxiliou na criação de experimentos, estudos e projetos por várias décadas do século XX, embora não fosse reconhecida nem bem definida (MCBRATNEY *et al.*, 2018). Nos anos 1960, os estudos pedométricos se concentraram no problema da classificação dos solos e aplicaram os métodos da taxonomia numérica. No final da década de 1970, começaram a tratar as propriedades do solo como processos aleatórios correlacionados espacialmente e utilizaram a geoestatística para suas análises e previsões (MCBRATNEY *et al.*, 2018). Assim, suas origens não remontam estritamente a pedologia, mas a um esforço coletivo e avanços de diversos campos do saber no início do século XX (MINASNY *et al.*, 2014; MCBRATNEY *et al.*, 2018).

Na seção 2.1 mencionamos que o termo pedometria foi cunhado em meados da década de 1980 e desde então foi formalmente reconhecido como um campo da ciência do solo.

Na década de 1990, os pedometristas começaram a experimentar técnicas matemáticas e estatísticas contemporâneas, como conjuntos fuzzy, fractais, wavelets e técnicas de mineração de dados para a descrição e previsão das propriedades do solo (MINASNY *et al.*, 2014).

Mais recentemente, além desses temas anteriores, a pedometria começou a tentar elucidar a pedogênese, quantificando as relações entre as propriedades individuais do solo e os fatores de controle (MCBRATNEY *et al.*, 2018).

Atualmente, a aplicação é amplamente focado no Mapeamento Digital de Solos (MDS) tratado aqui no tópico 2.1.2, para produzir mapas de propriedades do solo eficientemente. A pedometria foi proposta pela primeira vez como um grupo de trabalho dentro da Sociedade Internacional de Ciências do Solo em 1987. O grupo de trabalho foi formalmente estabelecido no Congresso Mundial de Ciência do Solo em Kyoto, Japão, 1990. Posteriormente, foi estabelecido como uma comissão da União Internacional of Soil Sciences em 2004. Três grupos de trabalho foram posteriormente formados a partir de pedometria: DSM, sensoriamento proximal do solo e monitoramento do solo. Para uma revisão abrangente sobre os primeiros estudos de pedometria (MINASNY *et al.*, 2014)

Embora a modelagem geoestatística da variação do solo e, posteriormente, o mapeamento digital do solo (MDS) tenham sido o foco principal dos pedometristas por quase três décadas, o MDS é apenas uma faceta da pedometria. A pedometria também abrange, entre outros esforços, o estudo de projetos de amostragem e sua otimização, a integração de modelos de processo com observações, a quantificação da incerteza na saída do modelo, o uso de métodos matemáticos para elucidar a variação do solo em múltiplas escalas e novas abordagens. (WADOUX *et al.*, 2021)

A pedometria está atualmente na mesma situação que muitos outros campos científicos em rápido crescimento, alimentados pelo rápido desenvolvimento da disponibilidade de dados digitais e novas ferramentas e técnicas em estatística, aprendizado de máquina e sistemas de informação (WADOUX *et al.*, 2021). Portanto, a pedometria contribuiu e ganhou com a estatística aplicada, geoestatística, SIG, amostragem ambiental e geomorfometria (MCBRATNEY *et al.*, 2018). Na próxima subseção 2.1.2 será explanado sobre uma das aplicabilidades mais relevantes da pedometria, o Mapeamento Digital dos Solos (MDS).

2.1.2 Mapeamento Digital dos Solos

A conexão entre pedosfera, biosfera e atmosfera sempre fez parte do imaginário das pessoas, se tornando objeto de investigação empírica (SILVA; LAMBERS, 2021). Assim sendo, ao longo do tempo foram elaborados modelos científicos para auxiliar o entendimento ou explicação da complexidade solo-paisagem no continuum espaço-tempo (HUANG *et al.*, 2012). Tais modelos são abstrações que formalizam como o espaço geográfico é discretizado em unidades espaciais menores para análise e comunicação (GRUNWALD *et al.*, 2006).

Os modelos fatoriais solo-paisagem são considerados mnemônicos para o procedimento de modelagem e mapeamento dos solos. Esses modelos foram inspirados por observações efetuadas pelo russo Dokuchaev e colaboradores durante a última parte do século XIX. Após cobriram áreas extensas na Rússia, reconheceram que os solos variam onde o material de origem, clima, relevo, drenagem e vegetação diferem (ALEXANDER, 2013; OSMAN, 2013).

Jenny (1994) popularizou esse entendimento na Equação 2.1, onde um solo (s) é produto da interação entre o clima (cl); organismo (o); relevo (r); material parental (p) e o tempo (t):

$$s = f (cl', o', r', p', t') \quad (2.1)$$

Similar ao clorpt, mais tarde o modelo scorpan (MCBRATNEY *et al.*, 2003) formulado na Equação 2.2 incorporou a dimensão espacial, explícita nos procedimentos operacionais da modelagem preditiva e mapeamento das propriedades/classes do solo.

$$S = f (s', c', o', r', p', a', n') + e \quad (2.2)$$

Em que o tipo ou atributo do solo em um local não visitado (S) pode ser previsto a partir de uma função ou modelo numérico (f), dados os fatores:

- s' : classes ou propriedades (auxiliares) do solo;
- c' : clima;
- o' : organismos, vegetação, fauna ou atividade humana;
- r' : relevo;
- p' : material parental;
- a' : idade, isto é, o fator tempo;

- n' : espaço, variação espacial;
- e : erro, que pode incluir as tendências espaciais que não são capturadas pelos fatores preditivos, erro estocástico espacialmente autocorrelacionado e ruído.

O MDS é considerada uma abordagem orientada a dados, que busca automatizar a construção de modelos analíticos e simulações numéricas para inferir as variações espaço-temporais dos tipos e propriedades dos solos a partir das relações entre as variáveis-alvo (amostras de campo geolocalizadas) e as covariáveis ambientais relacionadas (HENGL; MACMILLAN, 2019; LAGACHERIE; MCBRATNEY, 2006). Para tanto, se utiliza modelos estatísticos lineares simples, técnicas geoestatísticas, híbridas e avançadas/complexas para entender a partir de dados (baseados no modelo scorpan) (LAMICHHANE *et al.*, 2019; LIAKOS *et al.*, 2018; PADARIAN *et al.*, 2020).

Além disso, o MDS não se concentra apenas no mapeamento, mas também busca responder questões ambientais, como degradação da terra e mudanças climáticas globais, que exigem uma avaliação dos solos no contexto de mudanças no ecossistema e estressores ambientais (GRUNWALD, 2010).

Atualmente o MDS tem sido aliado a aprendizagem de máquina (ML na sigla em inglês), que é entendida como uma estrutura algorítmica unificada projetada para identificar modelos computacionais que descrevem com precisão dados empíricos e os fenômenos subjacentes a eles (WATT *et al.*, 2020).

ML é uma estrutura algorítmica unificada projetada para identificar modelos computacionais que descrevem com precisão dados empíricos e os fenômenos subjacentes a eles, com pouco ou nenhum envolvimento humano (WATT *et al.*, 2020). Trataremos a seguir de alguns algoritmos que já foram testados e empregados na predição de propriedades contínuas do solo.

Neste contexto, a aplicabilidade dos algoritmos de ML no MDS cresceu rapidamente na última década (PADARIAN *et al.*, 2020). Pois, elas conseguiram trabalhar com volumes grandes de variáveis e casos, formando uma análise contínua da variabilidade do solo, além de lidar melhor com relações não lineares complexas presentes nos dados de entrada (LAMICHHANE *et al.*, 2019; LIAKOS *et al.*, 2018; PADARIAN *et al.*, 2020). Neste contexto, estudos recentes em todo o mundo demonstra o potencial desses algoritmos para prever propriedades nutricionais dos solos (CHEN *et al.*, 2019; HOUNKPATIN *et al.*, 2018; ??; MAHMOUDZADEH *et al.*, 2020; SILATSA *et al.*, 2020; TAJIK *et al.*, 2020; WANG *et al.*, 2018; ZHOU *et al.*,

2020a; ZHU *et al.*, 2019).

2.2 BIOMA CAATINGA

A Caatinga está inserida na região Nordeste do Brasil (NEB) e em pequena parcela da região Sudeste (IBGE, 2019; MORO *et al.*, 2014; AB'SABER, 1977; PRADO, 2003). Constitui-se por uma área territorial de 862.818 km² (IBGE, 2019), correspondendo 10% do território brasileiro. Representa 54% do NEB, abarcando grandes áreas dos estados do Ceará (CE), Alagoas (AL), Bahia (BA), Maranhão (MA), Paraíba (PB), Pernambuco (PE), Piauí (PI), Rio Grande do Norte (RN) e Sergipe (SE). Além disso, o bioma também está presente em parte do norte do estado de Minas Gerais (MG), na região Sudeste do País (GANEM *et al.*, 2020).

A Caatinga apresenta uma ampla diversidade de ambientes conforme as formas e característica do relevo. Pode apresentar do estrato arbustivo ao arbórea, mata seca e úmida e formações com a presença de cactáceas e bromeliáceas (VELLOSO *et al.*, 2002).

Inserida entre a linha do Equador e o Trópico de Capricórnio, na Caatinga apresenta uma alta radiação solar anual (MORO *et al.*, 2016), influenciando na temperatura média que se concentra entre 25 °C e 30 °C, não existindo diferença entre os meses de inverno e verão. Quanto ao período chuvoso, grande parte da água que precipita não é armazenada devido a condição do solo pouco profundo, ocasionando alta evaporação. Quanto a umidade relativa, é baixa, com percentuais entorno de 50% (GANEM, 2017).

Inserida entre a linha do Equador e o Trópico de Capricórnio, isso proporciona a Caatinga uma alta radiação solar anual (MORO *et al.*, 2016), influenciando na temperatura média que se concentra entre 25 ° e 30 °, não existindo diferença entre os meses de inverno e verão. Quanto ao período chuvoso, grande parte da água que precipita não é armazenada devido a condição do solo pouco profundo, ocasionando altos índices de evapotranspiração. Quanto a umidade relativa, é baixa, com percentuais entorno de 50% (GANEM, 2017).

A Caatinga é um dos biomas mais degradados do país, principalmente pelo uso inadequado da vegetação e do solo. Acredita-se que a Caatinga perdeu metade de sua cobertura original, ocasionado pela ação antrópica para uso agropecuários, uso madeireira e consumo de lenha (SILVA *et al.*, 2017; ANTONGIOVANNI *et al.*, 2020).

A hipsometria da área de estudo, que apresenta principalmente elevações abaixo de 600 m. Como menciona Velloso et al. (2002), as formas de relevo são predominantes às superfícies aplainadas que variam de 5% a 20%. Porém, a Caatinga se apresenta uma ampla

diversidade de ambientes conforme as formas e característica do relevo. Pode-se apresentar do estrato arbustivo ao arbórea, mata seca e úmida e formações com a presença de cactáceas e bromeliáceas (VELLOSO et al., 2002).

A apresenta claramente a influência do relevo e orientação das vertentes na determinação de seus quantitativos. As chuvas mais escassas (<550 mm) se encontram sobrepostas aos patamares e depressões das províncias Borborema e São Francisco (GRAEFF, 2015). As chuvas mais abundantes (> 700 mm) estão mais próximas dos tabuleiros e planícies litorâneas, muito em função do comportamento dos sistemas atmosféricos e massas de ar que atuam nesse perímetro. Pode-se ressaltar ainda a presença de enclaves úmidos e com temperaturas mais baixas em regiões com elevações acentuadas e direcionadas para as circulações de nuvens vindas do litoral.

As temperaturas máximas variam de 30 °C e 36 °C e as mínimas entre 20 °C e 22 °C (ROSS, 2009). A variabilidade espacial das temperaturas acompanha as cotas altimétricas (Erro! Fonte de referência não encontrada.). As temperaturas inferiores a 23 graus estão situadas nos ambientes mais elevados, enquanto as mais altas na parte centro-oeste e extremo norte da área de estudo.

Conforme a classificação de Köppen (ALVARES *et al.*, 2013), a região é marcada principalmente pelo clima semiárido quente (Bsh) na parte central e mais próximo ao oceano pelo clima tropical quente e úmido (As) devido aos efeitos orográficos. Durante os meses de novembro a março, a atividade convectiva é influenciada principalmente pela instabilidade provocada pela atuação de ciclones, denominados de Vórtice Ciclônico de Altos Níveis (VCAN) originários no oceano Atlântico e sua trajetória normalmente é de leste a oeste (ALVES et al., 1996; FERREIRA; MELLO, 2005). No período de março a junho frequentemente ocorre propagação dos Complexos Convectivos de Mesoescala, que se originam de perturbações ao sul da ZCIT no oceano Atlântico. Eles costumam sofrer uma intensificação bastante acentuada quando interagem com a topografia, sobretudo com as serras da Borborema (GOMES FILHO et al., 2003).

Entre os meses de fevereiro a maio (quadrimestre mais chuvoso), têm-se a atuação da Zona de Convergência Intertropical (ZCIT), sendo o principal sistema atmosférico causador de chuvas na região (ALVES et al., 1996; FERREIRA; MELLO, 2005). As massas de ar que mais predominam na área são: a MEA (Massa de ar equatorial Atlântica), MTA (Massa de ar tropical Atlântica) e MEC (Massa de ar equatorial continental) (MENDONÇA; Dannio-

Oliveira, 2007).

Portanto, o clima predominantemente quente e semiárido, intensamente sazonal, com chuvas distribuídas concentradas em um período de três a seis meses. Os totais de chuva variam muito de ano para ano e, em intervalos de dez a vinte anos, caem a menos de metade da média, às vezes durante três a cinco anos seguidos. A maior incidência de secas ocorre ao norte do rio São Francisco do que ao sul, onde as chuvas são mais bem distribuídas ao longo da estação chuvosa (VELLOSO *et al.*, 2002).

De acordo com (IBGE, 2019), a Caatinga pode ser agrupada em quatro grandes províncias geológicas: Borborema (36,05%), São Francisco (28,85%), Cobertura Cenozoica (21,78%), Parnaíba (6,78%), Recôncavo-Tucano-Jatobá (3,43%), costeira e Margem Continental (1,73%) e Mantiqueira (0,20%) (Figura 3).

A Província Estrutural Borborema limita-se a oeste com a Bacia Sedimentar do Parnaíba, ao sul com o Cráton do São Francisco, a leste pelas bacias da margem continental e por Depósitos Quaternários Costeiros. Compreende um mosaico de remanescentes de micro continentes (crátons) e orógenos do Arqueano, Paleoproterozoico e Mesoproterozoico, amalgamados por faixas móveis brasileiras (IBGE, 2019).

A Província Mantiqueira é resultante do fechamento no Neoproterozoico, do paleo-oceano Adamastor (IBGE, 2019; SCHOBENHAUS; NEVES, 2003). Corresponde a uma faixa da região costeira localizada no sul da Bahia. Tem aproximadamente 3.000 km de extensão, 200 km de largura na parte sul e 600 km na parte norte, correspondendo a uma área de 376.078 km² do território brasileiro.

A Província Parnaíba envolve a Bacia Paleozoica do Parnaíba, as Bacias Mesozoicas Alpercatas (rifte), Grajau (sinéclise), parte da Espigão Mestre e o vulcanoplutonismo Mosquito-Sardinha (IBGE, 2019).

A Província Estrutural Recôncavo-Tucano-Jatobá tem a sua representação introduzida no rol de províncias estruturais brasileiras, tendo em vista tratar-se de uma extensa cobertura sedimentar que recobre, parcialmente, os limites de duas outras províncias estruturais: São Francisco e Borborema. Situa-se na Região Nordeste do Brasil, onde seus 43.809 km² de área exposta distribuem-se, principalmente, na Bahia, que abriga 85% da área total, seguida por Pernambuco, com 14%. Alagoas e Sergipe somados detêm menos de 1% da área dessa Província (IBGE, 2019).

A Província Estrutural São Francisco constitui de um fragmento continental antigo,

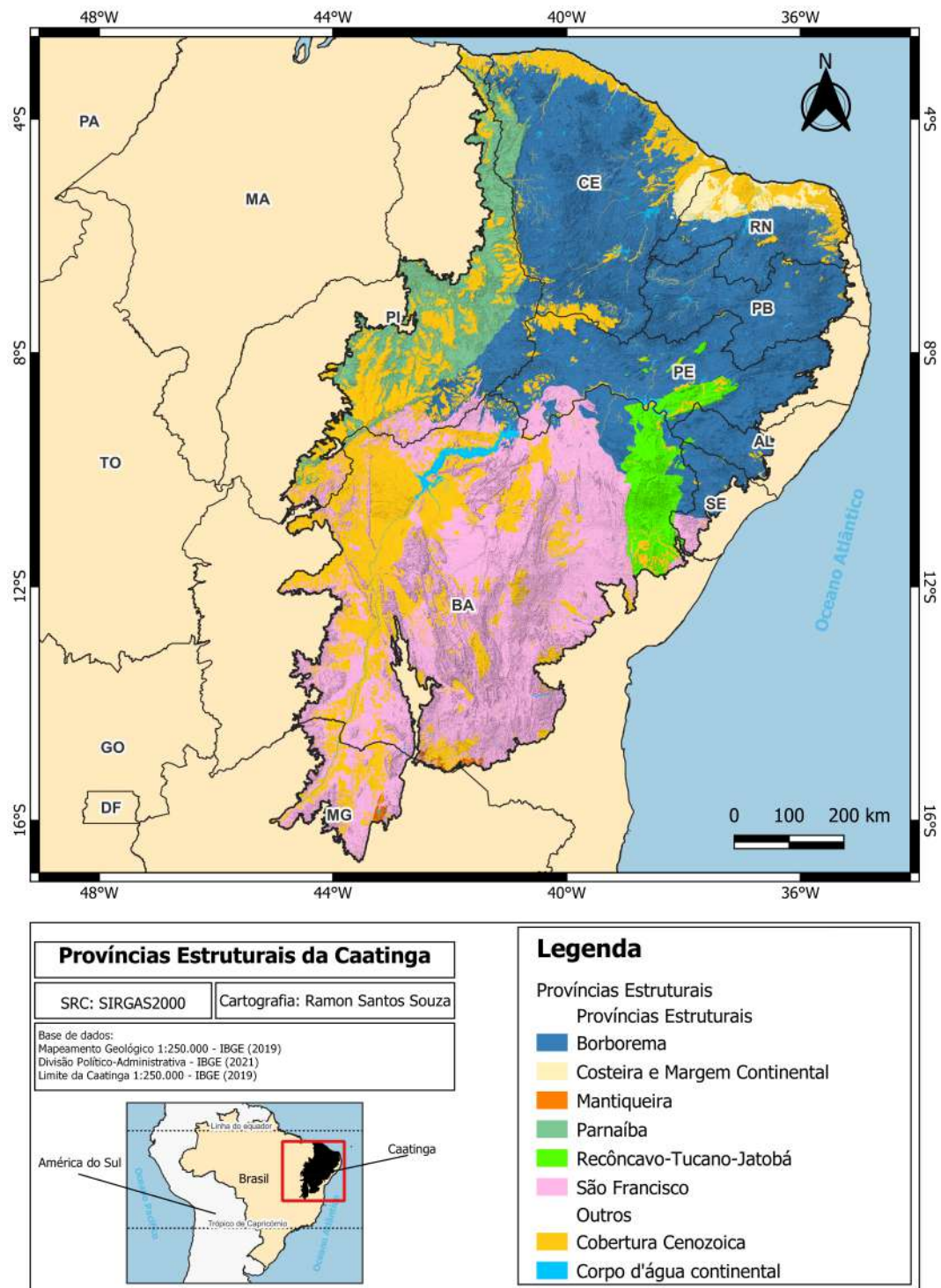
composto por núcleos do Arqueano, aglomerados pelos cinturões orogênicos do Riachiano. Os conjuntos de rochas da sua composição são remanescentes de crátons e cinturões móveis do Neoarqueano e do Riachiano (IBGE, 2019).

Já as coberturas cenozoicas estão reunidas sob essa denominação, as coberturas pós-gondwanicas continentais, são formadas pelo preenchimento das bacias e depressões continentais. Foram desenvolvidas no contexto da Plataforma Sul-Americana, durante o Cenozoico (Paleógeno, Neógeno e Quaternário), formada pelos depósitos correlativos dos eventos de aplainamento decorrentes das glaciações dessa Era, conservadas através das formações edafoestratigráficas detríticas (IBGE, 2019).

Conforme o IBGE (2019) os compartimentos do relevo são (Figura 5):

- As chapadas são feições características dos topos tabulares (planos) ou suavemente ondulados, entremeados por escarpas, geralmente, em rochas sedimentares;
- As depressões são formas feições aplainadas em relação a relevos do entorno como sendo mais elevados, geralmente elaboradas em quaisquer tipos de rochas;
- Patamares são feições planas ou onduladas, constituídas de degraus transicionais entre um planalto, chapada ou um conjunto serrano, sendo mais elevados, e comumente distribuídos entre uma depressão ou planície;
- Planaltos de características plana ou dissecada, de forma predominantemente homogênea, apresentam elevadas altitudes, apresentando ao menos em uma de suas bordas, as superfícies mais baixas;
- Planícies são feições planas ou suavemente onduladas, nas quais os processos de sedimentação excedem os de erosão, geralmente formadas em sedimentos aluvionares recentes, resultantes de acumulação fluvial, marinha e/ou lacustre, com predisposição a inundações periódicas, e/ou devido à ação eólica;
- Serras são feições de características acidentadas, constituídas em diferentes tipos de rochas, com aparência de cumeadas aguçadas e alongadas e vales estreitos cortados por vertentes de inclinações acentuadas;
- Tabuleiros são feições compostas por topos tabulares ou suavemente ondulados. Os tabuleiros diferenciam-se das chapadas por terem altitudes mais baixas e não se tratarem de formas planálticas, comumente situando-se em coberturas sedimentares inconsolidadas.

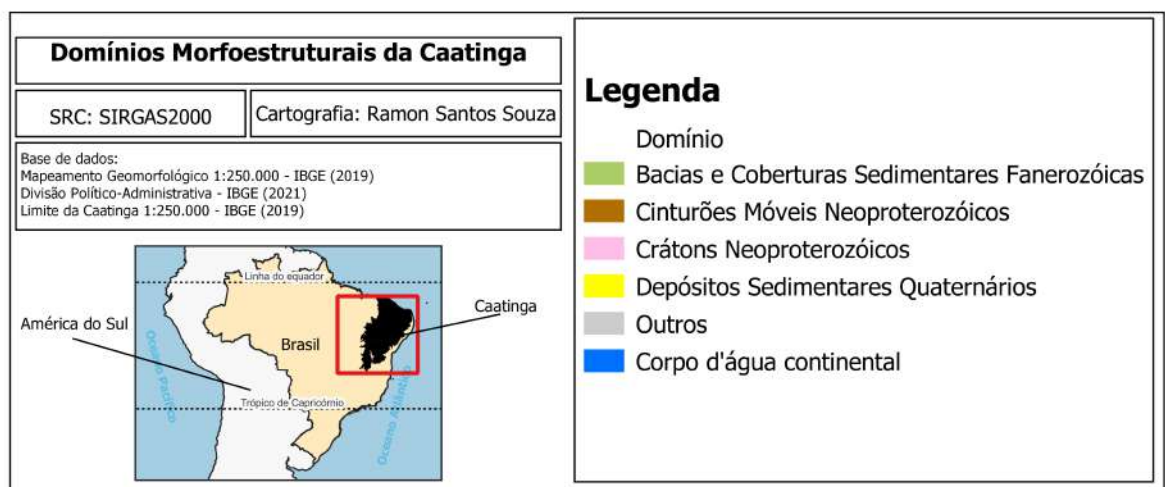
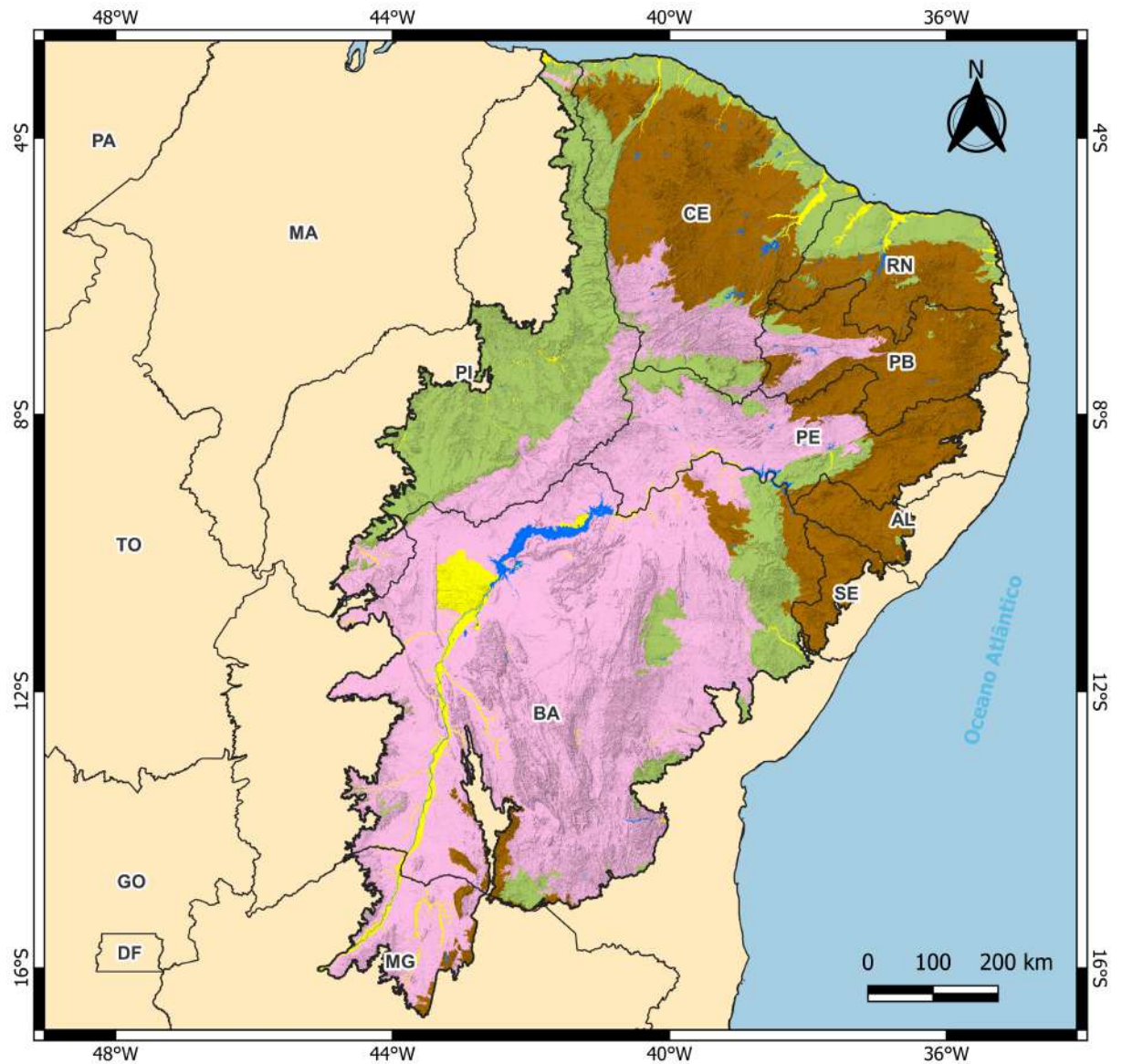
Figura 3 – Províncias estruturais da Caatinga.



Fonte: Elaborada pelo autor.

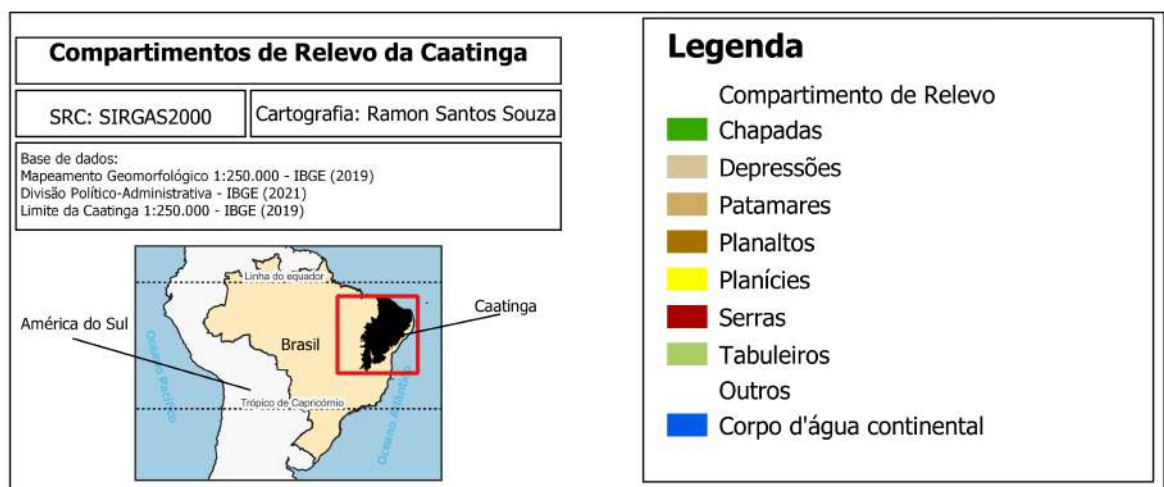
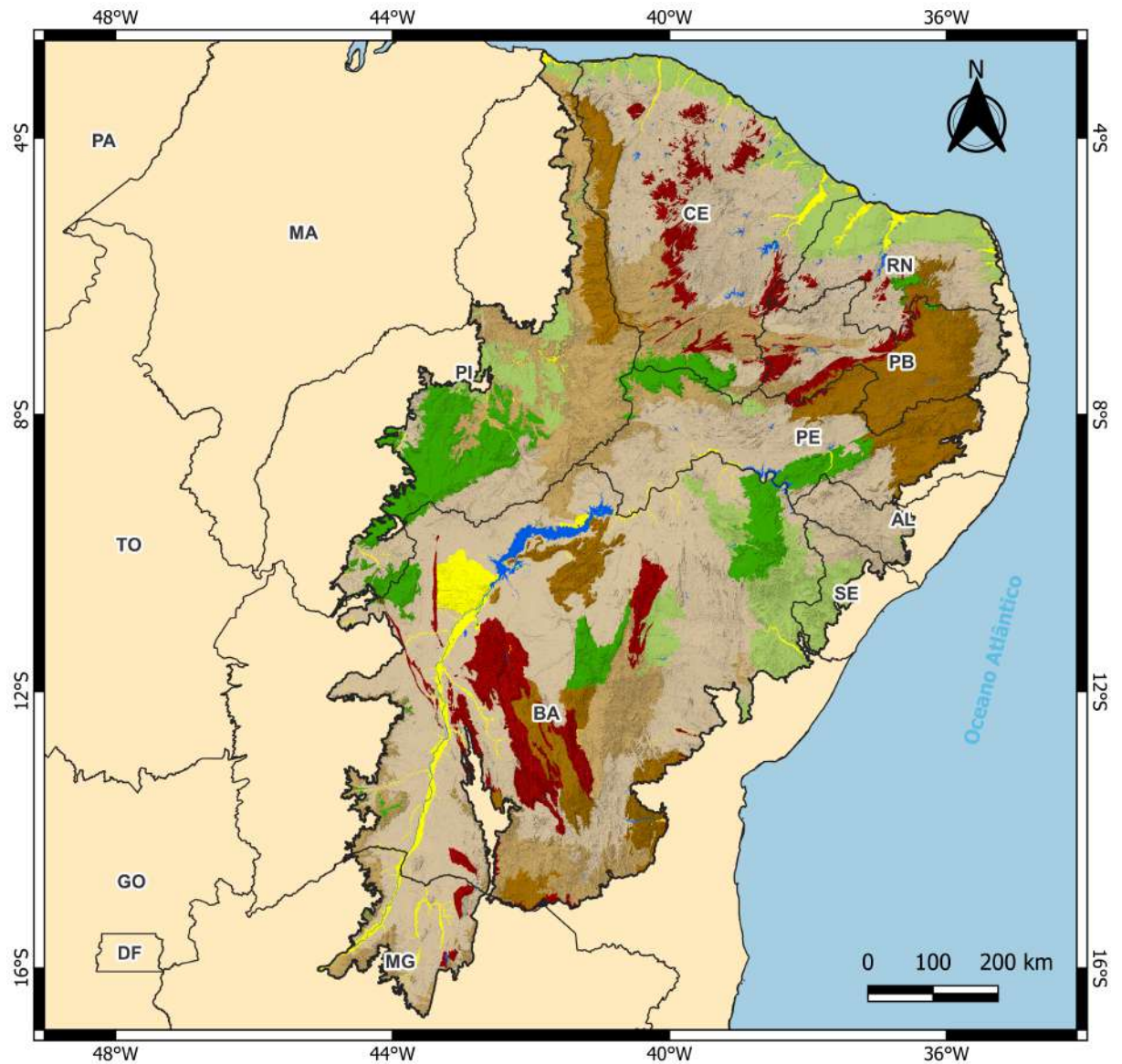
A Figura 4 demonstra os domínios morfoestruturais da caatinga, centrado principalmente pelos Cratóns neoproterozóicos, seguindo pelos cinturões móveis neoproterozóicos e as bacias e cobertura sedimentares sedimentares fanerózoicas.

Figura 4 – Domínios morfoestruturais da Caatinga.



Fonte: Elaborada pelo autor.

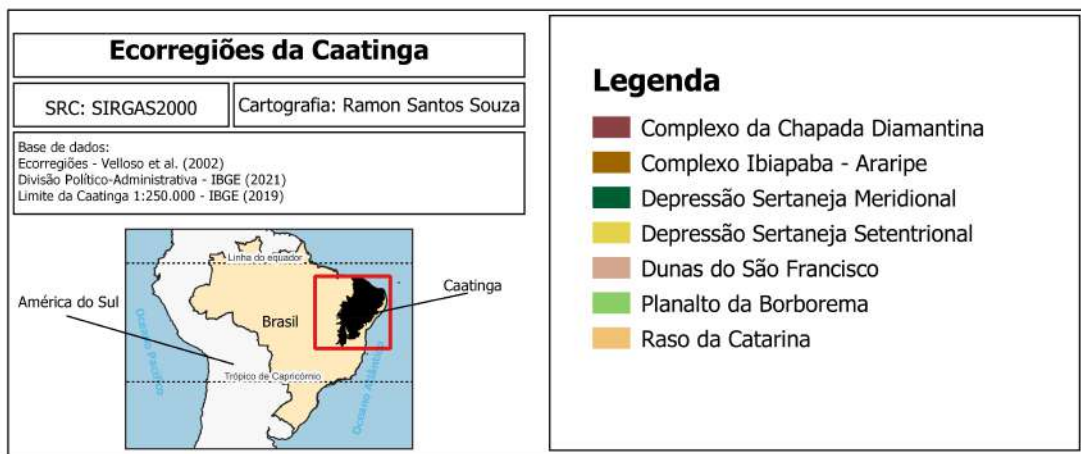
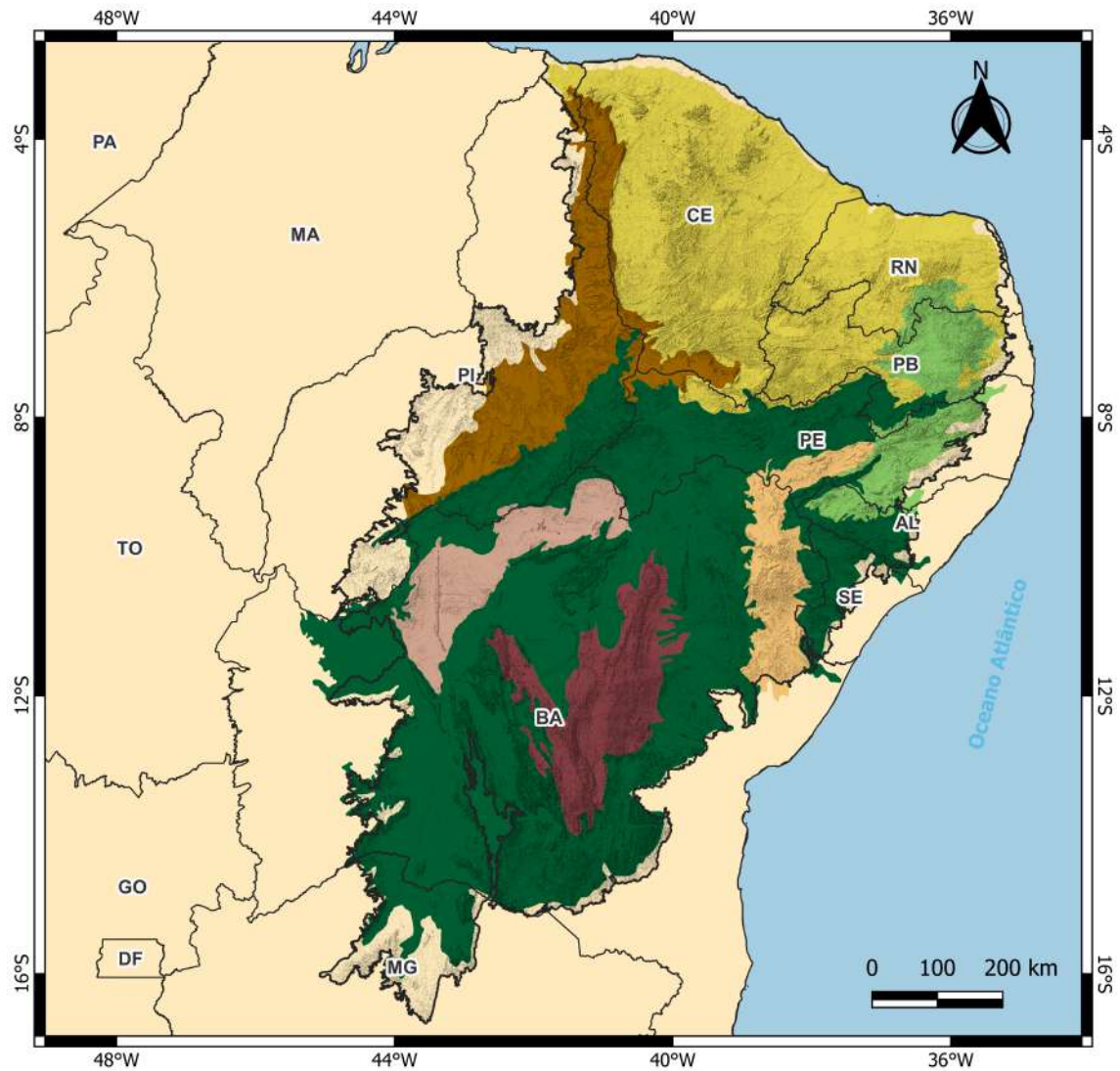
Figura 5 – Compartimentos de relevo da Caatinga.



Fonte: Elaborada pelo autor.

A Caatinga está dividida em oito ecorregiões identificadas e mapeadas por Velloso et al. (2002), como pode ser visto na Figura 6:

Figura 6 – Ecorregiões da Caatinga.



Fonte: Elaborada pelo autor.

Em descrição das ecorregiões Velosso et al., (2002) descreve dois conjuntos de paisagem características da Caatinga: “1 - superfícies de cimeira plana ou levemente colinosas, com predomínio de solos medianamente ou muito profundos, evoluídos sobre rochas sedimentares areníticas 2 - áreas deprimidas, marginais às cimeiras, ou situadas ao centro de extensas peneplanícies arrasadas, com solos variando de colúvios espessos a litólicos, evoluídos sobre granitoides, quartzitos ou calcários” (GRAEFFE, 2015, p.463).

O Complexo Ibiapaba-Araripe está localizado nas serras que compõem as chapadas do Araripe (fronteira Ceará – Pernambuco) e Ibiapaba (Ceará). Os solos (principalmente latossolos) são profundos, apresentando fertilidade baixa, geralmente arenosos e com uma boa drenagem, ocorrendo pouca água na superfície. Nos ambientes de encostas das chapadas há a presença de floresta pluvial e nos topos das chapadas concentram-se os cerradões.

A Depressão Sertaneja Setentrional está concentrada em grande parte do norte do bioma, abrangendo grande parte de Pernambuco, Paraíba, Rio Grande do Norte e Ceará e estendendo-se até a parte norte do Piauí. É limitada ao norte pelos Tabuleiros Costeiros (onde a formação geológica passa para os solos mais profundos da Formação Barreiras e a zona mais costeira tende a receber mais chuva), ao encontrar o oceano em alguns trechos do litoral do Rio Grande do Norte e Piauí).

O Planalto da Borborema é a Ecorregião mais a leste do bioma, alongada no sentido N-S em forma de arco, compreendendo partes do Rio Grande do Norte, Paraíba, Pernambuco e Alagoas. A altitude e o relevo são os fatores principais que determinam os limites da ecorregião. A parte norte do Planalto é circundada pela Depressão Sertaneja Setentrional. A leste, abaixo da Depressão Setentrional, o Planalto encontra a Zona da Mata da Paraíba, Pernambuco e Alagoas, apresentam clima seco, muito quente e semiárido. Apresenta um grande mosaico de solos que, junto com a umidade mais alta, propicia a variedade de vegetação, vão desde caatinga arbustiva aberta a arbórea (com gradações intermediárias), a matas secas e matas úmidas.

A Depressão Sertaneja Meridional ocupa a maioria do centro e sul do bioma, limítrofe com todas as outras ecorregiões, exceto a do Complexo de Campo Maior. Ao norte e noroeste encontra uma barreira de altitude (Serra dos Cariris e Complexo Ibiapaba - Araripe). A oeste faz fronteira com o início do Planalto Central, onde começa o cerrado. Ao sul e a leste encontra modificações de solo e precipitação, fazendo limite com o cerrado de Minas Gerais e a Zona da Mata da Bahia, Sergipe e Alagoas. Os solos são mais profundos que os da Depressão Sertaneja Setentrional, com latossolos (profundos, bem drenados, ácidos e com fertilidade na-

tural baixa) predominam nas partes oeste e sul. A estrutura da Caatinga é de arbustiva a arbórea, de porte mais alto que a da Depressão Setentrional.

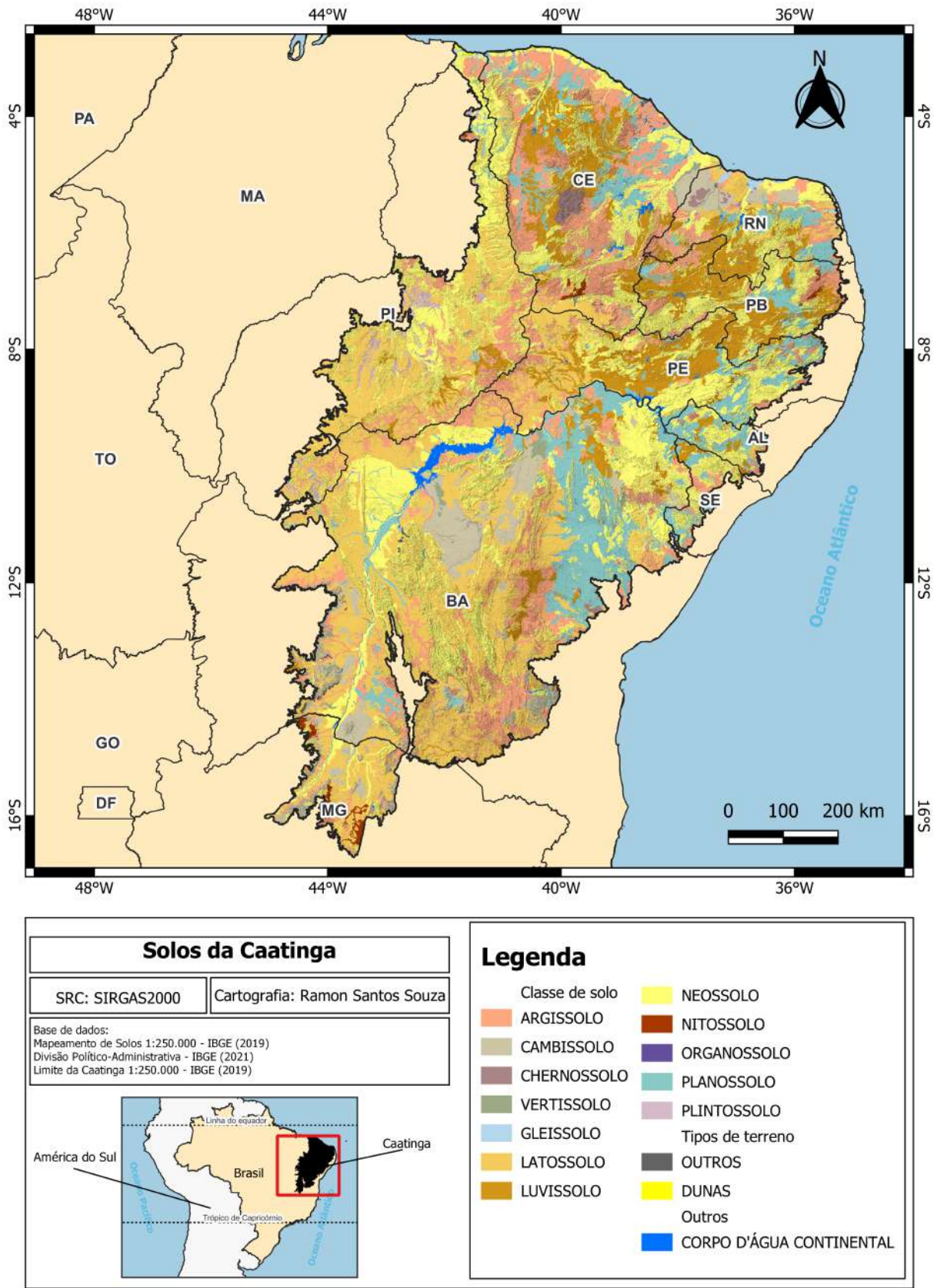
As Dunas do São Francisco ocupam a parte centro-oeste do bioma, inteiramente circundada pela ecorregião da Depressão Sertaneja Meridional. Seus limites estão na região das dunas de areias quartzosas, disjuntas, conectadas por um trecho de solos arenosos sem dunas, terminando onde os solos se modificam ao norte e oeste para latossolos. As dunas são formadas pelos ventos alísios de sudeste, que movimentam os sedimentos arenosos. A temperatura alta do solo (areia) é um fator importante, mas não há grande diferença de temperatura ambiente entre dia e noite. A Caatinga encontra-se agrupada em moitas densas, predominantemente arbustiva, geralmente com arvoretas e muita macambira (*Bromelia laciniosa*).

O Complexo da Chapada Diamantina está localizado na parte centro-sul do bioma, alongada no sentido N-S e em forma de “Y”, seguindo o alinhamento do divisor de águas da Chapada Diamantina. É inteiramente circundada pela ecorregião da Depressão Sertaneja Meridional. Os limites são explicados principalmente pelas mudanças de relevo, altitude e tipo de solo. É a parte mais alta do bioma Caatinga. Nos maciços e serras altas os solos são em geral rasos, pedregosos e pobres, predominando os solos litólicos e grandes afloramentos de rocha. É um mosaico que inclui caatinga com grande diversidade (abaixo de 1.000 m de altitude), cerrado, campos rupestres e diferentes tipos de mata (dá mais seca à mais úmida).

O Raso da Catarina está localizado em uma estreita e alongada faixa no sentido N-S, na parte centro-leste do bioma. A norte, oeste e leste limita com a Depressão Sertaneja Meridional. Uma ponta nordeste limita com o Planalto da Borborema e a parte sul com o Recôncavo Baiano, na Zona da Mata. O clima é semiárido, quente e seco, com precipitação média de 650 mm/ano na parte sul (BA) e período chuvoso de dezembro a julho. A Caatinga é predominantemente arbustiva, muito densa e menos espinhosa que a caatinga de solos cristalinos.

Ocorre uma expressiva na Caatinga pedodiversidade, variando de solos rasos, rochosos e relativamente férteis a profundos, arenosos e de baixa fertilidade natural (SAMPAIO, 2010). Registram-se neste bioma 12 das 13 ordens de solos do Sistema Brasileiro de Classificação de Solos (Figura 7), dominando Neossolos (28,51% da área total), seguidos por Latossolos (23,84%), Argissolos (16,51%), Luvisolos (10,91%), Planossolos (10,24%), Cambissolos (5,85%), Vertissolos, Plintossolos, Chernossolos, Nitossolo, Gleissolos e Organossolos (< 1% cada).

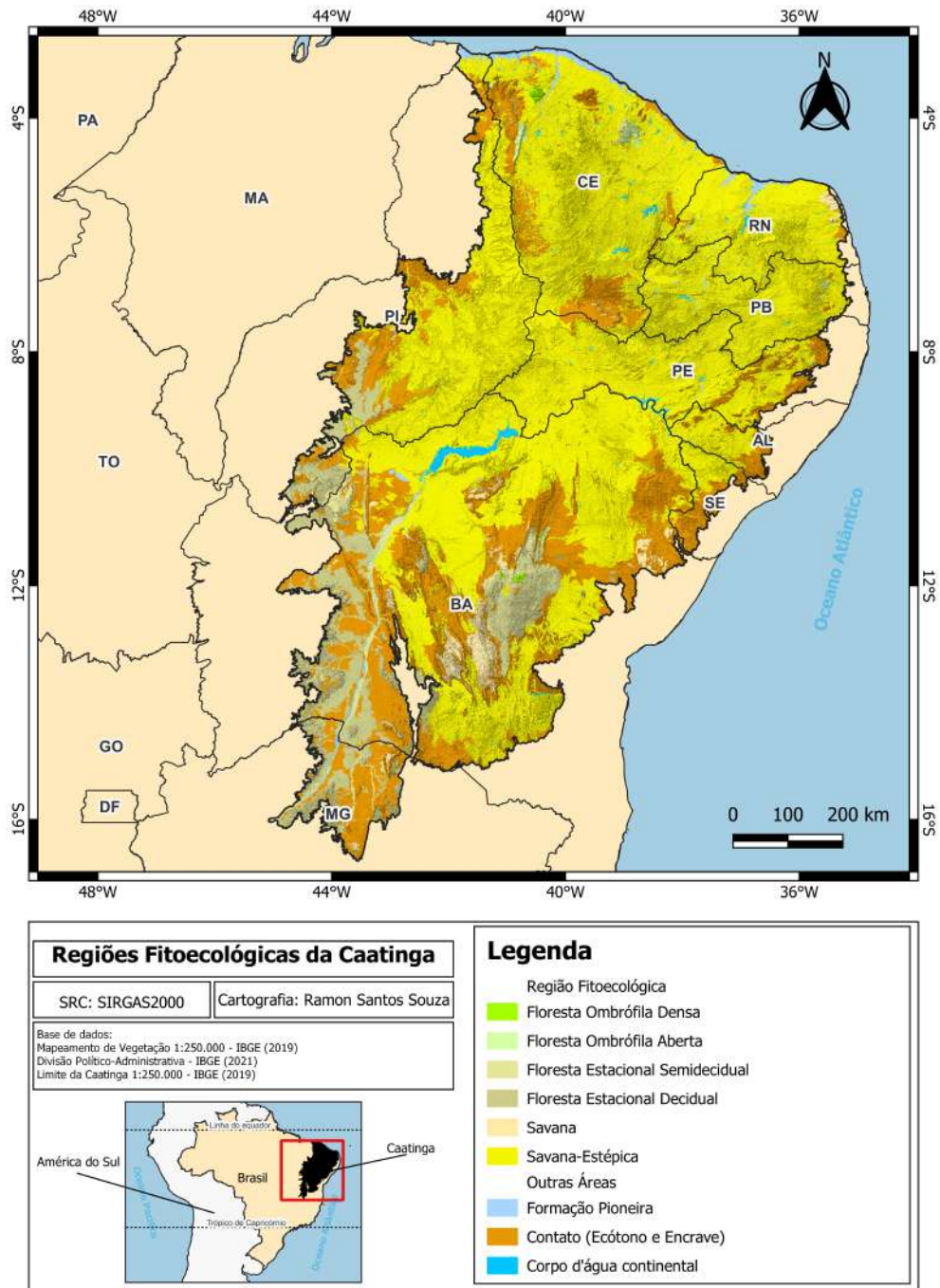
Figura 7 – Solos do bioma Caatinga.



Fonte: Elaborada pelo autor.

As regiões fitoecológicas dominantes são a Savana-Estépica (63,27%) e o Contato (Ecótono e Enclave) (22,56%) (IBGE, 2019). As espécies decíduas são predominantes nesse bioma, apresentando distribuição sobreposta à região de clima tropical semiárido e subúmido seco do Brasil (REDDV, 1983; SAMPAIO, 1995; VELLOSO; SAMPAIO; PAREYN, 2002) (Figura 8).

Figura 8 – Regiões fitoecológicas da Caatinga.



Fonte: Elaborada pelo autor.

RESULTADOS E DISCUSSÃO

3 INTEGRANDO IMAGENS GEDI E SENTINEL PARA ESTIMAR PARÂMETROS ESTRUTURAIS DE UMA FLORESTA SAZONALMENTE SECA DO BRASIL

3.1 INTRODUÇÃO

As florestas estão distribuídas em aproximadamente um terço da superfície terrestre (MORIN *et al.*, 2022). Os ambientes florestados suportam serviços ecossistêmicos essenciais como: estabilização do clima, ciclagens de nutrientes, conservação de nascentes, fornecimento de recursos para uma infinidade de formas de vida (MORIN *et al.*, 2022; MUSTHAFA *et al.*, 2023; REN *et al.*, 2023). Além disso, florestas tem grande valor econômico como matéria-prima para inúmeros produtos (MORIN *et al.*, 2022).

Linares-Palomino *et al.* (2011) afirmam que as Florestas Tropicais Sazonalmente Secas (FTSS) representam tipos de vegetação que já foram mais amplamente distribuídos nas Américas, estando atualmente encontrados em fragmentos distribuídos do noroeste do México ao norte da Argentina e sudoeste do Brasil em áreas separadas de tamanhos variados. A FTSS situada no domínio da caatinga (PENNINGTON *et al.*, 2018; PRADO, 2003; SARMIENTO; MONASTERIO, 1975) está localizada principalmente na região Nordeste do Brasil (NEB) e em uma pequena parte da região Sudeste (IBGE, 2019). É composto por um conjunto vegetativo adaptado a exposição de climas semiáridos e subúmidos secos, ocupando uma área de uma área de 862.818 km² (IBGE, 2019).

A estrutura da floresta é composta da disposição das partes da planta no espaço; distribuição vertical da folhagem ou estruturas ramificadas; padrão horizontal da altura das árvores ou distribuição das espécies (BELAND *et al.*, 2019). Nesta perspectiva, a complexidade estrutural captura a diversidade dos elementos da vegetação, desde a altura das árvores até a composição das espécies (CAMARRETTA *et al.*, 2020). Nesse contexto, estudos têm apontado a importância estrutural das florestas como indicador chave da biodiversidade dos ecossistemas (CAMARRETTA *et al.*, 2020) e primordial para estudos de mudanças globais (WANG, 2022).

A altura e cobertura do dossel é um parâmetro estrutural usado para representação visual e estimativa da biomassa, ajuda entender a circulação de carbono terrestre e é uma informação auxiliar para avaliar os recursos florestais e estabelecer modelos do sistema terrestre (ZHANG *et al.*, 2022; ZHOU *et al.*, 2023). Assim, ter informações da variabilidade desses parâmetros facilita o manejo florestal e a formulação de políticas (XI *et al.*, 2022).

O monitoramento em larga escala dos parâmetros estruturais das florestas é crucial

no esforço mundial para combater as mudanças climáticas, o declínio contínuo da biodiversidade e redução da pobreza de populações vulneráveis (NEUVILLE *et al.*, 2021; WALDELAND *et al.*, 2022). Embora os inventários florestais tradicionais, como levantamentos de campo, possam fornecer informações detalhadas da altura do dossel florestal com alta precisão, eles são demorados, trabalhosos e dificultam tanto a investigação rápida quanto os levantamentos em larga escala (XI *et al.*, 2022).

Assim, muitas variáveis estruturais são difíceis e demoradas para medir no campo e especialmente em ecossistemas remotos, altos, complexos, espacialmente variáveis ou altamente sensíveis. O LiDAR (do acrônimo inglês “*light detection and ranging*”) provou ser útil na obtenção de informações sobre a estrutura da floresta devido à sua velocidade, cobertura e capacidade de descrever atributos 3D em comparação com os métodos manuais existentes. Os dados posicionais 3D altamente detalhados fornecidos pelos sistemas LiDAR revolucionaram - e podem expandir ainda mais - a maneira como consideramos a estrutura do dossel na ciência do ecossistema florestal (BELAND *et al.*, 2019).

A tecnologia LiDAR tem contribuído significativamente para a estimativa de parâmetros estruturais da floresta (XI *et al.*, 2022). Esta tecnologia pode observar diretamente o dossel da floresta na direção vertical e pode fornecer informações tridimensionais sobre a estrutura da floresta (XI *et al.*, 2022). Neste contexto, o propósito desta investigação foi modelar as propriedades estruturais da vegetação em um modelo de regressão florestal aleatória. Os dados são processados no ambiente de análise geoespacial de computação em nuvem do GEE. Os produtos finais são os primeiros mapas de alta resolução (10 m) de altura do dossel e densidade de cobertura do dossel para o bioma Caatinga.

3.2 MATERIAL E MÉTODOS

3.2.1 Área de Estudo

A Caatinga está inserida na região Nordeste do Brasil (NEB) e em uma pequena parte da região Sudeste (IBGE, 2019; MORO *et al.*, 2014; AB’SABER, 1977; PRADO, 2003). Constitui-se por uma área de 862.818 km² (IBGE, 2019), correspondendo 10% do território brasileiro. A Caatinga é circundada pela Mata Atlântica e Cerrado, representando 54% do NEB, abarcando grandes áreas dos estados do Ceará (CE), Alagoas (AL), Bahia (BA), Maranhão (MA), Paraíba (PB), Pernambuco (PE), Piauí (PI), Rio Grande do Norte (RN) e Sergipe (SE).

Além disso, o bioma também está presente em parte do norte do estado de Minas Gerais (MG), na região Sudeste (GANEM *et al.*, 2020).

Inserida entre a linha do Equador e o Trópico de Capricórnio, isso proporciona a Caatinga uma alta radiação solar anual (MORO *et al.*, 2016), o que influencia na temperatura média que se concentra entre 25° e 30°, não existindo diferença entre os meses de inverno e verão. Quanto ao período chuvoso, grande parte da precipitação não é armazenada, devido a condição do solo pouco profundo, ocasionando altos índices de evaporação. Quanto a umidade relativa, é baixa, com percentuais entorno de 50% (GANEM, 2017).

A Caatinga é um dos biomas mais degradados do país, principalmente pelo uso inadequado da vegetação e do solo. Acredita-se que a Caatinga perdeu metade de sua cobertura original, ocasionado pela ação antrópica para uso agropecuários, uso madeireira e consumo de lenha (SILVA *et al.*, 2017; ANTONGIOVANNI *et al.*, 2020).

3.2.2 Levantamento de Dados

Os atributos estruturais da vegetação foram obtidos pelos *lasers GEDI* (*Global Ecosystem Dynamics Investigation*), utilizou-se a altura do dossel (98° percentil das métricas de altura relativa do GEDI L2A), cobertura total do dossel (GEDI L2B) e índice de área total (GEDI L2B). Primeiramente, os dados GEDI disponíveis para a Caatinga foram filtrados rejeitando todas as medições inválidas que atendessem aos requisitos de qualidade: “*quality flag*” = 1 e “*degrade flag*” = 0. Além disso, excluímos valores espúrios restantes que estivessem abaixo da percentil 1 e acima de 99. Após essa etapa foi gerada a mediana para todas as imagens disponíveis no ano de 2021 e finalmente foram transformados em pontos vetoriais para servir de entrada no modelo.

3.2.3 Covariáveis Ambientais

Neste estudo, usamos principalmente a plataforma de computação em nuvem *Google Earth Engine* (GEE): <<https://earthengine.google.com>>, que possibilita utilizar um grande banco de dados geoespacial em escala de petabytes disponíveis gratuitamente, aplicando-se uma variedade de algoritmos em alta performance (GORELICK *et al.*, 2017).

Dados ópticos do Sentinel-2 (S2), a banda C *Synthetic-Aperture-Radar* (SAR) do Sentinel-1 (S1) (10m), o Modelo Digital de Elevação (MDE) FABDEM (*Forest and Buildings removed Copernicus DEM*) (HAWKER *et al.*, 2022) e seus atributos derivados foram conside-

rados como variáveis independentes, enquanto os dados *LiDAR* do *Global Ecosystem Dynamic Investigation* (GEDI) foram considerados uma como variável dependente. Todos acessíveis via *Google Earth Engine* (GEE) usando o pacote *rgee* 1.1.6 na linguagem de programação R 4.2.2.

Para S2 foi considerada uma porcentagem máxima de nuvens definida em 60% entre janeiro e julho de 2021 para gerar imagens com medianas geométricas para cinco bandas (azul, verde, vermelho, borda vermelha, infravermelho próximo) na estação chuvosa. As imagens compostas geradas foram então usadas para produzir sete índices de vegetação na resolução espacial de 10 m:

O Índice de Vegetação por Diferença Normalizada (NDVI) (TUCKER, 1979), é calculado pela equação 4.1:

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad (3.1)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{RED} = banda do vermelho.

—

O Índice de Verdura de Diferença Normalizada (NDGI) (BANNARI *et al.*, 1995; YANG *et al.*, 2019), é calculado pela equação 3.2:

$$NDGI = \frac{\rho_{NIR} - \rho_{GREEN}}{\rho_{NIR} + \rho_{GREEN}} \quad (3.2)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{GREEN} = banda do verde.

O Índice de Taxa de queima normalizada (NBR) (JI *et al.*, 2011), é calculado pela equação 3.3:

$$NBR = \frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}} \quad (3.3)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{SWIR1} = banda do infravermelho médio.

Em seguida, foi gerado o índice do potencial de fluxo de dióxido de carbono (CFLUX), a partir da metodologia descrita por Rahman *et al.* (2000) a partir da seguinte equação 3.6:

$$PRI = \frac{\rho_{BLUE} - \rho_{GREEN}}{\rho_{BLUE} + \rho_{GREEN}} \quad (3.4)$$

$$sPRI = \frac{(PRI + 1)}{2} \quad (3.5)$$

$$CFLUX = NDVI * sPRI \quad (3.6)$$

Foi obtido a Fração de Cobertura de Vegetação (FVC), a partir da metodologia descrita por Carlson e Ripley (1997) a partir da seguinte equação 3.7

$$FCV = \left(\frac{NDVI - NDVI_S}{NDVI_{VEG} - NDVI_S} \right)^2 \quad (3.7)$$

- $NDVI_S$ índice da diferença normalizada do solo exposto;
- $NDVI_{VEG}$ índice da diferença normalizada de solo com cobertura vegetal;
- O valor de $NDVI_S$ adotado foi de 0,2 e o valor de $NDVI_{VEG}$ foi 0,6.

O Índice de vegetação ajustado ao solo (*Soil-adjusted vegetation index* (SAVI)) leva conta a relação entre as regiões espectrais vermelha e infravermelha, utilizando um fator de ajuste (L) com valores de 1, 0,5 e 0,3 para cobertura vegetal baixa, média e alta, respectivamente (HUETE, 1988), optou-se neste estudo por 0,5.

É calculado pela equação 3.8:

$$\frac{(1 + L)(\rho_{NIR} - \rho_{RED})}{L + \rho_{NIR} + \rho_{RED}} \quad (3.8)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{RED} = banda do vermelho;
- L = 0,5 é o fator de ajuste do solo utilizado.

O Índice de vegetação de clorofila verde (GCVI) (GITELSON, 2005), é computado pela equação 3.9:

$$GCVI = \frac{\rho_{NIR}}{\rho_{GREEN}} - 1 \quad (3.9)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{GREEN} = banda do verde.

O Índice de taxa de vegetação (RVI) (JORDAN, 1969), é determinado pela equação 3.10:

$$\rho_{NIR} \cdot \rho_{RED} \quad (3.10)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{RED} = banda do vermelho.

O Índice de vegetação de diferença (DVI) (TUCKER *et al.*, 1979), é obtido pela equação ??:

$$DVI = \rho_{NIR} - \rho_{RED} \quad (3.11)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{RED} = banda do vermelho.

O Índice de vegetação de clorofila (CVI) (VINCINI *et al.*, 2008), é dada pela equação 3.13:

$$CVI = \rho_{NIR} \times \left(\frac{\rho_{RED}}{\rho_{GREEN}^2} \right) \quad (3.12)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{GREEN} = banda do verde.

O Índice de vegetação verde atmosféricamente resistente (GARI) (GITELSON *et al.*, 1996), é dada pela equação 3.13:

$$GARI = \frac{\rho_{NIR} - (\rho_{GREEN} - (\rho_{BLUE} - \rho_{RED}))}{\rho_{NIR} - (\rho_{GREEN} + (\rho_{BLUE} - \rho_{RED}))} \quad (3.13)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{RED} = banda do vermelho.

O Índices de vegetação resistente à atmosfera (ARVI) (KAUFMAN; TANRÉ, 1996), é calculado pela equação 3.14:

$$ARVI = \frac{\rho_{NIR} - (2 \times (\rho_{RED} - \rho_{BLUE}))}{\rho_{NIR} + (2 \times (\rho_{RED} - \rho_{BLUE}))} \quad (3.14)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{RED} = banda do vermelho.

O Índice de Vegetação Aprimorada (EVI) (HUETE *et al.*, 2002), é dada pela equação 4.2:

$$EVI = 2.5 \times \left(\frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + 6 \times \rho_{RED} - 7.5 \times \rho_{BLUE} + 1} \right) \quad (3.15)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{RED} = banda do vermelho;
- ρ_{BLUE} = banda do azul.

O Índice de Água por Diferença Normalizada (NDWI) (GAO, 1996), é obtida pela equação 4.3:

$$NDWI = \frac{\rho_{NIR} - \rho_{SWIR1}}{\rho_{NIR} + \rho_{SWIR1}} \quad (3.16)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;

- ρ_{SWIR1} = banda do infravermelho médio.

O Índice de Vegetação Ajustada ao Solo Modificado (MSAVI2) (QI *et al.*, 1994), é gerado com base na equação 4.4:

$$\text{MSAVI2} = \frac{2 \times \rho_{\text{NIR}} + 1 - \sqrt{(2 \times \rho_{\text{NIR}} + 1)^2 - 8 \times (\rho_{\text{NIR}} - \rho_{\text{R}})}}{2} \quad (3.17)$$

A equação 4.5 gera o Índice de Vegetação de Diferença Verde (GDVI) (WU, 2014):

$$\text{GDVI} = \text{GREEN} - \text{RED} \quad (3.18)$$

O Índice de Vegetação por Diferença Normalizada Verde (GNDVI) (GITELSON *et al.*, 1996), é obtido conforme a equação 3.19:

$$\text{GNDVI} = \frac{\rho_{\text{NIR}} - \rho_{\text{GREEN}}}{\rho_{\text{NIR}} + \rho_{\text{GREEN}}} \quad (3.19)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{GREEN} = banda do verde.

O *Brightness* da transformação *Tasseled Cap* (TCB) (CRIST, 1985) (Figura 18h), é dada pela equação 3.20:

$$\text{TCB} = 0.3037 \times \rho_{\text{BLUE}} + 0.2793 \times \rho_{\text{GREEN}} + 0.4743 \times \rho_{\text{RED}} + 0.5585 \times \rho_{\text{NIR}} + 0.5082 \times \rho_{\text{SWIR1}} + 0.1863 \times \rho_{\text{SWIR2}} \quad (3.20)$$

Em que:

- ρ_{BLUE} = banda do azul;
- ρ_{GREEN} = banda do verde;
- ρ_{RED} = banda do vermelho;
- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{SWIR1} = banda do infravermelho médio;
- ρ_{SWIR2} = banda do infravermelho médio.

O *Greenness* da transformação *Tasseled Cap* (TCG) (CRIST, 1985), é dada pela equação 3.21:

$$\begin{aligned} \text{TCG} = & -0.2848 \times \rho_{\text{BLUE}} - 0.2435 \times \rho_{\text{GREEN}} - 0.5436 \times \rho_{\text{RED}} + 0.7243 \\ & \times \rho_{\text{NIR}} + 0.084 \times \rho_{\text{SWIR1}} - 0.18 \times \rho_{\text{SWIR2}} \end{aligned} \quad (3.21)$$

Em que:

- ρ_{BLUE} = banda do azul;
- ρ_{GREEN} = banda do verde;
- ρ_{RED} = banda do vermelho;
- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{SWIR1} = banda do infravermelho médio;
- ρ_{SWIR2} = banda do infravermelho médio.

O *Wettness* da transformação *Tasseled Cap* (TCW) (CRIST, 1985), é dada pela equação 3.22:

$$\begin{aligned} \text{TCW} = & 0.1509 \times \text{BLUE} + 0.1973 \times \text{GREEN} + 0.3279 \times \text{RED} + 0.3406 \\ & \times \text{NIR} + 0.7112 \times \text{SWIR1} + 0.4572 \times \text{SWIR2} \end{aligned} \quad (3.22)$$

Em que:

- ρ_{BLUE} = banda do azul;
- ρ_{GREEN} = banda do verde;
- ρ_{RED} = banda do vermelho;
- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{SWIR1} = banda do infravermelho médio;
- ρ_{SWIR2} = banda do infravermelho médio.

A partir do S1 foi obtida imagens de retroespalhamento nas polarizações vertical-vertical (VV) e vertical-horizontal (VH) no modo de faixa interferométrica larga (IW) foram usadas para gerar compósitos sazonais para janeiro e julho de 2021: média, mediana, mínima, máximo e os percentis 10, 25, 75, 90 para as polarizações VV, VH e relação VV/VH.

Os parâmetros de textura da matriz de co-ocorrência de nível de cinza (GLCM) foram gerados para a relação VV/VH (CONNERS *et al.*, 1984; HARALICK *et al.*, 1973). Os parâmetros foram calculados usando a função `glcmTexture` no GEE em uma janela móvel de 3 por 3 pixels. Selecionamos o:

- ASM: Segundo Momento Angular; mede o número de pares repetidos;

- CONTRAST: Contraste; mede o contraste local de uma imagem;
- CORR: Correlação; mede a correlação entre pares de pixels;
- VAR: Variância; mede a dispersão da distribuição dos níveis de cinza;
- IDM: Momento Inverso da Diferença; mede a homogeneidade;
- SVAR: Variância da soma;
- ENT: Entropia. Mede a aleatoriedade de uma distribuição de nível de cinza;
- DISS: Dissimilaridade;
- SHADE: Sombra de cluster.

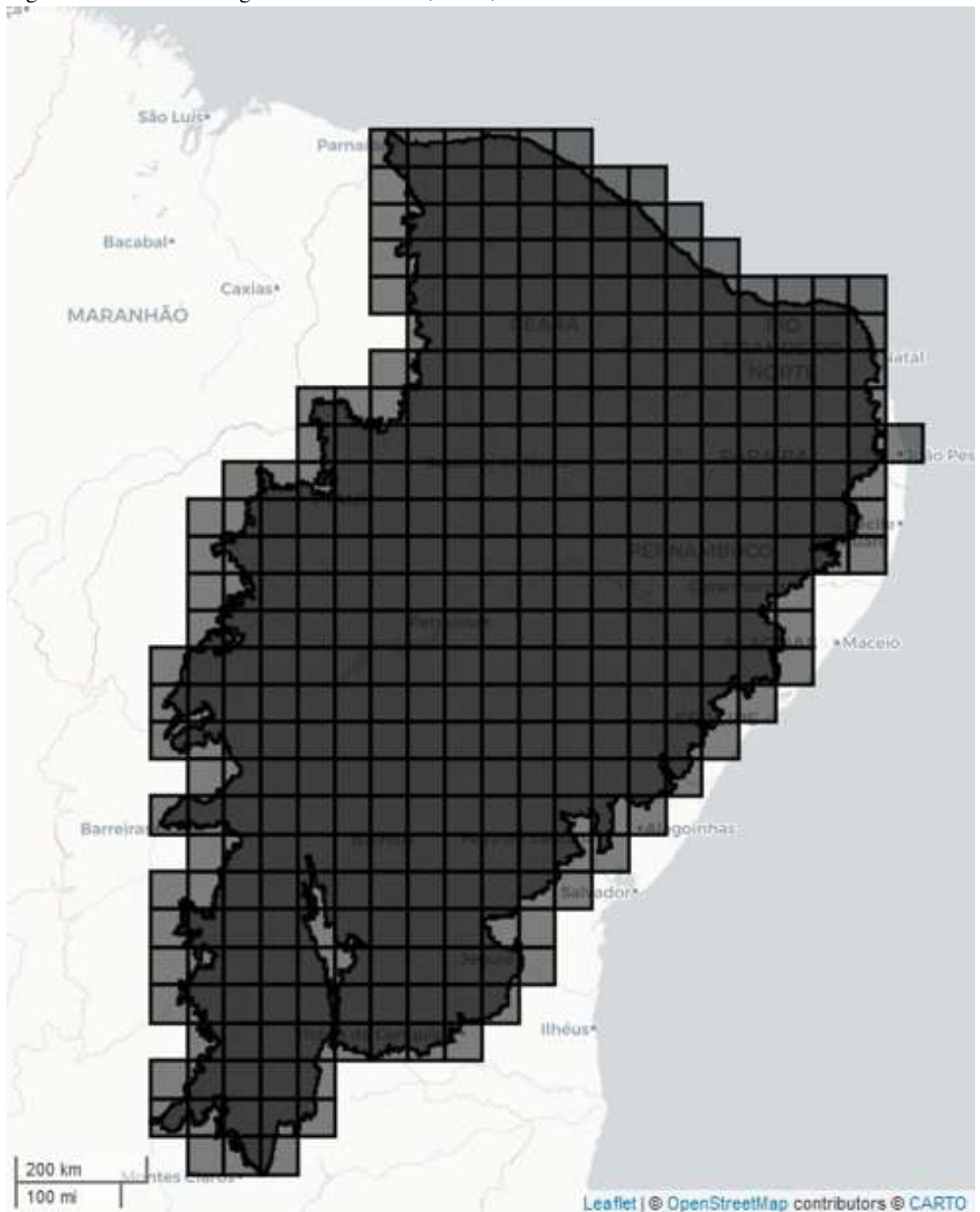
O MDE FABDEM possui a resolução espacial de 30 m, é um resultado da correção do novo *Copernicus GLO-30* DEM a partir de aprendizado de máquina removendo artefatos de florestas e edificações (HAWKER *et al.*, 2022). A partir software Saga 7.8.2 foi gerado o área de drenagem acumulada (SCA) usando o módulo *Flow Width and Specific Catchment Area* e o método de saída foi o *Multiple Flow Direction* (MFD) (QUINN *et al.*, 1991). Após isso foi importado o SCA no GEE para elaborar: i-As (Aspecto); ii-SI (Declividade); iii-El (Elevação); iv-SPI (Índice de Potência de Escoamento); v-TWI (Índice de Umidade Topográfica).

3.2.4 Treinamento, Avaliação e Predição Espacial

O algoritmo de regressão *Random Forest* (RF) (BREIMAN, 2001) foi usado para estimar as propriedades do solo e a estrutura florestal na Caatinga. No GEE, a saída do modelo de RF pode ser definida como regressão, usada para modelar variáveis contínuas, com base na média das previsões de todas as árvores de decisão para as versões *bootstrap* dos dados de treinamento (HASTIE *et al.*, 2013).

Para estimar as estruturas da vegetação primeiramente a área de estudo foi subdividida em 353 grids ($0,5^\circ \times 0,5^\circ$) para melhorar o desempenho do uso do recurso de processamento paralelo e evitar limites computacionais no GEE (??). Após isso, usamos uma janela móvel baseada em um grid e seus primeiros vizinhos (3x3 grids) para rodar a regressão RF. A modelagem de RF foi realizada usando a plataforma de computação em nuvem GEE aplicando a seguinte linha de código: `ee.Classifier.smileRandomForest.setOutputMode("REGRESSION")`. Os parâmetros do modelo foram as seguintes: “*decision trees*”= 100 (default), “*min leaf population*”= 1, “*variables per split*”= raiz quadrada do número de variáveis, “*bag fraction*”= 0,5, *max nodes* = (default) ilimitado e *seed*= 0.

Figura 9 – Estrutura em grids em folhas de $0,5^\circ \times 0,5^\circ$.



Fonte: Elaborada pelo autor.

Para avaliar a precisão dos dados previstos, foram utilizadas duas métricas de medição de acurácia: coeficiente de determinação (R^2) (Equação 3.23) e raiz quadrada do erro-médio (RMSE: *Root-Mean-Square Error*) (Equação 3.24):

$$R^2 = \frac{(\sum_{i=1}^n (Y_i - \bar{X})(X_i - \bar{X}))^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2 \sum_{i=1}^n (X_i - \bar{X})^2} \quad (3.23)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - X_i)^2}{n}} \quad (3.24)$$

Em que:

- X e Y representam valores medidos e previstos, respectivamente;
- n é o número de amostras;
- X_i e Y_i são os valores i emparelhados dos dados medidos e previstos, respectivamente;
- $\bar{X}$ e $\bar{Y}$ são a média dos dados previstos e observados, respectivamente.

3.3 RESULTADOS E DISCUSSÃO

3.3.1 Avaliação de Erro

A incerteza estatística da modelagem utilizando algoritmo *Random Forest* da altura do dossel (m), cobertura total do dossel (%) e índice de área total da planta (m^2/m^2) foi calculada com base na comparação de um conjunto teste contendo 20% das amostras GEDI (observações) com as previsões do modelo, calculando o coeficiente de determinação (R^2) e erro quadrático médio (RMSE). Como a modelagem foi implementada numa abordagem baseada em *grids*, para cada grid modelado e atributo GEDI, os critérios de erro foram calculados e agregados como valores médios e medianos apresentados na Tabela 1. Neste contexto, a maior precisão do modelo foi alcançada, de acordo com R^2 , para a altura do dossel (média: 37%, mediana: 38%) e RMSE equivalente a 2 m. Isso sugere que o mapa de altura do dossel tende a superestimar espécies de porte herbáceo á arbustivo (<2 m) (Figura 2). Em contraste, Sothe *et al.* (2022) obteve resultados parecidos no Canadá, com espécies < 5 m. A incerteza estatística da cobertura total do dossel modelado tem um valor médio e mediano para R^2 igual a 28% e RMSE de 12%. O índice de área total da planta teve um RMSE médio e mediando respectivamente de 0,46 e 0,43 m^2/m^2 . A cobertura total do dossel modelado é ligeiramente maior (média: 28%,

mediana: 28%) do que o índice de área total da planta (média: 27%, mediana: 26%) (Tabela 1). Resultados semelhantes foram encontrados em outros estudos que também usaram o GEDI (GUPTA; SHARMA, 2022; MA *et al.*, 2023; ZHOU *et al.*, 2023).

Tabela 1 – Erros do modelo do modelo *Random Forest*-RF para cada propriedade estrutural da vegetação Caatinga. RMSE (erro quadrático médio) e R^2 (coeficiente de determinação). Valores de média e mediana por critério de erro calculados para todos os *grids* modelados.

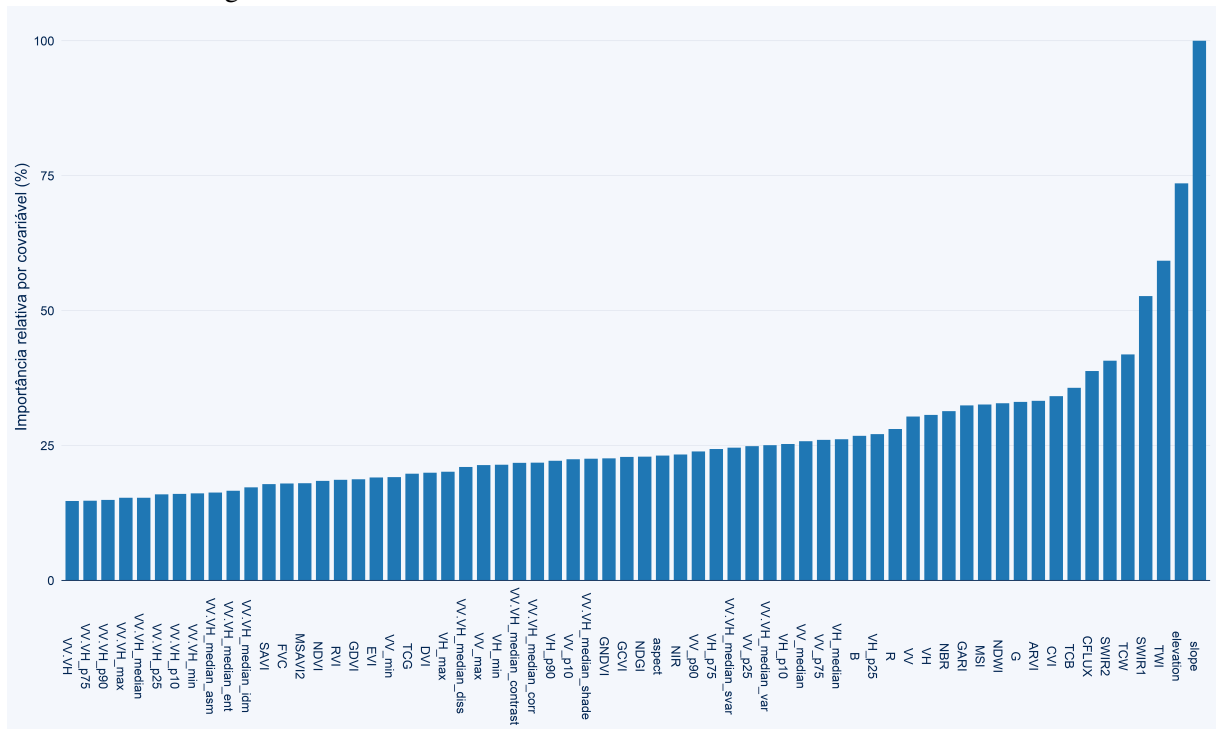
Propriedade estrutural	Média do R^2 (%)	Média do RMSE	Mediana do R^2 (%)	Mediana do RMSE
Cobertura total do dossel (cover) em %	28	12	28	12
Índice de área total da planta (pai) em m^2/m^2	27	0,46	26	0,43
Altura do dossel (rh98) em m	37	2,07	38	2,08

Fonte: Elaborada pelo autor.

3.3.2 Variáveis Importantes

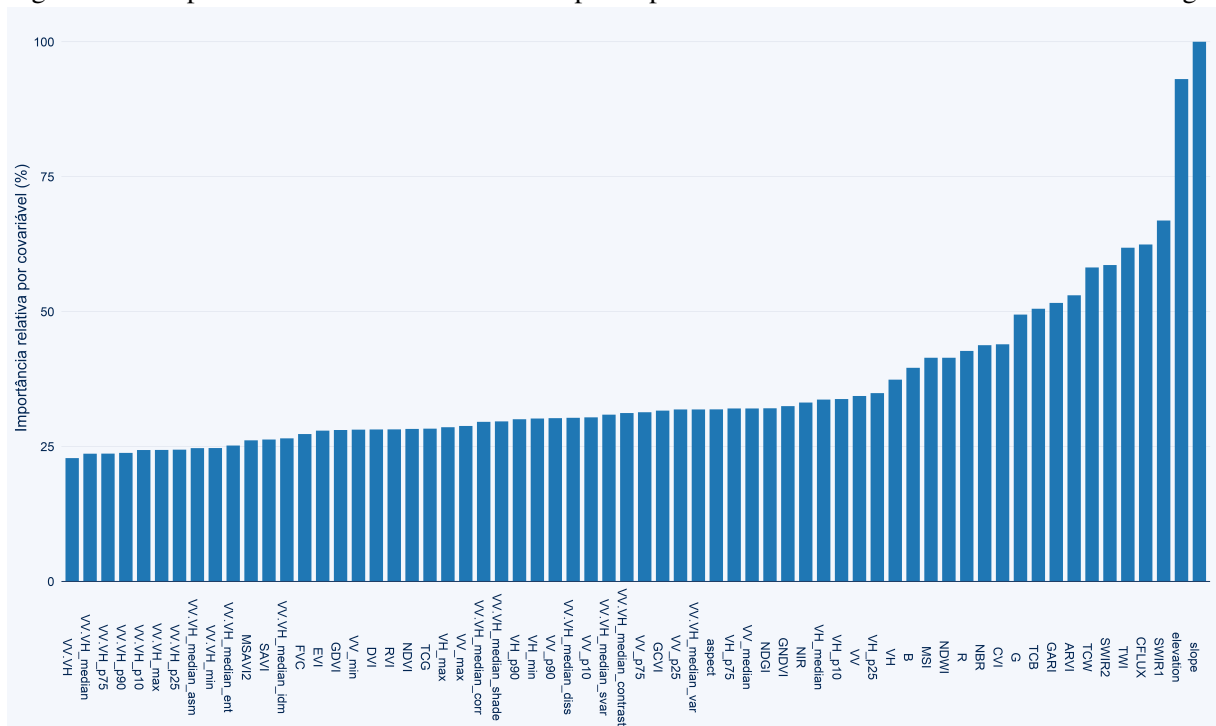
A Figura 10 apresenta o resultado médio da importância relativa para as regressões RF usadas para prever a altura superior do dossel (rh98), destaca-se a declividade (*slope*), elevação (*elevation*) e TWI como os preditores mais relevantes para estimar os parâmetros estruturais da vegetação. O SWIR1, TCW, SWIR2, CFLUX, TCB, CVI e ARVI que foram gerados a partir do satélite S2 também contribuíram significativamente na modelagem, esses índices espectrais refletem as condições e estados da vegetação no bioma Caatinga (Figura 10). Na Figura 12 observa-se que a declividade, elevação, SWIR1, CFLUX, TWI, SWIR2, TCW, ARVI e GARI são as variáveis mais importantes para estimar a cobertura total da vegetação, exibindo contribuições acima de 50%. Já na Figura 12, nota-se que respectivamente a declividade, elevação, TWI, CFLUX, SWIR1, ARVI, SWIR2 e GARI são as variáveis mais relevantes na previsão do índice de área total da planta.

Figura 10 – Importância relativa das covariáveis para a previsão da altura superior do dossel (rh98) na Caatinga.



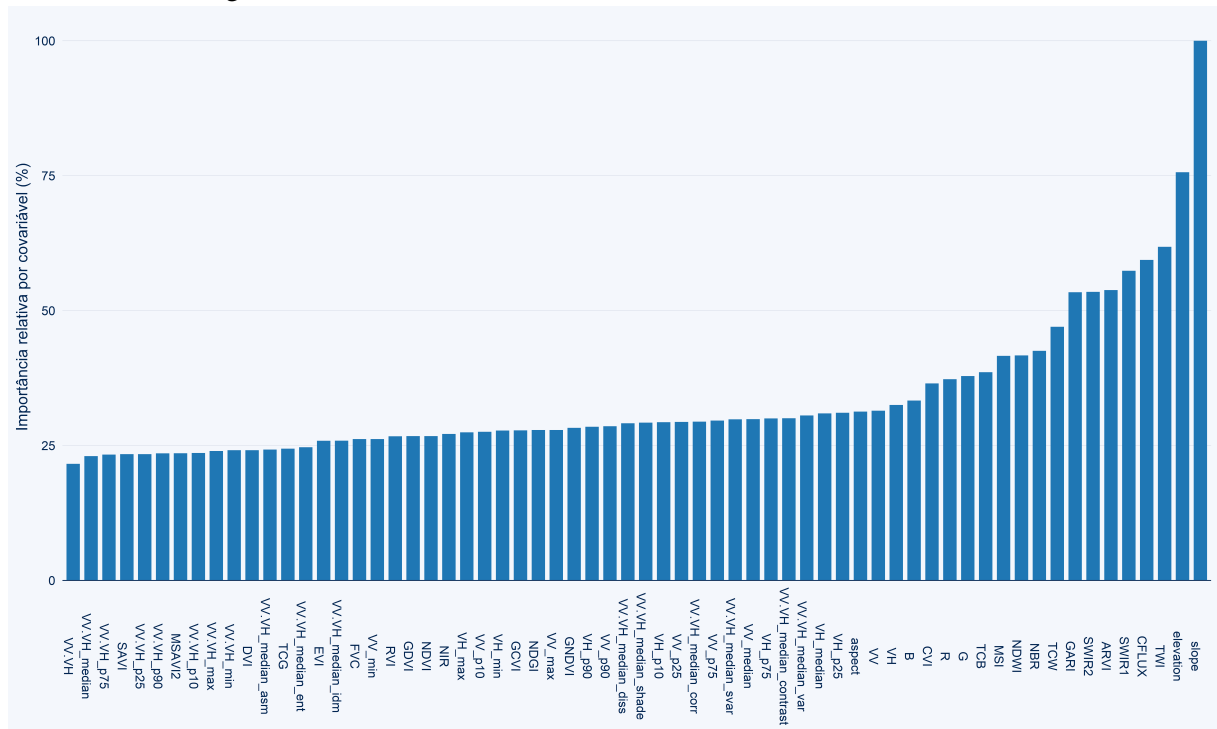
Fonte: Elaborada pelo autor.

Figura 11 – Importância relativa das covariáveis para a previsão da cobertura total do dossel na Caatinga.



Fonte: Elaborada pelo autor.

Figura 12 – Importância relativa das covariáveis para a previsão da índice de área total da planta na Caatinga.



Fonte: Elaborada pelo autor.

3.3.3 Atributos da Estrutura da Vegetação Modelada

Este estudo apresenta a primeiro mapeamento digital da estrutura da vegetação para o bioma Caatinga, obtida a partir de imagens de sensoriamento remoto de alta resolução utilizando o *Google Earth Engine*. As Figuras de 14-16 apresentam a altura do dossel, cobertura vegetal e o índice de área total da planta.

As alturas do dossel modeladas variaram de 2,8 a 17,7 m e um valor médio de $5,6 \pm 1,7$ m para a Caatinga, em conformidade com Andrade-Lima (1981) e Prado (2003). Observa-se que nos ambientes degradados do bioma, as alturas do dossel apresentaram valores inferiores a 4 m (Figura 14), estando localizados principalmente na parte central, sudeste e nordeste da área de estudo. Souza *et al.* (2015) identificaram a ocorrência de espécies vegetais com porte até ≈ 3 m (arbustivo alto) em ambientes degradados na bacia do Alto Curso na Paraíba, apresentou ainda, baixa densidade das espécies e longos trechos de solos expostos (Figura 13b). As alturas das copas mais elevadas (maiores de 7 m) estão dispersos em trechos de Caatinga preservada e úmida (Figura 14). Configura-se geralmente por um porte arbóreo de $\approx 5 - 20$ m, caracterizando assim uma fitofisionomia florestal (SOUZA *et al.*, 2015) (Figura 13a).

Figura 13 – Áreas de caatinga (a) preservada [8°7'23,95"S, 36°27'1,18"O] e (b) degradada [7°39'8,60"S, 36°41'26,33"O] na bacia do Alto Curso rio Paraíba no estado da Paraíba.

(a)



(b)

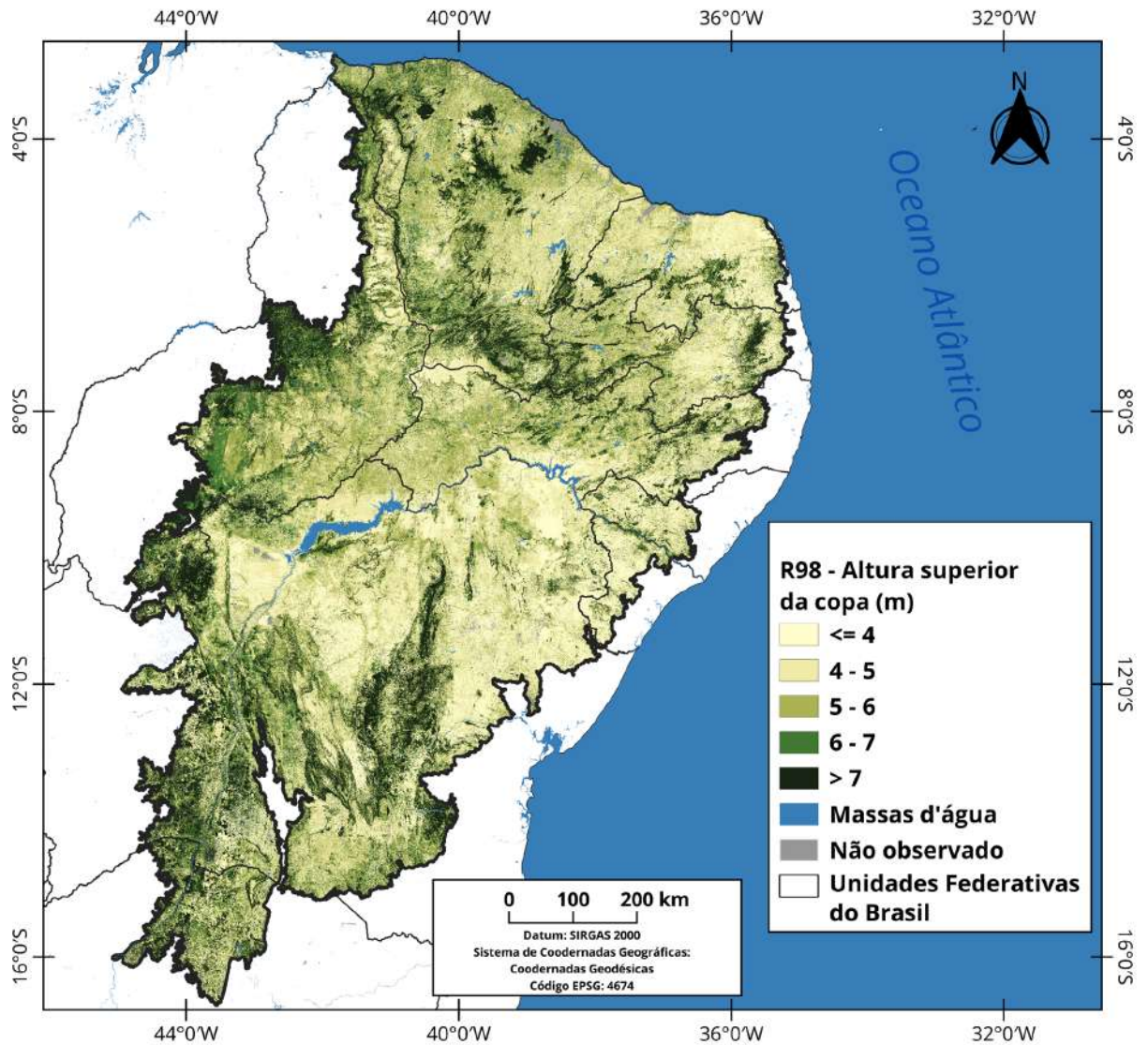


Fonte: Elaborada pelo autor.

A cobertura de dossel total modelada é apresentada na Figura 15 com densidades de cobertura de dossel variando de cerca de 0% (sem cobertura de dossel presente) a 87% (cobertura de dossel densa a fechada). A densidade média da cobertura do dossel é de 13% ($\pm 0,08$), destacando os níveis bastante esparsos das condições de cobertura do dossel na Caatinga.

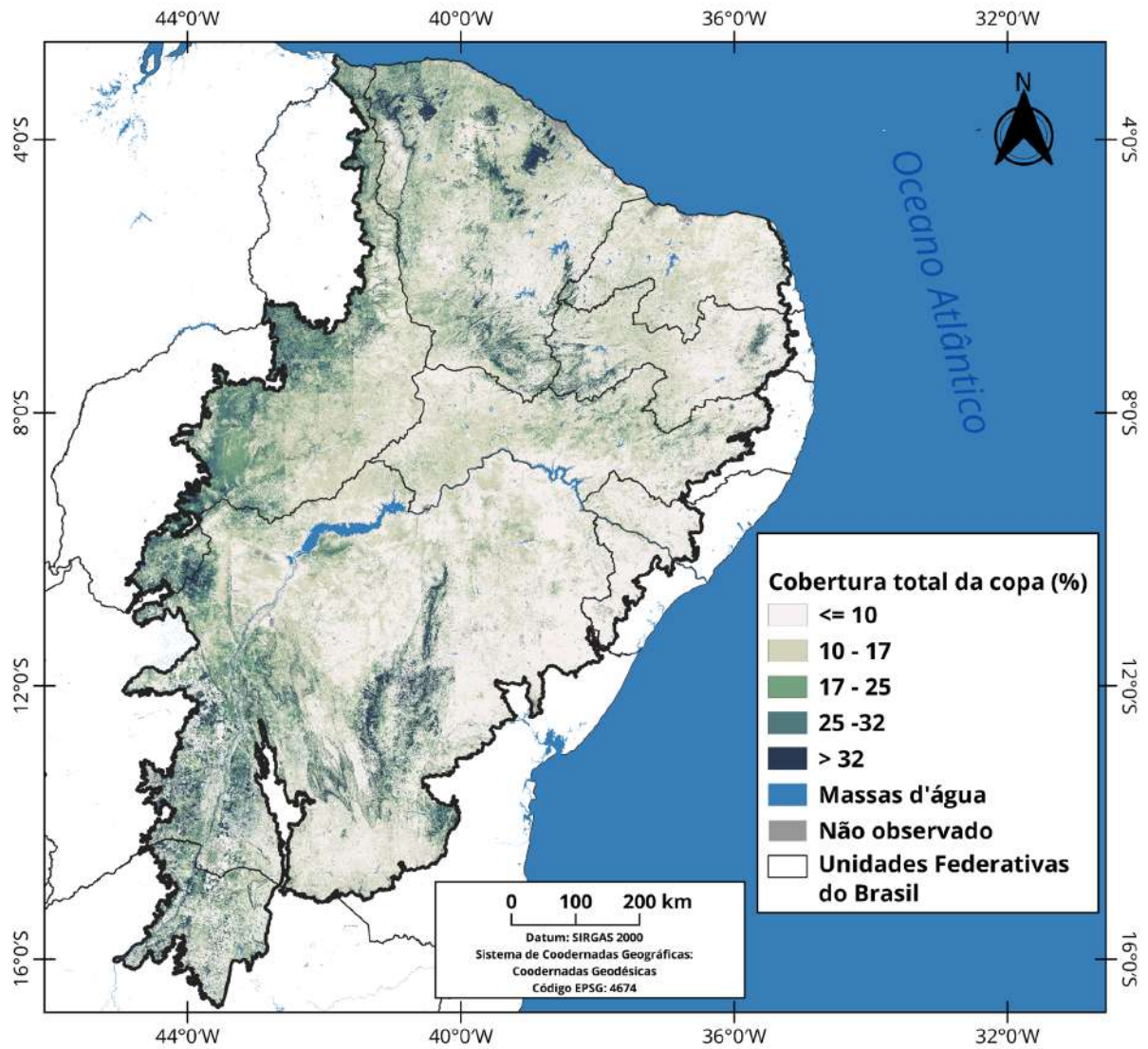
O mapa de índice de área total da planta modelado (Figura 16). Os valores variam de 0 m²/m² (nenhuma planta por área) a 4,4 m²/m² (alta densidade de área de planta), apresentam valor médio de 0,36 m²/m² ($\pm 0,33$) e intervalo de 4,38 m²/m². Contém variações espaciais semelhantes as demais propriedades estruturais analisadas, com predominâncias na área de valores inferiores a 0,33 m²/m². Novamente as áreas mais altas se destacam valores acima de 1,1 m²/m² (Figura 16).

Figura 14 – Altura superior do dossel (rh98) em m modelada para o bioma Caatinga.



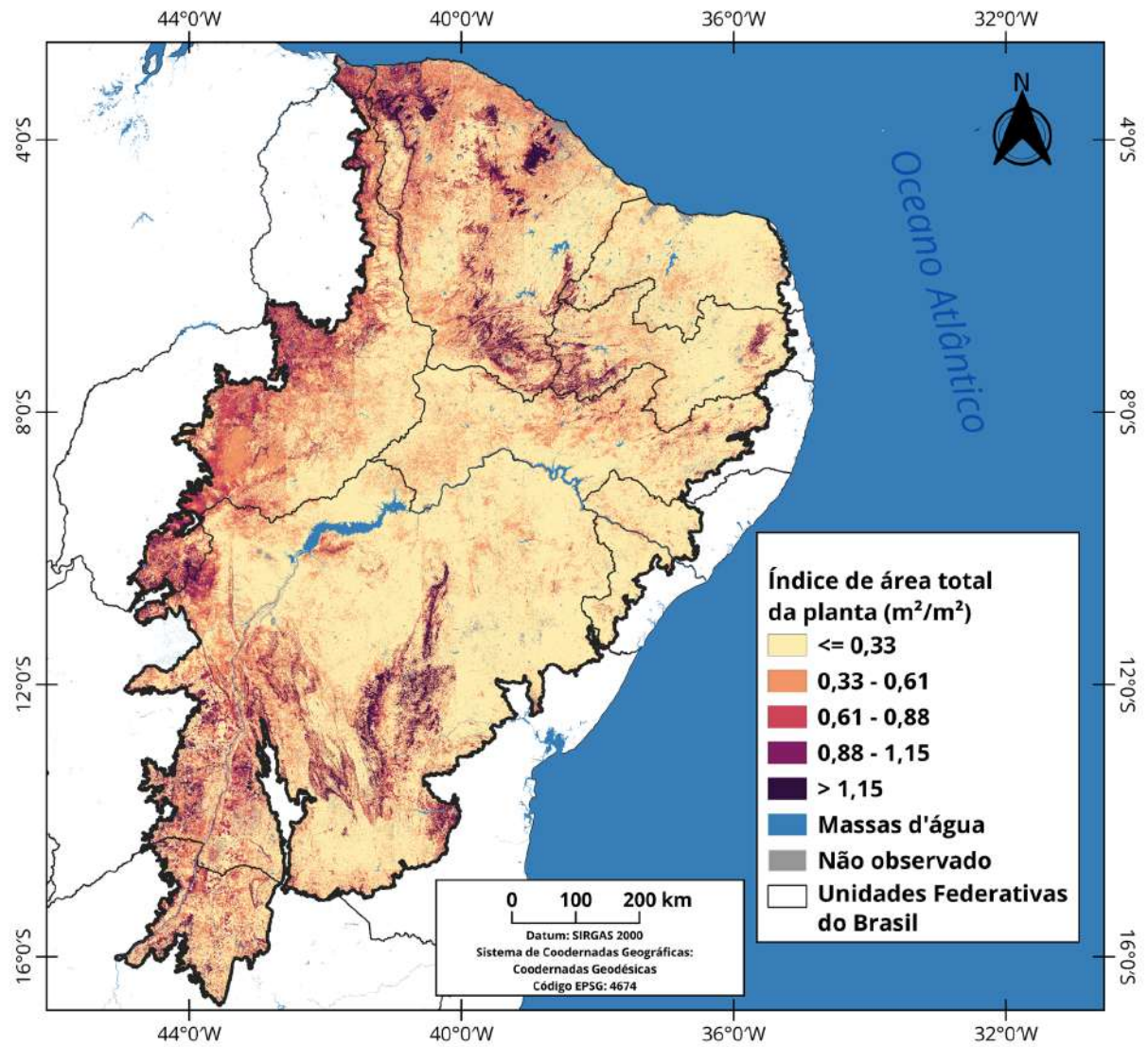
Fonte: Elaborada pelo autor.

Figura 15 – Cobertura total do dossel em % modelada para o bioma Caatinga.



Fonte: Elaborada pelo autor.

Figura 16 – Índice de área total da planta em m^2/m^2 modelada para o bioma Caatinga.



Fonte: Elaborada pelo autor.

4 MAPEAMENTO DIGITAL DAS PROPRIEDADES NUTRICIONAIS DO SOLO NO BIOMA CAATINGA

4.1 INTRODUÇÃO

Os nutrientes do solo desempenham papel fundamental no controle da produtividade biológica, na riqueza de espécies, na qualidade ambiental e na promoção da saúde de pessoas, animais e plantas nos diversos ecossistemas da Terra (ABEGAZ *et al.*, 2016; WANG *et al.*, 2018). A principal fonte desses nutrientes no solo é orgânica, ou seja, derivada da atividade biológica (FUJII *et al.*, 2018).

Dentre estes, o carbono (C), nitrogênio (N) e fósforo (P) são atributos nutricionais essenciais para a qualidade e fertilidade do solo (LAL, 2008; LAUSCH *et al.*, 2016; WAN *et al.*, 2018; WANG *et al.*, 2018; WANG *et al.*, 2019; ZHOU *et al.*, 2020a). Além disso, esses atributos do solo retroalimentar às mudanças climáticas, podendo atuar como fonte de emissão ou remoção de CO₂ da atmosfera (BARDGETT, 2011; SMITH, 2012; WANG *et al.*, 2019).

O estoque de carbono (C) armazenado no solo é estimado em 2500 Pg (1 Pg = 1015 g) a 1 m de profundidade (JOHNSTON *et al.*, 2004a; LAL, 2008b). É constituído de dois componentes: carbono orgânico do solo (COS) estimado em 1550 Pg e carbono inorgânico do solo (CIS) com cerca 950 Pg (BATJES, 1996), representando juntos mais de três quartos do C da Terra e 4,5 vezes a quantidade de C contida na vegetação (LAL, 2004a). O estoque de COS inclui húmus altamente ativo e carvão (carbono pirogênico) relativamente inerte, resultante de resíduos orgânicos em vários estágios de decomposição sintetizados microbiologicamente e/ou quimicamente (LAL, 2008b; SCHNITZER, 1991).

Os solos são responsáveis por manter o terceiro maior estoque global de carbono (JOHNSTON *et al.*, 2004b; LAL, 2008c). Por isso, O COS desempenha um papel vital no balanço de CO₂ e outros gases de efeito estufa (MAHMOUDZADEH *et al.*, 2020). Além disso, o manejo de COS influencia a capacidade de retenção e disponibilidade de água no solo (GAISSER; GRAEF; CORDEIRO, 2000; VALARINI *et al.*, 2003). Devido aos limitantes hídricos são importantes controladores de inúmeras funções e serviços ecossistêmicos centrados no solo em ecossistemas secos, incluindo o sequestro de CO₂ da atmosfera (LAL, 2004b, 2019). Até o momento sabe-se que 19% do estoque de COS mundial estão em zonas secas, variando entre 470±7 Pg a 1 m a 646±9 Pg a 2 m de profundidade (PLAZA *et al.*, 2018). Embora no estoque total de C possuir mais CIS 578±8 Pg a 1 m e 1237±15 (LAL, 2019).

A distribuição espacial dos nutrientes do solo é descontínua, apresentando diferenças gerais, constituindo variações no seu arranjo (WANG *et al.*, 2019). Investigar a variação espacial dos nutrientes do solo pode fornecer dados valiosos e orientação teórica para apoiar as atividades humanas na manutenção da segurança alimentar e melhorar a qualidade ambiental no contexto das mudanças ambientais globais, especialmente em ambientes ecológicos frágeis (GUAN *et al.*, 2017; SCHILLACI *et al.*, 2017; ZERAATPISHEH *et al.*, 2019).

Entretanto, os métodos tradicionais de mapeamento de solos têm-se demonstrado dependentes de grandes investimentos, além de serem pesquisas demoradas e de difícil atualização (MAHMOUDZADEH *et al.*, 2020; Taghizadeh-Mehrjardi *et al.*, 2020). Por isso, é essencial uma previsão precisa do teor de C, N, P com base em abordagens robustas e de baixo custo (KHEIR *et al.*, 2010). O mapeamento digital do solo (MDS) pode ser uma ferramenta útil para reduzir esse tipo de esforço e obter resultados razoáveis. Trata-se de um procedimento orientado a dados, que busca automatizar a construção de modelos analíticos e simulações numéricas capazes de inferir as variações espaço-temporal dos tipos e propriedades dos solos, relacionando as variáveis-alvo (geralmente amostras de campo geolocalizadas) e covariáveis ambientais (por exemplo, modelos digitais de elevação e seus atributos) (HENGL; MACMILLAN, 2019; LAGACHERIE; MCBRATNEY, 2006).

A aprendizagem de máquina (ML na sigla em inglês) é entendida como uma estrutura algorítmica unificada projetada para identificar modelos computacionais que descrevem com precisão dados empíricos e os fenômenos subjacentes a eles (WATT *et al.*, 2020). A aplicabilidade dos algoritmos de ML no MDS cresceu rapidamente na última década (PADARIAN *et al.*, 2020). Pois, elas foram capazes de trabalhar com volumes grandes de variáveis e casos, formando uma análise contínua da variabilidade do solo, além de lidar melhor com relações não lineares complexas presentes nos dados de entrada (LAMICHHANE *et al.*, 2019; LIAKOS *et al.*, 2018; PADARIAN *et al.*, 2020). Neste contexto, estudos recentes em todo o mundo demonstram o potencial desses algoritmos para prever propriedades nutricionais do solos (CHEN *et al.*, 2019; HOUNKPATIN *et al.*, 2018; LIESS *et al.*, 2016; MAHMOUDZADEH *et al.*, 2020; SILATSA *et al.*, 2020; SZATMÁRI; PÁSZTOR, 2019; TAJIK *et al.*, 2020; WANG *et al.*, 2018; ZHOU *et al.*, 2020a; ZHU *et al.*, 2019).

Na Caatinga ainda são escassos os trabalhos que buscaram compreender os controladores da variabilidade espacial dos C, N, P no solo e, menos ainda, os que exploraram o uso da modelagem por ML. Assim, existe a necessidade de desenvolver estudos em múltiplas

escalas espaciais e temporais, que forneça informações fundamentais ainda em falta. O presente trabalho pode gerar uma base teórica pertinente para preencher lacunas dos esquemas de gestão sustentável de ambientes ecológicos frágeis, como: (a) adequação do uso do solo; (b) restauração da fertilidade do solo; (c) orientação das atividades agropastoris; (d) provisão de serviços ecossistêmicos; (e) diagnóstico e recuperação de áreas desertificadas. Também poderá favorecer o desenvolvimento de previsões dos impactos potenciais das mudanças climáticas, para melhor subsidiarem estratégias capazes de aumentar a resiliência dos ecossistemas secos.

Neste último contexto, destacamos a iniciativa global “4 por 1000: os solos em prol da segurança alimentar e do clima”, lançada na 21^a Conferência das Partes da Convenção-Quadro das Nações Unidas sobre a Mudança do Clima (a COP-21 da Convenção do Clima, UNFCCC na sigla em inglês), que objetiva aumentar os estoques globais de matéria orgânica do solo em 0,4% por ano, como compensação pelas emissões globais de gases de efeito estufa por fontes antropogênicas (MINASNY *et al.*, 2017). É previsto ainda explicitamente a inclusão do C, destacando o papel central dos solos na neutralidade da degradação da terra em regiões secas, presente na meta 15.3 dos Objetivos de Desenvolvimento Sustentável (ODS) e da Organização das Nações Unidas (ONU), vinculada à Convenção das Nações Unidas para o Combate à Desertificação (UNCCD na sigla em inglês).

Neste contexto, objetivamos prever a variabilidade espacial de propriedades nutricionais do solo (C, N, C:N, P) na Caatinga a partir dados legados e avaliar a capacidade preditiva de variáveis ambientais e geográficas.

4.2 MATERIAL E MÉTODOS

4.2.1 Área de Estudo

A Caatinga está inserida na região Nordeste do Brasil (NEB) e em uma pequena parte da região Sudeste (IBGE, 2019; MORO *et al.*, 2014; AB’SABER, 1977; PRADO, 2003). Constitui-se por uma área de 862.818 km² (IBGE, 2019), correspondendo 10% do território brasileiro. A Caatinga é circundada pela Mata Atlântica e Cerrado, representando 54% do NEB, abarcando grandes áreas dos estados do Ceará (CE), Alagoas (AL), Bahia (BA), Maranhão (MA), Paraíba (PB), Pernambuco (PE), Piauí (PI), Rio Grande do Norte (RN) e Sergipe (SE). Além disso, o bioma também está presente em parte do norte do estado de Minas Gerais (MG), na região Sudeste (GANEM *et al.*, 2020).

Inserida entre a linha do Equador e o Trópico de Capricórnio, isso proporciona a Caatinga uma alta radiação solar anual (MORO *et al.*, 2016), o que influencia na temperatura média que se concentra entre 25°C e 30°C, não existindo diferença entre os meses de inverno e verão. Quanto ao período chuvoso, grande parte da precipitação não é armazenada, devido a condição do solo pouco profundo, ocasionando altos índices de evaporação. Quanto a umidade relativa, é baixa, com percentuais entorno de 50% (GANEM, 2017).

A Caatinga é um dos biomas mais degradados do país, principalmente pelo uso inadequado da vegetação e do solo. Acredita-se que a Caatinga perdeu metade de sua cobertura original, ocasionado pela ação antrópica para uso agropecuários, uso madeireira e consumo de lenha (SILVA *et al.*, 2017; ANTONGIOVANNI *et al.*, 2020).

4.2.2 Procedimentos Metodológico

O nosso fluxo de trabalho geral do mapeamento das propriedades do solo foi baseado no conceito de MDS – Mapeamento Digital do Solo (MCBRATNEY *et al.*, 2003), seguindo as orientações da literatura especializada (BOETTINGER *et al.*, 2010; GARG *et al.*, 2020; HARTEMINK *et al.*, 2008; HENGL; MACMILLAN, 2019; MALONE *et al.*, 2017; ZHANG *et al.*, 2016). Os métodos tradicionais de mapeamento de solos têm-se demonstrado dependentes de grandes investimentos, além de serem pesquisas demoradas e de difícil atualização (MAHMOUDZADEH *et al.*, 2020).

4.2.2.1 Levantamento de Dados

Os dados de treinamento de 1311 perfis adquiridos a partir do banco de dados Pronassolos: <<https://geoportal.cprm.gov.br/pronasolos/>>. Estes, foram complementados com novos levantamentos realizados de 2017-2022 através de amostragem simples nos horizontes superficiais do solo em diferentes sítios no Alto curso do rio Paraíba/PB. A localização foi registrada por meio de GPS e as análises físicas e químicas foram realizadas pelo Laboratório de Rotina de Solo e Planta/ Departamento de Solos da Universidade Federal de Viçosa (??).

As previsões das propriedades do solo foram realizadas na profundidade equidistante de 0-20 cm. Os dados foram harmonizados verticalmente ao longo de 200 cm de profundidade aplicando a função *spline* quadrática de preservação de massa de área igual (BISHOP *et al.*, 1999) do pacote *mpspline2* v. 0.1.6 disponível no R 4.2.2 (O'BRIEN, 2022). Para cada *spline*, usa-se um valor de parâmetro de suavização de 0,1, os valores são gerados a cada cm.

Em seguida, o spline calcula a média para cada propriedade do solo de acordo com as profundidades padrão.

4.2.2.2 Análise Laboratorial

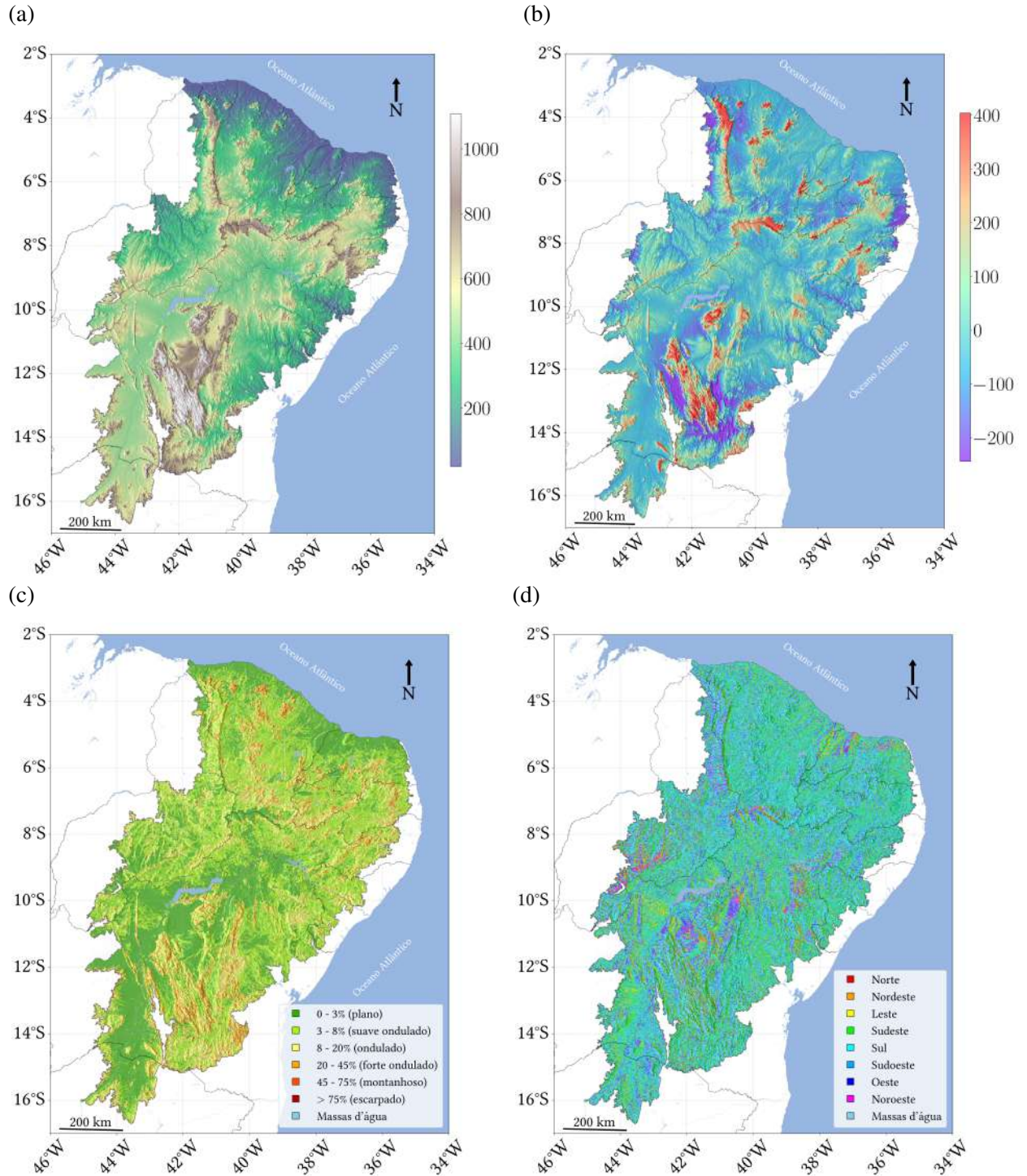
Em laboratório, as amostras coletadas dos horizontes do solo foram secas ao ar, destorroadas e tamisadas em peneira de 2 mm. O teor de C orgânico do solo foi determinado por digestão ácida de acordo com o método *Walkey-Black* (SILVA; MENDONÇA, 2007). O nitrogênio total (N) foi determinado pelo método de *Kjeldahl* e titulação (DONAGEMA *et al.*, 2011). O teor de fósforo disponível (PM) foi determinado por uma solução de extração de *Mehlich-1*. A capacidade de adsorção de P do solo foi determinada após agitação durante 1 hora com 2,5 g de solo em CaCl_2 0,01 M contendo 60 mg de P L-1. A suspensão foi filtrada e o restante P em solução (PREM) determinado por fotolorimetria (ALVAREZ *et al.*, 2000).

4.2.2.3 Covariáveis Preditoras

Para a gerar as covariáveis preditoras foi usado a plataforma de computação nuvem *Google Earth Engine* (GEE): <<https://earthengine.google.com>>, que possibilita utilizar um grande banco de dados geoespacial em escala de petabytes disponíveis gratuitamente, aplicando-se uma variedade de algoritmos em alta performance (GORELICK *et al.*, 2017).

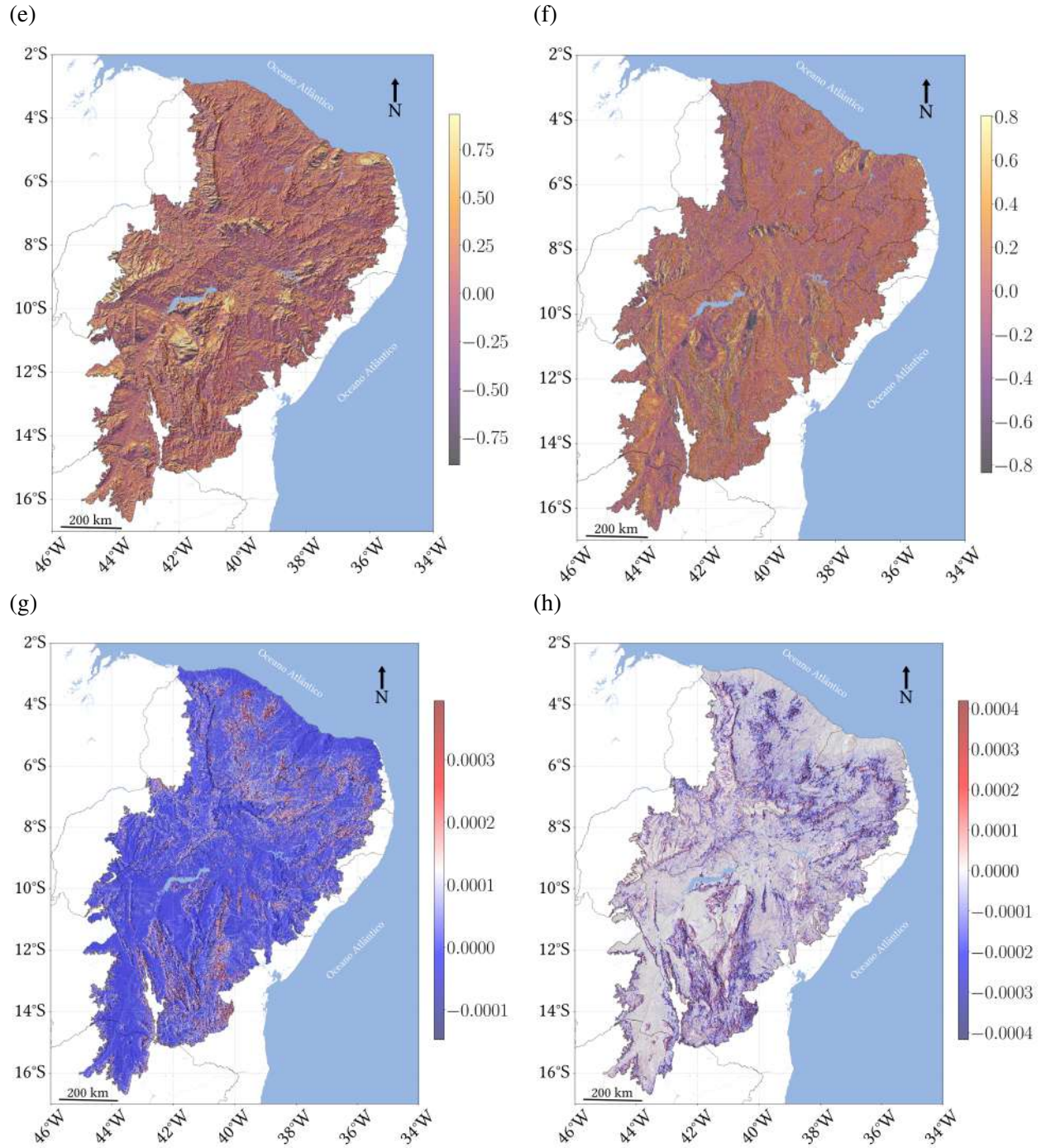
Os atributos do terreno foram óbitos a partir do Modelo Digital de Elevação (MDE) FABDEM (*Forest and Buildings removed Copernicus DEM*) (HAWKER *et al.*, 2022). Foi gerado o área de drenagem acumulada (SCA) usando o módulo *Flow Width and Specific Catchment Area* e o método de saída foi o *Multiple Flow Direction* (MFD) (QUINN *et al.*, 1991) no software Saga 7.8.2. O índice de posição topográfica (TPI) (Figura 17b) com 50 vizinhanças foi obtido com o algoritmo *focalMean* do GEE. Foi usado o pacote *Terrain Analysis in GEE* (TAGEE) para calcular os seguintes atributos derivados do MDE: declividade (Figura 17c), aspecto (Figura 17d), índice de exposição ao norte 17e, índice de exposição ao leste (Figura 17f), curvatura horizontal (Figura 17g) e a curvatura vertical (Figura 17h) (SAFANELLI *et al.*, 2020).

Figura 17 – Atributos derivados do modelo digital de elevação: (a) hipsometria (m), (b) índice de posição topográfica (TPI), (c) declividade, (d) aspecto.



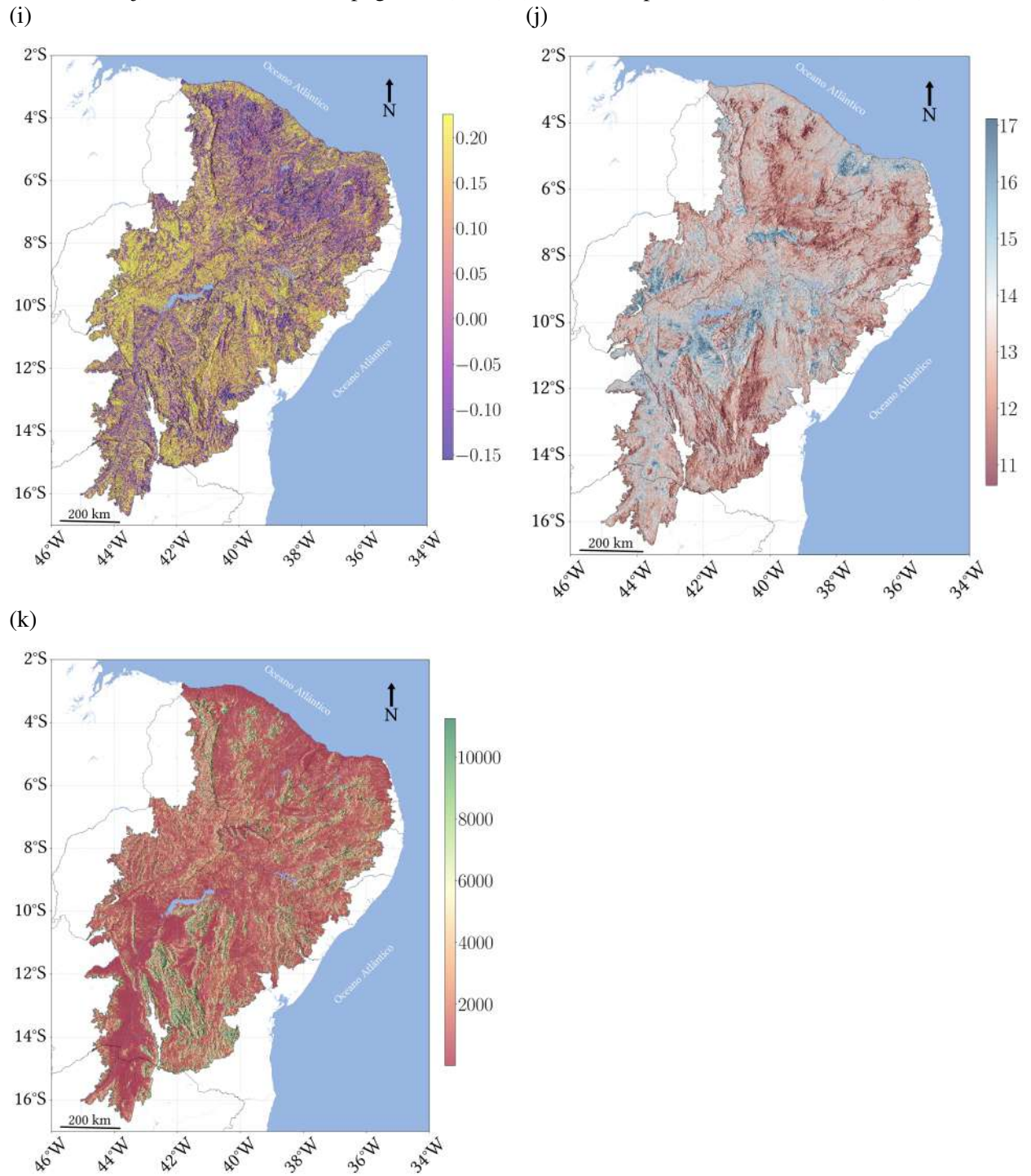
Fonte: Elaborada pelo autor.

Figura 17 – (Cont.) Atributos derivados do modelo digital de elevação: (e) índice de exposição ao norte, (f) índice de exposição ao leste, (g) curvatura horizontal, (h) curvatura vertical.



Fonte: Elaborada pelo autor.

Figura 17 – (Cont.) Atributos derivados do modelo digital de elevação: (i) índice de forma (*Shape Index*),
(j) índice de umidade topográfica (TWI), (k) índice de potência de escoamento (SPI).



Fonte: Elaborada pelo autor.

As variáveis bioclimáticas de 1970 a 2000 do *WorldClim – Global Climate Data* (FICK; HIJMAN, 2017) foram importadas para o GEE, pois não se encontra em seu banco de dados. Esse conjunto de dados climáticos globais tem resolução espacial de 1 km². Foram obtidos mais dois atributos do clima no catálogo da comunidade *awesome-GEE*: evapotranspiração de referência (ET0) (projects/sat-io/open-datasets/global_et0/global_et0_yearly) e índice de aridez (IA) (projects/sat-io/open-datasets/global_ai/global_ai_yearly). O IA e a ET0 foram relacionados aos processos de evapotranspiração e déficit de chuva para crescimento vegetativo potencial, baseado na implementação de uma equação de evapotranspiração de *Penman Monteith* para cultura de referência (ALLEN *et al.*, 1998; ZOMER *et al.*, 2022).

A temperatura da superfície terrestre (Ts: MOD11A1 e MOD11A2), produtividade primária bruta (GPP) (Figura 19c) e líquida (NPP) (Figura 19d) anual (MOD17A3HGF) foram obtidas do sensor MODIS (*Moderate-Resolution Imaging Spectroradiometer*). Foram gerados valores médios de LST, GPP e NPP para o período de 2000-2022.

Os dados de satélite incluem os seguintes conjuntos de dados/coleções-GEE ID: Landsat 5 TM (LANDSAT/LT05/C02/T1_L2), Landsat 7 ETM+ (LANDSAT/LE07/C02/T1_L2) e Landsat 8 OLI/TIRS (LANDSAT/LC08/C02/T1_L2) com resolução de 30 m e ciclo total de varredura global de 16 dias. Filtramos imagens de janeiro a julho com uma cobertura de nuvem terrestre máxima de 60%. Esses meses representam o pico da estação chuvosa a fim de refletir a cobertura vegetal no nível ótimo. As imagens do Landsat foram mascaradas para nuvens e sombras de nuvens usando a faixa de avaliação de qualidade do produto SR gerado a partir do algoritmo CFMASK. Posteriormente, foi calculada uma mediana geométrica para cada ano. Após isso, foi gerada uma mediana histórica (1985-2020). Os índices espectrais calculados no GEE foram:

O Índice de Vegetação por Diferença Normalizada (NDVI) (TUCKER, 1979) (Figura 18a), é calculado pela equação 4.1:

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad (4.1)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{RED} = banda do vermelho.

O Índice de Vegetação Aprimorada (EVI) (HUETE *et al.*, 2002) (Figura 18c), é dada pela equação 4.2:

$$EVI = 2.5 \times \left(\frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + 6 \times \rho_{RED} - 7.5 \times \rho_{BLUE} + 1} \right) \quad (4.2)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{RED} = banda do vermelho;
- ρ_{BLUE} = banda do azul.

O Índice de Água por Diferença Normalizada (NDWI) (GAO, 1996) (Figura 18g), é obtida pela equação 4.3:

$$NDWI = \frac{\rho_{NIR} - \rho_{SWIR1}}{\rho_{NIR} + \rho_{SWIR1}} \quad (4.3)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{SWIR1} = banda do infravermelho médio.

O Índice de Vegetação Ajustada ao Solo Modificado (MSAVI2) (QI *et al.*, 1994) (Figura 18b), é gerado com base na equação 4.4:

$$MSAVI2 = \frac{2 \times \rho_{NIR} + 1 - \sqrt{(2 \times \rho_{NIR} + 1)^2 - 8 \times (\rho_{NIR} - \rho_R)}}{2} \quad (4.4)$$

O Índice de Vegetação por Diferença Normalizada Verde (GNDVI) (GITELSON *et al.*, 1996) (Figura 18e), é obtido conforme a equação 3.19:

$$GDVI = \frac{\rho_{NIR} - \rho_{GREEN}}{\rho_{NIR} + \rho_{GREEN}} \quad (4.5)$$

Em que:

- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{GREEN} = banda do verde.

O Albedo (LIANG *et al.*, 2003) (Figura 18f), é calculado pela equação 4.6:

$$ALBEDO = 0.356 \times \rho_{BLUE} + 0 \times \rho_{GREEN} + 0.13 \times \rho_{RED} + 0.373 \times \rho_{NIR} + 0.085 \times \rho_{SWIR1} + 0.072 \times \rho_{SWIR2} - 0.0018 \quad (4.6)$$

Em que:

- ρ_{BLUE} = banda do azul;
- ρ_{GREEN} = banda do verde;
- ρ_{RED} = banda do vermelho;
- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{SWIR1} = banda do infravermelho médio;
- ρ_{SWIR2} = banda do infravermelho médio.

O *Wettness* da transformação *Tasseled Cap* (TCW) (CRIST, 1985) (Figura 18h), é dada pela equação 4.7:

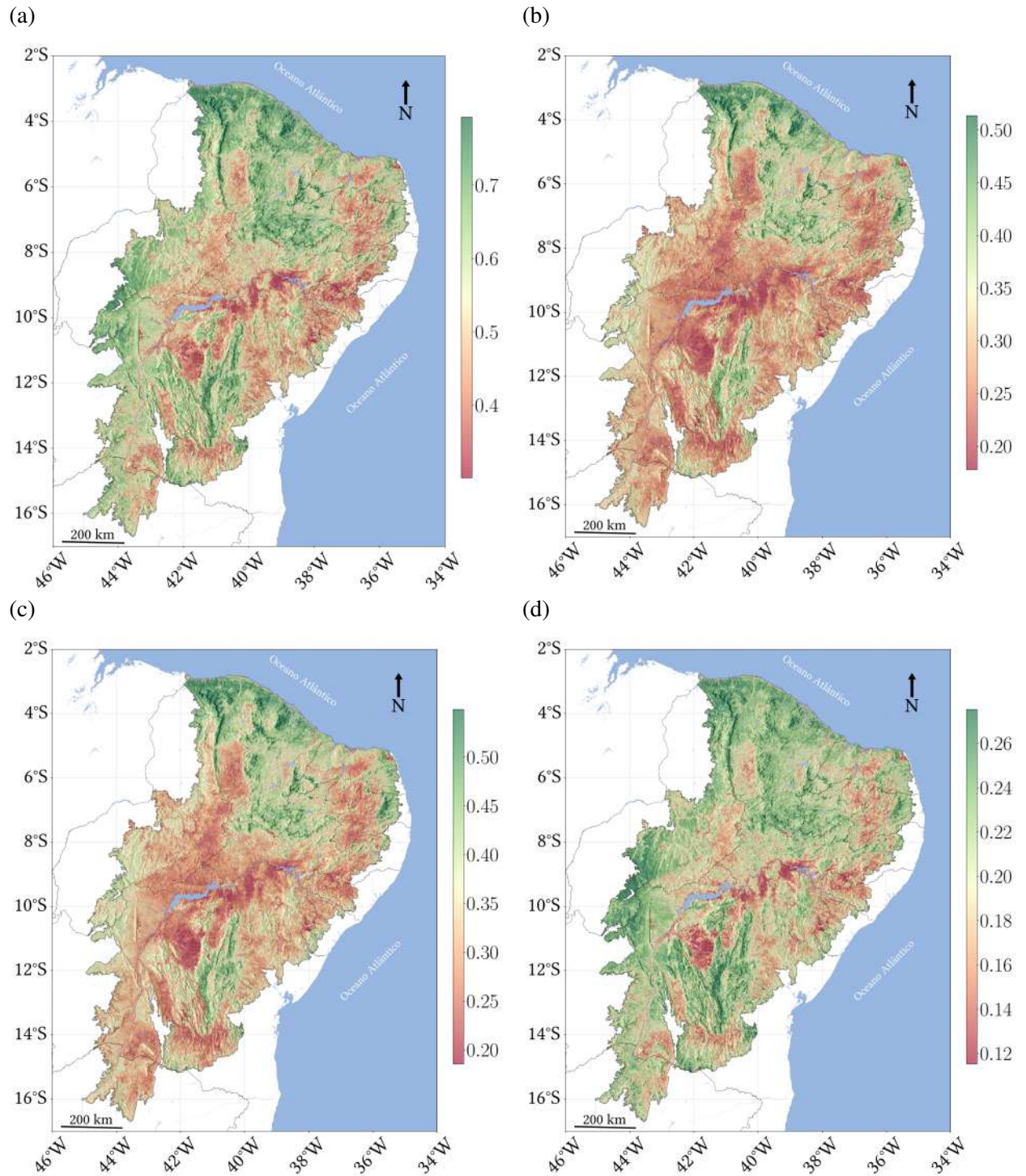
$$\text{TCW} = 0.1509 \times \text{BLUE} + 0.1973 \times \text{GREEN} + 0.3279 \times \text{RED} + 0.3406 \times \text{NIR} + 0.7112 \times \text{SWIR1} + 0.4572 \times \text{SWIR2} \quad (4.7)$$

Em que:

- ρ_{BLUE} = banda do azul;
- ρ_{GREEN} = banda do verde;
- ρ_{RED} = banda do vermelho;
- ρ_{NIR} = banda do infravermelho próximo da imagem do satélite;
- ρ_{SWIR1} = banda do infravermelho médio;
- ρ_{SWIR2} = banda do infravermelho médio.

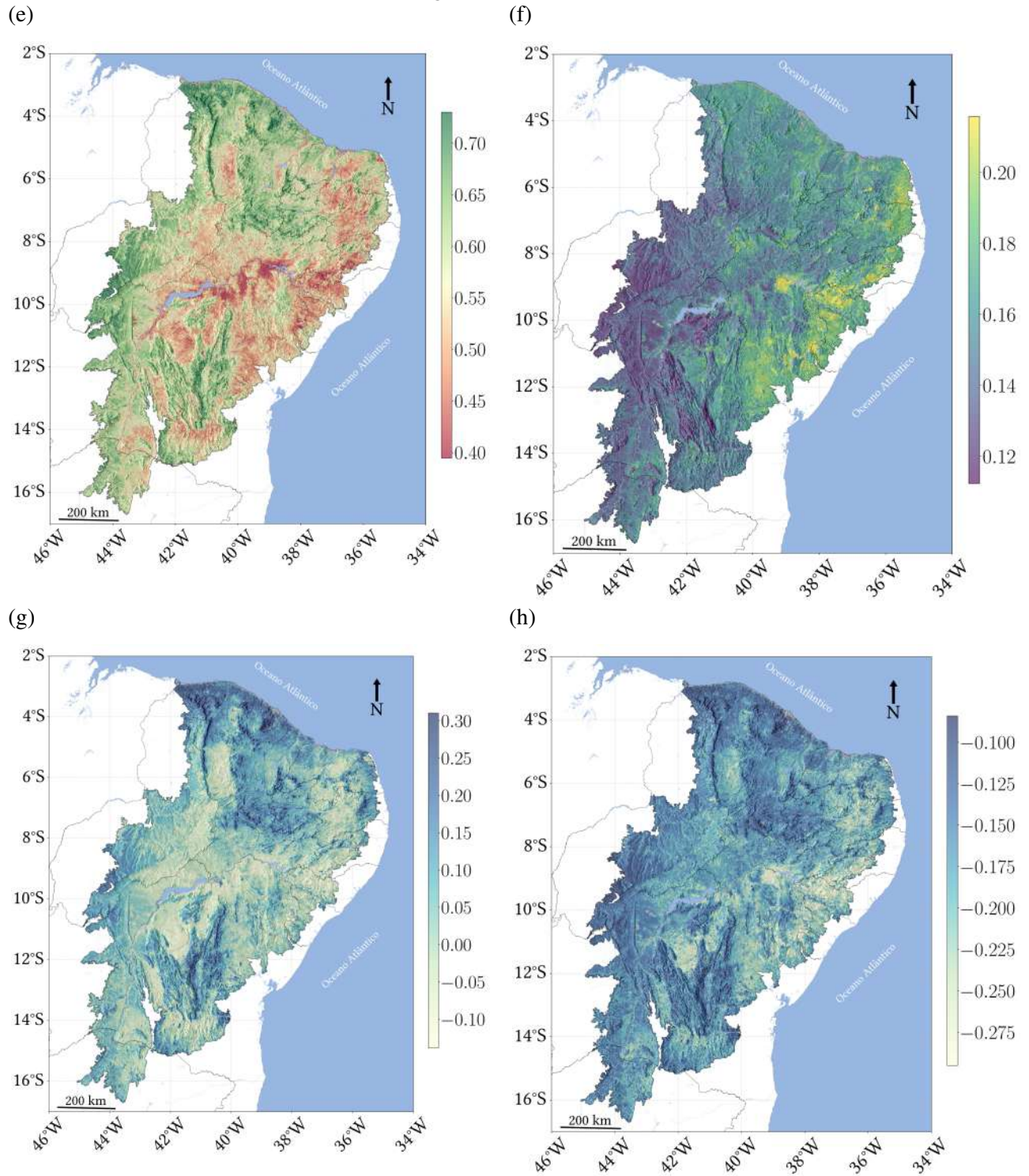
Todas as covariáveis foram, finalmente, reajustadas para 30 metros de resolução espacial, mascaradas com corpos hídricos, praias, dunas, areais, áreas urbanas e mineradoras, rochas expostas e outras superfícies não vegetadas, usando o uso e cobertura da terra do ano de 2021 do projeto MapBiomias <<https://mapbiomas.org/>> (SOUZA *et al.*, 2020).

Figura 18 – Mediana histórica (1985-2020) dos índices espectrais derivados dos satélites Landsat 5, 7 e 8: (a) NDVI, (b) MSAVI2, (c) EVI, (d) CFLUX.



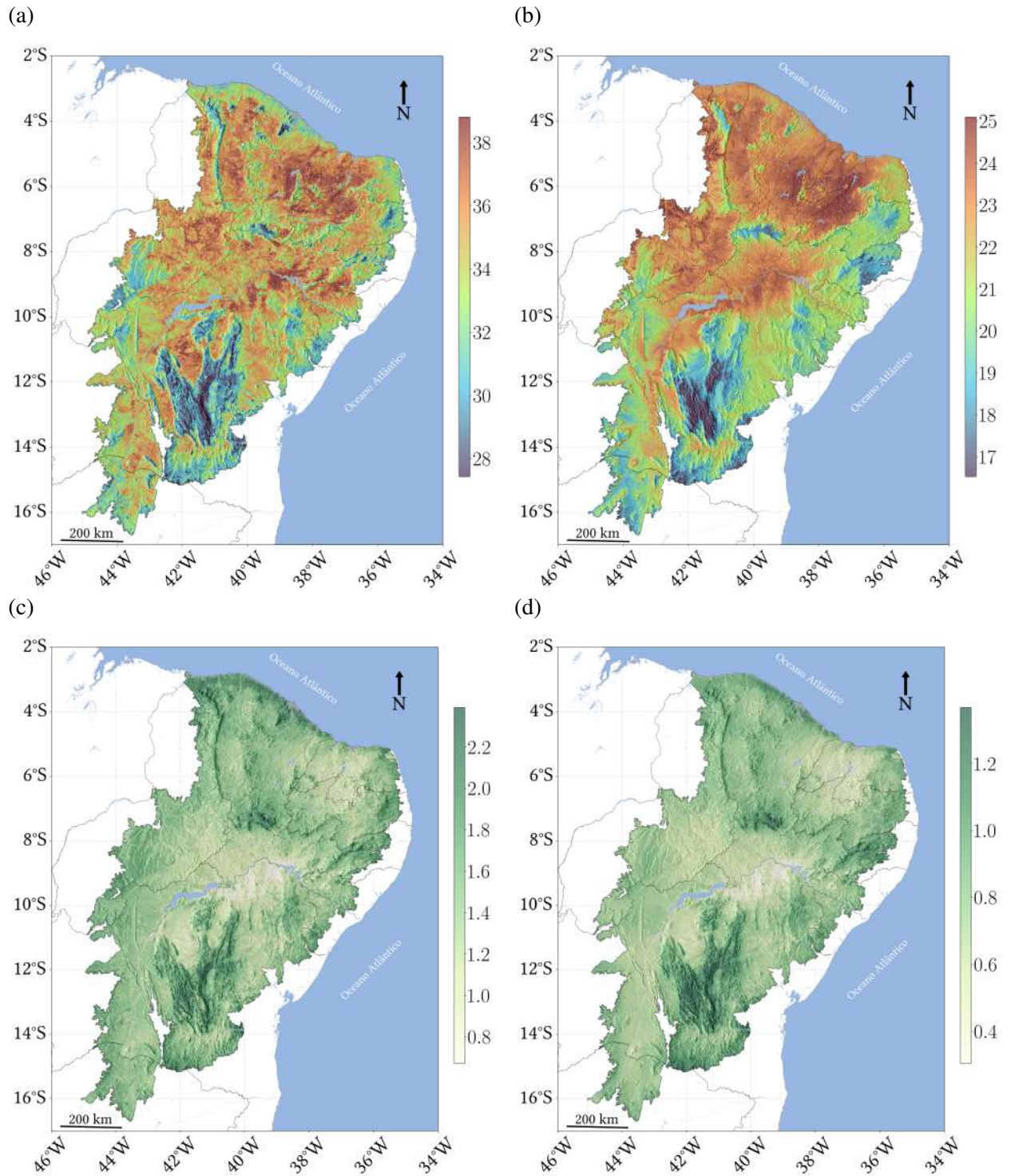
Fonte: Elaborada pelo autor.

Figura 18 – (Cont.) Mediana histórica (1985-2020) dos índices espectrais derivados dos satélites Landsat 5, 7 e 8: (e) GNDVI, (f) Albedo, (g) NDWI, (h) TCW.



Fonte: Elaborada pelo autor.

Figura 19 – Média histórica da: (a) temperatura da superfície (Ts) diária e (b) Ts da noturna (2000-2021), (c) GPP e (d) NPP (2001-2021).



Fonte: Elaborada pelo autor.

4.2.2.4 Treinamento, Avaliação e Predição Espacial

Os dados de solos antes de ser usados na modelagem foram transformados usando o logaritmo natural para se aproximar de uma distribuição normal. Logo após foram divididos em dados treino (75%) e teste (25%). As covariáveis que tiveram alta correlação de *Spearman* $>|0,95|$ foram descartadas. Se dois preditores tiverem uma correlação alta, remove-se o preditor com a maior correlação média (SIQUEIRA *et al.*, 2023).

O algoritmo de regressão *Random Forest* (RF) (BREIMAN, 2001) foi usado para estimar as propriedades do solo na Caatinga. O modelo de regressão RF, usado para modelar variáveis contínuas, usa a média das previsões de um conjunto de árvores de decisão para as versões *bootstrap* dos dados de treinamento (HASTIE *et al.*, 2013). Usamos a biblioteca *scikit-learn* 1.2.2 através da função *RandomForestRegressor* no *Python* 3.8.

Para cada propriedade do solo e intervalo de profundidade foi feita uma validação cruzada *k-fold* de 10 vezes. Primeiramente as amostras são divididas aleatoriamente em 10 subconjuntos (*k* grupos) de 10% das amostras totais. Cada conjunto de dados menos um (*k*-1) são usados para treinamento do modelo e, em seguida, avalia-se esse modelo no conjunto de dados que não foi usado para treinamento. Esse processo é repetido 10 vezes, com um subconjunto diferente reservado para avaliação (e excluído do treinamento) a cada vez.

O modelo de RF foi explanado com valores de *Shapley* (SHAPLEY, 1953; WADOUX; MOLNAR, 2022). Os valores de *Shapley* utiliza uma abordagem baseada na teoria dos jogos, ou seja, considera um jogo em que cada covariável é um jogador e a previsão é o pagamento; os valores de *Shapley* distribuem o ganho para os jogadores (covariáveis) de acordo com sua participação relativa no resultado. O ganho é a previsão para uma determinada unidade menos a previsão média, e os jogadores são as covariáveis que contribuem para a previsão e “colaboram” para receber um ganho (WADOUX *et al.*, 2023). Foi usado a biblioteca *SHAP* no *Python* 3.8.

Para avaliar a precisão dos dados previstos, foram utilizadas duas métricas de medição de acurácia: Coeficiente de Determinação (R^2) (Equação 4.8), Coeficiente de Correlação de Concordância de Lin (LCCC: (*Lin's Concordance Correlation Coefficient*)) (Equação 4.9), Erro Absoluto Médio (MAE: (*Mean Absolute Error*)) (Equação 4.10), Raiz Quadrada do Erro-Médio (RMSE: *Root-Mean-Square Error*) (Equação 4.11) e Razão de Desempenho para Distância Interquartil (RPIQ: *Ratio of Performance to InterQuartile distance*) (Equação 4.12):

$$R^2 = \frac{(\sum_{i=1}^n (Y_i - \bar{Y})(X_i - \bar{X}))^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2 \sum_{i=1}^n (X_i - \bar{X})^2} \quad (4.8)$$

$$LCCC = \frac{2 \cdot r \cdot (S_x \cdot S_y)}{S_x^2 + S_y^2 + (x_{\text{bar}} - y_{\text{bar}})^2} \quad (4.9)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - X_i| \quad (4.10)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - X_i)^2}{n}} \quad (4.11)$$

$$RPIQ = 1 - \frac{6 \sum_{i=1}^n (R(X_i) - R(Y_i))^2}{n(n^2 - 1)} \quad (4.12)$$

Em que:

- X e Y representam valores medidos e previstos, respectivamente;
- n é o número de amostras;
- X_i e Y_i são os valores i emparelhados dos dados medidos e previstos, respectivamente;
- $\bar{X}$ e $\bar{Y}$ são a média dos dados previstos e observados, respectivamente;
- S_x é o desvio padrão dos valores verdadeiros;
- S_y é o desvio padrão dos valores previstos;
- r é o coeficiente de correlação de Pearson entre os valores observados e os valores previstos (*pred*).

4.2.3 Análise Estatística

A média aritmética é calculado utilizando-se a equação 4.13:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i \quad (4.13)$$

Em que: n é o número total de valores e x_i cada valor, onde $i = 1, \dots, n$. Média aritmética é a soma dos valores x_i dividido pelo número total de valores n.

Elemento central para o qual se contém 50% das observações.

O Desvio Padrão (DP) é calculado usando-se a seguinte fórmula:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (4.14)$$

O Coeficiente de Variação (CV) é definido como a razão do desvio padrão σ pela média μ :

$$cv = \frac{\sigma}{\bar{x}} \quad (4.15)$$

A correlação de *Spearman* foi aplicada aos dados de solo e parâmetros estruturais da vegetação (vide Tópico ??)

4.3 RESULTADOS E DISCUSSÃO

4.3.1 Estatística Descritiva dos Dados Amostrais

As estatísticas descritivas são apresentadas na Tabela 2. Os resultados do estudo revelaram uma ampla variação nos atributos químicos do solo analisados. O teor de carbono (C) apresentou uma média de $5,68 \pm 8,76$ g/kg, com uma mediana de 3,71 g/kg. Observou-se uma grande dispersão, com valores mínimos de 0 g/kg e máximos de 147,6 g/kg. O coeficiente de variação (CV) de 154,22% indica uma alta heterogeneidade nos teores de carbono do solo.

Para o teor de nitrogênio (N), a média foi de $0,62 \pm 0,47$ g/kg com uma mediana de 0,50 g/kg. Os valores variaram de 0 g/kg a 5,68 g/kg, evidenciando uma significativa amplitude nos dados. O CV foi de 75,49%, sugerindo uma heterogeneidade nesse atributo (Tabela 2).

A razão C:N apresentou uma média de $7,51 \pm 3,19$ e a mediana de 7,12. A dispersão dos dados foi relativamente baixa, com um CV de 42,49%. Isso indica que, em geral, que C:N se manteve mais estável em comparação aos outros atributos analisados (Tabela 2).

Por fim, o fósforo (P) disponível no solo revelou uma média de $13,79 \pm 46,95$ mg/kg, com uma mediana de 1,86 mg/kg. Os valores mínimos e máximos foram 0 mg/kg e 563,32 mg/kg, respectivamente, indicando uma grande variação nos dados. O CV para o fósforo foi de 340,62%, evidenciando uma alta heterogeneidade nesse atributo (Tabela 2).

Os coeficientes de assimetria positivos superiores a 1 para todas as propriedades do solo indicam um enviesamento positivo. Isso revela haver mais observações abaixo da média do que acima dela, ou seja, há um número relativamente baixo de valores altos. Quanto à curtose, ela mede o grau de pico ou achatamento de uma distribuição. Nossos dados demonstraram ser leptocúrticos, com uma curtose acima de 3. Isso indica um alto grau de alongamento, com uma curva de frequências mais estreita do que a distribuição normal. Portanto, os valores tendem a se concentrar à direita e apresentam uma distribuição mais achatada.

Tabela 2 – Estatística descritiva das propriedades nutricionais do solo agregados em diferentes intervalos de profundidade na Caatinga.

	C	N	C:N	P
N	1079	985	984	720
Mínimo	0	0	1,17	0
1Q	2,08	0,33	5,26	1
Média	5,68	0,62	7,51	13,79
Mediana	3,71	0,50	7,12	1,86
3Q	6,42	0,75	9,20	4,98
Máximo	147,39	5,68	25,47	563,32
Desvio Padrão	8,76	0,47	3,19	46,95
CV	154,22	75,49	42,49	340,62
Assimetria	9,84	3,37	1,16	6,95
Curtose	135,79	21,63	2,97	60,09

Fonte: Elaborada pelo autor.

Nota: N: amostras; 1Q: primeira quartil; 3Q: terceira quartil; CV (%): coeficiente de variação.

4.3.2 Caracterização Geoquímica dos Solos

Uma síntese da estatística descritiva dos atributos C, N, C:N e P por classe de solos da Caatinga consta nas Tabelas de 3 a 6.

A variação no número de amostras para o C de acordo com as subordens indicou que enquanto o Neossolo Litólico, Luvisolo Crômico e Latossolo Amarelo, possuem um grande número de amostras (166, 146 e 127, respectivamente), outras subordens têm um número significativamente menor de amostras, como Plintossolo Argilúvico, Plintossolo Pétrico, Vertissolo Ebânico e Vertissolo Háplico, com apenas 2 ou 3 amostras (Tabela 3).

Ao analisar as medidas de tendência central, observa-se que essas estatísticas variam entre as diferentes subordens de solos. Isso indica haver diferenças nos níveis médios dos atributos avaliados em cada subordem. Entre os solos mais representativos, a média de C para

Luvissolo Crômico é de 8,14 g/kg, enquanto para Argissolo Amarelo é de 4,90 g/kg. O Gleissolo Sálico teve a maior média (9,55 g/kg), porém encontra-se entre os menos representativos.

Além disso, a variação dos valores mínimo e máximo nos dá uma ideia da faixa de variação dos atributos para cada subordem. Por exemplo, o valor mínimo para o atributo de C no Neossolo Litólico é de 0,19 g/kg, enquanto o valor máximo é de 140,85 g/kg. Para Luvissolo Crômico os valores de C foram de 0,06 g/kg a 147,39 g/kg. Essa diferença substancial mostra a amplitude de variação dos atributos em uma mesma subordem de solo.

Entre as subordens de solo avaliadas, algumas exibem um desvio padrão relativamente reduzido, indicando uma menor dispersão dos valores em relação à média. Essa observação sugere uma maior uniformidade nos atributos avaliados dentro dessas subordens. Ilustrando essa tendência, destacam-se o Argissolo Vermelho (DP= 3,81 g/kg), o Argissolo Vermelho-Amarelo (DP= 3,19 g/kg) e o Latossolo Amarelo (DP= 3,69 g/kg).

Por outro lado, há subordens que apresentam um desvio padrão mais elevado, indicando uma maior dispersão dos valores em relação à média. Essa maior variabilidade pode refletir diferenças significativas nas propriedades do solo dentro dessas subordens. Exemplos notáveis incluem o Luvissolo Crômico (DP= 17,07 g/kg), o Cambissolo Háplico (DP= de 6,25 g/kg) e o Gleissolo Sálico (DP= 6,28 g/kg).

Tabela 3 – Estatística descritiva do teor de carbono orgânico (g/kg) da camada superficial de solos por subordem de solo.

Subordem	N	Média	Mediana	Mínimo	Máximo	DP	CV
Argissolo Amarelo	22	4,90	2,78	0,21	30,04	6,34	129,27
Argissolo Vermelho	60	5,23	3,77	0,80	22,02	3,81	72,85
Argissolo Vermelho-Amarelo	95	4,35	3,66	0,07	16,02	3,19	73,26
Cambissolo Háplico	88	6,00	4,12	0,75	36,97	6,25	104,24
Chernossolo Rêndzico	4	5,65	5,31	1,92	10,05	3,48	61,62
Gleissolo Sáfico	3	9,55	6,02	5,84	16,80	6,28	65,70
Latossolo Amarelo	127	5,54	4,63	0,18	17,16	3,69	66,66
Latossolo Vermelho	16	7,07	5,69	1,08	20,00	5,09	71,94
Latossolo Vermelho-Amarelo	110	5,37	3,26	0,41	32,02	5,33	99,29
Luvissolo Crômico	146	8,14	3,46	0,06	147,39	17,07	209,64
Neossolo Flúvico	20	4,81	4,11	0,73	11,37	3,25	67,60
Neossolo Litólico	166	5,93	3,06	0,19	140,85	11,92	200,81
Neossolo Quartzarênico	36	4,23	3,39	0,59	23,55	4,09	96,55
Neossolo Regolítico	30	4,54	3,29	0,95	18,48	3,63	79,96
Nitossolo Vermelho	4	3,92	3,46	2,06	6,72	2,05	52,28
Planossolo Háplico	101	5,05	3,37	0,00	26,56	4,74	93,84
Planossolo Nátrico	14	4,78	3,18	0,61	16,02	4,60	96,29
Plintossolo Argilúvico	3	1,98	1,76	1,43	2,74	0,68	34,47
Plintossolo Pétrico	2	1,65	1,65	1,49	1,81	0,23	13,77
Vertissolo Ebânico	2	7,55	7,55	0,59	14,52	9,85	130,36
Vertissolo Háplico	30	4,43	3,86	0,20	21,20	4,01	90,49

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação.

A mediana do teor de N variou de 0,35 g/kg para o Argissolo Amarelo a 0,66 g/kg para o Vertissolo Ebânico. Isso indica uma variação considerável nos valores de nitrogênio nas diferentes subordens de solos (Tabela 4).

O teor mínimo de N foi de 0,13 g/kg para o Argissolo Amarelo, enquanto o teor máximo foi de 5,68 g/kg para o Latossolo Amarelo. Esses valores ilustram a ampla variação observada no teor de nitrogênio entre as subordens de solos (Tabela 4).

As subordens de solos com maiores desvios padrão foram o Latossolo Amarelo (0,69 g/kg) e o Cambissolo Háplico (0,53 g/kg), indicando maior variabilidade nos valores de teor de N. Em contraste, o Gleissolo Sáfico apresentou o menor desvio padrão, com 0,09 g/kg

(Tabela 4).

O CV, expresso em porcentagem, é uma medida relativa de variabilidade que considera o desvio padrão em relação à média. Os resultados mostraram que o Plintossolo Argilúvico teve o maior CV, com 104,90%, indicando uma grande variação no teor de N em relação à média. Por outro lado, o Gleissolo Sálico teve o menor CV, com apenas 22,25% (Tabela 4).

Tabela 4 – Estatística descritiva do teor de nitrogênio (g/kg) da camada superficial de solos por subordem de solo.

Subordem	N	Média	Mediana	Mínimo	Máximo	DP	CV
Argissolo Amarelo	22	0,48	0,35	0,13	1,70	0,33	69,57
Argissolo Vermelho	60	0,54	0,46	0,07	1,84	0,30	54,93
Argissolo Vermelho-Amarelo	91	0,61	0,53	0,11	2,13	0,42	69,19
Cambissolo Háplico	88	0,72	0,57	0,03	2,72	0,53	74,03
Chernossolo Rêndzico	4	0,56	0,50	0,30	0,96	0,28	49,83
Gleissolo Sálico	3	0,38	0,38	0,30	0,47	0,09	22,25
Latossolo Amarelo	127	0,71	0,53	0,09	5,68	0,69	97,52
Latossolo Vermelho	16	0,66	0,48	0,20	1,82	0,47	70,92
Latossolo Vermelho-Amarelo	110	0,75	0,64	0,08	4,39	0,56	75,22
Luvissolo Crômico	93	0,51	0,47	0,07	1,70	0,28	54,46
Neossolo Flúvico	20	0,44	0,40	0,07	1,16	0,27	61,02
Neossolo Litólico	132	0,55	0,48	0	2,39	0,37	67,14
Neossolo Quartzarênico	36	0,55	0,49	0,18	1,19	0,28	50,40
Neossolo Regolítico	30	0,60	0,50	0,20	2,44	0,42	69,77
Nitossolo Vermelho	4	0,62	0,55	0,29	1,09	0,34	54,20
Planossolo Háplico	100	0,63	0,50	0,10	2,65	0,45	71,20
Planossolo Nátrico	12	0,48	0,46	0,20	0,93	0,20	41,12
Plintossolo Argilúvico	3	0,55	0,35	0,10	1,20	0,58	104,90
Plintossolo Pétrico	2	0,61	0,61	0,50	0,73	0,16	26,36
Vertissolo Ebânico	2	0,66	0,66	0,42	0,90	0,35	52,40
Vertissolo Háplico	30	0,65	0,50	0,05	1,72	0,45	69,32

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação.

Dentre as subordens mencionadas, os Chernossolos Rêndzicos, Gleissolos Sálcos e Plintossolos Argilúvicos apresentam as maiores médias de C:N, com valores médios de 9,09, 8,63 e 8,55, respectivamente (Tabela 5). Os Chernossolos Rêndzicos possuem uma mediana para C:N de 10,12, com valores variando entre 4,11 e 12,02. Os Gleissolos Sálcos têm uma mediana de 9, com intervalo de valores entre 5,07 e 11,83 (Tabela 5). Já os Plintossolos Argilúvicos possuem mediana de 9,29 de C:N, com valores mínimos de 6,02 e máximos de 10,33. Por outro lado, o Latossolo Amarelo, com média de 8,44 e mediana de 7,95, possui uma faixa mais ampla, variando de 2 a 25,47. O Neossolo Litólico, com média de 7,93 e mediana de 7,44, apresenta valores mínimos de 1,76 e máximos de 16,92 de C:N (Tabela 5).

Tabela 5 – Estatística descritiva da razão C:N da camada superficial de solos por subordem de solo.

Subordem	N	Média	Mediana	Mínimo	Máximo	DP	CV
Argissolo Amarelo	22	6,15	6,32	2,41	14,92	2,88	46,84
Argissolo Vermelho	60	7,75	7,42	2,65	21,40	3,09	39,85
Argissolo Vermelho-Amarelo	91	7,74	6,83	1,26	17,51	3,44	44,52
Cambissolo Háplico	88	7,67	7,19	2,14	22,27	3,32	43,35
Chernossolo Rêndzico	4	9,09	10,12	4,11	12,02	3,65	40,10
Gleissolo Sílico	3	8,63	9	5,07	11,83	3,40	39,33
Latossolo Amarelo	127	8,44	7,95	2	25,47	3,71	43,97
Latossolo Vermelho	16	7,86	7,58	2,52	13	2,83	36,02
Latossolo Vermelho-Amarelo	110	6,27	5,72	1,72	22,95	2,82	45,05
Luvissolo Crômico	93	6,54	6,17	1,17	13,56	2,49	38,12
Neossolo Flúvico	20	7,21	7,51	3,44	10,81	1,95	27,01
Neossolo Litólico	131	7,93	7,44	1,76	16,92	3,25	40,99
Neossolo Quartzarênico	36	7,71	7,28	2,87	12,96	2,49	32,34
Neossolo Regolítico	30	7,34	7,49	2,46	12,33	2,61	35,55
Nitossolo Vermelho	4	7,16	4,69	3,12	16,15	6,06	84,54
Planossolo Háplico	100	7,77	7,44	2,57	23,94	3,33	42,85
Planossolo Nátrico	12	7,74	7,91	2,07	16,72	3,63	46,93
Plintossolo Argilúvico	3	8,55	9,29	6,02	10,33	2,25	26,31
Plintossolo Pétrico	2	7,62	7,62	4,76	10,49	4,05	53,17
Vertissolo Ebânico	2	7,20	7,20	5,75	8,65	2,05	28,50
Vertissolo Háplico	30	7,22	7,64	2,28	12,40	2,39	33,04

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação.

As maiores médias de P foram registrados nas seguintes subordens: Neossolo Flúvico, Neossolo Regolítico, Latossolo Vermelho-Amarelo, Neossolo Litólico, Argissolo Vermelho-Amarelo e Cambissolo Háplico (Tabela 6). Para o Neossolo Regolítico, com uma amostra de 23 observações, constatou-se uma média de 29,03 mg/kg, uma mediana de 3 mg/kg e uma variação que se estende de 0,97 mg/kg a 444,01 mg/kg, com um desvio-padrão de 92,15 mg/kg.

O Neossolo Flúvico, por sua vez, com uma amostra de 13, apresenta uma média de 31,84 mg/kg, uma mediana de 1,75 mg/kg e uma faixa que abrange de 0,28 mg/kg a 275,44 mg/kg, com um desvio-padrão de 78,84 mg/kg.

Para o Latossolo Vermelho-Amarelo, observou-se uma amostra de 94, com uma média de 21,36 mg/kg, uma mediana de 1,99 mg/kg e uma amplitude que varia de 0 mg/kg a 563,32 mg/kg, acompanhada de um desvio-padrão de 69,64 mg/kg.

O Neossolo Litólico, representado por uma amostra de 91, apresenta uma média de 20,67 mg/kg, uma mediana de 1,89 mg/kg e uma variação que se estende de 0 mg/kg a 509,09 mg/kg, com um desvio-padrão de 67,1 mg/kg.

Para o Argissolo Vermelho-Amarelo, com uma amostra de 64, foram encontrados uma média de 19,98 mg/kg, uma mediana de 1,35 mg/kg e uma faixa que compreende de 0 mg/kg a 239,3 mg/kg, juntamente com um desvio-padrão de 49,42 mg/kg.

Por fim, o Cambissolo Háplico, com uma amostra de 68, exibe uma média de 13,91 mg/kg, uma mediana de 1,97 mg/kg e uma amplitude que varia de 0,3 mg/kg a 296,89 mg/kg, com um desvio-padrão de 43,7 mg/kg. Essas medidas estatísticas proporcionam uma compreensão mais precisa das características e distribuição dos dados associados a cada subordem de solo mencionada, considerando a unidade de medida em mg/kg (Tabela 6).

Tabela 6 – Estatística descritiva do fósforo disponível (mg/kg) da camada superficial de solos por subordem de solo.

Subordem	N	Média	Mediana	Mínimo	Máximo	DP	CV
Argissolo Amarelo	13	3,20	1,28	0,65	15,62	4,05	126,63
Argissolo Vermelho	50	5,15	1,84	0,68	68,01	11,04	214,52
Argissolo Vermelho-Amarelo	64	19,98	1,35	0	239,30	49,42	247,34
Cambissolo Háplico	68	13,91	1,97	0,30	296,89	43,70	314,05
Chernossolo Rêndzico	3	4,68	4,12	0,92	9	4,07	86,93
Gleissolo Sáfico	2	2,65	2,65	1	4,30	2,33	88,07
Latossolo Amarelo	101	9,97	2	0,02	131,24	21,77	218,38
Latossolo Vermelho	9	3,54	1	0,53	17,04	5,42	153,30
Latossolo Vermelho-Amarelo	94	21,36	1,99	0	563,32	69,64	326,05
Luvissolo Crômico	58	4,62	2	0	41,68	7,08	153,14
Neossolo Flúvico	13	31,84	1,75	0,28	275,44	78,84	247,60
Neossolo Litólico	91	20,67	1,89	0	509,09	67,10	324,70
Neossolo Quartzarênico	29	9,47	2	0	166,21	30,57	322,95
Neossolo Regolítico	23	29,03	3	0,97	444,01	92,15	317,39
Nitossolo Vermelho	3	1,37	1,33	0,96	1,81	0,42	30,89
Planossolo Háplico	68	10	1,61	0	193,15	30,73	307,17
Planossolo Nátrico	9	6,24	1	0,03	41,59	13,52	216,81
Plintossolo Argilúvico	3	1	1	1	1	-	-
Plintossolo Pétrico	1	1	1	1	1	-	-
Vertissolo Ebânico	2	7,66	7,66	1	14,32	9,42	122,96
Vertissolo Háplico	16	4,45	1,77	0,66	33	8,01	180,16

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação.

A classificação dos solos e a litologia revelam padrões distintos de distribuição no bioma Caatinga. Dentre eles, destacam-se os solos poucos e moderadamente intemperizados ou que apresentam baixa atividade pedogenética, nos quais são encontrados as seguintes subordens e províncias geológicas com os maiores valores de C: o Cambissolo Háplico na Borborema, o Chernossolo Rêndzico nas áreas costeiras e margens continentais, e o Gleissolo Sáfico na cobertura cenozoica. Além disso, o Luvisolo Crômico apresenta média mais elevada tanto na cobertura cenozoica quanto na Borborema, alcançando um valor máximo de 147,39 g/kg. Outras ocorrências de médias altas incluem o Neossolo Flúvico na Borborema, o Neossolo Litólico na cobertura cenozoica e na Borborema, com um valor de 140,85 g/kg. Também são observados o Neossolo Quartzarênico, Neossolo Regolítico, o Nitossolo Vermelho e Planossolo Nátrico na Borborema, o Planossolo Háplico no São Francisco. Por fim, temos o Vertissolo Ebânico no São Francisco e o Vertissolo Háplico na Borborema (Tabela 7).

Os solos em estágio mais avançado de intemperismo revelam que o Argissolo Amarelo e o Latossolo Vermelho-Amarelo apresenta seu maior teor médio de C na província da Borborema. Já entre o Argissolo Vermelho-Amarelo e o Latossolo Vermelho, os teores mais elevados são encontrados na Cobertura Cenozoica. No caso dos Argissolos Vermelhos, os teores médios de C aumentam na província do São Francisco. Por fim, o Latossolo Amarelo exibe uma maior média de C na região da Parnaíba (Tabela 7).

Tabela 7 – Estatística descritiva do teor de carbono orgânico (g/kg) da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
PA							
Borborema	3	7,45	6,42	3,52	12,40	4,53	60,79
Cobertura Cenozoica	11	5,72	2,84	0,93	30,04	8,32	145,43
Parnaíba	1	2,71	2,71	2,71	2,71	-	-
São Francisco	7	2,85	1,94	0,21	8,95	2,89	101,42
PV							
Borborema	40	5,28	4,00	1,25	15,55	3,38	64,07
Cobertura Cenozoica	7	4,15	3,69	0,80	6,60	2,27	54,78
Costeira e Margem Continental	5	4,93	2,84	1,81	10,38	3,65	73,94
Parnaíba	4	4,79	4,50	2,17	8,01	2,42	50,47
São Francisco	4	7,54	2,86	2,41	22,02	9,66	128,10
PVA							
Borborema	42	4,60	4,15	0,07	12,87	2,98	64,81
Cobertura Cenozoica	10	5,08	4,47	0,61	11,69	3,63	71,29
Costeira e Margem Continental	1	2,08	2,08	2,08	2,08	-	-
Parnaíba	4	3,28	2,76	0,96	6,65	2,67	81,50
São Francisco	38	4,06	3,00	0,23	16,02	3,41	84,05
CX							
Borborema	3	7,71	5,14	2,38	15,60	6,97	90,52
Cobertura Cenozoica	32	5,61	4,12	0,75	24,00	5,23	93,35
Costeira e Margem Continental	19	6,21	4,50	1,44	36,97	7,97	128,36
São Francisco	34	6,10	3,79	0,80	28,14	6,26	102,78
MD							
Costeira e Margem Continental	4	5,65	5,31	1,92	10,05	3,48	61,62
GZ							
Cobertura Cenozoica	2	11,32	11,32	5,84	16,80	7,75	68,48
Costeira e Margem Continental	1	6,02	6,02	6,02	6,02	-	-

Fonte: Elaborada pelo autor.

Nota: Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sáfico (GZ); Luvisso Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

Tabela 7 – Cont. Estatística descritiva do teor de carbono orgânico (g/kg) da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
LA							
Borborema	12	5,89	5,44	1,16	11,32	3,17	53,77
Cobertura Cenozoica	58	5,66	5,47	0,73	14,64	3,48	61,38
Costeira e Margem Continental	1	1,60	1,60	1,60	1,60	-	-
Parnaíba	21	6,25	4,91	1,59	17,16	4,09	65,35
Recôncavo-Tucano-Jatobá	1	5,18	5,18	5,18	5,18	-	-
São Francisco	34	4,87	3,70	0,18	15,31	4,04	82,90
LV							
Cobertura Cenozoica	3	7,75	5,98	5,39	11,87	3,58	46,25
São Francisco	13	6,92	4,75	1,08	20,00	5,49	79,28
LVA							
Borborema	1	8,91	8,91	8,91	8,91	-	-
Cobertura Cenozoica	15	4,55	3,40	0,70	11,08	3,21	70,53
São Francisco	94	5,46	3,16	0,41	32,02	5,62	102,77
TC							
Borborema	132	8,58	3,42	0,06	147,39	17,89	208,60
Cobertura Cenozoica	1	9,88	9,88	9,88	9,88	-	-
Parnaíba	1	1,87	1,87	1,87	1,87	-	-
São Francisco	12	3,75	3,71	1,84	6,75	1,61	42,98
RY							
Borborema	2	7,08	7,08	2,91	11,26	5,90	83,38
Cobertura Cenozoica	12	4,92	4,57	1,52	11,37	3,00	61,07
Costeira e Margem Continental	1	0,73	0,73	0,73	0,73	-	-
São Francisco	5	4,46	3,90	1,36	9,31	3,12	70,09
RL							
Borborema	107	6,64	2,82	0,19	140,85	14,44	217,60
Cobertura Cenozoica	4	11,56	11,57	7,25	15,86	4,81	41,62
Parnaíba	24	3,81	2,38	0,67	23,60	4,70	123,48
Recôncavo-Tucano-Jatobá	3	7,09	8,73	1,16	11,38	5,31	74,85
São Francisco	28	4,14	3,54	0,23	14,16	3,33	80,50

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sáfico (GZ); Luvissolo Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

Tabela 7 – Cont. Estatística descritiva do teor de carbono orgânico (g/kg) da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
RQ							
Borborema	3	5,91	4,27	2,16	11,30	4,79	80,99
Cobertura Cenozoica	16	5,09	3,99	1,11	23,55	5,36	105,24
Parnaíba	5	2,51	2,71	0,59	4,81	1,85	73,67
Recôncavo-Tucano-Jatobá	8	3,90	3,39	1,40	9,23	2,30	59,08
São Francisco	4	2,39	1,62	1,44	4,87	1,66	69,41
RR							
Borborema	26	4,74	3,27	0,95	18,48	3,86	81,44
São Francisco	4	3,25	3,42	2,06	4,11	0,87	26,65
NV							
Borborema	1	6,72	6,72	6,72	6,72	-	-
São Francisco	3	2,99	2,79	2,06	4,13	1,05	35,15
SX							
Borborema	60	4,84	2,92	0,00	26,56	4,89	101,18
Cobertura Cenozoica	3	2,24	1,75	1,49	3,48	1,08	48,32
Recôncavo-Tucano-Jatobá	6	4,44	3,57	2,25	9,45	2,62	59,16
São Francisco	32	5,83	4,04	1,53	22,76	4,91	84,28
SN							
Borborema	12	5,10	3,56	0,61	16,02	4,92	96,51
Cobertura Cenozoica	2	2,86	2,86	2,49	3,22	0,51	17,93
FT							
Borborema	1	2,74	2,74	2,74	2,74	-	-
Cobertura Cenozoica	1	1,76	1,76	1,76	1,76	-	-
Parnaíba	1	1,43	1,43	1,43	1,43	-	-
FF							
Parnaíba	2	1,65	1,65	1,49	1,81	0,23	13,77
VE							
Cobertura Cenozoica	1	0,59	0,59	0,59	0,59	-	-
São Francisco	1	14,52	14,52	14,52	14,52	-	-

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sálco (GZ); Luvissolo Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

Tabela 7 – Cont. Estatística descritiva do teor de carbono orgânico (g/kg) da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
VX							
Borborema	6	6,14	3,86	1,63	21,20	7,47	121,70
Cobertura Cenozoica	12	3,84	3,89	0,20	8,30	2,66	69,36
Costeira e Margem Continental	3	5,36	6,39	0,58	9,13	4,36	81,34
Recôncavo-Tucano-Jatobá	7	3,13	2,84	0,66	5,28	1,66	52,87
São Francisco	2	6,05	6,05	3,41	8,69	3,74	61,80

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sáfico (GZ); Luvissolo Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

A Tabela 8 apresenta os resultados da análise exploratória do teor de N segundo as subordens e províncias geológicas. Destacam-se os seguintes padrões com as maiores médias para solos considerados pouco ou moderadamente intemperizados: Cambissolo Háplico, Neossolo Regolítico e Planossolo Háplico na província São Francisco; Gleissolo Sáfico, Luvissolo Crômico, Neossolo Quartzarênico, Vertissolo Ebânico e Vertissolo Háplico na Cobertura Cenozoica; Neossolo Flúvico, Nitossolo Vermelho e Planossolo Nátrico na Borborema; Plintossolo Argilúvico e Neossolo Litólico na Parnaíba.

Em relação aos solos em estágio mais avançado de intemperismo, é importante destacar onde ocorre as maiores médias de teor de N. Na província do São Francisco, os Argissolos Amarelos e Latossolos Amarelos. Na Cobertura Cenozoica, os Argissolos Vermelhos, Latossolos Vermelhos e Latossolos Vermelho-Amarelos mostram médias mais altas de N. Por fim, na província da Parnaíba, os Argissolos Vermelho-Amarelos apresentam a média mais elevada de N (vide Tabela 8).

Tabela 8 – Estatística descritiva do teor de nitrogênio (g/kg) da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
PA							
Borborema	3	0,30	0,30	0,29	0,31	0,01	3,71
Cobertura Cenozoica	11	0,42	0,33	0,13	0,84	0,20	48,72
Parnaíba	1	0,48	0,48	0,48	0,48	-	-
São Francisco	7	0,65	0,45	0,20	1,70	0,51	79,08
PV							
Borborema	40	0,52	0,44	0,07	1,13	0,25	48,19
Cobertura Cenozoica	7	0,70	0,49	0,20	1,84	0,58	82,85
Costeira e Margem Continental	5	0,42	0,37	0,19	0,69	0,18	43,79
Parnaíba	4	0,59	0,50	0,46	0,88	0,20	33,44
São Francisco	4	0,57	0,54	0,29	0,91	0,26	45,28
PVA							
Borborema	38	0,62	0,57	0,11	2,03	0,41	66,32
Cobertura Cenozoica	10	0,59	0,55	0,21	1,31	0,35	59,55
Costeira e Margem Continental	1	0,30	0,30	0,30	0,30	-	-
Parnaíba	4	0,70	0,28	0,19	2,06	0,91	128,80
São Francisco	38	0,61	0,52	0,11	2,13	0,41	67,06
CX							
Borborema	3	0,59	0,47	0,31	1,00	0,36	61,01
Cobertura Cenozoica	32	0,65	0,49	0,03	2,28	0,53	81,38
Costeira e Margem Continental	19	0,67	0,51	0,07	2,10	0,51	76,95
São Francisco	34	0,83	0,60	0,30	2,72	0,56	67,82
MD							
Costeira e Margem Continental	4	0,56	0,50	0,30	0,96	0,28	49,83
GZ							
Cobertura Cenozoica	2	0,39	0,39	0,30	0,47	0,12	31,40
Costeira e Margem Continental	1	0,38	0,38	0,38	0,38	-	-

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sáfico (GZ); Luvissolo Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

Tabela 8 – Cont. Estatística descritiva do teor de nitrogênio (g/kg) da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
LA							
Borborema	12	0,64	0,48	0,28	2,06	0,49	76,28
Cobertura Cenozoica	58	0,73	0,51	0,09	3,07	0,63	86,21
Costeira e Margem Continental	1	0,60	0,60	0,60	0,60	-	-
Parnaíba	21	0,60	0,51	0,15	1,42	0,39	64,88
Recôncavo-Tucano-Jatobá	1	0,63	0,63	0,63	0,63	-	-
São Francisco	34	0,76	0,54	0,10	5,68	0,97	127,83
LV							
Cobertura Cenozoica	3	0,89	0,89	0,20	1,57	0,69	77,31
São Francisco	13	0,61	0,46	0,20	1,82	0,43	69,92
LVA							
Borborema	1	0,80	0,80	0,80	0,80	-	-
Cobertura Cenozoica	15	0,83	0,63	0,26	3,21	0,74	89,64
São Francisco	94	0,73	0,64	0,08	4,39	0,53	72,90
TC							
Borborema	79	0,50	0,46	0,07	1,70	0,27	53,24
Cobertura Cenozoica	1	1,59	1,59	1,59	1,59	-	-
Parnaíba	1	0,32	0,32	0,32	0,32	-	-
São Francisco	12	0,52	0,50	0,30	0,99	0,20	38,08
RY							
Borborema	2	0,70	0,70	0,24	1,16	0,65	92,25
Cobertura Cenozoica	12	0,40	0,40	0,10	0,72	0,20	49,06
Costeira e Margem Continental	1	0,31	0,31	0,31	0,31	-	-
São Francisco	5	0,47	0,40	0,07	0,80	0,30	63,81
RL							
Borborema	73	0,52	0,41	0,00	2,39	0,37	71,09
Cobertura Cenozoica	4	0,54	0,57	0,30	0,72	0,17	32,30
Parnaíba	24	0,64	0,53	0,09	1,77	0,48	75,66
Recôncavo-Tucano-Jatobá	3	0,35	0,44	0,08	0,53	0,24	68,70
São Francisco	28	0,57	0,59	0,06	1,07	0,27	47,35

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sáfico (GZ); Luvissolo Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

Tabela 8 – Cont. Estatística descritiva do teor de nitrogênio (g/kg) da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
RQ							
Borborema	3	0,39	0,30	0,27	0,61	0,19	49,00
Cobertura Cenozoica	16	0,62	0,55	0,21	1,19	0,31	50,09
Parnaíba	5	0,44	0,28	0,18	0,92	0,30	69,01
Recôncavo-Tucano-Jatobá	8	0,61	0,52	0,30	0,95	0,26	43,29
São Francisco	4	0,46	0,51	0,22	0,61	0,18	39,73
Borborema	26	0,55	0,50	0,20	1,05	0,24	44,03
São Francisco	4	0,93	0,49	0,32	2,44	1,01	108,41
NV							
Borborema	1	1,00	1,09	1,09	1,09	1,09	-
São Francisco	3	3,00	0,46	0,51	0,29	0,59	0,15
SX							
Borborema	60	59,00	0,62	0,50	0,10	2,20	0,40
Cobertura Cenozoica	3	3,00	0,56	0,31	0,20	1,17	0,53
Recôncavo-Tucano-Jatobá	6	6,00	0,57	0,45	0,25	1,00	0,32
São Francisco	32	32,00	0,66	0,49	0,17	2,65	0,55
SN							
Borborema	12	10,00	0,50	0,46	0,30	0,93	0,20
Cobertura Cenozoica	2	2,00	0,37	0,37	0,20	0,55	0,25
FT							
Borborema	1	1,00	0,10	0,10	0,10	0,10	-
Cobertura Cenozoica	1	1,00	0,35	0,35	0,35	0,35	-
Parnaíba	1	1,00	1,20	1,20	1,20	1,20	-
FF							
Parnaíba	2	2,00	0,61	0,61	0,50	0,73	0,16
VE							
Cobertura Cenozoica	1	1,00	0,90	0,90	0,90	0,90	-
São Francisco	1	1,00	0,42	0,42	0,42	0,42	-

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sáfico (GZ); Luvisso Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

Tabela 8 – Cont. Estatística descritiva do teor de nitrogênio (g/kg) da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
VX							
Borborema	6	6,00	0,72	0,57	0,05	1,42	0,53
Cobertura Cenozoica	12	12,00	0,82	0,77	0,29	1,72	0,52
Costeira e Margem Continental	3	3,00	0,36	0,34	0,29	0,45	0,08
Recôncavo-Tucano-Jatobá	7	7,00	0,40	0,40	0,10	0,64	0,20
São Francisco	2	2,00	0,81	0,81	0,52	1,10	0,41

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sálico (GZ); Luvisso Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

A Tabela 9 exhibe o resumo descritivo da razão C:N segregado pelas subordens de solo e províncias geológicas. Pode-se observar que na província Borborema os valores médios mais elevados estão nos domínios de solos dos Cambissolos Háplicos, Nitossolos Vermelhos e Planossolos Nátricos. Os Gleissolos Sálicos e Neossolos Flúvicos tiveram maiores médias na província Costeira e Margem Continental. Já os Luvisso Crômicos, Vertissolos Ebânicos, Plintossolos Argilúvicos apresentaram maiores médias na Cobertura Cenozoica. Os Neossolos Litólicos, Neossolos Quartzarênicos na província Parnaíba. As médias da razão C:N para os Neossolos Regolíticos se destacaram na província São Francisco. Finalmente, os Planossolos Háplicos, Vertissolos Háplicos resultaram em maior média no Recôncavo-Tucano-Jatobá.

A média de Argissolos Amarelos e Latossolos Vermelhos foi registrada com a maior razão C:N na província Borborema. A média de Argissolos Vermelhos e Latossolos Vermelhos foi maior na Cobertura Cenozoica. Na Parnaíba, a média de C:N foi a maior para o Argissolo Vermelho-Amarelo. Os Latossolos Amarelos e Vermelho-Amarelo apresentaram médias superiores para as seguintes províncias, respectivamente Recôncavo-Tucano-Jatobá e São Francisco (vide Tabela 9).

Tabela 9 – Estatística descritiva da razão C:N da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
PA							
Borborema	3,00	7,69	8,68	5,20	9,19	2,17	28,21
Cobertura Cenozoica	11,00	6,32	6,12	2,41	14,92	3,42	54,20
Parnaíba	1,00	2,84	2,84	2,84	2,84	-	-
São Francisco	7,00	5,71	6,51	3,11	8,62	2,17	37,95
PV							
Borborema	40,00	7,68	7,22	3,96	14,73	2,68	34,96
Cobertura Cenozoica	7,00	10,24	10,00	4,05	21,40	5,53	54,04
Costeira e Margem Continental	5,00	7,16	7,55	5,52	8,50	1,33	18,55
Parnaíba	4,00	7,55	7,86	6,18	8,29	0,94	12,42
São Francisco	4,00	5,10	5,09	2,65	7,57	2,25	44,12
PVA							
Borborema	38,00	8,11	7,62	2,61	16,81	3,76	46,37
Cobertura Cenozoica	10,00	7,97	8,03	3,70	13,02	2,44	30,58
Costeira e Margem Continental	1,00	6,66	6,66	6,66	6,66	-	-
Parnaíba	4,00	10,04	9,94	6,09	14,17	3,76	37,49
São Francisco	38,00	7,09	6,26	1,26	17,51	3,31	46,66
CX							
Borborema	3,00	9,10	10,27	6,13	10,91	2,60	28,53
Cobertura Cenozoica	32,00	7,53	6,77	3,13	14,85	3,00	39,79
Costeira e Margem Continental	19,00	7,93	7,96	2,14	19,29	3,49	43,97
São Francisco	34,00	7,53	6,97	3,50	22,27	3,66	48,61
MD							
Costeira e Margem Continental	4,00	9,09	10,12	4,11	12,02	3,65	40,10
GZ							
Cobertura Cenozoica	2,00	7,04	7,04	5,07	9,00	2,78	39,55
Costeira e Margem Continental	1,00	11,83	11,83	11,83	11,83	-	-

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sáfico (GZ); Luvissolo Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

Tabela 9 – Cont. Estatística descritiva da razão C:N da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
LA							
Borborema	12,00	6,66	7,12	2,90	10,05	2,01	30,17
Cobertura Cenozoica	58,00	9,04	8,30	2,00	25,47	3,99	44,14
Costeira e Margem Continental	1,00	9,10	9,10	9,10	9,10	-	-
Parnaíba	21,00	7,61	7,68	4,01	13,30	2,29	30,15
Recôncavo-Tucano-Jatobá	1,00	16,97	16,97	16,97	16,97	-	-
São Francisco	34,00	8,30	7,83	2,66	17,95	4,02	48,40
LV							
Cobertura Cenozoica	3,00	9,67	8,94	7,06	13,00	3,04	31,41
São Francisco	13,00	7,44	6,78	2,52	11,49	2,73	36,73
LVA							
Borborema	1,00	3,64	3,64	3,64	3,64	-	-
Cobertura Cenozoica	15,00	6,80	6,19	3,12	14,02	2,74	40,37
São Francisco	94,00	6,21	5,56	1,72	22,95	2,84	45,79
TC							
Borborema	79,00	6,37	6,03	1,17	11,82	2,48	39,01
Cobertura Cenozoica	1,00	8,88	8,88	8,88	8,88	-	-
Parnaíba	1,00	7,23	7,23	7,23	7,23	-	-
São Francisco	12,00	7,39	6,89	4,64	13,56	2,57	34,77
RY							
Borborema	2,00	6,59	6,59	6,39	6,79	0,28	4,30
Cobertura Cenozoica	12,00	6,97	7,41	3,44	10,81	2,43	34,91
Costeira e Margem Continental	1,00	8,19	8,19	8,19	8,19	-	-
São Francisco	5,00	7,86	7,84	6,79	8,88	0,78	9,92
RL							
Borborema	72,00	7,69	7,41	2,39	14,80	2,78	36,09
Cobertura Cenozoica	4,00	5,32	5,78	1,76	7,96	2,69	50,56
Parnaíba	24,00	9,10	8,90	2,08	16,92	3,97	43,63
Recôncavo-Tucano-Jatobá	3,00	5,63	4,54	2,20	10,17	4,10	72,71
São Francisco	28,00	8,16	7,24	3,44	16,69	3,48	42,71

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sáfico (GZ); Luvissolo Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

Tabela 9 – Cont. Estatística descritiva da razão C:N da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
RQ							
Borborema	3,00	5,33	4,26	2,87	8,88	3,15	58,98
Cobertura Cenozoica	16,00	7,67	7,03	3,75	10,67	2,27	29,61
Parnaíba	5,00	8,09	6,84	5,75	11,80	2,73	33,70
Recôncavo-Tucano-Jatobá	8,00	7,75	8,07	5,00	11,19	2,42	31,22
São Francisco	4,00	9,06	8,41	6,47	12,96	2,86	31,52
RR							
Borborema	26,00	7,17	6,68	2,46	12,33	2,75	38,39
São Francisco	4,00	8,47	8,43	7,50	9,52	0,89	10,47
NV							
Borborema	1,00	16,15	16,15	16,15	16,15	-	-
São Francisco	3,00	4,17	4,15	3,12	5,23	1,06	0,15
SX							
Borborema	59,00	7,85	7,64	3,06	15,45	2,80	0,40
Cobertura Cenozoica	3,00	8,06	5,64	5,21	13,33	4,57	0,53
Recôncavo-Tucano-Jatobá	6,00	9,17	8,89	4,50	14,39	4,32	0,32
São Francisco	32,00	7,36	7,29	2,57	23,94	3,98	0,55
SN							
Borborema	10,00	7,75	8,29	2,07	16,72	4,02	0,20
Cobertura Cenozoica	2,00	7,73	7,73	7,56	7,90	0,25	0,25
FT							
Borborema	1,00	6,02	6,02	6,02	6,02	-	-
Cobertura Cenozoica	1,00	10,33	10,33	10,33	10,33	-	-
Parnaíba	1,00	9,29	9,29	9,29	9,29	-	-
FF							
Parnaíba	2,00	7,62	7,62	4,76	10,49	4,05	0,16
VE							
Cobertura Cenozoica	1,00	8,65	8,65	8,65	8,65	-	-
São Francisco	1,00	5,75	5,75	5,75	5,75	-	-

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sáfico (GZ); Luvissolo Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

Tabela 9 – Cont. Estatística descritiva da razão C:N da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
VX							
Borborema	6,00	7,13	6,83	4,53	11,18	2,51	0,53
Cobertura Cenozoica	12,00	6,49	6,61	2,28	9,67	2,35	0,52
Costeira e Margem Continental	3,00	6,14	6,66	2,67	9,08	3,24	0,08
Recôncavo-Tucano-Jatobá	7,00	8,80	8,48	6,31	12,40	1,96	0,20
São Francisco	2,00	7,92	7,92	7,71	8,14	0,30	0,41

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sáfico (GZ); Luvisso Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

Com base na Tabela 10, é possível obter as médias mais significativas de P para os Cambissolos Háplicos e Vertissolos Háplicos na região Costeira e Margem Continental. Já para os Gleissolos Sáficos, Neossolos Flúvicos e Neossolos Quartzarênicos, houve maiores valores médios de P na Cobertura Cenozoica. Os Luvisso Crômicos, Planossolos Háplicos e Planossolos Nátricos apresentaram as maiores médias de P na Borborema. Os Neossolos Litólicos tiveram a média de P mais elevada no Recôncavo-Tucano-Jatobá. Por fim, encontramos maiores médias para os Vertissolos Ebânicos, Neossolos Regolíticos e Nitossolos Vermelhos na região do São Francisco.

Em relação aos solos muito intemperizados, houve maiores valores médios para os Argissolos Amarelos e Latossolos Amarelos na Parnaíba. Os Argissolos Vermelhos e Latossolos Vermelhos-Amarelos ocorreu maiores médias na província geológica São Francisco. A média de P nos Argissolos Vermelhos-Amarelo foi maior na região Costeira e Margem Continental. Por último, os Latossolos Vermelhos na Cobertura Cenozoica (Tabela 10).

Tabela 10 – Estatística descritiva do fósforo disponível (mg/kg) da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
PA							
Borborema	1,00	2,93	2,93	2,93	2,93	-	-
Cobertura Cenozoica	9,00	2,91	1,00	0,65	15,62	4,85	166,66
Parnaíba	1,00	5,62	5,62	5,62	5,62	-	-
São Francisco	2,00	3,41	3,41	2,50	4,32	1,28	37,65
PV							
Borborema	33,00	3,19	1,97	0,68	12,83	3,43	107,49
Cobertura Cenozoica	7,00	1,58	1,06	0,99	3,45	0,90	57,07
Costeira e Margem Continental	5,00	8,05	1,90	1,00	33,47	14,23	176,77
Parnaíba	3,00	10,49	2,12	1,14	28,23	15,36	146,39
São Francisco	2,00	34,62	34,62	1,23	68,01	47,22	136,40
PVA							
Borborema	30,00	12,31	1,81	0,00	239,30	43,62	354,31
Cobertura Cenozoica	7,00	0,99	1,00	0,83	1,10	0,08	8,29
Costeira e Margem Continental	1,00	160,60	160,60	160,60	160,60	-	-
Parnaíba	3,00	52,00	2,00	1,00	153,01	87,47	168,21
São Francisco	23,00	25,47	1,12	0,08	181,22	50,25	197,27
CX							
Borborema	3,00	10,47	2,98	1,00	27,41	14,71	140,58
Cobertura Cenozoica	22,00	20,59	1,47	0,73	153,06	43,83	212,84
Costeira e Margem Continental	15,00	21,33	1,03	0,30	296,89	76,24	357,49
São Francisco	28,00	5,07	2,10	0,51	31,84	8,10	159,81
MD							
Costeira e Margem Continental	3,00	4,68	4,12	0,92	9,00	4,07	86,93
GZ							
Cobertura Cenozoica	1,00	4,30	4,30	4,30	4,30	-	-
Costeira e Margem Continental	1,00	1,00	1,00	1,00	1,00	-	-

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sáfico (GZ); Luvisso Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

Tabela 10 – Cont. Estatística descritiva do fósforo disponível (mg/kg) da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
LA							
Borborema	9,00	10,41	3,15	1,00	68,27	21,90	210,35
Cobertura Cenozoica	44,00	7,71	1,77	0,02	57,06	14,70	190,72
Costeira e Margem Continental	1,00	3,48	3,48	3,48	3,48	-	-
Parnaíba	19,00	13,04	2,88	0,68	131,24	31,23	239,48
São Francisco	28,00	11,52	1,98	0,15	113,12	24,48	212,42
LV							
Cobertura Cenozoica	2,00	11,87	11,87	6,71	17,04	7,31	61,55
São Francisco	7,00	1,16	0,99	0,53	2,69	0,72	62,44
LVA							
Borborema	1,00	6,52	6,52	6,52	6,52	-	-
Cobertura Cenozoica	11,00	10,02	3,01	0,87	81,96	23,96	239,10
São Francisco	82,00	23,06	1,99	0,00	563,32	73,99	320,84
TC							
Borborema	52,00	5,00	2,20	0,00	41,68	7,39	147,91
Parnaíba	1,00	2,01	2,01	2,01	2,01	-	-
São Francisco	5,00	1,26	1,06	1,00	2,00	0,42	33,67
RY							
Borborema	2,00	0,49	0,49	0,28	0,69	0,29	60,21
Cobertura Cenozoica	8,00	51,01	4,01	1,00	275,44	97,78	191,70
Costeira e Margem Continental	1,00	3,00	3,00	3,00	3,00	-	-
São Francisco	2,00	0,94	0,94	0,88	1,00	0,09	9,41
RL							
Borborema	50,00	23,67	2,14	0,00	509,09	76,28	322,32
Cobertura Cenozoica	1,00	69,03	69,03	69,03	69,03	-	-
Parnaíba	19,00	7,05	1,01	0,83	66,39	16,13	228,93
Recôncavo-Tucano-Jatobá	2,00	160,07	160,07	0,99	319,14	224,97	140,55
São Francisco	19,00	9,17	1,60	0,56	136,17	30,82	336,21

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sáfico (GZ); Luvisso Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

Tabela 10 – Cont. Estatística descritiva do fósforo disponível (mg/kg) da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
RQ							
Borborema	2,00	1,07	1,07	1,00	1,15	0,10	9,73
Cobertura Cenozoica	12,00	17,70	2,02	0,58	166,21	47,16	266,47
Parnaíba	5,00	4,91	2,30	0,84	13,36	5,11	104,12
Recôncavo-Tucano-Jatobá	7,00	4,79	3,00	1,00	14,67	5,09	106,33
São Francisco	3,00	0,65	0,89	0,00	1,07	0,57	87,71
RR							
Borborema	21,00	10,61	3,00	0,97	64,31	18,29	172,40
São Francisco	2,00	222,51	222,51	1,00	444,01	313,26	140,79
NV							
Borborema	1,00	1,33	1,33	1,33	1,33	-	-
São Francisco	2,00	1,38	1,38	0,96	1,81	0,60	43,05
SX							
Borborema	42,00	8,51	1,15	0,00	148,87	25,91	304,48
Cobertura Cenozoica	3,00	0,88	1,00	0,58	1,07	0,26	29,54
Recôncavo-Tucano-Jatobá	3,00	2,69	1,68	1,03	5,37	2,34	86,92
São Francisco	20,00	15,61	2,95	0,65	193,15	42,68	273,42
SN							
Borborema	7,00	7,62	0,91	0,03	41,59	15,29	200,68
Cobertura Cenozoica	2,00	1,40	1,40	1,00	1,80	0,57	40,47
FT							
Borborema	1,00	1,00	1,00	1,00	1,00	-	-
Cobertura Cenozoica	1,00	1,00	1,00	1,00	1,00	-	-
Parnaíba	1,00	1,00	1,00	1,00	1,00	-	-
FF							
Parnaíba	1,00	1,00	1,00	1,00	1,00	-	-
VE							
Cobertura Cenozoica	1,00	1,00	1,00	1,00	1,00	-	-
São Francisco	1,00	14,32	14,32	14,32	14,32	-	-

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sáfico (GZ); Luvissolo Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

Tabela 10 – Cont. Estatística descritiva do fósforo disponível (mg/kg) da camada superficial, segundo classificação de solo e província geológica.

Subordem/Província	N	Média	Mediana	Mínimo	Máximo	DP	CV
VX							
Borborema	2,00	1,50	1,50	1,00	2,00	0,71	47,20
Cobertura Cenozoica	6,00	2,68	1,53	0,66	8,65	3,00	111,92
Costeira e Margem Continental	2,00	20,84	20,84	8,67	33,00	17,20	82,56
Recôncavo-Tucano-Jatobá	5,00	1,49	1,00	0,82	2,63	0,80	53,96
São Francisco	1,00	2,94	2,94	2,94	2,94	-	-

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão; CV (%): coeficiente de variação. Argissolo Amarelo (PA); Argissolo Vermelho (PV); Argissolo Vermelho-Amarelo (PVA); Latossolo Amarelo (LA); Latossolo Vermelho (LV); Latossolo Vermelho-Amarelo (LVA); Cambissolo Háplico (CX); Chernossolo Rêndzico (MD); Gleissolo Sáfico (GZ); Luvissolo Crômico (TC); Neossolo Flúvico (RY); Neossolo Litólico (RL); Neossolo Quartzarênico (RQ); Neossolo Regolítico (RR); Nitossolo Vermelho (NV); Planossolo Háplico (SX); Planossolo Nátrico (SN); Vertissolo Ebânico (VE); Vertissolo Háplico (VX); Plintossolo Argilúvico (FT); Plintossolo Pétrico (FF).

4.3.3 Relação entre Propriedades Biogeoquímicas do Solo e Estrutura da Vegetação

A Tabela 11 apresenta as estatísticas descritivas do teor de carbono (C) nos diferentes estratos da vegetação. Em relação ao estrato Arbustivo Alto, que corresponde a uma altura entre 1,5 e 3 metros, a média de teor de carbono foi de 1,93 g/kg, com uma mediana de 1,84 g/kg. O valor mínimo observado foi de 0,07 g/kg, enquanto o máximo foi de 3,87 g/kg. A variação dos dados, representada pelo desvio padrão, foi de 1,90 g/kg.

No estrato Arbóreo Baixo, com altura de 3 a 5 metros, a média de teor de carbono foi de 5,86 g/kg, e a mediana foi de 3,71 g/kg. O valor mínimo encontrado foi de 0,06 g/kg, enquanto o máximo foi de 147,39 g/kg. A dispersão dos dados foi representada pelo desvio padrão de 9,45 g/kg.

No estrato Arbóreo Intermediário, com altura de 5 a 10 metros, a média de teor de carbono foi de 5,58 g/kg, com uma mediana de 3,75 g/kg. O valor mínimo registrado foi de 0 g/kg, e o valor máximo foi de 140,85 g/kg. O desvio padrão foi de 8,24 g/kg.

No estrato Arbóreo Alto, com altura de 10 a 20 metros, a média de teor de carbono foi de 4,78 g/kg, com uma mediana de 3,02 g/kg. O valor mínimo observado foi de 0,70 g/kg, enquanto o máximo foi de 16,84 g/kg. A variação dos dados, representada pelo desvio padrão, foi de 4,21 g/kg.

Esses resultados mostram a variabilidade do teor de carbono nos diferentes estratos da vegetação, indicando que o estrato Arbóreo Intermediário apresentou a maior mediana de teor de carbono, seguido pelo estrato Arbóreo Baixo e, em seguida, pelo estrato Arbóreo Alto.

Tabela 11 – Estatística descritiva do carbono orgânico (g/kg) da camada superficial de solos por estratos da vegetação.

Estratos	N	Média	Mediana	Mínimo	Máximo	DP
1,5 - 3 m (Arbustivo Alto)	3	1,93	1,84	0,07	3,87	1,90
3 - 5 m (Arbóreo Baixo)	518	5,86	3,71	0,06	147,39	9,45
5 - 10 m (Arbóreo Intermediário)	532	5,58	3,75	0	140,85	8,24
10 -20 m (Arbóreo Alto)	26	4,78	3,02	0,70	16,84	4,21

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão.

O resumo dos dados apresentados na Tabela 12 mostra as estatísticas descritivas do teor de N em diferentes estratos da vegetação. No estrato Arbustivo Alto, a média de teor de N foi de 0,70 g/kg, com uma mediana de 0,41 g/kg. O valor mínimo observado foi de 0,41 g/kg,

enquanto o máximo foi de 1,29 g/kg. O desvio padrão foi de 0,50 g/kg.

No estrato Arbóreo Baixo, a média de teor de N foi de 0,60 g/kg e a mediana foi de 0,48 g/kg. O valor mínimo encontrado foi de 0,03 g/kg, enquanto o máximo foi de 5,68 g/kg. A dispersão dos dados foi representada pelo desvio padrão de 0,48 g/kg.

No estrato Arbóreo Intermediário, a média de teor de N foi de 0,64 g/kg, com uma mediana de 0,51 g/kg. O valor mínimo registrado foi de 0 g/kg, e o valor máximo foi de 4,39 g/kg. O desvio padrão foi de 0,46 g/kg.

No estrato Arbóreo Alto, a média de teor de nitrogênio foi de 0,58 g/kg, com uma mediana de 0,51 g/kg. O valor mínimo observado foi de 0,20 g/kg, enquanto o máximo foi de 1,28 g/kg. A variação dos dados, representada pelo desvio padrão, foi de 0,29 g/kg.

Esses dados revelam a variabilidade dos teores de N nos diferentes estratos da vegetação. Nota-se que o estrato Arbustivo Alto apresentou a maior média de teor de N, seguido pelo estrato Arbóreo Intermediário e, em seguida, pelo estrato Arbóreo Baixo.

Tabela 12 – Estatística descritiva do teor de nitrogênio (g/kg) da camada superficial de solos por estratos da vegetação.

Estratos	N	Média	Mediana	Mínimo	Máximo	DP
1,5 - 3 m (Arbustivo Alto)	3	0,70	0,41	0,41	1,29	0,50
3 - 5 m (Arbóreo Baixo)	475	0,60	0,48	0,03	5,68	0,48
5 - 10 m (Arbóreo Intermediário)	484	0,64	0,51	0	4,39	0,46
10 -20 m (Arbóreo Alto)	23	0,58	0,51	0,20	1,28	0,29

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão.

A Tabela 13 refere-se à razão C:N em diferentes estratos da vegetação. No estrato Arbustivo Alto, a média da razão C:N foi de 0,70, com uma mediana de 0,41. O valor mínimo observado foi de 0,41, enquanto o máximo foi de 1,29. A variação dos dados, representada pelo desvio padrão, foi de 0,50.

No estrato Arbóreo Baixo, com altura de 3 a 5 metros, a média da razão C:N foi de 0,60, e a mediana foi de 0,48. O valor mínimo encontrado foi de 0,03, enquanto o máximo foi de 5,68. O desvio padrão de 0,48.

No estrato Arbóreo Intermediário, com altura de 5 a 10 metros, a média da razão C:N foi de 0,64, com uma mediana de 0,51. O valor mínimo registrado foi de 0, e o valor máximo foi de 4,39. O desvio padrão foi de 0,46.

No estrato Arbóreo Alto, com altura de 10 a 20 metros, a média da razão C:N foi

de 0,58, com uma mediana de 0,51. O valor mínimo observado foi de 0,20, enquanto o máximo foi de 1,28. A variação dos dados, representada pelo desvio padrão, foi de 0,29.

Esses dados revelam a variabilidade das razões C:N nos diferentes estratos da vegetação. Nota-se que o estrato Arbustivo Alto apresentou a maior média de razão C:N, seguido pelo estrato Arbóreo Intermediário e, em seguida, pelo estrato Arbóreo Baixo. A análise dessas estatísticas descritivas é importante para compreender a distribuição e a variação das razões C:N nos diferentes estratos da vegetação, o que pode fornecer informações sobre a qualidade e a decomposição da matéria orgânica presente nesses estratos.

Tabela 13 – Estatística descritiva da razão C:N da camada superficial de solos por estratos da vegetação.

Estratos	N	Média	Mediana	Mínimo	Máximo	DP
1,5 - 3 m (Arbustivo Alto)	3	0,70	0,41	0,41	1,29	0,50
3 - 5 m (Arbóreo Baixo)	475	0,60	0,48	0,03	5,68	0,48
5 - 10 m (Arbóreo Intermediário)	484	0,64	0,51	0	4,39	0,46
10 -20 m (Arbóreo Alto)	23	0,58	0,51	0,20	1,28	0,29

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão.

O resumo dos dados apresentados na tabela 14 refere-se ao teor de P em diferentes estratos da vegetação, expresso em mg/kg. No estrato Arbustivo Alto, a média de teor de P foi de 0,70 mg/kg, com uma mediana de 0,41 mg/kg. O valor mínimo observado foi de 0,41 mg/kg, enquanto o máximo foi de 1,29 mg/kg. A variação dos dados, representada pelo desvio padrão, foi de 0,50 mg/kg.

No estrato Arbóreo Baixo, a média de teor de fósforo foi de 0,60 mg/kg, e a mediana foi de 0,48 mg/kg. O valor mínimo encontrado foi de 0,03 mg/kg, enquanto o máximo foi de 5,68 mg/kg. O desvio padrão de 0,48 mg/kg.

No estrato Arbóreo Intermediário, a média de teor de fósforo foi de 0,64 mg/kg, com uma mediana de 0,51 mg/kg. O valor mínimo registrado foi de 0 mg/kg, e o valor máximo foi de 4,39 mg/kg. O desvio padrão foi de 0,46 mg/kg.

No estrato Arbóreo Alto, a média de teor de fósforo foi de 0,58 mg/kg, com uma mediana de 0,51 mg/kg. O valor mínimo observado foi de 0,20 mg/kg, enquanto o máximo foi de 1,28 mg/kg. A variação dos dados, representada pelo desvio padrão, foi de 0,29 mg/kg.

Esses dados revelam a variabilidade dos teores de fósforo nos diferentes estratos da vegetação. Observa-se que o estrato Arbustivo Alto apresentou a maior média de teor de fósforo, seguido pelo estrato Arbóreo Intermediário e, em seguida, pelo estrato Arbóreo Baixo. A

análise dessas estatísticas descritivas é importante para compreender a distribuição e a variação dos teores de fósforo nos diferentes estratos da vegetação, o que pode fornecer informações sobre a disponibilidade desse nutriente para as plantas e sua influência na produtividade e no desenvolvimento da vegetação.

Tabela 14 – Estatística descritiva do fósforo disponível (mg/kg) da camada superficial de solos por estratos da vegetação.

Estratos	N	Média	Mediana	Mínimo	Máximo	DP
1,5 - 3 m (Arbustivo Alto)	3	0,70	0,41	0,41	1,29	0,50
3 - 5 m (Arbóreo Baixo)	475	0,60	0,48	0,03	5,68	0,48
5 - 10 m (Arbóreo Intermediário)	484	0,64	0,51	0	4,39	0,46
10 -20 m (Arbóreo Alto)	23	0,58	0,51	0,20	1,28	0,29

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão.

A Tabela 15 apresenta o teor de carbono (C) em diferentes coberturas da vegetação. Na cobertura de 0 a 10% (classificada como Muito Esparsada/Isolada), a média de teor de C foi de 5,97 g/kg, com uma mediana de 3,51 g/kg. O valor mínimo observado foi de 0,06 g/kg, enquanto o máximo foi de 147,39 g/kg. O desvio padrão foi de 10,60 g/kg.

Na cobertura de 10 a 30% (classificada como Esparsado/Moderadamente denso), a média de teor de carbono foi de 5,30 g/kg, e a mediana foi de 3,92 g/kg. O valor mínimo encontrado foi de 0 g/kg, enquanto o máximo foi de 76,87 g/kg. Com desvio padrão de 5,72 g/kg.

Na cobertura > 30% (classificada como Denso), a média de teor de carbono foi de 5,41 g/kg, com uma mediana de 3,79 g/kg. O valor mínimo registrado foi de 0,18 g/kg, e o valor máximo foi de 20,02 g/kg. O desvio padrão foi de 4,71 g/kg.

Esses dados revelam a variabilidade dos teores de carbono nas diferentes coberturas da vegetação. Observa-se que a cobertura de 0 a 10% apresentou a maior média de teor de carbono, seguida pela cobertura > 30% e, em seguida, pela cobertura de 10 a 30%.

O resumo dos dados apresentados na Tabela 16 refere-se ao teor de N em diferentes coberturas da vegetação. Na cobertura de 0 a 10%, a média de teor de N foi de 0,62 g/kg, com uma mediana de 0,51 g/kg. O valor mínimo observado foi de 0,05 g/kg, enquanto o máximo foi de 5,68 g/kg. A variação dos dados, representada pelo desvio padrão, foi de 0,48 g/kg.

Na cobertura de 10 a 30%, a média de teor de nitrogênio foi de 0,61 g/kg, e a mediana foi de 0,49 g/kg. O valor mínimo encontrado foi de 0 g/kg, enquanto o máximo foi de

Tabela 15 – Estatística descritiva do carbono orgânico (g/kg) da camada superficial de solos por cobertura da vegetação.

Cobertura	N	Média	Mediana	Mínimo	Máximo	DP
0 - 10% (Muito Esparsada/Isolada)	605	5,97	3,51	0,06	147,39	10,60
10 - 30% (Esparsado/Moderadamente denso)	419	5,30	3,92	0	76,87	5,72
>30 (Denso)	55	5,41	3,79	0,18	20,02	4,71

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão.

3,21 g/kg. A dispersão dos dados foi representada pelo desvio padrão de 0,47 g/kg.

Na cobertura > 30% (classificada como Denso), foram analisadas 53 amostras. A média de teor de nitrogênio foi de 0,66 g/kg, com uma mediana de 0,60 g/kg. O valor mínimo registrado foi de 0,10 g/kg, e o valor máximo foi de 2,13 g/kg. O desvio padrão foi de 0,38 g/kg.

Esses dados revelam a variabilidade dos teores de nitrogênio nas diferentes coberturas da vegetação. Observa-se que a cobertura > 30% apresentou a maior média de teor de nitrogênio, seguida pela cobertura de 0 a 10% e, em seguida, pela cobertura de 10 a 30%.

Tabela 16 – Estatística descritiva do teor de nitrogênio (g/kg) da camada superficial de solos por cobertura da vegetação.

Cobertura	N	Média	Mediana	Mínimo	Máximo	DP
0 - 10% (Muito Esparsada/Isolada)	547	0,62	0,51	0,05	5,68	0,48
10 - 30% (Esparsado/Moderadamente denso)	385	0,61	0,49	0,00	3,21	0,47
>30 (Denso)	53	0,66	0,60	0,10	2,13	0,38

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão.

A Tabela 17 refere-se ao teor de razão carbono-nitrogênio (C:N) em diferentes coberturas da vegetação. Na cobertura de 0 a 10%, a média da razão C:N foi de 7,39, com uma mediana de 6,83. O valor mínimo observado foi de 1,72, enquanto o máximo foi de 25,47. A variação dos dados, representada pelo desvio padrão, foi de 3,24.

Na cobertura de 10 a 30%, a média da razão C:N foi de 7,69 e a mediana foi de 7,37. O valor mínimo encontrado foi de 1,17, enquanto o máximo foi de 21,40. A dispersão dos dados foi representada pelo desvio padrão de 3,14.

Na cobertura > 30% (classificada como Denso), a média da razão C:N foi de 7,47, com uma mediana de 7,23. O valor mínimo registrado foi de 2,08, e o valor máximo foi de

15,57. O desvio padrão foi de 3,08.

Esses dados revelam a variabilidade das razões C:N nas diferentes coberturas da vegetação. Observa-se que a cobertura de 10 a 30% apresentou a maior média de razão C:N, seguida pela cobertura > 30% e, em seguida, pela cobertura de 0 a 10%.

Tabela 17 – Estatística descritiva da razão C:N da camada superficial de solos por cobertura da vegetação.

Cobertura	N	Média	Mediana	Mínimo	Máximo	DP
0 - 10% (Muito Esparsada/Isolada)	546	7,39	6,83	1,72	25,47	3,24
10 - 30% (Esparsado/Moderadamente denso)	385	7,69	7,37	1,17	21,40	3,14
>30 (Denso)	53	7,47	7,23	2,08	15,57	3,08

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão.

O teor de P (Tabela 18) na cobertura de 0 a 10% teve uma média de teor de fósforo foi de 14,71 mg/kg, com uma mediana de 1,85 mg/kg. O valor mínimo observado foi de 0 mg/kg, enquanto o máximo foi de 563,32 mg/kg. A variação dos dados, representada pelo desvio padrão, foi de 49,64 mg/kg.

Na cobertura de 10 a 30%, a média de teor de fósforo foi de 13,13 mg/kg, e a mediana foi de 1,80 mg/kg. O valor mínimo encontrado foi de 0 mg/kg, enquanto o máximo foi de 509,09 mg/kg. A dispersão dos dados foi representada pelo desvio padrão de 45,49 mg/kg.

Na cobertura > 30%, foram analisadas 44 amostras. A média de teor de fósforo foi de 9,39 mg/kg, com uma mediana de 2 mg/kg. O valor mínimo registrado foi de 0,51 mg/kg, e o valor máximo foi de 181,22 mg/kg. O desvio padrão foi de 27,67 mg/kg.

Esses dados revelam a variabilidade dos teores de fósforo nas diferentes coberturas da vegetação. Observa-se que a cobertura de 0 a 10% apresentou a maior média de teor de fósforo, seguida pela cobertura de 10 a 30% e, em seguida, pela cobertura > 30%.

Tabela 18 – Estatística descritiva do fósforo disponível (mg/kg) da camada superficial de solos por cobertura da vegetação.

Cobertura	N	Média	Mediana	Mínimo	Máximo	DP
0 - 10% (Muito Esparsada/Isolada)	402	14,71	1,85	0	563,32	49,64
10 - 30% (Esparsado/Moderadamente denso)	274	13,13	1,80	0	509,09	45,49
>30 (Denso)	44	9,39	2	0,51	181,22	27,67

Fonte: Elaborada pelo autor.

Nota: N: amostras; DP: desvio-padrão.

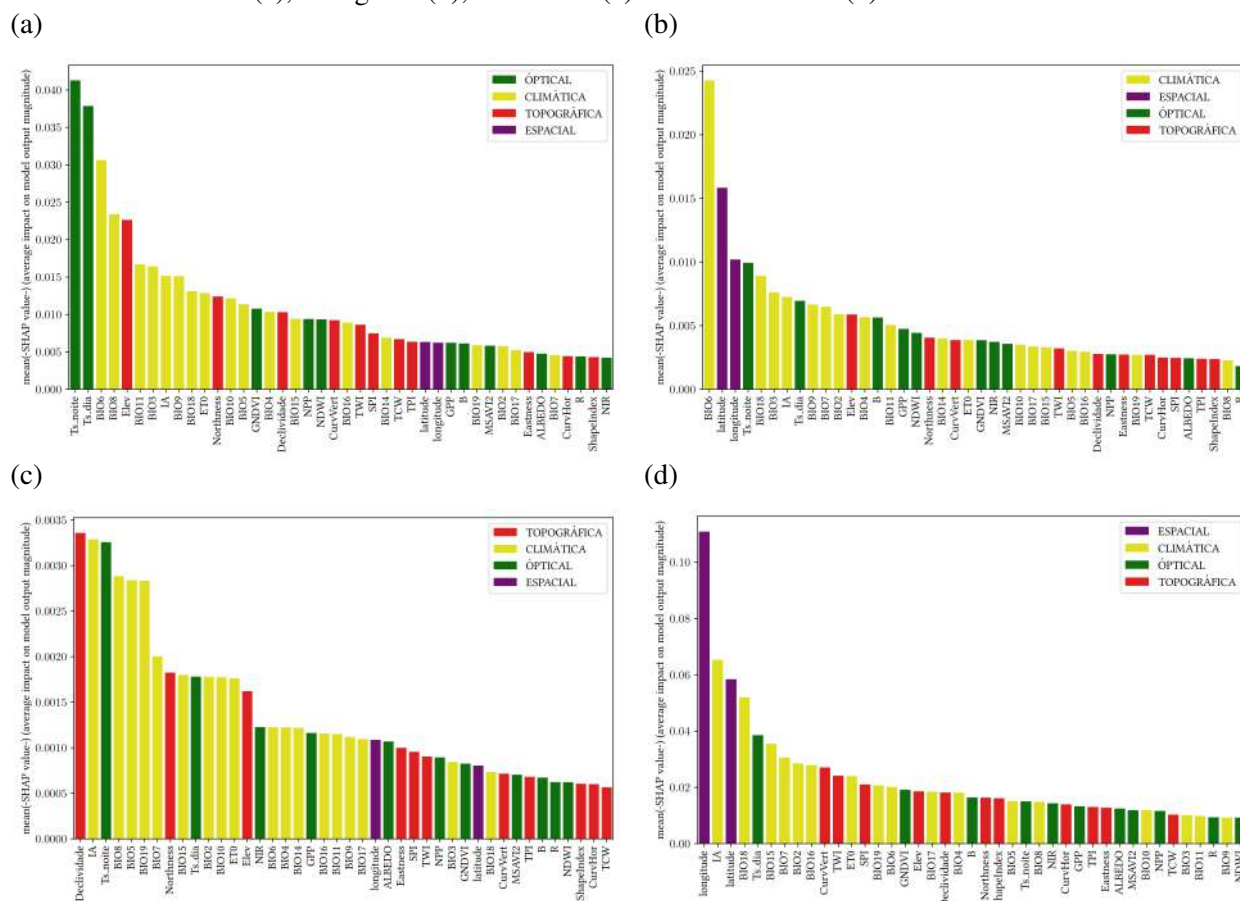
4.3.4 Variáveis importantes calculadas com valores SHAP

A magnitude das contribuições das variáveis preditoras na previsão dos nutrientes do solo, utilizando o modelo Random Forest e com base no conjunto de dados de calibração. A análise dos valores médios absolutos de Shapley revela as variáveis mais importantes para cada nutriente. Na Figura 20a, observa-se que as temperaturas diárias (Ts_dia) e noturnas (Ts_noite) da superfície do solo são as variáveis mais relevantes na previsão do teor de C. Além disso, a temperatura mínima durante o mês mais frio (BIO6), latitude e longitude também apresentam uma contribuição significativa na previsão do teor de N (Figura 20b).

Para a razão C:N, a Figura 20c revela que a declividade, o índice de aridez (IA) e a temperatura noturna do solo (Ts_noite) são as variáveis mais importantes, exibindo uma ampla gama de valores de Shapley. O que corrobora outros estudos que usaram atributos de clima e terreno como variáveis preditoras (REDDY *et al.*, 2021).

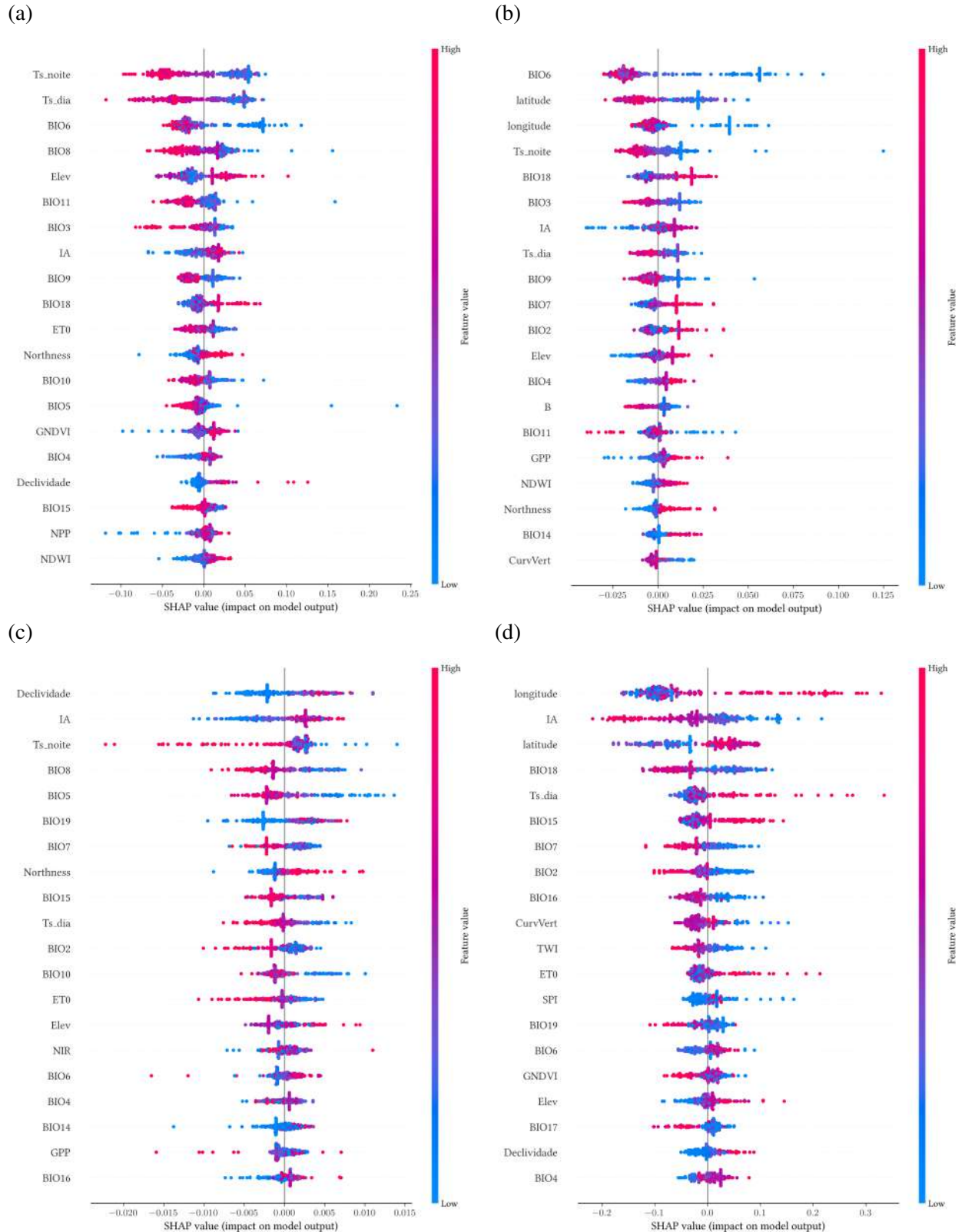
Por fim, na Figura 20d, nota-se que a longitude é a variável mais relevante na previsão do P, evidenciando sua maior influência. Esses resultados destacam as principais variáveis preditoras que contribuem para as estimativas dos nutrientes do solo na região da Caatinga, proporcionando informações valiosas para compreender os fatores que influenciam essas características do solo.

Considerando as principais covariáveis, a Figura 21a ilustra como o modelo capturou a relação inversa entre as variáveis Ts_dia, Ts_noite, BIO6 e a temperatura média do quarto trimestre mais chuvoso (BIO8) com o teor de carbono (C), e uma relação direta com a elevação (Elev). A Figura 21b revela uma relação inversa entre BIO6, latitude, longitude e Ts_noite, enquanto mostra uma correlação positiva entre a precipitação do trimestre mais quente (BIO18) e o teor de nitrogênio (N). Por sua vez, a razão C:N apresentou uma relação positiva com a declividade, índice de aridez (IA) e precipitação do trimestre mais frio (BIO19), enquanto teve uma relação negativa com Ts_noite, BIO8 e temperatura máxima do mês mais quente (BIO5) (Figura 21c). Em relação ao fósforo (P), houve uma relação positiva com a longitude, latitude, Ts_dia e a sazonalidade da precipitação (coeficiente de variação, BIO15), e uma relação negativa com o IA e BIO18 (Figura 21d).



Fonte: Elaborada pelo autor.

Figura 21 – Os valores de Shapley para cada ponto individual no conjunto de dados de calibração representam a contribuição da covariável para a previsão dos teores de carbono (a), nitrogênio (b), razão C:N (c) e fósforo no solo (d) em um determinado local. A escala de cores utilizada reflete o valor da covariável normalizada no intervalo de 0 a 1.



Fonte: Elaborada pelo autor.

4.3.5 Avaliação e Precisão do Modelo

Os resultados apresentados na (Figura 22) referem-se ao desempenho dos modelos de regressão usando a técnica de Random Forest (RF) utilizando a validação cruzada de 10 conjuntos (10cv) para a modelagem de C, N, C:N e P do solo na região da Caatinga.

Para o C, o modelo RF apresentou um coeficiente de determinação (R^2) de 0,24, o que significa que aproximadamente 24% da variação nos níveis de carbono do solo foi explicada pelo modelo. O coeficiente de correlação de concordância de Lin (LCCC) foi de 0,40, indicando uma concordância moderada entre os valores observados e os previstos pelo modelo RF. O erro médio quadrático (RMSE) foi de 13,90 g/kg, enquanto o erro absoluto médio (MAE) foi de 4,81 g/kg. Quanto ao índice de qualidade preditiva relativa (RPIQ) foi de 0,56, sugerindo um ajuste moderado do modelo RF (Figura 22a).

No caso de N, o modelo RF apresentou um R^2 de 0,17, indicando que cerca de 17% da variação nos níveis de nitrogênio do solo foi explicada pelo modelo. O LCCC foi de 0,31, evidenciando uma concordância baixa entre os valores observados e os valores previstos pelo modelo RF. O RMSE foi de 0,90 g/kg, enquanto o MAE foi de 0,50 g/kg, apontando para erros médios quadráticos e absolutos moderados entre os valores observados e os valores previstos pelo modelo RF. O RPIQ foi de 0,82, indicando um ajuste razoável entre a variabilidade dos valores observados e os previstos pelo modelo RF (Figura 22b).

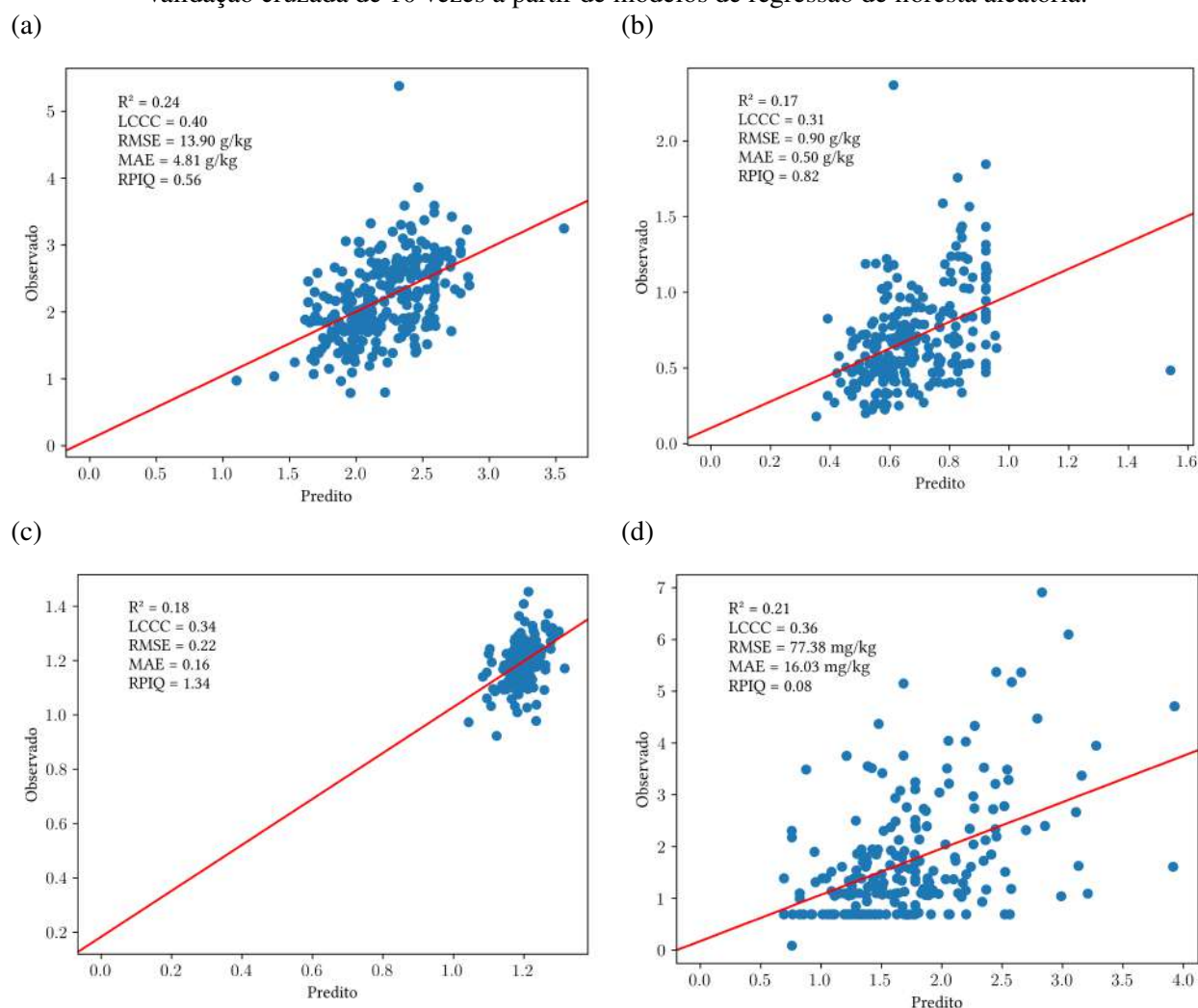
Quanto a razão C:N, o modelo RF apresentou um R^2 de 0,18, indicando que aproximadamente 18% da variação na razão C:N do solo foi explicada pelo modelo. O LCCC foi de 0,34, revelando uma concordância moderada entre os valores observados e os previstos pelo modelo RF. O RMSE foi de 0,22, enquanto o MAE foi de 0,16, sugerindo erros médios quadráticos e absolutos baixos entre os valores observados e os valores previstos pelo modelo RF. O RPIQ foi de 1,34, indicando um bom ajuste entre a variabilidade dos valores observados e a variabilidade dos valores previstos pelo modelo RF (Figura 22c).

Para o P, o modelo RF apresentou um R^2 de 0,21, o que significa que cerca de 21% da variação nos níveis de fósforo do solo foi explicada pelo modelo. O LCCC foi de 0,36, indicando uma concordância moderada entre os valores observados e os valores previstos pelo modelo RF. O RMSE foi de 77,38 mg/kg. O MAE foi de 16,03 mg/kg. No entanto, o RPIQ foi de apenas 0,08, sugerindo um ajuste relativamente baixo do modelo RF (Figura 22d).

O modelo *Random Forest* tem sido comumente utilizada para Modelagem Digital de Solos (MDS) (HENGL *et al.*, 2015; ZHANG *et al.*, 2019; ZERAATPISHEH *et al.*, 2019; ZHU

et al., 2019; ZHOU *et al.*, 2020b). Este modelo ensemble é robusto, pois se beneficia de um conjunto de árvores de decisão para melhorar a capacidade de generalização e determinar uma previsão mais precisa e com baixa variância (BREIMAN, 2001). Podemos justificar o baixo desempenho nos testes de validação dos modelos, devido à escassez de dados pedológicos e sua baixa representatividade para refletir à heterogeneidade do bioma caatinga. Ademais, o uso de dados legados implica incertezas desconhecidas devido à medição, localização posicional e mudanças temporais das informações do solo (GOMES *et al.*, 2019).

Figura 22 – Comparação entre os valores previstos e observados para os teores de carbono (a), nitrogênio (b), razão C:N (c) e fósforo no solo (d) na faixa de profundidade de 0-20 cm, utilizando validação cruzada de 10 vezes a partir de modelos de regressão de floresta aleatória.



Nota: R^2 foi realizada com os valores transformados em ln.

Fonte: Elaborada pelo autor.

4.3.6 Previsão das Propriedades Nutricionais do Solo

As Figuras de 23 - 26 apresentam os mapas de distribuição dos nutrientes do solo gerados pelo modelo Random Forest (RF) para o bioma Caatinga no intervalo de 0-20 cm (Figuras 23 - 26). O padrão espacial dos nutrientes dos solos apresenta um padrão irregular, apresentando inúmeras manchas de pequenos e grandes valores de C, N, C:N e P em toda a área de estudo, apontando para uma variabilidade pronunciada em pequenas distâncias.

Os mapas finais mostraram uma relação estreita entre os aspectos topoclimáticos e os parâmetros de C, N, C:N e P. Dessa forma, é possível relacionar os padrões espaciais dos nutrientes com os compartimentos geomorfológicos na Caatinga (vide Figura 5).

As áreas que correspondem as chapadas, patamares, planaltos, serras e tabuleiros apresentaram concentrações maiores de C (> 10 g/kg), N (> 1 g/kg) e C:N (> 9). Esses ambientes apresentam temperaturas amenas e um aumento significativo na precipitação anual na Caatinga. Evidenciando, além disso, que os valores mais elevados de todos os nutrientes estão localizados nas áreas de chuvas mais intensas ao longo do ano (Figuras 23 - 25). A Figura 26 sugere maior valor de P distribuído principalmente na província geológica da Borborema.

O acúmulo de C e N é maior no clima mais úmido da Caatinga, que pode ser acompanhado por maior biomassa vegetal acima do solo, com predominância de florestas decíduas e subdecíduas, quando comparados com solos sob clima semiárido na Caatinga. No clima seco sub-úmido, solos sob influência de florestas decíduas apresentaram maior C do que solos sob influência de florestas hipoxerófitas e hiperxerófilas, respectivamente, vegetação predominante de clima semiárido Filho *et al.* (2022). Portanto, com o aumento da aridez, a proporção de mudanças conforme a quantidade de matéria orgânica do solo diminui e as quantidades de ânions carbonato, sulfato e nitrato na solução do solo aumentam (STRAWN *et al.*, 2020).

A disponibilidade de P dos solos também são favorecidos na condição mais úmida (FILHO, 2023). Assim sendo, a disponibilidade de P é um controle estequiométrico da decomposição do COS em ambientes tropicais (MANZONI *et al.*, 2010). O P desempenha um papel importante em uma variedade de processos celulares e no nível da planta, como a estimulação da germinação de sementes, rendimento de culturas e qualidade do solo. Além disso, aumenta a capacidade de fixação de N das plantas leguminosas. Portanto, P é essencial em todas as etapas do desenvolvimento, desde a germinação até a maturidade (MALHOTRA *et al.*, 2018).

Como esperado, a distribuição do N do solo está altamente correlacionada com o COS, uma vez que o N é o principal componente da matéria orgânica do solo. Dada esta

relação, fica bastante claro que a cobertura vegetal e o clima são os principais impulsionadores da distribuição de nitrogênio na área estudada conforme observado por Ballabio *et al.* (2019).

O C fornece a energia e matriz para os compostos nos quais é fixado o nitrogênio. O N, por sua vez, estão entre os elementos que governam a taxa de crescimento e fotossíntese das plantas. Alterações do solo que alteram a relação C:N de um equilíbrio natural podem promover ou inibir a decomposição microbiana da matéria orgânica do solo (STRAWN *et al.*, 2020).

Em geral, os valores de teores mais baixos de C (<10 g/kg) (Figura 23), N (< 1 g/kg) (Figura 24), razão C:N (< 9) (Figura 25) e P (<2,7 mg/kg) (Figura 26) estão distribuídos sob o clima semiárido predominante com chuvas limitadas. Encontrando-se no norte dos estados do Ceará e Piauí, na Bahia, na divisa com Pernambuco e Piauí, e em áreas desertificadas da Caatinga, onde são comuns temperaturas elevadas (acima de 24°C).

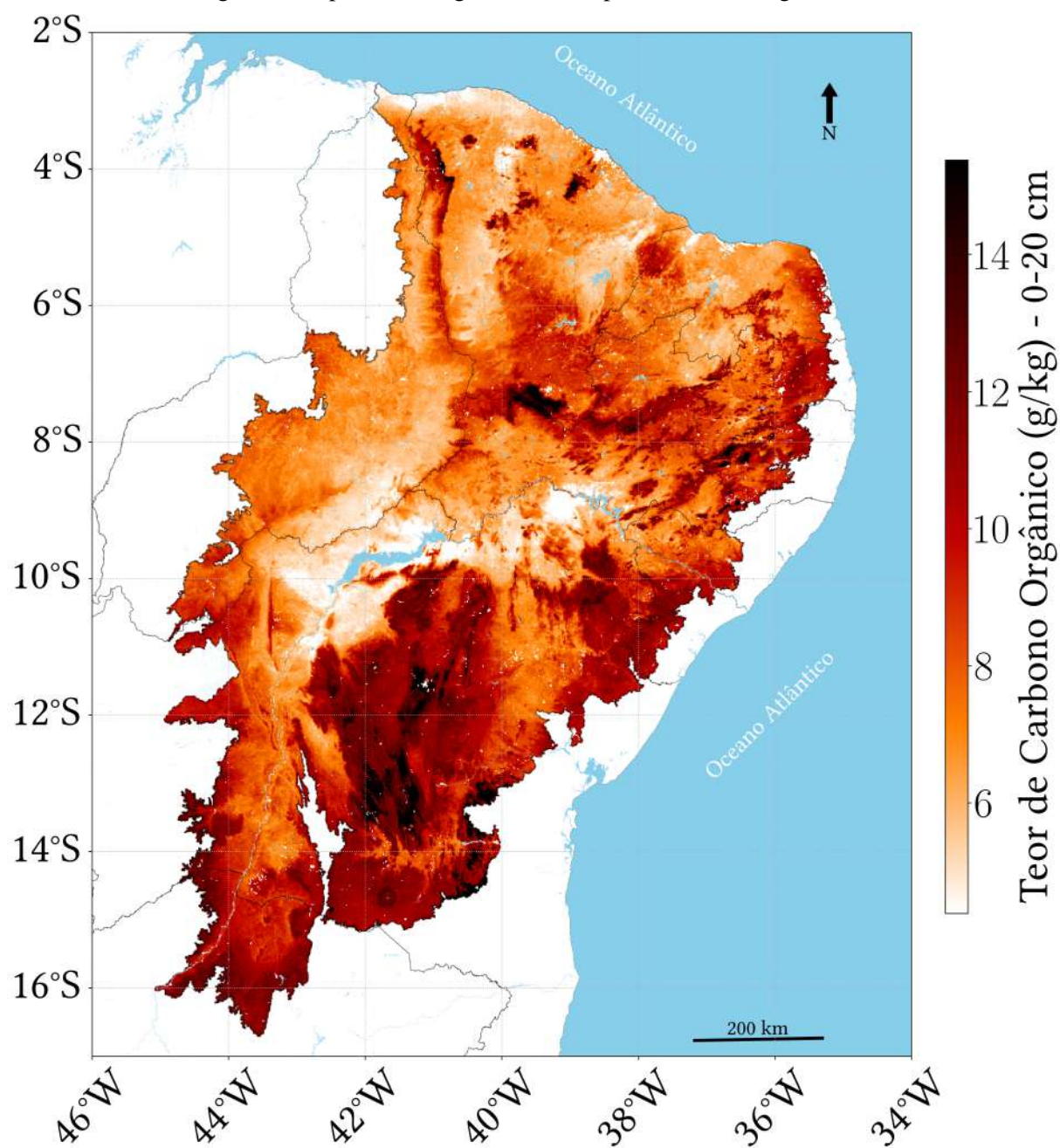
O baixo teor de C e N pode ser causado pelo déficit hídrico, baixa umidade do solo, baixa umidade do ar e elevadas temperaturas na Caatinga. Pois essas condições ambientais afetam o solo, interferindo na atividade microbiana (ANDRADE *et al.*, 2020). Na interface solo-planta-atmosfera, a fonte de carbono orgânico do solo é proveniente principalmente da fotossíntese ou da produtividade primária líquida, assimilados podem ser transferidos diretamente para as raízes através do floema ou são convertíveis em biomassa e posteriormente depositado no solo via decomposição da matéria orgânica (KUTSCH *et al.*, 2010). Assim, os diferentes componentes da planta alocados a partir da fotossíntese líquida que atuam como insumos para impulsionar a decomposição e a formação de C orgânico no solo (COS). A formação e conversão do COS é influenciada pela água no solo, que afetam os processos biogeoquímicos, fornecendo o ambiente físico, o meio de reação ou a matéria reagente (SUN *et al.*, 2019).

Não obstante, estudos revelaram que o C pode variar com as diferenças entre composição das comunidades vegetativas no bioma Caatinga, os solos com alto teor de C estão associados à maior biodiversidade e à ocorrência de remanescentes de fragmentos de floresta semidecídua em clima semiárido (SOUZA *et al.*, 2022; RAMOS *et al.*, 2023). O que demonstra a sua importância para o equilíbrio e manutenção da qualidade ambiental em regiões sequeiras. A conversão dos resquícios de caatingas mais densas em áreas agropecuárias reduz significativamente os estoques de carbono entre os diferentes tipos de solos do bioma (SANTANA *et al.*, 2019). Neste contexto, o uso da terra tem mostrado um maior potencial de reduzir o carbono orgânico total, reflexo de uma má gestão secular dos recursos naturais na Caatinga (ANDRADE *et al.*, 2020).

Na Caatinga a demanda da população por recursos madeireiros, desmatamentos em ciclos curtos, práticas tradicionais de gerenciamento do uso da terra como queimadas, remoção de resíduos orgânicos, lavoura convencional e cultivo limpo, monoculturas e pecuária são práticas que, associadas à baixa produção natural de biomassa, contribuem para a redução do conteúdo de matéria orgânica dos solos (ALTHOFF *et al.*, 2018; ANDRADE *et al.*, 2020). No entanto, ressalta-se que isso pode indicar um potencial risco ao aquecimento global, tendo em vista a sensibilidade de redução da matéria orgânica em áreas de vegetação nativa pode aumentar as perdas de estoque de carbono, contribuindo para um feedback positivo sobre as mudanças climáticas e comprometendo o desempenho produtivo dos sistemas da região (MAIA *et al.*, 2019).

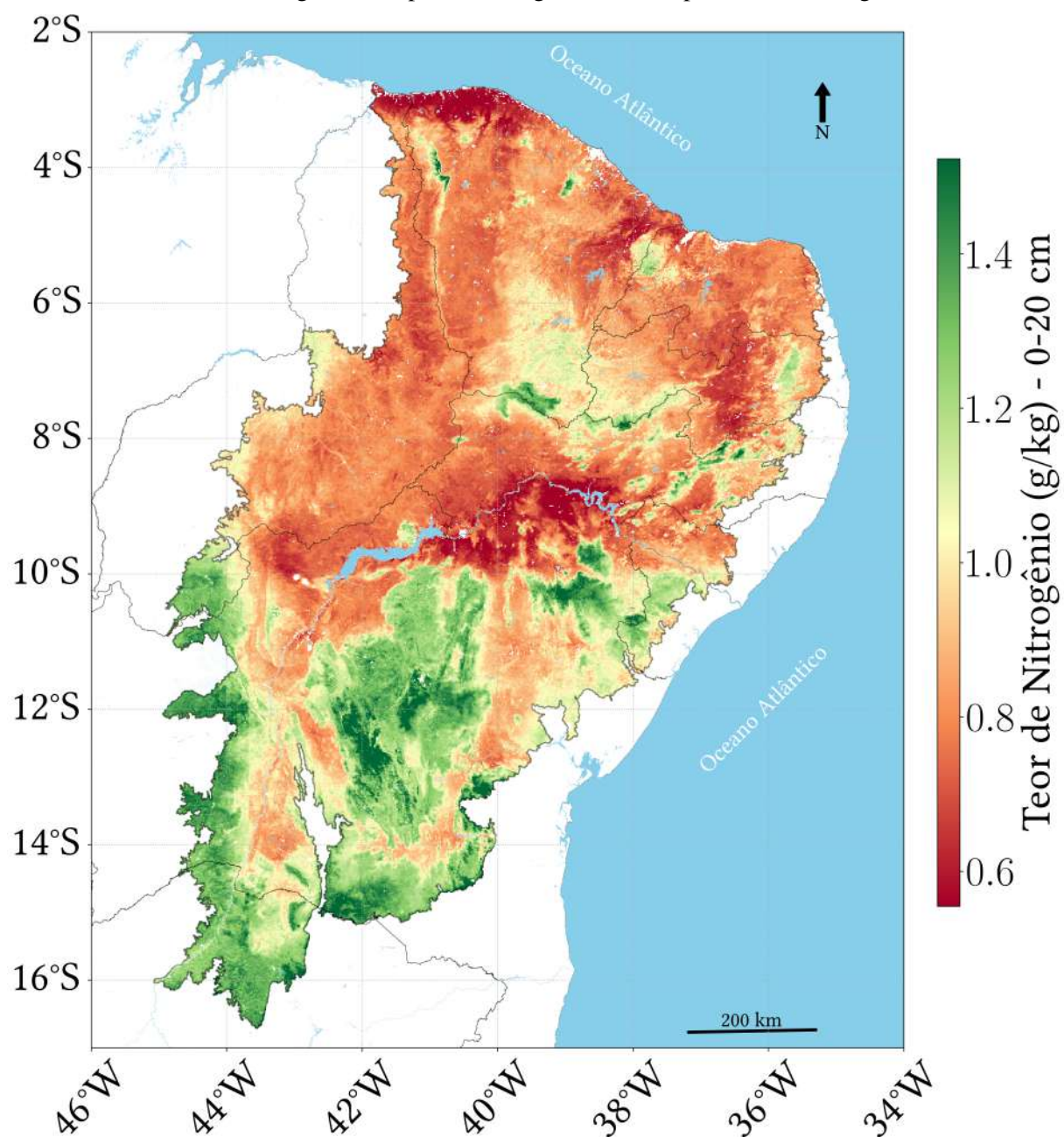
No caso da pecuária, está também demonstrou ter um impacto negativo sobre os estoques de carbono orgânico do solo, diminuindo a resiliência do ecossistema e a fertilidade (SCHULZ *et al.*, 2016). Mesmo assim, dependendo do tipo de solo, a exploração sustentável da Caatinga para produzir biomassa e proteína pode ser adotado em áreas de cobertura vegetal aberta e pastagem, como por exemplo, em Neossolos que apresenta um estoque carbono orgânico reduzido (ANDRADE *et al.*, 2020).

Figura 23 – Teor de carbono orgânico (g/kg) para o intervalo de profundidade de 0-20 cm estimado usando a abordagem de mapeamento digital do solo a partir de dados legados de solo.



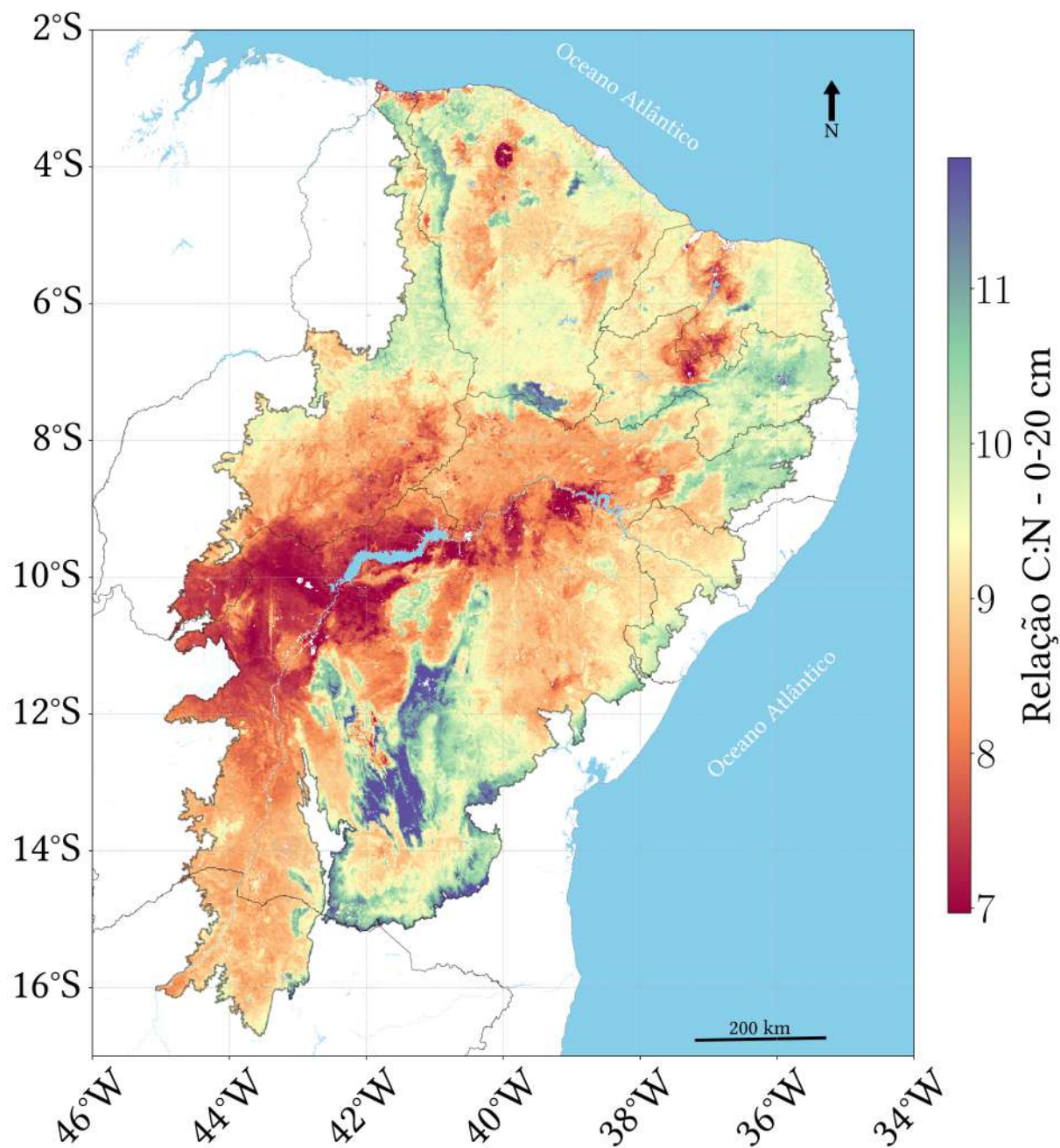
Fonte: Elaborada pelo autor.

Figura 24 – Teor de carbono nitrogênio (g/kg) para o intervalo de profundidade de 0–20 cm estimado usando a abordagem de mapeamento digital do solo a partir de dados legados de solo.



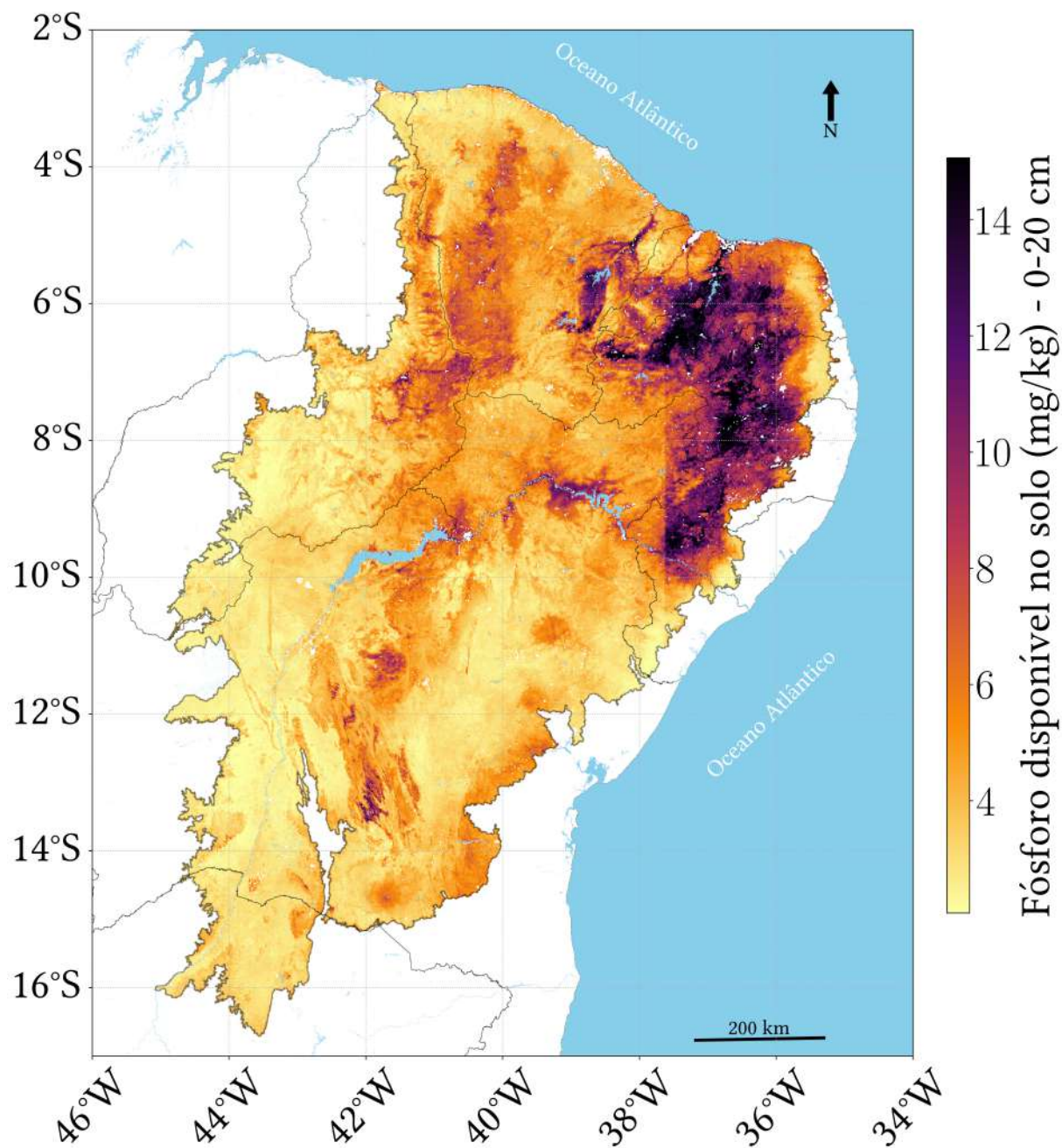
Fonte: Elaborada pelo autor.

Figura 25 – Razão C:N para o intervalo de profundidade de 0–20 cm estimado usando a abordagem de mapeamento digital do solo a partir de dados legados de solo.



Fonte: Elaborada pelo autor.

Figura 26 – Fósforo disponível no solo para o intervalo de profundidade de 0–20 cm estimado usando a abordagem de mapeamento digital do solo a partir de dados legados de solo.



Fonte: Elaborada pelo autor.

5 CONSIDERAÇÕES FINAIS

A partir dos resultados obtidos foi possível confirmar a hipótese de que com o uso da Mapeamento Digital dos Solos pode ser uma ferramenta rápida e razoavelmente confiável para modelar o teor de C, N, C:N e P, além dos parâmetros estruturais da vegetação (altura e densidade) para o bioma Caatinga.

A avaliação das métricas para todos os modelos de nutriente dos solos não foram satisfatórias. Isso acontece porque as informações de solo usadas na modelagem são limitadas, devido aos métodos de amostragem diferirem e não utilizarem coordenadas precisas. Vale salientar que os dados utilizados são os únicos disponíveis na área de estudo.

O uso de dados ópticos *LiDAR GEDI* unidos aos satélites alta resolução (*Sentinel 1* e *2*) apresenta uma moderada precisão para estimar os parâmetros da Caatinga. Como mencionado, esses são os primeiros produtos cartográficos que estimaram a altura e densidade da vegetação nesse bioma. Isso é de fundamental importância para a gestão das florestas e uma fonte de dados para relacionar com outros componentes da paisagem, além de compreender ciclos biogeoquímicos que estão na interface solo-planta-atmosfera.

Identificamos que os atributos do relevo e mesoclimáticos influenciaram significativamente nos nutrientes do solo e nos parâmetros estruturais da vegetação, assim como revelado em trabalhos anteriores no bioma caatinga.

Os estratos Arbustivo Alto e Arbóreo Intermediário tiveram os melhores valores para C, N, C:N e P. Esses estratos são bons representativos da condição de caatinga preservada. Não foi possível identificar um padrão em relação aos nutrientes do solo e cobertura do dossel, o que é relevante, pois os produtos oriundos de índices espectrais da vegetação gerados por sensores orbitais que conseguem ver a densidade do dossel são amplamente usados para identificar áreas degradadas e desertificadas do bioma Caatinga.

As áreas desflorestadas/degradadas têm baixos valores de C e N. Isso mostra a relação desses parâmetros do solo para avaliar a qualidade do ambiente. O que gera a necessidade de incluir esses elementos nos indicadores de áreas desertificadas.

Este trabalho gerou informações relevantes no âmbito do bioma, trazendo novos *insights* para entender a estrutura florestal da Caatinga. Pois, até onde se sabe, ainda não tínhamos dados em escala regional dos parâmetros estimados na presente pesquisa.

As amostras localizadas nos enclaves úmidos e subúmidos e nas bases dos afloramentos rochosos registraram os maiores teores médios de C e N e C:N, demonstrando a im-

portância dessas áreas para a manutenção e dinâmica desses ecossistemas, funcionando como zona-tampão das mudanças e variações climáticas. Portanto, essas áreas são essenciais para as políticas ambientais e para o orçamento global de C nas regiões semiáridas.

A definição de novos pontos de coleta de amostragem pode melhorar os resultados obtidos facilmente modeladas a partir de algoritmos de *machine learnig* em arquitetura de computação em nuvem disponível no *Google Earth Engine* (GEE) ou em outras plataformas disponíveis. Os dados dessas pesquisas serão importantes para estudos futuros como projeções climáticas.

REFERÊNCIAS

- ABEGAZ, Assefa *et al.* Spatial and temporal dynamics of soil organic carbon in landscapes of the upper blue Nile basin of the Ethiopian highlands. **Agriculture, Ecosystems & Environment**, v. 218, p. 190–208, fev. 2016. ISSN 01678809.
- AB’SABER, Aziz Nacib. Os domínios morfoclimáticos na América do Sul: Primeira aproximação. **Geomorfologia**, n. 52, p. 1–22, 1977.
- ALEXANDER, Earl B. **Soils in Natural Landscapes**. 0. ed. Boca Raton: CRC Press, 2013. ISBN 978-0-429-16817-8.
- ALLEN, Richard G *et al.* **Crop Evapotranspiration-Guidelines for Computing Crop Water Requirements-FAO Irrigation and Drainage Paper 56**. Rome: Food and Agriculture Organization, 1998.
- ALTHOFF, Tiago Diniz *et al.* Adaptation of the Century model to simulate C and N dynamics of Caatinga dry forest before and after deforestation. **Agriculture, Ecosystems & Environment**, v. 254, p. 26–34, fev. 2018. ISSN 01678809.
- ALVARES, Clayton Alcarde; STAPE, Jose Luiz; SENTELHAS, Paulo Cesar. Köppen’s climate classification map for Brazil. **Meteorol. Z.**, 2013.
- ALVAREZ, VVH de *et al.* Determinação e uso do fósforo remanescente. **Boletim Informativo da Sociedade Brasileira de Ciência do Solo**, v. 25, n. 1, p. 27–34, 2000.
- ANDRADE, Eunice *et al.* Land-use effect on soil carbon and nitrogen stock in a seasonally dry tropical forest. **Agronomy**, v. 10, n. 2, p. 158, jan. 2020. ISSN 2073-4395.
- Andrade-Lima, Dárdano de. The Caatinga dominium. **Revista brasileira de Botânica**, v. 4, p. 149–163, 1981.
- ANTONGIOVANNI, Marina *et al.* Chronic anthropogenic disturbance on Caatinga dry forest fragments. **Journal of Applied Ecology**, v. 57, n. 10, p. 2064–2074, out. 2020. ISSN 0021-8901, 1365-2664.
- ARNOLD, Richard W. Concepts of soils. In: _____. **Soils: Basic Concepts and Future Challenges**. Cambridge; New York: Cambridge University Press, 2006. p. 1–10.
- BALLABIO, Cristiano *et al.* Mapping LUCAS topsoil chemical properties at European scale using Gaussian process regression. **Geoderma**, v. 355, p. 113912, dez. 2019. ISSN 00167061.
- BANNARI, A. *et al.* A review of vegetation indices. **Remote Sensing Reviews**, v. 13, n. 1-2, p. 95–120, ago. 1995. ISSN 0275-7257.
- BANWART, Steven A. *et al.* Soil functions: Connecting Earth’s critical zone. **Annual Review of Earth and Planetary Sciences**, v. 47, n. 1, p. 333–359, maio 2019. ISSN 0084-6597, 1545-4495.
- BARDGETT, Richard D. Plant-soil interactions in a changing world. **F1000 Biology Reports**, v. 3, ago. 2011. ISSN 1757594X.
- BELAND, Martin *et al.* On promoting the use of lidar systems in forest ecosystem research. **Forest Ecology and Management**, v. 450, p. 117484, out. 2019. ISSN 03781127.

BISHOP, T.F.A.; MCBRATNEY, A.B.; LASLETT, G.M. Modelling soil attribute depth functions with equal-area quadratic smoothing splines. **Geoderma**, v. 91, n. 1-2, p. 27–45, ago. 1999. ISSN 00167061.

BOETTINGER, Janis L. *et al.* (Ed.). **Digital Soil Mapping: Bridging Research, Environmental Application, and Operation**. Dordrecht: Springer Netherlands, 2010. ISBN 978-90-481-8862-8 978-90-481-8863-5.

BRADY, Nyle C.; WEIL, Ray R. **Elementos Da Natureza e Propriedades Dos Solos**. 3. ed. Porto Alegre: Bookman, 2013.

BREIMAN, Leo. Random forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001. ISSN 08856125.

BREVIK, Eric C.; HARTEMINK, Alfred E. Early soil knowledge and the birth and development of soil science. **CATENA**, v. 83, n. 1, p. 23–33, out. 2010. ISSN 03418162.

CAMARRETTA, Nicolò *et al.* Monitoring forest structure to guide adaptive management of forest restoration: A review of remote sensing approaches. **New Forests**, v. 51, n. 4, p. 573–596, jul. 2020. ISSN 0169-4286, 1573-5095.

CARLSON, Toby N.; RIPLEY, David A. On the relation between NDVI, fractional vegetation cover, and leaf area index. **Remote Sensing of Environment**, v. 62, n. 3, p. 241–252, dez. 1997. ISSN 00344257.

CHEN, Di *et al.* Mapping dynamics of soil organic matter in croplands with MODIS data and machine learning algorithms. **Science of The Total Environment**, v. 669, p. 844–855, jun. 2019. ISSN 00489697.

CHESWORTH, Ward (Ed.). **Encyclopedia of Soil Science**. Dordrecht, Netherlands: Springer, 2008. (Encyclopedia of Earth Sciences Series). ISBN 978-1-4020-3994-2.

CONNERS, Richard W.; TRIVEDI, Mohan M.; HARLOW, Charles A. Segmentation of a high-resolution urban scene using texture operators. **Computer Vision, Graphics, and Image Processing**, v. 25, n. 3, p. 273–310, mar. 1984. ISSN 0734189X.

CRIST, Eric P. A TM tasseled cap equivalent transformation for reflectance factor data. **Remote Sensing of Environment**, v. 17, n. 3, p. 301–306, jun. 1985. ISSN 00344257.

DINTWE, Kebonye; OKIN, Gregory S. Soil organic carbon in savannas decreases with anthropogenic climate change. **Geoderma**, v. 309, p. 7–16, jan. 2018. ISSN 00167061.

DONAGEMA, Guilherme Kangussú *et al.* **Manual de Métodos de Análise de Solo**. 2. ed. Rio de Janeiro: Embrapa Solos, 2011.

ELBASIOUNY, Heba *et al.* Plant nutrition under climate change and soil carbon sequestration. p. 20, 2022.

EMBRAPA. **Sistema Brasileiro de Classificação de Solos**. Brasília: Embrapa, 2018.

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, out. 2017. ISSN 0899-8418, 1097-0088.

FILHO, José de Souza Oliveira. How do climate and land-use affect soil carbon and nitrogen stocks and the chemical properties of floodplain soils in tropical drylands? **CATENA**, v. 231, p. 107289, out. 2023. ISSN 03418162.

FILHO, José De Souza Oliveira *et al.* How does increasing humidity in the environment affect soil carbon and nitrogen stocks and the C/N ratio in tropical drylands? Evidence from northeastern Brazil. **CATENA**, v. 213, p. 106208, jun. 2022. ISSN 03418162.

FUJII, Kazumichi *et al.* Plant–soil interactions maintain biodiversity and functions of tropical forest ecosystems. **Ecological Research**, v. 33, n. 1, p. 149–160, jan. 2018. ISSN 0912-3814, 1440-1703.

GANEM, Khalil Ali *et al.* Mapeamento da vegetação da caatinga a partir de dados ópticos de observação da terra – oportunidades e desafios. **Revista Brasileira de Cartografia**, v. 72, p. 829–854, dez. 2020. ISSN 1808-0936, 0560-4613.

GANEM, Roseli Senna. **Caatinga: estratégias de conservação**. Brasília: Câmara dos Deputados, 2017.

GAO, Bo-cai. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. **Remote Sensing of Environment**, v. 58, n. 3, p. 257–266, dez. 1996. ISSN 00344257.

GARG, Pradeep Kumar *et al.* **Digital Mapping of Soil Landscape Parameters: Geospatial Analyses Using Machine Learning and Geomatics**. Singapore: Springer Singapore, 2020. v. 72. (Studies in Big Data, v. 72). ISBN 9789811532375 9789811532382.

GITELSON, Anatoly A. Remote estimation of canopy chlorophyll content in crops. **Geophysical Research Letters**, v. 32, n. 8, p. L08403, 2005. ISSN 0094-8276.

GITELSON, Anatoly A.; KAUFMAN, Yoram J.; MERZLYAK, Mark N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. **Remote Sensing of Environment**, v. 58, n. 3, p. 289–298, dez. 1996. ISSN 00344257.

GOMES, Lucas Carvalho *et al.* Modelling and mapping soil organic carbon stocks in brazil. **Geoderma**, v. 340, p. 337–350, abr. 2019. ISSN 00167061.

GORELICK, Noel *et al.* Google earth engine: Planetary-scale geospatial analysis for everyone. **Remote Sensing of Environment**, v. 202, p. 18–27, dez. 2017. ISSN 00344257.

GRUNWALD, S. Current state of digital soil mapping and what is next. In: BOETTINGER, Janis L. *et al.* (Ed.). **Digital Soil Mapping**. Dordrecht: Springer Netherlands, 2010. p. 3–12. ISBN 978-90-481-8862-8 978-90-481-8863-5.

GRUNWALD, S. *et al.* Are current scientific visualisation and virtual reality techniques capable to represent real soil-landscapes? In: LAGACHERIE, P.; MCBRATNEY, A. B.; VOLTZ, M. (Ed.). **Developments in Soil Science**. Amsterdam: Elsevier, 2006. v. 31, p. 571–634. ISBN 978-0-444-52958-9.

GUAN, Fengying *et al.* Spatial variability of soil nitrogen, phosphorus and potassium contents in moso bamboo forests in yong'an city, china. **CATENA**, v. 150, p. 161–172, mar. 2017. ISSN 03418162.

GUPTA, Rajit; SHARMA, Laxmi Kant. Mixed tropical forests canopy height mapping from spaceborne LiDAR GEDI and multisensor imagery using machine learning models. **Remote Sensing Applications: Society and Environment**, v. 27, p. 100817, ago. 2022. ISSN 23529385.

HARALICK, Robert M.; SHANMUGAM, K.; DINSTEN, Its'Hak. Textural features for image classification. **IEEE Transactions on Systems, Man, and Cybernetics**, SMC-3, n. 6, p. 610–621, nov. 1973. ISSN 0018-9472, 2168-2909.

HARTEMINK, Alfred E.; MCBRATNEY, A. B.; Mendonça-Santos, Maria de Lourdes (Ed.). **Digital Soil Mapping with Limited Data**. Dordrecht ; London: Springer, 2008. ISBN 978-1-4020-8591-8.

HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, Jerome. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. 2. ed. New York: Springer, 2013.

HAWKER, Laurence *et al.* A 30 m global map of elevation with forests and buildings removed. **Environmental Research Letters**, v. 17, n. 2, p. 024016, fev. 2022. ISSN 1748-9326.

HENGL, Tomislav *et al.* Mapping soil properties of africa at 250 m resolution: Random forests significantly improve current predictions. **PLOS ONE**, v. 10, n. 6, p. e0125814, jun. 2015. ISSN 1932-6203.

HENGL, Tomislav; MACMILLAN, Robert A. **Predictive Soil Mapping with R**. Wageningen: OpenGeoHub foundation, 2019. ISBN 978-0-359-30635-0.

HOUNKPATIN, Ozias K.L. *et al.* Soil organic carbon stocks and their determining factors in the dano catchment (southwest burkina faso). **CATENA**, v. 166, p. 298–309, jul. 2018. ISSN 03418162.

HUANG, Pan Ming; LI, Yuncong; SUMNER, Malcolm E. **Handbook of Soil Sciences**. 2. ed. New York: CRC Press, 2012. ISBN 13: 978-1-4398-0306-6.

HUETE, A.R. A soil-adjusted vegetation index (SAVI). **Remote Sensing of Environment**, v. 25, n. 3, p. 295–309, ago. 1988. ISSN 00344257.

HUETE, A *et al.* Overview of the radiometric and biophysical performance of the MODIS vegetation indices. **Remote Sensing of Environment**, v. 83, n. 1-2, p. 195–213, nov. 2002. ISSN 00344257.

IBGE, Instituto Brasileiro de Geografia e Estatística. **Províncias Estruturais, Compartimentos de Relevô, Tipos de Solos e Regiões Fitoecológicas**. Rio de Janeiro: IBGE, Coordenação de Recursos Naturais e Estudos Ambientais, 2019.

JENNY, Hans. **Factors of Soil Formation: A System of Quantitative Pedology**. New York: Dover, 1994. ISBN 978-0-486-68128-3.

JL, Lei *et al.* On the terminology of the spectral vegetation index (NIR - SWIR)/(NIR + SWIR). **International Journal of Remote Sensing**, v. 32, n. 21, p. 6901–6909, nov. 2011. ISSN 0143-1161, 1366-5901.

JOHNSON, William M. The pedon and the polypedon. **Soil Science Society of America Journal**, v. 27, n. 2, p. 212–215, mar. 1963. ISSN 03615995.

JORDAN, Carl F. Derivation of leaf-area index from quality of light on the forest floor. **Ecology**, v. 50, n. 4, p. 663–666, jul. 1969. ISSN 00129658.

KAUFMAN, Yoram J.; TANRÉ, Didier. Strategy for direct and indirect methods for correcting the aerosol effect on remote sensing: From AVHRR to EOS-MODIS. **Remote Sensing of Environment**, v. 55, n. 1, p. 65–79, jan. 1996. ISSN 00344257.

KHEIR, Rania Bou *et al.* Predictive mapping of soil organic carbon in wet cultivated lands using classification-tree based models: The case study of denmark. **Journal of Environmental Management**, v. 91, n. 5, p. 1150–1160, maio 2010. ISSN 03014797.

KUTSCH, W. L.; BAHN, M.; HEINEMEYER, A. Soil carbon relations: An overview. In: **Soil Carbon Dynamics**. Cambridge: Cambridge University Press, 2010. p. 1–15.

LAGACHERIE, P.; MCBRATNEY, A.B. Spatial soil information systems and spatial soil inference systems: Perspectives for digital soil mapping. In: LAGACHERIE, P.; MCBRATNEY, A. B.; VOLTZ, M. (Ed.). **Developments in Soil Science**. Amsterdam: Elsevier, 2006. v. 31, p. 3–22. ISBN 978-0-444-52958-9.

LAL, R. Sequestration of atmospheric CO₂ in global carbon pools. **Energy & Environmental Science**, v. 1, n. 1, p. 86, 2008. ISSN 1754-5692, 1754-5706.

LAMICHHANE, Sushil; KUMAR, Lalit; WILSON, Brian. Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review. **Geoderma**, v. 352, p. 395–413, out. 2019. ISSN 00167061.

LAUSCH, A. *et al.* Linking earth observation and taxonomic, structural and functional biodiversity: Local to ecosystem perspectives. **Ecological Indicators**, v. 70, p. 317–339, nov. 2016. ISSN 1470160X.

LIAKOS, Konstantinos *et al.* Machine learning in agriculture: A review. **Sensors**, v. 18, n. 8, p. 2674, ago. 2018. ISSN 1424-8220.

LIANG, Shunlin *et al.* Narrowband to broadband conversions of land surface albedo: II. Validation. **Remote Sensing of Environment**, v. 84, n. 1, p. 25–41, jan. 2003. ISSN 00344257.

LIESS, Mareike; SCHMIDT, Johannes; GLASER, Bruno. Improving the spatial prediction of soil organic carbon stocks in a complex tropical mountain landscape by methodological specifications in machine learning approaches. **PLOS ONE**, v. 11, n. 4, p. e0153673, abr. 2016. ISSN 1932-6203.

LIN, H. Earth's critical zone and hydopedology: Concepts, characteristics, and advances. **Hydrol. Earth Syst. Sci.**, p. 21, 2010.

Linares-Palomino, Reynaldo; Oliveira-Filho, Ary T; PENNINGTON, R Toby. Neotropical seasonally dry forests: Diversity, endemism, and biogeography of woody plants. In: **Seasonally Dry Tropical Forests: Ecology and Conservation**. Washington: Island Press, 2011. p. 3–21.

MA, Xiao *et al.* Mapping fine-scale building heights in urban agglomeration with spaceborne lidar. **Remote Sensing of Environment**, v. 285, p. 113392, fev. 2023. ISSN 00344257.

MAHMOUDZADEH, Hamid *et al.* Spatial prediction of soil organic carbon using machine learning techniques in western iran. **Geoderma Regional**, v. 21, p. e00260, jun. 2020. ISSN 23520094.

MAIA, Stoécio Malta Ferreira *et al.* Soil organic carbon temperature sensitivity of different soil types and land use systems in the brazilian semi-arid region. **Soil Use and Management**, v. 35, n. 3, p. 433–442, set. 2019. ISSN 0266-0032, 1475-2743.

MALHOTRA, Hina *et al.* Phosphorus nutrition: Plant growth in response to deficiency and excess. In: HASANUZZAMAN, Mirza *et al.* (Ed.). **Plant Nutrients and Abiotic Stress Tolerance**. Singapore: Springer Singapore, 2018. p. 171–190. ISBN 978-981-10-9043-1 978-981-10-9044-8.

MALONE, Brendan P.; MINASNY, Budiman; MCBRATNEY, Alex B. **Using R for Digital Soil Mapping**. Cham: Springer International Publishing, 2017. (Progress in Soil Science). ISBN 978-3-319-44325-6 978-3-319-44327-0.

MANZONI, Stefano *et al.* Stoichiometric controls on carbon, nitrogen, and phosphorus dynamics in decomposing litter. **Ecological Monographs**, v. 80, n. 1, p. 89–106, fev. 2010. ISSN 0012-9615.

MCBRATNEY, A.B; SANTOS, M.L Mendonça; MINASNY, B. On digital soil mapping. **Geoderma**, v. 117, n. 1-2, p. 3–52, nov. 2003. ISSN 00167061.

MCBRATNEY, A. B. **Introduction to Pedometrics: A Course of Lectures, CSIRO Division of Soils Technical Memorandum 53/1986**. Glen Osmond: CSIRO Division of Soils, 1986.

MCBRATNEY, Alex. B.; MINASNY, Budiman; STOCKMANN, Uta (Ed.). **Pedometrics**. Cham: Springer International Publishing, 2018. (Progress in Soil Science). ISBN 978-3-319-63437-1 978-3-319-63439-5.

MENDONÇA, Francisco; Dannio-Oliveira, Inês Moresco. **Climatologia: noções básicas e climas do brasil**. São Paulo: Oficina de Textos, 2007. ISBN 978-85-86238-54-3.

MILLER, B.A.; SCHAETZL, R.J. The historical role of base maps in soil geography. **Geoderma**, v. 230–231, p. 329–339, out. 2014. ISSN 00167061.

MINASNY, B. *et al.* Pedometrics. In: **Reference Module in Earth Systems and Environmental Sciences**. [S.l.]: Elsevier, 2014. p. B9780124095489091636. ISBN 978-0-12-409548-9.

MINASNY, Budiman *et al.* Soil carbon 4 per mille. **Geoderma**, v. 292, p. 59–86, abr. 2017. ISSN 00167061.

MORIN, David *et al.* Improving heterogeneous forest height maps by integrating GEDI-based forest height information in a multi-sensor mapping process. **Remote Sensing**, v. 14, n. 9, p. 2079, abr. 2022. ISSN 2072-4292.

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, jun. 2016. ISSN 0006-8101, 1874-9372.

_____. A catalogue of the vascular plants of the caatinga phytogeographical domain: A synthesis of floristic and phytosociological surveys. **Phytotaxa**, v. 160, n. 1, p. 1, fev. 2014. ISSN 1179-3163, 1179-3155.

MUSTHAFA, Mohamed; SINGH, Gulab; KUMAR, Praveen. Comparison of forest stand height interpolation of GEDI and icesat-2 LiDAR measurements over tropical and sub-tropical forests in india. **Environmental Monitoring and Assessment**, v. 195, n. 1, p. 71, jan. 2023. ISSN 0167-6369, 1573-2959.

NEUVILLE, Romain; BATES, Jordan Steven; JONARD, François. Estimating forest structure from UAV-mounted LiDAR point cloud using machine learning. **Remote Sensing**, v. 13, n. 3, p. 352, jan. 2021. ISSN 2072-4292.

NRC, National Research Council. **Basic Research Opportunities in Earth Science**. Washington, D.C.: National Academies Press, 2001. 9981 p. ISBN 978-0-309-07133-8.

O'BRIEN, Lauren. **Mpspline2: Mass-preserving Spline Functions for Soil Data**. 2022.

OSMAN, Khan Towhid. **Soils: Principles, Properties and Management**. Dordrecht: Springer Netherlands, 2013. ISBN 978-94-007-5662-5 978-94-007-5663-2.

PADARIAN, José; MINASNY, Budiman; MCBRATNEY, Alex B. Machine learning and soil sciences: A review aided by machine learning tools. **SOIL**, v. 6, n. 1, p. 35–52, fev. 2020. ISSN 2199-398X.

PENNINGTON, R. Toby; LEHMANN, Caroline E.R.; ROWLAND, Lucy M. Tropical savannas and dry forests. **Current Biology**, v. 28, n. 9, p. R541–R545, maio 2018. ISSN 09609822.

PEREIRA, Brunno Cesar. Relação solo-paisagem e sua aplicabilidade: uma ferramenta fundamental para o entendimento da caracterização da paisagem. **Cadernos do Leste**, v. 20, n. 20, 2020. ISSN 1679-5806.

PRADO, Darién E. As caatingas da América do Sul. **Ecologia e conservação da Caatinga**, v. 2, p. 3–74, 2003.

QI, J. *et al.* A modified soil adjusted vegetation index. **Remote Sensing of Environment**, v. 48, n. 2, p. 119–126, maio 1994. ISSN 00344257.

QUINN, P. *et al.* The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. **Hydrological Processes**, v. 5, n. 1, p. 59–79, jan. 1991. ISSN 08856087, 10991085.

RAHMAN, AF *et al.* Modeling CO2 flux of boreal forests using narrow-band indices from AVIRIS imagery. In: **AVIRIS Workshop**. [S.l.: s.n.], 2000.

RAMOS, Maiara Bezerra *et al.* The role of chronic anthropogenic disturbances in plant community assembly along a water availability gradient in Brazil's semiarid Caatinga region. **Forest Ecology and Management**, v. 538, p. 120980, jun. 2023. ISSN 03781127.

REDDY, Nagarjuna N. *et al.* Legacy data-based national-scale digital mapping of key soil properties in india. **Geoderma**, v. 381, p. 114684, jan. 2021. ISSN 00167061.

REN, Chunying *et al.* Quantifying temperate forest diversity by integrating GEDI LiDAR and multi-temporal sentinel-2 imagery. **Remote Sensing**, v. 15, n. 2, p. 375, jan. 2023. ISSN 2072-4292.

RICHTER, Daniel deB. Humanity's transformation of earth's soil: Pedology's new frontier. **Soil Science**, v. 172, n. 12, p. 957–967, dez. 2007. ISSN 0038-075X.

ROBINSON, D. A. *et al.* On the value of soil resources in the context of natural capital and ecosystem service delivery. **Soil Science Society of America Journal**, v. 78, n. 3, p. 685–700, maio 2014. ISSN 03615995.

ROSS, Jurandyr Luciano Sanches. **Ecogeografia Do Brasil: Subsídios Para o Planejamento Ambiental**. São Paulo: Oficina de textos, 2009.

SAFANELLI, José *et al.* Terrain analysis in google earth engine: A method adapted for high-performance global-scale analysis. **ISPRS International Journal of Geo-Information**, v. 9, n. 6, p. 400, jun. 2020. ISSN 2220-9964.

SANTANA, Mônica da Silva *et al.* Carbon and nitrogen stocks of soils under different land uses in pernambuco state, brazil. **Geoderma Regional**, v. 16, p. e00205, mar. 2019. ISSN 23520094.

SANTOS, Maria de Lourdes Mendonça; MANZATTO, Celso Vainer. Pedologia ou pedometria? Modelo discreto ou contínuo? Mapeamento tradicional ou digital? **BOLETIM INFORMATIVO - Sociedade Brasileira de Ciência do Solo**, v. 32, n. 1, p. 21–26, 2007.

SARMIENTO, G.; MONASTERIO, M. A critical consideration of the environmental conditions associated with the occurrence of savanna ecosystems in tropical america. In: GOLLEY, F; MEDINA, E (Ed.). **Tropical Ecological Systems**. New York: Springer Verlag, 1975. p. 223–250.

SCHAETZL, Randall J.; ANDERSON, Sharon. **Soils: Genesis and Geomorphology**. Cambridge ; New York: Cambridge University Press, 2005. ISBN 978-0-521-81201-6.

SCHILLACI, Calogero *et al.* Spatio-temporal topsoil organic carbon mapping of a semi-arid mediterranean region: The role of land use, soil texture, topographic indices and the influence of remote sensing data to modelling. **Science of The Total Environment**, v. 601–602, p. 821–832, dez. 2017. ISSN 00489697.

SCHOBENHAUS, Carlos; NEVES, Benjamim Bley de Brito. A geologia do brasil no contexto da plataforma sul-americana. In: BIZZI, L. A. *et al.* (Ed.). **Geologia, tectônica e recursos minerais do Brasil**. [S.l.]: CPRM, 2003. p. 5–25.

SCHULZ, Katharina *et al.* Grazing deteriorates the soil carbon stocks of caatinga forest ecosystems in brazil. **Forest Ecology and Management**, v. 367, p. 62–70, maio 2016. ISSN 03781127.

SHAPLEY, Lloyd S. A value for n-person games. **Contributions to the theory of games (AM-28)**, Princeton University Press, v. 2, p. 307–318, 1953.

SHUKLA, Manoj. **Soil Physics: An Introduction**. [S.l.: s.n.], 2014. ISBN 978-1-4822-1686-8.

SILATSA, Francis B.T. *et al.* Assessing countrywide soil organic carbon stock using hybrid machine learning modelling and legacy soil data in cameroon. **Geoderma**, v. 367, p. 114260, maio 2020. ISSN 00167061.

SILVA, Ivo Ribeiro da; MENDONÇA, Eduardo de Sá. Matéria orgânica do solo. In: NOVAIS, RF *et al.* (Ed.). **Fertilidade Do Solo**. Viçosa: Sociedade Brasileira de Ciência do Solo, 2007. p. 275–374.

SILVA, José Maria Cardoso da; LEAL, Inara R.; TABARELLI, Marcelo (Ed.). **Caatinga**. Cham: Springer International Publishing, 2017. ISBN 978-3-319-68338-6 978-3-319-68339-3.

SILVA, Lucas C. R.; LAMBERS, Hans. Soil-plant-atmosphere interactions: Structure, function, and predictive scaling for climate change mitigation. **Plant and Soil**, v. 461, n. 1-2, p. 5–27, abr. 2021. ISSN 0032-079X, 1573-5036.

SIQUEIRA, Rafael G. *et al.* Machine learning applied for antarctic soil mapping: Spatial prediction of soil texture for maritime antarctica and northern antarctic peninsula. **Geoderma**, v. 432, p. 116405, abr. 2023. ISSN 00167061.

SMITH, Pete. Soils and climate change. **Current Opinion in Environmental Sustainability**, v. 4, n. 5, p. 539–544, nov. 2012. ISSN 18773435.

SOTHE, Camile *et al.* Spatially continuous mapping of forest canopy height in Canada by combining GEDI and icesat-2 with PALSAR and sentinel. **Remote Sensing**, v. 14, n. 20, p. 5158, out. 2022. ISSN 2072-4292.

SOUZA, Bartolomeu Israel de; QUEIROZ, Rubens Teixeira; CARDOSO, Eini Celly Moraes. Degradação e riscos à desertificação no Alto Curso do rio Paraíba - PB/Brasil. **Revista da Anpege**, v. 11, n. 16, p. 201–222, 2015. ISSN 1679768X.

SOUZA, Carlos M. *et al.* 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, p. 2735, ago. 2020. ISSN 2072-4292.

SOUZA, José João Lelis Leal de *et al.* Organic carbon rich-soils in the brazilian semiarid region and paleoenvironmental implications. **CATENA**, v. 212, p. 106101, maio 2022. ISSN 03418162.

STAFF, Soil Survey. **Keys to Soil Taxonomy**. 12. ed. Washington, DC: USDA-Natural Resources Conservation Service, 2014.

STRAWN, Daniel G.; BOHN, Hinrich L.; O'CONNOR, George A. **Soil Chemistry**. 5. ed. Hoboken, NJ: John Wiley & Sons, 2020.

SUN, Pengcheng *et al.* Remote sensing and modeling fusion for investigating the ecosystem water-carbon coupling processes. **Science of The Total Environment**, v. 697, p. 134064, dez. 2019. ISSN 00489697.

SZATMÁRI, Gábor; PÁSZTOR, László. Comparison of various uncertainty modelling approaches based on geostatistics and machine learning algorithms. **Geoderma**, v. 337, p. 1329–1340, mar. 2019. ISSN 00167061.

Taghizadeh-Mehrjardi, Ruhollah *et al.* Improving the spatial prediction of soil organic carbon content in two contrasting climatic regions by stacking machine learning models and rescanning covariate space. **Remote Sensing**, v. 12, n. 7, p. 1095, mar. 2020. ISSN 2072-4292.

TAJIK, Samaneh; AYOUBI, Shamsollah; ZERAATPISHEH, Mojtaba. Digital mapping of soil organic carbon using ensemble learning model in mollisols of hyrcanian forests, northern iran. **Geoderma Regional**, v. 20, p. e00256, mar. 2020. ISSN 23520094.

TOTSCHKE, Kai U. *et al.* Biogeochemical interfaces in soil: The interdisciplinary challenge for soil science. **Journal of Plant Nutrition and Soil Science**, v. 173, n. 1, p. 88–99, fev. 2010. ISSN 1436-8730, 1522-2624.

TUCKER, C.J. *et al.* Monitoring corn and soybean crop development with hand-held radiometer spectral data. **Remote Sensing of Environment**, v. 8, n. 3, p. 237–248, ago. 1979. ISSN 00344257.

TUCKER, Compton J. Red and photographic infrared linear combinations for monitoring vegetation. **Remote Sensing of Environment**, v. 8, n. 2, p. 127–150, 1979.

VELLOSO, Agnes L.; SAMPAIO, Everardo V. S. B.; PAREYN, Frans G. C. **Ecorregiões: propostas para o bioma Caatinga**. Recife: Associação Plantas do Nordeste; Instituto de Conservação Ambiental The Nature Conservancy do Brasil, 2002.

VINCINI, M.; FRAZZI, E.; D’ALESSIO, P. A broad-band leaf chlorophyll vegetation index at the canopy scale. **Precision Agriculture**, v. 9, n. 5, p. 303–319, out. 2008. ISSN 1385-2256, 1573-1618.

VOGT, Daniel John; TILLEY, Joel P.; EDMONDS, Robert L. **Soil and Plant Analysis for Forest Ecosystem Characterization**. [S.l.]: DE GRUYTER, 2015. ISBN 978-3-11-029029-5.

WADOUX, Alexandre M.J.-C. *et al.* Ten challenges for the future of pedometrics. **Geoderma**, v. 401, p. 115155, nov. 2021. ISSN 00167061.

WADOUX, Alexandre M.J.-C.; MOLNAR, Christoph. Beyond prediction: Methods for interpreting complex models of soil variation. **Geoderma**, v. 422, p. 115953, set. 2022. ISSN 00167061.

WADOUX, Alexandre M. J.-C.; SABY, Nicolas P. A.; MARTIN, Manuel P. Shapley values reveal the drivers of soil organic carbon stock prediction. **SOIL**, v. 9, n. 1, p. 21–38, jan. 2023. ISSN 2199-398X.

WALDELAND, Anders U.; TRIER, Øivind Due; SALBERG, Arnt-Børre. Forest mapping and monitoring in africa using sentinel-2 data and deep learning. **International Journal of Applied Earth Observation and Geoinformation**, v. 111, p. 102840, jul. 2022. ISSN 15698432.

WAN, Siang; MOU, Xiaojie; LIU, Xingtu. Effects of reclamation on soil carbon and nitrogen in coastal wetlands of liaohhe river delta, china. **Chinese Geographical Science**, v. 28, n. 3, p. 443–455, jun. 2018. ISSN 1002-0063, 1993-064X.

WANG, Bin *et al.* Estimating soil organic carbon stocks using different modelling techniques in the semi-arid rangelands of eastern australia. **Ecological Indicators**, v. 88, p. 425–438, maio 2018. ISSN 1470160X.

WANG, Jingxin. **Forest and Biomass Harvest and Logistics**. Cham: Springer International Publishing, 2022. ISBN 978-3-031-12945-2 978-3-031-12946-9.

WANG, Kefeng *et al.* Changes in soil organic carbon and microbial carbon storage projected during the 21st century using TRIPLEX-MICROBE. **Ecological Indicators**, v. 98, p. 80–87, mar. 2019. ISSN 1470160X.

WANG, Miaomiao *et al.* Soil nutrients and stoichiometric ratios as affected by land use and lithology at county scale in a karst area, southwest china. **Science of The Total Environment**, v. 619–620, p. 1299–1307, abr. 2018. ISSN 00489697.

_____. Influencing factors on soil nutrients at different scales in a karst area. **CATENA**, v. 175, p. 411–420, abr. 2019. ISSN 03418162.

WARRICK, A. W. Spatial variability. In: _____. **Environmental Soil Physics**. San Diego, CA: Academic Press, 1998. p. 655–675.

WATT, Jeremy; BORHANI, Reza; KATSAGGELOS, Aggelos. **Machine Learning Refined: Foundations, Algorithms, and Applications**. 2. ed. Cambridge: Cambridge University Press, 2020. ISBN 978-1-108-69093-5 978-1-108-48072-7.

WILDING, Larry P.; LIN, Henry. Advancing the frontiers of soil science towards a geoscience. **Geoderma**, v. 131, n. 3-4, p. 257–274, abr. 2006. ISSN 00167061.

WU, Weicheng. The Generalized Difference Vegetation Index (GDVI) for Dryland Characterization. **Remote Sensing**, v. 6, n. 2, p. 1211–1233, jan. 2014. ISSN 2072-4292.

XI, Zhilong *et al.* Forest canopy height mapping by synergizing icesat-2, sentinel-1, sentinel-2 and topographic information based on machine learning methods. **Remote Sensing**, v. 14, n. 2, p. 364, jan. 2022. ISSN 2072-4292.

YANG, Wei *et al.* A semi-analytical snow-free vegetation index for improving estimation of plant phenology in tundra and grassland ecosystems. **Remote Sensing of Environment**, v. 228, p. 31–44, jul. 2019. ISSN 00344257.

ZERAATPISHEH, Mojtaba *et al.* Digital mapping of soil properties using multiple machine learning in a semi-arid region, Central Iran. **Geoderma**, v. 338, p. 445–452, mar. 2019. ISSN 00167061.

ZHANG, Gan-Lin *et al.* (Ed.). **Digital Soil Mapping across Paradigms, Scales and Boundaries**. Singapore: Springer Singapore, 2016. (Springer Environmental Science and Engineering). ISBN 978-981-10-0414-8 978-981-10-0415-5.

ZHANG, Nan *et al.* Forest height mapping using feature selection and machine learning by integrating multi-source satellite data in baoding city, north china. **Remote Sensing**, v. 14, n. 18, p. 4434, set. 2022. ISSN 2072-4292.

ZHANG, Yue *et al.* Mapping stocks of soil total nitrogen using remote sensing data: A comparison of random forest models with different predictors. **Computers and Electronics in Agriculture**, v. 160, p. 23–30, maio 2019. ISSN 01681699.

ZHOU, Tao *et al.* High-resolution digital mapping of soil organic carbon and soil total nitrogen using DEM derivatives, sentinel-1 and sentinel-2 data based on machine learning algorithms. **Science of The Total Environment**, v. 729, p. 138244, ago. 2020. ISSN 00489697.

_____. Mapping soil organic carbon content using multi-source remote sensing variables in the heihe river basin in china. **Ecological Indicators**, v. 114, p. 106288, jul. 2020. ISSN 1470160X.

ZHOU, Xiaocheng *et al.* Improving GEDI forest canopy height products by considering the stand age factor derived from time-series remote sensing images: A case study in fujian, china. **Remote Sensing**, v. 15, n. 2, p. 467, jan. 2023. ISSN 2072-4292.

ZHU, Meng *et al.* Soil organic carbon in semiarid alpine regions: The spatial distribution, stock estimation, and environmental controls. **Journal of Soils and Sediments**, v. 19, n. 10, p. 3427–3441, out. 2019. ISSN 1439-0108, 1614-7480.

ZOMER, Robert J.; XU, Jianchu; TRABUCCO, Antonio. Version 3 of the global aridity index and potential evapotranspiration database. **Scientific Data**, v. 9, n. 1, p. 409, jul. 2022. ISSN 2052-4463.