



UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA AGRÍCOLA

RENATO LÓPEZ NORONHA

**MODELOS PREDITIVOS DE ESTOQUE DE CARBONO DO
SOLO EM TERRA PRETA ARQUEOLÓGICA – NATURAL E
TRANSFORMADA**

CAMPINAS

2019

RENATO LÓPEZ NORONHA

**MODELOS PREDITIVOS DE ESTOQUE DE
CARBONO DO SOLO EM TERRA PRETA ARQUEOLÓGICA
– NATURAL E TRANSFORMADA**

Dissertação apresentada à Faculdade de Engenharia Agrícola da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Agrícola, na Área de Água e Solos.

Orientador: Prof. Dr. Zigomar Menezes de Souza

Coorientador: Prof. Dr. Stanley Robson de Medeiros Oliveira

ESTE EXEMPLAR CORRESPONDE Á VERSÃO
FINAL DA DISSERTAÇÃO DEFENDIDA PELO
ALUNO RENATO LÓPEZ NORONHA, E
ORIENTADO PELO PROF. DR. ZIGOMAR
MENEZES DE SOUZA

CAMPINAS

2019

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Elizangela Aparecida dos Santos Souza - CRB 8/8098

L881m López Noronha, Renato, 1986-
Modelos preditivos de estoque de carbono do solo em terra preta
arqueológica - natural e transformada / Renato López Noronha. – Campinas,
SP : [s.n.], 2019.

Orientador: Zigomar Menezes de Souza.
Coorientador: Stanley Robson de Medeiros Oliveira.
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade
de Engenharia Agrícola.

1. Solos - Manejo. 2. Mineração de dados. 3. Árvores de decisão. 4.
Aprendizado de máquina. 5. Terra preta - Amazonas. I. Souza, Zigomar
Menezes de, 1970-. II. Oliveira, Stanley Robson de Medeiros. III. Universidade
Estadual de Campinas. Faculdade de Engenharia Agrícola. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Predicted models for soils carbon stock on archaeological black
lands - natural and transformed

Palavras-chave em inglês:

Soil - Management

Data mining

Decision trees

Machine learning

Black Earth - Amazonas

Área de concentração: Água e Solo

Titulação: Mestre em Engenharia Agrícola

Banca examinadora:

Zigomar Menezes de Souza [Orientador]

Fabio Cesar da Silva

Edson Eiji Matsura

Data de defesa: 26-02-2019

Programa de Pós-Graduação: Engenharia Agrícola

Identificação e informações acadêmicas do(a) aluno(a)

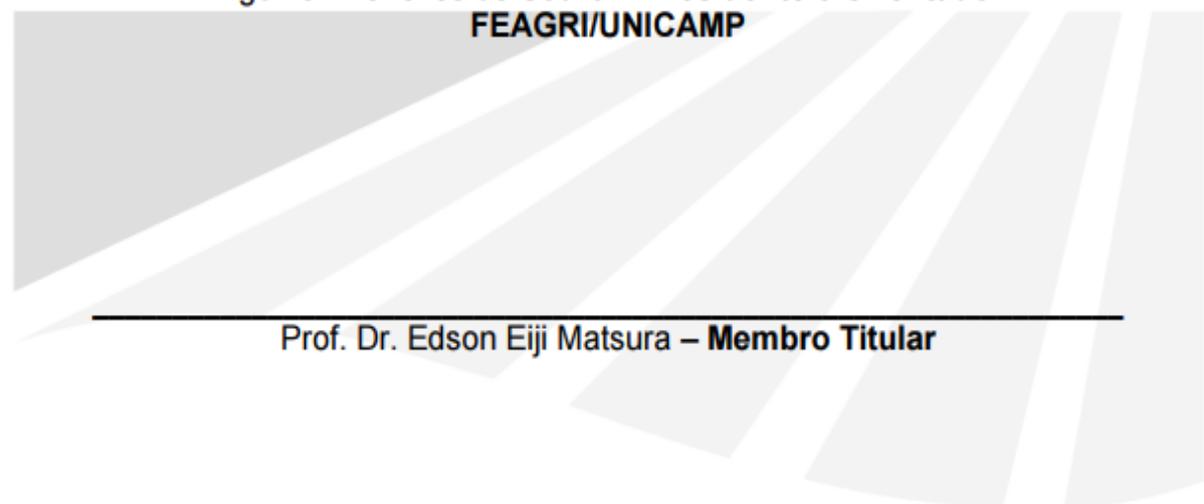
- ORCID do autor: <https://orcid.org/0000-0001-7967-4291>

- Currículo Lattes do autor: <http://lattes.cnpq.br/0762196214755973>

Este exemplar corresponde à redação final da **Dissertação de Mestrado** defendida por Renato López Noronha, aprovada pela Comissão Julgadora em 26 de Fevereiro de 2019, na Faculdade de Engenharia Agrícola da Universidade Estadual de Campinas.

FEAGRI

Zigomar Menezes de Souza – **Presidente e Orientador**
FEAGRI/UNICAMP



Prof. Dr. Edson Eiji Matsura – **Membro Titular**

Faculdade de
Engenharia Agrícola
Unicamp

Dr. Fábio Cesar da Silva – **Membro Titular**

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no processo de vida acadêmica do discente.

DEDICATÓRIA

Dedico a minha esposa Cris e aos nossos filhos Lucca Inácio e Vagner Luan que durante todos estes anos foram a força que me manteve firme.

AGRADECIMENTOS

Primeiramente a Deus, que sempre me acompanhou em todas as jornadas de minha vida. Que continue me dando luz para seguir os caminhos certos.

A minha esposa, que semeou esse sonho no meu coração e foi a fortaleza durante os momentos mais difíceis, por todo amor e apoio incondicional.

Aos meus filhos Lucca Inácio e Vagner Luan, que com o brilho dos seus olhos alegam meus dias.

Agradecer aos meus pais, Manuel e Nancy, que sempre fizeram tudo o que podiam (e o que não também) para eu seguir em frente.

Aos meus irmãos, Debra, Brenda e Leonardo, os quais sempre levo nos meus pensamentos.

Agradeço também aos meus sogros: Oscar Elias Anaya e Silvia Fernandez, e ao meu cunhado: Oscar Gonzalo Anaya, todos eles que são minha outra família.

Ao meu orientador Prof. Dr. Zigomar Menezes de Souza pela oportunidade, compromisso e pela paciência.

Ao meu coorientador Dr. Stanley Robson de Medeiros Oliveira pelo apoio ímpar.

À equipe de pesquisa da Universidade Federal do Amazonas – UFAM, liderado pelo Prof. Milton César Costa Campos e com apoio do Prof. Marcelo Dayron Rodrigues Soares, pelo fornecimento do banco de dados produto da coleta do material.

À Unicamp e FEAGRI pela infraestrutura fornecida, com agradecimento especial à Ritinha, pelo exemplo de profissionalismo e bom coração.

A minha colega Camila Viana Vieira Farhate pelo acompanhamento e orientações, e também aos demais colegas e funcionários do laboratório de solos – FEAGRI.

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

A FAPESP, pelo financiamento do projeto (Projeto: 2015/24280-8).

Este trabalho é resultado de esforço que não seria possível sem a ajuda de tantas pessoas que não caberia em uma folha de papel, mas sou agradecido por toda ajuda que recebi.

Meu sincero muito obrigado!

EPÍGRAFE

*“Sucesso não é o final, falhar não é fatal:
é a coragem para continuar que conta”.*

Winston Churchill

RESUMO

Na Amazônia, os solos conhecidos como Terra Preta Arqueológica apresentam horizonte A antrópico e estão associados à prolongada ocupação humana, realizada por sociedades indígenas do período pré-colombiano, nos quais possuem atributos químicos e físicos são superiores em relação aos de outros solos da Amazônia, configurando como grande reservatório de carbono orgânico. Entretanto, a conversão desses ecossistemas naturais em ambientes de cultivo proporciona alterações na dinâmica de carbono no solo, que frequentemente, provocam o declínio no teor de carbono orgânico do solo. Dessa forma, essa pesquisa tem como objetivo utilizar técnicas de mineração de dados para gerar modelos preditivos sobre o efeito do uso do solo no estoque de carbono em áreas de Terra Preta Arqueológica, naturais e transformadas. O experimento foi realizado nos municípios de Novo Aripuanã e Apuí, estado de Amazonas, Brasil em áreas com floresta natural, pastagem, feijão, café e cacau, nas profundidades de 0,00-0,05, 0,05-0,10 e 0,10-0,20 m. Após a etapa de coleta dos dados e análise laboratorial foram aplicadas duas técnicas de mineração de dados para modelagem preditiva, sendo a indução por meio de árvores de decisão e *Random Forest*. Ao final da pesquisa, os resultados mostraram que os atributos, uso do solo e areia, foram as variáveis mais importantes para prever o estoque de carbono em Terra Preta Arqueológica sob diferentes usos. Por meio da técnica de indução de árvore de decisão foi gerado um modelo preditivo de estoque de carbono do solo em Terra Preta Arqueológica, com alta taxa de precisão: 83%, 74% e 81%, usando sete (07), sete (07) e quatro (04) regras, nas profundidades estudadas, 0,00-0,05 m, 0,05-0,10 m e 0,10-0,20 m, respectivamente. O algoritmo *Random Forest* indicou que o cálcio, a densidade do solo e uso do solo foram as variáveis mais importantes que afetaram o estoque de carbono em Terra Preta Arqueológica sob diferentes usos e, técnica *Random Forest* originou modelos com bom ajuste ($R^2 = 0,71, 0,71$ e $0,77$) para valores preditos e observados.

Palavras-Chave: Solos - manejo, técnicas de mineração de dados, árvore de decisão, aprendizado de máquina, Terra preta - Amazonas.

ABSTRACT

In the Amazon, soils known as Archaeological Black Earth have anthropic A-horizon and are associated with prolonged human occupation, performed by indigenous societies of the pre-Columbian period, in which they have chemical and physical attributes are superior to those of other Amazonian soils, becoming as a large reservoir of organic carbon. However, the conversion of these natural ecosystems into cultivated environments provides changes in soil carbon dynamics, which often lead to a decline in soil organic carbon content. Thus, this research had as objective to elaborate predictive models of soil carbon stock in Archaeological black earths transformed by action of agricultural crops, using techniques of data mining. The experiment was conducted in the counties of Novo Aripuanã and Apuí, state of Amazonas - Brazil, in areas of coffee, cocoa, beans, pasture and forest. The samples were collected at 0.00-0.05, 0.05-0.10 and 0.10-0.20 m. of depth. . After the data collection and laboratory analysis, two data mining techniques were applied for predictive modeling: Decision trees and Random Forest. At the end of the research, a predictive model of soil carbon stock was generated through the decision tree technique, with a high accuracy: 83%, 74% and 81%, using seven (07), seven (07) and four (04) rules in the studied depths, 0.00-0.05 m, 0.05-0.10 m and 0.10-0.20 m, respectively. The results showed that the variables: 'land uses' and 'sand' were the most important variables to predict soil carbon stock in Archaeological black earth under different uses. The Random Forest algorithm builded models with good fit ($R^2 = 0.71, 0.71$ and 0.77) for predicted and observed values, and indicated calcium, soil density and land uses were the most important variables affected soil carbon stock in Archaeological black earth under different uses.

Key words: Soil - Management, Data mining, Decision tree, Machine learning, Black Earth - Amazonas.

LISTA DE ILUSTRAÇÕES

Figura 1. Processo de descoberta de conhecimento em bancos de dados.	22
Figura 2. Localização da área de estudo com Terra Preta Arqueológica sob diferentes usos, nos municípios de Novo Aripuanã e Apuí, Amazonas, Brasil.....	26
Figura 3. Árvore de decisão resultante da utilização do método CFS para seleção de atributos e o algoritmo J48 para indução de árvore de decisão, na profundidade de 0,00-0,05 m.	35
Figura 4. Árvore de decisão resultante da utilização do método CFS para seleção de atributos e o algoritmo J48 para indução de árvore de decisão, na profundidade de 0,05-0,10 m.	36
Figura 5. Árvore de decisão resultante da utilização do método CFS para seleção de atributos e o algoritmo J48 para indução de árvore de decisão, na profundidade de 0,10-0,20 m.	37
Figura 6. Localização da área de estudo com Terra Preta Arqueológica sob diferentes usos, nos municípios de Novo Aripuanã e Apuí, Amazonas, Brasil.....	45
Figura 7. Diagrama conceitual do algoritmo Random Forest.	48
Figura 8 Ordem de importância dos atributos selecionados para uso no modelo gerado pelo algoritmo de Random Forest para explicar o estoque de carbono no solo em Terra Preta Arqueológica, (8A) 0,00-0,05 m, (8B) 0,05-0,10 m e (8C) 0,10-0,20 m. Ds = densidade do solo (kg dm^{-3}); Macro = macroporosidade ($\text{m}^3 \text{ m}^{-3}$); Micro = microporosidade ($\text{m}^3 \text{ m}^{-3}$); VTP = volume total de poros ($\text{m}^3 \text{ m}^{-3}$); Umidade = umidade volumétrica ($\text{kg}^1 \text{ kg}^{-1}$); Acidez = acidez potencial ($\text{mmol}_c \text{ dm}^{-3}$); Ca = cálcio ($\text{mmol}_c \text{ dm}^{-3}$); Mg = magnésio ($\text{mmol}_c \text{ dm}^{-3}$); K = potássio ($\text{mmol}_c \text{ dm}^{-3}$); P = fósforo (mg dm^{-3}); Al^{3+} = alumínio ($\text{mmol}_c \text{ dm}^{-3}$).....	49
Figura 9 Validação do modelo de regressão, os eixos x e y representam, respectivamente, os valores observados e previstos do estoque de carbono no solo em Terra Preta Arqueológica, (7A) 0,00-0,05 m, (7B) 0,05-0,10 m e (7C) 0,10-0,20 m. Corr = correlação; RMSE = Root-Mean-Square Error / distância média quadrática mínima; R ² : R quadrado / coeficiente de determinação.....	50

LISTA DE TABELAS

Tabela 1. Atributos físicos e químicos de Terras Pretas Arqueológicas transformadas e naturais, que compõe o banco de dados.	27
Tabela 2. Estoque de carbono do solo do solo (Mg ha^{-1}) de acordo com as classes (baixa e alta) e seus respectivos limites, nas profundidades de 0,00-0,05, 0,05-0,10 e 0,10-0,20 m.....	27
Tabela 3. Matriz de erros para análise da acurácia, de um classificador binário (árvore de decisão).....	29
Tabela 4. Classificação da qualidade do desempenho do modelo de árvore de decisão, por meio da estatística Kappa.	29
Tabela 5. Quantidade de Outliers identificados, pelo método de bloxplot, na base de dados em Terras Pretas Arqueológicas.	30
Tabela 6. Estoque de carbono do solo (Mg ha^{-1}) em Terra Preta Arqueológica sob diferentes sistemas de uso (valores médios de 88 amostras por cada sistema de uso).....	31
Tabela 7. Desempenho do classificador J48, na geração do modelo preditivo de estoque de carbono em Terra Preta Arqueológica, sob diferentes sistemas de uso.....	32
Tabela 8. Relação entre o número mínimo de objetos (minNumObj), taxa de precisão, estatística Kappa e número de regras nas três profundidades estudadas.....	34
Tabela 9. Atributos físicos e químicos de Terras Pretas Arqueológicas naturais e transformadas, que compõe o banco de dados.	47

SUMÁRIO

CAPÍTULO 1: Introdução e revisão bibliográfica	13
1. Introdução	13
2. Hipótese do estudo	15
3. Objetivos	16
3.1. Geral	16
3.2. Específicos	16
3.3. Organização da dissertação	16
4. Revisão bibliográfica	17
4.1. Aspectos gerais da Terra Preta Arqueológica	17
4.2. Estoque de carbono no solo em Terra Preta Arqueológica	18
4.3. Mineração de dados	20
CAPÍTULO 2: Classificação do estoque de carbono do solo em Terra Preta Arqueológica sob diferentes usos	23
5. Introdução	24
6. Material e métodos	25
6.1. Descrição da área experimental	25
6.2. Descrição do banco de dados	26
6.3. Metodologia de indução de Árvore de decisão	27
6.4. Seleção de atributos	28
6.5. Indução e validação do modelo de classificação	28
7. Resultados	30
8. Discussão	37
9. Conclusões	41
CAPÍTULO 3: <i>Random Forest</i> para predição do estoque de carbono em Terra Preta Arqueológica sob diferentes usos	42
10. Introdução	43
11. Material e Métodos	44
11.1. Descrição da área experimental	44
11.2. Coleta de solo para determinação dos atributos físicos e químicos do solo	45
11.3. Atributos físicos do solo	46
11.4. Atributos químicos do solo	46
11.5. Descrição do banco de dados	46
11.6. Indução de <i>Random Forest</i>	47
11.7. Validação do modelo de regressão	48
12. Resultados	48
13. Discussões	50
14. Conclusões	52
15. Considerações finais	53
16. Referências bibliográficas	54

CAPÍTULO 1: Introdução e revisão bibliográfica

1. Introdução

O solo é uma coleção de corpos naturais constituídos por partes sólidas, líquidas e gasosas, tridimensionais, dinâmicos, formados por materiais minerais e orgânicos que ocupam a maior parte do manto superficial das extensões continentais do nosso planeta, contém matéria viva e pode ser vegetado na natureza onde ocorre e, eventualmente, tem sido modificado por interferências antrópicas (SANTOS *et al.*, 2018). Na Amazônia, os solos que apresentam horizonte A antrópico estão associados à prolongada ocupação humana pré-histórica, realizada por sociedades indígenas do período pré-colombiano (SANTOS *et al.*, 2018), as quais formaram extensos depósitos de resíduos, ou simplesmente rejeitos, que alteraram significativamente os atributos dos solos (AQUINO *et al.*, 2016).

A formação das Terras Pretas Arqueológicas (TPA) ocorre em três etapas: formação do carvão - refere-se à (1) formação do carbono pirogênico, altamente estável que contribui com a fertilidade do solo; (2) incorporação de nutrientes - nessa etapa os nutrientes são incorporados ao solo por diferentes fontes: excrementos humanos e animais, cinzas, resíduos de combustão incompleta e carvão, biomassa de plantas aquáticas e terrestres; e (3) ação dos microrganismos - esses são responsáveis pela ciclagem de nutrientes, agindo tanto na decomposição da matéria orgânica como na imobilização de nutrientes do solo (GLASER, 2007).

Pesquisas tem demonstrado (CAMPOS *et al.*, 2012; MARQUES *et al.*, 2016; CUNHA *et al.*, 2017; SOAREZ *et al.*, 2018), que os atributos químicos e físicos das TPA são superiores em relação aos de outros solos da Amazônia, figurando como grande reservatório de carbono orgânico. Entretanto, com a modificação da cobertura da vegetação nativa por meio das ações antrópicas atuais, é possível que haja perdas massivas de carbono do solo o que leva a degradação do solo (LA SCALA JÚNIOR *et al.*, 2012).

A conversão dos ecossistemas naturais em ambientes de cultivo vem alterando a dinâmica de carbono no solo, especialmente na Amazônia, essas modificações frequentemente provocam o declínio no teor de carbono orgânico do solo, em função da redução do aporte, perdas por erosão e por decomposição da matéria orgânica (HICKMANN; COSTA, 2012).

Em adição, no cenário atual, a investigação do mecanismo de sequestro de carbono destes solos apresenta uma posição interessante frente a comunidade científica,

uma vez que, o biocarvão, presente nas TPA (MANGRICH *et al.*, 2011), é geralmente bastante estável (TRAZZI *et al.*, 2018) e, por conseguinte, fornece um alto potencial no sequestro de carbono (IPPOLITO, *et al.*, 2012). Este aspecto combinado com a alta fertilidade (MANGRICH *et al.*, 2011) é exclusivo destes solos. Assim, as TPA podem atuar como um modelo de solo, para entender melhor a fertilidade a longo prazo e o potencial de sequestro de carbono (MAIA, 2010), em resposta as práticas de gestão antropogênicas.

Com tudo, o conhecimento do comportamento do estoque de carbono em áreas de TPA é fundamental e, utilizar técnicas ainda não exploradas se apresenta como uma alternativa atrativa para aprofundar o conhecimento do melhor manejo desse solo. Uma alternativa é o uso de técnicas de mineração de dados, que têm sido empregadas para extração de padrões e tendências em bancos de dados. O constante crescimento do volume de dados nas mais diversas áreas do conhecimento tem impulsionado e exigido da comunidade acadêmica o desenvolvimento de técnicas que possibilitem a extração de informações desses dados para apoiar a tomada de decisão.

Neste contexto, a Descoberta do Conhecimento em Bases de Dados (*Knowledge Discovery in Databases - KDD*) se apresenta como uma tecnologia detentora de ferramentas eficientes para transformar informações contidas em grandes volumes de dados em conhecimento útil. A descoberta de conhecimento em bases de dados é definida como um processo de identificação de padrões válidos, não conhecidos, potencialmente úteis e interpretáveis em dados armazenados (FAYYAD *et al.*, 1996). Uma das etapas do processo de KDD é a aplicação das técnicas de mineração de dados que em muitas vezes é usada como sinônimo do KDD (HAN *et al.*, 2011). Há pesquisas em solo com essa abordagem tem utilizado diferentes técnicas de mineração de dados: *árvores de decisão* e *Random Forest* (LIMA *et al.*, 2017; FARHATE *et al.*, 2018).

Dentre as técnicas de mineração de dados, o método com uso de árvores de decisão tem várias vantagens (QUINLAN, 1996): (a) são fáceis de entender e interpretar; (b) pessoas conseguem entender modelos em árvore depois de uma simples explicação, ajudando na documentação e entendimento do analista com o usuário; (c) está apto a lidar tanto com dados nominais (numéricos) ou com dados categóricos (escala não numérica); (d) é um modelo "caixa-branca"; (e) é robusto, e tem bom desempenho com grandes quantidades de informação em pouco tempo, porém, entre as desvantagens pode-se destacar que o estudo com variáveis categóricas com vários níveis, o ganho de informação gerado pela árvore de decisão tem uma tendência de favorecer os atributos com mais

níveis, uma outra desvantagem é que a árvore de decisão pode ser instável, pois uma pequena mudança nos dados pode modificar amplamente a estrutura da árvore.

Por um outro lado, Rodriguez-Galiano *et al.* (2012) relatam as principais vantagens para o uso do algoritmo *Random Forest*, apresentado por Breiman (2001): (a) tem natureza não-paramétrica; (b) é um dos algoritmos de aprendizado de máquina com alta taxa de acurácia; (c) pode lidar com milhares de variáveis de entrada sem a necessidade prévia de redução de dimensionalidade; (d) identifica quais são as variáveis mais importantes na classificação; (e) é robusto na presença de ruído e de variáveis sem importância; (f) pode ser altamente flexível para realizar vários tipos de análise de dados, incluindo regressão, classificação e aprendizagem não supervisionada. Quanto às desvantagens podem-se citar que a diferencia das árvores de decisão, neste caso o controle sobre as interações dos atributos é mínimo o que aumenta a complexidade do resultado final.

Desta forma é grande o interesse de desenvolver trabalhos científicos que estudem o comportamento de carbono em Terra Preta Arqueológica, gerando informações que possam ser aproveitados de uma forma mais viável tanto para fins de produção agrícola como ambiental. Diante do exposto, o objetivo do presente estudo foi elaborar modelos preditivos de estoque de carbono no solo em Terras Pretas Arqueológicas naturais e transformadas pela ação de cultivos agrícolas, utilizando técnicas de mineração de dados, capazes de auxiliar o processo de tomada de decisão sobre possíveis práticas de manejo com potencial para conservar ou aumentar o estoque de carbono do solo em Terra Preta Arqueológica.

2. Hipótese do estudo

Propõe-se nessa pesquisa a seguinte hipótese de trabalho: a utilização de técnicas de mineração de dados aplicadas a um conjunto de dados composto por atributos físicos e químicos do solo, utilizando como atributo resposta o estoque de carbono em Terra Preta Arqueológica, fornecerá modelos preditivos, de classificação e regressão, com elevada precisão e acurácia, capazes de auxiliar o processo de tomada de decisão para conservar ou aumentar o estoque de carbono do solo.

3. Objetivos

3.1. Geral

Elaborar modelos preditivos de estoque de carbono no solo em Terras Pretas Arqueológicas naturais e transformadas pela ação de cultivos agrícolas, utilizando técnicas de mineração de dados.

3.2. Específicos

Para o cumprimento do objetivo geral, são apresentados a seguir os objetivos específicos:

- i. Avaliar a eficiência de diferentes métodos de seleção de atributos, por meio de algoritmos de classificação e regressão na predição do estoque de carbono em Terra Preta Arqueológica – naturais e transformadas.
- ii. Identificar por meio de seleção, quais são os atributos físicos e químicos do solo mais importantes para predição do estoque de carbono em Terra Preta Arqueológica – naturais e transformadas.

3.3. Organização da dissertação

Este estudo foi conduzido de forma que o primeiro capítulo foi desenvolvido com o propósito de realizar uma visão ampla a respeito do tema de interesse, por meio da introdução, definição da hipótese, do objetivo geral e dos objetivos específicos. Seguindo por uma revisão bibliográfica com os seguintes tópicos: aspectos gerais da Terra Preta Arqueológica, estoque de carbono no solo, mineração de dados e a descrição das técnicas usadas nesta dissertação.

É importante salientar que o banco de dados analisado nesta pesquisa, foi produto de trabalho de campo efetuado pela equipe de pesquisa coordenada pelo Prof. Dr. Milton César Costa Campos e o Prof. Dr. Marcelo Dayron Rodrigues Soares, que vem desenvolvendo diversas pesquisas em Terra Preta Arqueológica na região Amazônica.

Os resultados da pesquisa são apresentados no segundo e terceiro capítulo, em estrutura de artigos científicos, sendo que o segundo capítulo corresponde a predição de estoque de carbono usando a classificação por meio da indução de árvore de decisão e o terceiro corresponde a abordagem *Random Forest*, que nada mais é que uma predição usando um método de regressão. Finalmente, as considerações gerais durante o desenvolvimento desta pesquisa são apresentadas no último capítulo.

4. Revisão bibliográfica

4.1. Aspectos gerais da Terra Preta Arqueológica

No passado pensava-se que antes da chegada dos europeus, a Amazônia era uma enorme floresta úmida sem presença humana, povoado apenas por pequenos grupos de caçadores nômades ou coletores afastados uns dos outros por centenas de quilômetros de vegetação densa. Mas na atualidade se conhece que, por séculos (de 2.500 a 3.000 anos atrás), essas civilizações prosperaram nesta região da América, sendo que para uma civilização nascer e subsistir durante séculos, ela deve obrigatoriamente ter uma agricultura poderosa, porém, qual foi o segredo para essas civilizações prosperar no meio desse imenso bioma verde e úmido?

Os solos da região Amazônica são representados em sua maioria por Latossolos e Argissolos, que são referenciados como altamente intemperizados, de acidez elevada, de baixa fertilidade natural e baixa capacidade de troca de cátions, além de serem facilmente alterados quimicamente com a substituição da floresta primária por cultivos intensivos (CUNHA *et al.*, 2007; NOVOTNY *et al.*, 2009; GALVÃO *et al.*, 2015).

Em contraste aos solos de terra pouco fértil, existem as Terras Pretas da Amazônia, que também são denominadas de Terra Preta Arqueológica ou Antropogênica (TPA) ou Terra Preta de Índio (TPI), obtido graças à aplicação de uma tecnologia milenar nativa dos povos indígenas da floresta amazônica. As TPA apresentam alta fertilidade natural e com propriedades marcantes como coloração escura e a presença de fragmentos de cerâmica e/ou líticos e artefatos indígenas incorporados à matriz dos horizontes superficiais (KÄMPF; KERN, 2005; GLASER; BIRK, 2012).

Sua formação é, ainda, motivo de controvérsias entre os pesquisadores, porém, Sombroek *et al.* (2003) formalizaram um modelo abrangente e simples sobre os fatores que afetariam sua formação. O modelo do aterro da cozinha sustenta que TPA seriam o resultado da decomposição de restos de comida, lixo doméstico, resíduos de combustão e excrementos humanos concentrados em áreas de eliminação de resíduos. Embora aceita, esta relação ainda gera uma série de dúvidas, entre as quais, o questionamento sobre se a sua formação foi intencional ou como mera consequência da ocupação humana (NEVES *et al.*, 2003; ARROYO-KALIN, 2017).

As TPA passaram por um longo processo de formação com influência externa do ambiente como clima e vegetação, atuando como agentes controladores do meio. Segundo Foster *et al.* (2003) e Silva *et al.* (2016) as mudanças na cobertura vegetal e as transformações do ambiente alteram a dinâmica natural do carbono no solo. As TPA

ocorrem praticamente em toda a região amazônica e, as informações a respeito do tamanho dos sítios de existência variam de menos de um hectare, em terraços elevados, paralelos aos rios, a 400 hectares (SOMBROEK *et al.*, 2003).

Essas manchas ocorrem sobre várias classes de solos, pois Santos *et al.* (2012) verificaram que os horizontes antropogênicos variaram entre as texturas franco-arenosa, franco-siltosa e franco-argilosa, modificando para os horizontes diagnósticos subsuperficiais que foram classificados como franco-siltosa, franca, argilosa e muito argilosa. Barros *et al.* (2016) relata que apesar das diferentes classes texturais dos solos de TPA, todos apresentaram uma boa estruturação (porosidade total > 50%).

Quanto aos atributos físicos, estes solos apresentam estrutura granular (SANTOS *et al.*, 2013), relacionada a ação do carbono pirogênico sobre a formação e estabilização dos agregados (CUNHA *et al.*, 2007). São considerados bem drenados, com porosidade total elevada e boa aeração (CAMPOS *et al.*, 2011; SANTOS *et al.*, 2013), sendo a baixa densidade do solo outra característica comum em TPA (TEIXEIRA; MARTINS, 2003; CAMPOS *et al.*, 2012).

Em relação aos atributos químicos estes solos: em comparação com as áreas vizinhas imediatas, apresentam um pH mais básico, maior capacidade de troca de cátions, alta saturação por bases e altas concentrações de cálcio, fósforo, magnésio, potássio, bário, cobre, manganês, estrôncio, zinco e carbono, entre outros (GLASER; BIRK, 2012).

Atualmente as TPA vêm ganhando destaque, devido à possibilidade de os cientistas descobrirem e replicarem os processos que lhe deram origem, para que no futuro estes processos sejam utilizados como possíveis recuperadores de solos degradados no mundo. Apesar de sua importância para o desenvolvimento agrícola da região, estas também são fundamentais para o conhecimento da pré-história da Amazônia e da mudança global do clima atual, por conta de sua grande quantidade de carbono orgânico armazenado, atuando assim como um dreno de dióxido de carbono (CO₂).

4.2. Estoque de carbono no solo em Terra Preta Arqueológica

No contexto mais amplo dos solos tropicais da Amazônia - que em geral perdem seus nutrientes facilmente e, sem proteção conservacionista, são muito suscetíveis à erosão - a fertilidade incomum dessas terras parece sugerir que está diante de um fenômeno natural excepcional. No entanto, essas terras são o resultado da ação centenária dos processos pedogenéticos em superfícies terrestres enriquecidos pelas práticas habitacionais das sociedades pré-colombianas da Amazônia.

As mudanças climáticas devido ao aumento da emissão de gases pelo homem modificam o regime hídrico e a temperatura global o que pode influenciar a produtividade das culturas (CARVALHO *et al.*, 2010). Sendo que, os solos são o maior reservatório terrestre de carbono (SCHARLEMANN *et al.*, 2014), excedendo toda a biomassa por um fator de 3 ou 4, uma vez que o carbono se encontra mais concentrado na parte superficial do solo (LAL, 2002; BERNOUX *et al.*, 2005; NOVOTNY *et al.*, 2012), o que ressalta seu papel central no ciclo global do carbono e nas mudanças climáticas (HESTERBERG; VERGÜTZ, 2017).

O ciclo do carbono começa com a fixação CO₂ atmosférico por meio da fotossíntese pelas plantas e alguns microrganismos (ORJUELA, 2018). Desse modo, o solo apresenta-se como uma solução quando se pensa em poluição por CO₂, pois ele promove o sequestro de carbono, que significa transformar o CO₂ atmosférico em carbono estocado no solo, tornando-se um componente da matéria orgânica (TONON; BORSZOWSKI, 2015).

A matéria orgânica tem influência direta e indiretamente aos atributos do solo (VIANA *et al.*, 2011) e, o seu estudo em agroecossistema brasileiro é base para que se alcance a sustentabilidade agrícola (CUNHA *et al.*, 2011). A matéria orgânica desempenha efeitos positivos sobre a estrutura do solo, pois o aumento do seu conteúdo também potencializa a atividade biológica, contribuindo para uma maior estabilidade dos agregados (ZHU *et al.*, 2016), além de aumentar a capacidade de retenção de cátions do solo, estabilizar a temperatura, manter a umidade, dentre outros benefícios.

Assim, para aumentar a produtividade desses solos é necessário incrementar o conteúdo de carbono. A importância do carbono orgânico do solo tem sido muito relatada na literatura (BHARALI *et al.*, 2016; DAS *et al.*, 2016; OMARI *et al.*, 2016; FORTE *et al.*, 2017). A sua manutenção é vital para a continuidade da produção e sustentabilidade dos sistemas agrícolas a longo prazo e, o manejo da matéria orgânica em condições tropicais exerce um papel chave no aumento do carbono orgânico do solo (DAS *et al.*, 2016).

De acordo com Niero *et al.* (2010) a qualidade do solo pode ser avaliada por meio de atributos relacionados à sua capacidade de fornecer nutrientes às plantas, dar suporte ao crescimento e desenvolvimento de raízes e propiciar adequada estabilidade estrutural para resistir à erosão e reter água para as plantas.

Os solos tropicais caracterizam-se, sobretudo, como solos antigos, com mineralogia composta principalmente por caulinita e gibbsita. Tratam-se de solos

intemperizados, com baixo pH, baixa concentração de óxidos de ferro e de matéria orgânica. Estes solos, em sua maioria, podem apresentar grande plasticidade em relação à coesão, tornando-se adensados quando secos (MOURA *et al.*, 2012). Portanto, estes solos possuem uma baixa fertilidade natural, o que dificulta o seu manejo, sendo que existem estudos que reportam que os estoques de carbono no solo são influenciados pelos sistemas de manejo (BARRETO *et al.*, 2014; MASCARENHAS *et al.*, 2017).

As áreas de TPA apresentam atributos físicos e químicos superiores em relação aos outros tipos de solos da Amazônia (SOARES *et al.*, 2018) e considerando os altos teores de carbono, figura como grande reservatório de matéria orgânica (CAMPOS *et al.*, 2012), pois foi descrito por Novotny *et al.* (2009) que o carbono destes solos apresenta a capacidade de ser até seis vezes mais estáveis do que das áreas adjacentes.

A recalcitrância do carbono pirogênico (principal constituinte da matéria orgânica das TPA) lhe confere uma alta estabilidade, fazendo com que este seja um importante reservatório de carbono estável no solo, podendo desta forma mitigar a emissão de CO₂ para a atmosfera, além de desempenhar importante papel na fertilidade dos solos (CUNHA *et al.*, 2010).

No que diz respeito ao horizonte A antrópico das TPA, muitas questões ainda permanecem sem resposta, principalmente sobre a sua gênese, distribuição e características da sua matéria orgânica (CUNHA *et al.*, 2007), sendo propícia a exploração com novos métodos que permitam entender melhor todo esse fenômeno.

Considerando esse interesse da comunidade científica para entender as TPA, ocorreu um grande crescimento na capacidade de gerar e coletar dados. Nesse contexto é importante salientar que os estudos da formação e comportamento das TPA, colocam essas terras como modelo de solo em resposta as práticas antropogênicas desde o ponto de vista da enorme capacidade de estocar carbono orgânico. Sabendo que os métodos de extração de conhecimento evoluíram, a análise desse cenário com novas técnicas se apresenta como uma excelente oportunidade de contribuição com a comunidade científica.

4.3. Mineração de dados

Os métodos clássicos de análise de dados, que usam modelos estatísticos como normalmente estão acostumados a formular, funcionam adequadamente sob suposições específicas. No entanto, fatores como o avanço tecnológico geraram um enorme crescimento na capacidade de gerar e coletar dados. Essas grandes massas de

dados, possuem características particulares (alto número de variáveis de vários tipos, falta de dados, entre outros) que dificulta a tomada de decisões.

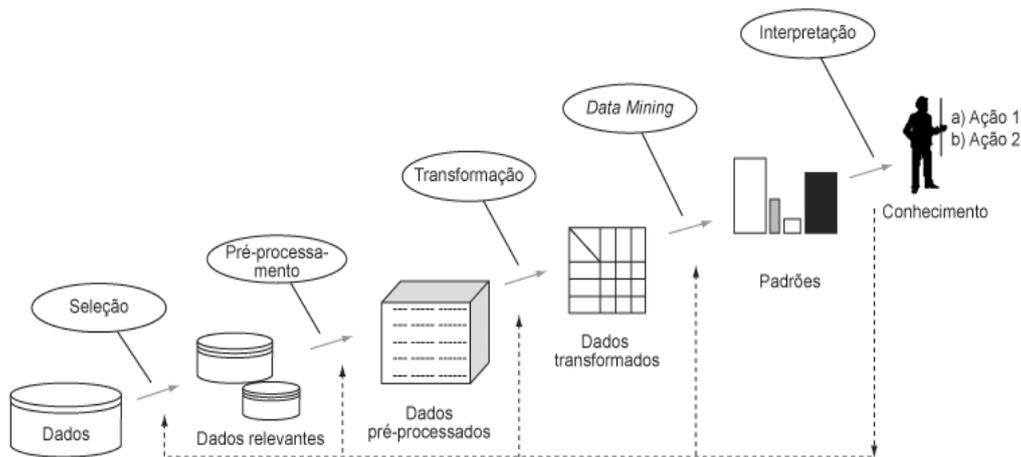
Atualmente, na área da agricultura, há uma grande quantidade de dados que é obtida do solo e das culturas, tal quantidade de dados que vem aumentando aceleradamente nos últimos anos, e as técnicas tradicionais de exploração de dados não são mais adequadas para tratar todo esse volume. De acordo com Witten *et al.* (2011) o montante de dados armazenados nos bancos de dados ao redor do mundo se duplica a cada 20 meses, aumentando assim a distância entre a geração e o entendimento dessas informações.

Nesse sentido, a mineração de dados surge como ferramenta capaz de elucidar os padrões implícitos nessas grandes massas de dados, ela é conhecida como a ciência de extrair conhecimentos úteis de grandes repositórios de dados a partir da aplicação de técnicas da estatística, inteligência artificial, aprendizado de máquina, recuperação da informação, entre outros (HAND, 2007). Conforme a definição de Witten *et al.* (2011), a mineração de dados faz parte do processo de descoberta de conhecimento em banco de dados (*Knowledge Discovery Databases - KDD*), sendo que os termos KDD e Mineração de Dados foram entendidos por muitos pesquisadores como sinônimos até 1995 (LEMOS *et al.*, 2005).

Porém, a mineração de dados é considerada a principal etapa do processo de descoberta de conhecimento em banco de dados (GOLDSCHMIDT; PASSOS, 2005). Estima-se que a mineração de dados represente até 25% do processo de descoberta de conhecimento em banco de dados (SHIMABUKURU, 2004). Na atualidade, são vastas as referências que definem e padronizam as fases da descoberta de conhecimento nas bases de dados, no entanto, todos apresentam basicamente a mesma estrutura (LAROSE, 2005; HAND, 2007). A Figura 1, apresenta as fases do processo de KDD.

O processo de mineração de dados integra técnicas oriundas de diferentes áreas do conhecimento, como tecnologia de bancos de dados, estatística, computação, processamento e visualização de dados, entre outros (HAN *et al.*, 2011). Por exemplo, Berson *et al.* (2000), relata que aproximadamente 5% de todas as relações podem ser encontradas por esses modelos estatísticos e a mineração de dados pode descobrir outras relações anteriormente desconhecidas, ou seja, os 95% restantes.

Figura 1. Processo de descoberta de conhecimento em bancos de dados.



Adaptado de Fayyad *et al.*, 1996.

De acordo com Han *et al.* (2011), as tarefas de mineração de dados podem ser divididas em preditivas, que consistem em se fazer previsões a partir de inferências sobre os dados em questão, ou descritivas, com o foco na caracterização das propriedades dos dados. Nesta pesquisa será abordada duas tarefas preditivas da mineração de dados, a classificação e a regressão.

A classificação é uma função de aprendizado que mapeia dados de entrada, ou conjunto de dados de entrada, em um número finito de classes. O objetivo de um algoritmo de classificação é encontrar alguma correlação entre os atributos e uma classe, de modo que o processo de classificação possa usá-la para prever a classe de um exemplo novo (REZENDE, 2003). A tarefa de regressão é conceitualmente similar à de classificação, sendo que a principal diferença é que o atributo a ser predito é contínuo em vez de discreto, isto é, número e não classe (REZENDE, 2003).

Para realização de cada tarefa, diferentes técnicas podem ser utilizadas, sendo que não há aquela ideal, com cada uma delas apresentando vantagens e desvantagens. A análise do problema em questão definirá a escolha de determinada técnica, sendo que se pode aplicar mais de um método para solucionar um mesmo problema.

CAPÍTULO 2: Classificação do estoque de carbono do solo em Terra Preta Arqueológica sob diferentes usos

Resumo

Na Amazônia, existem solos conhecidos como Terra Preta Arqueológica, que estão associados à prolongada ocupação humana, os quais apresentam maior qualidade física e química em relação aos outros solos típicos na mesma região, logo compreender o mecanismo de sequestro de carbono destes solos vem sendo discutido com mais intensidade pela comunidade científica. Nesse sentido, o objetivo do trabalho foi desenvolver um modelo preditivo utilizando métodos de seleção de variáveis e indução de árvores de decisão para prever o estoque de carbono no solo em área de Terras Pretas Arqueológicas sob diferentes usos. O experimento foi realizado em áreas de café, cacau, feijão, pastagem e floresta e, as amostras foram coletadas nas camadas de 0,00-0,05, 0,05-0,10 e 0,10-0,20 m. Quatro abordagens para seleção de atributos foram utilizadas para identificar as variáveis mais relevantes, para o modelo proposto, *Correlation based Feature Selection - CFS*, *Qui-quadrado (χ^2)*, *Wrapper* e sem seleção de atributos. Para a classificação dos dados foi utilizada a técnica de indução de árvore de decisão disponível no software *Weka*. Os resultados mostraram que os atributos, uso do solo e areia, foram as variáveis mais importantes para prever o estoque de carbono em Terra Preta Arqueológica sob diferentes usos. Por meio da técnica de indução de árvore de decisão foi gerado um modelo preditivo de estoque de carbono do solo em Terra Preta Arqueológica, com alta taxa de precisão: 83%, 74% e 81%, usando sete (07), sete (07) e quatro (04) regras, nas profundidades estudadas, 0,00-0,05 m, 0,05-0,10 m e 0,10-0,20 m, respectivamente.

Palavras-Chave: Manejo do solo, mineração de dados, árvore de decisão, seleção de atributos, solos antrópicos.

5. Introdução

Na Amazônia, os solos que apresentam horizonte A antrópico estão associados à prolongada ocupação humana pré-histórica, realizada por sociedades indígenas do período pré-colombiano (SANTOS *et al.*, 2018), as quais formaram extensos depósitos de resíduos orgânicos, ou simplesmente rejeitos, que alteraram significativamente os atributos dos solos (AQUINO *et al.*, 2016).

De acordo com Campos *et al.* (2012a) os atributos químicos e físicos das Terra Preta Arqueológica (TPA) são superiores em relação aos de outros solos da Amazônia, pois a maioria dos solos Amazônicos possuem baixa fertilidade. Entretanto, existem várias manchas de solos antrópicos com elevada fertilidade, denominadas regionalmente de TPA (SCHAEFER *et al.*, 2004; SILVA *et al.*, 2012; SCHMIDT *et al.*, 2014). As cores escuras dos horizontes antrópicos são devidas a elevada concentração de algumas formas de carbono de origem predominantemente pirogênico (GLASER, 2007).

Estes solos apresentam estrutura granular (SANTOS *et al.*, 2013), relacionada a ação do carbono pirogênico sobre a formação e estabilização dos agregados (CUNHA *et al.*, 2007). São considerados bem drenados, com porosidade total elevada e boa aeração (CAMPOS *et al.*, 2011; SANTOS *et al.*, 2013), sendo a baixa densidade do solo outra característica comum em TPA (TEIXEIRA; MARTINS, 2003; CAMPOS *et al.*, 2012a). Além disso, segundo Glaser *et al.* (2001), as TPA podem apresentar em média até seis vezes mais carbono orgânico estável que solos não antropogênicos, figurando, portanto, como um grande reservatório de carbono orgânico.

Entretanto, a conversão dos ecossistemas naturais em ambientes de cultivo vem alterando a dinâmica de carbono no solo e, essas modificações frequentemente provocam o declínio no teor de carbono orgânico do solo, em função da redução do aporte, perdas por erosão e por decomposição da matéria orgânica (HICKMANN; COSTA, 2012). Assim, o conhecimento do comportamento do estoque de carbono em áreas de TPA é fundamental e, utilizar técnicas ainda não exploradas se apresenta como uma alternativa promissora. Dentre dessas opções, pode-se utilizar técnicas de mineração de dados, as quais tem sido empregada para extração de padrões e tendências em bancos de dados (HAN *et al.*, 2011).

Uma das técnicas de mineração de dados mais conhecidas na literatura é a árvore de decisão (BREIMAN *et al.*, 1984; QUINLAN, 1996). A representação simbólica e interpretável do modelo obtido por meio de uma árvore de decisão pode ser utilizada tanto para predição como para conhecer a estrutura preditiva do problema. O modelo

gerado é representado graficamente por nós e ramos, parecido com uma árvore, mas no sentido invertido. O nó raiz fica no topo da estrutura e os nós internos contêm um teste de valor sobre um dos atributos mais relevantes e, os resultados desses testes formam os ramos. Os nós folha, nas extremidades, referem-se às classes da variável resposta e representam o resultado da predição obtida pelo modelo (WITTEN *et al.*, 2011).

Dessa forma, o objetivo deste capítulo foi desenvolver um modelo preditivo utilizando métodos de seleção de variáveis e indução de árvores de decisão para prever o estoque de carbono no solo em área de Terra Preta Arqueológica sob diferentes usos.

6. Material e métodos

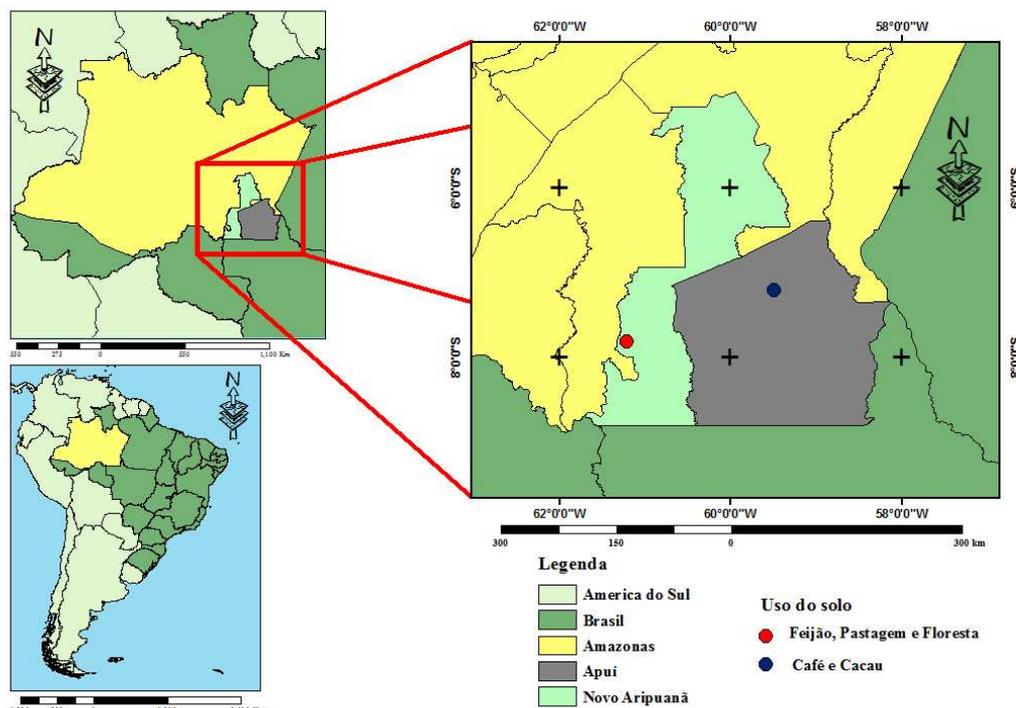
6.1. Descrição da área experimental

O estudo foi conduzido em propriedades rurais na região sul do estado do Amazonas, em áreas com cinco diferentes usos e presença de horizonte A antrópico, conforme Santos *et al.* (2018). As áreas de floresta e pastagem estão localizadas no município de Novo Aripuanã (com coordenadas de 07° 51' 30" S e 61° 18' 01" W) e, as áreas de café, cacau e feijão estão localizadas no município de Apuí (com coordenadas de 07° 12' 05" S e 59° 39' 35" W). O solo foi classificado como Latossolo Amarelo eutrófico argissólico, segundo SiBCS (SANTOS *et al.*, 2018).

O clima da região segundo Köppen é tipo Am (chuvas do tipo monção), caracterizado por apresentar um período seco de pequena duração, com precipitações entre 2.250 e 2.750 mm, temperaturas médias anuais variando entre 25 e 27 °C e umidade relativa do ar entre 85 e 90% (ALVARES *et al.*, 2014). A maior ocorrência de chuvas se dá na parte leste da área e, o período chuvoso geralmente inicia-se em outubro, prolongando-se até junho.

Foram selecionadas cinco áreas de estudos sobre terra preta arqueológica, sob os seguintes usos: (1) Floresta (Mata amazônica) preservada por mais de vinte e cinco anos; (2) Pastagem com brachiaria (*Brachiaria brizanta*); (3) Feijão guandu (*Cajanus cajan*), ressaltando que as duas últimas foram cultivadas com milho, feijão e melancia, durante os últimos 25 anos; (4) Cacau (*Theobroma cacao*), vem sendo cultivada há quatorze anos com essa cultura e, nos primeiros seis anos abrigou os cultivos de arroz, milho, feijão e melancia; (5) Café (*Coffea arabica*) vem sendo cultivada há seis anos. Os solos nestas áreas, nos municípios de Novo Aripuanã e Apuí, foram classificados como Latossolo Amarelo eutrófico argissólico (SANTOS *et al.*, 2018).

Figura 2. Localização da área de estudo com Terra Preta Arqueológica sob diferentes usos, nos municípios de Novo Aripuanã e Apuí, Amazonas, Brasil.



6.2. Descrição do banco de dados

O conjunto de dados original foi composto por 22 atributos do solo, sendo 21 atributos preditivos, formados por 13 atributos físicos, 07 atributos químicos, um atributo relacionado ao uso do solo e um atributo resposta ou atributo meta, que se refere ao estoque de carbono do solo, sendo o alvo da classificação (Tabela 1). Os dados foram agregados em três conjuntos de acordo com a profundidade avaliada, 0,00-0,05, 0,05-0,10 e 0,10-0,20 m, de forma que cada conjunto de dados apresentava 440 observações.

No conjunto de dados foi realizado uma análise estatística exploratória, com a finalidade de retirar os valores discrepantes e entender o comportamento do banco de dados, depois foi feito uma análise estatística descritiva com a finalidade de compreender melhor a variabilidade dos atributos em relação ao estoque de carbono do solo. Também foi realizada uma análise estatística no intuito de observar as diferenças dos atributos entre os diferentes cenários abordados nesta pesquisa.

Tabela 1. Atributos físicos e químicos de Terras Pretas Arqueológicas transformadas e naturais, que compõe o banco de dados.

Atributo	Descrição	Unidade	Tipo	Método	
Uso do solo	Floresta, feijão, pastagem, café e cacau	-	Preditivo	-	
Físico	Areia	g kg ⁻¹	Preditivo	Teixeira <i>et al.</i> (2017)	
	Silte	g kg ⁻¹	Preditivo		
	Argila	g kg ⁻¹	Preditivo		
	Umidade volumétrica	kg kg ⁻¹	Preditivo		
	Densidade do solo	kg dm ⁻³	Preditivo		
	Macroporosidade	m ³ m ⁻³	Preditivo		
	Microporosidade	m ³ m ⁻³	Preditivo		
	Volume total de poros	%	Preditivo		
	Diâmetro médio geométrico	Mm	Preditivo		
	Diâmetro médio ponderado	Mm	Preditivo		Kemper e Chepil (1965)
	Agregados > 2,0 mm	mm (%)	Preditivo		
	Agregados < 2,0 mm	mm (%)	Preditivo		Stolf <i>et al.</i> 2014
Resistência do solo à penetração	MPa	Preditivo			
Químico	pH	H ₂ O	Preditivo	Raij <i>et al.</i> (2001)	
	Acidez potencial	mmol _c dm ⁻³	Preditivo		
	Alumínio	mmol _c dm ⁻³	Preditivo		
	Fósforo	mg dm ⁻³	Preditivo		
	Potássio	mmol _c dm ⁻³	Preditivo		
	Cálcio	mmol _c dm ⁻³	Preditivo		
	Magnésio	mmol _c dm ⁻³	Preditivo		
	Estoque de carbono no solo	Mg ha ⁻¹	Resposta		Sisti <i>et al.</i> (2004)

6.3. Metodologia de indução de Árvore de decisão

Para identificar diferentes níveis de estoque de carbono do solo do solo foi realizada a discretização do atributo meta em categorias. Para isso, os valores de estoque de carbono do solo foram ordenados de forma crescente e divididos igualmente em duas classes: baixa e alta (Tabela 2).

Tabela 2. Estoque de carbono do solo do solo (Mg ha⁻¹) de acordo com as classes (baixa e alta) e seus respectivos limites, nas profundidades de 0,00-0,05, 0,05-0,10 e 0,10-0,20 m.

Classe	Limites
	0,00-0,05 m
Baixa	[29; 207]
Alta	[208; 574]
	0,05-0,10 m
Baixa	[28; 171]
Alta	[172; 418]

Classe	0,10-0,20 m		Limites
Baixa			[22; 154]
Alta			[155; 340]

6.4. Seleção de atributos

Devido ao grande número de atributos gerados no pré-processamento de dados, uma seleção de atributos foi utilizada visando melhorar o desempenho do algoritmo de aprendizagem, simplificar os modelos de predição e reduzir o custo computacional e finalmente, oferecer um melhor entendimento sobre os resultados encontrados. Para isso, foram avaliadas quatro abordagens para seleção de atributos:

- i. Sem seleção de atributos, onde ocorreu a utilização de todos os atributos, caracterizando-se pela ausência de seleção.
- ii. Seleção de características da correlação (CFS) que identifica um conjunto de atributos pouco correlacionados entre si e bem correlacionados com o atributo meta.
- iii. Método do Qui-quadrado (χ^2), o qual se fundamenta no conceito de independência estatística. Para isso, os atributos são avaliados individualmente usando a medida χ^2 , com relação à classe de interesse. Quanto maior o valor de χ^2 , mais provável é a associação entre as variáveis predictoras e o atributo meta.
- iv. Método Wrapper, esta abordagem avalia os atributos utilizando estimativas de precisão providas pelos algoritmos de aprendizado predeterminados (FREITAS, 1998).

6.5. Indução e validação do modelo de classificação

Os modelos induzidos foram avaliados utilizando o método de validação cruzada em 10 partes e três métricas: (i) taxa de acerto (acurácia); (ii) o número de regras geradas, que em geral são associadas à facilidade de interpretação do modelo; (iii) a estatística *Kappa*, que é uma medida de concordância entre as classes preditas e observadas pelo classificador.

Como resultado da indução do modelo de árvore de decisão, obtém-se a conhecida matriz de erros ou matriz de confusão (Tabela 3), amplamente utilizada em análise estatística de concordância (HAN *et al.*, 2011).

Tabela 3. Matriz de erros para análise da acurácia, de um classificador binário (árvore de decisão).

		PREDITO		
		Classe A	Classe B	Total
VERDADEIRO	Classe A	VP	FN	P
	Classe B	FP	VN	N
Total		P'	N'	P+N

VP = verdadeiros positivos; FP = falsos positivos; FN = falsos negativos; VN = verdadeiros negativos.

Na coluna *Total* da Tabela 3, *P* é o valor total de casos positivos e *N* é o total de casos negativos existentes no conjunto de treinamento. Na linha *Total*, *P'* é o total de casos que o modelo classificou como casos positivos e *N'* é o total de casos classificados como negativos.

A partir da matriz de confusão, é possível extrair as métricas de avaliação de desempenho. A taxa de acerto ou acurácia (*accuracy*) é a porcentagem de exemplos que foram classificados corretamente pelo classificador e, pode ser expressa na Equação 1.

$$\text{Taxa de acerto} = \frac{VP + VN}{P + N} \quad (1)$$

Para descrever a medida de concordância entre as classes preditas e observadas, que infere o número esperado de acertos (utilizando uma classificação ao acaso) do número real de acertos do classificador, utiliza-se a medida *Kappa*. Seus valores variam de 0 a 1, representando resultados de classificação ruins e excelentes, respectivamente (WITTEN *et al.*, 2011).

O coeficiente *Kappa* é calculado levando-se em consideração todas as classes. Uma possível interpretação do desempenho dos modelos a partir da estatística *Kappa* foi introduzida por Landis e Koch (1977), sendo expressa na Tabela 4.

Tabela 4. Classificação da qualidade do desempenho do modelo de árvore de decisão, por meio da estatística Kappa.

Coeficiente <i>kappa</i>	Qualidade
< 0,00	Péssima
0,00-0,20	Ruim
0,21-0,40	Razoável
0,41-0,60	Boa
0,61-0,80	Muito Boa
0,81-1,00	Excelente

Adaptado de Landis e Koch (1977).

Para a classificação dos dados foi utilizado o método de árvore de decisão binária induzida no software Weka 3.6. O algoritmo de indução utilizado foi o J48, amplamente conhecido como C4.5, desenvolvido por Quinlan (1993).

7. Resultados

Inicialmente foi realizada uma análise por meio de gráficos do tipo *boxplot*, em que foram identificados a distribuição dos dados e os respectivos valores discrepantes para cada sistema de uso do solo, com a finalidade de diminuir o efeito tendencioso de tais valores (Tabela 5). Foram retirados 3,21% dos dados da profundidade 0,00-0,05 m, 2,45% de 0,05-0,10 m e 2,61% de 0,10-0,20 m. Os demais valores foram mantidos para a geração dos modelos de árvore de decisão.

Tabela 5. Quantidade de *Outliers* identificados, pelo método de *bloxplot*, na base de dados em Terras Pretas Arqueológicas.

Profundidade (m)	Uso do solo					SUB TOTAL
	Floresta	Cacau	Café	Feijão	Pastagem	
0,00-0,05	59	63	70	39	78	309
0,05-0,10	53	49	56	32	37	227
0,10-0,20	48	51	54	44	45	242
TOTAL						778

Na Tabela 6 são apresentados os valores da média do estoque de carbono do solo, produto da análise estatística descritiva dos dados para as profundidades de 0,00-0,05, 0,05-0,10 e 0,10-0,20 m, respectivamente. Tendo como foco a variável resposta, observa-se que o uso do solo com cacau e café apresentaram maiores valores médios para o estoque de carbono do solo do que os outros sistemas de uso de solo (feijão, pastagem e floresta), para todas as profundidades estudadas.

Na profundidade 0,00-0,05 m (Tabela 6), o uso do solo com cacau apresentou a maior média para o estoque de carbono do solo (378,94 Mg ha⁻¹), seguido do café (272,50 Mg ha⁻¹), feijão (225,35 Mg ha⁻¹), pastagem (159,29 Mg ha⁻¹) e floresta (158,90 Mg ha⁻¹). Para 0,05-0,10 m, existe um grupo com valores altos, em que o uso do solo com cacau apresenta a maior média de estoque de carbono (279,50 Mg ha⁻¹), seguido do café (233,17 Mg ha⁻¹) e, um outro grupo com valores mais baixos, floresta (162,02 Mg ha⁻¹), pastagem (146,84 Mg ha⁻¹) e feijão guandu (122,43 Mg ha⁻¹).

Tendência similar foi encontrada na profundidade 0,10-0,20 m (Tabela 6), onde o uso do solo com café apresenta o maior valor médio de estoque de carbono (223,03 Mg ha⁻¹), seguido pelo cacau (216,59 Mg ha⁻¹) e, outro grupo formado por floresta, feijão e pastagem (127,56 Mg ha⁻¹, 126,41 Mg ha⁻¹ e 124,30 Mg ha⁻¹, respectivamente) em ordem decrescente.

Tabela 6. Estoque de carbono do solo (Mg ha⁻¹) em Terra Preta Arqueológica sob diferentes sistemas de uso (valores médios de 88 amostras por cada sistema de uso).

Profundidade (m)	Uso do solo									
	Cacau		Café		Feijão		Pastagem		Floresta	
	Média	DP	Média	DP	Média	DP	Média	DP	Média	DP
0,00-0,05	378,9	74,92	272,5	88,28	225,45	63,40	159,3	32,64	158,9	65,73
0,05-0,10	279,5	28,03	233,2	76,49	122,4	33,48	146,8	39,41	162,0	35,31
0,10-0,20	216,6	36,79	223,0	37,21	126,4	56,56	124,3	1,24	127,6	45,69

DP = desvio padrão.

Em relação aos métodos de seleção de atributos, observa-se na Tabela 7, que a utilização de todos os atributos presentes no banco de dados, para as três profundidades, gerou modelos com alta taxa de acerto (81%, 75% e 80%, respectivamente) e estatística *Kappa* considerada boa (0,62, 0,50 e 0,61, respectivamente), porém, as árvores geradas nestes modelos contêm elevado número de regras (24, 35 e 38, respectivamente), o que faz com que outras possibilidades sejam exploradas.

Tabela 7. Desempenho do classificador J48, na geração do modelo preditivo de estoque de carbono em Terra Preta Arqueológica, sob diferentes sistemas de uso.

Método de seleção	Atributos selecionados	Acurácia (%)	Coefficiente Kappa	Número de regras
0,00-0,05 m				
Sem seleção	Todos	81	0,62	24
CFS	Uso do solo, areia e Ca	83	0,67	8
x^2	Todos menos: > 2 mm, < 2 mm e DMG	81	0,62	20
Wrapper	Uso do solo e Ca	86	0,72	7
0,05-0,10 m				
Sem seleção	Todos	75	0,50	35
CFS	RSP, Ds, Micro, VTP, Silte, Areia, Umidade, H ⁺ + Al ³⁺ , Ca e Mg	74	0,47	15
x^2	Todos menos: Macro	77	0,53	32
Wrapper	Uso do solo	79	0,58	5
0,10-0,20 m				
Sem seleção	Todos	80	0,61	38
CFS	Ds, Micro, Silte, Areia, Umidade, H ⁺ + Al ³⁺ , Ca e Mg	81	0,61	15
x^2	Todos menos: Macro	80	0,61	38
Wrapper	Uso do solo	81	0,63	5

RSP = resistência do solo à penetração (MPa); Ds = densidade do solo (kg dm^{-3}); Macro = macroporosidade ($\text{m}^3 \text{m}^{-3}$); Micro = microporosidade ($\text{m}^3 \text{m}^{-3}$); VTP = volume total de poros ($\text{m}^3 \text{m}^{-3}$); Umidade = umidade volumétrica ($\text{kg}^1 \text{kg}^{-1}$); DMG = diâmetro médio geométrico (mm); > 2mm = classe de diâmetro maior que 2,0 mm; < 2,0 mm = classe de diâmetro menor que 2,0 mm; H⁺ + Al³⁺ = acidez potencial (mmolc dm^{-3}); Ca = cálcio (mmolc dm^{-3}); Mg = magnésio (mmolc dm^{-3}).

Dentre os métodos de seleção de atributos avaliados, o método *Wrapper* foi o que obteve os melhores resultados, pois das 21 variáveis preditivas disponíveis no conjunto de dados, para a profundidade de 0,00-0,05 m, o método selecionou apenas duas variáveis e, obteve uma taxa de precisão de 86%, com coeficiente *Kappa* de 0,72 e sete regras geradas (Tabela 7). Na profundidade 0,05-0,10 m, o método *Wrapper* selecionou apenas uma variável, com uma taxa de precisão de 79% e, o coeficiente *Kappa* de 0,58 e, com cinco regras geradas.

Para a profundidade de 0,10-0,20 m, o método *Wrapper* apresentou um desempenho superior em relação aos outros métodos, selecionando apenas uma variável com uma taxa de precisão de 81%, coeficiente *Kappa* de 0,63 e com cinco regras geradas (Tabela 7). Em contrapartida, apesar do método *Wrapper* apresentar melhores métricas, o fato de ter selecionado apenas dois, um e um atributo nas três profundidades estudadas, 0,00-0,05, 0,05-0,10 e 0,10-0,20 m, respectivamente, faz com que seja avaliada a possibilidade de escolher um outro método, pois a seleção de poucos atributos leva a perda de informação importante que poderia ajudar em aumentar a eficiência na tomada de decisões a partir do conhecimento gerado.

Nesse sentido, o método de seleção de atributos CFS torna-se a melhor opção, uma vez que apresenta valores relevantes para as métricas avaliadas, com taxa de precisão de 83%, 74% e 81%, estatística *Kappa* de 0,67, 0,47 e 0,61 e, número de regras de 8, 15 e 15 nas três profundidades estudadas, respectivamente (Tabela 7).

Em seguida, verificou-se que o modelo de árvore de decisão, gerado com atributos selecionados pelo método CFS, mantém a taxa de acerto, estatística *Kappa* relativamente constante, considerando uma variação do número mínimo de objetos por folhas entre 2 e 35 (Tabela 8). Observou-se uma diminuição significativa do número de regras para os níveis iniciais de poda nas profundidades de 0,05-0,10 e 0,10-0,20 m, diferente do que acontece na profundidade de 0,00-0,05 m, conforme aumenta o número mínimo de objetos.

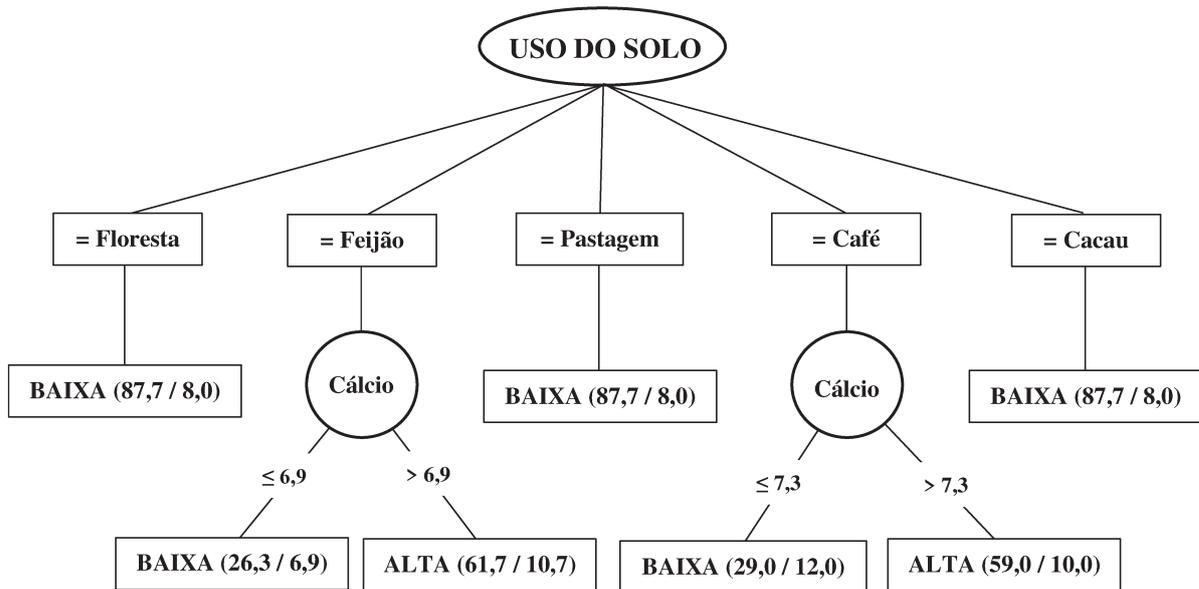
Tabela 8. Relação entre o número mínimo de objetos (minNumObj), taxa de precisão, estatística *Kappa* e número de regras nas três profundidades estudadas.

minNumObj	0,00-0,05 m			0,05-0,10 m			0,10-0,20 m		
	Acurácia	<i>Kappa</i>	# Regras	Acurácia	<i>Kappa</i>	# Regras	Acurácia	<i>Kappa</i>	# Regras
3	83,67	0,67	7	74,32	0,49	15	81,09	0,62	4
5	83,90	0,68	7	73,86	0,48	13	81,55	0,63	4
7	84,13	0,68	7	75,91	0,52	11	81,32	0,63	4
10	84,58	0,69	7	74,32	0,49	7	81,32	0,63	4
12	84,58	0,69	7	73,41	0,47	7	81,55	0,63	4
15	84,58	0,69	7	74,55	0,49	3	82,00	0,64	4
20	85,03	0,70	7	74,32	0,49	3	81,55	0,63	4
25	85,26	0,71	7	75,91	0,52	3	80,64	0,61	4

Dessa forma, considera-se que os melhores registros nas métricas avaliadas foram alcançadas utilizando 12 como número mínimo de objetos, isto é, pelo menos 12 instâncias do conjunto de treinamento garantem a geração das regras presentes no modelo (Tabela 8). Para este caso, verifica-se acurácia de 85%, 73% e 82%, estatística *Kappa* de 0,69, 0,47 e 0,63 e número de regras de 7, 7 e 4, nas três profundidades estudadas, respectivamente.

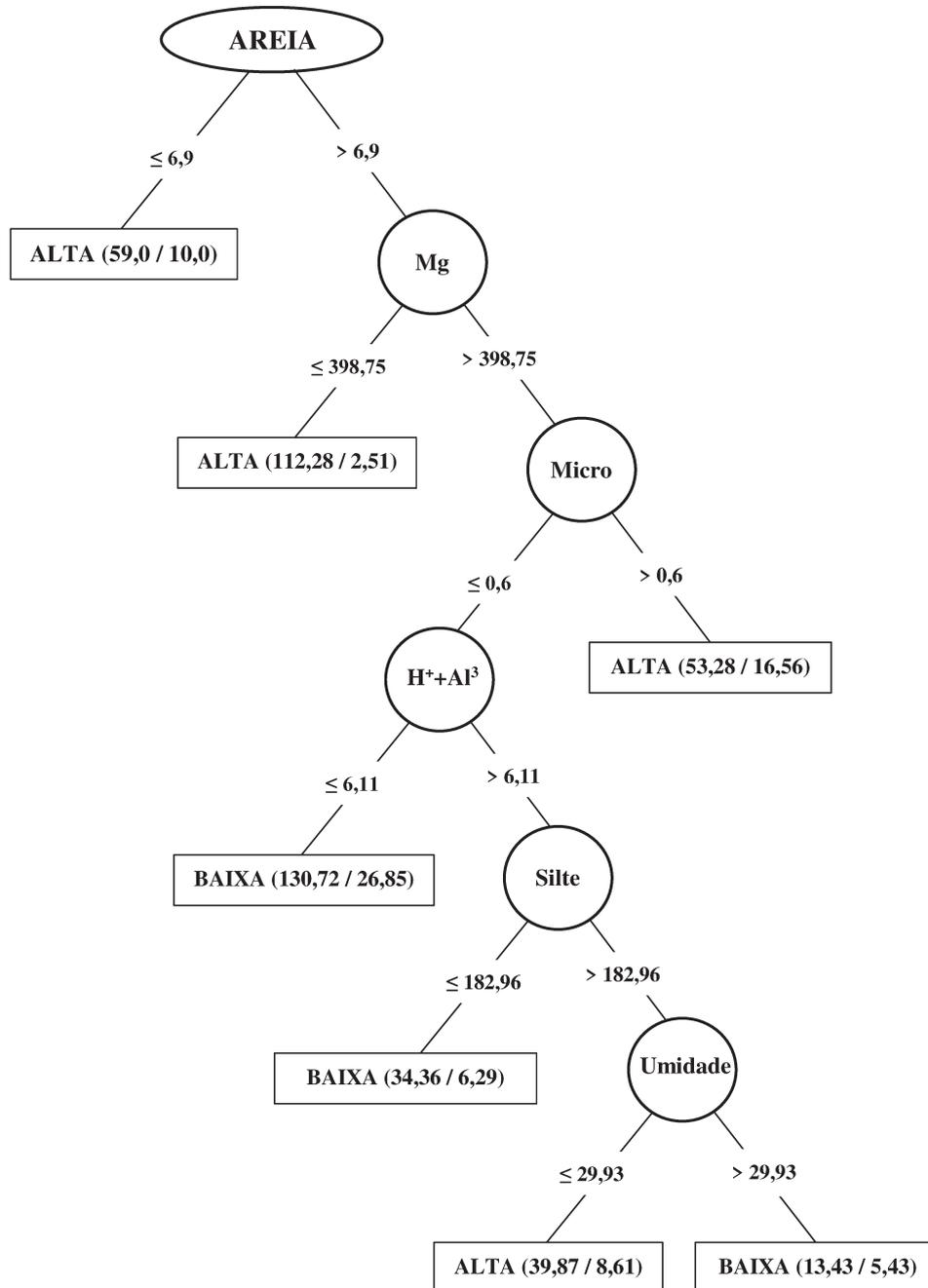
Após determinado o parâmetro de poda para geração do modelo de árvore de decisão, foram construídas as árvores de decisão, para cada uma das três profundidades estudadas. Na Figura 2, encontra-se o modelo resultante da combinação do método de seleção CFS para escolha de atributos e o algoritmo J48 para indução de árvore de decisão na profundidade de 0,00-0,05 m, verificando-se que a variável mais importante (nó 1) foi o atributo uso de solo. Observa-se que o segundo atributo mais importante foi cálcio (Ca) e que a árvore de decisão apresenta duas ramificações, a primeira vinculada ao uso do solo com feijão e, a outra vinculada ao uso do solo com café.

Figura 3. Árvore de decisão resultante da utilização do método CFS para seleção de atributos e o algoritmo J48 para indução de árvore de decisão, na profundidade de 0,00-0,05 m.



O modelo resultante da combinação do método de seleção CFS e o algoritmo J48 para indução de árvore de decisão na profundidade de 0,05-0,10 m, demonstrou uma conformação com maior diversidade de atributos, sendo a areia o atributo que fica no topo da árvore, acompanhado dos atributos magnésio, microporosidade, acidez potencial, silte e umidade volumétrica em ordem decrescente (Figura 3). Por outro lado, verifica-se a conformação de uma sub-árvore a partir do nó acidez potencial, desde o qual está estabelecida uma relação direta com os atributos silte e umidade volumétrica.

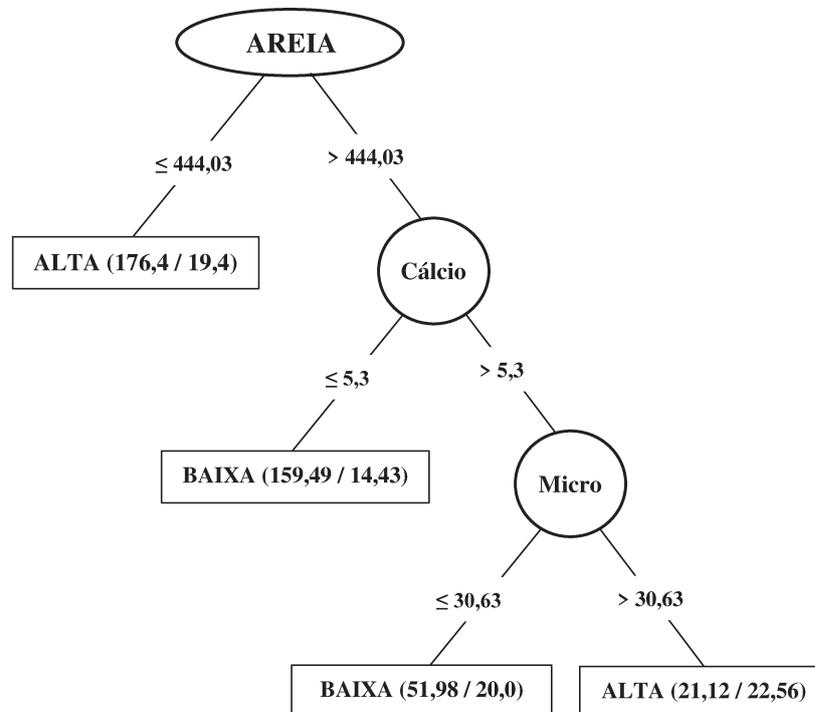
Figura 4. Árvore de decisão resultante da utilização do método CFS para seleção de atributos e o algoritmo J48 para indução de árvore de decisão, na profundidade de 0,05-0,10 m.



Na árvore que mostra o modelo resultante da combinação do método de seleção CFS e o algoritmo J48 para indução de árvore de decisão na profundidade de 0,10-0,20 m, verifica-se que a árvore é governada por apenas quatro regras, ocorrendo a participação de três atributos, areia, cálcio e microporosidade (Figura 4). Por exemplo, a regra mais simples da árvore seria, se o valor do atributo areia é igual ou menor a 444 Mg ha⁻¹ o modelo classifica o estoque de carbono como ALTA. Por outro lado, se o atributo areia for maior que 444 Mg ha⁻¹ e o tributo cálcio for menor ou igual a 5,3 mmol_c dm⁻³, a classificação do estoque de carbono

será BAIXA, mas se o valor supera os 5,3 $\text{mmol}_e \text{dm}^{-3}$ para ser classificado depende da combinação com o valor do atributo microporosidade, portanto, se o valor é menor ou igual a 30,63 $\text{m}^3 \text{m}^{-3}$ será classificado como BAIXA, do contrário será classificado como ALTA.

Figura 5. Árvore de decisão resultante da utilização do método CFS para seleção de atributos e o algoritmo J48 para indução de árvore de decisão, na profundidade de 0,10-0,20 m.



8. Discussão

Um dos primeiros passos para obter uma análise coerente é a detecção de *Outliers* (BEN-GAL, 2005), pois os valores discrepantes detectados podem levar a erros de especificação do modelo, estimativas de parâmetros tendenciosas e resultados incorretos. Sendo que os *outliers* influem nos cálculos de média, desvio padrão, histogramas, podendo distorcer conclusões e generalizações sobre o conjunto de dados analisados (TRIOLA, 2008), portanto, é importante identificar os valores discrepantes antes da modelagem e análise (LIU *et al.*, 2004).

Os resultados da análise estatística descritiva em todas as profundidades estudadas, mostram um cenário novo a respeito do esperado para o estoque de carbono no solo. Pois acreditava-se que o uso do solo com floresta apresentasse o maior estoque de carbono, porém, verifica-se que os dados apresentam a tendência que o uso de solo com cacau e café apresentam valores médios superiores para o estoque de carbono que os outros usos de solo (feijão, pastagem e floresta).

Essa condição deve-se possivelmente às transformações dos resíduos vegetais em carbono são mais difíceis em condições naturais (floresta), do que em condições sob cultivos

(cacau e café). Estes resultados devem-se possivelmente a estabilidade dos sistemas de cultivo, espaçamento da cultura, maior acúmulo de serapilheira e alto aporte de carbono orgânico, conforme observado por Campos *et al.* (2016).

Outra observação importante foi uma tendência de diminuição dos valores dos atributos estudados conforme aumenta a profundidade. Esse comportamento está relacionado ao maior aporte orgânico na superfície do solo, devido à queda natural e das podas da parte aérea das plantas e a renovação das raízes superficiais. De esse modo, a distribuição do carbono ao longo do perfil do solo pode ser variável, em função do tipo de vegetação.

Neste caso, os maiores teores de carbono obtidos nas camadas superficiais podem ser explicados devido à superfície do solo sofrer maior interferência da matéria orgânica depositada, promovendo processos mais intensos na ciclagem de nutrientes (VITAL *et al.*, 2004), essa tendência também foi identificada por Kato *et al.* (2010) que verificaram maior teor de carbono orgânico nas camadas superficiais, ao avaliarem o efeito de diferentes usos do solo, sofrendo redução com a profundidade. Corroborando os resultados obtidos por Cunha *et al.* (2007), Rossi *et al.* (2012), Campos *et al.* (2012a) e Mota Júnior *et al.* (2017), que encontraram maiores teores de carbono nos horizontes A nas TPA.

Ao respeito da seleção de atributos, observa-se um melhor desempenho do método *Correlation-based feature selection* (CFS). Hall (2000), afirma que esse método foi desenvolvido para trabalhar a seleção de atributos na predição do atributo alvo com valor numérico, concordante com o cenário no qual se desenvolve a mineração de dados nesta pesquisa.

A análise da árvore gerada na profundidade de 0,00-0,05 m evidencia que o atributo com maior ganho de informação, conhecido como atributo "*split*", na classificação do estoque de carbono em Terras Pretas Arqueológicas foi o uso do solo. Esse resultado está de acordo com a literatura, que relata que o uso e manejo do solo tem efeito direto sobre os atributos físicos e químicos do solo, além de interferências relevantes no teor de carbono orgânico do solo (MARCHINI *et al.*, 2015; PEREIRA; THOMAZ, 2015; COLOMBO *et al.*, 2017).

Em concordância, Cunha *et al.* (2017), avaliaram os atributos físicos e estoque de carbono do solo em áreas de Terra Preta Arqueológica na região Amazônica nas mesmas profundidades dessa pesquisa e, observaram que tanto o carbono orgânico quanto o estoque de carbono apresentaram alterações de acordo com os respectivos usos do solo, com valores médios superiores na camada superficial 0,00-0,05 m, passando a ter valores menores na profundidade de 0,05-0,10 m, finalmente não havendo diferença estatística entre os usos na profundidade 0,10-0,20 m. Em adição, Campos *et al.* (2012a), em caracterização física e

química de Terras Pretas Arqueológicas, encontraram maiores teores de carbono nos horizontes A.

Um outro fato a considerar para explicar o resultado encontrado é a relação carbono e nitrogênio (C/N) em cada cenário considerado para esta pesquisa, pois quanto maior a relação C/N menor tende a ser a decomposição da matéria orgânica e conseqüentemente a influência do uso do solo no estoque de carbono. Acosta *et al.* (2014) relataram que a relação C/N tem sido apontada por muitos trabalhos como principal atributo no processo de decomposição, o que vem de encontro com os resultados obtidos por Moitinho *et al.* (2015), os quais observaram que o FCO₂ e a relação C/N são inversamente proporcionais, ou seja, o aumento na taxa de um atributo provoca o decréscimo do outro.

Logo, para as profundidades de 0,05-0,10 e 0,10-0,20 m, observa-se que o atributo areia foi selecionado como o mais importante para a geração do modelo preditivo. Estando de acordo com os resultados obtidos por Saiz *et al.* (2012), ao estudarem as variações do estoque de carbono no solo, identificaram que os teores de areia foram mais determinantes que os teores de argila.

Assim como Marques *et al.* (2016) informaram que os solos arenosos apresentam grande potencial de armazenamento temporário de carbono, mas que podem ser facilmente perdidos para o ambiente em situações de exploração devido as características físicas do solo. Esta relação do estoque de carbono com a textura do solo foi apresentada por Lal (2007), que afirmam que solos mais arenosos tendem a apresentar maior perda de carbono, sendo explicado pela maior lixiviação, menor agregação das partículas do solo e baixa atividade da areia.

Por outro lado, em pesquisa num cenário similar a nossa pesquisa, com TPA sob cultivo de feijão e pastagem, Cunha *et al.* (2017), afirmaram que os maiores teores de areia sob as áreas de feijão guandu e pastagem se devem provavelmente a fortes chuvas que promovem grandes enxurradas, por essas apresentarem áreas mais abertas, com impacto das gotas de chuva sobre o solo em relação à área sob floresta. Segundo Corado Neto *et al.* (2015), perdas de carbono orgânico podem estar associadas a condições de baixa cobertura vegetal e elevado escoamento superficial da água, estando este último associado à conformação do terreno em certa topografia e ao regime pluviométrico anual da região de estudo de 2.300 mm (CAMPOS *et al.*, 2012b).

A dominância da fração areia também foi encontrada por Campos *et al.* (2012b), estudando a caracterização de horizontes antrópicos na região de Apuí-AM, associando esta característica granulométrica à formação destes solos e materiais constituintes. Outros autores afirmam que os solos de TPA, em geral, apresentam horizonte A antrópico mais arenoso (SOUZA *et al.*, 2009; SANTOS *et al.*, 2011; SILVA *et al.*, 2012), concordando por sua vez

com Santos *et al.* (2017), que encontraram uma relação entre o estoque de carbono com a granulometria.

Em estudos numa área com similar uso do solo que as estudadas nessa pesquisa, Soares *et al.* (2015), encontraram que a fração areia foi dominante em todas as profundidades, justificando o fato da natureza grosseira do material de origem, oriundo dos arenitos da formação Içá. Também, Campos *et al.* (2011), caracterizando Terras Pretas Arqueológicas na região de Manicoré-AM, Silva *et al.* (2011) em amostras de solos de TPA na Amazônia Central e, Teixeira *et al.* (2009) observaram que a fração areia foi dominante em todos os perfis analisados em relação aos solos circundantes.

Essa relação de areia com carbono orgânico pode ser explicada abordando o fracionamento físico da matéria orgânica, que permite o estudo dos compartimentos mais humificados ligados à fração mineral do solo, conhecida como fração pesada (FP). Nesta FP, são encontrados os maiores estoques de carbono do solo, cuja dinâmica estaria intimamente relacionada à textura (FELLER; BEARE, 1997; FREIXO *et al.*, 2002).

Para Teixeira *et al.* (2010), a textura mais arenosa nas TPA parece estar relacionada com o uso intenso do fogo e a disponibilidade de material orgânico, que combinados poderiam formar partículas organominerais estáveis do tamanho da fração areia. O efeito do fogo na textura do solo foi comprovado por Ketterings *et al.* (2000) estudando a mudança da mineralogia e textura do solo causada por incêndio em Sumatra, Indonésia e, também por Thomaz e Fachin (2014), que em estudo com queimada natural em campo e experimento com temperatura controlada em laboratório, constataram aumento no teor de areia no solo.

Contudo, os resultados gerados para as três profundidades nos cenários estudados mostraram que as técnicas de mineração de dados permitiram o desenvolvimento de um modelo eficiente para a classificação do estoque de carbono em TPA, utilizando métodos de seleção de atributos e o algoritmo J48 para indução de árvore de decisão. Demonstrando que o uso de mineração de dados é uma abordagem promissora para ser usada na área ciência do solo por mostrar resultados consistentes com a literatura, assim como foi comprovado em outros estudos na área (LIMA *et al.*, 2017; FARHATE *et al.*, 2018).

Finalmente, neste estudo, as variáveis mais importantes para prever estoque de carbono no solo em TPA, corresponderam ao uso do solo e areia, demonstrando que a mudança de uso do solo tem efeitos relevantes no teor de carbono orgânico do solo (MARCHINI *et al.*, 2015; PEREIRA; THOMAZ, 2015; COLOMBO *et al.*, 2017), sendo que existem estudos que reportam que os estoques de carbono no solo são influenciados pelos sistemas de manejo (BARRETO *et al.*, 2014; CUNHA *et al.*, 2017, MASCARENHAS *et al.*, 2017).

Desse modo, pode-se inferir nas TPA, que mesmo com o grande potencial de armazenamento de carbono, essas terras são mais susceptíveis a ter maiores perdas de carbono (LAL, 2007; MARQUES *et al.* 2016), ao apresentar dominância da fração areia na textura do solo (SOUZA *et al.*, 2009; SANTOS *et al.*, 2011; SILVA *et al.*, 2012; CAMPOS *et al.* 2012b; SOARES *et al.*, 2015; SANTOS *et al.*, 2017).

9. Conclusões

Os modelos resultantes neste estudo, classificam as variáveis mais importantes para prever estoque de carbono no solo em Terras Pretas Arqueológicas corresponderam a variável de uso do solo e variável areia.

Por meio da técnica de indução de árvore de decisão foi gerado um modelo preditivo de estoque de carbono do solo em Terra Preta Arqueológica, com alta taxa de precisão: 83%, 74% e 81%, usando sete (07), sete (07) e quatro (04) regras, nas profundidades estudadas, 0,00-0,05 m, 0,05-0,10 m e 0,10-0,20 m, respectivamente.

CAPÍTULO 3: *Random Forest* para predição do estoque de carbono em Terra Preta Arqueológica sob diferentes usos

Resumo

Os solos amazônicos são, em sua maioria, altamente intemperizados, ácidos, de baixa fertilidade e capacidade de troca de cátions. Porém, em contraste, existem também solos com atributos físicos e químicos superiores, denominados Terras Pretas Arqueológicas, que estão associados à ocupação por sociedades indígenas do período pré-colombiano e, podem atuar como um modelo, para entender melhor a fertilidade a longo prazo e o potencial de sequestro de carbono. Na atualidade existem ferramentas como o algoritmo *Random Forest*, que é uma técnica de mineração de dados usada para identificar atributos em ordem de importância que explicam a variação em um atributo alvo, como o estoque de carbono do solo em Terra Preta Arqueológica. Este trabalho teve como objetivo elaborar um modelo de regressão para predição de estoque de carbono do solo em Terra Preta Arqueológica com diferentes usos, por meio do algoritmo *Random Forest*. Cinco áreas com diferente sistema de manejo de solo foram selecionadas: café, cacau, pastagem, feijão e floresta, em cada área, foi montada uma grade de amostragem com 88 pontos, com 03 profundidades e foram avaliados os atributos físicos e químicos, sendo o estoque de carbono do solo o atributo alvo. Para interpretação dos dados, utilizou-se o algoritmo *Random Forest*, baseado na combinação de árvores de decisão preditas (aprendizado de máquina) aonde cada árvore depende dos valores de um vetor aleatório amostrado independentemente com a mesma distribuição para todas as árvores da floresta. Os resultados indicaram que o algoritmo *Random Forest* originou modelos com bom ajuste ($R^2 = 0,71, 0,71$ e $0,77$) para valores preditos e observados, aonde: o teor de cálcio, a densidade do solo e o uso do solo foram as variáveis mais importantes que afetaram o estoque de carbono em Terra Preta Arqueológica sob diferentes usos, nas três profundidades estudadas, 0,00-0,05, 0,05-0,10 e 0,10-0,20 m, respectivamente.

Palavras chave: Mineração de dados, solos antrópicos, modelo de regressão, aprendizado de máquina.

10. Introdução

Em meio à vasta região Amazônica ocorrem áreas em que a característica original do solo foi modificada por processos antrópicos, tais solos são conhecidos como Terra Preta Arqueológica (TPA), podendo ser comumente encontrados na paisagem Amazônica (CAMPOS *et al.*, 2011). As áreas de TPA apresentam atributos físicos e químicos superiores em relação aos de outros tipos de solos da Amazônia, elas apresentam uma elevada fertilidade natural (CUNHA *et al.*, 2017), quando comparado aos solos circunvizinhos que predominam na Amazônia (SOARES *et al.*, 2018), na grande maioria, extremamente pobre em nutrientes.

Os altos teores de matéria orgânica, colocam as TPA como um grande reservatório de carbono orgânico (CAMPOS *et al.*, 2012), como descrito por Novotny *et al.* (2009), o carbono destes solos apresenta a capacidade de ser até seis vezes mais estáveis do que das áreas adjacentes. No entanto, o uso de práticas agrícolas inadequada e de forma intensiva promovem alterações nas propriedades originais das diferentes classes de solos (ROZANE *et al.*, 2010; VIANA *et al.*, 2011; ROSSETTI; CENTURION, 2015).

Um desses atributos que pode ser alterado pelo manejo é a matéria orgânica do solo (MOS), a qual tem influência direta e indiretamente aos outros atributos do solo (VIANA *et al.*, 2011). A MOS desempenha efeitos positivos sobre a estrutura, além de aumentar a capacidade de retenção de cátions do solo, estabilizar a temperatura, manter a umidade, dentre outros benefícios (ZHU *et al.*, 2016), que são encontrados nas TPA.

O interesse da comunidade científica por conhecer cada vez melhor a dinâmica da Terra Preta Arqueológica e o seu potencial para estocar carbono no solo, fez com que a informação gerada ao respeito aumente gerando um cenário ímpar de quantidade de dados. Recentemente, com o avanço da área da inteligência artificial, técnicas de aprendizado de máquina têm se mostrado promissoras em relação aos métodos estatísticos de regressão para analisar bases de dados (MARTINS *et al.*, 2017).

O aprendizado de máquina é uma área de pesquisa da inteligência artificial que visa o desenvolvimento de programas de computador com a capacidade de aprender a executar uma dada tarefa com sua própria experiência (FACELI *et al.*, 2011). Isso leva ao desenho de programas capazes de aprender por si só, utilizando-se um conjunto de dados que representam experiências passadas.

Entre os vários algoritmos de aprendizado de máquina, o emergente algoritmo *Random Forest* (RF) proposto por Leo Breiman e Cutler Adele em 2001 (BREIMAN, 2001), foi considerado como um dos métodos de previsão mais precisos para regressão, pois pode modelar interações complexas entre variáveis de entrada e é relativamente robusto em relação a valores atípicos.

O algoritmo de RF apresenta várias vantagens, as quais são: (a) natureza não paramétrica; (b) um dos algoritmos de aprendizagem mais precisos disponíveis; (c) lida com milhares de variáveis de entrada, não sendo necessário excluir variáveis; (d) método eficaz para estimar dados perdidos com precisão quando uma grande proporção de dados estão faltando; (e) robustez na presença de ruído e variáveis sem importância; (f) alta flexibilidade para realizar vários tipos de análise de dados (RODRIGUEZ-GALIANO *et al.*, 2012).

Neste sentido, o uso da abordagem de *Random Forest* pode auxiliar na tomada de decisão do uso e manejo adequado do solo em Terra Preta Arqueológica, baseando-se na identificação dos atributos mais importantes do solo e, sua correlação com o estoque de carbono. Assim, o objetivo desse trabalho foi elaborar um modelo de regressão para predição de estoque de carbono do solo em Terra Preta Arqueológica com diferentes usos, por meio da abordagem *Random Forest*.

11. Material e Métodos

11.1. Descrição da área experimental

O estudo foi conduzido em propriedade rurais na região sul do estado do Amazonas, em áreas com cinco diferentes usos e presença de horizonte A antrópico, conforme Santos *et al.* (2018). As áreas de floresta e pastagem estão localizadas no município de Novo Aripuanã (com coordenadas de 07° 51' 30" S e 61° 18' 01" W), e as áreas de café, cacau e feijão estão localizadas no município de Apuí (com coordenadas de 07° 12' 05" S e 59° 39' 35" W).

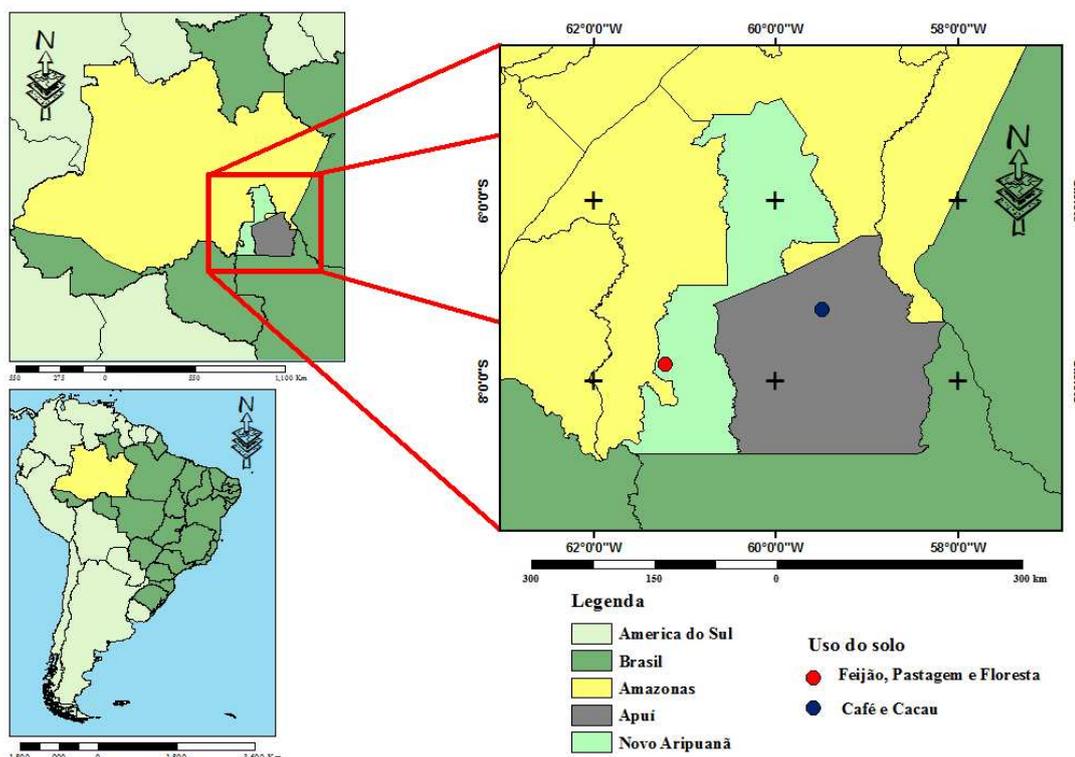
O clima da região segundo a classificação de Köppen é tipo Am (chuvas do tipo monção), caracterizado por apresentar um período seco de pequena duração, com precipitações entre 2.250 e 2.750 mm, temperaturas médias anuais variando entre 25 e 27 °C e umidade relativa do ar entre 85 e 90% (ALVARES *et al.*, 2014).

Com relação à geologia, a área de estudo está localizada sobre saprolitos de Granitos Rondonianos que se caracterizam por apresentar muscovita, biotita, adamelitos e granodioritos, de origem intrusiva cratogênica, em forma de “stocks” e batólitos. Os solos presentes na região são os Latossolos Amarelos e Latossolos Vermelho-Amarelos (BRASIL, 1978). A vegetação característica desta região é a floresta tropical densa constituída por árvores adensadas e multiestratificadas entre 20 a 50 metros de altura, com clima úmido, elevadas temperaturas e alta precipitação.

Foram selecionadas cinco áreas de estudos sobre terra preta arqueológica, sob os seguintes usos: (1) Floresta (Mata amazônica) preservada por mais de vinte e cinco anos; (2) Pastagem com brachiaria (*Brachiaria brizanta*); (3) Feijão guandu (*Cajanus cajan*), ressaltando que as duas últimas foram cultivadas com milho, feijão e melancia, durante os últimos 25 anos;

(4) Cacau (*Theobroma cacao*), vem sendo cultivada há quatorze anos com essa cultura e, nos primeiros seis anos abrigou os cultivos de arroz, milho, feijão e melancia; (5) Café (*Coffea arabica*) vem sendo cultivada há seis anos. Os solos nestas áreas, nos municípios de Novo Aripuanã e Apuí, foram classificados como Latossolo Amarelo eutrófico argissólico (SANTOS et al., 2018).

Figura 6. Localização da área de estudo com Terra Preta Arqueológica sob diferentes usos, nos municípios de Novo Aripuanã e Apuí, Amazonas, Brasil.



11.2. Coleta de solo para determinação dos atributos físicos e químicos do solo

Em cada área, foram delimitadas malhas amostrais com as dimensões 80 x 56, 80 x 56, 80 x 56, 100 x 42 e 88 x 42 m, em cada malha foram coletados 88 pontos amostrais distribuídos em espaçamentos variáveis de 8 x 8, 8 x 8, 8 x 8, 10 x 6 e 8 x 6 m, para as áreas com floresta, pastagem, feijão, café e cacau, respectivamente.

A coleta de solo foi realizada nos pontos de cruzamento da malha, nestes foram abertas pequenas trincheiras onde foram coletadas três amostras nas profundidades de 0,00-0,05, 0,05-0,10 e 0,10-0,20 m, totalizando 264 amostras de solo por malha. Após a coleta o material foi devidamente acondicionado, identificado e transportado ao laboratório onde foram realizadas as análises.

11.3. Atributos físicos do solo

A análise granulométrica foi realizada pelo método da pipeta com solução de Hidróxido de Sódio 0,1 (NaOH) com agente dispersante e agitação mecânica em aparato de alta rotação por 15 minutos (TEIXEIRA *et al.*, 2017). Para as determinações da densidade do solo (Ds), macroporosidade (MaP), microporosidade (MiP) e volume total de poros (VTP), as amostras foram coletadas com anéis volumétricos e saturadas por meio da elevação gradual de uma lâmina de água. Após atingirem o equilíbrio em um potencial matricial de 0,006 MPa, foram novamente pesadas e, em seguida, foram feitas as medidas da resistência do solo à penetração (RP), utilizando-se um penetrógrafo eletrônico modelo MA-933, marca Marconi. Posteriormente, as amostras foram levadas à estufa a 105 °C para determinação da Ds e VTP, pelo método do anel volumétrico e, a MaP foi determinada pela diferença entre VTP e MiP (TEIXEIRA *et al.*, 2017).

A umidade volumétrica foi obtida pela relação entre a umidade gravimétrica e a densidade do solo (TEIXEIRA *et al.*, 2017). A separação e estabilidade dos agregados, diâmetro médio geométrico (DMG) e diâmetro médio ponderado (DMP), foram determinadas segundo Kemper e Chepil (1965), com modificações, nas seguintes classes de diâmetro: >2,0 e < 2,0 mm.

11.4. Atributos químicos do solo

Foram determinados os valores de acidez ativa (pH em H₂O), cálcio (Ca²⁺), magnésio (Mg²⁺), potássio (K⁺), fósforo disponível (P), alumínio trocável (Al³⁺), acidez potencial (H⁺+Al³⁺) e carbono orgânico (CO). Para realizar a caracterização química as amostras foram secas ao ar, destorroadas e passadas em peneira com malha de 2,0 mm de abertura. O cálcio, magnésio e alumínio trocáveis foram extraídos por KCl, o potássio trocável e fósforo disponível, foram extraídos por Mehlich-1 e determinados por fotometria de chama e calorimetria, respectivamente. A acidez potencial (H⁺+Al³⁺) foi extraída com solução tamponada a pH 7,0 de acetato de cálcio utilizando-se metodologia proposta por Teixeira *et al.* (2017).

O carbono total foi determinado pelo método de Walkley-Black (1934) modificado por Yeomans e Bremner (1988) e, o estoque de carbono foi obtido pela fórmula proposta por Veldkamp (1994) e posteriormente corrigido pela fórmula matemática apresentada por Sisti *et al.* (2004).

11.5. Descrição do banco de dados

O conjunto de dados original foi composto por 22 atributos, sendo 21 atributos preditivos, formados por 01 atributo relacionado ao uso do solo, 13 atributos físicos do solo, 07

atributos químicos do solo e, um atributo resposta ou atributo meta, que se refere ao estoque de carbono do solo que é o alvo da classificação (Tabela 1).

Tabela 9. Atributos físicos e químicos de Terras Pretas Arqueológicas naturais e transformadas, que compõe o banco de dados.

Tipo de atributo	Descrição	Unidade	Tipo
Uso do solo	Floresta, feijão, pastagem, café e cacau	-	Preditivo
Físico	Areia	g kg ⁻¹	Preditivo
	Silte	g kg ⁻¹	Preditivo
	Argila	g kg ⁻¹	Preditivo
	Umidade volumétrica	kg kg ⁻¹	Preditivo
	Resistência do solo à penetração	MPa	Preditivo
	Densidade do solo	kg dm ⁻³	Preditivo
	Macroporosidade	m ³ m ⁻³	Preditivo
	Microporosidade	m ³ m ⁻³	Preditivo
	Volume total de poros	%	Preditivo
	Diâmetro médio geométrico	mm	Preditivo
	Diâmetro médio ponderado	mm	Preditivo
	Agregados > 2,0 mm	mm (%)	Preditivo
	Agregados < 2,0 mm	mm (%)	Preditivo
Químico	pH	H ₂ O	Preditivo
	Acidez potencial	mmol _c dm ⁻³	Preditivo
	Alumínio	mmol _c dm ⁻³	Preditivo
	Fósforo	mg dm ⁻³	Preditivo
	Potássio	mmol _c dm ⁻³	Preditivo
	Cálcio	mmol _c dm ⁻³	Preditivo
	Magnésio	mmol _c dm ⁻³	Preditivo
	Estoque de carbono no solo	Mg ha ⁻¹	Resposta

O agrupamento dessas variáveis compõe um banco de dados com 22 atributos, provenientes de 88 pontos de coleta (instâncias) para cada cultura (05 = floresta, pastagem, feijão, café e cacau) em 03 profundidades diferentes (0,00-0,05, 0,05-0,10 e 0,10-0,20 m).

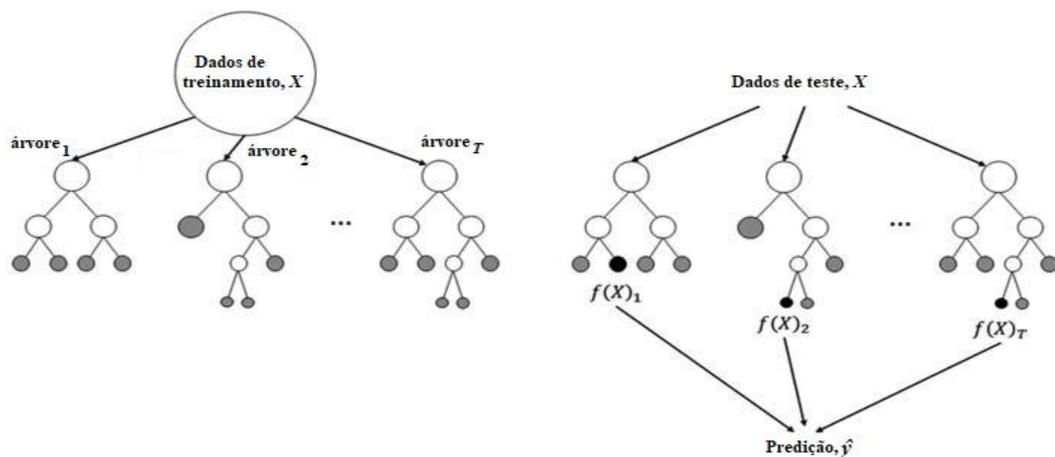
11.6. Indução de *Random Forest*

Neste caso, foi utilizada uma técnica que visa trabalhar a geração de um modelo considerando a variável resposta de tipo numérico. O modelo é composto de uma combinação de predições de árvores de decisão, de tal modo que cada árvore depende dos valores de um vetor independente amostrado de forma aleatória e com a mesma distribuição para todas as árvores da floresta (BREIMAN, 2001).

Consiste em um conjunto de árvores de decisão combinadas para resolver problemas de regressão. O *Random Forest* treina cada árvore de forma independente, usando

uma amostra aleatória dos dados. Essa aleatoriedade ajuda a tornar o modelo mais robusto do que uma única árvore de decisão e menos propenso a se adaptar aos dados de treinamento. Desse modo é gerada uma lista dos atributos mais importantes no desenvolvimento da floresta, que são determinados pela importância acumulada do atributo nas divisões dos nós de cada árvore da floresta (JAMES *et al.*, 2013).

Figura 7. Diagrama conceitual do algoritmo Random Forest.



Adaptado de Mennit *et al.* (2014).

Na Figura 5, observa-se que, à esquerda, as árvores são treinadas independentemente pelo particionamento recursivo binário de uma amostra inicial dos dados de entrada, X . Na direita, os dados de teste são trabalhados em cada árvore e a estimativa de resposta é a média sobre todas as previsões individuais na floresta (MENNIT *et al.*, 2014).

11.7. Validação do modelo de regressão

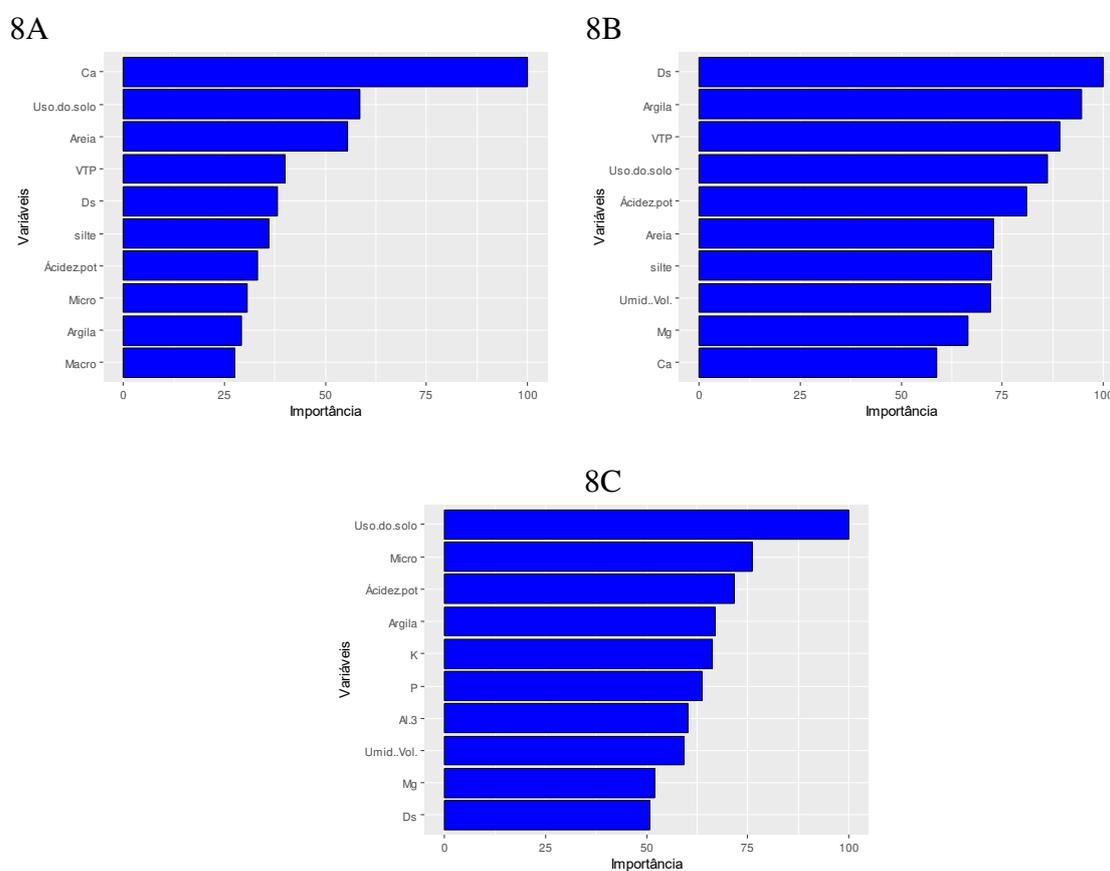
Para validação do modelo gerado pelo algoritmo *Random Forest* foram utilizados 75% dos dados para valores observados e 25% para testes, sendo os resultados expressos graficamente por meio da correlação entre os dados observados e preditos. A precisão do modelo foi avaliada por meio do coeficiente de determinação (R^2).

12. Resultados

Na profundidade de 0,00-0,05 m, observa-se que entre os atributos mais importantes, o atributo Ca (100%) que sobressai dentre o grupo e, os atributos uso do solo e areia apresentaram ordem de importância de mais de 50% na escala do modelo *Random Forest* para explicar o estoque de carbono no solo em TPA com diferentes usos (Figura 7A). Na profundidade 0,05-0,10 m, verifica-se que são 8 variáveis que apresentam ordem de importância maior que 70%, destacando-se a densidade do solo (Ds) e argila com ordens

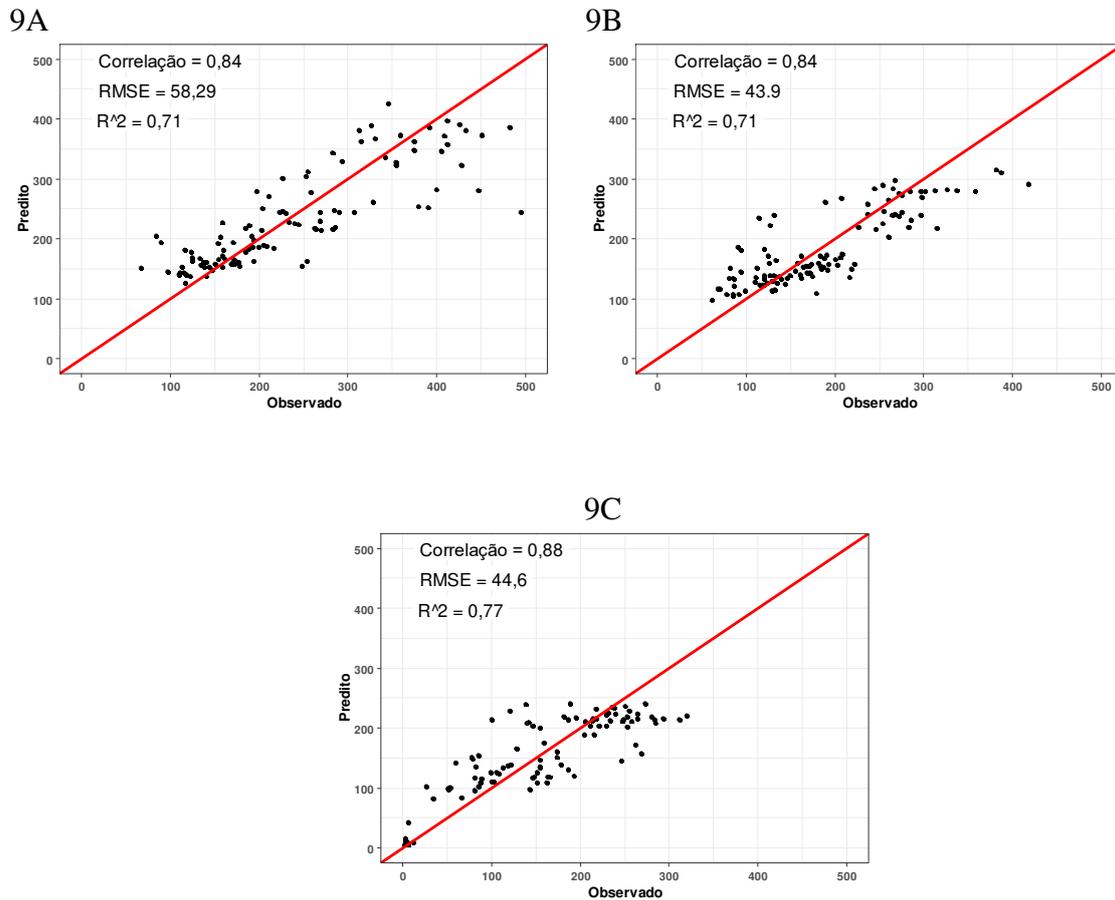
superiores a 95% (Figura 7B). Na profundidade 0,10-0,20 m, o atributo uso do solo (100%) que sobressai dentre outros 7 atributos com ordem de importância superior a 60% (Figura 7C).

Figura 8 Ordem de importância dos atributos selecionados para uso no modelo gerado pelo algoritmo de Random Forest para explicar o estoque de carbono no solo em Terra Preta Arqueológica, (8A) 0,00-0,05 m, (8B) 0,05-0,10 m e (8C) 0,10-0,20 m. Ds = densidade do solo (kg dm^{-3}); Macro = macroporosidade ($\text{m}^3 \text{ m}^{-3}$); Micro = microporosidade ($\text{m}^3 \text{ m}^{-3}$); VTP = volume total de poros ($\text{m}^3 \text{ m}^{-3}$); Umidade = umidade volumétrica ($\text{kg}^1 \text{ kg}^{-1}$); Acidez = acidez potencial ($\text{mmol}_c \text{ dm}^{-3}$); Ca = cálcio ($\text{mmol}_c \text{ dm}^{-3}$); Mg = magnésio ($\text{mmol}_c \text{ dm}^{-3}$); K = potássio ($\text{mmol}_c \text{ dm}^{-3}$); P = fósforo (mg dm^{-3}); Al^{3+} = alumínio ($\text{mmol}_c \text{ dm}^{-3}$).



Os modelos de validação usando $\frac{1}{4}$ dos dados observados, apresentam os resultados de coeficiente de determinação com $R^2 = 0,71$, $0,71$ e $0,77$, entre os dados observados e os preditos, para as profundidades de 0,00-0,05, 0,05-0,10 e 0,10-0,20 m, respectivamente, demonstrando o alto poder de precisão dos modelos gerados (Figura 9A, 9B e 9C).

Figura 9 Validação do modelo de regressão, os eixos x e y representam, respectivamente, os valores observados e previstos do estoque de carbono no solo em Terra Preta Arqueológica, (7A) 0,00-0,05 m, (78) 0,05-0,10 m e (7C) 0,10-0,20 m. Corr = correlação; RMSE = Root-Mean-Square Error / distância média quadrática mínima; R²: R quadrado / coeficiente de determinação.



13. Discussões

A importância do atributo Ca, na profundidade de 0,05-0,10 m, é compatível com informação apresentada em diferentes pesquisas, que descrevem que os solos da Amazônia geralmente são ácidos e com baixa disponibilidade de nutrientes. Entretanto, são encontrados no mesmo ambiente, manchas de solos escuros, com elevados valores de matéria orgânica e nutrientes (fósforo e cálcio), além de fragmentos cerâmicos arqueológicos (COSTA *et al.*, 2004; COSTA *et al.*, 2009; BARROS *et al.*, 2012; AQUINO *et al.*, 2016).

Em relação ao estoque de carbono, encontra-se na literatura que as TPA são caracterizadas pela elevada fertilidade e estoques de carbono (TEIXEIRA *et al.*, 2009). O cálcio, em forma de carbonato de cálcio (CaCO₃), pode ser um elemento chave na estabilização do carbono, conforme mostrado nos modelos teóricos apresentados por Archanjo *et al.* (2013).

É possível também explicar a importância do atributo Ca no fato que os solos de TPA apresentem altas concentrações de fósforo ligado ao cálcio (P-Ca) e, segundo Souza *et al.* (2009), o enriquecimento do solo de TPA com a forma de P-Ca está relacionada com a presença da apatita biogênica (fosfatos de cálcio), que é oriunda de ossos de animais e espinhas de peixes (SCHAEFER *et al.*, 2004).

Conforme Woods (2009), a cor enegrecida ocorre em virtude do conteúdo elevado de matéria orgânica, dos teores de carbonatos de cálcio e da concentração e estado da oxidação do ferro e do manganês, também influenciada por presença de carvões, cinzas e materiais terrosos oxidados. Santos *et al.* (2018) encontraram valores do Ca^{2+} trocável mais elevados nos horizontes superficiais, o que está de acordo com Steiner *et al.* (2004), os quais afirmaram que as TPA exibem elevados níveis de nutrientes, principalmente de Ca e P.

Na profundidade de 0,05-0,10 m, destaca-se o atributo densidade do solo, o qual já foi descrito como um dos indicadores físicos que têm sido usados com frequência para examinar os impactos causados pelos sistemas de manejo (RAMOS *et al.*, 2010; PEZARICO *et al.*, 2013). De igual modo, Alho *et al.* (2014) relatam que alguns atributos físicos são mais sensíveis às mudanças que outros, destacando dentre estes atributos a densidade do solo, teor de carbono orgânico e estoque de carbono, o que oferece uma visão da relação que existe entre estes atributos do solo. Conforme concluído por Silva *et al.* (2016), verificaram que a densidade do solo se apresentou dependente da quantidade de carbono orgânico e estoque de carbono nos solos de TPA.

Conforme relatado na pesquisa de Barros *et al.* (2016), pode estabelecer a relação de densidade do solo (Ds) e o conteúdo de matéria orgânica no solo, a partir da relação direta da Ds com o volume de poros, que por sua vez, foram positivamente relacionados ao conteúdo de matéria orgânica no solo. Por sua vez, Pignataro Netto *et al.* (2009), estudando atributos do solo em um Latossolo Vermelho-Amarelo sob pastagens com diferentes históricos de uso, verificaram que os atributos físicos que mais influenciam na redução da qualidade do solo foram a Ds, porosidade total e resistência do solo à penetração. A baixa Ds na TPA, segundo Steinbeiss *et al.* (2009), é resultante dos elevados teores de carbono orgânico e de intensa atividade biológica (fauna e raízes), que constroem canais, cavidades e galerias.

Sendo assim, a Ds que é um dos atributos que sofre maiores alterações se submetidos a ações antrópicas, pode ser considerado um dos indicadores da caracterização de TPA e, o uso de atributos do solo identificadores de diferentes ambientes se constitui em uma ferramenta fundamental para direcionar práticas que reduzam o depauperamento do solo (OLIVEIRA *et al.*, 2015).

A capacidade das TPA de manter o alto teor de carbono orgânico ocorre, possivelmente, devido às características químicas e da resistência do material à decomposição microbiana, com isso seus efeitos atuam sobre a agregação do solo, levando a formação de agregados estáveis e atuando sobre o armazenamento de água, porosidade e densidade do solo (MOREIRA, 2007; SILVA *et al.*, 2016).

Finalmente, na profundidade de 0,10-0,20 m, em que o modelo determinou o atributo uso do solo como o atributo mais importante, é possível encontrar na literatura informação que demonstra a relação desse atributo com o estoque de carbono do solo. De acordo com os dados obtidos, verificou-se que o estoque de carbono foi afetado de modo significativo pelos sistemas de uso e manejo do solo em cinco diferentes sistemas estudados: cacau, café, feijão, pastagem e floresta, em concordância com os dados obtidos por Campos *et al.* (2016) que descrevem que os sistemas de uso e manejo dos solos são fatores que podem influenciar no estoque de carbono.

Também em concordância, Cunha *et al.* (2017), avaliaram os atributos físicos e estoque de carbono do solo em áreas de Terra Preta Arqueológica na região Amazônia e, observaram que tanto o carbono orgânico quanto o estoque de carbono apresentaram alterações de acordo com os respectivos usos do solo.

Quando comparado com solos sob diferentes usos, Oliveira *et al.* (2015), destacam que a TPA com sua alta fertilidade natural, diferiu dos demais usos (mandioca, pastagem e cana-de-açúcar). Os estoques de carbono no solo variam em função do tipo de solo, profundidade, clima, bioma e, principalmente, uso e manejo da terra (PARRON *et al.*, 2015). Entretanto, práticas adequadas de manejo, que visam a manutenção ou mesmo o acúmulo de carbono no sistema solo-planta, podem atenuar os efeitos do aquecimento global (CARVALHO *et al.*, 2010).

Assim, mudanças no uso da terra alteram os processos biogeoquímicos do solo, com reflexos no estoque de carbono e no fluxo de gases entre o solo e a atmosfera. Dependendo das características da área e do sistema de manejo adotado, essas alterações podem representar uma mudança no papel do solo como reservatório de carbono (ROSA *et al.*, 2014).

14. Conclusões

O teor de cálcio, a densidade do solo e o uso do solo foram as variáveis mais importantes que afetaram o estoque de carbono em Terra Preta Arqueológica sob diferentes usos, nas três profundidades estudadas (0,00-0,05 m, 0,05-0,10 m e 0,10-0,20 m).

A utilização do algoritmo *Random Forest* viabilizou a geração de um modelo que apresentou um ajuste de $R^2 = 0,71$, $0,71$ e $0,77$, nas três profundidades estudadas, 0,00-0,05, 0,05-0,10 e 0,10-0,20 m, respectivamente.

15. Considerações finais

Os resultados do estudo foram considerados satisfatórios e a hipótese da pesquisa foi confirmada, porém, alguns outros atributos poderiam ser considerados em análises futuras como a precipitação, temperatura ambiental, dados geoespaciais, atributos microbiológicos entre outros, que contribuam com informação relevante para conseguir compreender ainda melhor como ocorre o fenômeno de estoque de carbono em Terra Preta Arqueológica.

As técnicas de mineração de dados abordadas neste trabalho foram de fácil aplicação e, a interpretação dos resultados foram dedutivas e de leitura simples conforme descrito na literatura, o que confirma o potencial que têm essas técnicas para lidar com um grande número de dados independente dos atributos a serem predito seja discreto ou contínuo.

Considera-se também, importante o fato de dar continuidade a coleta de dados ao longo do tempo, com a finalidade de compreender com maior detalhe o comportamento temporal do estoque de carbono em Terra Preta Arqueológica. Permitindo desse modo obter informação que auxilie na tomada de decisão sobre possíveis práticas de manejo com potencial para conservar ou aumentar o estoque de carbono do solo em Terra Preta Arqueológica.

16. Referências bibliográficas

- ACOSTA, J. A. A.; AMADO, T. J. C.; SILVA, L. S.; SANTI, A.; WEBER, M. A. Decomposição da fitomassa de plantas de cobertura e liberação de nitrogênio em função da quantidade de resíduos aportada ao solo sob sistema plantio direto. **Ciência Rural**, v.44, n.5, p.801-809, 2014.
- ALHO, L. C.; CAMPOS, M. C. C.; SILVA, D. M. P.; MANTOVANELLI, B. C.; SOUZA, Z. M. Variabilidade espacial de estabilidade de agregados e estoque de carbono em Cambissolo e Argissolo. **Pesquisa Agropecuária Tropical**, v.44, n.3, p.246-254, 2014.
- ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; GONÇALVES, J. L. M.; SPAROVEK, K. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, Stuttgart, v.22, n.6, p.711-728, 2014.
- AQUINO, R. E.; MARQUES JUNIOR, J.; CAMPOS, M. C. C.; OLIVEIRA, I. A.; SOUZA BAHIA, A. S. R.; SANTOS, L. A. C. Characteristics of color and iron oxides of clay fraction in Archeological Dark Earth in Apuí region, southern Amazonas. **Geoderma**, v.262, p.35-44, 2016.
- ARCHANJO, B. S.; ARAUJO, J. R.; SILVA, A. M.; CAPAZ, R. B.; FALCÃO, N. P. S.; JORIO, A.; ACHETE, C. A. Chemical analysis and molecular models for calcium-oxygen-carbon interactions in Black Carbon Found in fertile Amazonian Anthrosols. **Environmental Science & Technology**, v.48, n.13, p.7445-7452, 2013.
- ARROYO-KALIN, M. **Las tierras antrópicas amazónicas: algo más que un puñado de tierra**. En: Las Siete Maravillas de la Amazonía precolombina. ROSTAIN, S.; BETANCOURT, C.J. (Eds.). La Paz: EIAA/BAS/Plural Publicaciones. p.99-117, 2017.
- BARRETO P. A. B.; GAMA-RODRIGUES, E. F.; GAMA-RODRIGUES A. C. Carbono das frações da matéria orgânica em solos sob plantações de eucalipto de diferentes idades. **Scientia Forestalis**, v.42, n.104, p.581-590, 2014.
- BARROS, K.R. M.; LIMA, H. V.; RODRIGUES, S.; KERN, D. C. Distribuição da porosidade textural e estrutural em solos de Terra Preta Arqueológica. **Revista Ciência Agronômica**, v.47, n.4, p.609-615, 2016.
- BARROS, K. R. M.; LIMA, H. V.; CANELLAS, L. P.; KERN, D. C. Fracionamento químico da matéria orgânica e caracterização física de Terra Preta de Índio. **Revista de Ciências Agrárias**, v.55, n.1, p.44-51, 2012.
- BEN-GAL, I. **Outlier detection**. In: Data mining and knowledge discovery handbook. Boston: Springer, p.131-146, 2005.
- BERNOUX, M.; CERRI, C. C.; VOLKOFF, B.; CARVALHO, M. C. S.; FELLER, C.; CERRI, C. E. P.; ESCHENBRENNER, V.; PICCOLO, M. C.; FEIGL, B. Gases do efeito estufa e estoques de carbono nos solos: inventário do Brasil. **Cadernos de Ciência & Tecnologia**, v.22, n.1, p.235-246, 2005.
- BERSON, A.; SMITH, S.; THEARLING, K. **Building data mining applications for CRM**. New York: McGraw-Hill, 449 p. 2000.

BHARALI, A.; BARUAH, K. K.; GOGOI, N. Changes in organic carbon pool in a tropical soil planted to rice in relation to photosynthetic carbon fixation. **Australian Journal of Crop Science**, v.10, n.8, p.1197-1206, 2016.

BRASIL. Ministério das Minas e Energia. Projeto **Radambrasil**, folha SB. 20, Purus. Rio de Janeiro: Departamento Nacional da Produção Mineral, 1978.

BREIMAN, L. Random forests. **Machine learning**, v.45, n.1, p.5-32, 2001.

BREIMAN, L.; FRIEDMAN, J. H.; OLSHEN, R. A.; STONE, C. J. **Classification and regression trees**. Monterey: Wadsworth and Brooks, 358 p, 1984.

CAMPOS, M. C. C.; ALHO, L. C.; SILVA, D. A. P.; SILVA, M. D. R.; CUNHA, J. M.; SILVA, D. M. P. Distribuição espacial do efluxo de CO₂ em área de terra preta arqueológica sob cultivo de cacau e café no município de Apuí, AM, Brasil. **Revista Ambiente & Água**, v.11, n.4, p.788-798, 2016.

CAMPOS, M. C. C.; SANTOS, L. A. C.; SILVA, D. M. P.; MANTOVANELLI, B. C.; SOARES, M. D. R. Caracterização física e química de terras pretas arqueológicas e de solos não antropogênicos na região de Manicoré. **Revista Agro@mbiente On-line**, v.6, n.2, p.102-109, 2012a.

CAMPOS, M. C. C.; RIBEIRO, M. R.; SOUZA JÚNIOR, V. S. DE; RIBEIRO FILHO, M. R.; ALMEIDA, M. C. Topossequência de solos na transição campos naturais-floresta na região de Humaitá, Amazonas. **Acta Amazônica**, v.42, n.3, p.387-398, 2012b.

CAMPOS, M. C. C.; RIBEIRO, M. R.; SOUZA JÚNIOR, V. S.; RIBEIRO FILHO, M. R.; SOUZA, R. V. C. C.; ALMEIDA, M. C. Caracterização e classificação de terras pretas arqueológicas na Região do Médio Rio Madeira. **Bragantia**, v.70, n.3, p.598-609, 2011.

CARVALHO, J. L. N.; AVANZI, J. C.; SILVA, M. L. N.; MELLO, C. R. D.; CERRI, C. E. P. Potencial de sequestro de carbono em diferentes biomas do Brasil. **Revista Brasileira de Ciência do Solo**, v.34, n.2, p.277-290, 2010.

COLOMBO, G. A.; LOPES, M. B. S.; DOTTO, M. C.; CAMPESTRINI, R.; LIMA, S. O. Atributos físicos de um Latossolo Vermelho-Amarelo distrófico sob diferentes sistemas de manejo no cerrado tocantinense. **Campo Digital**, v.12, n.1, p.21-29, 2017.

CORADO NETO, F. C.; SAMPAIO, F. M. T.; VELOSO, M. E. C.; MATIAS, S. S. R.; ANDRADE, F. R.; LOBATO, M. G. R. Variabilidade espacial dos agregados e carbono orgânico total em Neossolo Litólico eutrófico no município de Gilbués, PI. **Revista de Ciências Agrárias**, v.8, n.1, p.75-83, 2015.

COSTA, J. A.; KERN, D. C.; COSTA, M. L.; RODRIGUES, T. E.; KÄMPF, N.; LEHMANN, J.; FRAZÃO, F. J. L. **Geoquímica das Terras Pretas Amazônicas**. In: TEIXEIRA, W. G.; KERN, D. C.; MADARI, B. E.; LIMA, H. N.; WOODS, W. As terras pretas de índio da Amazônia: sua caracterização e uso deste conhecimento na criação de novas áreas. Manaus: Embrapa Amazônia Ocidental, p.162-171, 2009.

COSTA, M. L.; KERN, D. C.; PINTO, A. H. E.; SOUZA, J. R. T. The ceramic artifacts in Archaeological Black Earth (Terra Preta) from lower Amazon region, Brazil: mineralogy. **Acta amazônica**, v.34, n.2, p.165-78, 2004.

CUNHA, E.Q.; STONE L.F.; MOREIRA, J.A.A.; FERREIRA, E.P.B.; DIDONET, A.D.; LEANDRO, W.M. Sistemas de preparo do solo e cultura de cobertura na produção orgânica de feijão e milho. **Revista Brasileira de Ciência do Solo**, v. 35, p. 589-602, 2011.

CUNHA, J. F.; CASARIN, V.; PROCHNOW, L. I. Balanço de nutrientes na agricultura brasileira. **Informações Agronômicas**, v.130, p.1-11, 2010.

CUNHA, J. M.; GAIO, D. C.; CAMPOS, M. C. C.; SOARES, M. D. R.; SILVA, D. M. P.; LIMA, A. F. L. Atributos físicos e estoque de carbono do solo em áreas de Terra Preta Arqueológica da Amazônia. **Revista Ambiente & Água**, v.12, n.2, p.263-281, 2017.

CUNHA, T. J. F.; MADARI, B. E.; BENITES, V. M.; CANELLAS, L. P.; NOVOTNY, E. H.; MOUTTA, R. O.; TROMPOWSKY, P. M.; SANTOS, G. A. Fracionamento químico da matéria orgânica e características de ácidos húmicos de solos com horizonte A antrópico da Amazônia (Terra Preta). **Acta Amazônica**, v.37, n.1, p.91-98, 2007.

DAS, B.; CHAKRABORTY, D.; SINGH, V. K.; AHMED, M.; SINGH, A. K.; BARMAN, A. Evaluating fertilization effects on soil physical properties using a soil quality index in an intensive rice-wheat cropping system. **Pedosphere**, v.26, n.6, p.887-894, 2016.

FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. C. P. L. F. **Inteligência artificial: uma abordagem de aprendizado de máquina**. Rio de Janeiro: LTC, 375 p. 2011.

FARHATE, C. V. V.; SOUZA, Z. M.; OLIVEIRA, S. R. D. M.; CARVALHO, J. L.; LA SCALA, N.; SANTOS, A. P. G. Classification of soil respiration in areas of sugarcane renewal using decision tree. **Scientia Agricola**, v.75, n.3, p.216-224, 2018.

FAYYAD, U.; PIATETSKI-SHAPIRO, G.; SMYTH, P. From data mining to knowledge discovery: an overview. In: *Advances in knowledge discovery & data mining*. Menlo Park: **American Association for Artificial Intelligence**, p.1-34, 1996.

FELLER, C.; BEARE, M. H. Physical control of soil organic matter dynamics in the tropics. **Geoderma**, v.79, n.1-4, p.69-116, 1997.

FORTE, A.; FAGNANO, M.; FIERRO, A. Potential role of compost and green manure amendment to mitigate soil GHGs emissions in mediterranean drip irrigated maize production systems. **Journal of Environmental Management**, v.192, n.1, p.68-78, 2017.

FOSTER, D.; SWANSON, F.; ABER, J.; BURKE, I.; BROKAW, N.; TILMAN, D.; KNAPP, A. The importance of land-use legacies to ecology and conservation. **BioScience**, v.53, n.1, p.77-88, 2003.

FREITAS, A. A. **Data mining and knowledge discovery with evolutionary algorithms**. New York, Springer-Verlag, 264 p, 1998.

FREIXO, A. A.; MACHADO P. L. O. A.; GUIMARÃES C. M.; SILVA C. A.; FADIGAS, F. S. Estoques de carbono e nitrogênio e distribuição de frações orgânicas de latossolo do cerrado sob diferentes sistemas de cultivo. **Revista Brasileira de Ciência do Solo**, v.26, n.2, p.425-464, 2002.

GALVÃO, J. R.; FERNANDES, A. R.; PINHEIRO, D. P.; MELO, N. C. Adubação potássica em híbridos de sorgo forrageiro cultivados em sistemas de manejo do solo na Amazônia oriental. **Revista Caatinga**, v.28, n.4, p.70-79, 2015.

GLASER, B. Prehistorically modified soils of central Amazonia: A model for sustainable agriculture in the twenty-first century. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v.362, p.187-196, 2007.

GLASER, B.; BIRK, J. J. State of the scientific knowledge on properties and genesis of Anthropogenic Dark Earths in Central Amazonia (Terra Preta de Índio). **Geochimica et Cosmochimica Acta**, v.82, n.1, p.39-51, 2012.

GLASER, B.; HAUMAIER, L.; GUGGENBERGER, G.; ZECH, W. The “Terra Preta” phenomenon: a model for sustainable agriculture in the humid tropics. **Naturwissenschaften**, v.88, n.1, p.37-41, 2001.

GOLDSCHMIDT, R.; PASSOS, E. **Data mining: um guia prático**. Rio de Janeiro: Elsevier, 2005. 261p

HALL, M. A. **Correlation-based feature selection for discrete and numeric class machine learning**. In: Proceedings of the Seventeenth International Conference on Machine Learning, p.359-366, 2000.

HAN, J.; KAMBER, M.; PEI, J. **Data mining: concepts and techniques**. 3. Ed. San Francisco: Morgan Kaufmann, p.702, 2011.

HAND, D. J. Principles of data mining. **Drug safety**, v.30, n.7, p.621-622, 2007.

HESTERBERG, D.; VERGÜTZ, L. Radiação síncrotron na agricultura e ciência do solo. **Ciência e Cultura**, v.69, n.3, p.52-55, 2017.

HICKMANN, C.; COSTA, L. M. Estoque de carbono no solo e agregados em Argissolo sob diferentes manejos de longa duração. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.16, n.10, p.1055-1061, 2012.

IPPOLITO, J. A., LAIRD, D. A., BUSSCHER, W. J. Environmental benefits of biochar. **Journal of environmental quality**, v.41, n.4, p.967-972, 2012.

JAMES, G.; HASTIE, T.; TIBSHIRANI, R. **An introduction to statistical learning: with applications in R**. London: Ed. Springer, 429 p. 2013.

KÄMPF, N.; KERN, D. C. **O solo como registro da ocupação humana pré-histórica na Amazônia**. In: VIDALTORRADO, P.; ALLEONI, L. R. F.; COOPER, M.; SILVA, A. P.; CARDOSO, E. J. (Eds.). Tópicos em ciência do solo. Viçosa: **Sociedade Brasileira de Ciência do Solo**, p.277-320, 2005.

KATO, E.; RAMOS, M. L. G.; VIEIRA, D. F. A.; MEIRA, A. D.; MOURÃO, V. C. Propriedades físicas e teor de carbono orgânico de um Latossolo Vermelho-Amarelo do cerrado, sob diferentes coberturas vegetais. **Bioscience Journal**, v.26, n.5, p.732-738, 2010.

KEMPER, W.D.; CHEPIL, W.S. Size distribution of aggregates. In: BLACK, C.A. (Ed.). **Methods of soil analysis**. Madison: **American Society Agronomy**, 1965. p.499-510.

KETTERINGS, Q. M.; BIGHAM, J. M.; LAPERCHE, V. Changes in soil mineralogy and texture caused by slash and burn fires in Sumatra, Indonesia. **Soil Science Society of America Journal**, v.64, p.1108-1117, 2000.

- LA SCALA JÚNIOR, N.; FIGUEIREDO, E. B.; PANOSSO, A. R. A review on soil carbon accumulation due to the management change of major Brazilian agricultural activities. **Brazilian Journal of Biology**, São Carlos, v.72, n.3, p.775-785, 2012.
- LAL, R. Carbon management in agricultural soils. **Mitigation and Adaptation Strategies for Global Change**, v.12, n.2, p.303-322, 2007.
- LAL, R. Soil carbon dynamics in cropland and rangeland. **Environmental Pollution**, v.116, n.3, p.353-362, 2002.
- LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics**, v.33, n.1, p.159-174, 1977.
- LAROSE, D. T. **Discovering knowledge in data: an introduction to data mining**. [S.l.]: John Wiley & Sons, 31 p. 2005.
- LEMOS, E. P.; STEINER, M. T. A.; NIEVOLA, J. C. Análise de crédito bancário por meio de redes neurais e árvores de decisão: uma aplicação simples de data mining. **Revista de Administração**, v.40, n.3, p.225-234, 2005.
- LIMA, E. S.; SOUZA, Z. M.; MONTANARI, R.; OLIVEIRA, S. R. M.; LOVERA, L. H.; FARHATE, C. V. V. Classification of the initial development of eucalyptus using data mining techniques. **Cerne**, v.23, n.2, p.201-208, 2017.
- LIU, H.; SHAH, S.; JIANG, W. On-line outlier detection and data cleaning. **Computers & Chemical Engineering**, v.28, n.9, p.1635-1647, 2004.
- MAIA, C. M. B. F. **Finos de carvão: fonte de carbono estável e condicionador de solos**. 1. (Ed.). Colombo: Embrapa Florestas, 36 p. 2010.
- MANGRICH, A. S.; MAIA, C. M. B. F.; NOVOTNY, E. H. Biocarvão - As terras pretas de índios e o sequestro de carbono. **Ciência Hoje**, v.47, p.48-52, 2011.
- MARCHINI, D. C.; LING, T. C.; ALVES, M. C.; CRESTANA, S.; SOUTO FILHO, S. N.; ARRUDA, O. G. Matéria orgânica, infiltração e imagens tomográficas de Latossolo em recuperação sob diferentes tipos de manejo. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.19, n.6, p.574-580, 2015.
- MARQUES, J. D. O.; LUIZÃO, F. J.; TEIXEIRA, W. G.; VITEL, C. M.; MARQUES, E. M. A. Soil organic carbon, carbon stock and their relationships to physical attributes under forest soils in central Amazonia. **Revista Árvore**, v.40, n.2, p.197-208, 2016.
- MARTINS, A. P. M.; DEBASTIANI, A. B.; PELISSARI, A. L.; MACHADO, S. D. A.; SANQUETTA, C. R. Estimativa do afilamento do fuste de araucária utilizando técnicas de inteligência artificial. **Floresta e Ambiente**, v.24, p.1-10, 2017.
- MASCARENHAS, A. R. P.; SCCOTI, M. S. V.; MELO, R. R.; CORRÊA, F. L. O.; SOUZA, E. F. M.; ANDRADE, R. A.; MÜLLER, M. W. Atributos físicos e estoques de carbono do solo sob diferentes usos da terra em Rondônia, Amazônia Sul-Occidental. **Pesquisa Florestal Brasileira**, v.37, n.89, p.19-27, 2017.
- MENNITT, D.; SHERRILL, K.; FRISTRUP, K. A geospatial model of ambient sound pressure levels in the contiguous United States. **The Journal of the Acoustical Society of America**, v.135, n.5, p.2746-2764, 2014.

MOITINHO, M. R.; PADOVAN, M. P.; PANOSSO, A. R.; TEIXEIRA, D. B., FERRAUDO, A. S.; LA SCALA JÚNIOR, N. On the spatial and temporal dependence of CO₂ emission on soil properties in sugarcane (*saccharum spp.*) production. **Soil and Tillage Research**, v.148, p.127-132, 2015.

MOREIRA, A. Fertilidade, matéria orgânica e substâncias húmicas em solos antropogênicos da Amazônia Ocidental. **Bragantia**, v.66, n.2, p.307-315, 2007.

MOTA JÚNIOR, P. C.; CAMPOS, M. C. C.; MANTOVANELLI, B. C.; FRANCISCON, U.; CUNHA, J. M. D. Variabilidade espacial de atributos físicos do solo em terra preta de índio sob cultivo de café Conilon. **Coffee Science**, v.12, n.2, p.260-271, 2017.

MOURA, E. G.; OLIVEIRA, A. K. C.; COUTINHO, G.; PINHEIRO, K. M.; AGUIAR, A. C. F. Management of a cohesive tropical soil to enhance rootability and increase the efficiency of nitrogen and potassium use. **Soil Use and Management**, v.28, n.3, p.370-377, 2012.

NEVES, E. G.; PETERSEN, J. B.; BARTONE, R. N.; SILVA, C. A. Historical and socio-cultural origins of Amazonian dark earth. In: **Amazonian Dark Earths**. Dordrecht: Springer, p. 29-50, 2003.

NIERO, L. A. C.; DECHEN, S. C. F.; COELHO, R. M.; DE MARIA, I. C. Avaliações visuais como índice de qualidade do solo e sua validação por análises físicas e químicas em um Latossolo Vermelho distroférico com usos e manejos distintos. **Revista Brasileira de Ciência do Solo**, v.34, n.4, p.1271-1282, 2010.

NOVOTNY, E. H.; HAYES, M. H. B.; MADARI, B. E.; BONAGAMBA, T. J.; AZEVEDO, E. R.; SOUZA, A. A.; SONG, G.; NOGUEIRA, C. M.; MANGRICH, A. S. Lessons from the Terra Preta de Índios of the Amazon Region for the utilisation of charcoal for soil amendment. **Journal of the Brazilian Chemical Society**, v.20, n.6, p.1003-1010, 2009.

NOVOTNY, E. H.; MADARI B. E.; MAIA C. M. B.; MANGRICH A. S. O potencial do biocarvão (carbono pirogênico) no sequestro de carbono; na ciclagem de nutrientes; no crescimento das plantas e no estímulo de processos microbiológicos. In Embrapa Florestas- Artigo em **Anais...** Viçosa: Sociedade Brasileira de Ciência do solo, 2012.

OLIVEIRA, I. A.; CAMPOS, M. C. C.; FREITAS, L.; SOARES, M. D. R. Caracterização de solos sob diferentes usos na região sul do Amazonas. **Acta amazônica**, v.45, n.1, p.1-12, 2015.

OMARI, R. A.; AUNG, H. P.; HOU, M.; YOKOYAMA, T.; ONWONA-AGYEMAN, S.; OIKAWA, Y.; FUJII, Y.; BELLINGRATH-KIMURA, S. D.. Influence of different plant materials in combination with chicken manure on soil carbon and nitrogen contents and vegetable yield. **Pedosphere**, v.26, n.4, p.510-521, 2016.

ORJUELA, H. B. El carbono orgánico del suelo y su papel frente al cambio climático. **Revista de Ciencias Agrícolas**, v.35, n.1, p.82-96, 2018.

PARRON, L. M.; RACHWAL, M. F. G.; MAIA, C. M. B. F. **Serviços ambientais em sistemas agrícola e florestais do Bioma Mata Atlântica**. In: PARRON, L. M.; RACHWAL, M. F. G.; MAIA, C. M. B. F. Estoques de carbono no solo como indicador de serviços ambientais. Brasília: Embrapa, 2015. p.71-83.

PEREIRA, A. A.; THOMAZ, E. L. Atributos químicos do solo em áreas sob diferentes sistemas de uso e manejo no município de Reserva-PR. **Caminhos de Geografia**, v.16, n.55, p.186-194, 2015.

PEZARICO, C. R.; VITORINO, A. C. T.; MERCANTE, F. M.; DANIEL, O. Indicadores de qualidade do solo em sistemas agroflorestais. **Revista de Ciências Agrônômicas**, v.56, n.1, p.40-47, 2013

PIGNATARO NETTO, I. T.; KATO, E.; GOEDERT, W. J. Atributos físicos e químicos de um Latossolo Vermelho-Amarelo sob pastagens com diferentes históricos de uso. **Revista Brasileira de Ciência do Solo**, v.33, n.5, p.1441-1448, 2009.

QUINLAN, J. R. C4.5: **Programs for Empirical Learning**. San Francisco, Califórnia: Morgan Kaufmann, p 302, 1993.

QUINLAN, J. R. Improved use of continuous attributes in C4.5. **Journal of Artificial Intelligence Research**, v.4, p.77-90, 1996.

RAIJ, B. V; ANDRADE, J.C.; CANTARELLA, H.; QUAGGIO, J.A. **Análise química para avaliação da fertilidade de solos tropicais**. Campinas: Instituto Agrônômico, 2001, 285 p.

RAMOS, F. T.; MONARI, Y. C.; NUNES, M. C. M.; CAMPOS, D. T. S.; RAMOS, D. T. Indicadores de qualidade em um Latossolo Vermelho-Amarelo sob pastagem extensiva no pantanal matogrossense. **Caatinga**, v.23, n.1, p.112-120, 2010.

REZENDE, S. O. **Sistemas inteligentes: fundamentos e aplicações**. Barueri: Editora Manole Ltda, 2003. 455 p.

RODRIGUEZ-GALIANO, V. F.; GHIMIRE, B.; ROGAN, J.; CHICA-OLMO, M.; RIGOL-SANCHEZ, J. P. An assessment of the effectiveness of a random forest classifier for land-cover classification. **Journal of Photogrammetry and Remote Sensing**, v.67, p.93-104, 2012.

ROSA, R.; SANO, E. E.; ROSENDO, J. D. Estoque de carbono em solos sob pastagens cultivadas na bacia hidrográfica do rio Paranaíba. **Sociedade & Natureza**, v.26, n.2, p.333-351, 2014.

ROSSETTI, K. V.; CENTURION, J. F. Estoque de carbono e atributos físicos de um Latossolo em cronosequência sob diferentes manejos. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.19, n.3, p.252-258, 2015.

ROSSI, C. Q.; PEREIRA, M. G.; GIÁCOMO, S. G.; BETTA, M.; POLIDORO, J. C. Frações orgânicas e índice de manejo de carbono do solo em Latossolo Vermelho sob plantio de soja no cerrado goiano. **Revista Brasileira de Ciências Agrárias**, v.7, n.2, p.233-241, 2012.

ROZANE, D. E.; CENTURION, J. F.; ROMUALDO, L. M.; TANIGUCHI, C. A. K.; TRABUCO, M.; ALVES, A. U. Estoque de carbono e estabilidade de agregados de um Latossolo Vermelho distrófico, sob diferentes manejos. **Bioscience Journal**, v.26, n.1, p.24-32, 2010.

SAIZ, G.; BIRD, M. I.; DOMINGUES, T.; SCHRODT, F.; SCHWARZ, M.; FELDPAUSCH, T. R.; VEENENDAAL, E.; DJAGBLETEY, G.; HIEN, F.; COMPAORE, H.; DIALLO, A.; LLOYD, J. Variation in soil carbon stocks and their determinants across a precipitation gradient in West Africa. **Global Change Biology**, v.18, n.5, p.1670-1683, 2012.

SANTOS, H. G.; JACOMINE, P.K.T.; ANJOS, L.H.C.; OLIVEIRA, V.A.; LUMBRERAS, J. F.; COELHO, M.R.; ALMEIDA, J.A.; ARAUJO FILHO, J.C.; OLIVEIRA, J.B.; CUNHA, T.J.F. **Sistema Brasileiro de Classificação de Solos**. 5ª Ed. Revista e Ampliada, Brasília: Embrapa, 353 p, 2018.

SANTOS, F. A. S.; PIERANGELI, M. A. P.; SILVA, F. L.; SERAFIM, M. E.; SOUSA, J. B.; OLIVEIRA, E. B. Dinâmica do carbono orgânico de solos sob pastagens em campos de murundus. **Scientia Agraria**, v.18, n.2, p.43-53, 2017.

SANTOS, L. A. C.; CAMPOS, M. C. C., AQUINO, R. E.; BERGAMIN, A. C., SILVA, D. M. P.; MARQUES JUNIOR, J.; FRANÇA, A. B. C. Caracterização de terras pretas arqueológicas no sul do estado do Amazonas. **Revista Brasileira de Ciência do Solo**, v.37, n.4, p.825-836, 2013.

SANTOS, L. A. C.; CAMPOS, M. C. C.; BERGAMIN, A. C.; SILVA, D. M. P.; MENDONÇA JÚNIOR, A. F. Caracterização física de seis sítios de terras pretas Arqueológicas na Região de Apuí-AM. **Revista Verde de Agroecologia e Desenvolvimento Sustentável**, v.6, n.4, p.167-174, 2012.

SCHAEFER, C. E. G. R.; LIMA, H. N.; GILKES, R. J.; MELLO, J. W. V. Micromorphology and electron microprobe analysis of phosphorus and potassium forms of an Indian Black Earth (IBE) Anthrosol from Western Amazonia. **Australian Journal of Soil Research**, v.42, p.401-409, 2004.

SCHARLEMANN, J. P. W.; TANNER, E. V. J.; HIEDERER, R.; KAPOS, V. Global soil carbon: understanding and managing the largest terrestrial carbon pool. **Carbon Management**, v.5, n.1, p.81-91, 2014.

SCHMIDT, M. J.; PY-DANIEL, A. R.; PAULA MORAES, C.; VALLE, R. B.; CAROMANO, C. F.; TEXEIRA, W. G.; SILVA, R. D. S. Dark earths and the human built landscape in Amazonia: a widespread pattern of anthrosol formation. **Journal of Archaeological Science**, v.42, p.152-165, 2014.

SHIMABUKURU, H. M. Visualização temporal em uma plataforma software extensível e adaptável. **Tese** - Instituto de Ciência Matemática e de Computação. Universidade de São Paulo, São Carlos, 2004.

SILVA, D. M. P.; CAMPOS, M. C. C.; ALHO, L. C.; CUNHA, J. M.; PAULA NETO, de. P. Variação espacial da estabilidade de agregados e estoque de carbono e terra preta arqueológica sob cultivo de cacau. **Revista Brasileira de Agroecologia**, v.11, n.3, p.179-187, 2016.

SILVA, A. K. T.; GUIMARÃES, J. T. F.; LEMOS, V. P.; COSTA, M. L.; KERN, D. C. Mineralogia e geoquímica de perfis de solo com Terra Preta Arqueológica de Bom Jesus do Tocantins, sudeste da Amazônia. **Acta Amazônica**, v.42, n.4, p.477-490, 2012.

SILVA, F. W. R.; LIMA, H.N.; TEIXEIRA, W. G.; MOTTA, M. B.; SANTANA, R. M. Caracterização química e mineralogia de solos antrópicos (Terras Pretas de Índio) na Amazônia Central. **Revista Brasileira de Ciência do Solo**, v.35, n.3, p.673-681, 2011.

SISTI, C. P.; SANTOS, H. P.; KOHHANN, R.; ALVES, B. J.; URQUIAGA, S.; BODDEY, R. M. Change in carbon and nitrogen stocks in soil under 13 years of conventional or zero tillage in southern Brazil. **Soil and Tillage Research**, v.76, n.1, p.39-58, 2004.

SOARES, M. D. R.; CAMPOS, M. C. C.; CUNHA, J. M.; WECKNER, F. C.; BRITO FILHO, E.; MANTOVANELI, B. C.; LEITE, A. F. L. Variabilidade espacial da estabilidade dos agregados e matéria orgânica do solo em terra preta arqueológica sob pastagem. **Gaia Scientia**, v.12, n.2, p.125-133, 2018.

SOARES, M. D. R.; CAMPOS, M. C. C.; SOUZA, Z. M.; BRITO, W. B. M.; FRANCISCON, U.; CASTIONE, G. A. F. Variabilidade espacial dos atributos físicos do solo em área de Terra Preta Arqueológica sob pastagem em Manicoré, AM. **Amazonian Journal of Agricultural and Environmental Sciences**, v.58, n.4, p.434-441, 2015.

SOMBROEK, W. I. M.; RUIVO, M. D. L.; FEARNSTIDE, P. M.; GLASER, B.; LEHMANN, J. **Amazonian dark earths as carbon stores and sinks**. In Amazonian dark earths. Dordrecht: Springer, p.125-139, 2003.

SOUZA, K. W. D.; LIMA, H. N.; SCHAEFER, C. E. G.; TEIXEIRA, W. G.; PULROLNIK, K.; CORRÊA, G. R. Phosphorous forms in cultivated Indian Black Earth (Anthrosols) of varying texture in the Brazilian Amazon. **Revista Brasileira de Ciência do Solo**, v.33, n.5, p.1347-1355, 2009.

STEINBEISS, S.; GLEIXNER, G.; ANTONIETTI, M. Effect of biochar amendment on soil carbon balance and soil microbial activity. **Soil Biology and Biochemistry**, v.41, n.6, p.1301-1310, 2009.

STEINER, C.; TEIXEIRA, W. G.; LEHMANN, J.; ZECH, W. **Microbial response to charcoal amendments of highly weathered soils and Amazonian dark earths in Central Amazonia – preliminary results**. In: GLASER, B.; WOODS, W. I. (Eds.). Amazonian dark earths: explorations in space and time. New York: Springer, p.195-212, 2004.

STOLF, R.; MURAKAMI, J. H.; BRUGNARO, C.; SILVA, L. G.; SILVA, L. C. F.; MARGARIDO, L.A.C. Penetrômetro de impacto Stolf - programa computacional de dados em EXCEL-VBA. **Revista Brasileira de Ciência do Solo**, v.38, n.3, p.774-782, 2014.

TEIXEIRA, P. C.; DONAGEMMA, G. K.; FONTANA, A.; TEIXEIRA, W. G. **Manual de Métodos de Análise de Solos**. 3ª Ed. Revista e Ampliada, Brasília: Embrapa, 2017. 573p.

TEIXEIRA, W. G.; KERN, D. C.; MADARI, B. E.; LIMA, H. N.; WOODS, W. **As Terras Pretas de Índio da Amazônia: sua caracterização e uso deste conhecimento na criação de novas áreas**. Manaus: Editora da Universidade Federal do Amazonas (EDUA): Embrapa Amazônia Ocidental, 2010.

TEIXEIRA, W. G.; MARTINS, G. C. **Soil physical characterization**. In: LEHMANN, J.; KERN, D.C.; GLASER, B.; WOODS, W.I. Amazonian Dark Earths: origin, properties, management. Dordrecht: Kluwer Academic Publishers, Cap.15, p.271-286, 2003.

TEIXEIRA, W. G.; MARTINS, G. C., MACEDO, R. S.; NEVES JÚNIOR, A. F.; MOREIRA, A.; BENITES, V. M.; STEINER, C. N. **As propriedades físicas e hídricas dos horizontes antrópicos das Terras Pretas Arqueológicas**. In: As Terras Pretas de índio da Amazônia: Sua Caracterização e Uso deste Conhecimento na Criação de Novas Áreas. TEIXEIRA, W. G.; KERN, D. C.; MADARI, B. E.; LIMA, H. N.; WOODS, W. (Eds.). Embrapa Amazônia Ocidental, Manaus, AM. p.242-250, 2009.

THOMAZ, E. L.; FACHIN, P. A. Effects of heating on soil physical properties by using realistic peak temperature gradients. **Geoderma**, v.230, p.243-249, 2014.

TONON, L.; BORSZOWSKI, P. R. Perda de carbono do solo devido a conversão de vegetação natural para sistema convencional e estimativa de recuperação utilizando cenário através de um modelo unicompartmental. **Revista TechnoEng**, v.1, n.5, p.1-19, 2015.

TRAZZI, P. A., HIGA, A. R., DIECKOW, J., MANGRICH, A. S.; HIGA, R. C. V. (Biochar: Reality and potential use in forestry. **Ciência Florestal**, v.28, n.2, p.875-887, 2018.

TRIOLA, M. F. **Introdução à estatística**. 10. Ed. Rio de Janeiro: LTC, 696 p, 2008.

VELDKAMP, E. Organic carbon turnover in three tropical soils under pasture after deforestation. **Soil Science Society of America Journal**, v.58, n.1, p.175-180, 1994.

VIANA, E. T.; BATISTA, M. A.; TORMENA, C. A.; COSTA, A. C. S. INOUE, T. T. Atributos físicos e carbono orgânico em Latossolo Vermelho sob diferentes sistemas de uso e manejo. **Revista Brasileira de Ciência do Solo**, v.35, n.3, p.2105-2114, 2011.

VITAL, A. R. T.; GUERRINI, I. A.; FRANKEN, W. K.; FONSECA, R. C. B. Produção de serrapilheira e ciclagem de uma floresta estacional semidecidual em zona ripária. **Revista Árvore**, v.28, n.6, p.793-800, 2004.

WALKLEY, A.; BLACK, I. A. An examination of the Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. **Soil Science**, v.37, n.1, p.29-38, 1934.

WITTEN, I. H.; FRANK, E.; HALL, M. A. **Data mining: practical machine learning tools and techniques**. 3. Ed. San Francisco: Morgan Kaufmann, 664 p., 2011.

WOODS, W. **Os solos e as ciências humanas: Interpretação do passado**. In: As Terras Pretas de índio da Amazônia: Sua caracterização e uso deste conhecimento na criação de novas áreas. (Eds.). TEIXEIRA, W. G.; KERN, D. C.; MADARI, B. E.; LIMA, H. N.; WOODS, W. Embrapa Amazônia Ocidental, Manaus, AM, p.62-71, 2009.

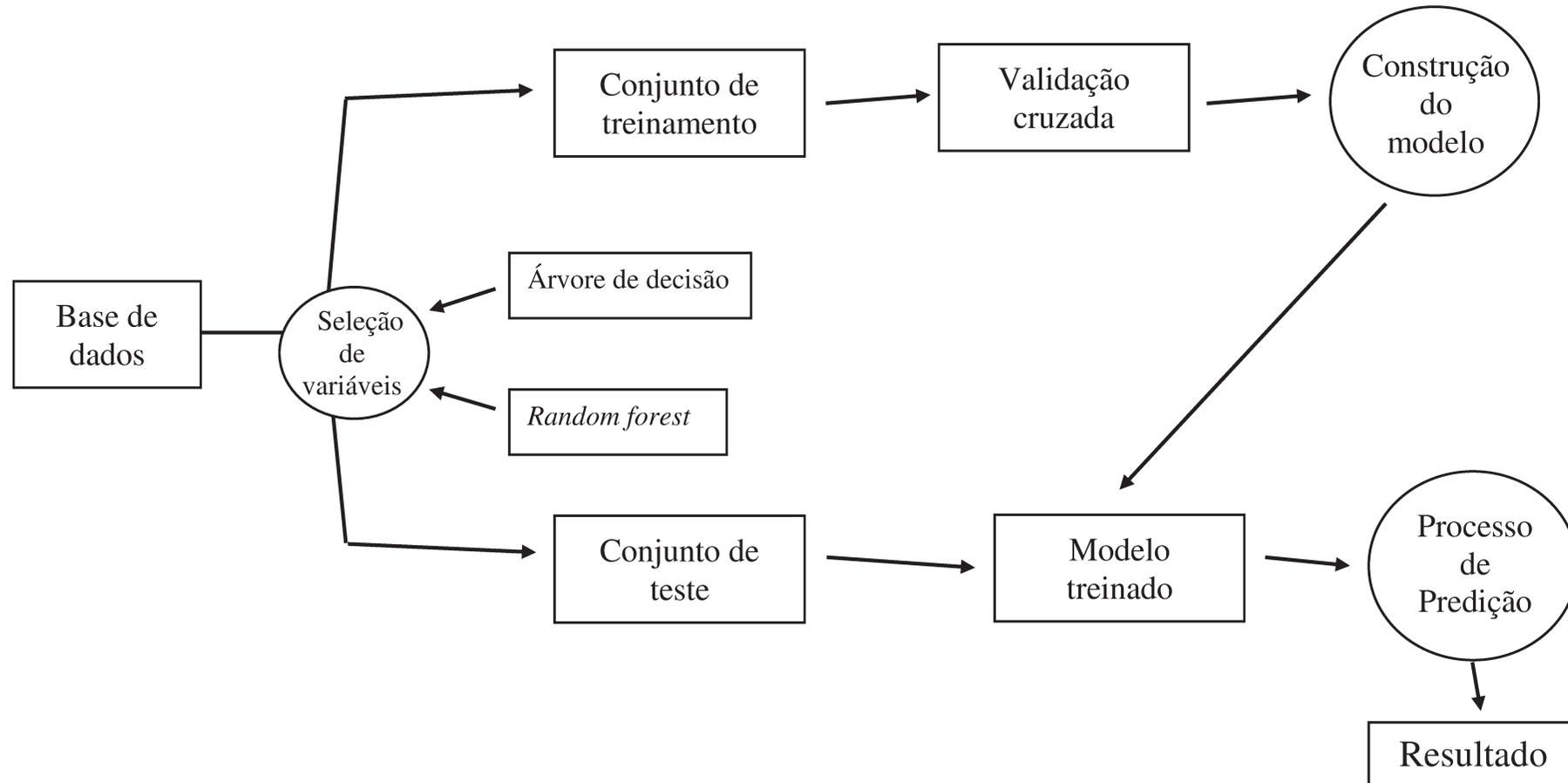
YEOMANS, J. C.; BREMNER, J. M. A rapid and precise method for routine determination of organic carbon in soil. **Communications in Soil Science and Plant Analysis**, v.19, n.13, p.1467-1476, 1988.

ZHU, F.; LI, X.; XUE, S.; HARTLEY, W.; WU, C.; HAN, F. Natural plant colonization improves the physical condition of bauxite residue over time. **Environmental Science and Pollution Research**, v.23, n.22, p.22897-22905, 2016.

APÊNDICES

Apêndice 1. Processo de Predição de estoque de carbono do solo em terra preta arqueológica.

|----- Etapa de preparação de dados -----|----- Etapa de construção da árvore de decisão -----|----- Etapa de predição -----|



Apêndice 2. Descrição geral do perfil pedológico – Terra Preta Arqueológica.

PERFIL: TERRA PRETA DE ÍNDIO

CLASSIFICAÇÃO:

SITUAÇÃO, DECLIVIDADE E COBERTURA VEGETAL SOBRE O PERFIL: Descrito e coletado no topo, suave ondulado 0 a 2,5%.

ALTITUDE: 120 metros

LITOLOGIA:

FORMAÇÃO GEOLÓGICA:

PERÍODO:

MATERIAL ORIGINÁRIO:

PEDREGOSIDADE: não pedregoso

ROCHOSIDADE: não rochosa

RELEVO LOCAL: Plano.

RELEVO REGIONAL: Suave Ondulado

SUPERFÍCIE GEOMÓRFICA:

POSIÇÃO DA PAISAGEM: Topo

EROSÃO: Nula

DRENAGEM: Boa

VEGETAÇÃO PRIMÁRIA: Floresta

USO ATUAL: cultivo de café, cacau, feijão e pastagem.

CLIMA:

DESCRITO E COLETADO: CAMPOS, M. C. C & SANTOS, L. A. C.

B - DESCRIÇÃO MORFOLÓGICA

Ap₁-0 – 19 cm, (10YR 2/2, úmida); franco argila arenosa; moderada muito pequena a pequena blocos angulares e sub angular; muito friável não plástico e não pegajosa; transição gradual e plana.

BA-19 - 41 cm; (10YR 4/3, úmida); franco argilosa; moderada pequena a médio blocos angulares e sub angulares; friável ligeiramente plástico e ligeiramente pegajosa; transição gradual e plana.

Bt₁-41 - 67 cm; (10YR 5/6, úmida); franco argilosa; forte pequena a médio blocos angulares e sub angulares; firme ligeiramente plástico e ligeiramente pegajosa; transição gradual e plana.

Bt₂-67 - 104 cm; (10YR 5/8, úmida); argila; forte pequena a médio blocos angulares e sub angulares; muito firme plástico e pegajosa; transição gradual e plana.

Bt₃-104 - 147⁺ cm; (10YR 5/8, úmida); argila; forte pequena a médio blocos angulares e sub angulares; firme plástico e pegajosa;

RAÍZES – Muitas, muito finas e finas no horizonte Ap₁; poucas, muito finas e finas nos horizontes BA e Bt₁, raras e muito finas no horizonte Bt₂ e Bt₃.

C – ANÁLISES FÍSICAS E QUÍMICAS

Horizonte		Composição Granulométrica					pH		BASES TROCAVEIS				SOMA BASES	Al ³⁺	H+Al	CTC	V	m
Símb.	Prof.	Areia Grossa	Areia Fina	Silte	Argila	S/A	Água	Ca ²⁺	Mg ²⁺	K ⁺	Na ⁺	P						
(cm)		g.kg ⁻¹						mmol/Kg				mg/d ³	mmol/Kg			%		
Ap ₁	0 – 19	223,99	107,99	304,09	363,92	0,83	4,3	16,0	5,0	0,80	0,001	65,0	21,8	1,20	109,0	130,8	16,7	2,76
BA	19 – 41	104,15	155,72	266,59	473,53	0,57	4,2	11,0	4,0	0,10	0,05	34,0	15,1	1,05	64,0	79,1	19,1	1,32
Bt ₁	41 – 67	65,87	118,83	391,44	423,85	2,18	4,2	11,0	4,0	0,10	0,02	41,0	15,1	0,65	42,0	57,1	26,4	1,13
Bt ₂	67 – 104	59,15	107,75	233,89	599,20	0,39	4,2	10,0	4,0	0,10	0,02	41,0	14,1	0,45	34,0	48,1	29,3	0,93
Bt ₃	104 – 147 ⁺	55,84	99,78	298,48	545,89	0,54	4,3	9,0	4,0	0,10	0,03	32,0	13,1	0,40	34,0	47,1	27,8	0,84

ATAQUE SULFURICO (H ₂ SO ₄ 1:1)			Fe _d	Fe _{ox}	SiO ₂	SiO ₂	Al ₂ O ₃	Fe _d /Fe _{ox}	CO	MO
SiO ₂	Al ₂ O ₃	Fe ₂ O ₃			/Al ₂ O ₃	/Al ₂ O ₃ +Fe ₂ O ₃	/Fe ₂ O ₃			
g.Kg ⁻¹					Ki	Kr		g/dm ³		
-	-	-	3,29	0,24	-	-	-	13,43	18,56	32,0
-	-	-	4,89	0,14	-	-	-	35,28	6,38	11,0
-	-	-	5,42	0,17	-	-	-	31,62	4,06	7,0
-	-	-	2,35	0,16	-	-	-	15,14	3,48	6,0
-	-	-	2,51	1,00	-	-	-	2,50	3,48	6,0

Apêndice 3. Estatística descritiva dos atributos físicos e químicos da base de dados usada na predição de estoque de carbono em Terra preta arqueológica, na profundidade 0,00 – 0,05 m.

Estatística descritiva	RSP	Ds	Macro	Micro	VTP	Silte	Argila	Areia	Umid (Vol)	DMG	DMP	> 2mm	< 2mm	pH	Ac. pot	Al ⁺³	P	K ⁺	Ca ⁺²	Mg ⁺²	Est. de C.
Cacau																					
Media	1.13	0.90	19.96	49.38	69.01	574.96	202.94	219.65	49.35	2.71	3.15	90.37	11.15	6.23	6.55	0.06	106.40	0.02	19.47	2.13	378.94
Mínimo	0.45	0.69	10.20	40.40	56.95	515.16	103.92	115.18	40.40	1.80	2.77	76.87	0.07	4.49	1.16	0.04	30.38	0.00	10.50	0.25	191.53
Máximo	1.99	1.15	32.78	57.03	85.05	638.00	288.08	326.30	57.03	3.68	3.45	99.93	44.96	7.53	15.02	0.10	549.23	0.05	26.25	5.00	573.76
DP	0.37	0.10	5.64	3.58	4.77	26.58	41.71	47.91	3.57	0.39	0.16	5.34	7.88	0.66	3.45	0.02	77.43	0.01	3.33	1.07	74.92
Var.	0.14	0.01	31.80	12.80	22.75	706.69	1739.86	2295.22	12.78	0.15	0.03	28.53	62.08	0.43	11.90	0.00	5995.49	0.00	11.11	1.14	5613.61
Café																					
Media	1.18	1.09	22.16	37.41	59.64	617.52	17.16	364.65	5.15	2.77	3.14	92.04	6.70	5.36	8.49	0.09	34.34	0.01	8.75	2.47	272.50
Mínimo	0.42	0.80	11.77	32.87	47.59	543.98	13.43	276.63	2.71	1.21	2.86	83.36	1.21	4.34	2.31	0.04	10.26	0.01	2.50	0.50	129.95
Máximo	2.26	1.36	29.83	42.29	68.99	687.89	21.80	442.39	9.06	3.45	3.36	102.63	16.97	6.61	14.03	0.19	79.99	0.03	17.50	4.50	494.98
DP	0.42	0.11	4.62	2.00	4.56	32.26	1.92	34.70	1.44	0.36	0.11	3.92	3.66	0.47	2.42	0.04	15.07	0.00	3.23	0.91	88.28
Var.	0.18	0.01	21.38	3.99	20.80	1040.92	3.70	1204.10	2.08	0.13	0.01	15.38	13.42	0.22	5.87	0.00	227.23	0.00	10.43	0.83	7793.73
Feijão																					
Media	0.54	1.48	18.38	31.74	50.09	141.01	96.77	763.28	31.74	1.44	2.05	55.22	44.90	6.19	6.06	0.23	58.81	0.08	7.45	1.32	225.35
Mínimo	0.16	1.25	7.91	21.66	39.74	56.91	34.12	666.97	21.66	0.41	0.68	9.20	8.46	5.12	3.80	0.10	8.50	0.03	4.70	0.60	53.23
Máximo	1.06	1.69	33.13	41.78	64.32	260.88	142.80	824.20	41.78	2.83	3.12	91.55	98.42	7.47	9.90	0.60	135.48	0.16	10.70	2.00	389.77
DP	0.18	0.09	5.29	4.24	3.75	44.17	22.87	35.22	4.24	0.56	0.59	20.08	20.30	0.45	1.38	0.14	32.71	0.03	1.20	0.33	63.40
Var.	0.03	0.01	27.97	18.00	14.03	1950.59	523.11	1240.75	18.00	0.31	0.35	403.39	412.00	0.20	1.91	0.02	1070.14	0.00	1.44	0.11	4019.24
Pastagem																					
Media	0.48	1.23	27.29	31.19	57.84	179.98	78.21	741.92	31.19	2.61	3.08	90.57	9.43	6.38	6.35	0.08	80.53	0.09	6.46	1.60	159.29
Mínimo	0.11	0.98	8.86	23.13	46.42	90.96	33.52	675.92	23.13	1.89	2.71	77.19	2.61	6.01	2.64	0.06	28.89	0.01	4.70	0.70	96.26
Máximo	1.03	1.51	46.14	41.08	68.36	254.67	126.04	799.74	41.08	3.11	3.28	97.39	22.81	6.69	11.22	0.10	168.87	0.18	8.70	2.50	253.99
DP	0.21	0.11	7.35	3.76	4.76	34.23	19.56	28.52	3.76	0.29	0.14	4.81	4.81	0.16	1.73	0.01	32.13	0.04	0.94	0.38	32.64
Var.	0.05	0.01	54.03	14.14	22.68	1171.67	382.74	813.67	14.14	0.08	0.02	23.12	23.12	0.03	2.98	0.00	1032.44	0.00	0.89	0.14	1065.25
Floresta																					
Media	0.40	1.27	26.46	30.95	56.99	203.39	94.60	700.07	30.95	2.70	3.09	90.80	9.20	5.76	5.27	0.10	24.09	0.06	6.67	1.38	158.90
Mínimo	0.09	1.05	18.58	25.45	47.09	139.36	55.69	626.62	25.45	1.94	2.82	78.81	2.20	5.04	2.64	0.10	0.00	0.00	3.40	0.50	51.74
Máximo	0.76	1.47	34.86	36.31	65.56	270.45	131.69	771.60	36.31	3.15	3.29	97.80	21.19	6.58	8.25	0.10	61.27	0.13	9.70	2.20	657.60
DP	0.13	0.09	3.58	2.40	4.10	29.01	15.80	31.39	2.40	0.28	0.12	4.30	4.30	0.34	1.35	0.00	12.81	0.03	1.31	0.39	65.73
Var.	0.02	0.01	12.80	5.76	16.83	841.62	249.62	985.35	5.76	0.08	0.01	18.46	18.46	0.11	1.82	0.00	164.21	0.00	1.72	0.15	4319.89

RSP = Resistência do solo a penetração (MPa); Ds = Densidade do solo (kg dm⁻³); Macro = Macroporosidade (m³ m⁻³); Micro = Microporosidade (m³ m⁻³); VTP = Volume total de poros (%); Umid (Vol) = Umidade volumétrica (kg kg⁻¹); DMG = Diâmetro médio geométrico (mm); DMP = Diâmetro médio ponderado (mm); > 2mm = classe de diâmetro: > 2,0 mm; < 2mm = classe de diâmetro: < 2,0 mm; pH = Acidez ativa (H₂O); Ac. Pot = Acidez Potencial (mmolc dm⁻³); Al⁺³ = Alumínio trocável (mmolc dm⁻³); P = fósforo disponível (mg dm⁻³); K⁺ = Potássio (mmolc dm⁻³); Ca⁺² = Cálcio (mmolc dm⁻³); Mg⁺² = Magnésio (mmolc dm⁻³); Est. de Carb = Estoque de carbono no solo (Mg ha⁻¹).

Apêndice 4. Estatística descritiva dos atributos físicos e químicos da base de dados usada na predição de estoque de carbono em Terra preta arqueológica, na profundidade 0,05 – 0,10 m.

Estatística descritiva	RSP	Ds	Macro	Micro	VTP	Silte	Argila	Areia	Umid (Vol)	DMG	DMP	> 2mm	< 2mm	pH	Ác. pot	Al+3	P	K	Ca	Mg	Est. de C.
Cacau																					
Media	1.12	0.95	19.07	46.70	66.21	542.24	260.49	195.16	46.70	2.55	3.02	85.90	14.10	6.14	9.16	0.08	109.61	0.02	11.60	4.16	279.50
Mínimo	0.35	0.73	10.23	38.03	59.09	464.00	166.72	110.63	38.03	1.65	2.53	69.40	0.71	4.78	0.99	0.03	35.22	0.01	3.40	0.95	113.94
Máximo	1.79	1.18	31.45	56.99	75.32	599.81	388.00	299.40	56.99	3.48	3.50	99.29	30.60	7.78	19.64	0.17	198.32	0.04	21.50	8.65	418.37
DP	0.34	0.09	4.86	4.15	3.56	30.26	47.49	38.74	4.17	0.40	0.20	7.00	7.00	0.79	4.97	0.03	33.63	0.01	3.95	1.60	49.96
Var.	0.11	0.01	23.60	17.24	12.67	915.52	2254.99	1501.16	17.37	0.16	0.04	49.05	49.05	0.62	24.71	0.00	1130.99	0.00	15.62	2.58	2495.96
Café																					
Media	1.11	1.16	18.85	38.45	57.57	565.84	13.98	419.92	4.10	2.70	3.08	88.45	10.16	5.08	8.64	0.27	30.30	0.00	4.87	1.59	233.17
Mínimo	0.39	0.85	6.38	33.16	48.19	494.08	9.33	353.96	2.15	1.85	2.67	76.46	1.72	4.12	2.64	0.03	5.14	0.00	1.25	0.25	81.00
Máximo	1.90	1.43	28.90	47.57	69.35	627.41	18.72	493.43	7.46	3.51	3.37	98.13	33.12	6.55	15.02	0.98	77.56	0.02	10.50	3.00	395.89
DP	0.32	0.12	5.08	3.16	4.55	29.60	2.13	30.00	1.12	0.31	0.14	5.08	5.42	0.53	2.96	0.25	18.39	0.01	2.46	0.62	76.49
Var.	0.10	0.02	25.84	9.99	20.72	876.08	4.55	900.04	1.26	0.10	0.02	25.78	29.42	0.28	8.74	0.06	338.03	0.00	6.05	0.39	5851.22
Feijão																					
Media	0.66	1.42	19.42	31.73	51.22	143.80	86.95	769.72	31.84	2.26	2.79	80.39	19.66	6.47	5.30	0.26	62.80	0.04	6.64	0.90	122.43
Mínimo	0.28	1.13	9.14	25.24	43.10	22.63	38.00	672.05	25.24	1.14	1.93	51.49	3.34	5.35	1.65	0.10	12.33	0.01	3.20	0.30	35.53
Máximo	1.19	1.65	30.47	40.04	61.51	248.87	147.48	865.61	41.84	3.18	3.26	96.66	50.59	7.68	9.08	0.40	133.08	0.09	10.80	1.60	229.67
DP	0.21	0.11	4.58	3.59	4.09	40.75	24.76	38.44	3.75	0.48	0.33	11.34	11.44	0.49	1.58	0.07	34.18	0.02	1.78	0.31	35.20
Var.	0.04	0.01	20.98	12.88	16.72	1660.74	612.82	1477.95	14.03	0.23	0.11	128.53	130.96	0.24	2.50	0.00	1168.57	0.00	3.18	0.09	1238.72
Pastagem																					
Media	0.75	1.41	15.63	33.48	49.58	172.53	80.59	746.32	33.48	2.23	2.85	82.71	17.29	6.38	5.23	0.11	52.66	0.03	6.00	0.91	146.84
Mínimo	0.19	1.15	6.67	25.95	42.24	106.03	54.48	681.32	25.95	0.61	1.97	52.87	1.86	5.83	2.31	0.10	11.94	0.01	4.10	0.40	67.77
Máximo	1.37	1.65	26.34	42.46	58.95	240.32	126.20	817.69	42.46	3.19	3.30	98.14	47.13	6.72	9.08	0.13	108.16	0.07	8.60	1.60	291.88
DP	0.25	0.10	4.62	3.22	4.01	30.48	16.83	27.03	3.22	0.59	0.33	11.15	11.15	0.22	1.39	0.01	23.35	0.02	0.93	0.28	42.19
Var.	0.06	0.01	21.36	10.37	16.07	928.97	283.12	730.43	10.37	0.35	0.11	124.30	124.30	0.05	1.92	0.00	545.25	0.00	0.87	0.08	1779.90
Floresta																					
Media	0.41	1.36	22.91	28.22	51.37	196.59	103.51	699.87	28.22	2.05	2.65	75.56	24.59	5.89	6.48	0.12	0.40	0.03	5.99	0.95	162.02
Mínimo	0.15	1.14	16.71	23.30	43.70	105.19	63.93	598.39	23.30	0.88	1.93	49.40	0.00	5.02	1.65	0.00	0.00	0.01	2.90	0.30	27.54
Máximo	0.71	1.57	30.49	33.72	59.68	306.05	134.61	792.32	33.72	2.95	3.20	94.39	62.51	6.98	10.40	0.20	0.95	0.04	8.50	1.80	648.00
DP	0.12	0.08	3.01	2.28	3.26	38.76	14.62	41.03	2.28	0.48	0.35	12.18	13.00	0.49	1.63	0.04	0.21	0.01	1.23	0.35	65.01
Var.	0.02	0.01	9.09	5.22	10.61	1502.47	213.76	1683.34	5.22	0.23	0.12	148.37	168.93	0.24	2.65	0.00	0.04	0.00	1.51	0.12	4226.23

RSP = Resistência do solo a penetração (MPa); Ds = Densidade do solo (kg dm⁻³); Macro = Macroporosidade (m³ m⁻³); Micro = Microporosidade (m³ m⁻³); VTP = Volume total de poros (%); Umid (Vol) = Umidade volumétrica (kg kg⁻¹); DMG = Diâmetro médio geométrico (mm); DMP = Diâmetro médio ponderado (mm); > 2mm = classe de diâmetro: > 2,0 mm; < 2mm = classe de diâmetro: < 2,0 mm; pH = Acidez ativa (H₂O); Ac. Pot = Acidez Potencial (mmol_c dm⁻³); Al⁺³ = Alumínio trocável (mmol_c dm⁻³); P = fósforo disponível (mg dm⁻³); K⁺ = Potássio (mmol_c dm⁻³); Ca⁺² = Cálcio (mmol_c dm⁻³); Mg⁺² = Magnésio (mmol_c dm⁻³); Est. de Carb = Estoque de carbono no solo (Mg ha⁻¹).

Apêndice 5. Estatística descritiva dos atributos físicos e químicos da base de dados usada na predição de estoque de carbono em Terra preta arqueológica, na profundidade 0,10 – 0,20 m.

Estatística descritiva	RSP	Ds	Macro	Micro	VTP	Silte	Argila	Areia	Umid (Vol)	DMG	DMP	> 2mm	< 2mm	pH	Ác. pot	Al+3	P	K	Ca	Mg	Est. de C.
Cacau																					
Media	0.89	0.95	22.38	43.37	65.75	537.87	274.20	187.93	43.37	2.57	2.50	75.57	24.43	5.71	11.54	0.20	135.75	0.02	10.47	2.07	216.59
Mínimo	0.25	0.73	10.64	29.71	53.05	432.47	175.64	112.76	29.71	1.29	1.26	3.43	1.19	4.58	1.32	0.04	48.29	0.00	2.75	0.38	99.50
Máximo	1.89	1.25	34.13	53.21	76.03	654.68	373.72	286.45	53.21	3.53	3.98	98.81	96.57	7.95	22.11	1.06	550.19	0.33	20.38	5.63	339.53
DP	0.37	0.10	5.25	3.84	4.31	37.58	47.98	36.11	3.83	0.43	0.50	12.33	12.33	0.75	4.99	0.22	62.13	0.04	4.29	1.17	46.19
Var.	0.14	0.01	27.54	14.71	18.59	1412.38	2302.22	1304.12	14.71	0.19	0.25	152.04	152.04	0.57	24.86	0.05	3860.19	0.00	18.41	1.36	2133.25
Café																					
Media	1.29	1.23	16.61	38.42	55.03	638.52	17.16	344.32	4.18	2.23	3.04	80.76	15.31	4.98	9.90	0.58	32.70	0.00	3.92	1.33	223.03
Mínimo	0.48	0.77	7.00	33.77	47.01	541.23	-189.28	239.77	2.06	1.31	2.07	50.87	2.01	3.96	2.81	0.04	5.94	0.00	1.25	0.25	85.05
Máximo	2.21	1.45	26.70	49.40	69.58	809.25	27.87	444.03	12.38	3.87	3.37	94.78	64.05	6.80	21.95	2.19	119.78	0.01	10.50	3.25	319.34
DP	0.44	0.14	5.24	2.82	4.63	36.07	22.42	32.50	1.68	0.39	0.21	8.01	9.57	0.55	3.50	0.51	21.48	0.00	2.24	0.52	49.94
Var.	0.19	0.02	27.44	7.97	21.44	1301.32	502.64	1056.03	2.81	0.15	0.04	64.09	91.53	0.30	12.26	0.26	461.52	0.00	5.02	0.27	2494.05
Feijão																					
Media	0.57	1.47	18.38	31.89	50.27	142.38	96.77	760.89	31.89	1.44	2.05	55.22	44.90	6.29	4.98	0.36	51.15	0.04	5.45	0.81	126.41
Mínimo	0.16	1.06	7.91	21.66	39.74	56.91	34.12	651.99	21.66	0.41	0.68	9.20	8.46	4.84	0.00	0.00	5.52	0.01	2.10	0.30	26.25
Máximo	1.18	1.75	33.13	45.52	66.43	264.41	142.80	824.20	45.52	2.83	3.12	91.55	98.42	8.17	9.41	0.70	150.40	0.28	10.10	1.90	337.94
DP	0.22	0.11	5.29	4.46	4.10	45.80	22.87	38.30	4.46	0.56	0.59	20.08	20.30	0.61	1.74	0.12	29.36	0.04	1.75	0.38	65.27
Var.	0.05	0.01	27.97	19.91	16.84	2097.87	523.11	1467.22	19.91	0.31	0.35	403.39	412.00	0.37	3.02	0.01	862.29	0.00	3.07	0.14	4260.62
Pastagem																					
Media	0.68	1.46	15.76	33.38	49.14	149.65	90.98	759.37	33.38	1.59	2.27	62.89	37.11	5.81	4.02	0.08	61.25	0.01	0.01	5.11	124.30
Mínimo	0.30	1.28	7.86	24.32	42.21	79.45	38.60	646.26	24.32	0.42	0.75	11.74	4.38	5.20	0.50	0.06	7.33	0.01	0.01	2.60	21.95
Máximo	1.35	1.66	25.20	43.34	58.77	294.46	177.18	828.59	43.34	3.04	3.23	95.62	88.26	6.37	11.72	0.27	163.35	0.08	0.08	8.40	245.89
DP	0.23	0.09	4.19	3.57	3.64	37.34	22.75	31.13	3.57	0.66	0.60	20.39	20.39	0.26	1.96	0.03	36.96	0.01	0.01	1.08	38.80
Var.	0.05	0.01	17.59	12.74	13.22	1393.98	517.78	969.02	12.74	0.44	0.36	415.90	415.90	0.07	3.83	0.00	1366.37	0.00	0.00	1.17	1505.21
Floresta																					
Media	0.39	1.37	22.07	28.26	50.33	196.46	86.03	717.50	28.26	1.34	1.97	51.73	48.27	5.78	6.03	0.30	27.93	0.02	5.56	0.80	127.56
Mínimo	0.11	1.14	10.66	23.46	40.82	118.96	20.09	592.52	23.46	0.49	0.90	5.95	9.02	5.15	2.31	0.10	2.36	0.01	1.20	0.30	23.98
Máximo	0.96	1.64	31.72	33.63	58.16	331.68	132.93	779.48	33.63	2.75	3.09	90.98	94.05	6.84	13.04	1.10	103.13	0.04	9.00	1.90	342.07
DP	0.13	0.10	3.65	2.65	3.99	39.40	17.25	39.12	2.65	0.51	0.53	18.53	18.53	0.43	1.88	0.20	18.20	0.01	1.72	0.33	55.30
Var.	0.02	0.01	13.34	7.00	15.90	1552.00	297.65	1530.08	7.00	0.26	0.28	343.51	343.51	0.18	3.55	0.04	331.40	0.00	2.97	0.11	3058.42

RSP = Resistência do solo a penetração (MPa); Ds = Densidade do solo (kg dm⁻³); Macro = Macroporosidade (m³ m⁻³); Micro = Microporosidade (m³ m⁻³); VTP = Volume total de poros (%); Umid (Vol) = Umidade volumétrica (kg kg⁻¹); DMG = Diâmetro médio geométrico (mm); DMP = Diâmetro médio ponderado (mm); > 2mm = classe de diâmetro: > 2,0 mm; < 2mm = classe de diâmetro: < 2,0 mm; pH = Acidez ativa (H₂O); Ac. Pot = Acidez Potencial (mmol_c dm⁻³); Al⁺³ = Alumínio trocável (mmol_c dm⁻³); P = fósforo disponível (mg dm⁻³); K⁺ = Potássio (mmol_c dm⁻³); Ca⁺² = Cálcio (mmol_c dm⁻³); Mg⁺² = Magnésio (mmol_c dm⁻³); Est. de Carb = Estoque de carbono no solo (Mg ha⁻¹).