

**UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL**

**EMIÇÃO DE CO₂ DO SOLO E SUA RELAÇÃO COM CO₂ ATMOSFÉRICO
EM ÁREAS AGRÍCOLAS DO ESTADO DO MATO GROSSO DO SUL: UMA
ABORDAGEM DE APRENDIZADO DE MÁQUINA**

LETÍCIA ROBERTA DE LIMA

JABOTICABAL, 2024

**UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL**

**EMIÇÃO DE CO₂ DO SOLO E SUA RELAÇÃO COM CO₂ ATMOSFÉRICO
EM ÁREAS AGRÍCOLAS DO ESTADO DO MATO GROSSO DO SUL: UMA
ABORDAGEM DE APRENDIZADO DE MÁQUINA**

Letícia Roberta de Lima

Orientador: Prof. Dr. Alan Rodrigo Panosso
Coorientador: Me. Kleve Freddy Ferreira Canteral

Trabalho de Conclusão de Curso
apresentado à Faculdade de Ciências
Agrárias e Veterinárias – Unesp,
Câmpus de Jaboticabal, como parte das
exigências para Graduação em
Engenharia Agrônoma.

**JABOTICABAL – SP
2º semestre/2024**

L732e

Lima, Letícia Roberta de

Emissão de CO₂ do solo e sua relação com CO₂ atmosférico em áreas agrícolas do mato grosso do sul: uma abordagem de aprendizado de máquina / Letícia Roberta de Lima. -- , 2024

31 p. : il., tabs., fotos, mapas

Trabalho de conclusão de curso (-) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal,

Orientador: Alan Rodrigo Panosso

Coorientador: Kleve Freddy Ferreira Canteral

1. Dioxido de carbono. 2. Inteligencia artificial Processamento de dados. 3. Mudanças climaticas. 4. Mineração de dados (Computação). 5. Solo e clima. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

LETÍCIA ROBERTA DE LIMA

TÍTULO DO TRABALHO ACADÊMICO:

Emissão de CO₂ do solo e sua relação com CO₂ atmosférico em áreas agrícolas do estado do Mato Grosso do Sul: uma abordagem de aprendizado de máquina.

Trabalho de Iniciação Científica apresentada à Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal, para obtenção do título de Bacharel em Engenharia Agrônoma.

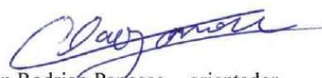
Orientador: Prof. Dr. Alan Rodrigo Panosso
Coorientador: MSc. Kleve Freddy Ferreira Canteral

Área de Concentração: Agronomia / Ciência do Solo

Data da defesa: 5 / 7 / 2024

(X) Aprovado
() Reprovado

Banca Examinadora:



Prof. Dr. Alan Rodrigo Panosso - orientador
UNESP – Faculdade de Ciências Agrárias e Veterinárias – Câmpus de Jaboticabal



Dr.ª Mara Regina Moutinho
Grupo de pesquisa CSME da FCAV/Unesp



Dr.ª Ludhanna Marinho Veras
SEBRAE (Barretos-SP e Presidente Prudente-SP)

Aprovado por ato “ad referendum” do presidente do Conselho do Departamento em: 5 / 7 / 2024.



Prof. Dr. Danisio Prado Munari
Chefe do Departamento de Ciências Exatas FCAV/Unesp

DEDICATÓRIA

Dedico este trabalho à minha família e ao meu orientador, Prof. Dr. Alan Rodrigo Panosso.

AGRADECIMENTOS

Em primeiro lugar, agradeço a Deus e Nossa Senhora Aparecida, por me concederem saúde, força e perseverança para superar os desafios ao longo desta jornada acadêmica.

Ao meu orientador, Prof. Dr. Alan Rodrigo Panosso, pela orientação, paciência e valiosas contribuições que enriqueceram este trabalho. Seus incentivos deixaram a jornada inspiradora, tornando o processo possível e o acolhimento com conselhos que foram essenciais para minha formação acadêmica, profissional e pessoal. Ao meu coorientador e amigo Me. Kleve Freddy Ferreira Canteral, por se mostrar solícito sempre (em todos os horários possíveis) e pelo incansável apoio emocional e profissional.

Aos meus pais, Aline Roberta Ingissa de Lima e Otávio Henrique de Lima, por todo o amor, dedicação, apoio incondicional e por sempre acreditarem em meu potencial. As minhas irmãs, Rafaela Caroline de Lima e Laís Maria de Lima, pela compreensão, incentivo nos momentos mais difíceis, compartilhamento de muitas risadas e carinho aos finais de semana.

Aos meus avós por todos os momentos de cuidado e por sempre me incentivarem a ir atrás dos meus sonhos, sempre viveram na minha lembrança, o amor de vocês é a minha maior herança. A minha amada madrinha, Izabel Aparecida de Lima, por ter me auxiliado nos estudos desde cedo e ter sempre me concedido amor materno.

À minha velha amiga, Maria Rita Nunes Ribeiro, eterna parceira de truco, que me concedeu carinho, atenção, cuidado e amor desde à infância; demonstrando sempre o seu entusiasmo e admiração pelas minhas ideias. A minha amiga Julia di Donato Santos, que esteve comigo durante esse período, sempre escutando e apoiando minhas escolhas. Pelo afeto ao longo desses anos, por estar comigo nas horas boas e ruins sempre melhorando os meus dias.

À minha querida namorada, Treyce Stephane Cristo Tavares, minha gratidão pelo apoio, paciência e carinho. Sua memorável presença constante e encorajamento foram fundamentais para que eu pudesse superar os desafios e acreditar no meu potencial profissional. Obrigada por ser o meu amor lindo, por acreditar em mim, me auxiliar e ser minha fonte de inspiração.

À Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP,
pelo auxílio financeiro para desenvolvimento desta pesquisa - NÚMERO DO
PROCESSO: 2021/02487-0.

Finalmente, a todos aqueles que, direta ou indiretamente, contribuíram
para a realização deste trabalho, meu sincero agradecimento.

LISTA DE FIGURAS

Figura 1. Mapa das áreas de estudo localizadas no município de Selvíria, MS. a) Estado do Mato Grosso do Sul, Brasil Central; b) Município de Selvíria; c) Áreas de florestas plantadas: pinus, eucalipto e mata ciliar, respectivamente.	5
Figura 2. Resumo do processo de desenvolvimento dos modelos para prever a emissão de CO ₂ do solo, por meio dos algoritmos de aprendizado de máquina. R ² = coeficiente de determinação; R ² _{ajust} = coeficiente de determinação ajustado; MSE = erro quadrático médio; RMSE = raiz quadrada do erro-médio; MAE = média dos erros absolutos.	9
Figura 3. Mapa de calor de Pearson apresentando uma matriz de correlação entre as variáveis de emissão de dióxido de carbono, atributos atmosféricos e características físico-químicas do solo de todas as áreas estudadas para N = 4. FCO ₂ = Emissão de CO ₂ do solo (μmol CO ₂ m ⁻² s ⁻¹); Ts = Temperatura do solo (°C); Us = Umidade do solo (% do volume); MO = Matéria orgânica (g dm ⁻³); Macro = Macroporosidade do solo (m ³ m ⁻³); VTP = Volume total de poros (%); ARG = Argila (g kg ⁻¹); Tmed = Temperatura média do solo (°C); Tmax = Temperatura máxima do solo (°C); Tmin = Temperatura mínima do solo (°C); Umed = Umidade média do solo (% do volume); Umax = Umidade máxima do solo (% do volume); Umin = Umidade mínima do solo (% do volume); PkPa = Pressão atmosférica (kPa); Rad = Radiação (W m ⁻²); Eto = Evapotranspiração de referência (mm dia ⁻¹); Velmax = Velocidade máxima do vento (m s ⁻¹); Velmin = Velocidade mínima do vento (m s ⁻¹); Dir_vel = Direção da velocidade do vento (°); chuva = Precipitação diária (mm dia ⁻¹); inso = Insolação (hora dia ⁻¹); xCO ₂ = Concentração do dióxido de carbono atmosférico (μmol CO ₂ m ⁻² s ⁻¹); SIF = Fluorescência Induzida pelo Sol (W m ⁻² sr ⁻¹ μm ⁻¹).	12
Figura 4. Análise das correlações negativas entre FCO ₂ e XCO ₂	13
Figura 5. Análise das correlações positivas entre FCO ₂ e SIF.	14
Figura 6. Análise das correlações negativas entre SIF e XCO ₂	14
Figura 7. Resultado da análise de regressão entre os valores observado e o predito para FCO ₂ com o Árvores de Decisão (DT).	15
Figura 8. Resultado da análise de regressão entre os valores observado e o predito para FCO ₂ com o Extreme Gradient Boosting (XGBoost).	16
Figura 9. Resultado da análise de regressão entre os valores observado e o predito para FCO ₂ com o Floresta Aleatória (RF).	16

LISTA DE TABELAS

Tabela 1. Equações da métrica. Observe que nas equações tem-se que: n é o número de observações; y_i é a i -ésima observação; $\hat{y}_i$ é o valor estimado; y_i é o valor observado; valor médio previsto..... 10

Tabela 2. Critérios estatísticos para avaliação do desempenho..... 15

Tabela 1. Equações da métrica. Observe que nas equações tem-se que: n é o número de observações; y_i é a i -ésima observação; $\hat{y}_i$ é o valor estimado; y_i é o valor observado; valor médio previsto..... 10

Tabela 2. Critérios estatísticos para avaliação do desempenho..... 15

SUMÁRIO

RESUMO	x
ABSTRACT	xi
1. INTRODUÇÃO E JUSTIFICATIVA	1
2. OBJETIVOS	2
3. REVISÃO BIBLIOGRÁFICA	2
3.1. Gases do efeito estufa (GEE)	2
3.2. Fatores ambientais e atributos do solo	3
4. MATERIAL E MÉTODOS	5
4.1. Amostragem de Campo	5
4.2. Determinação da emissão de CO ₂ do solo, temperatura e umidade do solo.....	6
4.3. Determinação de variáveis químicas e físicas do solo	6
4.4. Variáveis climáticas e orbitais	7
4.5. Concentrações de CO ₂ atmosférico	8
4.6. Modelos de aprendizado de máquina (ML)	8
4.6.1. Floresta aleatória (RF)	9
4.6.2. Extreme Gradient Boosting (XGBoost)	10
4.7. Métricas de desempenho do modelo	10
5. RESULTADOS E DISCUSSÃO	11
5.1. Correlações	11
5.2. Desempenho dos modelos	14
6. CONCLUSÃO	17
REFERÊNCIAS	19

EMIÇÃO DE CO₂ DO SOLO E SUA RELAÇÃO COM CO₂ ATMOSFÉRICO EM ÁREAS AGRÍCOLAS DO ESTADO DO MATO GROSSO DO SUL: UMA ABORDAGEM DE APRENDIZADO DE MÁQUINA

RESUMO

As projeções indicam um contínuo crescimento do agronegócio brasileiro e consequente incremento nas emissões de gases do efeito estufa (GEE) advindas desse setor. O dióxido de carbono (CO₂) representa cerca de 66% das emissões totais de GEE do planeta, sendo o carbono orgânico do solo um dos principais reservatórios terrestres para o armazenamento e o intercâmbio de carbono (C) atmosférico, uma vez que, dependendo do uso e manejo de solos agrícolas, podem atuar como fontes ou sumidouros desse carbono. Modelar a dinâmica do carbono em áreas agrícolas é uma ação estratégica para diminuir incertezas associadas aos processos de mitigação de GEE e melhorar a capacidade de análises para construção de cenários mais acurados. Nas últimas décadas, técnicas de inteligência artificial e mineração de dados têm sido aplicadas com sucesso na modelagem de inúmeros atributos em ciência do solo. Assim, com o presente estudo objetivou-se avaliar o desempenho preditivo de algoritmos de aprendizado de máquina para o estudo da relação de emissão de dióxido de carbono no solo (FCO₂) e o CO₂ atmosférico (xCO₂) em áreas agrícolas na região do estado do Mato Grosso do Sul (MS), a partir de uma série temporal de 2015 a 2017. Os algoritmos de *Machine Learning* utilizados foram: árvore de decisão (DT), árvores de decisão utilizando os algoritmos *Random Forest* (RF) e *Extreme Gradient Boosting* (XGBoost). Entre os três algoritmos de modelagem para simular a emissão de dióxido de carbono no solo (FCO₂) neste estudo, os modelos de regressão RF e XGBoost foram os modelos de melhor desempenho.

Palavras-chave: inteligência artificial, respiração do solo, mineração de dados, mudanças climáticas.

SOIL CO₂ EMISSION AND ITS RELATIONSHIP WITH ATMOSPHERIC CO₂ IN AGRICULTURAL AREAS OF MATO GROSSO DO SUL STATE: A MACHINE LEARNING APPROACH

ABSTRACT

Projections indicate a continuous growth of Brazilian agribusiness and a consequent increase in greenhouse gas (GHG) emissions from this sector. Carbon dioxide (CO₂) represents about 66% of the planet's total GHG emissions, with soil organic carbon being one of the main terrestrial reservoirs for the storage and exchange of atmospheric carbon (C), as agricultural soils can act as sources or sinks of this carbon depending on their use and management. Modeling carbon dynamics in agricultural areas is a strategic action to reduce uncertainties associated with GHG mitigation processes and improve analysis capabilities for constructing more accurate scenarios. In recent decades, artificial intelligence and data mining techniques have been successfully applied in modeling numerous attributes in soil science. Thus, the present study aimed to evaluate the predictive performance of machine learning algorithms for studying the relationship between soil carbon dioxide emission (FCO₂) and atmospheric CO₂ (xCO₂) in agricultural areas in the state of Mato Grosso do Sul (MS), based on a time series from 2015 to 2017. The machine learning algorithms used were: decision tree (DT), decision trees using the algorithms Random Forest (RF) and Extreme Gradient Boosting (XGBoost). Among the three modeling algorithms to simulate soil carbon dioxide emission (FCO₂) in this study, the RF and XGBoost regression models were the best performing models.

Key Word: artificial intelligence, soil respiration, data mining, climate change.

1. INTRODUÇÃO E JUSTIFICATIVA

O processo de perda de carbono do solo para a atmosfera é denominado respiração do solo ou emissão de CO₂ do solo (FCO₂), resultante da atividade microbiana (oxidação química) e respiração das raízes, considerada a segunda maior fonte de CO₂ para atmosfera, atrás apenas dos oceanos. A FCO₂ é um componente essencial do ciclo biogeoquímico do carbono e, portanto, de fundamental importância para o armazenamento de carbono no solo. Os níveis de carbono orgânico nos solos são resultantes das complexas interações entre variáveis relacionadas aos processos de produção e transporte do gás do solo para a atmosfera. Pequenos incrementos em suas taxas podem ser suficientes para um ecossistema mudar de sumidouro para fonte de carbono para a atmosfera (LAGANIÈRE; ANGERS and PARÉ (2010); FARHATE et al. (2018); SARAGOSA ROSSI et al. (2022)). Por exemplo, a transformação dos ecossistemas naturais por meio do uso e ocupação do solo altera os processos biofísicos e intensificam a perda de carbono via FCO₂ para a atmosfera que se devem às mudanças nos atributos do solo que depende, principalmente, das condições edafoclimáticas (RAICH and SCHLESINGER (1992); THANGAVEL et al. (2018)).

Ampliar o fomento e a adoção de tecnologias e métodos de produção com baixas emissões é fundamental para garantir a redução das emissões absolutas do setor, que continuam aumentando (SEEG11 (2023)). E para tal, se faz necessário o uso de técnicas de análises estáticas de dados para amparar tomadas de decisões em implementação de políticas públicas e privadas, visando mitigar os efeitos das mudanças causadas pelos gases do efeito estufa (GEE).

Devido à grande relevância e a preocupação mundial com relação ao aumento das emissões do GEE, diversos pesquisadores pelo mundo, como (FREITAS et al. (2018)), (MORAIS FILHO et al. (2021)) e (CANTERL (2020)) estão estudando esses processos e uma linha de pesquisa que está crescente é a estimativa e a quantificação das emissões do CO₂ do solo, por meio da modelagem matemática e estatística. A não linearidade das relações existentes entre as variáveis utilizadas nos estudos da emissão dos GEE resultam em processos de grande complexidade, sendo necessário o uso de técnicas

matemáticas e estatísticas que consigam extrair informações de maneira sólida e estruturada.

Assim, a caracterização de como os diferentes usos e manejos dos solos brasileiros afetam a dinâmica FCO_2 é de grande importância para a determinação de forma quantitativa, do potencial de mitigação desse gás, refletindo no impacto dessas práticas nas mudanças climáticas do planeta. Além disso, o melhor entendimento do fenômeno e suas interações são essenciais para a construção de cenários com menores incertezas associadas, podendo assim, ser utilizados na elaboração de estratégias para apoiar ações de mitigação, especialmente em regiões tropicais, onde as emissões são frequentemente mais altas.

A hipótese do presente estudo é que a dinâmica temporal da emissão de CO_2 do solo e do CO_2 atmosférico pode ser descrita como um fenômeno com estruturas multidimensionais que evoluem no tempo, aptas a serem modeladas por algoritmos de aprendizado de máquina. Assim, objetivou-se avaliar o desempenho dos algoritmos árvore de decisão (DT) e árvores de decisão usando *Random Forest* e *Extreme Gradient Boosting* na predição da FCO_2 e sua relação com o CO_2 atmosférico em áreas agrícolas na região do estado do Mato Grosso do Sul. Especificamente, pretende-se usar a abordagem de aprendizado de máquina para investigar os padrões temporais de FCO_2 e CO_2 atmosféricos avaliando as fontes de incerteza e de previsão em áreas agrícolas.

2. OBJETIVOS

O presente estudo objetivou-se avaliar o desempenho preditivo de algoritmos de aprendizado de máquina para o estudo da relação de emissão de dióxido de carbono no solo (FCO_2) e o CO_2 atmosférico (xCO_2) em áreas agrícolas na região do estado do Mato Grosso do Sul (MS), a partir de uma série temporal de 2015 a 2017.

3. REVISÃO BIBLIOGRÁFICA

3.1. Gases do efeito estufa (GEE)

As concentrações atmosféricas de gases do efeito estufa (GEE) têm aumentado a níveis preocupantes. O último boletim da *World Meteorological Organization* (WMO), publicado em novembro de 2023, aponta que as concentrações atmosféricas de dióxido de carbono (CO₂), metano (CH₄) e óxido nitroso (N₂O), atingiram novas máximas no ano de 2023, com CO₂ a 417,9 ppm, CH₄ a 1.923 ppb e N₂O a 335,8 ppb. Esses valores representam, respectivamente, 150%, 264% e 124% dos níveis atmosféricos no período pré-industrial (WMO et al. (2023)). Ainda de acordo com o referido boletim, o desmatamento e demais mudanças no uso da terra registraram média de emissão de 4,5 Gt de CO₂ ano⁻¹ no período de 2012-2021, em comparação, as emissões devido à queima de combustíveis fósseis e produção de cimento foram projetadas em 36,8 Gt de CO₂ em 2022.

No Brasil as emissões de GEE advêm principalmente dos setores da agropecuária associados às mudanças no uso e cobertura da terra (LULUCF- *Land Use, Land Change and Forestry*). No ano de 2020, de acordo com o relatório de estimativas anuais de emissões de gases de efeito estufa, esses setores emitiram juntos cerca de 622,29 Mt de CO₂eq, ou seja, 66,5% das emissões totais de GEE do país (MCTI (2022)). Vislumbra-se, portanto, um cenário preocupante quanto ao compromisso nacional de mitigação das emissões, instituído pela Política Nacional sobre a Mudança do Clima (LEI, 12.187/2009).

O Índice Anual de Gases de Efeito Estufa (AGGI) mostra que de 1990 a 2022, o forçamento radiativo por gases de efeito estufa de longa duração (LLGHGs – *Long-Lived Green House Gases*) aumentou 49%, com o CO₂ sendo responsável por cerca de 78% desse aumento (AGGI (2023)). De acordo com os dados da CAIT/WRI (2022), para as emissões globais, em 2021 o Brasil ocupou o 5º lugar entre os maiores emissores de gases de efeito estufa, com 4,2% do total mundial. Somando-se as emissões de CO₂ da agropecuária com a parcela de LULUCF, conclui-se que a atividade rural, seja direta ou indiretamente, respondeu por 60,63% das emissões do Brasil no ano de 2021 (CAIT/WRI (2022)).

3.2. Fatores ambientais e atributos do solo

Os fatores ambientais e atributos do solo que controlam a magnitude da FCO_2 continuam a ser de difícil separação e interpretação, mesmo após décadas de pesquisa, ainda são poucos modelos matemáticos robustos desenvolvidos para prever o efeito de fatores bióticos e abióticos na dinâmica do carbono do solo (SIERRA et al. (2015); FARHATE et al. (2018)). As razões para essa dificuldade devem-se às numerosas interações entre tais fatores (covariáveis) e FCO_2 , em combinação com diferentes escalas de variação no tempo e no espaço (GRAF et al. (2012); BICALHO et al. (2014)). Na escala temporal a emissão de CO_2 do solo é controlada por atributos como a umidade do solo (LEON et al. (2014); SANTOS et al. (2019), porosidade livre de água (WARNER et al. (2018)), oxigenação do solo (CHEN and CHEN (2010); VICENTINI et al. (2019)) e variáveis climáticas e ambientais como precipitação, temperatura radiação solar e índice de vegetação melhorado (WARNER et al. (2019)). Algoritmos de aprendizado de máquina supervisionado, como as máquinas de vetores de suporte (support vector machines - SVM), árvores de classificação/regressão (CARTs) e redes neurais artificiais (RNAs) são consideradas ferramentas promissoras e robustas, para interpretar dados de alta variabilidade e indicar relações não-lineares no espaço multivariado (MCBRATNEY et al. (2000); WARNER et al. (2019)). Estas técnicas têm sido satisfatoriamente empregadas no desenvolvimento de modelos preditivos para diversas variáveis nas ciências climatológicas, ambientais e agrárias (YILMAZ and KAYNAR (2011); BESALATPOUR et al. (2013); REICHSTEIN et al. (2019); SHADRIN et al. (2020)).

Desde a década de 1960, os métodos matemáticos e estatísticos estão sendo utilizados em pesquisas da ciência do solo, com o intuito de promover o melhor entendimento das dinâmicas do solo e gerando ferramentas de avaliação que auxiliam na resolução de problemas agrícolas e ambientais (MCBRATNEY; GRUIJTER, 1992). Em função desta demanda, tem-se observado o aumento do uso das técnicas de *Machine Learning* associadas a agricultura, devido a sua grande robustez são gerados resultados que podem auxiliar na compreensão da variabilidade espaçotemporal do CO_2 , por meio dos modelos de *Machine Learning* como máquina de vetores de suporte (SVM) e árvores de decisão (*Random Forest* e *Extreme Gradient Boosting*). Há pesquisas atuais referentes

a emissão de CO₂ do solo, como exemplo, aprendizado de máquina para a previsão da emissão do CO₂ do solo (TAVARES et al., 2018), no entanto, faltam estudos sobre a previsão da CO₂ utilizando variáveis oriundas do solo, planta e atmosfera, usando técnicas de aprendizado de máquina.

4. MATERIAL E MÉTODOS

4.1. Amostragem de Campo

O conjunto de dados apresentado foi derivado de quatro experimentos de campo realizados entre novembro de 2015 e maio de 2017. Estes experimentos foram conduzidos em áreas florestais que consistiam em plantações de eucalipto (*Eucalyptus camaldulensis*) e pinus (*Pine caribaea* var. *hondurensis*) estabelecidas em 1986, áreas reflorestadas com 35 espécies nativas distribuídas de forma aleatória, e uma área destinada à integração lavoura-pecuária-floresta (ILPF).

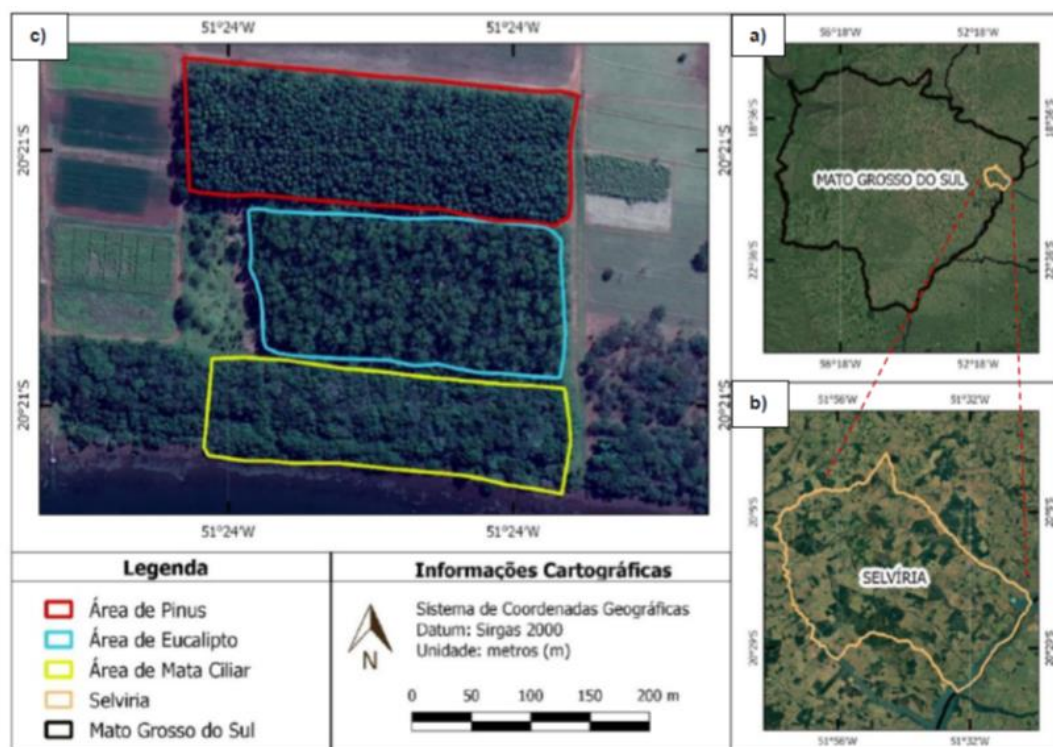


Figura 1. Mapa das áreas de estudo localizadas no município de Selvíria, MS. a) Estado do Mato Grosso do Sul, Brasil Central; b) Município de Selvíria; c) Áreas de florestas plantadas: pinus, eucalipto e mata ciliar, respectivamente.

Os experimentos foram instalados na Fazenda Experimental da Faculdade de Engenharia da UNESP, situada no município de Selvíria, estado de Mato

Grosso do Sul, Brasil. Esta área específica está localizada na microrregião de Paranaíba, com coordenadas geográficas de 20° 20' 53,41" Sul e 51° 23' 55,50" Oeste, e uma elevação de 354 metros acima do nível do mar. O solo foi classificado como Latossolo vermelho-escuro de textura média. A topografia da região é caracterizada como moderadamente plana e ondulada, e tem origem em solos basálticos, conforme definido pela Taxonomia de Solos (STAFF, 2014). Em termos climáticos, a região é classificada como tropical úmido (Aw) segundo classificação climática de acordo com Köppen, caracterizada por invernos secos, a média anual de precipitação é de 1.370 mm, temperatura média de 23,5 °C e umidade relativa variando entre 70 e 80%.

4.2. Determinação da emissão de CO₂ do solo, temperatura e umidade do solo

Em todos os ensaios a emissão de CO₂ do solo foi monitorada utilizando o sistema LI-COR (LI-8100; LI-COR Bioscience, Nebraska, EUA), este sistema opera em um modo de medição que acompanha as variações na concentração de CO₂ dentro de uma câmara para solos, utilizando espectroscopia no espectro infravermelho. A câmara foi acoplada sobre colares de policloreto de vinila (PVC), que foram previamente inseridos no solo em cada ponto de amostragem a uma profundidade de 0,03 m. O fluxo de CO₂ foi calculado em cada ponto através de um ajuste da concentração de CO₂ do ar dentro da câmara, usando uma regressão exponencial ao longo do tempo após o fechamento da câmara. A temperatura do solo foi registrada simultaneamente às medições de respiração do solo, utilizando um termômetro digital do tipo espeto. A umidade do solo foi determinada usando um sistema de Reflectometria no Domínio do Tempo (TDR) (Hydrosense TM, Campbell Scientific, Austrália), que consiste em duas sondas de 0,12 m que são inseridos no solo a 0,10 m dos colares de PVC.

4.3. Determinação de variáveis químicas e físicas do solo

Os dados referentes aos atributos físicos e químicos do solo são oriundos de amostras de solo coletadas na profundidade de 0,0-0,10 m e 0,10-0,20 m. Foram realizadas as seguintes análises de rotina: pH, determinação do teor de matéria orgânica do solo (MOS), fósforo disponível (P), potássio (K⁺), cálcio

(Ca²⁺) e magnésio (Mg²⁺) segundo o método de (RAIJ et al., 2001). A partir dos teores desses elementos foram calculados os seguintes atributos: soma de bases (SB), acidez potencial (H+Al), capacidade de troca de cátions (CTC) e a saturação por bases (V%). O carbono orgânico foi determinado pelo método da combustão úmida, via colorimétrica de RAIJ et al. (1987) e, a determinação do teor de nitrogênio por meio de digestão sulfúrica por MALAVOLTA; VITTI and OLIVEIRA (1997). As amostras indeformadas foram coletadas com amostrador adaptado a cilindros com dimensões médias de 0,05 m de diâmetro interno e 0,04 m de altura (EMBRAPA, 1997). A macroporosidade (Macro) e microporosidade (Micro) foram determinadas por meio de mesa de tensão com 0,60 m de altura de coluna d'água em amostras previamente saturadas. O volume de água retido na amostra nesta condição corresponde à Macro e a Micro foi determinada após a retirada dos anéis da estufa à 105 °C num período de 24h e posterior pesagem. A porosidade total (PT) foi calculada pela soma dos macroporos e microporos e a densidade do solo (Ds) utilizando-se metodologia de EMBRAPA (1997). Uma vez que as amostras de solo foram coletadas em uma camada fixa, os estoques de carbono (EstC) foram ajustados para as mudanças na densidade do solo que ocorrem após as mudanças do uso da terra. Para isso, este estudo utilizou a metodologia descrita por ELLERT and BETTANY (1995) e MORAES et al. (1996) para corrigir os estoques de carbono do solo em uma profundidade de massa equivalente, ou seja, a profundidade do solo das áreas manejadas que contém a mesma massa de solo como a camada correspondente (0 – 0,20 m) na área de vegetação nativa (área de referência específica para cada região). Os cálculos da camada de solo equivalente foram realizados de acordo com o apresentado por CARVALHO et al. (2009) e SEGNINI et al. (2013).

4.4. Variáveis climáticas e orbitais

As variáveis climáticas utilizadas no presente estudo foram: precipitação diária (mm dia⁻¹); velocidade do vento (m s⁻¹); evapotranspiração por Penman Monteith (ET_o PN^M) (mm dia⁻¹); radiação fotossinteticamente ativa (PAR) (μmol m⁻² s⁻¹); radiação global (GR) (MJ m² dia⁻¹); pressão atmosférica (P_{atm}) (kpa);

temperatura atmosférica (T_{atm}) ($^{\circ}\text{C}$) e o índice de vegetação melhorado (EVI). A radiação PAR é a radiação incidente na faixa de ondas de 400 a 700 nm, que pode ser absorvida pelo sistema fotossintético das plantas. Seu valor é responsável por aproximadamente 50% da radiação solar. Tais informações foram adquiridas no período compreendido de novembro de 2015 a maio de 2017 a partir de estações meteorológicas próximas às regiões dos ensaios experimentais.

4.5. Concentrações de CO_2 atmosférico

A fim de uma determinação acurada da concentração de CO_2 atmosférico foi associado dados de sensores orbitais com os dados coletados em campo, empregando-se um processo que estabeleceu um padrão entre os dados dos sensores e os dados coletados em campo a partir das coordenadas obtidas na plataforma CO_2 Virtual Science Data Environment.

Os dados foram obtidos de sensores orbitais do Orbiting Carbon Observatory-2 (OCO-2) utilizando a espaçonave LEOStar-2 que carrega um único instrumento. Ele incorpora três espectrômetros de alta resolução que fazem medições coincidentes da luz solar refletida no CO_2 do infravermelho próximo a 1,61 e 2,06 micrômetros e na banda A do oxigênio molecular (O_2) a 0,76 micrômetros. Cada banda possui 1.016 elementos espectrais, foi lançado em julho de 2014, com resolução espacial de $2,25 \times 1,29$ km, resolução temporal de 16 dias, na versão OCO2_L2_Lite_FP_10r de acordo com NASA/OCO-2 (2019); utilizamos durante o intervalo de maio de 2015 a novembro de 2017.

4.6. Modelos de aprendizado de máquina (ML)

O aprendizado de máquina (ML) é um fragmento da inteligência artificial (IA) que busca prever processos de classificação, regressão e decisão, por meio de diversos modelos matemáticos (ALPAYDIN, 2020). Na Figura 2 observa-se um resumo das etapas realizadas para a determinação da construção dos algoritmos de ML.

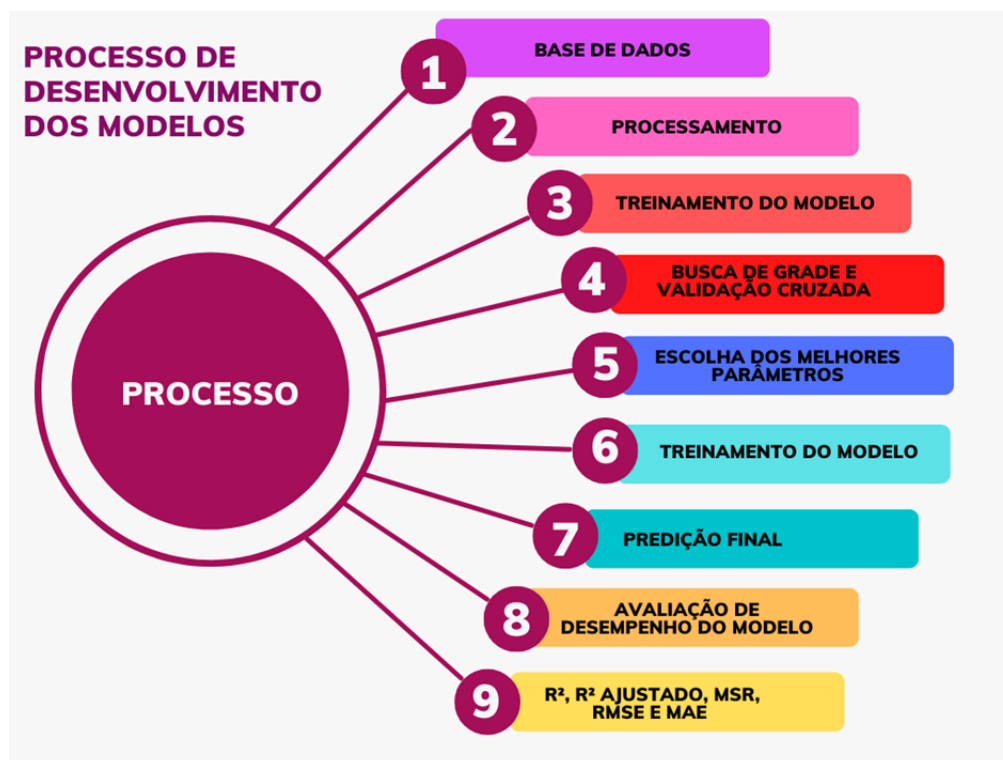


Figura 2. Resumo do processo de desenvolvimento dos modelos para prever a emissão de CO₂ do solo, por meio dos algoritmos de aprendizado de máquina. R^2 = coeficiente de determinação; R^2_{ajust} = coeficiente de determinação ajustado; MSE = erro quadrático médio; RMSE = raiz quadrada do erro-médio; MAE = média dos erros absolutos.

4.6.1. Floresta aleatória (RF)

São modelos estruturados em árvores utilizados tanto para classificação, quanto para regressão, sendo compostos por uma combinação de árvores preditoras (floresta) que são geradas a partir de um vetor aleatório, amostrado de forma independente e com a mesma distribuição para todas as árvores na floresta (BREIMAN, 2001). Na construção de uma árvore de regressão o modelo baseia-se na divisão das variáveis explicativas, conforme suas semelhanças e dessemelhanças e o seu resultado final é formado pela média dos resultados de todas as árvores, é uma técnica estatística não paramétrica de aprendizado (DE CAMPOS, 1979). No seu processo de aprendizagem, três hiper-parâmetros precisam ser otimizados: o número de árvores cultivadas na regressão (n_{tree}); o número mínimo de dados em cada nó terminal ($nodesize$) e o número de diferentes preditoras testadas em cada nó ($mtry$) (LIAW; WIENER, 2002). Para

a primeira abordagem são adotados o número de estimadores das árvores na floresta foi igual a 300 e a profundidade máxima da árvore foi igual a 10. O número de covariáveis usadas em cada árvore foi um terço do número total de amostras.

4.6.2. Extreme Gradient Boosting (XGBoost)

Baseia-se no conceito de obtenção de algoritmos de classificação/regressão fortes, a partir de transformação em algoritmos de baixo desempenho (CHEN; GUESTRIN, 2016). Assim como no algoritmo Random Forest, ocorre a distribuição dos pesos às árvores individuais para, posteriormente, gerar um consenso a partir da recombinação das mesmas. Nessa abordagem, o peso de um classificador será maior quanto pior for o desempenho de seu antecessor, previamente ajustado ao longo do processo de aprendizado (FRIEDMAN, 1999; HASTIE; TIBSHIRANI; FRIEDMAN, 2001). Esse método gera uma sequência de árvores que usam a informação contida no treinamento de suas antecessoras, ou seja, o algoritmo vai "aprendendo com o erro". Os hiperparâmetros são ajustados para o número de árvores usadas para construir o modelo, redução, interação.

4.7. Métricas de desempenho do modelo

O desempenho preditivo dos modelos foi avaliado por meio dos indicadores estatísticos: coeficiente de correlação (r – equação 1), coeficiente de determinação (R^2 – equação 2), erro quadrático médio (MSE – equação 3), raiz do erro quadrático médio (RMSE - equação 4), erro médio absoluto (MAE - equação 5), erro absoluto médio percentual (MAPE - equação 6).

Tabela 1. Equações da métrica. Observe que nas equações tem-se que: n é o número de observações; y_i é a i -ésima observação; $\hat{y}_i$ é o valor estimado; $\bar{y}$ é o valor observado; valor médio previsto.

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{(\sum (x_i - \bar{x})^2)(\sum (y_i - \bar{y})^2)}} \quad (1)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (2)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (5)$$

$$MAPE = \left[\frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \right] \quad (6)$$

5. RESULTADOS E DISCUSSÃO

5.1. Correlações

As correlações da FCO₂ foram influenciadas pelos atributos XCO₂ e SIF, devido à complexidade do processo, foi realizado um estudo mais aprofundado para essas variáveis (*Figura 3*).

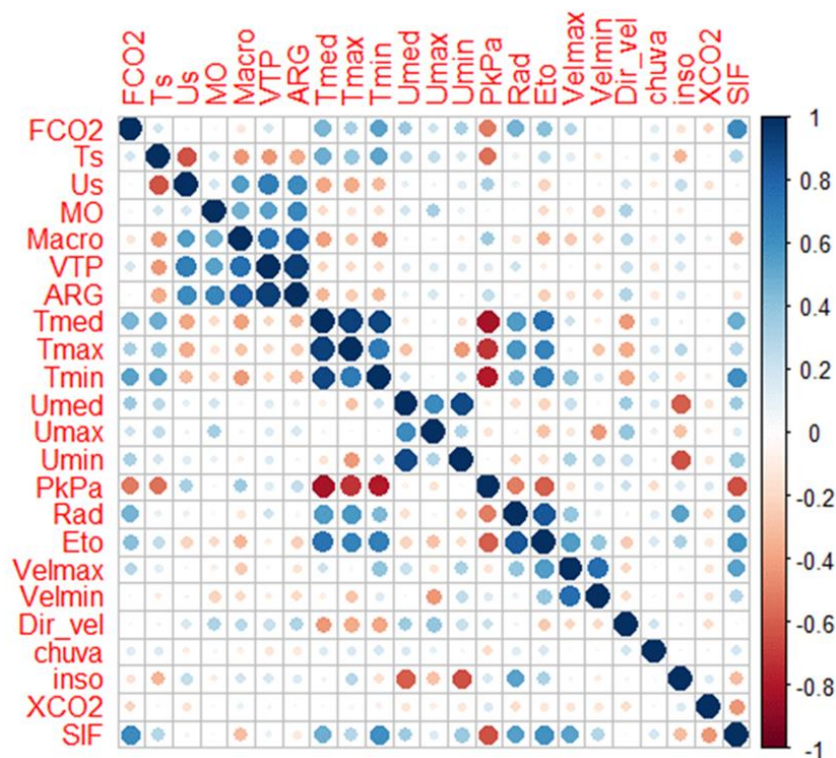


Figura 3. Mapa de calor de Pearson apresentando uma matriz de correlação entre as variáveis de emissão de dióxido de carbono, atributos atmosféricos e características físico-químicos do solo de todas as áreas estudadas para $N = 4$. FCO₂ = Emissão de CO₂ do solo ($\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$); Ts = Temperatura do solo ($^{\circ}\text{C}$); Us = Umidade do solo (% do volume); MO = Matéria orgânica (g dm^{-3}); Macro = Macroporosidade do solo ($\text{m}^3 \text{ m}^{-3}$); VTP = Volume total de poros (%); ARG = Argila (g kg^{-1}); Tmed = Temperatura média do solo ($^{\circ}\text{C}$); Tmax = Temperatura máxima do solo ($^{\circ}\text{C}$); Tmin = Temperatura mínima do solo ($^{\circ}\text{C}$); Umed = Umidade média do solo (% do volume); Umax = Umidade máxima do solo (% do volume); Umin = Umidade mínima do solo (% do volume); PkPa = Pressão atmosférica (kPa); Rad = Radiação (W m^{-2}); Eto = Evapotranspiração de referência (mm dia^{-1}); Velmax = Velocidade máxima do vento (m s^{-1}); Velmin = Velocidade mínima do vento (m s^{-1}); Dir_vel = Direção da velocidade do vento ($^{\circ}$); chuva = Precipitação diária (mm dia^{-1}); inso = Insolação (hora dia^{-1}); xCO₂ = Concentração do dióxido de carbono atmosférico ($\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$); SIF = Fluorescência Induzida pelo Sol ($\text{W m}^{-2} \text{ sr}^{-1} \mu\text{m}^{-1}$).

Três preditores (SIF e XCO₂), (SIF e FCO₂) e (FCO₂ e XCO₂); foram moderadamente correlacionados ($R^2 = 0,18, 0,47$ e $0,32$; respectivamente). Quando as variáveis apresentaram correlações negativas, observou-se um aumento no índice de uma variável resultava em uma diminuição proporcional na outra. Por outro lado, nas relações positivas, tem-se o aumento diretamente proporcional para as duas variáveis. Devido à complexidade do processo, foi realizado um estudo mais aprofundado para avaliarmos a interação entre os atributos FCO₂, XCO₂ e SIF, apresentados nas Figura 4, Figura 5 e Figura 6.

A FCO_2 apresentou correlações negativas com a XCO_2 (Figura 4), confirmando que conforme vai diminuindo as emissões de dióxido de carbono no solo (FCO_2) aumenta diretamente proporcional a concentração de CO_2 . A FCO_2 também apresentou correlações positivas com a SIF (Figura 5), constatando que conforme vai aumentando as emissões de dióxido de carbono no solo (FCO_2) a fluorescência induzida pelo sol (SIF) aumenta diretamente proporcional, devido a atividade fotossintética realizada pela vegetação. A SIF apresentou correlações negativas com a XCO_2 (Figura 6), por isso, entende-se que a diminuição da fluorescência induzida pelo sol é devido a menor atividade fotossintética, devido a diminuição da área vegetal, como exemplo, o que leva a concentração de CO_2 atmosférico aumentar significativamente.

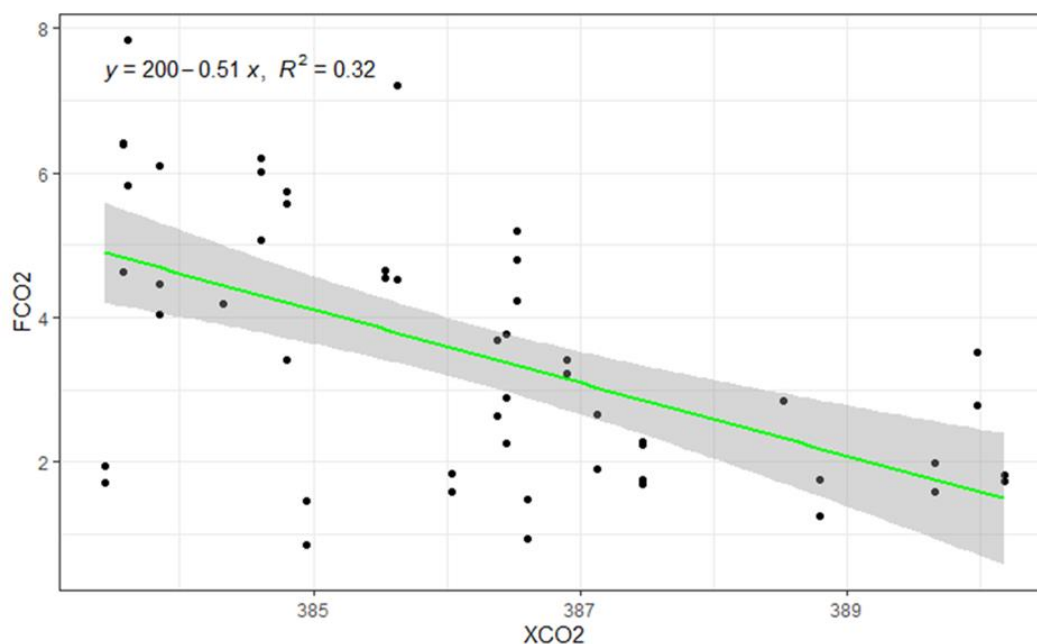


Figura 4. Análise das correlações negativas entre FCO_2 e XCO_2 .

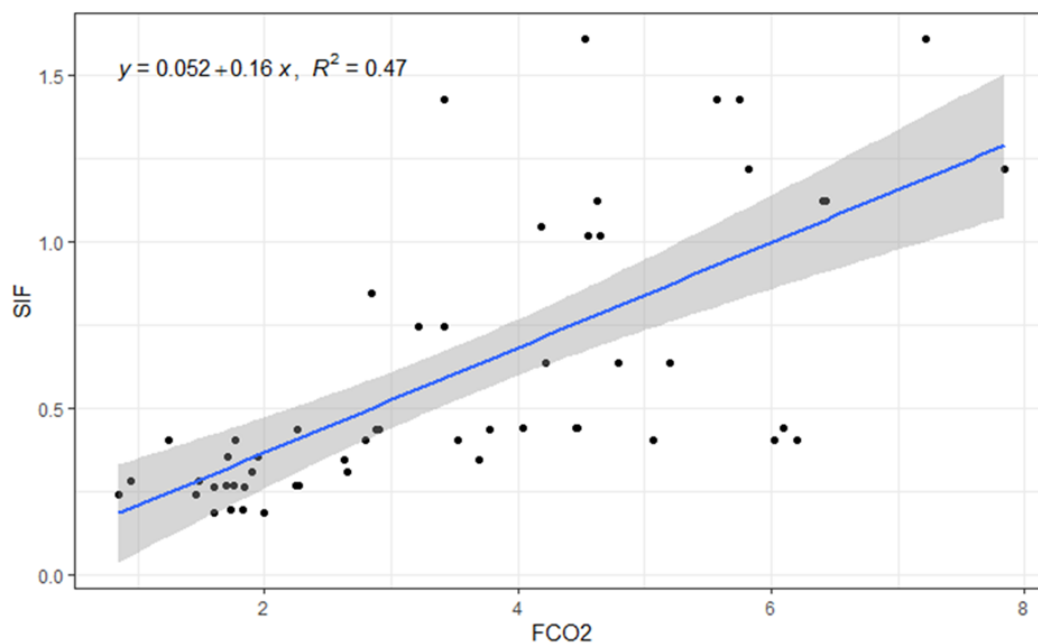


Figura 5. Análise das correlações positivas entre FCO_2 e SIF.

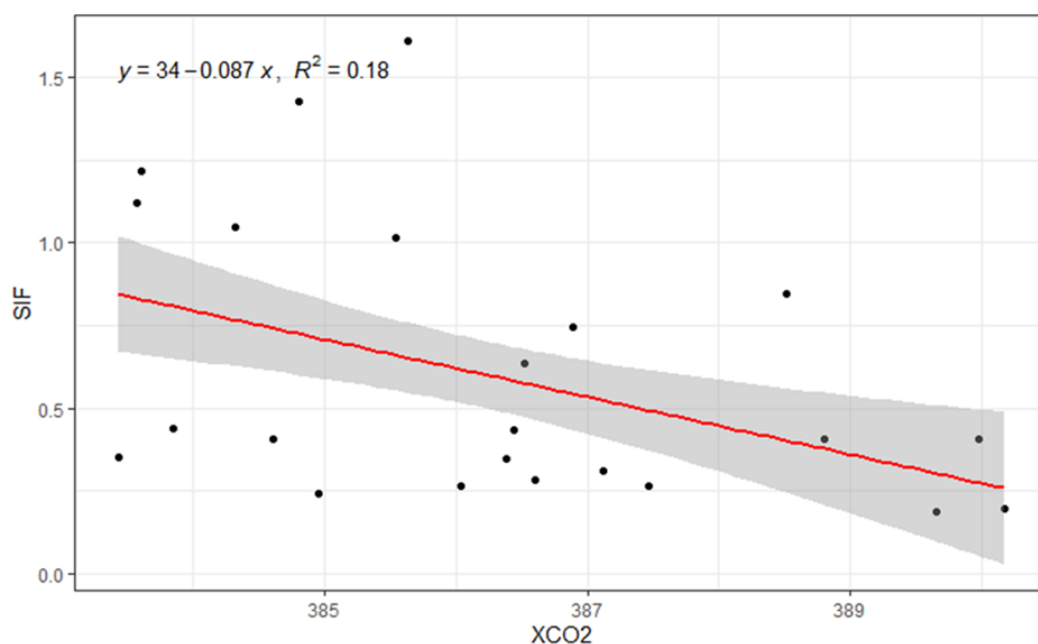


Figura 6. Análise das correlações negativas entre SIF e XCO_2 .

5.2. Desempenho dos modelos

Dentre os parâmetros do modelo, foram usados três critérios estatístico para avaliar o desempenho dos modelos resultantes: coeficiente de determinação, erro absoluto médio percentual e raiz do erro quadrático médio (R^2 , MAPE e RMSE, respectivamente). A Tabela 2 apresenta as métricas de

avaliação para os modelos preditivos distintos de Árvore de Decisão, *Extreme Gradient Boosting* e Floresta Aleatória, respectivamente. SILVA (2022) encontrou valores de RMSE iguais a $0,38 \mu\text{mol m}^{-2} \text{s}^{-1}$ para XGBoost e $0,23 \mu\text{mol m}^{-2} \text{s}^{-1}$ em RF para previsão da FCO_2 do solo. Corroborando com nosso resultado dos quais a Floresta Aleatória ($1,01 \mu\text{mol m}^{-2} \text{s}^{-1}$) foi superior ao *Extreme Gradient Boosting* ($1,00 \mu\text{mol m}^{-2} \text{s}^{-1}$).

Tabela 2. Critérios estatísticos para avaliação do desempenho.

	R^2	MAPE	RMSE
Árvores de decisão	0,66	27,03%	$1,17 \mu\text{mol m}^{-2} \text{s}^{-1}$
<i>Extreme Gradient Boosting</i>	0,70	26,76%	$1,00 \mu\text{mol m}^{-2} \text{s}^{-1}$
Floresta aleatória	0,69	26,14%	$1,01 \mu\text{mol m}^{-2} \text{s}^{-1}$

Os resultados desta análise de modelagem indicam que o modelo DT foi o modelo de menor desempenho.

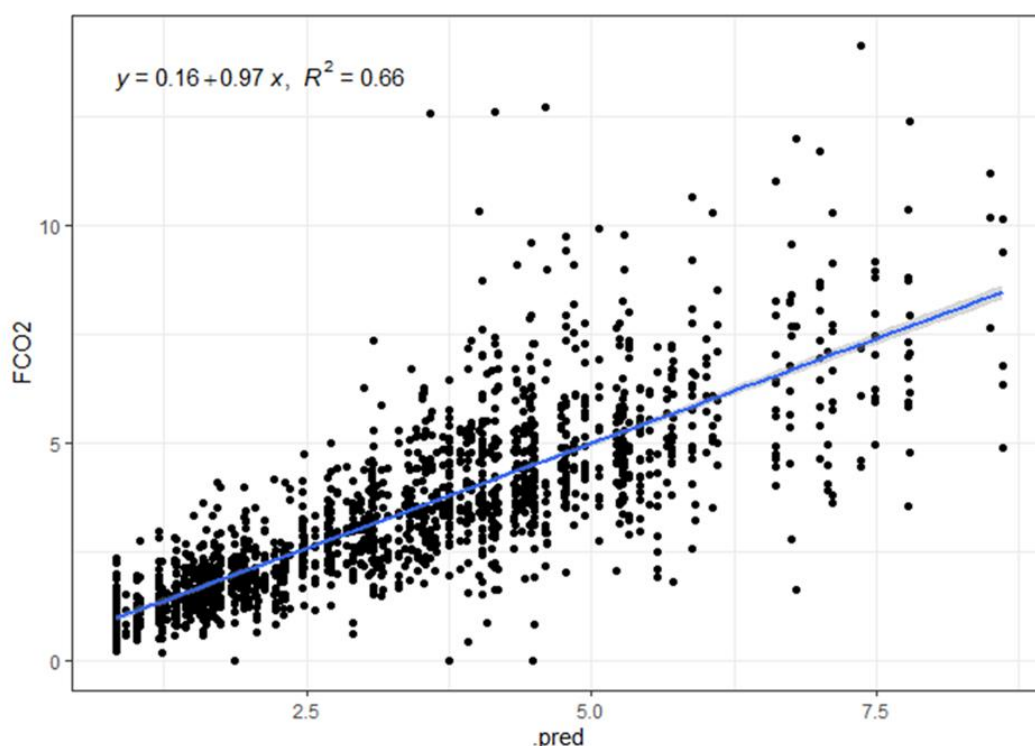


Figura 7. Resultado da análise de regressão entre os valores observado e o predito para FCO_2 com o Árvore de Decisão (DT).

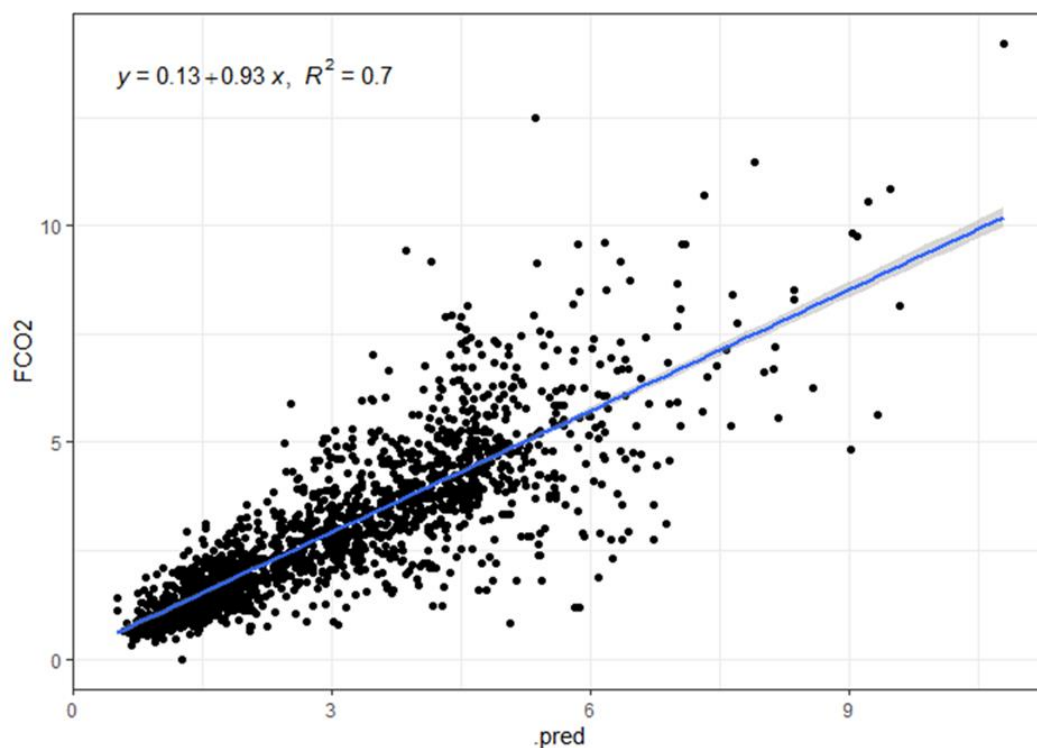


Figura 8. Resultado da análise de regressão entre os valores observado e o predito para FCO₂ com o *Extreme Gradient Boosting* (XGBoost).

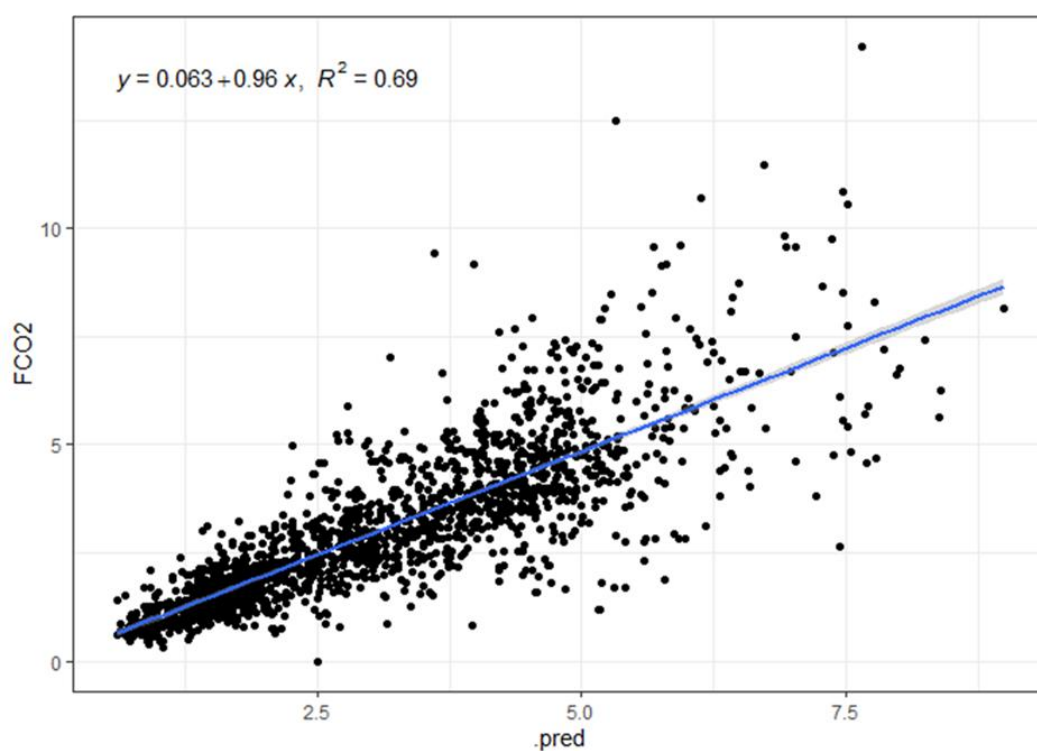


Figura 9. Resultado da análise de regressão entre os valores observado e o predito para FCO₂ com o Floresta Aleatória (RF).

Entre os três algoritmos de modelagem para simular a emissão de dióxido de carbono no solo (FCO_2) neste estudo, os modelos de regressão RF e XGBOOST foram os modelos de melhor desempenho quando usados para prever a emissão de dióxido de carbono no solo (FCO_2) em um conjunto de dados. HAMRANI.; AKBARZADEH. and MADRAMOOTOO. (2020) relatam um R^2 de 0,96 e 0,75 durante o treinamento e teste usando um modelo de RF para prever fluxos de CO_2 e ADJUIK and DAVIS (2022) relata um R^2 de 0,82. A combinação de árvores cultivadas usando variáveis aleatórias no caso do algoritmo RF aumenta a precisão porque as variáveis são escolhidas aleatoriamente em cada nó (BREIMAN, 2001). O algoritmo de regressão XGBoost e o modelo de RF resultou em métricas de desempenho comparativamente altas no conjunto de dados de treinamento quando comparado aos algoritmos de regressão DT.

6. CONCLUSÃO

Os melhores modelos que a predição da FCO_2 , em função dos atributos do sistema solo-planta-atmosfera são o *Random Forest* e o *Extreme Gradient Boosting* (XGBoost).

Estudos nessa linha de pesquisa são necessários para colaborar com a compreensão em relação a mitigação não apenas do CO_2 do solo, mas também dos demais gases agentes do efeito estufa. Processos como a mineração dos dados, seleção, organização e as escolhas dos melhores parâmetros nos algoritmos de RF são decisórios para obtenção de modelos com as melhores acurácias e precisões.

Dessa forma, se conclui que estudos como este são de extrema importância para o setor da agricultura, e poderes públicos e privados, pois as informações geradas podem ser utilizadas no gerenciamento e nas tomadas de decisões do processo produtivo e financeiro, sendo uma alternativa para a caracterização não só da FCO_2 , mas também de outros atributos do sistema solo-planta-atmosfera, promovendo a redução de custos, o uso eficiente dos

recursos infinitos e finitos, rentabilidade, produtividade e maior sustentabilidade para o sistema produtivo como um todo.

REFERÊNCIAS

ADJUIK, T. A.; DAVIS, S. C. Machine Learning Approach to Simulate Soil CO₂ Fluxes under Cropping Systems. **Agronomy**, v. 12, n. 1, p. 197, 2022.

AGGI, 2023. THE NOAA ANNUAL GREENHOUSE GAS INDEX (AGGI), NOAA Global Monitoring Laboratory, R/GML, 325 Broadway, Boulder, CO 80305-3328 ed.

ALPAYDIN, E. **Introduction to machine learning**. MIT press, 2020. p.

BESALATPOUR, A. A.; AYOUBI, S.; HAJABBASI, M. A.; MOSADDEGHI, M. R.; SCHULIN, R. Estimating wet soil aggregate stability from easily available properties in a highly mountainous watershed. **Catena**, v. 111, n., p. 72-79, 2013.

BICALHO, E.; PANOSSO, A.; TEIXEIRA, D.; MIRANDA, J. G.; PEREIRA, G.; LA SCALA JR, N. Spatial variability structure of soil CO₂ emission and soil attributes in a sugarcane area. **Agriculture, Ecosystems & Environment**, v. 189, n., p. 206–215, 2014.

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

CAIT/WRI, 2022. Climate Watch Historical GHG Emissions, In: Institute, W. R. (Ed.), Washington, DC: World Resources Institute. ed.

CANTERL, K. F. F. **Aprendizado de máquina na modelagem temporal da emissão de co₂ do solo em áreas agrícolas no cerrado brasileiro**. Faculdade de Ciências Agrárias e Veterinárias – UNESP, 2020. p.

CARVALHO, J. L. N.; CERRI, C. E. P.; FEIGL, B. J.; PICCOLO, M. C.; GODINHO, V. P.; CERRI, C. C. Carbon sequestration in agricultural soils in the Cerrado region of the Brazilian Amazon. **Soil & Tillage Research**, v. 103, n. 2, p. 342-349, 2009.

CHEN, G.; CHEN, Z.-M. Carbon emissions and resources use by Chinese economy 2007: A 135-sector inventory and input-output embodiment. **Communications in Nonlinear Science and Numerical Simulation**, v. 15, n., p. 3647-3732, 2010.

CHEN, T.; GUESTRIN, C., 2016. XGBoost: A Scalable Tree Boosting System, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing Machinery, San Francisco, California, USA, pp. 785–794.

DE CAMPOS, H. **Estatística experimental não-paramétrica**. ESALQ, 1979. p.

ELLERT, B. H.; BETTANY, J. R. Calculation of organic matter and nutrients stored in soils under contrasting management regimes. **Canadian Journal of Soil Science**, v. 75, n., p. 529-538, 1995.

EMBRAPA Manual de métodos de análise de solo. **Brasília: Ministério da Agricultura e do Abastecimento / EMBRAPA-CNPq**, v. 2, n., p. 212, 1997.

FARHATE, C. V. V.; SOUZA, Z. M. D.; OLIVEIRA, S. R. D. M.; TAVARES, R. L. M.; CARVALHO, J. L. N. Use of data mining techniques to classify soil CO₂ emission induced by crop management in sugarcane field. **PLoS ONE**, v. 13, n. 3, p. e0193537, 2018.

FREITAS, L. P. S.; LOPES, M. L. M.; CARVALHO, L. B.; PANOSSO, A. R.; LA SCALA JUNIOR, N.; FREITAS, R. L. B.; MINUSSI, C. R.; LOTUFO, A. D. P. Forecasting the spatiotemporal variability of soil CO₂ emissions in sugarcane areas in southeastern Brazil using artificial neural networks. **Environ Monit Assess**, v. 190, n. 12, p. 741, 2018.

FRIEDMAN, J. Greedy function approximation: the gradient boosting machine. **Technical report, Stanford University**, v., n., p., 1999.

GRAF, A.; HERBST, M.; WEIHERMÜLLER, L.; HUISMAN, J.; PROLINGHEUER, N.; BORNEMANN, L.; VEREECKEN, H. Analyzing spatiotemporal variability of heterotrophic soil respiration at the field scale using orthogonal functions. **Geoderma**, v. 181–182, n., p. 91–101, 2012.

HAMRANI, A.; AKBARZADEH, A.; MADRAMOOTOO, C. A. Machine learning for predicting greenhouse gas emissions from agricultural soils. **Science**, v. 741, n., p., 2020.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. The Elements of Statistical Learning. **Data mining, Inference and Prediction**, Springer Verlag, New York, v., n., p., 2001.

LAGANIÈRE, J.; ANGERS, D. A.; PARÉ, D. Carbon accumulation in agricultural soils after afforestation: a meta-analysis. **Global Change Biology**, v. 16, n. 1, p. 439-453, 2010.

LEON, E.; VARGAS, R.; BULLOCK, S.; LOPEZ, E.; PANOSSO, A.; LA SCALA JR, N. Hot spots, hot moments, and spatio-temporal controls on soil CO₂ efflux in a water-limited ecosystem. **Soil Biology and Biochemistry**, v. 77, n., p. 12–21, 2014.

LIAW, A.; WIENER, M. Classification and regression by randomForest. **R News**, v. 2/3, n., p. 18-22, 2002.

MALAVOLTA, E.; VITTI, G. C.; OLIVEIRA, S. A. D. **Avaliação do estado nutricional das plantas: princípios e aplicações**. Piracicaba: Potafós, 1997. p.

MCBRATNEY; GRUIJTER, J. J. D.

A continuum approach to soil classification by modified fuzzy k-means with extragrades. **Journal of Soil Science**, v. 43, n., p. 159-175, 1992.

MCBRATNEY, A. B.; ODEH, I. O. A.; BISHOP, T. F. A.; DUNBAR, M. S.; SHATAR, T. M. An overview of pedometric techniques for use in soil survey. **Geoderma**, v. 97, n., p. 293-327, 2000.

MCTI, 2022. Estimativas anuais de emissões de gases de efeito estufa no brasil, In: Ministro da ciência, t., inovações e comunicações (Ed.), 6º ed, Brasília

MORAES, J. F. L.; VOLKOFF, B.; CERRI, C. C.; BERNOUX, M. Soil properties under Amazon forest and changes due to pasture installation in Rondonia, Brazil. **Geoderma**, v. 70, n., p. 63-81, 1996.

MORAIS FILHO, L. F. F.; MENESES, K. C.; SANTOS, G. A. A.; BICALHO, E. D. S.; ROLIM, G. S.; LA SCALA, N., JR. xCO₂ temporal variability above Brazilian agroecosystems: A remote sensing approach. **J Environ Manage**, v. 288, n., p. 112433, 2021.

NASA/OCO-2, 2019. OCO-2 Science Team / Michael Gunson, Annmarie Eldering, Efemérides da nave espacial OCO-2 Nível 0, Processamento retrospectivo V10r, Greenbelt, MD, EUA, Goddard Earth Sciences Data and Information Services Center (GES DISC).

RAICH, J. W.; SCHLESINGER, W. H. The global carbon dioxide flux in soil respiration and its relationship to vegetation and climate. **Tellus B: Chemical and Physical Meteorology**, v., n., p., 1992.

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

RAIJ, B. V.; QUAGGIO, J. A.; CANTARELLA, H.; FERREIRA, M. E. L.; BATAGLIA, O. C. **Análise química de solos para fins de fertilidade**. Campinas: Fundação Cargill, 1987. p.

REICHSTEIN, M.; CAMPS-VALLS, G.; STEVENS, B.; JUNG, M.; DENZLER, J.; CARVALHAIS, N.; PRABHAT Deep learning and process understanding for data-driven Earth system science. **Nature**, v. 566, n., p. 195-204, 2019.

SANTOS, G. D. A.; MOITINHO, M.; SILVA, B.; XAVIER, C.; TEIXEIRA, D.; CORÁ, J.; LA SCALA JR, N. Effects of long-term no-tillage systems with different succession cropping strategies on the variation of soil CO₂ emission. **Science of The Total Environment**, v. 686, n., p., 2019.

SARAGOSA ROSSI, F.; DE ARAÚJO SANTOS, G.; MARIA, L.; LOURENÇONI, T.; PELISSARI, T.; DELLA SILVA, J. L.; OLIVEIRA JÚNIOR, J. W.; SILVA, A. D. A.; LIMA, M.; TEODORO, P.; RIBEIRO, L.; DE OLIVEIRA-JÚNIOR, J.; LA SCALA JR, N.; SILVA JUNIOR, C. A. Carbon dioxide spatial variability and dynamics for contrasting land uses in central Brazil agricultural frontier from remote sensing data. **Journal of South American Earth Sciences**, v. 116, n., p. 103809, 2022.

SEEG11, 2023. Análise das emissões de gases de efeito estufa e suas implicações para as metas climáticas do Brasil.

SEGNINI, A.; CARVALHO, J. L. N.; BOLONHEZI, D.; MILORI, D.; SILVA, W. T. L. D.; SIMOES, M. L.; CANTARELLA, H.; MARIA, I. C. D.; MARTIN-NETO, L. Carbon stock and humification index of organic matter affected by sugarcane straw and soil management. **Scientia Agricola**, v. 70, n. 5, p. 321-326, 2013.

SHADRIN, D.; PUKALCHIK, M.; KOVALEVA, E.; FEDOROV, M. Artificial intelligence models to predict acute phytotoxicity in petroleum contaminated soils. **Ecotoxicology and Environmental Safety**, v. 194, n., p., 2020.

SIERRA, C.; TRUMBORE, S.; DAVIDSON, E.; VICCA, S.; JANSSENS, I. Sensitivity of decomposition rates of soil organic matter with respect to simultaneous changes in temperature and moisture. **Journal of Advances in Modeling Earth Systems**, v. 7, n., p., 2015.

SILVA, P. A. D. **Utilização de diferentes técnicas de machine learning no processo de determinação da emissão da CO₂ do solo, em função dos atributos do sistema solo-planta-atmosfera.** Faculdade de Ciências Agrárias e Veterinárias – UNESP, Câmpus de Jaboticabal, 2022. 212p.

STAFF, S. S. Soil Survey Field and Laboratory Methods Manual., In: Soil Survey Investigations Report, V. (Ed.), **R. Burt and Soil Survey Staff (ed.):** U.S. Department of Agriculture, Natural Resources Conservation Service., 2014. p.

TAVARES, R.; OLIVEIRA, S.; BARROS, F.; VIEIRA, C.; SOUZA, Z.; LA SCALA JR, N. Prediction of soil CO₂ flux in sugarcane management systems using the Random Forest approach. **Scientia Agricola**, v. 75, n., p. 281-287, 2018.

THANGAVEL, R.; KANCHIKERIMATH, M.; SUDHARSANAM, A.; AYYANADAR, A.; KARUNANITHI, R.; DESHMUKH, N. A.; VANAIO, N. S. Evaluating organic carbon fractions, temperature sensitivity and artificial neural network modeling of CO₂ efflux in soils: Impact of land use change in subtropical India (Meghalaya). **Ecological Indicators**, v. 93, n., p. 129-141, 2018.

VICENTINI, M. E.; PINOTTI, C.; HIRAI, W.; MORAES, M.; MONTANARI, R.; TEIXEIRA FILHO, M.; MILORI, D.; LA SCALA JR, N.; PANOSSO, A. CO₂ emission and its relation to soil temperature, moisture, and O₂ absorption in the

reforested areas of Cerrado biome, Central Brazil. **Plant and Soil**, v. 444, n., p. 1-19, 2019.

WARNER, D.; BOND-LAMBERTY, B.; JIAN, J.; STELL, E.; VARGAS, R. Spatial Predictions and Associated Uncertainty of Annual Soil Respiration at the Global Scale. **Global Biogeochemical Cycles**, v. 33, n., p., 2019.

WARNER, D.; VARGAS, R.; SEYFFERTH, A.; INAMDAR, S. Transitional slopes act as hotspots of both soil CO₂ emission and CH₄ uptake in a temperate forest landscape. **Biogeochemistry**, v. 138, n., p., 2018.

WMO; FREDERIC CHEVALLIER; ANDREW CROTWELL; CHRISTOPH GERBIG; ARMIN JORDAN; XIN LAN; ZOË LOH; INGRID LUIJKX; JOHN MILLER; YOUSUKE SAWA; OKSANA TARASOVA; ALEX VERMEULEN; RAY WEISS; THORSTEN WARNEKE; YVER., C., 2023. WMO Greenhouse Gas Bulletin (GHG Bulletin) - No. 19: The State of Greenhouse Gases in the Atmosphere Based on Global Observations through 2022, World Meteorological Organization (WMO)

YILMAZ, I.; KAYNAR, O. Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils. . **Expert Systems with Applications**, v. 38, n., p. 5958-5966, 2011.