

UNIVERSIDADE ESTADUAL PAULISTA (UNESP)
"JÚLIO DE MESQUITA FILHO"

Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal
Departamento de Engenharia e Ciências Exatas
(FCAV - DECE)

Projeto de Pesquisa
Iniciação Científica

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

Candidata (Bolsista): Letícia Roberta de Lima
Responsável: Prof. Dr. Alan Rodrigo Panosso

Jaboticabal – SP – Brasil
Julho de 2021

SUMÁRIO

Resumo:	3
1 INTRODUÇÃO E JUSTIFICATIVA	3
1.1 Emissão de gases do efeito estufa e a emissão de CO ₂ no Brasil.....	3
1.2 Inteligência artificial e aprendizado de máquina.....	5
2 HIPÓTESE E OBJETIVOS	7
3 MATERIAL E MÉTODOS.....	7
3.1 Conjunto de dados.....	7
3.2 Determinação da emissão de CO ₂ , captura de O ₂ , temperatura e umidade do solo	8
3.3 Concentração de CO ₂ atmosférico.....	8
3.4 Determinação dos atributos físicos e químicos do solo.....	9
3.5 Covariáveis ambientais.....	9
3.6 Forma de análise dos resultados.....	10
3.6.1 Pré-processamento.....	11
3.6.2 Aprendizado de máquina.....	12
3.6.3 Medidas estatísticas de avaliação de desempenho dos diferentes modelos	13
4 CRONOGRAMA DE ATIVIDADES.....	14
5 RESULTADOS PRELIMINARES.....	15
6 REFERÊNCIAS BIBLIOGRÁFICAS	18

ABORDAGEM DE APRENDIZADO DE MÁQUINA SOBRE A RELAÇÃO ENTRE A EMISSÃO DE CO₂ NO SOLO E O CO₂ ATMOSFÉRICO EM ÁREAS AGRÍCOLAS NO BRASIL CENTRAL

Resumo:

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 terrestre 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 a diminuição das 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, o objetivo da proposta será 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 em áreas agrícolas na região do estado do Mato Grosso do Sul (MS), a partir de uma série temporal de 2015 a 2017. As técnicas utilizadas serão: máquina de vetores de suporte (SVM) e árvores de decisão (*Random Forest* e *Gradient Boosting Machines*). De maneira geral 70-80% das observações serão utilizadas para aprendizagem (processo de treinamento) dos modelos e 30-20% para validação. Novos experimentos estão sendo realizados em campo para validação dos resultados em áreas agrícolas. A acurácia dos modelos será determinada por meio da correlação de Pearson (r), coeficiente de determinação (R²), erro quadrático médio (RMSE), erro médio (ME), índice de concordância (d), coeficiente de confiança (c) e menor erro percentual absoluto médio (MAPE). Espera-se que essa abordagem contribua para melhorar o entendimento da dinâmica do FCO₂ e CO₂ atmosférico em diferentes regiões, usos e manejos dos solos no Brasil central, produzindo cenários com menores incertezas que possam servir como sustentáculo nas tomadas de decisões focadas na mitigação das emissões de CO₂ em áreas agrícolas.

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

1 INTRODUÇÃO E JUSTIFICATIVA

1.1 Emissão de gases do efeito estufa e a emissão de CO₂ no Brasil

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 2020, aponta que as concentrações atmosféricas de dióxido de carbono (CO₂), metano (CH₄) e óxido nitroso (N₂O), atingiram novas

máximas no ano de 2019, com CO₂ a 410,5 ppm, CH₄ a 1877 ppb e N₂O a 332 ppb. Esses valores representam, respectivamente, 148%, 260% e 123% dos níveis atmosféricos no período pré-industrial (WMO, 2020). Ainda de acordo com o referido boletim, o desmatamento e demais mudanças no uso da terra registraram média de emissão de 5,5 Gt de CO₂ ano⁻¹ no período de 2009-2018, em comparação, as emissões devido à queima de combustíveis fósseis e produção de cimento foram projetadas em 36,7 Gt de CO₂ em 2019.

A emissão de GEE pelo Brasil é devido, principalmente aos setores *Agropecuária e Mudança do Uso da Terra e Florestas*. No ano de 2016 esses setores emitiram juntos cerca de 730,1 Mt de CO₂eq, ou seja, 56% das emissões totais de GEE do país (MINISTRO DA CIÊNCIA TECNOLOGIA INOVAÇÕES E COMUNICAÇÕES, 2019). Ainda em 2016, as emissões líquidas decorrentes do subsetor conversões de uso da terra foram de 254,3 Mt de CO₂, representando um aumento 27,1% em comparação a 2015, atribuído ao crescente desmatamento nos biomas Amazônia e Mata Atlântica (MINISTRO DA CIÊNCIA TECNOLOGIA INOVAÇÕES E COMUNICAÇÕES, 2019). Em adição, as últimas projeções do Ministério da Agricultura, Pecuária e Abastecimento indicam um crescimento do agronegócio brasileiro para os próximos 10 anos que será associado, consequentemente, à novas altas nas taxas de emissões de GEE (MINISTÉRIO DA AGRICULTURA PECUÁRIA E ABASTECIMENTO, 2018). 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). Nesse contexto, muitas práticas agrícolas têm potencial de mitigar as emissões GEE, das quais as mais proeminentes são: o melhor manejo dos solos agrícolas e pastagens, a restauração de solos degradados e orgânicos, a mudança no uso da terra e agroflorestas (WARNER et al., 2019).

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. 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 os resultados 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 (LAGANIERE; ANGERS; PARE, 2010; FARHATE et al., 2018). 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; SCHLESINGER, 1992; THANGAVEL et al., 2018).

O Índice Anual de Gases de Efeito Estufa (AGGI) mostra que de 1990 a 2019, o forçamento radiativo por gases de efeito estufa de longa duração (LLGHGs) aumentou 45%, com o CO₂ sendo responsável por cerca de 80% desse aumento (WMO, 2020). De acordo com os dados da oitava edição do Sistema de Estimativas de

Emissões e Remoções de Gases de Efeito Estufa (SEEG), para as emissões globais, o Brasil ocupa o 6º lugar entre os maiores emissores de gases de efeito estufa, com 3,2% do total mundial. Somando-se as emissões de CO₂ da agropecuária com a parcela das emissões dos demais setores relacionados ao setor agro, o SEEG conclui que a atividade rural – seja direta ou indiretamente – respondeu por 72% das emissões do Brasil no ano de 2019 (SEEG, 2020). Ampliar o fomento e a adoção de tecnologias e métodos de produção com baixas emissões continua fundamental para garantir a redução das emissões absolutas do setor, que continuam subindo (SEEG, 2020). E para tal, se faz necessário uso de técnicas de análises estáticas de dados para amparar tomadas de decisões para a implementação de políticas públicas e privadas, visando a diminuição dos efeitos das mudanças causadas pelos GEE.

Os fatores ambientais e atributos do solo que controlam a magnitude de FCO₂ 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 no equilíbrio 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 FCO₂, 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 CO₂ do solo é controlada por atributos como a umidade do solo (LEON et al., 2014; SANTOS et al., 2019a), porosidade livre de água (WARNER et al., 2018), oxigenação do solo (CHEN; 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).

Assim, a caracterização de como os diferentes usos e manejos dos solos brasileiros afetam a dinâmica FCO₂ 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.

1.2 Inteligência artificial e aprendizado de máquina

Inteligência artificial (IA) é a capacidade de um algoritmo assimilar informações para realizar tarefas características da inteligência humana, como por exemplo: reconhecer objetos e sons, contextualizar a linguagem, aprender com o ambiente e resolver problemas. Pode ser entendida como o potencial da máquina de tomar a melhor decisão possível dada a quantidade de informação disponível, com habilidade de adaptação a diferentes situações (KUHN; JOHNSON, 2013; CHOLLET; ALLAIRE, 2017). Atualmente a IA tem sido utilizada em vários campos da ciência, principalmente nas áreas da saúde, da indústria e ambiental (GE et al., 2017; CHIAVEGATTO FILHO et al., 2018; DI MININ et al., 2018; SANTOS et al., 2019b). Como estratégia para o aprendizado, os dados a

respeito de um problema são fornecidos à máquina, bem como as respostas esperadas, e o aprendizado consiste na geração das regras de decisão, que é a saída do sistema computacional, ou seja, a máquina aprende por conta própria as regras para a tomada de decisão (BRUCE; BRUCE, 2017).

Portanto, o aprendizado de máquina se concentra no desenvolvimento de algoritmos eficientes que escalonam grandes quantidades de dados visando otimizar modelos preditivos para, a partir da compreensão da estrutura dos dados, compor as regras de predição. Em resumo, aprendizado de máquina é um campo de pesquisa estatística para o treinamento de algoritmos computacionais que dividem, classificam e transformam uma base de dados para maximizar a capacidade de classificar, prever, agrupar ou descobrir padrões em um conjunto de dados de destino (REICHSTEIN 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; KAYNAR, 2011; BESALATPOUR et al., 2013; REICHSTEIN et al., 2019; SHADRIN et al., 2020).

Um grande número de estudos tem sido conduzido para analisar e compreender a relação entre a emissão de GEE e o aumento da temperatura global usando abordagens estatísticas convencionais. Entretanto, essas técnicas seguem suposições de modelagem probabilística, em que os resultados podem ser associados a incertezas de várias magnitudes (KHAN; KHAN, 2019). As técnicas de aprendizado de máquina oferecem grande potencial para refinar as estimativas a respeito ciclo de carbono em escala global (TRAMONTANA et al., 2016; REICHSTEIN et al., 2019). Essas técnicas permitem a superação dos desafios estatísticos impostos por um número cada vez maior de covariáveis espaciais disponíveis, muitos das quais podem estar correlacionados entre si, ou ter relações não-lineares, incluso a dependência espacial, violando a importante suposição de dados distribuídos de forma idêntica e independente impostos pela abordagem probabilística (REICHSTEIN et al., 2019; WARNER et al., 2019). Utilizando RNAs para prever FCO₂ em áreas de cana-de-açúcar, Farhate et al. (2018), observaram que a arquitetura de rede Multilayer Perceptron (MLP) com o método de seleção de atributos Wrapper apresentou alta precisão para classificação das taxas de perda de carbono no solo via FCO₂. Similarmente, Freitas et al. (2018), obtiveram resultados satisfatórios ao estimarem a variabilidade espaço-temporal de FCO₂ em áreas de cultivo de cana-de-açúcar por meio de rede *Multilayer Perceptron* (MLP) e pelo algoritmo de retropropagação do erro, com ótimo desempenho preditivo, indicado pelos valores preditos próximos aos valores medidos experimentalmente, conforme o erro percentual absoluto médio (MAPE) (18,29%) e o coeficiente de determinação obtidos ($R^2 = 0,92$).

2 HIPÓTESE E OBJETIVOS

A hipótese da presente proposta é que a dinâmica temporal da emissão de CO₂ do solo e do CO₂ atmosférico pode ser descrita como um fenômeno com estruturas multidimensionais que evoluem no tempo, aptas a seres modeladas por algoritmos de aprendizado de máquina. Assim, o objetivo será avaliar o desempenho preditivo dos principais algoritmos de aprendizado de máquina para FCO₂ e sua relação com o CO₂ atmosférico em áreas agrícolas na região do estado do Mato Grosso do Sul. As técnicas utilizadas serão: máquina de vetores de suporte (SVM) e árvores de decisão (*Random Forest* e *Gradient Boosting Machines*). Especificamente, pretende-se usar a abordagem de aprendizado de máquina para investigar os padrões temporais de FCO₂ e CO₂ atmosféricos no Brasil central avaliando as fontes de incerteza e de previsão em áreas agrícolas. Espera-se que a proposta, além de impactar na formação do aluno bolsista, traga avanços para a geração de produtos e processos com potencial de auxílio na tomada de decisão sobre mitigação das emissões de CO₂.

3 MATERIAL E MÉTODOS

3.1 Conjunto de dados

A presente proposta irá compilar achados anteriores e alguns resultados ainda não publicados sobre a emissão de CO₂ do solo (FCO₂) a partir de 4 experimentos realizados durante 3 anos (Tabela 1) no Mato Grosso do Sul. Os experimentos selecionados representam usos da terra típicos da região específica onde os processos envolvidos na mudança do uso ocorreram e, em sua maioria, em áreas próximas dentro da localidade. Nesses estudos a variabilidade de FCO₂ foi determinada ao longo do tempo e do espaço, bem como de demais variáveis como temperatura, umidade do solo e atributos físicos e químicos do solo.

Tabela1. Informações geográficas, clima e histórico das áreas onde foram conduzidos os experimentos de variabilidade espaço-temporal da emissão de CO₂ e atributos físicos e químicos do solo ao longo de 3 anos no Brasil central.

Cidade	Estado	Coordenadas Geográficas	Elevação (m) ¹	Clima ²	Cultura	Manejo	TC	Ano	Solo	N
Selvíria	MS	20°20' S; 51°24' O	362	Aw	Eucalipto	Reflorestamento	29	2015-2017	Latossolo Vermelho distroférico	102
Selvíria	MS	20°20' S; 51°24' O	362	Aw	Pinus	Reflorestamento	29	2015-2017	Latossolo Vermelho distroférico	15
Selvíria	MS	20°20' S; 51°24' O	362	Aw	Silvipastoril	ILPF	29	2015-2017	Latossolo Vermelho distroférico	15
Selvíria	MS	20°20' S; 51°24' O	362	Aw	Nativas Cerrado	Reflorestamento	29	2015-2016	Latossolo Vermelho distroférico	15

TC = Tempo de conversão em anos; N = número de pontos amostrais; Ano = ano de condução do experimento.

¹ Elevação em metros acima do nível do mar.

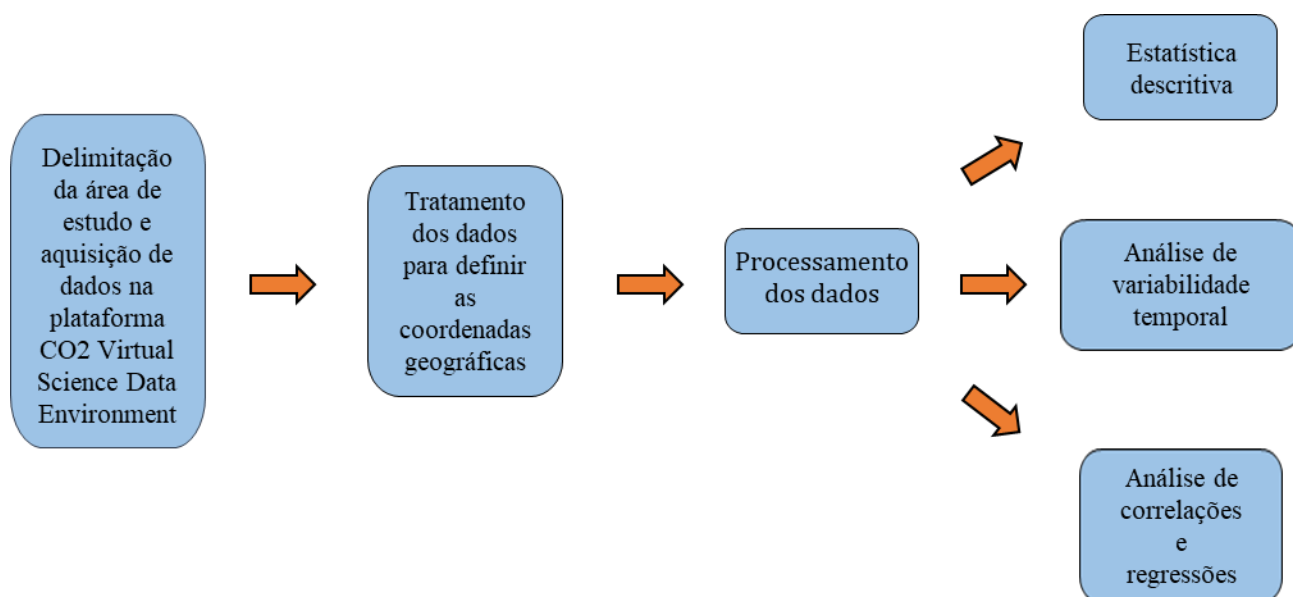
² Classificação climática de acordo com Köppen.

3.2 Determinação da emissão de CO₂, captura de O₂, temperatura e umidade do solo

Em todos os ensaios a emissão de CO₂ do solo foi registrada por meio do sistema LI-COR (LI-8100), que em seu modo de medição o sistema monitora as mudanças na concentração de CO₂ dentro da câmara para solos por meio de espectroscopia na região do infravermelho. Essa câmara é acoplada sobre colares de policloreto de vinila (PVC) previamente inseridos no solo em cada ponto amostral na profundidade de 3 cm. O fluxo de CO₂ foi computado em cada ponto por um ajuste da concentração de CO₂ do ar dentro da câmara em função de uma regressão exponencial no tempo após o fechamento da mesma. A temperatura do solo foi monitorada, concomitantemente às avaliações de respiração do solo, utilizando-se um sensor de temperatura que é parte integrante do sistema ao LI-8100. Tal sensor consiste de uma haste de 0,2 m que foi inserida no interior do solo próximos ao local onde foram previamente instalados os colares de PVC para a avaliação da emissão de CO₂. Em alguns ensaios, a temperatura do solo foi medida com auxílio de um termômetro digital do tipo espeto. A umidade do solo foi determinada por meio de um equipamento de TDR (Time Domain Reflectometry - Hydrosense TM, Campbell Scientific, Austrália).

3.3 Concentração de CO₂ atmosférico

Afim de uma determinação acurada da concentração de CO₂ atmosférico, o presente estudo visa coligir dados de sensores orbitais com os dados coletados em campo, empregando um processo, descrito no Fluxograma 1, que estabelecerá um padrão entre os dados dos sensores e os dados coletados em campo a partir das coordenadas obtidas na plataforma CO₂ Virtual Science Data Environment.



Fluxograma 1. Fluxograma de aquisição, processamento e análise dos dados.

3.4 Determinação dos atributos físicos e químicos do solo

Os dados referentes aos atributos físico e químicos do solo foram determinados na profundidade de 0 – 0,10 e 0,10-0,20 m. Foram realizadas as seguintes análises de rotina: pH, determinação do teor de matéria orgânica do solo (MO), fósforo disponível (P), potássio (K), cálcio (Ca), magnésio (Mg), soma de bases (SB), acidez potencial (H+Al). Os teores de cálcio, magnésio, potássio trocáveis e fósforo disponíveis serão extraídos utilizando-se o método da resina trocadora de íons (RAIJ, 2001); a capacidade de troca de cátions (CTC) e a saturação por base (V%) também serão calculados. O carbono orgânico será pelo método da combustão úmida, via colorimétrica (RAIJ et al., 1987), a determinação do teor de nitrogênio será realizada por meio de digestão sulfúrica (MALAVOLTA; VITTI; 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 (Ma) e microporosidade (Mi) 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 Ma. A Mi foi determinada após a retirada dos anéis da estufa à 105 °C num período de 24 h e posterior pesagem. Já a porosidade total foi calculada pela soma dos macroporos e microporos. Também foi determinado 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 (Ds) que ocorrem após as mudanças do uso da terra. Para isso, será utilizada a metodologia descrita por Ellert; 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).

3.5 Covariáveis ambientais

As seguintes covariáveis ambientais serão selecionadas: precipitação; velocidade do vento (ms⁻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}) (°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. As variáveis serão adquiridas a partir de estações meteorológicas próximas às regiões de ensaio, e a partir de banco de dados disponibilizados, oriundos de sensoriamento remoto em alguns casos, e incorporadas à base de dados.

3.6 Forma de análise dos resultados

Todos os modelos de aprendizado de máquina e etapas de pré-processamento e validação dos modelos serão realizado a partir de implmentações na linguagem R (R DEVELOPMENT CORE TEAM, 2021), cuja maior vantagem é ser livre, de código aberto, amplamente utilizada e de alta longevidade, o que garantirá a inspeção e reprodutibilidade dos resultados, facilitando a divulgação para a comunidade científica. As interface de aplicação das técnicas serão utilizados os seguintes pacotes (coleções de funções implementadas dentro da linguagem R): *Caret* e *Keras*, para aprendizado de máquina e aprendizado profundo respectivamente. Esses pacotes são conjunto de funções implementadas para simplificação de todo o processo de criação e avaliação dos modelos preditivos além de ser recomendado para a visualição dos resultados. A Figura 5 apresenta as etapas do processo de data mining, o qual será utilizado para o processo de aprendiadio de máquina. O protocolo adotado será o CRISP-DM (*Cross Industry Standard Process fo Data Mining*) o qual adaptado de Chapman et al. (2000) e constituirá das seguintes etapas:

- a) Compreensão: definição dos objetivos e necessidades iniciais do projeto que se pretende atender com a modelagem e mineração de dados, compreende as etapa iniciais do projeto;
- b) Entendimento dos dados: dados serão de diversas fontes e localidades, alguns provenientes de sensores estarão em diferentes formatos, assim, dever-se-á compreender e entender as estruturas dos diferentes tipos de dados junto às suas especificidades com o principal objetivo de identificação de problemas e atestar a qualidade dos dados;
- c) Preparação dos dados: incorporação dos dados no banco de dados final (https://arpanosso.github.io/FCO2_machinelearning/index.html). Para isso, a estrutura do banco de dados poderá ser modificada, e adaptada, possíveis eliminações de covariáveis, inclusão de novas variáveis, mudanças de escala, eliminação de inconsistências e possivelmente de dados faltantes;
- d) Modelagem: Aprendizado de máquina propriamente dita. Atendimento das exigências específicas de cada técnica de aprendizado;
- e) Avaliação: após o desenvolvimento do modelo, serão revisados os processos e verificado se os objetivos foram atingidos para ao final da etapa ser tomada a decisão de utilização ou não dos resultados. Para avaliação dos modelos, será primeiramente necessário a separação do banco de dados em duas partes, sendo uma para o processo de aprendizado e outra para o teste dos modelos gerados a qual será feita pelo processo de validação cruzada.
- f) Implantação: Proposição de aplicação e distribuição do conhecimento gerado durante o processo será realizado por meio de *web app* desenvolvido em R (pacote *Shiny*). Os resultados também

serão organizados para divulgação de forma que este possa ser utilizado pela comunidade (https://arpanosso.github.io/FCO2_machinelearning/index.html).

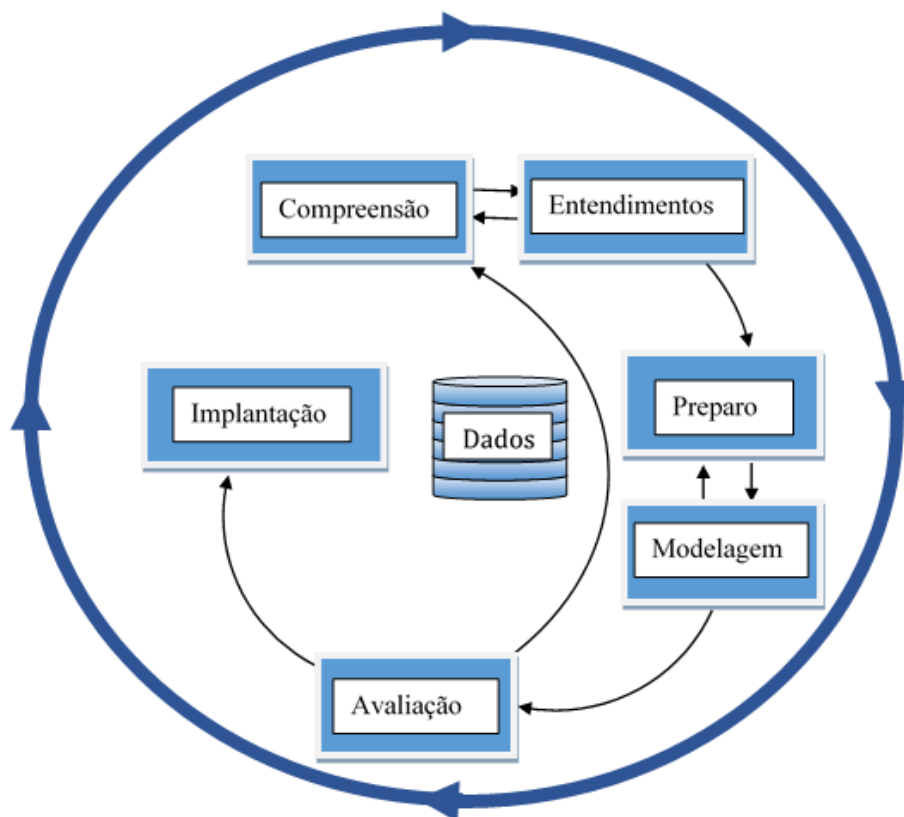


Figura 1. Etapas do modelo de referência adaptado CRISP-DM, onde as sequencias das fases não serão rígidas uma vez que o movimento entre etapas é requerido (CHAPMAN et al., 2000).

3.6.1 Pré-processamento

Inicialmente, cada covariável será avaliada por meio da análise de estatística descritiva dos dados, bem como a avaliação/remoção de outliers multivariados por meio da distância de Mahalanobis ao nível de 1% de significância. Para garantir a qualidade dos dados entendida nessa abordagem como a adequação por meio de transformação recomendada, eliminação e padronização de valores, será realizada a análise de componentes principais (ACP). ACP é uma abordagem de aprendizado não supervisionado, e também será realizada com o objetivo de reduzir um grande número de covariáveis originais para um pequeno número de variáveis transformadas, e não correlacionadas onde serão considerados os componentes principais que apresentarem autovalores superiores à unidade (Kaiser, 1958). Deve-se ressaltar que ACP terá o objetivo de maximizar a representação da variabilidade dos dados, onde as variáveis alvo (FCO₂ e CO₂ atmosférico) não serão utilizadas no processo. Antes da análise, os dados serão padronizados para média igual a zero e variância unitária, para

garantir que as covariáveis estejam em mesma escala, evitando uma maior representatividade de uma característica apenas devido à ordem de grandeza de sua unidade de medida.

3.6.2 Aprendizado de máquina

Máquina de vetores de suporte

A máquina de vetores de suporte (SVR) é um algoritmo de aprendizado de máquina, supervisionado para problemas de classificação e regressão. É conhecido popularmente como um método não linear. Engloba funções ou classificadores que implementam modificações espaciais nos dados, levando-os a um plano onde a classificação se torna mais fácil. Para certo dado de treinamento representado por $(x_1, y_1); \dots; (x_N, y_N)$, onde x é um vetor de entrada e y é um vetor de saída correspondente, a função de regressão do SVM (máquina de vetores de suporte) pode ser expressa como:

$$f(x) = \omega \cdot \phi(x) + b \quad (1)$$

onde ω é um vetor de peso, b é um viés e ϕ apresenta uma função de transferência não linear.

O objetivo de ϕ é mapear de maneira não linear os vetores de entrada para um espaço de característica de alta dimensão. A precisão do modelo SVR com Função Base Radial (RBF) depende das seleções dos parâmetros C (custo), e regularização positivo que corresponde aos parâmetros do peso no problema de otimização para minimizar o erro empírico; a variável de penalização do treinamento ϵ (épsilon) e do parâmetro do kernel γ (gamma). Serão inicialmente selecionadas a regressão SVM tipo 1 e função de Base Radial (RBF) de Kernel, escolhida para os dois modelos. Os parâmetros (C ; γ e ϵ) serão ajustados para os modelos. Assim, um grande número de combinações desses parâmetros será testado. O algoritmo de validação cruzada terá o valor C constante da capacidade estimado como 1. O algoritmo de validação cruzada será utilizado para obter melhores estimativas de treinamento de parâmetros do modelo SVR escolhido. O número máximo de iterações escolhida inicialmente será de 10 e erro de 0,001.

Random Forest

São modelos estruturados em árvore 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. Na construção de uma árvore de regressão o modelo baseia-se na divisão das variáveis explicativas, conforme suas semelhanças e dissimilaridades e o seu resultado final é formado pela média dos resultados de todas as árvores. É uma técnica estatística não paramétrica de aprendizagem. 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*). Para a primeira abordagem, serão adotados o número de estimadores das árvores na floresta foi igual a 300 e a profundidade máxima da árvore será igual a 10. O número de covariáveis usadas em cada árvore será um terço do número total de amostras.

Gradient Boosting Machines

Baseia-se no conceito de obtenção de algoritmos de classificação/regressão fortes, a partir de transformação em algoritmos de baixo desempenho. Assim como o algoritmo de *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. Esse método funciona aplicando-se a partir da aplicação sequencial dos classificadores a versões reponderadas do conjunto de treinamento, dando maior peso aos registros/observações classificados erroneamente no passo anterior e menos peso aos classificados corretamente. Assim, gera-se uma sequência de árvores que usam a informação contida no treinamento de suas antecessoras, o seja, o algoritmo vai "aprendendo com o erro". Os hiperparâmetros ajustados para serão o número de árvores usadas para construir o modelo, redução, interação

3.6.3 Medidas estatísticas de avaliação de desempenho dos diferentes modelos

Os diferentes modelos serão comparados pelos seguintes 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).

$$r = \frac{\sum \sigma_{ht_i} \times \sigma_{hp_i} - \bar{\sigma}_{ht} \times \bar{\sigma}_{hp}}{(n-1) S_{\sigma_{ht}} S_{\sigma_{hp}}} \quad (1)$$

$$R^2 = 1 - \frac{\sum_i (\sigma_{hp_i} - \sigma_{ht_i})^2}{\sum_i (\sigma_{hp_i} - \bar{\sigma}_{hp_i})^2} \quad (2)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (\sigma_{ht_i} - \sigma_{hp_i})^2 \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\sigma_{ht_i} - \sigma_{hp_i})^2} \quad (4)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\sigma_{ht_i} - \sigma_{hp_i}| \quad (5)$$

$$MAPE = \left[\frac{1}{n} \sum_{i=1}^n \left| \frac{\sigma_{ht_i} - \sigma_{hp_i}}{\sigma_{ht_i}} \right| \right] \times 100 \quad (6)$$

Observe que σ_{ht} , σ_{hp} são, respectivamente valor observado e estimado, n número de observações, $\bar{\sigma}_{ht}$, $\bar{\sigma}_{hp}$ valores médios previsto.

4 CRONOGRAMA DE ATIVIDADES

O projeto será desenvolvido junto ao Departamento de Engenharia e Ciências Exatas da UNESP de Jaboticabal com auxílio do programa de pós-graduação em Agronomia (Ciência do Solo) da Unesp de Jaboticabal. A bolsista será responsável por várias etapas do projeto, sobre supervisão do pesquisador responsável e dos pesquisadores associados, como: preparo do banco de dados; mineração de dados; aplicação dos algoritmos de aprendizagem; revisão de literatura, confecção de relatórios científicos e apresentação de resultados em congressos e simpósios. Será realizado a busca, atualização e padronização dos dados de variáveis ambientais. Finalmente, ficarão sobre nossa supervisão as etapas de redação de trabalhos acadêmicos e científicos além da construção da base de dados e disponibilização pública de web app para previsão de FCO₂ em áreas agrícolas no Brasil. Devemos ressaltar que durante o período de pandemia e consequente restrição da universidade, todas as reuniões e supervisões serão realizadas via GoogleMeet, e o projeto será compartilhado via sistema em nuvem para controle de versão Git (<https://arpanosso.github.io/oco2/index.html>), que permitirá o compartilhamento do projeto entre os discentes envolvidos.

Tabela 1. Cronograma de atividades previstas para o período de vigência da proposta atual, 2021 a 2022, período total de 1 ano (4 trimestres).

ATIVIDADES	PERÍODO: (ANO) / - TRIMESTRE -			
	2021		2022	
	- 3º -	- 4º -	- 1º -	- 2º -
Revisão de literatura				X
Avaliação da emissão de CO ₂ , captura de O ₂ temperatura e umidade do solo	X	X	X	
Preparo do Banco de dados para mineração e incorporação de variáveis ambientais	X	X		
Aplicação dos algoritmos de aprendizado de máquina			X	X
Confecção de dissertações, teses; Publicação de trabalho científico e programa computacional para previsão de FCO ₂ .				X
Divulgação de Resultados e preparo de relatório científico		X		X
Atualização de disponibilização da base de dados			X	X

5 RESULTADOS PRELIMINARES

Nessa seção serão apresentados alguns resultados promissores a respeito da utilização das técnicas de aprendizado de máquina para dados de emissão de CO₂ do solo (FCO₂) e a captura de oxigênio pelo solo (FO₂) em áreas de floresta plantada no Cerrado sul-mato-grossense, município de Selvíria em Mato Grosso do Sul (MS) (Figura 2). As áreas foram: Pinus (**pi**), Eucalipto (**eu**) e área reflorestada com espécies nativas (**ns**), todas com histórico de 30 anos de conversão do solo de culturas anuais (soja, milho e algodão). As técnicas utilizadas foram redes neurais artificiais (RNA), máquina de vetores de suporte (SVM) e árvore de decisão (utilizando algoritmo *random forest* - RP). Na etapa de modelagem foram utilizados fatores ambientais e do solo, totalizando 23 covariáveis. A avaliação do fluxo de gases foi realizada entre os anos de 2015 e 2019 (proposta atual em andamento 301606/2017-0).

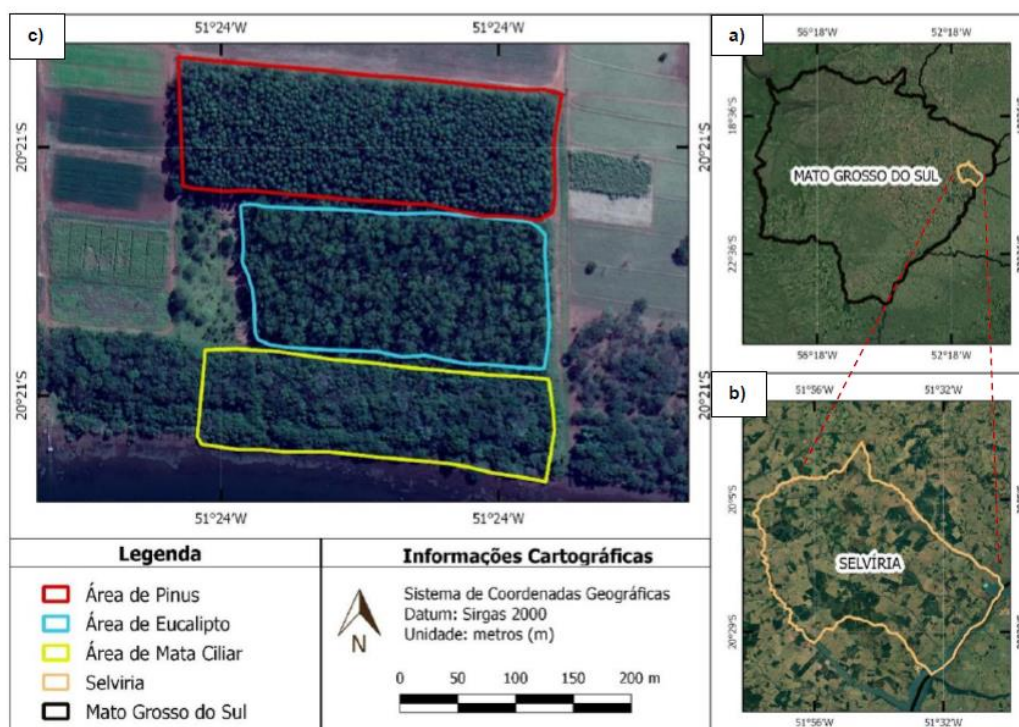


Figura 2. Mapa das áreas de estudo localizadas no município de Selvíria, MS. a) Estado do Mato Grosso do Sul, Brasil Central; b) Município de Selvíria; c) Áreas de florestas plantadas: pinus, eucalipto e mata ciliar, respectivamente.

A variabilidade temporal dos fluxos de gases em áreas de floresta plantada foi governada pelo índice de vegetação melhorado (EVI), temperatura do solo (Ts), teor de matéria orgânica do solo (MO), radiação solar (Rs) e macroporosidade (Macro). Ts e MO foram as mais importantes para modelagem de FCO₂, enquanto Ts e Macro governaram FO₂. Foram testados vários modelos e, de maneira geral, as covariáveis que sempre apresentaram importância relativa no processo de aprendizagem foram: o uso da área, a umidade do solo, temperatura do ar, a

radiação solar e pH do solo. Curiosamente, o aprendizado por árvore de decisão (RF) foi aquele que apresentou o melhor desempenho preditivo para FCO₂ (Figura 3) e o pior para FO₂, o qual foi melhor modelado por redes neurais artificiais. Os resultados preliminares indicam que o desempenho dos modelos foi sensível à seleção das covariáveis de entrada evidenciando a estrutura multidimensional da dinâmica desses gases no solo, indicando a necessidade da incorporação da abordagem espacial do fenômeno.

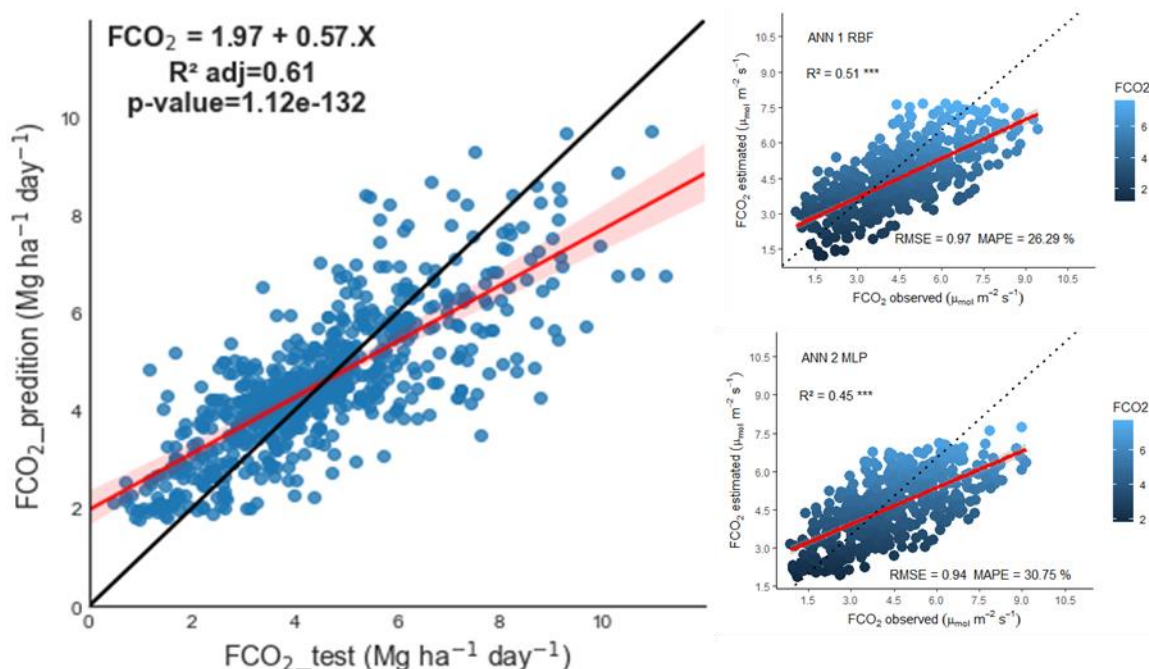


Figura 3. Desempenho do algoritmo Random Forest na previsão da variabilidade temporal de FCO₂ em áreas de floresta plantada (a) e avaliação do desempenho de FCO₂ com duas arquiteturas de RNA, rede neural artificial com função de base radial (b) e de neural artificial de multicamadas-perceptron (c).

Resultados promissores, foram observados a respeito da variabilidade espacial de FCO₂ e áreas de Eucalipto durante o período avaliado, ano de 2015. A Figura 7 apresenta a importância relativa das covariáveis atributos do solo para a previsão de FCO₂ em locais não amostrados de um campo de aproximadamente 1 ha. O aprendizado foi realizado a partir do algoritmo *Gradient Boosting Machines*, para comparação com o algoritmo de krigagem ordinária, considerando um modelo isotrópico de dependência espacial. O aprendizado de máquina foi altamente eficiente para geração de mapas de FCO₂ a partir dos mapas da covariáveis. O algoritmo apresentou baixos valores de RMSE e altos valores de R². Os valores RMSE foram 0,12; 0,10; 0,086 para os dias 8 de outubro, 29 de outubro e 18 de novembro do ano de 2015, respectivamente, sendo que para todos os dias foram observados valores de precisão superiores a 98%. Na dinâmica espaço temporal as covariáveis grau de humificação da matéria orgânica do solo (H_{LIFS}), pH, teor de matéria orgânica do solo (SOM) e teor de acidez potencial do solo (H + Al) foram aquelas que explicaram a maior parte da variação do FCO₂ no período. Com exceção de 29 de outubro,

conforme exemplificado pelo gráfico, a umidade do solo (SWC) foi a variável que regeu FCO₂. No período esta data marcou a transição da estação seca para o início da estação chuvosa, assim os resultados indicam quo potencial da técnica para geração de mapas de alta qualidade na modelagem sazonal de FCO₂.

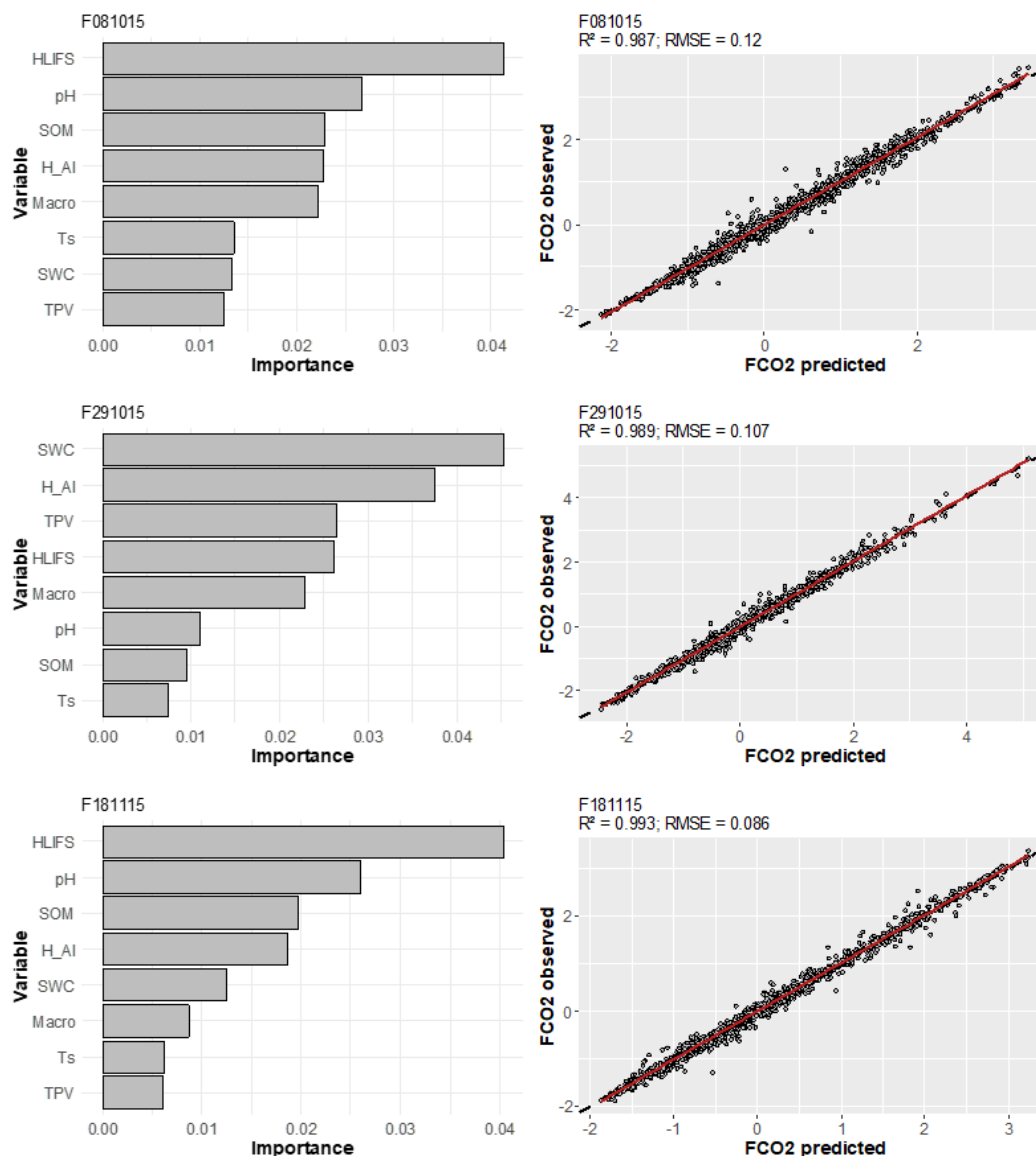


Figura 4. Importância dos atributos físicos e químicos do solo gerados pelo algoritmo *Gradient Boosting Machines*: temperatura do óleo (Ts); teor de água do solo (SWC); matéria orgânica do solo (SOM); potencial de hidrogênio (pH); acidez potencial (H + Al); índice de humificação da matéria orgânica do solo (H_{LIFS}); microporosidade (Micro); volume total de poros (TPV); coeficiente de determinação (R^2) e erro quadrático médio (RMSE).

6 REFERÊNCIAS BIBLIOGRÁFICAS

- 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. S.; PANOSSO, A. R.; TEIXEIRA, D. D. B.; MIRANDA, J. G. V.; PEREIRA, G. T.; LA SCALA, 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.
- BRUCE, P.; BRUCE, A. **Practical Statistics for Data Scientists**. United States of America: O'Reilly Media, Inc., 2017. 562 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.
- CHAPMAN, P.; CLINTON, J.; KERBER, R.; KHABAZA, T.; REINARTZ, T.; SHEARER, C. R.; WIRTH, R., 2000. CRISP-DM 1.0: Step-by-step data mining guide.
- CHEN, G. Q.; 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. 11, p. 3647-3732, 2010.
- CHIAVEGATTO FILHO, A. D. P.; DOS SANTOS, H. G.; DO NASCIMENTO, C. F.; MASSA, K.; KAWACHI, I. Overachieving Municipalities in Public Health: A Machine-learning Approach. **Epidemiology**, v. 29, n. 6, p. 836-840, 2018.
- CHOLLET, F.; ALLAIRE, J. J. **Deep Learning with R**. 2017. 341 p.
- DI MININ, E.; FINK, C.; TENKANEN, H.; HIIPALA, T. Machine learning for tracking illegal wildlife trade on social media. **Nature Ecology & Evolution**, v. 2, n. 3, p. 406-407, 2018.
- 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. 4, p. 529-538, 1995.
- EMBRAPA. **Manual de métodos de análise de solo**. 2 ed. Brasília: Ministério da Agricultura e do Abastecimento / EMBRAPA-CNPq, 1997. 212 p.
- FARHATE, C. V. V.; DE SOUZA, Z. M.; OLIVEIRA, S. R. D.; 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., 2018.
- GE, Z.; SONG, Z.; DING, S. X.; HUANG, B. Data Mining and Analytics in the Process Industry: The Role of Machine Learning. **IEEE Access**, v. 5, n., p. 20590-20616, 2017.
- GRAF, A.; HERBST, M.; WEIHERMULLER, L.; HUISMAN, J. A.; PROLINGHEUER, N.; BORNEMANN, L.; VEREECKEN, H. Analyzing spatiotemporal variability of heterotrophic soil respiration at the field scale using orthogonal functions. **Geoderma**, v. 181, n., p. 91-101, 2012.
- KHAN, M. Z.; KHAN, M. F. Application of ANFIS, ANN and fuzzy time series models to CO₂ emission from the energy sector and global temperature increase. **International Journal of Climate Change Strategies and Management**, v. 11, n. 5, p. 622-642, 2019.
- KUHN, M.; JOHNSON, K., 2013. Applied predictive modeling.
- LAGANIERE, J.; ANGERS, D. A.; PARE, 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. R.; 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. 0, p. 12-21, 2014.
- MALAVOLTA, E.; VITTI, G. C.; OLIVEIRA, A. S. **Avaliação do estado nutricional das plantas: princípios e aplicações**. Piracicaba: Potafós, 1997. p.
- 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. 3-4, p. 293-327, 2000.

- MINISTÉRIO DA AGRICULTURA PECUÁRIA E ABASTECIMENTO. **Brasil Projeções do Agronegócio 2017/2018 a 2027/2028**. Brasília: Ministério da Agricultura, Pecuária e Abastecimento, 2018. p.
- MINISTRO DA CIÊNCIA TECNOLOGIA INOVAÇÕES E COMUNICAÇÕES. **Estimativas anuais de emissões de gases do efeito estufa no Brasil**. Brasília: Ministro da Ciência, Tecnologia, Inovações e Comunicações, 2019. p.
- 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. 1, p. 63-81, 1996.
- R DEVELOPMENT CORE TEAM. **A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- RAICH, J. W.; SCHLESINGER, W. H. The global carbon-dioxide flux in soil respiration and its relationship to vegetation and climate. **Tellus Series B-Chemical and Physical Meteorology**, v. 44, n. 2, p. 81-99, 1992.
- RAIJ, B. V. **Análise química para avaliação da fertilidade de solos tropicais**. Campinas: Instituto Agrônômico, 2001. 285 p.
- RAIJ, B. V.; DE ANDRADE, J. C.; CANTARELLA, H.; QUAGGIO, J. A. **Análise química do solo para fins de fertilidade**. Campinas: Fundação Cargill, 1987. 170 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. 7743, p. 195-204, 2019.
- SANTOS, G. A. D.; MOITINHO, M. R.; SILVA, B. D.; XAVIER, C. V.; TEIXEIRA, D. D.; CORA, J. E.; LA SCALA, 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. 413-424, 2019a.
- SANTOS, H. G. D.; NASCIMENTO, C. F. D.; IZBICKI, R.; DUARTE, Y. A. D. O.; PORTO CHIAVEGATTO FILHO, A. D. Machine learning para análises preditivas em saúde: exemplo de aplicação para predizer óbito em idosos de São Paulo, Brasil. **Cadernos de Saúde Pública**, v. 35, n., p., 2019b.
- SEGNINI, A.; CARVALHO, J. L. N.; BOLONHEZI, D.; MILORI, D.; DA SILVA, W. T. L.; SIMOES, M. L.; CANTARELLA, H.; DE MARIA, I. C.; 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. A.; TRUMBORE, S. E.; DAVIDSON, E. A.; 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. 1, p. 335-356, 2015.
- 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.
- TRAMONTANA, G.; JUNG, M.; SCHWALM, C. R.; ICHII, K.; CAMPS-VALLS, G.; RADULY, B.; REICHSTEIN, M.; ARAIN, M. A.; CESCATTI, A.; KIELY, G.; MERBOLD, L.; SERRANO-ORTIZ, P.; SICKERT, S.; WOLF, S.; PAPALE, D. Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. **Biogeosciences**, v. 13, n. 14, p. 4291-4313, 2016.
- VICENTINI, M. E.; PINOTTI, C. R.; HIRAI, W. Y.; DE MORAES, M. L. T.; MONTANARI, R.; FILHO, M. C. M. T.; MILORI, D. M. B. P.; JÚNIOR, N. L. S.; PANOSSO, A. R. 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., n., p., 2019.
- WARNER, D. L.; 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. 12, p. 1733-1745, 2019.

- WARNER, D. L.; 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. 2, p. 121-135, 2018.
- WMO. **WMO Greenhouse Gas Bulletin (GHG Bulletin) - No. 16: The State of Greenhouse Gases in the Atmosphere Based on Global Observations through 2019**. World Meteorological Organization, World Data Centre for Greenhouse Gases, Tokyo, Japan. https://library.wmo.int/doc_num.php?explnum_id=10437.
- 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. 5, p. 5958-5966, 2011.

Jaboticabal-SP, 30 de julho de 2021



Discente Letícia Roberta de Lima



Prof. Dr. Alan Rodrigo Panosso