

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL**

**APRENDIZADO DE MÁQUINA NA MODELAGEM
TEMPORAL DA EMISSÃO DE CO₂ DO SOLO EM ÁREAS
AGRÍCOLAS NO CERRADO BRASILEIRO**

**Kleve Freddy Ferreira Canteral
Engenheiro Ambiental e de Energias Renováveis**

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL**

**APRENDIZADO DE MÁQUINA NA MODELAGEM
TEMPORAL DA EMISSÃO DE CO₂ DO SOLO EM ÁREAS
AGRÍCOLAS NO CERRADO BRASILEIRO**

Kleve Freddy Ferreira Canteral

Orientador: Prof. Dr. Alan Rodrigo Panosso

Coorientador: Prof. Dr. Nelson José Peruzzi

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias – UNESP, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Agronomia (Ciência do Solo).

2020

C229a

Canteral, Kleve Freddy Ferreira

Aprendizado de máquina na modelagem temporal da variabilidade temporal da emissão de CO₂ do solo em áreas agrícolas no Cerrado brasileiro / Kleve Freddy Ferreira Canteral. -- Jaboticabal, 2020

41 p. : il., tabs., mapas

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientador: Alan Rodrigo Panosso

Coorientador: Nelson José Peruzzi

1. Respiração do Solo. 2. Mudança climática. 3. Reflorestamento. 4. Modelagem ambiental. I. Título.

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

Essa ficha não pode ser modificada.



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal




CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: APRENDIZADO DE MÁQUINA NA MODELAGEM TEMPORAL DA EMISSÃO DE CO₂ DO SOLO EM ÁREAS AGRÍCOLAS NO CERRADO BRASILEIRO


AUTOR: KLEVE FREDDY FERREIRA CANTERAL

ORIENTADOR: ALAN RODRIGO PANOSSO

Aprovado como parte das exigências para obtenção do Título de Mestre em AGRONOMIA (CIÊNCIA DO SOLO), pela Comissão Examinadora:

Prof. Dr. ALAN RODRIGO PANOSSO 
Departamento de Ciências Exatas / FCAV / UNESP - Jaboticabal

Prof. Dr. DANIEL DE BORTOLI TEIXEIRA 
UNIMAR / Marília/SP

Prof.ª Dr.ª TERESA CRISTINA TARLE PISSARRA 
Engenharia Rural / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal - UNESP

Jaboticabal, 29 de julho de 2020

DADOS CURRICULARES DO AUTOR

KLEVE FREDDY FERREIRA CANTERAL - Filho de Maria Cleide dos Santos Ferreira e Kleber Freddy Eslao Canteral, nasceu em Vigia de Nazaré, Pará, no dia 01 de outubro de 1996. Ingressou no curso de Engenharia Ambiental e Energias Renováveis, em março de 2014, pela Universidade Federal Rural da Amazônia (UFRA), Câmpus de Belém, Pará. Foi bolsista pelo Programa de Educação Tutorial em Ciência do Solo (PET Solos) durante o período de julho de 2016 a fevereiro de 2019, no qual desenvolveu atividades voltadas à educação ambiental, reflorestamento e sustentabilidade de solos no contexto da agricultura familiar na Amazônia Oriental. Foi membro do Plano de Logística Sustentável (PLS) da UFRA (2016 a 2018), atuando na dimensão de resíduos sólidos e coleta seletiva. Em março de 2019, ingressou no curso de Mestrado em Agronomia (Ciência do Solo), da Faculdade de Ciências Agrárias e Veterinárias (FCAV-UNESP), Câmpus de Jaboticabal-SP, no Departamento de Ciências Exatas, sob a orientação do Prof. Dr Alan Rodrigo Panosso. Atua principalmente nos seguintes temas: emissão de CO₂ do solo e modelagem matemática aplicada às ciências agrárias.

“É fácil se posicionar no meio da multidão, mas é preciso coragem para se posicionar sozinho.”

(Mahatma Gandhi)

“Se você não é livre para ser você mesmo na questão mais importante de todas as atividades - a expressão do amor - então a vida, em si mesma, perde seu sentido.”

(Harvey Milk)

DEDICO

À minha mãe, Maria Cleide dos Santos Ferreira;
minhas irmãs, Cleidiane Ferreira e Eduarda
Ferreira; e aos amados sobrinhos, João Miguel
Ferreira e Luna Victória Ferreira, que eu tanto amo.

OFEREÇO

Aos meus bondosos avós, Maria Ernestina
dos Santos (*in memoriam*) e João Pereira
Ferreira (*in memoriam*). Vocês são minha
fonte infinita de inspiração.

AGRADECIMENTOS

À minha família, especialmente à minha mãe, Maria Cleide dos Santos Ferreira, por todo o suporte, amor incondicional e confiança depositada a mim. Às minhas irmãs, Cleidiane Ferreira e Maria Eduarda Ferreira, pelo apoio, amizade e por compreenderem a minha ausência em muitos momentos.

Aos meus pequenos e amados sobrinhos, João Miguel Ferreira e Luna Victória Ferreira, que são os meus maiores motivos de felicidade.

Aos meus avós, Maria Ernestina dos Santos (*in memoriam*) e João Pereira Ferreira (*in memoriam*), por serem meus moldes de caráter e humildade e por continuarem sendo o meu alicerce para enfrentar as dificuldades diárias.

Ao meu orientador, Alan Rodrigo Panosso, pelas contribuições científicas, companheirismo, aprendizado e por me motivar ao longo deste trabalho.

Ao meu coorientador, Nelson José Peruzzi, pela disponibilidade em colaborar com este trabalho e pelos valiosos ensinamentos repassados.

Aos professores membros do Exame Geral de Qualificação: Newton La Scala Júnior e Gener Tadeu Pereira. E aos professores membros da Banca de Defesa: Teresa Cristina Tarlé Pissarra e Daniel de Bortoli Teixeira, pela disponibilidade, sugestões e críticas, que certamente contribuíram muito para a qualidade final deste estudo.

Ao Programa de Educação Tutorial em Ciência do Solo (PET Solos), por ter me proporcionado uma vivência multidisciplinar e ter possibilitado o meu crescimento pessoal e profissional. Ao tutor deste grupo, Prof. Dr. Mário Lopes da Silva Júnior, pela transmissão de valores e conhecimentos que perpassam o âmbito acadêmico, além de todo o incentivo e amizade durante os anos de graduação.

Aos amigos, Marcilene Sarah e Fernando Moraes, por terem me motivado a prestar o processo seletivo e me fornecerem todo o suporte inicial durante os primeiros meses de adaptação em Jaboticabal. Vocês são incríveis!!!

Aos (maravilhosos) amigos que fiz durante a graduação: Manuella Raiol, Alessandra Araújo, Letícia Vaz, Eduarda Reis, Samara Pinheiro, João Duarte, Gabriel Martins, Adrienne Braga, Ana Renata Moraes, Maynara Gomes, Larissa Miranda, Mila Façanha, Amanda Segtowich e Diego Pureza. Sou grato pela amizade de todos e espero que este laço tão fraterno se estenda por muitos anos.

Aos amigos que a FCAV-UNESP me apresentou: Luís Fernando Vieira, João de Deus, Eliane Nascimento, Solange Celestino, Jonas Júnior, Vanessa Dantas e

especialmente à Kássia Ferreira, com quem eu dividi apartamento neste período e que me possibilitou uma convivência tão enriquecedora e agradável, tornando a rotina acadêmica menos árdua.

Ao meu namorado, Clebson Praxedes, pelo companheirismo, apoio e incentivo, que foram indispensáveis para a realização deste trabalho.

À Maria Elisa Vicentini, pela amizade, carinho e contribuição intelectual. Teu apoio foi fundamental para que esta pesquisa tivesse andamento, expresso aqui meus sinceros agradecimentos.

Aos amigos e colegas de departamento: Bruna Oliveira, Kárita Almeida, Kamila Meneses, Katharine Batista, Nayane Maia, Tatiana Santos, Ludhanna Veras, Kátia Noronha, Paulo Alexandre, Deise Nogueira, Lígia Correa, Angélica Silva, Aline Moreno, Wanderson Lucena, Washington Pereira e Gustavo André pela amizade, bons momentos partilhados e pelo suporte científico.

Às funcionárias e amigas do Departamento de Engenharia e Ciências Exatas: Maria José Trizólio (Zezé), Adriana Takakura e Shirley de Sousa. Obrigado por serem tão receptivas e tornarem o convívio mais leve e agradável.

Ao Programa de Pós-Graduação em Agronomia (Ciência do Solo) pela oportunidade de realizar esta pesquisa e por todo o conhecimento que me foi transmitido por meio dos docentes da FCAV/UNESP.

À Coordenação de Aperfeiçoamento Pessoal de Nível Superior – CAPES, pela concessão de bolsa de mestrado, o que possibilitou a minha permanência em Jaboticabal e consequente execução deste trabalho.

À Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP (processos nº 2008/58187-0 e 2016/03861-5) pelo apoio financeiro.

A todos aqueles que não mencionei, mas que direta ou indiretamente contribuíram no decurso deste trabalho, prestando ajuda, incentivo ou palavras reconfortantes.

Muito obrigado!

SUMÁRIO

	Página
RESUMO	XII
ABSTRACT	XIII
LISTA DE ABREVIATURAS E SIGLAS	XIV
LISTA DE TABELAS	XV
LISTA DE FIGURAS	XVI
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	3
2.1. Gases de efeito estufa em áreas de reflorestamento.....	3
2.2. Fluxos de CO ₂ associados às propriedades do solo.....	4
2.3. Análise de correlação canônica (ACC).....	5
2.4. Modelagem da emissão de CO ₂ do solo por meio de redes neurais artificiais.....	7
2.5. Arquiteturas das redes neurais artificiais.....	8
2.6. Modelagem da emissão de CO ₂ do solo por meio do sistema de inferência adaptativo neuro-difuso.....	10
3. MATERIAL E MÉTODOS	11
3.1. Amostragem de campo.....	11
3.2. Determinação da emissão de CO ₂ do solo, temperatura e umidade do solo.....	12
3.3. Determinação de variáveis químicas e físicas do solo.....	13
3.4. Mineração dos dados.....	14
3.4.1. Análise de correlação canônica (ACC).....	15
3.4.2. Redes neurais artificiais (RNAs).....	15
3.4.3. Sistema de Inferência Adaptativo Neuro-Difuso (ANFIS).....	17
3.4.4. Critérios para avaliação dos modelos.....	18
4. RESULTADOS	19
4.1. Análise de Correlação Canônica (ACC).....	19
4.2. Redes Neurais Artificiais (RNAs).....	22
4.3. Sistema de Inferência Adaptativo Neuro-Difuso (ANFIS).....	25
5. DISCUSSÃO	27
5.1. Análise de correlação canônica associada às propriedades físico-químicas dos solos.....	27
5.2. Modelagem da FCO ₂ usando RNAs.....	29
5.3. Abordagem ANFIS para modelar a emissão de CO ₂ do solo.....	31

6. CONCLUSÕES.....32
7. REFERÊNCIAS.....33

APRENDIZADO DE MÁQUINA NA MODELAGEM TEMPORAL DA EMISSÃO DE CO₂ DO SOLO EM ÁREAS AGRÍCOLAS NO CERRADO BRASILEIRO

RESUMO – As mudanças no uso da terra e floresta afetam diretamente a dinâmica de carbono entre o sistema solo-atmosfera, portanto assumem importante papel no âmbito das mudanças climáticas. A aplicação de aprendizado de máquina é uma alternativa promissora para modelar sistemas complexos não lineares, como a emissão de CO₂ de solo (FCO₂), fornecendo informações que auxiliem no planejamento e gestão de terras na mitigação do efeito estufa adicional. Nesse contexto, o objetivo deste trabalho foi avaliar o desempenho preditivo de redes neurais artificiais (RNAs), usando três diferentes arquiteturas (GRNN, MLP e RBF), e do sistema de inferência adaptativo neuro-difuso (ANFIS) na modelagem da variabilidade temporal da FCO₂ em diferentes usos do solo na região do Cerrado brasileiro e indicar a técnica mais eficiente. O conjunto de dados deste estudo foi composto por 14 variáveis físicas e químicas do solo em áreas de eucalipto, pinus e reflorestamento com árvores de espécies nativas (mata nativa), totalizando 1500 observações ao longo de 193 dias de estudo, entre novembro de 2015 a maio de 2016. As variáveis de entrada das RNAs e ANFIS foram otimizadas pela análise de correlação canônica (ACC). Os ajustes dos modelos foram testados por análise de variância (teste F ao nível de 1% de significância) e a precisão e acurácia foram avaliadas com base no erro médio absoluto (MEA), erro quadrático médio (RMSE), erro percentual absoluto médio (MAPE), coeficiente de correlação de Pearson (r), índice de concordância (d) e coeficiente de confiança (c). A ACC apresentou alto coeficiente de determinação canônico (R²) (0,764) e as cargas canônicas cruzadas indicaram diferentes variáveis para constituir os modelos. As RNAs mostraram melhor desempenho preditivo que a abordagem ANFIS, exibindo baixos erros e altas correlações associadas ao modelo global durante a fase de calibração (R² = 0,79; RMSE = 0,70 μmol m⁻² s⁻¹) e validação (R² = 0,65; RMSE = 0,85 μmol m⁻² s⁻¹). O desempenho dos modelos individuais e global foi altamente sensível à seleção das variáveis de entrada e a topologia GRNN apontou a maior precisão das RNAs. Nesse contexto, a modelagem da FCO₂ mostrou desempenhos satisfatórios e estes dados podem ser usados na elaboração de relatórios sobre fontes e sumidouros de GEE e assim traçar estratégias para apoiar a mitigação climática, especialmente em regiões tropicais, onde as emissões de CO₂ do solo são frequentemente mais altas.

Palavras-chave: respiração do solo, mudança climática, reflorestamento, modelagem ambiental, RNA, ANFIS.

MACHINE LEARNING IN TEMPORAL MODELING OF SOIL CO₂ EMISSION IN AGRICULTURAL AREAS IN THE BRAZILIAN CERRADO

ABSTRACT - Changes in land and forest use directly affect the carbon dynamics between the soil-atmosphere system and therefore assume an important role in the context of climate change. The machine learning application is a promising alternative for modeling complex non-linear systems, such as CO₂ emission from soil (FCO₂), providing information to assist in land planning and management in mitigating the additional greenhouse effect. In this context, the objective of this work was to evaluate the predictive performance of artificial neural networks (ANNs), using three different architectures (GRNN, MLP and RBF), and the adaptive neuro-fuzzy inference system (ANFIS) in modeling the temporal variability of the FCO₂ in different land uses in the Brazilian Cerrado region and indicate the most efficient technique. The data set of this study was composed of 14 physical and chemical variables of the soil in areas of eucalyptus, pine and reforestation with trees of native species (native forest), totaling 1500 observations over 193 study days, from November 2015 to May 2016. The RNA and ANFIS input variables were optimized by the canonical correlation analysis (ACC). The adjustments of the models were tested by analysis of variance (F test at the level of 1% significance) and precision and accuracy were assessed based on the mean absolute error (MEA), mean square error (RMSE), mean absolute percentage error (MAPE), Pearson's correlation coefficient (r), agreement index (d) and confidence coefficient (c). The ACC presented a high canonical determination coefficient (R²) (0.764) and the crossed canonical loads indicated different variables to constitute the models. The ANNs showed better predictive performance than the ANFIS approach, showing low errors and high correlations associated with the global model during the calibration phase (R² = 0.79; RMSE = 0.70 μmol m⁻² s⁻¹) and validation (R² = 0.65; RMSE = 0.85 μmol m⁻² s⁻¹). The performance of the individual and global models was highly sensitive to the selection of the input variables and the GRNN topology pointed out the greater accuracy of the ANNs. In this context, FCO₂ modeling has shown satisfactory performance and these data can be used in reporting on GHG sources and sinks and thus outline strategies to support climate mitigation, especially in tropical regions, where CO₂ emissions from soil are often taller.

Keywords: soil respiration, climate change, reforestation, environmental modeling, RNA, ANFIS.

LISTA DE ABREVIATURAS E SIGLAS

- ACC** – Análise de Correlação Canônica
- ANFIS** – Sistema de Inferência Adaptativo Neuro-Difuso
- BP** – *Backpropagation*
- c** – Índice de confiança
- C/N** – Carbono/Nitrogênio
- Ca** – Cálcio
- CH₄** – Metano
- CO₂** – Dióxido de carbono
- CTC** – Capacidade de troca de cátions
- d** – Índice de concordância
- Ds** – Densidade do solo
- EstC** – Estoque de carbono
- EstN** – Estoque de nitrogênio
- FCO₂** – Emissão de CO₂ do solo
- FIS** – Sistema de Inferência Difuso
- GEE** – Gases de Efeito Estufa
- GRNN** – Redes Neurais de Regressão Generalizada
- H + AI** – Acidez potencial
- H_{LIFS}** – Grau de humificação da matéria orgânica do solo
- H_{LIFS.DS}** – Interação entre o grau de humificação da matéria orgânica do solo e a densidade do solo
- IA** – Inteligência Artificial
- IPCC** – Painel Intergovernamental de Mudanças Climáticas
- IR** – Índice de Redundância
- K** – Potássio
- Macro** – Macroporosidade do solo
- MAPE** – Erro padrão médio absoluto
- MEA** – Erro médio absoluto
- Mg** – Magnésio
- Micro** – Microporosidade do solo
- MLP** – *Multilayer perceptrons*
- MOS** – Matéria orgânica do solo

N₂O – Óxido nitroso

P – Fósforo

pH – Potencial hidrogeniônico

PLA – Porosidade livre de água

PNN – Redes Neurais Probabilísticas

PVTE – Proporção de variação total explicada

r – Coeficiente de correlação de Pearson

R² – Coeficiente de determinação

RBF – Função de Base Radial

RMSE – Erro quadrático médio da raiz

RNA – Redes neurais artificiais

Ts – Temperatura do solo

UNFCCC – Convenção-Quadro das Nações Unidas sobre a Mudança do Clima

Us – Umidade do solo

VTP – Volume total de poros

LISTA DE TABELAS

Tabela 1. Estatística descritiva: média, erro padrão (EP) e coeficiente de variação (CV) da emissão de CO ₂ do solo, temperatura e atributos químicos e físicos do solo na camada de 0-0,10m em áreas de Eucalipto, Pinus e Mata Nativa. Fonte: Vicentini et al., 2019 (Adaptado).....	14
Tabela 2. Autovalores e correlações canônicas para os atributos dinâmicos e físico-químicos do solo.	20
Tabela 3. Correlações entre os atributos estudados incluindo as cargas canônicas e cargas canônicas cruzadas.	22
Tabela 4. Desempenho dos índices estatísticos para as fases de calibração e validação das redes neurais artificiais.	23
Tabela 5. Desempenho dos índices estatísticos para as fases de calibração e validação do sistema de inferência neuro-difuso adaptativo (ANFIS).	26

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.	12
Figura 2. a) Sistema LI-8100 interligado à câmara de solos; b) câmara para solo inserida sobre o colar de PVC; c) colar de PVC inserido no solo na área de pinus. .	13
Figura 3. Raciocínio difuso tipo 3 (a) e ANFIS equivalente (b). Fonte: Adaptado (Jang, 1993).	17
Figura 4. Matriz de correlação entre os atributos relacionados à dinâmica do carbono e as propriedades físico-químicas do solo.	20
Figura 5. Gráfico de dispersão do primeiro par de variáveis canônicas padronizadas. Atributos dinâmicos do solo (U1): FCO ₂ , EstC, H _{LIFS,DS} ; atributos físico-químicos do solo (V1): P, pH, H + Al, CTC, EstN, CN, Us, Ts, macro, micro, areia, silte e PLA. ..	21
Figura 6. Desempenho do fluxo de CO ₂ do solo na fase de calibração das redes neurais artificiais (RNAs).	24
Figura 7. Desempenho do fluxo de CO ₂ do solo na fase de validação das redes neurais artificiais (RNAs).	25
Figura 8. Modelo de emissão de dados reais e modelados do fluxo de CO ₂ do solo na fase de calibração do sistema de inferência adaptativo neuro-difuso (ANFIS).	26
Figura 9. Modelo de emissão de dados reais e modelados do fluxo de CO ₂ do solo na fase de validação do sistema de inferência adaptativo neuro-difuso (ANFIS).	27

1. INTRODUÇÃO

A degradação do meio ambiente e as mudanças climáticas são importantes questões de interesse global no século XXI (IPCC, 2014; UNFCCC, 2013). A influência antrópica é um fator preeminente no desmatamento das florestas, mudanças de uso da terra e, conseqüentemente, no âmbito das mudanças climáticas (Noojipady et al., 2017). Estima-se que as atividades humanas tenham causado cerca de 1 °C de aquecimento global acima dos níveis pré-industriais, com uma variação provável de 0,8 a 1,2 °C (Olsson et al., 2019).

As mudanças climáticas são causadas pela emissão de gases de efeito estufa (GEE) (IPCC, 2014), sendo o dióxido de carbono (CO₂) o principal responsável pelo efeito estufa adicional, representando cerca de 60% das emissões totais dos GEE (Alexander et al., 2013). Globalmente, a maior fonte de CO₂ para a atmosfera advém da queima de combustíveis fósseis. Contudo, as mudanças de uso da terra correspondem a aproximadamente 22% das emissões globais de gases de efeito estufa. O desmatamento das florestas e a mudança e uso do solo impulsionam o aumento da emissão de GEE, em especial o CO₂, uma vez que os solos florestais armazenam grandes quantidades de carbono (EstC), principalmente em áreas de Cerrado (Noojipady et al., 2017). Assim, o reflorestamento pode auxiliar na mitigação das elevadas taxas de GEE para a atmosfera, visto que o cultivo de florestas, especialmente em regiões tropicais e consorciada com práticas agrícolas, tem sido apontado como alternativa eficaz no estoque de carbono (EstC) (Pulrolnik et al., 2009; De Figueiredo et al., 2017).

O bioma Cerrado brasileiro, também conhecido como savana brasileira, é o segundo maior ecossistema da América do Sul e abrange uma área de aproximadamente 2 milhões de km² na parte central do país, possuindo relevância em âmbito global devido a sua biodiversidade e altas taxas de endemismo (Anach et al., 2018; De Oliveira et al., 2019). Contudo, a expansão agrícola ocorrida no século XX promoveu uma perda superior a 50% da cobertura vegetal nativa deste bioma e em 2014, particularmente no estado do Mato Grosso do Sul (MS), essa perda ultrapassou 70% de sua vegetação (Soares-Filho et al., 2014; Sousa et al., 2015). Nesse contexto, a adoção de sistemas integrados foi apontada como uma estratégia eficiente na gestão de terras para restaurar pastagens degradadas (De Figueiredo et al., 2017) e evitar mais desmatamentos (IPCC, 2014).

A emissão de CO₂ do solo (FCO₂) é resultante de complexas interações entre o clima e os atributos químicos, físicos e biológicos do solo (Ball e Smith, 1991; Kang et al., 2000) e varia de acordo com o tempo, condições ambientais e características dos solos (Carvalho et al., 2018; Freitas et al., 2018; Tavanti et al., 2020). Vários estudos têm sido conduzidos para analisar e compreender a relação entre os GEE e a temperatura global usando abordagens estatísticas convencionais. No entanto, essas técnicas seguem suposições de modelagem probabilística, em que os resultados podem ser associados a grandes erros (Tavares et al., 2018; Khan e Khan, 2019). Por outro lado, técnicas de inteligência artificial (IA), como RNAs e ANFIS, oferecem modelos eficazes quando aprendem e simulam processos (Zendehboudi et al., 2018). Nos últimos anos, essas técnicas têm demonstrado desempenho satisfatório ao modelar problemas complexos e ajustar modelos em funções contínuas não-lineares no âmbito das ciências agrárias (Farhate et al., 2018; Freitas et al., 2018; Altikat et al., 2020).

Redes neurais artificiais (RNAs) são abstrações e simulações da estrutura e função do sistema nervoso biológico (Prieto et al., 2016) que visam identificar as relações intrínsecas entre os dados por meio do processo de aprendizado e o uso de processadores chamados neurônios (Haykin, 1999). Por ter a capacidade de reconhecer padrões nos conjuntos de dados, estudos demonstraram que as RNAs têm melhor desempenho preditivo do que técnicas de regressão linear múltipla (Altikat et al., 2020) e ANFIS, sendo capaz de exibir uma precisão 40% maior (Yilmaz e Kaynar, 2011; Kashi et al., 2014).

Apesar do Cerrado brasileiro ser o segundo maior bioma da América do Sul e ter contribuições notórias a respeito das mudanças climáticas (Brasil, 2016), existem poucas pesquisas voltadas ao uso de inteligência artificial para fins de modelagem do padrão de variabilidade temporal da FCO₂ a partir da correlação entre os atributos físicos e químicos do solo, especialmente com técnicas de RNAs e ANFIS. Tais técnicas podem corroborar para melhor entendimento da dinâmica FCO₂ em áreas de florestas plantadas no Cerrado brasileiro, provendo informações sobre as fontes e sumidouros de GEE em áreas de florestas tropicais, auxiliando na formulação de planos estratégicos que potencializem uma economia de baixo carbono, a fim de mitigar as mudanças climáticas e promover o aumento da resiliência socioambiental. Assim a hipótese básica deste estudo é que a aplicação de técnicas de aprendizado de máquina: RNAs e ANFIS gera modelos preditivos para explicar o padrão de

variabilidade temporal da emissão de CO₂ do solo, a partir de atributos físico-químicos do solo. Nesse contexto, o objetivo do trabalho foi avaliar o desempenho preditivo de redes neurais artificiais (RNAs), usando três diferentes arquiteturas (GRNN, MLP e RBF), e do sistema de inferência adaptativo neuro-difuso (ANFIS) na modelagem da variabilidade temporal da FCO₂ em diferentes usos do solo na região do Cerrado brasileiro e indicar a técnica mais eficiente.

2. REVISÃO DE LITERATURA

2.1. Gases de efeito estufa em áreas de reflorestamento

O efeito estufa é um processo natural, resultante da elevada concentração de gases na atmosfera, como o dióxido de carbono (CO₂), o metano (CH₄), o óxido nitroso (N₂O) e vapores d'água, os quais veem aumentando rapidamente desde a era pré-industrial (IPCC, 2014). As maiores contribuições de gases de efeito estufa (GEE) na atmosfera advém da queima de combustíveis fósseis e de mudanças no uso da terra (Vermeulen et al., 2012; Gilbert, 2012; Smith et al., 2014) e embora as emissões de combustíveis fósseis tenham começado antes da era pré-industrial, somente a partir deste período que estas emissões se tornaram uma fonte dominante de origem antropogênica para a atmosfera e sua participação relativa começou a aumentar (Le Quéré et al., 2015).

As mudanças no uso da terra e floresta, sobretudo vinculadas ao desmatamento, afetam diretamente a entrada e saída de carbono entre os compartimentos ambientais e, portanto, assumem importante papel como fontes de emissões globais de GEE provenientes de atividades antrópicas (Le Quéré et al., 2015). Em países de florestas tropicais, como o Brasil, as emissões de carbono associadas ao desmatamento correspondem a uma grande proporção das emissões totais do país (42% das emissões de CO₂ em 2010) (Brasil, 2016), pois as florestas armazenam grandes quantidades de carbono no solo, principalmente em áreas de Cerrado (Noojipady et al., 2017). Neste contexto, solos de florestas tropicais representam um compartimento chave no processo de emissão e sequestro de carbono, pois podem atuar como fontes ou sumidouros de carbono, dependendo do uso e manejo dos mesmos (Pulrolnik et al., 2009; Funk et al., 2015; Pourbakhtiar et al., 2017).

O bioma Cerrado brasileiro é o segundo maior ecossistema da América do Sul e abrange uma área de aproximadamente 2 milhões de km² na parte central do país,

possuindo relevância em âmbito global devido a sua biodiversidade e altas taxas de endemismo (Anach et al., 2018; De Oliveira et al., 2019). Apesar disto, nas últimas décadas, este bioma sofreu mudanças significativas quanto a sua cobertura vegetal e uso da terra. No período de 1997 a 2010 houve a conversão de aproximadamente 1.587.626 hectares de áreas de florestas desmatadas para pastos limpos entre (Bertani et al., 2015) e em 2014, particularmente no estado do Mato Grosso do Sul (MS), essa perda ultrapassou 70% da vegetação no Cerrado (Soares-Filho et al., 2014; Sousa et al., 2015).

Além disso, o tipo de floresta exerce influência direta sobre as propriedades físicas, químicas e biológicas do solo (Zhou et al., 2013). Segundo Ribeiro e Walter (2008), a vegetação do Cerrado brasileiro é classificada em onze diferentes tipos fisionômicos (mata ciliar, mata de galeria, mata seca, cerradão, cerrado sentido restrito, parque cerrado, palmeiral, vereda, campo sujo, campo limpo e campo rupestre), com base na estrutura, formas de crescimento dominante, possíveis composições florísticas e fatores edáficos. Outros estudos ambientais vinculados às emissões de GEE têm sido reportados em múltiplas áreas de aplicação: mapeamento florestal (Chen et al., 2017); índices de riscos de incêndios florestais (Anezakis et al., 2016; Iliadis et al., 2018); uso e cobertura do solo (Saah et al., 2019); solos (Castelli et al., 2019).

2.2. Fluxos de CO₂ associados às propriedades do solo

A emissão de CO₂ do solo (FCO₂), também conhecida como respiração do solo, é resultante de complexas interações entre atributos relacionados aos processos de produção e transporte de gases do solo para a atmosfera (Ball e Smith, 1991; Kang et al., 2000) e, portanto, representa um importante papel no ciclo do carbono (C) global (IPCC, 2007). Diferentes práticas agrícolas e mudanças de uso da terra desempenham impactos significativos na dinâmica de carbono orgânico do solo e, conseqüentemente, nas emissões de GEE (Paustian et al., 2000; Poch et al., 2005; Lee et al. 2006).

Dentre estes atributos, a temperatura do solo (Ts) e a umidade do solo (Us) são variáveis que exercem grande influência sobre a variabilidade temporal da emissão de CO₂ do solo (FCO₂). A Us está fortemente associada aos ciclos biogeoquímicos, pois constitui um importante fator ambiental que condiciona a disponibilidade de oxigênio no solo e, assim, afeta a dinâmica entre o movimento de

água e gases no interior do solo, está fortemente associada aos ciclos biogeoquímicos (Bubier et al., 2003; Blodau et al., 2004).

No entanto, condições ambientais são altamente sensíveis e isto tem direta implicação na variabilidade dos efeitos ambientais sobre a emissão de CO₂ do solo (Hu et al., 2011). Neste sentido, a dinâmica do solo e características texturais são importantes fatores, visto que a respiração do solo é fortemente sensível às alterações no conteúdo de água no solo e aos teores de areia, silte e argila (Kainiemi et al., 2014).

Atributos físicos do solo exercem diferentes efeitos sobre o fluxo de CO₂ do solo. A porosidade do solo tem impacto direto no transporte gasoso no solo, permitindo tanto a entrada de oxigênio (necessária à atividade aeróbica dos microrganismos) quanto a emissão de CO₂, um subproduto dessa atividade (Teixeira et al. al., 2012; Almeida et al., 2018). Deste modo, alterações neste atributo tendem a refletir diretamente nas emissões de CO₂ do solo. A fração microporos contribui para a retenção de gases, vapores ou materiais em solução na superfície de corpos sólidos devido adesão molecular (Brady e Weil, 2013), sendo este potencial de adesão um fator que pode dificultar as trocas gasosas do sistema solo-atmosfera, assim diminuindo as emissões de CO₂ (Silva et al., 2019).

Panosso et al. (2011) estudaram a FCO₂ e sua relação com propriedades do solo, em áreas cultivadas com cana-de-açúcar, nos sistemas de manejo de cana queimada (CQ) e crua (CC). O estudo apontou que a média de emissão de CO₂ foi significativamente maior ($p < 0,01$) na área de CQ ($2,74 \mu\text{mol m}^{-2} \text{s}^{-1}$ em CQ e $2,07 \mu\text{mol m}^{-2} \text{s}^{-1}$ em CC). Os autores destacam que os teores de pH, fósforo disponível, soma de bases e capacidade de troca de cátions, textura (areia, silte e argila) e interação entre a densidade do solo e o grau de humificação da matéria orgânica do solo ($H_{\text{LIFS.DS}}$) foram propriedades que apresentaram diferenças significativas entre as áreas. Em estudo semelhante, Tavares et al. (2015) apontaram que a densidade do solo, macroporosidade, pH e saturação por bases foram os atributos do solo que mais influenciaram a emissão de CO₂.

2.3. Análise de correlação canônica (ACC)

A análise de correlação canônica (ACC) foi inicialmente proposta por Hotelling (1936) e constitui uma abordagem de estatística multivariada, sendo considerada como uma generalização do método de regressão múltipla (Rencher, 2002; Manly, 2004). O principal objetivo desta técnica é identificar e quantificar associações entre

dois conjuntos de variáveis aleatórias quantitativas (Cooley & Lohnes, 1962; Johnson & Wichern, 2007), resumindo a informação de cada conjunto de variáveis observadas em funções lineares (Jobson, 1992).

Deste modo, a ACC visa maximizar a correlação existente entre o conjunto de variáveis de saída (Hair et al., 2005), fornecendo escores que explicam a máxima variação total possível (Cruz e Regazzi, 1994). Inicialmente, é determinado o par de combinações lineares que apresenta a maior correlação canônica e, em seguida, determina-se o par de variáveis com a segunda maior correlação e, assim, sucessivamente.

Para compreensão do modelo teórico, considere o seguinte vetor $\mathbf{X}_{((p+q) \times 1)}$, com matriz de variâncias e covariâncias amostrais \mathbf{S} e vetor de médias amostrais. Sejam os vetores $\mathbf{X}_{(p \times 1)}^{(1)}$ e $\mathbf{X}_{(q \times 1)}^{(2)}$ como sendo partições do vetor original \mathbf{X} , representando o primeiro grupo com p variáveis e o outro com q , respectivamente (Morrison, 1967). Outra forma de representação do vetor de médias $\bar{\mathbf{X}}_{((p+q) \times 1)}$ (Eq. 1) e da matriz de variância e covariância $\mathbf{S}_{(p+q) \times (p+q)}$ (Eq. 2) de $\mathbf{X}_{((p+q) \times 1)}$ são descritas abaixo:

$$\bar{\mathbf{X}}_{((p+q) \times 1)} = \begin{bmatrix} \bar{\mathbf{X}}^{(1)} \\ \bar{\mathbf{X}}^{(2)} \end{bmatrix} \quad (1)$$

$$\mathbf{S}_{(p+q) \times (p+q)} = \begin{bmatrix} \mathbf{S}_{(p \times p)}^{11} & \mathbf{S}_{(p \times q)}^{12} \\ \mathbf{S}_{(q \times p)}^{21} & \mathbf{S}_{(q \times q)}^{22} \end{bmatrix} \quad (2)$$

As covariâncias entre os pares de variáveis pertencentes aos grupos $\mathbf{X}^{(1)}$ e $\mathbf{X}^{(2)}$ estão contidas em $\mathbf{S}_{(p \times q)}^{12}$ e $\mathbf{S}_{(q \times p)}^{21}$, sendo uma matriz a transposta da outra, assim, os elementos p e q dessas matrizes medem par a par a associação entre as variáveis de ambos os grupos (Johnson & Wichern, 2007).

Estudos como os de Carvalho et al. (2018), utilizaram uma abordagem de estatística multivariada para caracterizar solos sob diferentes usos da terra, no sul da Amazônia brasileira. A partir da ACC, os autores destacam que foi possível identificar e quantificar as associações entre os grupos ligados tanto aos atributos biológicos, quanto físico-químicos do solo. Esta análise evidenciou, sobretudo, a sensibilidade do carbono da biomassa microbiana às alterações iniciais nos atributos físico-químicos do solo, como a saturação das concentrações de alumínio, cálcio e magnésio no solo,

bem como as concentrações de argila e silte. Tavanti et al. (2020) também usaram a técnica de correlação canônica para explorar as interrelações entre dois grupos distintos de variáveis: físico-químicas do solo (V1) e a fração estável dos estoques orgânicos de carbono (U1) e observaram que o componente canônico (U1) explicou aproximadamente 88% da variação total no grupo de atributos relacionados às frações de carbono orgânico, e a componente canônica (V1) representou 17,64% da variação total nos atributos físico-químicos do solo.

2.4. Modelagem da emissão de CO₂ do solo por meio de redes neurais artificiais

Desde a década de 1960, métodos matemáticos e estatísticos têm sido desenvolvidos no estudo da ciência do solo para melhorar a compreensão de processos importantes do solo e também para atuar como ferramentas de avaliação de problemas agrícolas e ambientais (Mc Bratney et al., 2002). No entanto, somente nos últimos 30 anos, devido a não linearidade dos estudos, levou-se à aplicação de métodos mais robustos de modelagem, tais como modelos lineares generalizados (MLG), árvores de decisão e modelos paramétricos não lineares, como redes neurais artificiais (RNAs) e sistema de inferência adaptativo neuro-difuso (ANFIS), a fim de corroborar na compreensão da variabilidade do solo no espaço e tempo (Mc Bratney et al., 2000). Segundo Khan e Khan (2019), um grande número de estudos tem sido conduzido para analisar e compreender a relação entre os GEE e a temperatura global usando abordagens estatísticas convencionais. No entanto, essas técnicas seguem suposições de modelagem probabilística, em que os resultados podem ser associados a grandes erros.

De acordo com Deh Kiani et al. (2010), a aplicação de modelos constitui uma maneira eficaz para quantificar o efeito das práticas agrícolas na emissão de GEE e uma RNA bem treinada pode servir de modelo preditivo para tal finalidade. RNAs constituem uma importante ferramenta para modelar sistemas não lineares e que demandam maior complexidade para serem descritos por métodos analíticos (Haykin, 1999). Além de serem capazes de determinar um padrão de reconhecimento, diagnóstico e previsão, as RNAs são sistemas computacionais paralelos, constituídos por neurônios artificiais (Binoti et al., 2013), cujo comportamento assemelha-se ao cérebro humano, no qual o conhecimento é adquirido por meio do processo de aprendizagem e posterior armazenamento das informações (Haykin, 2001).

Modelos de regressão baseados em RNAs têm se mostrado eficazes para explicar a dinâmica FCO₂. Safari et al. (2016), compararam a técnica de regressão linear múltipla (RLM) e RNAs para simular a FCO₂ em lavouras de trigo na Nova Zelândia. Os resultados demonstraram que as RNAs podem prever FCO₂ das áreas de produção de trigo sob diferentes condições, possuindo melhor desempenho que o método de RLM. Lentzsch et al. (2005), ao avaliarem a biomassa microbiana do solo (BMS) em 30 locais na Alemanha, utilizando um modelo de regressão e uma RNA com um conjunto de 89 pontos em uma rede neural multicamadas feedforward (configuração com dois ou quatro neurônios escondidos e um neurônio de saída), constataram que ambos os modelos demonstram ser promissores para a predição da BMS. Song et al. (2014) observaram que uma RNA com arquitetura de rede Multilayer Perceptron (MLP) com três camadas, treinada com o algoritmo de retropropagação e o método de gradiente descendente, mostrou-se uma abordagem efetiva para estimar FCO₂ em florestas chinesas usando atributos do solo e variáveis meteorológicas.

Utilizando RNAs para prever a FCO₂, Farhate et al. (2018) observaram que a abordagem MLP com o método de seleção de atributos Wrapper possui maior precisão da classe positiva e conseqüente menor custo de classificação, tornando-a mais vantajosa. Similarmente, Freitas et al. (2018) obtiveram resultados satisfatórios ao estimarem a FCO₂ pela rede MLP e pelo algoritmo de retropropagação, visto que os valores preditos se aproximaram dos valores medidos experimentalmente, conforme o MAPE (18,29%) e o coeficiente de determinação obtidos (0,92), assim ratificando o potencial uso de redes neurais na agricultura.

2.5. Arquiteturas das redes neurais artificiais

As RNAs apresentam múltiplas arquiteturas e estas descrevem como os neurônios são estruturados. (Yilmaz e Kaynar, 2011). A primeira distinção que deve ser estabelecida é com relação ao número de camadas, pois RNAs de camada única possuem apenas uma camada de entrada e de saída. Enquanto que redes de multicamadas possuem uma ou mais camadas ocultas entre a entrada e saída (Roman et al., 2020).

A *Multilayer Perceptrons* (MLP) é uma rede neural artificial composta por mais de um perceptron treinada pelo algoritmo *backpropagation* (BP), o qual é responsável por realizar os ajustes dos pesos e de viés em relação ao erro, que pode ser medido de diversas maneiras. Eles são compostos por uma camada de entrada para receber

o sinal, uma camada de saída que toma uma decisão ou previsão sobre a entrada, e entre esses dois, um número arbitrário de camadas ocultas que são o verdadeiro mecanismo computacional do MLP (Haykin, 1999). As MLPs são frequentemente aplicadas a problemas de aprendizagem supervisionados: treinam em um conjunto de pares entrada-saída e aprendem a modelar a correlação entre essas entradas e saídas.

A rede neural de regressão generalizada (GRNN) foi introduzida por Nadaraya e Watson (1964) e redescoberta por Specht (1991). Esta arquitetura é uma generalização de redes de função de base radial (RBFN) e redes neurais probabilísticas (PNN) que podem executar regressões lineares e não lineares (Wachowiak et al, 2001). Além disso, estas redes possuem estrutura fixa: uma camada de entrada; duas intermediárias e uma saída e seu treinamento é do tipo supervisionado, no qual a medida de desempenho usada corresponde ao conjunto de respostas desejadas usando um critério de erro conhecido (Specht, 1991). Essas redes *feedforward* usam arquiteturas de funções básicas que podem aproximar qualquer função arbitrária entre vetores de entrada e saída diretamente de amostras de treinamento e podem ser usadas para interpolação multidimensional (Specht, 1991; Wachowiak et al., 2001; Watson, 1964). As GRNN têm a vantagem de requerem apenas um parâmetro livre e de serem treinadas com facilidade. Estudos envolvendo a aplicação de GRNN para resolver problemas de previsão, controle, modelagem de processos de plantas ou problemas gerais de mapeamento têm sido reportados (Patterson, 1998; Rutkowski, 2004).

As GRNN pertencem à classe de redes neurais amplamente usadas para o mapeamento contínuo de funções, sendo sua principal função estimar uma superfície de regressão linear ou não linear em variáveis dependentes (vetores de saída desejados) a partir de variáveis independentes (vetores de entrada) (Amiri et al., 2010). Outra importante vantagem desta tipologia corresponde à sua simplicidade e procedimentos de aproximação rápida. Ademais, diferentemente das redes baseadas em retropropagação (BP-NN), as GRNN não convergem para mínimos locais (Specht, 1991).

A Função de Base Radial (RBF) constitui outra arquitetura de redes neurais artificiais. Estas redes são uma classe principal de modelo de rede neural, em que a distância entre o vetor de entrada e um vetor de protótipo determina a ativação de uma unidade oculta (Ghodsi e Schuurmans, 2003). Esta tipologia apresenta uma

estrutura simples em termos da direção do fluxo de informações, possuindo uma camada de entrada, uma camada oculta e uma camada de saída (Han et al., 2011). O treinamento das RBF normalmente ocorre em duas etapas: a primeira corresponde aos parâmetros da função base (pertencentes às unidades ocultas) que são determinados por cluster e na segunda etapa são determinados os pesos da camada final por mínimos quadrados (Ghodsi e Schuurmans, 2003).

2.6. Modelagem da emissão de CO₂ do solo por meio do sistema de inferência adaptativo neuro-difuso

O sistema de inferência adaptativo neuro-difuso (ANFIS) foi introduzido por Jang em 1993 por meio da incorporação do Sistema de Inferência Fuzzy (FIS) na estrutura de redes adaptativas (Jang, 1993). O FIS é uma estrutura baseada na teoria de conjuntos e regras difusas se-então, a qual possui três componentes principais, sejam eles: uma base de regras, um banco de dados e um mecanismo de raciocínio. Por outro lado, as redes adaptáveis são estruturas de redes que possuem vários nós conectados por meio de links direcionais e as saídas desses nós adaptáveis dependem de parâmetros modificados pertencentes a estes nós (Jang, Sun e Mizutani, 1997).

A arquitetura ANFIS é composta por cinco camadas, sendo elas: camada difusa; camada de produto; camada normalizada; camada desfavorável e camada de saída total. Cada camada executa uma tarefa específica no sistema de inferência difusa (Jang, 1993). No ANFIS, as capacidades de aprendizagem de uma RNA e as capacidades de raciocínio da lógica difusa são combinadas para fornecer capacidades de previsão aprimoradas, em comparação com o uso de uma única metodologia. Deste modo, o objetivo do ANFIS é encontrar um modelo ou mapeamento que associe corretamente os valores de entrada aos valores de destino (Yilmaz e Kaynar, 2011). Os sistemas híbridos visam à obtenção de um sistema mais eficiente na predição, aprendizado, estimação de parâmetros e na capacidade de generalização e esta abordagem vem sendo gradativamente aplicada nos setores ambientais e agrícolas (Silva et al., 2014).

O método ANFIS tem sido empregado para a resolução de diferentes problemas. Chen et al. (2017) utilizaram o modelo híbrido ANFIS para o mapeamento de suscetibilidade e deslizamentos de terra em uma área propensa a riscos socioambientais na cidade de Hanyuan, China. Os autores concluíram que esta

abordagem atende satisfatoriamente ao objetivo proposto, pois a modelagem de riscos socioambientais é um problema de magnitude não linear, portanto, múltiplas variáveis atuam de forma conjunta no sistema. Similarmente, Alhindawi et al. (2019) utilizaram dados equivalentes a 20 anos para estimar GEE emitidos por transportes rodoviários nos Estados Unidos da América (EUA) por meio do método ANFIS. As variáveis de entrada foram compostas pela proporção veículo-quilômetro por modo (VKM) e o número de veículos de transportes (NTV), sendo observados eficácia da metodologia para tal finalidade.

Estudando a aplicação de RNAs e ANFIS para a estimativa de gases (NO_x , HC, CO, CO_2 e PM) provenientes de transportes coletivos movidos a biodiesel, Gopalakrishnan et al. (2011) obtiveram modelos com baixos erros associados (RMSE < 1,40) e elevados coeficientes de determinação ($R^2 > 0,47$). Najafi et al. (2018) também obtiveram bom desempenho preditivo ao modelar fatores que afetam o rendimento e custo da produção de biodiesel (RMSE < 3,22; $R^2 > 0,93$).

3. MATERIAL E MÉTODOS

3.1. Amostragem de campo

O conjunto de dados é resultado de um estudo de campo conduzido no período de novembro de 2015 a maio de 2016, em áreas de florestas plantadas de eucalipto: 3 hectares contendo uma população básica de eucalipto (*Eucalyptus camaldulensis*) em espaçamento de 4 x 4 m; pinus: população formada por 3 hectares de pinheiro (*Pinus caribaea* var. *Hondurensis*) em espaçamento 3 x 3 m; e reflorestamento com espécies nativas: 21 espécies distribuídas aleatoriamente em espaçamento de 3 x 2 m. Tais áreas são localizadas na Fazenda Experimental da Faculdade de Engenharia (UNESP), na cidade de Selvíria, Mato Grosso do Sul, Brasil Central, às margens do rio Paraná (Figura 1) cujas coordenadas centrais são 20° 20' 53,41" Sul e 51° 23' 55,50" Oeste, com 354 m acima do nível do mar. O solo foi classificado como Latossolo Vermelho distrófico. A topografia da região é caracterizada como moderadamente plana e ondulada. O clima da região é classificado de acordo com o sistema internacional de Köppen, tropical úmido (Aw), com estações chuvosas no verão e seca no inverno, com médias anuais de precipitação de 1.370 mm, temperatura de 23,5 °C e umidade relativa entre 70 e 80%.

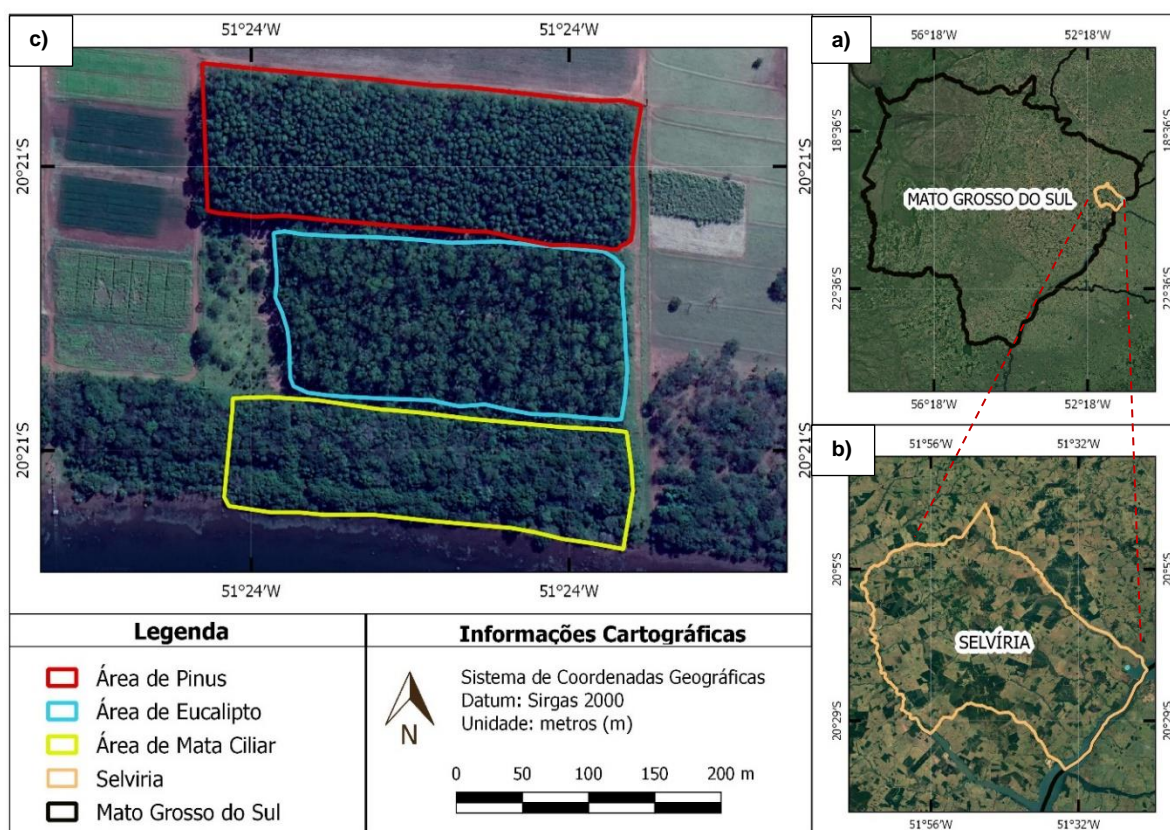


Figura 1. Mapa das áreas de estudo localizadas no município de Selvíria, MS. a) Estado do Mato Grosso do Sul, Brasil Central; b) Município de Selvíria; c) Áreas de florestas plantadas: pinus, eucalipto e mata ciliar, respectivamente.

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

As medições para cada parcela experimental foram feitas por 20 dias, ao longo de 193 dias de estudo, entre novembro de 2015 e maio de 2016. As avaliações para as áreas de estudo foram realizadas no mesmo dia durante a manhã (7 às 12 h). As emissões de CO₂ do solo foram registradas utilizando um sistema de fluxo de solo (LI-8100; LI-COR Bioscience, Nebraska, EUA) (Figura 2-a). Em cada área foram estabelecidos 25 pontos de amostragem usando colares de cloreto de polivinila (PVC) (Figura 2-c) com diâmetro de 0,10 m e altura de 0,085 m; os colares permaneceram fixos durante todo o experimento. A temperatura do solo (Ts) foi determinada nos pontos estudados, com auxílio de um termômetro digital e o conteúdo de água do solo (SWC) foi obtido por meio de um equipamento de reflectometria no domínio do tempo (Hydrosense™; Campbell Scientific, Austrália) (Figura 2.b) contendo duas hastas metálicas de 0,12 m, inseridas perpendicularmente ao solo durante o experimento próximo aos pontos onde a emissão de CO₂ foi medida.

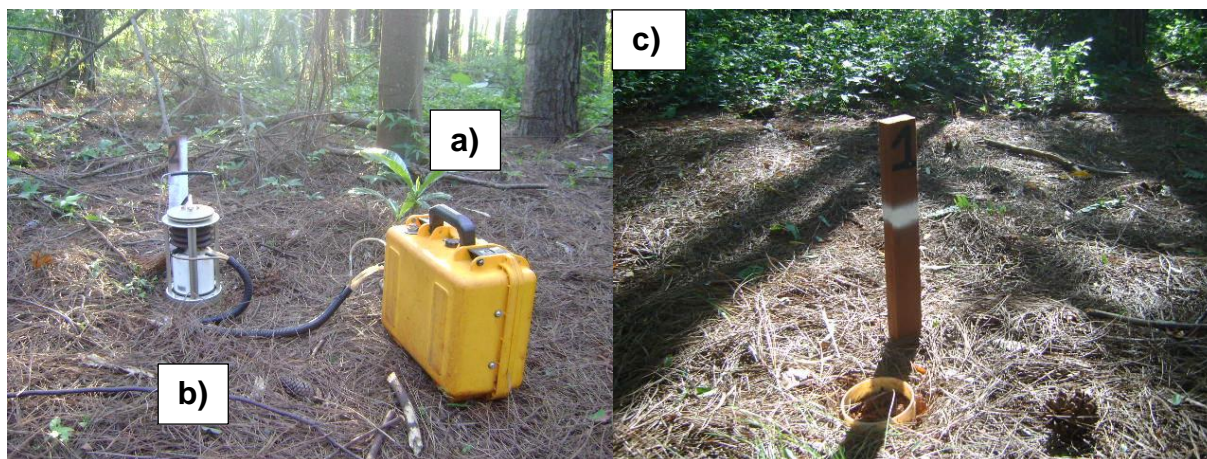


Figura 2. a) Sistema LI-8100 interligado à câmara de solos; b) câmara para solo inserida sobre o colar de PVC; c) colar de PVC inserido no solo na área de pinus.

3.3. Determinação de variáveis químicas e físicas do solo

Amostras de solo foram coletadas na profundidade de 0,0 a 0,10 m, após 193 dias de avaliação. As análises químicas (fósforo (P), potássio (K), cálcio (Ca), magnésio (Mg), alumínio (Al) e acidez potencial (H + Al), matéria orgânica do solo (MOS) e pH) foram determinadas segundo Raij et al. (2001). A capacidade de troca de cátions (CTC), soma de bases (SB) e saturação de base (V%) foram calculadas. Para as análises dos atributos físicos do solo, coletaram-se amostras de solo das profundidades de 0,0 a 0,10 m. A densidade do solo (Ds) foi determinada utilizando amostras de núcleo indeformadas. Foram calculadas a macroporosidade (macro), microporosidade (micro) e o volume total de poros (VTP) (Embrapa 1997).

A análise granulométrica (areia, silte e argila) foi realizada pelo método proposto por Embrapa (1997). A porosidade livre de água (PLA) foi calculada pela diferença entre o VTP e umidade do solo (Us). O estoque de carbono foi calculado com base na massa equivalente do solo (Carvalho et al., 2009). Os estoques de nitrogênio (EstN) foram calculados de maneira semelhante. Os teores de carbono orgânico e nitrogênio foram determinados pelos métodos descritos por Tedesco et al. (1995) e Bataglia et al. (1983), respectivamente. Enquanto o índice de humificação da matéria orgânica do solo (H_{LIFS}) foi determinado usando espectroscopia de fluorescência induzida por laser (LIFS; Milori et al. 2006). Neste estudo, usamos a interação entre a Ds e H_{LIFS} ($H_{LIFS.DS}$), pois expressa melhor a quantidade de ácidos húmicos em cada ponto amostral. Os procedimentos de preparação e análise das amostras de solo foram realizados de acordo com os métodos descritos por Santos et al. (2015).

3.4. Mineração dos dados

O conjunto de dados foi composto por 14 variáveis independentes: atributos químicos (EstC, $H_{LIFS.DS}$, P, pH, H + Al, CTC, EstN, CN) e físicos do solo (Ts, macro, micro, areia, silte e PLA) e uma variável dependente: FCO_2 , totalizando 1500 observações ao longo de 193 dias de estudo, entre novembro de 2015 a maio de 2016. Inicialmente, cada atributo foi avaliado por meio da análise de estatística descritiva dos dados, incluindo média, erro padrão da média e coeficiente de variação (Tabela 1), bem como a avaliação/remoção de outliers multivariados por meio da distância de Mahalanobis ao nível de 1% de significância. Posteriormente, foi avaliada a hipótese de normalidade multivariada, seguida da aplicação da análise de correlação canônica (ACC) para identificar associações entre grupos de variáveis, assim como otimizar as técnicas empregadas: redes neurais artificiais (RNAs) e sistema de inferência neuro-difuso adaptativo (ANFIS).

Tabela 1. Estatística descritiva: média, erro padrão (EP) e coeficiente de variação (CV) da emissão de CO_2 do solo, temperatura e atributos químicos e físicos do solo na camada de 0-0,10m em áreas de Eucalipto, Pinus e Mata Nativa. Fonte: Vicentini et al., 2019 (Adaptado).

Atributos	Eucalipto			Pinus			Mata Nativa		
	Média	EP	CV	Média	EP	CV	Média	EP	CV
FCO_2	5,61a	0,095	37,81	4,06b	0,069	38,26	5,53a	0,083	33,44
Ts	26,52a	0,046	3,84	25,99a	0,058	4,99	25,77a	0,059	5,13
EstC	20,67a	0,172	18,60	13,92c	0,086	13,88	17,50b	0,090	11,51
$H_{LIFS.DS}$	54713,41b	577,202	23,59	74913,21a	1041,645	31,09	50348,11b	329,316	14,63
P	6,52a	0,516	38,87	6,32a	0,51	39,75	6,16a	0,46	36,96
pH	4,312a	0,078	8,89	4,00b	0,02	2,96	4,44a	0,057	6,42
H + Al	55,68b	2,612	23,01	74,80a	3,90	25,57	48,80b	1,88	18,94
CTC	86,99a	3,102	17,49	83,92a	3,63	21,24	82,98a	1,76	10,41
EstN	1,47a	0,021	32,37	0,90b	0,004	8,98	1,57a	0,007	10,22
CN	22,04a	1,673	169,8	15,46a	0,082	11,86	11,16a	0,051	10,23
Macro	0,08a	0,006	39,71	0,04b	0,003	45,67	0,05b	0,007	67,64
Micro	0,32b	0,006	10,79	0,37b	0,004	5,99	0,394a	0,006	8,12
Areia	610,94b	9,884	7,93	648,28a	5,243	3,97	542,39c	5,195	4,70
Silte	56,47a	1,704	14,80	25,76b	3,660	69,79	65,21a	2,770	20,84
PLA	28,32a	0,197	15,53	27,42a	0,170	13,90	23,37b	0,234	22,34

N: 25; médias seguidas da mesma letra (minúscula) nas colunas não diferem entre si (teste Tukey; $p < 0,05$).

FCO_2 : emissão de CO_2 do solo ($\mu\text{mol m}^{-2} \text{s}^{-1}$); Ts: temperatura do solo ($^{\circ}\text{C}$); EstC: estoque de carbono do solo (mg ha^{-1}); $H_{LIFS.DS}$: interação entre o grau de humificação da matéria orgânica e a densidade do solo (unidade arbitrária); P: fósforo (mg dm^{-3}); H + Al: acidez potencial (mmolc dm^{-3}); CTC: capacidade de troca de cátions (mmolc dm^{-3}); EstN: estoque de nitrogênio do solo (mg ha^{-1}); CN: relação carbono/nitrogênio (g kg^{-1}); Macro: macroporosidade do solo ($\text{m}^3 \text{m}^{-3}$); Micro: microporosidade do solo ($\text{m}^3 \text{m}^{-3}$); Teores de areia e silte (g kg^{-1}); PLA: porosidade livre de água.

3.4.1. Análise de correlação canônica (ACC)

A ACC foi aplicada para verificar as associações existentes entre os atributos químicos e físicos do solo (Cruz e Regazzi, 1994), além de otimizar as variáveis de entrada nas RNAs e ANFIS, uma vez que a seleção de entrada dos dados é um passo crucial na implementação destas técnicas (Noori et al., 2010). O estudo dessas associações ocorre por meio de variáveis canônicas, que são construídas a partir de combinações lineares de variáveis ligadas à dinâmica do solo, representadas por U (Eq. 3), e as físico-químicas representadas por V (Eq. 4). Neste estudo, as variáveis dinâmicas do solo relacionam-se a FCO₂, EstC e H_{LIFS.DS}. Enquanto as variáveis físico-químicas correspondem ao P, pH, H + Al, CTC, EstN, CN, Us, Ts, macro, micro, areia, silte e PLA.

$$U = \mathbf{a}^T X^{(1)} \quad (3)$$

$$V = \mathbf{b}^T X^{(2)} \quad (4)$$

Sendo **a** e **b** vetores não nulos dos coeficientes dessas combinações lineares escolhidos para maximizar a correlação entre as variáveis do par canônico, garantindo assim que as variáveis canônicas não sejam correlacionadas, ou seja, sejam ortogonais com o par restante. Os autovalores, seus respectivos autovetores e, conseqüentemente, as correlações e variáveis canônicas foram obtidos a partir das variáveis padronizadas (média 0 e variância 1), por meio da substituição das matrizes de variâncias e covariâncias pelas respectivas matrizes de correlação das variáveis originais.

3.4.2. Redes neurais artificiais (RNAs)

No treinamento das RNAs, a seleção dos neurônios da camada oculta é essencial para o desempenho dos modelos (Boateng et al., 2019) e o número ideal de neurônios da camada oculta geralmente deve ser encontrado usando uma abordagem de tentativa e erro (Maier e Dandy, 2001). Portanto, para fins de otimização dos modelos, as variáveis de entrada das RNAs foram compostas predominantemente pelos atributos físico-químicos do solo que expressaram maiores cargas canônicas cruzadas, geradas pela ACC.

Nesta pesquisa, modelos de RNA com três arquiteturas diferentes: rede neural de regressão generalizada (GRNN), *multilayer perceptrons* (MPL) e redes de função de base radial (RBF) foram treinados usando o programa Statistica7® no modo *Intelligent Problem Solver*. A rede GRNN é uma generalização das RBF e de redes

neurais probabilísticas (PNN). Essa topologia de rede possui estrutura fixa, sendo: uma camada de entrada; duas intermediárias e uma saída; e seu treinamento é do tipo supervisionado, no qual a medida de desempenho usada corresponde ao conjunto de respostas desejadas usando um critério de erro conhecido (Specht, 1991). As redes MLPs são compostas por mais de um perceptron e treinadas pelo algoritmo *backpropagation* (BP), o qual é responsável por realizar os ajustes dos pesos e de bias em relação ao erro, que pode ser medido de diversas maneiras (MEA e RMSE, por exemplo). Eles são compostos por uma camada de entrada para receber o sinal, uma camada de saída que toma uma decisão ou previsão sobre a entrada, e entre esses dois, um número arbitrário de camadas ocultas que são o verdadeiro mecanismo computacional do MLP (Haykin, 1999). Já as redes RBFs apresentam uma estrutura simples em termos de direção do fluxo de informações, possuindo uma camada de entrada, uma camada oculta e uma camada de saída (Han et al., 2011).

Foram construídos modelos individuais para cada uma das áreas de estudo (eucalipto, pinus e mata nativa) com 500 observações, cada; e um modelo global (combinação das três áreas) contendo 1500 observações. Destes dados, 75% foi usado para calibração e 25% para validação dos modelos. As 10 RNAs com menor erro quadrático médio (RMSE) para cada modelo foram apresentadas aleatoriamente pelo programa Statistica7® e foram selecionadas as 3 melhores redes de cada topologia para avaliar seus respectivos desempenhos preditivos. Este procedimento foi realizado para verificar se a complexidade das RNAs afeta a precisão dos modelos. O critério para interromper o algoritmo de calibração utilizado foi o número total de ciclos igual a 50 (Haykin, 2001). A função de ativação utilizada foi sigmoide (ou logística) (Eq. 5). O método de tentativa e erro comumente usados para determinar o número de neurônios na camada oculta foi reduzido, pois a ACC indicou as variáveis que melhor atendiam aos modelos avaliados.

$$f(v) = \frac{1}{1+e^{-av}} \quad (5)$$

Onde:

$f(v)$ = função sigmoide de ativação;

a = estimativa do parâmetro que determina a inclinação sigmoide;

v = potencial de ativação da função.

3.4.3. Sistema de Inferência Adaptativo Neuro-Difuso (ANFIS)

O sistema de inferência adaptativo neuro-difuso (ANFIS) baseado em rede adaptável foi introduzido pela primeira vez por Jang, em 1993. Este método combina os recursos de aprendizado de uma RNA com o raciocínio da lógica difusa a fim de fornecer recursos avançados de previsão e tornar a abordagem mais robusta do que normalmente seria com apenas uma destas técnicas (Yilmaz e Kaynar, 2011; Ghiasi et al., 2016). No ANFIS, o objetivo é encontrar um modelo ou mapeamento que associe corretamente valores de entrada aos valores de saída. O ANFIS é um modelo constituído de cinco camadas (Figura 3) e nesta estrutura conectada os nós de entrada e saída representam os valores de treinamento e os valores previstos, respectivamente, e nas camadas ocultas, existem nós que funcionam como funções de associação e regras (Yilmaz e Kaynar, 2011).

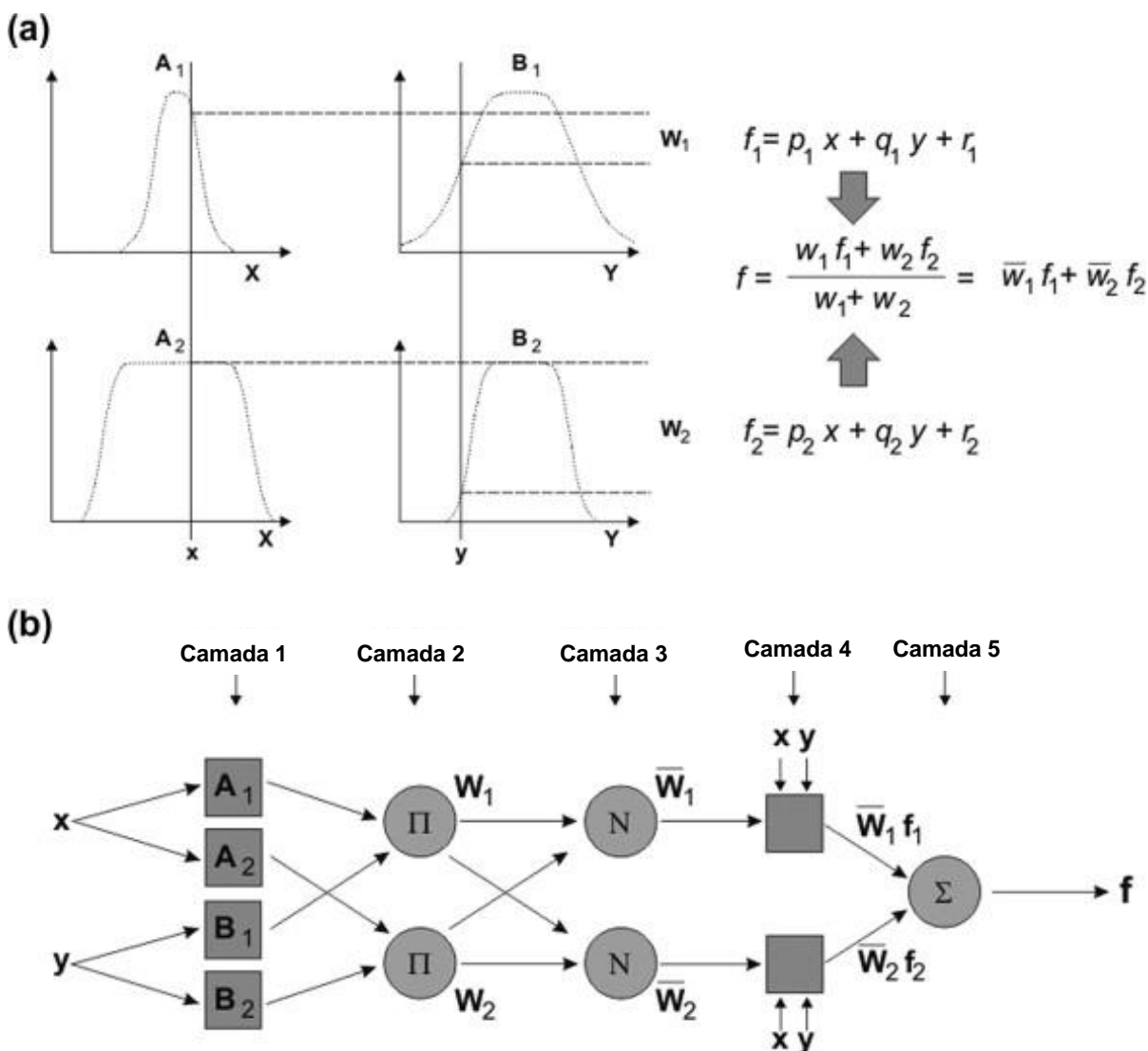


Figura 3. Raciocínio difuso tipo 3 (a) e ANFIS equivalente (b). Fonte: Adaptado (Jang, 1993).

Para simplificar, assume-se que o modelo ANFIS possui duas entradas x e y e uma saída f . O modelo envolve duas regras difusas baseadas no tipo de Takagi e Sugeno, que podem ser expressas da seguinte forma:

$$\text{Regra 1: Se } x \text{ é } A_1 \text{ e } y \text{ é } B_1, \text{ então } f_1 = p_1x + q_1y + r_1, \quad (6)$$

$$\text{Regra 2: Se } y \text{ é } A_2 \text{ e } x_2 \text{ é } B_2, \text{ então } f_2 = p_2x + q_2y + r_2, \quad (7)$$

Conforme demonstra a Figura 3b, diferentes camadas de ANFIS possuem nós distintos. Cada nó em uma camada é fixo ou adaptável (Jang, 1993). Diferentes camadas com seus nós associados são descritas abaixo:

1ª camada: Cada nó i nesta camada é um nó adaptável. Os parâmetros desta camada são denominados de parâmetros de premissa.

2ª camada: Cada nó nesta camada é um nó fixo rotulado Π , cuja saída é o produto de todos os sinais recebidos. Cada saída do nó representa a força de disparo de uma regra.

3ª camada: Cada nó nesta camada é um nó fixo rotulado como N . O nó i^{th} calcula a proporção da força de disparo das regras i . Assim, as saídas dessa camada são chamadas de forças de disparos normalizados.

4ª camada: Cada nó i nesta camada é um nó adaptável. Os parâmetros nesta camada são referidos como parâmetros consequentes.

5ª camada: o nó único nessa camada é um nó fixo chamado Σ , que calcula a saída geral como a soma de todos os sinais recebidos.

A abordagem ANFIS é treinada por um algoritmo de aprendizado híbrido, o qual consiste em uma combinação de descida de gradiente e o método dos mínimos quadrados (Jang, 1993). É importante destacar que, no processo de treinamento do ANFIS, a quantidade de variáveis de entradas pode dificultar a execução e aplicabilidade do sistema, pois se o número de entradas exceder cinco camadas, o número de regras difusas e o tempo computacional aumentará exponencialmente (Kaab et al., 2019). Neste estudo, o ANFIS foi implementado por meio do programa R versão 1.2.5033 (R Core Team, 2020) usando o pacote `frbs`.

3.4.4. Critérios para avaliação dos modelos

A capacidade preditiva dos modelos foi avaliada por meio do desempenho dos índices estatísticos obtidos. Foi calculado o erro médio absoluto (MEA) (Eq. 8), erro quadrático médio da raiz (RMSE) (Eq. 9), erro padrão médio absoluto (MAPE) (Eq.

10), coeficiente de correlação de Pearson (r) (Eq. 11), índice de concordância (d) (Eq. 12) (Willmott, 1981) e coeficiente de confiança (c) (Eq. 13). Ademais, a análise de variância das RNAs foi realizada pelo teste F (nível de probabilidade de 0,01).

$$MEA = \frac{\sum_{i=1}^N (Y_{obs1} - Y_{est1})}{N} \quad (8)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Y_{obs1} - Y_{est1})^2}{N}} \quad (9)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left(\frac{|(A_i - P_i)|}{A_i} 100 \right) \quad (10)$$

$$r = \frac{(N \sum_{i=1}^N (X_{obs1} Y_{obs1}) - \sum_{i=1}^N X_{obs1} \sum_{i=1}^N Y_{obs1} \left((N \sum_{i=1}^N Y_{obs1})^2 - (\sum_{i=1}^N Y_{obs1})^2 \right))}{\sqrt{\left((N \sum_{i=1}^N X_{obs1})^2 - (\sum_{i=1}^N X_{obs1})^2 \right) \left((N \sum_{i=1}^N Y_{obs1})^2 - (\sum_{i=1}^N Y_{obs1})^2 \right)}} \quad (11)$$

$$d = 1 - \frac{\sum_{i=1}^N (Y_{obs1} - Y_{est1})^2}{\sum_{i=1}^N (|Y_{est1} - \bar{Y}| + |Y_{obs1} - \bar{Y}|)^2} \quad (12)$$

$$c = r d \quad (13)$$

Onde: A_i = valor real; P_i = valor previsto; N = número de dados; X_{obs1} = valores observados de X ; Y_{obs1} = valores observados de Y ; \bar{Y} = média observada de \bar{Y} .

4. RESULTADOS

4.1. Análise de Correlação Canônica (ACC)

As correlações lineares obtidas a partir dos atributos dinâmicos do carbono e físico-químicos do solo (Figura 4) apresentaram correlação positiva ($p < 0,05$) entre a FCO_2 com as variáveis estoque de C, pH, estoque de N, macroporosidade e silte e correlações negativas ($p < 0,05$) com os teores de $H_{LIFS.DS}$, $H + Al$ e areia. O estoque de C apresentou correlações positivas ($p < 0,05$) com o teor de P, pH, CTC, estoque de N, macroporosidade e silte e negativas ($p < 0,05$) com o grau de humificação da matéria orgânica, acidez potencial, microporosidade e areia. Por outro lado, o grau de humificação da matéria orgânica esteve correlacionado positivamente ($p < 0,05$) com a acidez potencial e areia, enquanto apresentou correlação negativa ($p < 0,05$) com o teor de P, pH, CTC, estoque de N, macroporosidade e silte.

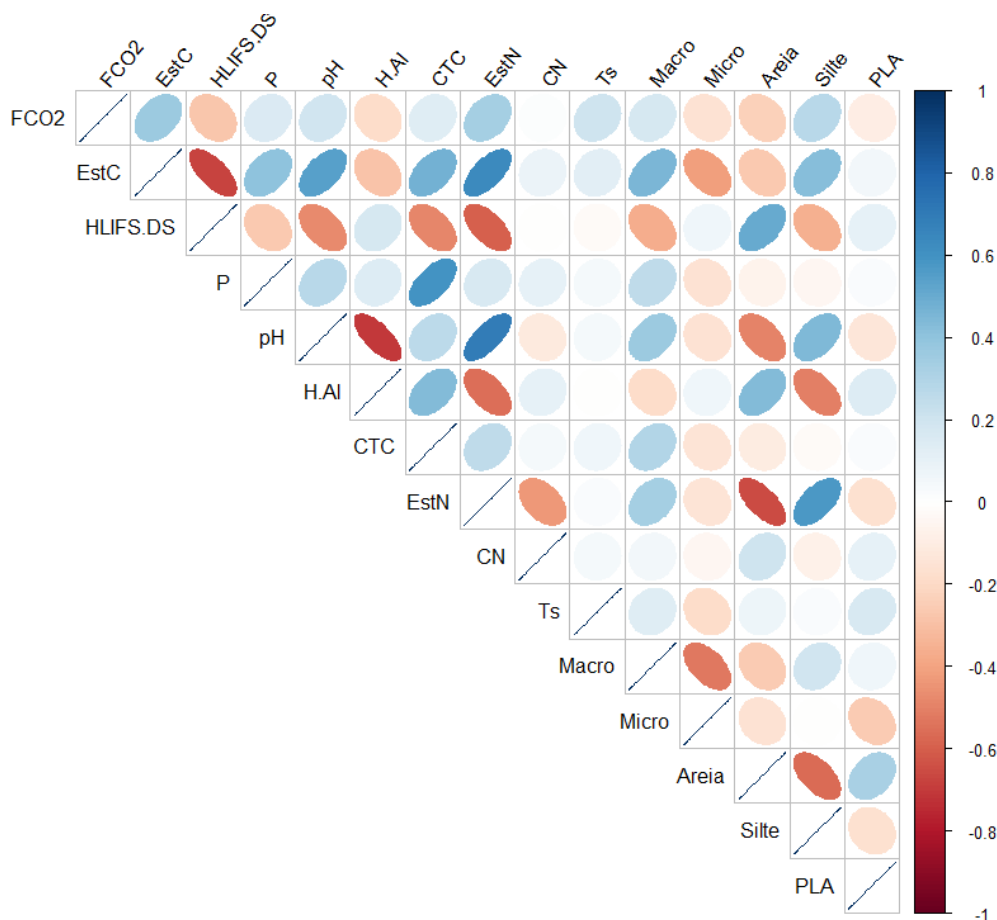


Figura 4. Matriz de correlação entre os atributos relacionados à dinâmica do carbono e as propriedades físico-químicas do solo.

A análise de correlação canônica (ACC), com seus respectivos R^2 canônico e teste de significância são apresentados na Tabela 2. As correlações canônicas (0,874, 0,589 e 0,312) e R^2 canônicos (0,764, 0,347 e 0,097) se mostraram significativos ($p < 0,01$) e sugerem que as variáveis atuam de forma interdependente. O primeiro par de correlação canônica explica aproximadamente 52% da variação conjunta dos atributos, sendo, portanto, o mais importante para a interpretação. Por meio do gráfico de dispersão dos valores padronizados das variáveis estudadas (Figura 5), pode-se inferir que estas propriedades apresentam alta associação linear positiva (0,88).

Tabela 2. Autovalores e correlações canônicas para os atributos dinâmicos e físico-químicos do solo.

Pares de variáveis canônicas	Correlação canônica	R^2 canônico (autovalores)	χ^2	Graus de Liberdade	p-valor
(U1, V1)	0,874	0,764	2939,50	36	< 0,0001
(U2, V2)	0,589	0,347	787,66	22	< 0,0001
(U3, V3)	0,312	0,097	152,57	10	< 0,0001

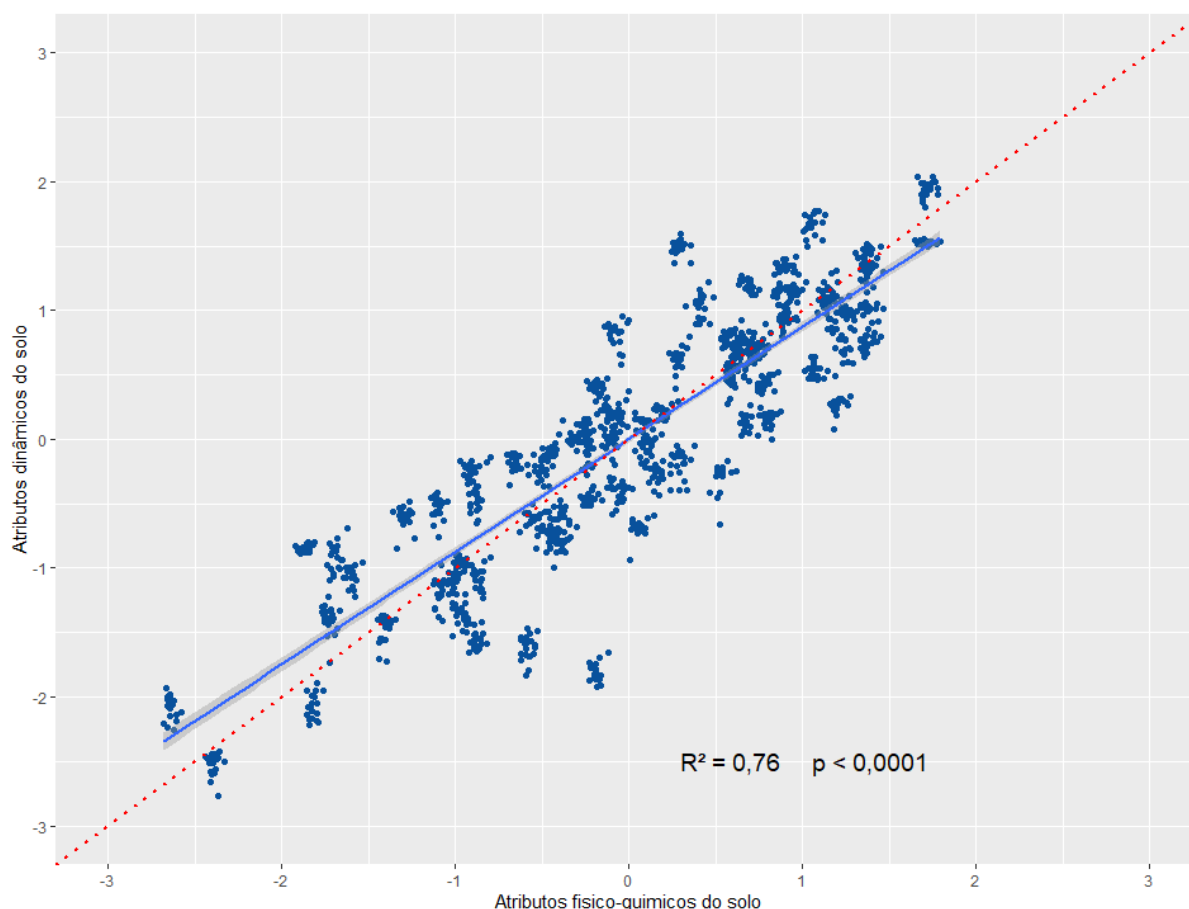


Figura 5. Gráfico de dispersão do primeiro par de variáveis canônicas padronizadas. Atributos dinâmicos do solo (U1): FCO_2 , EstC, $\text{H}_{\text{LIFS.DS}}$; atributos físico-químicos do solo (V1): P, pH, H + Al, CTC, EstN, CN, Us, Ts, macro, micro, areia, silte e PLA.

As proporções totais explicadas pelas propriedades vinculadas à dinâmica do carbono no solo (U1) foi de 56,79%. As variáveis relacionadas aos processos físicos e químicos do solo (V1) explicaram separadamente 20,59% da variação total deste grupo. O segundo par de variáveis (U2 e V2) explicou 16,47% e 10,03%, respectivamente, da proporção da variância total. Ao passo que o terceiro par de variáveis (U3 e V3) foi responsável, respectivamente, pela variação de 26,74% e 6,74%. Esses valores auxiliam na interpretação das variáveis canônicas, pois quanto maior o valor absoluto de uma carga canônica, maior a associação entre a variável original e o respectivo componente canônico.

Para as variáveis físico-químicas do solo (Tabela 3), as maiores correlações canônicas em módulo foram observadas para o estoque de N (-0,654), pH (-0,545), CTC (-0,480) e Macro (-0,461). Similarmente, as maiores associações entre as cargas canônicas cruzadas dos atributos físico-químicos e dinâmicos do carbono foram reportadas para o estoque de N (-0,748), pH (-0,624), CTC (-0,549) e Macro (-0,528).

Tabela 3. Correlações entre os atributos estudados incluindo as cargas canônicas e cargas canônicas cruzadas.

Variável	Cargas canônicas			Cargas canônicas cruzadas		
	U1	U2	U3	V1	V2	V3
FCO ₂	-0,460	0,004	0,888	-0,402	0,002	0,277
H _{LIFS.DS}	-0,993	0,053	-0,107	-0,868	0,031	-0,033
EstC	0,711	0,701	0,051	0,622	0,413	0,016
PVTE (%)	56,79	16,47	26,74			
IR	51,69					
	Cargas canônicas			Cargas canônicas cruzadas		
	V1	V2	V3	U1	U2	U3
P	-0,399	0,030	-0,027	-0,456	0,051	-0,087
pH	-0,545	-0,124	-0,059	-0,624	-0,210	-0,189
H + Al	0,296	-0,045	-0,057	0,338	-0,076	-0,182
CTC	-0,480	-0,200	-0,097	-0,549	-0,340	-0,311
EstN	-0,654	-0,184	0,038	-0,748	-0,313	0,122
CN	-0,080	0,080	-0,024	-0,092	0,135	-0,076
Ts	-0,138	0,092	0,155	-0,158	0,156	0,497
Macro	-0,461	-0,050	-0,040	-0,528	-0,085	-0,129
Micro	0,400	-0,312	0,034	0,458	-0,530	0,110
Areia	0,305	0,412	-0,104	0,349	0,699	-0,332
Silte	-0,440	-0,072	0,079	-0,503	-0,122	0,255
PLA	-0,029	0,189	-0,127	-0,033	0,320	-0,406
PVTE (%)	20,59	10,03	6,74			
IR	19,86					

FCO₂: emissão de CO₂ do solo; H_{LIFS.DS}: interação entre o grau de humificação da matéria orgânica e a densidade do solo; EstC: estoque de carbono do solo; P: fósforo; H + Al: acidez potencial; CTC: capacidade de troca de cátions; EstN: estoque de nitrogênio do solo; CN: relação carbono/nitrogênio; Ts: temperatura do solo; Macro: macroporosidade do solo; Micro: microporosidade do solo; PVTE (%): proporção de variação total explicada; IR: índice de redundância.

4.2. Redes Neurais Artificiais (RNAs)

A seleção de variáveis de entrada não contemplou os mesmos atributos para cada modelo avaliado, sejam eles: eucalipto (EstC, H_{LIFS.DS}, pH, H + Al, areia, silte e macro); pinus (EstC, H_{LIFS.DS}, PLA, Ts, micro, macro e pH); mata (EstC, H_{LIFS.DS}, Ts, areia, CTC, PLA e silte); e global (EstC, H_{LIFS.DS}, EstN, pH, CTC, H + Al, PLA, Ts, micro, macro, silte, areia). O modelo global precisou de mais variáveis de entrada que as demais áreas para atingir melhor precisão e acurácia, uma vez que este modelo é composto pela junção das três referidas áreas e a diferença da dinâmica de carbono destes ambientes influi diretamente nos atributos físicos e químicos do solo, que por sua vez afetam a dinâmica do fluxo de CO₂ do solo.

Os modelos gerados pelas redes neurais artificiais (RNAs) foram estatisticamente significativos ($p < 0,01$). A Tabela 4 apresenta os índices estatísticos para as três melhores redes obtidas para os modelos individuais e global. Observa-se que os melhores modelos foram obtidos por meio do teste de diferentes arquiteturas de RNAs, e de acordo com os testes, as melhores foram do tipo GRNN, com exceção da área de Pinus, que foi melhor modelada pela topologia RBF.

Tabela 4. Desempenho dos índices estatísticos para as fases de calibração e validação das redes neurais artificiais.

Área de eucalipto										
Topologia	Fase	MEA	RMSE	MAPE	r	d	Classificação (c)	F	Valor-p	
RNA 1 GRNN 7-250-2 ^a	Calibração	0,73	0,80	0,15	0,89	0,94	0,83	Muito bom	931,60	< 0,0001
	Validação	0,93	1,01	0,19	0,77	0,87	0,67	Bom	161,90	< 0,0001
RNA 2 RBF 7-29-1	Calibração	0,84	0,94	0,16	0,85	0,92	0,78	Muito bom	635,20	< 0,0001
	Validação	0,87	0,98	0,18	0,83	0,91	0,75	Bom	250,90	< 0,0001
RNA 3 MLP 7-8-1	Calibração	0,85	0,89	0,17	0,84	0,91	0,76	Muito bom	564,20	< 0,0001
	Validação	0,94	0,91	0,19	0,78	0,87	0,67	Bom	174,50	< 0,0001
Área de pinus										
RNA 4 RBF 7-65-1	Calibração	0,70	0,69	0,21	0,74	0,84	0,62	Mediano	298,50	< 0,0001
	Validação	0,79	0,81	0,21	0,71	0,82	0,58	Sufrível	116,00	< 0,0001
RNA 5 GRNN 7-250-2	Calibração	0,72	0,56	0,21	0,79	0,83	0,66	Bom	411,80	< 0,0001
	Validação	0,88	0,71	0,25	0,64	0,75	0,48	Mau	80,28	< 0,0001
RNA 6 MLP 7-11-1	Calibração	0,81	0,67	0,24	0,65	0,76	0,49	Mau	173,70	< 0,0001
	Validação	0,94	0,56	0,25	0,60	0,67	0,41	Péssimo	66,32	< 0,0001
Área de mata nativa										
RNA 7 GRNN 7-250-2	Calibração	0,66	0,67	0,13	0,86	0,90	0,78	Muito bom	667,40	< 0,0001
	Validação	0,85	0,81	0,17	0,80	0,87	0,70	Bom	207,60	< 0,0001
RNA 8 RBF 7-74-1	Calibração	0,83	0,86	0,17	0,79	0,88	0,69	Bom	414,60	< 0,0001
	Validação	0,96	1,10	0,20	0,77	0,87	0,67	Bom	166,20	< 0,0001
RNA 9 MLP 7-6-1	Calibração	0,83	0,79	0,18	0,73	0,83	0,60	Sufrível	259,70	< 0,0001
	Validação	0,96	0,82	0,19	0,67	0,78	0,52	Sufrível	95,12	< 0,0001
Modelo Global										
RNA 10 GRNN 12-750-2	Calibração	0,66	0,70	0,14	0,89	0,93	0,83	Muito bom	2718,0	< 0,0001
	Validação	0,80	0,85	0,19	0,81	0,89	0,71	Bom	646,00	< 0,0001
RNA 11 RBF 12-100-1	Calibração	0,77	0,85	0,17	0,84	0,91	0,77	Muito bom	1805,0	< 0,0001
	Validação	0,90	1,00	0,21	0,77	0,87	0,68	Bom	517,10	< 0,0001
RNA 12 MLP 12-10-1	Calibração	0,91	0,90	0,21	0,76	0,85	0,65	Mediano	975,30	< 0,0001
	Validação	0,96	0,86	0,23	0,75	0,83	0,62	Mediano	436,50	< 0,0001

^a Sequência de números indica o número de variáveis de entradas, n^o de neurônios na primeira camada e n^o de neurônios na segunda camada, respectivamente; MEA: erro médio absoluto; RMSE: erro quadrático médio da raiz; MAPE: erro padrão médio absoluto; r: correlação de Pearson; d: índice de concordância de Willmott; c: coeficiente de confiança; p-valor: probabilidade de ocorrência de H₀.

Na fase de calibração, as RNAs apresentaram índices de erros semelhantes, apontando baixo MEA (0,66 a 0,91 $\mu\text{mol m}^{-2} \text{s}^{-1}$), RMSE com valores variando de 0,56 a 0,94 $\mu\text{mol m}^{-2} \text{s}^{-1}$ e MAPE com variação entre 0,13 a 0,25 (Tabela 4). Quanto a precisão dos modelos, todos mostraram alta correlação e os índices de concordância foram superiores a 0,75 para todas as redes testadas (Willmott, 1981), o que demonstra um ajuste muito bom entre os valores observados e estimados nos respectivos modelos, conforme o índice de classificação (Camargo e Sentelhas, 1997).

A Figura 6 apresenta o gráfico de desempenho para a fase de calibração para os modelos individuais e global. A área de mata apresentou os menores erros absolutos e percentuais, ao passo que a área de eucalipto e o modelo global exibiram os maiores coeficientes de determinação (0,79).

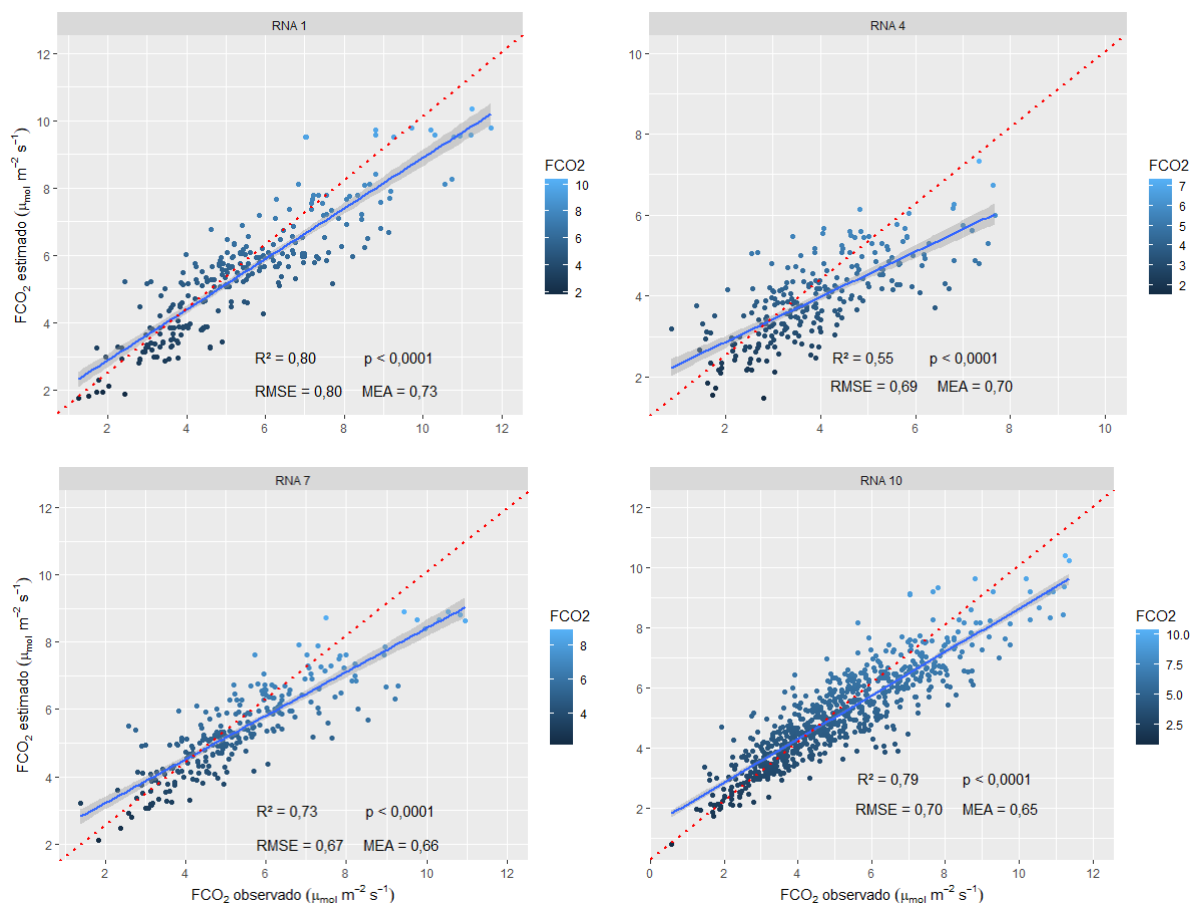


Figura 6. Desempenho do fluxo de CO₂ do solo na fase de calibração das redes neurais artificiais (RNAs).

Apesar dos índices de erros apresentarem pouca variação durante a fase de validação, o modelo global exibiu os maiores erros percentuais e absolutos (Tabela 4). Contudo, as maiores correlações foram associadas ao modelo global juntamente com a área de eucalipto. Conforme o desempenho das RNAs na fase de validação (Figura 7), verifica-se que o modelo global apresentou o maior coeficiente de determinação (0,65), índice de concordância (0,89) e o coeficiente de confiança (0,71), este último ratifica um bom ajuste entre os valores observados e estimados (Tabela 4).

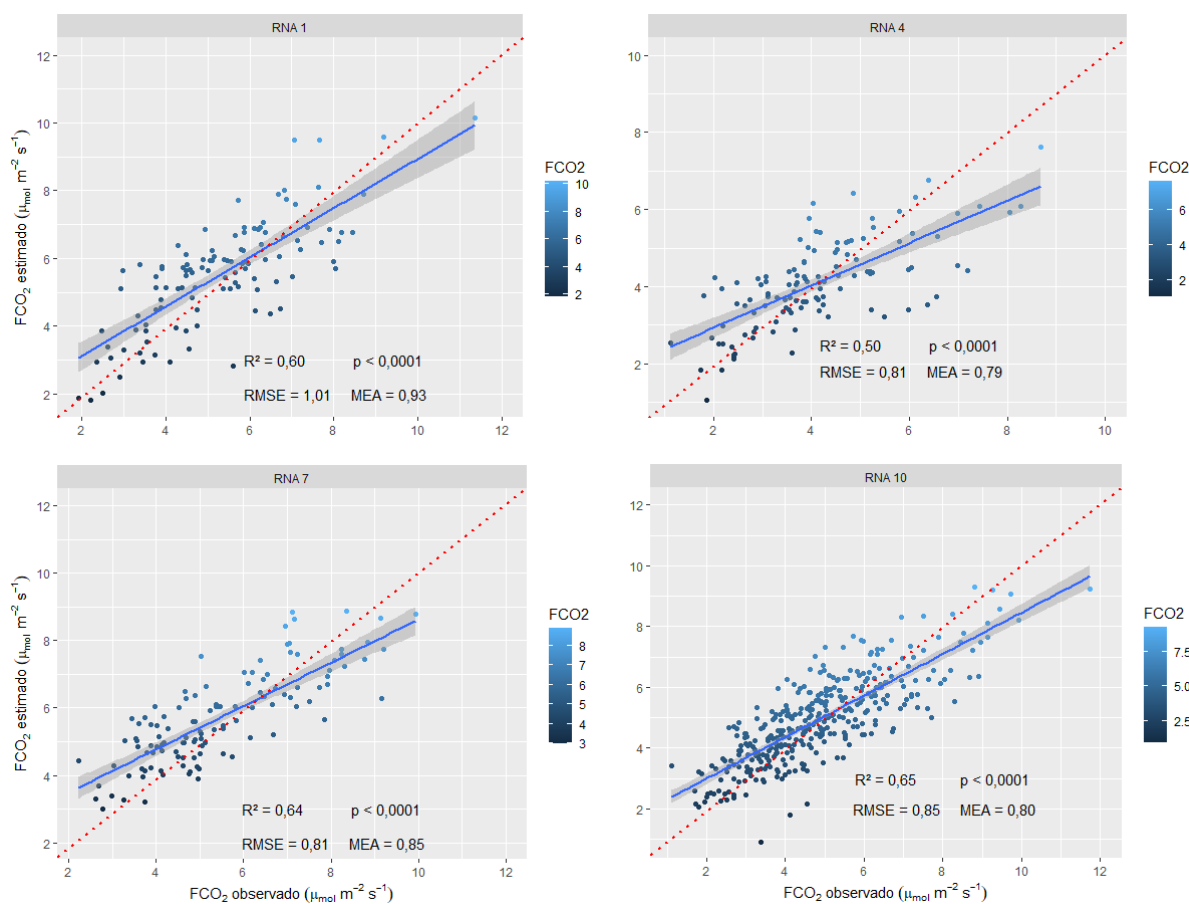


Figura 7. Desempenho do fluxo de CO₂ do solo na fase de validação das redes neurais artificiais (RNAs).

4.3. Sistema de Inferência Adaptativo Neuro-Difuso (ANFIS)

Os parâmetros estatísticos do método ANFIS são apresentados na Tabela 4. De maneira geral, durante a fase de calibração dos dados, observa-se que os índices de erros e as correlações foram semelhantes para todos os modelos, apresentando baixos valores de RMSE (1,71 a 2,16 $\mu\text{mol m}^{-2} \text{s}^{-1}$), MEA (1,36 a 1,69 $\mu\text{mol m}^{-2} \text{s}^{-1}$) e elevados MAPE (31,16 a 37,93 %) e paralelamente apresentaram correlações variando de 0,77 a 0,81. É válido ressaltar que a maior precisão dos modelos derivou da utilização de diferentes variáveis de entrada para cada área estudada, sendo composta por atributos químicos (EstC, EstN, H_{LIFS_DS} e CTC) e, predominantemente, por atributos físicos do solo (PLA, silte, areia, macro e microporosidade).

Tabela 5. Desempenho dos índices estatísticos para as fases de calibração e validação do sistema de inferência neuro-difuso adaptativo (ANFIS).

Modelos	Fase	MEA	RMSE	MAPE	r
ANFIS 1	Calibração	1,54	2,01	31,16	0,81
Eucalipto	Validação	1,37	1,80	28,84	0,84
ANFIS 2	Calibração	1,36	1,71	37,93	0,77
Pinus	Validação	1,58	1,98	50,46	0,73
ANFIS 3	Calibração	1,69	2,16	32,91	0,78
Mata nativa	Validação	1,50	1,94	31,40	0,80
ANFIS 4	Calibração	1,47	1,88	33,45	0,80
Global	Validação	1,51	1,93	35,64	0,80

A Figura 8 ilustra o gráfico de desempenho para a fase de calibração dos modelos individuais e global. Nesta fase, os maiores erros absolutos foram observados para a área de mata nativa (MEA = 1,69 $\mu\text{mol m}^{-2} \text{s}^{-1}$; RMSE = 2,16 $\mu\text{mol m}^{-2} \text{s}^{-1}$; Tabela 5) e elevados erros percentuais foram associados à área de pinus (MAPE = 37,93%). Entretanto, a área de eucalipto e o modelo global exibiram as maiores correlações (0,81 e 0,80), respectivamente.

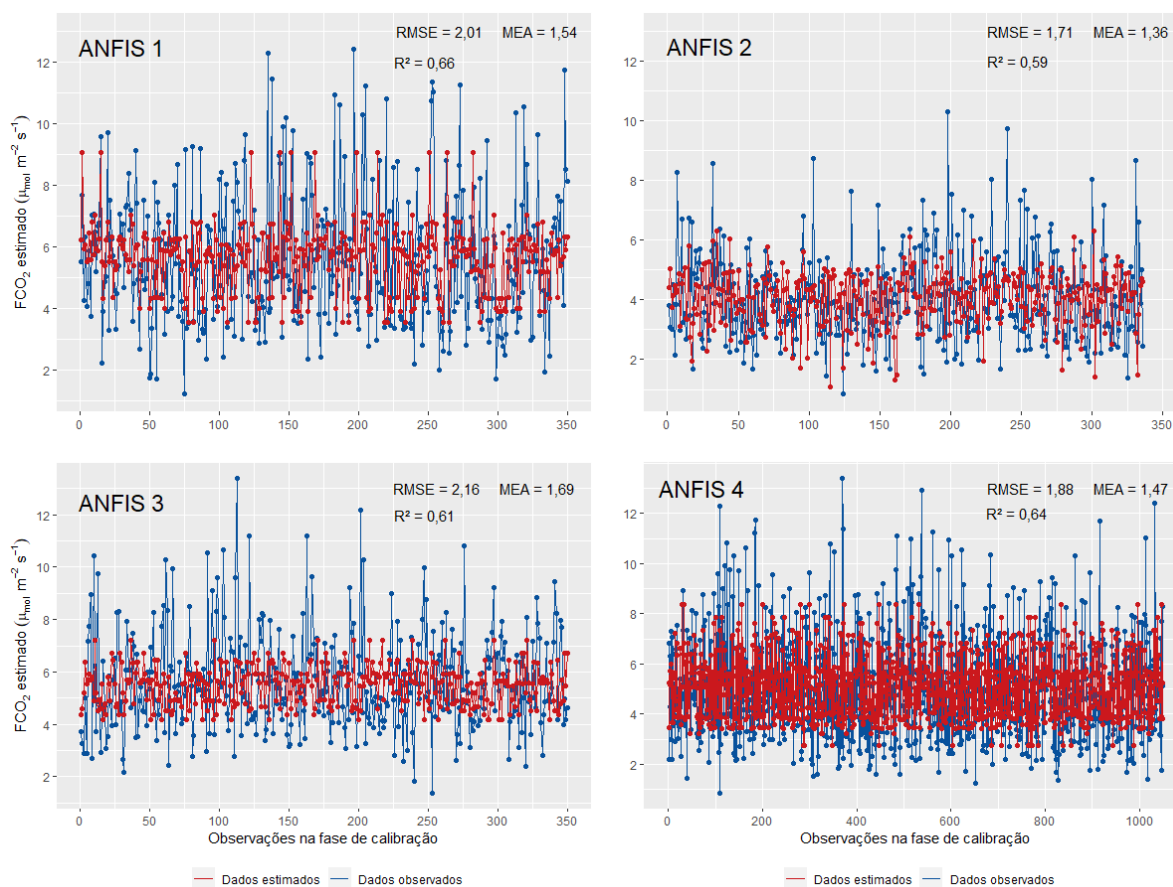


Figura 8. Modelo de emissão de dados reais e modelados do fluxo de CO₂ do solo na fase de calibração do sistema de inferência adaptativo neuro-difuso (ANFIS).

Na fase de validação, a área de eucalipto apresentou os melhores desempenhos estatísticos (MEA = 1,37 $\mu\text{mol m}^{-2} \text{s}^{-1}$; RMSE = 1,80 $\mu\text{mol m}^{-2} \text{s}^{-1}$; MAPE = 28,84%; $r = 0,84$). Diversamente, os maiores erros e a menor correlação foram apontados para a área de pinus (Tabela 5). O desempenho dos modelos individuais e global para a fase de validação é representado na Figura 9.

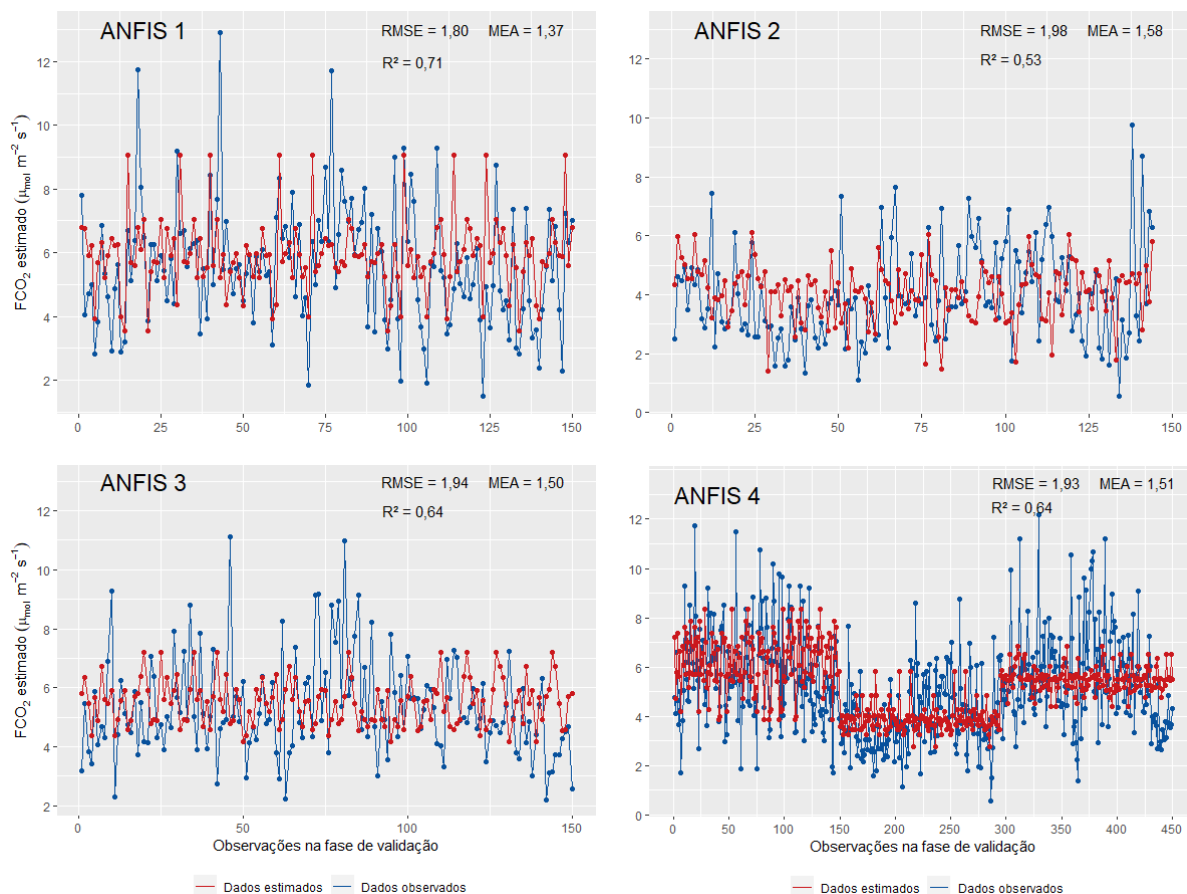


Figura 9. Modelo de emissão de dados reais e modelados do fluxo de CO₂ do solo na fase de validação do sistema de inferência adaptativo neuro-difuso (ANFIS).

5. DISCUSSÃO

5.1. Análise de correlação canônica associada às propriedades físico-químicas dos solos

A dinâmica do solo nas áreas de florestas plantadas: pinus, eucalipto e reflorestamento com espécies de mata nativa indica a complexidade das sequências e transformações químicas e físicas que ocorrem nestes ambientes. A maior correlação significativa ($p < 0,05$) existente entre qualquer variável ligada ao grupo dinâmica do carbono do solo e às variáveis físico-químicas correspondeu a 0,64 (EstC e EstN) (Figura 4). No geral, o estoque de C apresentou as maiores correlações com a componente físico-química do solo, visto que o adequado manejo do solo, ao prover

maior acúmulo e ciclagem de matéria orgânica, pode maximizar significativamente os estoques de C ao longo do tempo (Souza et al., 2018). Li et al. (2012) observaram que a dinâmica do EstC está intimamente ligada à dinâmica de EstN, isso porque o aumento do estoque de N reduz a limitação de nitrogênio e contribui com o sequestro de C a longo prazo. Portanto, caso o sequestro de C não seja acompanhado de um ganho simultâneo de N, isto é, a relação estequiométrica entre estes elementos não seja equivalente, o sistema sofrerá um desbalanço (Luo et al., 2004), reduzindo a taxa de sequestro de C e intensificando as emissões de CO₂ do solo.

A interação entre a densidade do solo e o grau de humificação da matéria orgânica do solo ($H_{LIFS.DS}$) expressa melhor a quantidade de ácidos húmicos em cada ponto amostral. Essa interação pode ser entendida como U1 e apresentou os maiores valores em módulo da correlação canônica (-0,993), seguido pelo EstC (0,711) e FCO₂ (-0,460) (Tabela 3). A $H_{LIFS.DS}$ apresentou carga canônica alta e negativa (-0,868), indicando que os processos dinâmicos que ocorrem no solo em função de elevados teores deste atributo apresentaram uma tendência similar de emissão de CO₂ e de estoque de carbono. Por outro lado, as variáveis responsáveis pelos maiores valores em V1 foram EstN (-0,654), pH (-0,545), CTC (-0,480) Macro (-0,461) e silte (0,440). Este padrão revela que U1 pode ser interpretada por $H_{LIFS.DS}$, enquanto que V1 pode ser entendida como a combinação de atributos químicos e físicos do solo. Na prática, solos com maiores valores de $H_{LIFS.DS}$ tendem a possuir baixos teores de silte e elevados valores de EstN, pH, CTC e macro, visto que as propriedades físicas estão intimamente ligadas ao processo de transporte de gases no interior do solo.

De forma semelhante, as cargas canônicas cruzadas (U1) tiveram altas correlações canônicas e corresponderam aos mesmos atributos das cargas canônicas: EstN (-0,748), pH (-0,624), CTC (-0,549) macro (-0,528) e silte (-0,503) (Tabela 3). Propriedades químicas do solo, como o pH e CTC, estão diretamente relacionadas com a MOS e, conseqüentemente, ligadas ao EstC, atuando na taxa de agregação e dinâmica da água no solo e na ciclagem e retenção de nutrientes do solo (Six et al., 2006), limitando o crescimento microbiano e a taxa de respiração (Lamparter et al., 2009).

A Ts atua em importantes processos, sejam eles de natureza fisiológica (evapotranspiração) ou bioquímica (aumento da atividade enzimática) e, portanto, corresponde a um dos principais fatores ambientais que governam as taxas de respiração do solo, sobretudo se tratando da variabilidade temporal da FCO₂

(Wallenstein et al., 2011; Flanagan et al., 2013). Apesar disto, o presente estudo não encontrou correlação significativa ($p > 0,05$; Figura 4) entre as emissões e a temperatura do solo. Contudo, as cargas canônicas cruzadas desta variável foram as mais altas em U3 (0,497), sugerindo que a componente 3 é o atributo que exerce maior influência na dinâmica FCO_2 .

Adicionalmente, a análise de redundância demonstrou que 51,69% da variação total dos atributos dinâmicos do carbono foi explicado pela variação dos atributos físico-químicos do solo para a avaliação da primeira correlação canônica. Em contrapartida, apenas 19,86% da variação total dos atributos físico-químicos é explicado pela componente dinâmica de carbono do solo (Tabela 2). Estes resultados refletem o quanto a dinâmica de carbono no solo é influenciada pelas propriedades físico-químicas, sobretudo o EstN, CTC, H + Al.

5.2. Modelagem da FCO_2 usando RNAs

No geral, as arquiteturas GRNN forneceram os melhores ajustes para os modelos estudados. Esta tendência pode ser explicada, pois a maior vantagem da topologia GRNN está relacionada à sua capacidade de aprender o problema mais rapidamente e assim convergir para a superfície de regressão óptica com a mesma rapidez e com grande número de conjuntos de dados (Pandey e Mishra, 2017). Além disso, a rede GRNN forneceu os maiores números de camadas ocultas aos modelos e o número de neurônios ligados a essas camadas resulta maior complexidade da rede (Fernandes et al., 2019) para fins de reconhecimento de padrões não lineares entre conjuntos de dados cujas variáveis não expressam correlação direta entre elas ou dependam de outras múltiplas variáveis (Maciel et al., 2012). Essa observação realça a complexidade de modelar a variável FCO_2 , visto que esta propriedade é resultante de complexas interações entre os compartimentos ambientais (Ball e Smith, 1991; Kang et al., 2000; Freitas et al., 2018) e no presente estudo foram usadas apenas variáveis vinculadas às propriedades físico-químicas do solo.

Nossos achados convergem com a pesquisa de Pandey e Mishra (2017), que usaram duas redes neurais diferentes (GRNN e RBF) para avaliar o rendimento de batatas. De acordo com sua conclusão, ambas as RNAs apresentam desempenho preciso, mas o GRNN forneceu melhores resultados que o RBF. Usando a mesma técnica, com as arquiteturas MLP, BPNN e GRNN, e modelos de regressão múltipla (RLM e MNLR) para modelar as emissões de CO_2 de reservatórios, Chen et al. (2018)

concluíram que as RNAs forneceram melhor capacidade preditiva que os modelos de regressão múltipla e a rede que obteve melhor desempenho foi a GRNN. Usando RNAs para prever a FCO₂ em áreas de cana-de-açúcar no estado de SP, Freitas et al. (2018) testaram diferentes arquiteturas de redes: GRNN e MLP. Entretanto, diferente do presente estudo, os melhores ajustes e menores erros foram reportados para a topologia MLP (MAPE = 18,29%), pois a GRNN apresentou o maior MAPE (43,02%).

Neste estudo, a variável EstC foi incluída como variável de entrada para os modelos de RNAs, devido aos altos valores apontados nas cargas canônicas cruzadas, sugerindo importância desta variável para a obtenção de melhores desempenhos nas redes. Em geral, o aumento da fixação de carbono no solo, obtido pelo efeito das práticas agrícolas de manejo e conservação, é observado a longo prazo (Maia et al., 2013). No Cerrado brasileiro são encontradas grandes reservas de carbono no solo, assim como na biomassa subterrânea, visto que a vegetação deste bioma aloca recursos substanciais para a produção de raízes (De Miranda et al., 2014). Portanto, os estoques de carbono associados a estes dois compartimentos provavelmente representam um relevante papel nas emissões nacionais de gases de efeito estufa no Brasil.

A porosidade livre de água (PLA) também apontou desempenhos satisfatórios para os modelos de RNAs. Este atributo atua principalmente na aeração do solo, afetando tanto a entrada do oxigênio necessário para a realização da atividade dos microrganismos aeróbicos, quanto a saída do CO₂, que é um subproduto final dessa atividade. Além disso, verificou-se que os maiores valores médios de macroporosidade (0,082 m³ m⁻³) associados aos menores teores de microporosidade (0,31 m³ m⁻³) na área de eucalipto estão intimamente ligados a FCO₂. Estes atributos podem explicar os maiores padrões temporais de emissão de CO₂ para a área de eucalipto (5,61 μmol m⁻² s⁻¹), uma vez que o sistema de poros condiciona o armazenamento e o movimento de água e gás no interior dos solos (Wick et al., 2012).

As menores emissões médias de CO₂ foram associadas às áreas de Pinus (4,06 μmol m⁻² s⁻¹). Um fator contribuinte para isto pode estar relacionado aos baixos níveis de pH no solo (4,06), juntamente com os elevados níveis de acidez potencial desta área (74,80 mmol_c dm⁻³), que provavelmente afetaram a atividade biológica do solo, assim resultando em menor respiração radicular e, conseqüentemente, menos respiração do solo. Outra hipótese para explicar a relação entre o potencial de acidez

do solo e as maiores taxas de emissões associadas, relacionam-se à tendência de dissociação de CO₂ no interior dos solos, pois o ácido carbônico produzido pela dissolução de CO₂ na água é um importante agente acidificante nos sistemas naturais (Reuss e Johnson, 1986; Tossell, 2009).

5.3. Abordagem ANFIS para modelar a emissão de CO₂ do solo

Nesta abordagem, as variáveis de entrada foram compostas predominantemente por atributos físicos do solo, especialmente os teores de macro e microporosidade do solo. Tavares et al. (2015; 2016) indicaram que estas propriedades exibem comportamentos antagônicos. Diferentemente de solos com maiores teores de microporos, solos que apresentam maiores valores de macroporos oferecem uma rota menos tortuosa para as moléculas de CO₂, assim permitindo maior respiração do solo e transporte de CO₂ do solo para a atmosfera. Outros estudos também apontaram a importância das propriedades físicas dos solos, especialmente a microporosidade, macroporosidade (Wick et al., 2012) e a Ds, para explicar a emissão de CO₂, indicando a importância da porosidade do solo para a atividade microbiana e as trocas gasosas (Panosso et al., 2011; Teixeira et al., 2013; Moitinho et al., 2015).

Sabe-se que o grau de evolução da MOS, a atividade biológica do solo e a capacidade de produzir formas de nitrogênio assimiláveis podem ser avaliadas por meio da razão C/N e, dentre as propriedades químicas do solo, este atributo se mostrou eficiente para constituir os modelos empregados e melhorar o desempenho dos mesmos. Observou-se que as emissões de CO₂ na área de mata nativa (5,53 $\mu\text{mol m}^{-2} \text{s}^{-1}$) foram acompanhadas por baixas taxas de C/N (32,78) e diversos estudos relataram que proporções mais altas de C/N resultam em taxas mais baixas de degradação orgânica do carbono (Vesterdal et al., 2008), influenciando sobre a dinâmica da respiração do solo (Dorodnikov et al., 2011).

As capacidades preditivas dos modelos utilizando a abordagem de RNAs e ANFIS se mostraram eficazes e observa-se que diferentes atributos do solo se mostraram mais adequados que outros para serem usados como variáveis preditoras nos respectivos modelos. Isso realça a heterogeneidade das áreas estudadas, uma vez que os atributos possuem diferentes sensibilidades às áreas de aplicação deste trabalho. Pesquisas envolvendo a aplicação de RNA e ANFIS para modelar gases (NO_x, HC, CO, CO₂ e PM) (Gopalakrishnan et al., 2011) e fatores que afetam o

rendimento e custo da produção de biodiesel (Najafi et al., 2018) também mostraram bom desempenho preditivo.

Contudo, de acordo com os índices estatísticos e dependendo da generalização, alguns modelos são mais adequados para determinados conjuntos de dados do que outros (Fernandes et al., 2020). Assim, com base nos desempenhos estatísticos das RNAs (Tabela 4) e ANFIS (Tabela 5), este estudo apontou que as RNAs forneceram maior precisão que os modelos derivados do ANFIS. Como resultado, as RNAs forneceram desempenhos preditivos mais confiáveis, portanto devem ser indicadas para modelar a FCO₂ usando atributos físico-químicos do solo. Adicionalmente, outro importante fator na escolha dos melhores modelos é a semelhança de precisão entre as fases de calibração e validação, pois quanto maior a diferença entre os índices estatísticos durante estas fases, o modelo tende a ser menos generalista e mais propensos a erro (Sargent, 2013).

6. CONCLUSÕES

Este estudo mostrou que as RNAs e ANFIS forneceram modelos preditivos para explicar os padrões de variabilidade temporal da FCO₂ a partir de atributos físicos e químicos do solo. No entanto, o modelo global das GRNN apontou os maiores desempenhos preditivos para a modelagem da emissão de CO₂ do solo. Estes resultados sugerem alta aplicabilidade para modelar a FCO₂ e podem ser usados na elaboração de relatórios sobre fontes e sumidouros de GEE e assim traçar estratégias para apoiar a mitigação climática, especialmente em regiões tropicais, onde as emissões são frequentemente mais altas.

Além disso, este método pode ser uma alternativa eficaz e de baixo custo especialmente para países subdesenvolvidos, pois, com exceção da H_{LIFS.DS}, a maioria dos atributos que constitui os modelos são comumente determinados em análises de rotina do solo. Assim, não seria necessário o uso de equipamentos de alto custo para determinar a emissão de CO₂ do solo.

Notavelmente, o desempenho dos modelos individuais e global, para ambas as abordagens, foi altamente sensível à seleção das variáveis de entrada. Isso evidencia a heterogeneidade entre os ambientes estudados e sua influência na variabilidade dos atributos físicos e químicos, que por sua vez afetam diretamente o fluxo de CO₂ no solo.

7. REFERÊNCIAS

ALEXANDER, M. A.; SCOTT, J. D.; MAHONEY, K.; BARSUGLI, J. Greenhouse gas-induced changes in summer precipitation over Colorado in NARCCAP regional climate models. **Journal of Climate**, v. 26, n. 21, p. 8690-8697, 2013. <https://doi.org/10.1175/JCLI-D-13-00088.1>.

ALTIKAT, S.; Gulbe, A.; Kucukerdem, H. K.; Altikat, A. Applications of artificial neural networks and hybrid models for predicting CO₂ flux from soil to atmosphere. <https://doi.org/10.1007/s13762-020-02799-6>.

ANACHE, J. A.; FLANAGAN, D. C.; SRIVASTAVA, A.; WENDLAND, E. C. Land use and climate change impacts on runoff and soil erosion at the hillslope scale in the Brazilian Cerrado. **Science of The Total Environment**, v. 622, p. 140-151, 2018. <https://doi.org/10.1016/j.scitotenv.2017.11.257>.

BALL, B. C.; SMITH, K. A. Gas movement. In: Smith KA, Mullins CE (Eds.) *Soils analysis: physical methods*. New York: Marcel Dekerp, p. 511-549. 1991.

BATAGLIA, O. C.; FURLANI, A. M. C.; TEIXEIRA, P. F.; FURLANI, P. R.; GALLO, J. R. (1983). **Métodos de análises químicas de plantas**. Campinas, Instituto Agrônômico, p 48. (Boletim Técnico, 78).

BESALATPOUR, A. A.; AYOUBI, S.; HAJABBASI, M. A.; MOSADDEGHI, M.; SCHULIN, R. Estimating wet soil aggregate stability from easily available properties in a highly mountainous watershed, **CATENA**, v. 111, p. 72-79, 2013. <https://doi.org/10.1016/j.catena.2013.07.001>

BRADY, N. C.; WEIL, R. R. **Elementos da natureza e propriedades dos solos**. (3rd edn.), Bookman, Porto Alegre (2013), p. 704.

BOATENG, E. B.; PILLAY, M.; DAVIS, P. Predicting the level of safety performance using an artificial neural network. In: **International Conference on Human Systems Engineering and Design: Future Trends and Applications**. Springer, Cham, 2018. p. 705-710. https://doi.org/10.1007/978-3-030-02053-8_107.

BORDONAL, R. O.; LAL, R.; RONQUIM, C. C.; DE FIGUEIREDO, E. B.; CARVALHO, J. L. N.; MALDONADO JR, W.; MILORI, D. M. B. P.; LA SCALA JR, N. Changes in quantity and quality of soil carbon due to the land-use conversion to sugarcane (*Saccharum officinarum*) plantation in southern Brazil. **Agriculture, ecosystems & environment**, v. 240, p. 54-65, 2017. <https://doi.org/10.1016/j.agee.2017.02.016>.

BRASIL. Third National Communication of Brazil to the United Nations Framework **Convention on Climate Change Brasilia**. [S.l.: s.n.], 2016. [Disponível em: <https://unfccc.int/resource/docs/natc/branc3v3.pdf>. Acesso em: 25. agosto. 2019.

CAMARGO, A. P.; SENTELHAS, P. C. Avaliação do desempenho de diferentes métodos de estimativa da evapotranspiração potencial no Estado de São Paulo, Brasil. **Revista Brasileira de agrometeorologia**, v. 5, n. 1, p. 89-97, 1997.

CARVALHO, J. L. N.; CERRI, C. E. P.; FEIGL, B. J.; PÍCCOLO, M. D. C.; GODINHO, V. P.; CERRI, C. C. Carbon sequestration in agricultural soils in the Cerrado region of the Brazilian Amazon. **Soil and Tillage Research**, v. 103, n. 2, p. 342-349, 2009. <https://doi.org/10.1016/j.still.2008.10.022>.

CHEN, Z.; YE, X.; HUANG, P. Estimating carbon dioxide (CO₂) emissions from reservoirs using artificial neural networks. **Water**, v. 10, n. 1, p. 26, 2018. <https://doi.org/10.3390/w10010026>.

CRUZ, C. D.; REGAZZI, A. J. (1994). Modelos Biométricos Aplicados Ao Melhoramento Genético. Universidade Federal de Viçosa, Viçosa.

DE CARVALHO, M. A. C.; PANOSSO, A. R.; TEIXEIRA, E. E. R.; ARAÚJO, E. G.; BRANCAGLIONI, V. A.; DALLACORT, R. Multivariate approach of soil attributes on the characterization of land use in the southern Brazilian Amazon. **Soil and Tillage Research**, v. 184, p. 207-215, 2018. <https://doi.org/10.1016/j.still.2018.08.004>.

DE FIGUEIREDO, E. B.; JAYASUNDARA, S.; DE OLIVEIRA BORDONAL, R.; BERCHIELLI, T. T.; REIS, R. A.; WAGNER-RIDDLE, C.; LA SCALA JR, N. Greenhouse gas balance and carbon footprint of beef cattle in three contrasting pasture-management systems in Brazil. **Journal of cleaner production**, v. 142, p. 420-431, 2017.

DE FIGUEIREDO, E. B. D.; PANOSSO, A. R.; BORDONAL, R. D. O.; TEIXEIRA, D. D. B.; BERCHIELLI, T. T.; LA SCALA JR, N. Soil CO₂-C emissions and correlations with soil properties in degraded and managed pastures in Southern Brazil. **Land Degradation & Development**, v. 28, n. 4, p. 1263-1273, 2017. <https://doi.org/10.1002/ldr.2524>.

DE MIRANDA, S. D. C.; BUSTAMANTE, M.; PALACE, M.; HAGEN, S.; KELLER, M.; FERREIRA, L. G. Regional variations in biomass distribution in Brazilian savanna woodland. **Biotropica**, v. 46, n. 2, p. 125-138, 2014. <https://doi.org/10.1111/btp.12095>.

DE OLIVEIRA, V. A.; DE MELLO, C. R.; BESKOW, S.; VIOLA, M. R.; SRINIVASAN, R. Modeling the effects of climate change on hydrology and sediment load in a headwater basin in the Brazilian Cerrado biome. **Ecological Engineering**, v. 133, p. 20-31, 2019. <https://doi.org/10.1016/j.ecoleng.2019.04.021>

DORODNIKOV, M.; KUZYAKOV, Y.; FANGMEIER, A.; WIESENBERG, G. L. B. C and N in soil organic matter density fractions under elevated atmospheric CO₂: Turnover vs. Stabilization. **Soil Biology and Biochemistry**, Oxford, v. 43, n. 3, p. 579–589, 2011. <https://doi.org/10.1016/j.soilbio.2010.11.026>.

DOS SANTOS MACIEL, L.; BALLINI, R.; DA SILVEIRA, R. L. F. Apreçamento de opções sobre taxa de câmbio R \$/US \$ negociadas no Brasil: uma comparação entre os modelos Black e redes neurais artificiais. **Revista de Administração**, v. 47, n. 1, p. 96-111, 2012. <https://doi.org/10.5700/rausp1028>.

EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária. **Manual de métodos de análise de solo**. 3. ed. Rio de Janeiro: Embrapa Solos. 2017.

FARHATE, C. V. V.; DE SOUZA, Z. M.; DE MEDEIROS OLIVEIRA, S. R.; 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, 2018. <https://doi.org/10.1371/journal.pone.0193537>

FERNANDES, M. M. H.; COELHO, A. P.; DA SILVA, M. F.; BERTONHA, R. S.; DE QUEIROZ, R. F.; FURLANI, C. E. A.; FERNANDES, C. Estimation of soil penetration resistance with standardized moisture using modeling by artificial neural networks. **Catena**, v. 189, p. 104505, 2020. <https://doi.org/10.1016/j.catena.2020.104505>.

FERNANDES, M. M. H.; COELHO, A. P.; FERNANDES, C.; DA SILVA, M. F.; MARTA, C. C. D Estimation of soil organic matter content by modeling with artificial neural networks. **Geoderma**, v. 350, p. 46-51, 2019. <https://doi.org/10.1016/j.geoderma.2019.04.044>.

FLANAGAN, L. B.; SHARP, E. J.; LETTS, M. G. Response of plant biomass and soil respiration to experimental warming and precipitation manipulation in a Northern Great Plains grassland. **Agricultural and Forest Meteorology**, v. 173, p. 40-52, 2013. <https://doi.org/10.1016/j.agrformet.2013.01.002>.

FREITAS, L. P. S.; LOPES, M. L. M.; CARVALHO, L. B.; PANOSSO, A. R.; LA SCALA JR, 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. **Environmental monitoring and assessment**, v. 190, n. 12, p. 741, 2018. <https://doi.org/10.1007/s10661-018-7118-0>.

FUNK, R.; PASCUAL, U.; JOOSTEN, H.; DUFFY, C.; PAN, G.; LA SCALA, N.; GOTTSCHALK, P.; BANWART, S. A.; BATJES, N.; CAI, Z.; SIX, J.; NOELLEMAYER, E. From potential to implementation: an innovation framework to realize the benefits of soil carbon. Soil carbon: **Science, management and policy for multiple benefits**, p. 47-59, 2015.

GHIASI, M. M.; ARABLOO, M.; MOHAMMADI, A. H.; BARGHI, T. Application of ANFIS soft computing technique in modeling the CO₂ capture with MEA, DEA, and TEA aqueous solutions. **International Journal of Greenhouse Gas Control**, v. 49, p. 47-54, 2016. <https://doi.org/10.1016/j.ijggc.2016.02.015>.

HAYKIN, S. **Redes neurais: princípios e prática**. (2 ed.), Bookman, Porto Alegre (2001).

IPCC - Intergovernmental Panel on Climate Change. Climate change 2007: **The physical science basis. Summary for policymakers**, Geneva, Suíça. Disponível em:<http://www.ipcc.ch/publications_and_data/publications_ipcc_fourth_assessment_report_wg1_report_the_physical_science_basis.htm>. Acesso em 29 set. 2007.

IPCC, 2014. **Intergovernmental panel on climate change**. In: Smith, P., Bustamante M. (Eds.), Working Group III Contribution to the IPCC 5th Assessment Report “Climate Change 2014: Mitigation of Climate Change” that was Accepted but Not Approved in Detail by the 12th Session of Working Group III and the 39th Session of the IPCC on April 12 2014 in Berlin, Germany (2014).

JANG, J. S. R. ANFIS: adaptive-network-based fuzzy inference system. **IEEE transactions on systems, man, and cybernetics**. <http://doi.org/10.1109/21.256541>
Johnson, D. W. Acid deposition and the acidification of soils and water. *Ecological studies (USA)*. v. 59., 1986.

KAAB, A.; SHARIFI, M.; MOBLI, H.; NABAVI-PELESARAEI, A.; CHAU, K. W. Combined life cycle assessment and artificial intelligence for prediction of output energy and environmental impacts of sugarcane production. **Science of the Total Environment**, v. 664, p. 1005-1019, 2019. <https://doi.org/10.1016/j.scitotenv.2019.02.004>

KANG, S., KIM, S., OH, S., LEE, D. Predicting spatial and temporal patterns of soil temperature based on topography, surface cover, and air temperature. **Forest Ecology and Management**, v. 136, n. 1-3, p. 173-184, 2000. [https://doi.org/10.1016/S0378-1127\(99\)00290-X](https://doi.org/10.1016/S0378-1127(99)00290-X)

KASHI, H.; EMAMGHOLIZADEH, S.; GHORBANI, H. Estimation of soil infiltration and cation exchange capacity based on multiple regression, ANN (RBF, MLP), and ANFIS models. **Communications in soil science and plant analysis**, v. 45, n. 9, p. 1195-1213, 2014. <https://doi.org/10.1080/00103624.2013.874029>.

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. <https://doi.org/10.1108/IJCCSM-01-2019-0001>

LAMPARTER, A.; BACHMANN, J.; GOEBEL, M. O.; WOCHE, S. K. Carbon mineralization in soil: Impact of wetting–drying, aggregation and water repellency. **Geoderma**, v. 150, n. 3-4, p. 324-333, 2009. <https://doi.org/10.1016/j.geoderma.2009.02.014>.

LI, D.; NIU, S.; LUO, Y. Global patterns of the dynamics of soil carbon and nitrogen stocks following afforestation: a meta-analysis. **New Phytologist**, v. 195, n. 1, p. 172-181, 2012. <https://doi.org/10.1111/j.1469-8137.2012.04150.x>

MAIA, S. M. F.; CARVALHO, J. L. N.; CERRI, C. E. P.; LAL, R.; BERNOUX, M.; GALDOS, M. V.; CERRI, C. C. Contrasting approaches for estimating soil carbon changes in Amazon and Cerrado biomes. **Soil and Tillage Research**, v. 133, p. 75-84, 2013. <https://doi.org/10.1016/j.still.2013.06.002>.

MAIER, H. R.; DANDY, G.C. Neural network based modelling of environmental variables: a systematic approach. **Mathematical and Computer Modelling**, v. 33, n. 6-7, p. 669-682, 2001. [https://doi.org/10.1016/S0895-7177\(00\)00271-5](https://doi.org/10.1016/S0895-7177(00)00271-5).

MILORI, D. M. B. P.; GALETI, H. V. A.; MARTIN-NETO, L.; DIECKOW, J.; GONZÁLEZ-PÉREZ, M.; BAYER, C.; SALTON, J. Organic matter study of whole soil samples using laser-induced fluorescence spectroscopy. *Soil Science Society of America Journal*. v. 70, n. 1, p. 57-63, 2006. <https://doi.org/10.2136/sssaj2004.0270>.

MOITINHO, M. R.; PADOVAN, M. P.; PANOSSO, A. R.; TEIXEIRA, D. D. B.; FERRAUDO, A. S.; LA SCALA JR, N. On the spatial and temporal dependence of CO₂ emission on soil properties in sugarcane (*Saccharum spp.*) production. **Soil and Tillage Research**, v. 148, p. 127-132, 2015. <https://doi.org/10.1016/j.still.2014.12.012>.

NAJAFI, B.; ARDABILI, S. F.; SHAMSHIRBAND, S.; CHAU, K. W.; RABCZUK, T. Application of ANNs, ANFIS and RSM to estimating and optimizing the parameters that affect the yield and cost of biodiesel production. **Engineering Applications of Computational Fluid Mechanics**, v. 12, n. 1, p. 611-624, 2018. <https://doi.org/10.1080/19942060.2018.1502688>.

NOOJIPADY, P.; MORTON, C. D.; MACEDO, N. M.; VICTORIA, C. D.; HUANG, C.; GIBBS, K. H.; BOLFE, L. E. Forest carbon emissions from cropland expansion in the Brazilian Cerrado biome. **Environmental Research Letters, Bristol**, v. 12, n. 2, p. 025004, 2017. <https://doi.org/10.1088/1748-9326/aa5986>

NOORI, R.; HOSHYARIPOUR, G.; ASHRAFI, K.; ARAABI, B. N. Uncertainty analysis of developed ANN and ANFIS models in prediction of carbon monoxide daily concentration. **Atmospheric Environment**, v. 44, n. 4, p. 476-482, 2010. <https://doi.org/10.1016/j.atmosenv.2009.11.005>.

OLSSON, L. H.; BARBOSA, S.; BHADWAL, A.; COWIE, K.; DELUSCA, D.; FLORES-RENTERIA, K.; HERMANS, E.; JOBBAGY, W.; KURZ, D.; LI, D.J.; SONWA, L. Stringer, 2019: Land Degradation. In: *Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems* [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi,

M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, J. Malley, (eds.)]. In press.

PANDEY, A.; MISHRA, A. Application of artificial neural networks in yield prediction of potato crop. **Russian Agricultural Sciences**, v. 43, n. 3, p. 266-272, 2017. <https://doi.org/10.3103/S1068367417030028>.

PANOSSO, A. R.; MARQUES JR, J.; MILORI, D. M. B. P.; FERRAUDO, A. S.; BARBIERI, D. M.; PEREIRA, G. T.; LA SCALA JR, N. Soil CO₂ emission and its relation to soil properties in sugarcane areas under Slash-and-burn and Green harvest. **Soil and Tillage Research**, v. 111, n. 2, p. 190-196, 2011. <https://doi.org/10.1016/j.still.2010.10.002>.

PANOSSO, A. R.; MARQUES JR, J.; PEREIRA, G. T.; LA SCALA JR, N. Spatial and temporal variability of soil CO₂ emission in a sugarcane area under green and slash-and-burn managements. **Soil and Tillage Research**, v. 105, n. 2, p. 275-282, 2009. <https://doi.org/10.1016/j.still.2009.09.008>

PULROLNIK, K.; BARROS, N. F. D.; SILVA, I. R.; NOVAIS, R. F.; BRANDANI, C. B. Estoques de carbono e nitrogênio em frações lábeis e estáveis da matéria orgânica de solos sob eucalipto, pastagem e cerrado no Vale do Jequitinhonha-MG. **Revista Brasileira de Ciência do Solo**, v. 33, n. 5, p. 1125-1136, 2009. <http://dx.doi.org/10.1590/S0100-06832009000500006>.

RAIJ, J. C.; ANDRADE, H.; CANTARELLA, J. A.; QUAGGIO. **Análise química para avaliação da fertilidade de solos tropicais** (1 ed.), Instituto Agrônomo de Campinas, Campinas (2001).

RIZA, L. S.; BERGMEIR, C. N.; HERRERA TRIGUERO, F.; BENÍTEZ SÁNCHEZ, J. M. frbs: Fuzzy rule-based systems for classification and regression in R. **Journal of statistical software**. v. 65, n. 6, 2015. <https://doi.org/10.18637/jss.v065.i06>.

SARGENT, R. G. Verification and validation of simulation models. **Journal of simulation**, v. 7, n. 1, p. 12-24, 2013. <https://doi.org/10.1057/jos.2012.20>.

SCHIASSI, M. C. E. V.; DE SOUZA, V. R.; LAGO, A. M. T.; CAMPOS, L. G.; QUEIROZ, F. Fruits from the Brazilian Cerrado region: Physico-chemical characterization, bioactive compounds, antioxidant activities, and sensory evaluation. **Food chemistry**, v. 245, p. 305-311, 2018. <https://doi.org/10.1016/j.foodchem.2017.10.104>

SILVA, E. N. S. D.; MONTANARI, R.; PANOSSO, A. R.; CORREA, A. R.; TOMAZ, P. K.; FERRAUDO, A. S. Variabilidade de atributos físicos e químicos do solo e produção de feijoeiro cultivado em sistema de cultivo mínimo com irrigação. **Revista Brasileira de Ciência do Solo**, v. 39, n. 2, p. 598-607, 2015. <http://doi.org/10.1590/01000683rbc20140429>

SILVA, B. O.; MOITINHO, M. R., DE ARAUJO SANTOS, G. A., TEIXEIRA, D. D. B., FERNANDES, C.; LA SCALA JR, N. Soil CO₂ emission and short-term soil pore class distribution after tillage operations. **Soil and Tillage Research**, v. 186, p. 224-232, 2019. <https://doi.org/10.1016/j.still.2018.10.019>.

SIX, J.; FREY, S. D.; THIET, R. K.; BATTEN, K. M. Bacterial and fungal contributions to carbon sequestration in agroecosystems. **Soil Science Society of America Journal**, v. 70, n. 2, p. 555-569, 2006. <https://doi.org/10.2136/sssaj2004.0347>.

SOARES-FILHO, B.; RAJÃO, R.; MACEDO, M.; CARNEIRO, A.; COSTA, W.; COE, M.; RODRIGUES, H.; ALENCAR, A. Cracking Brazil's forest code. **Science**, v. 344, n. 6182, p. 363-364, 2014. <http://doi:/10.1126/science.1246663>.

SOUSA, R. F. D.; BRASIL, E. P. F.; FIGUEIREDO, C. C. D.; LEANDRO, W. M. Soil organic matter fractions in preserved and disturbed wetlands of the cerrado biome. **Revista Brasileira de Ciência do Solo**, v. 39, n. 1, p. 222-231, 2015. <http://dx.doi.org/10.1590/01000683rbc20150048>

SOUZA, L. H. C.; MATOS, E. D. S.; MAGALHÃES, C. A. D. S.; DE LA TORRE, É. R.; LAMAS, F. M.; LAL, R. Soil carbon and nitrogen stocks and physical properties under no-till and conventional tillage cotton-based systems in the Brazilian Cerrado. **Land Degradation & Development**, v. 29, n. 10, p. 3405-3412, 2018. <https://doi-org.ez87.periodicos.capes.gov.br/10.1002/ldr.3105>.

TAVANTI, R. F. R.; MONTANARI, R.; PANOSSO, A. R.; LA SCALA JR, N.; NETO, M. C.; DA SILVA FREDDI, O.; GONZÁLEZ, A. P.; DE CARVALHO, M. A. C.; SOARES, M. B.; TAVANTI, T. R.; GALINDO, F. S. What is the impact of pasture reform on organic carbon compartments and CO₂ emissions in the Brazilian Cerrado? **CATENA**, v. 194, p. 104702, 2020. <https://doi.org/10.1016/j.catena.2020.104702>.

TAVARES, R. L. M.; OLIVEIRA, S. R. D. M.; BARROS, F. M. M. D.; FARHATE, C. V. V.; SOUZA, Z. M. D.; JUNIOR, S.; LA SCALA JR, N. Prediction of soil CO₂ flux in sugarcane management systems using the Random Forest approach. **Scientia Agricola**, v. 75, n. 4, p. 281-287, 2018. <http://dx.doi.org/10.1590/1678-992x-2017-0095>.

TAVARES, R. L. M.; SIQUEIRA, D. S.; PANOSSO, A. R.; CASTIONI, G. A. F.; SOUZA, Z.M.; LA SCALA JR. N. Soil management of sugarcane fields affecting CO₂ fluxes. **Scientia Agricola**, v. 73, n. 6, p. 543-551, 2016. <https://doi.org/10.1590/0103-9016-2015-0174>.

TAVARES, R. L. M.; SOUZA, Z. M.; SIQUEIRA, D. S.; LA SCALA JR. N.; PANOSSO, A.R., Campos, M.C.C. Soil CO₂ emission in sugarcane management systems. **Acta Agriculturae Scandinavica, Section B—Soil & Plant Science**, v. 65, n. 8, p. 755-762, 2015. <https://doi.org/10.1080/09064710.2015.1061048>.

TEDESCO, M. J.; GIANELLO, C.; BISSANI, C. A.; BOHNEN, H.; WOLKWEISS, S. J. (1995) Análises de solo, plantas e outros materiais, 2nd edn. Universidade Federal do Rio Grande do Sul, Porto Alegre

TEIXEIRA, D. D. B.; BICALHO, E. D. S.; PANOSSO, A. R.; CERRI, C. E. P.; PEREIRA, G. T.; JÚNIOR, S.; LA SCALA JR, N. Spatial variability of soil CO₂ emission in a sugarcane area characterized by secondary information. **Scientia Agricola**, v. 70, n. 3, p. 195-203, 2013. <http://dx.doi.org/10.1590/S0103-90162013000300008>.

TEIXEIRA, D. D. B.; BICALHO, E. D. S.; PANOSSO, A. R.; PERILLO, L. I.; IAMAGUTI, J. L.; PEREIRA, G. T.; LA SCALA JR, N. Uncertainties in the prediction of spatial variability of soil CO₂ emissions and related properties. **Revista Brasileira de Ciência do Solo**, v. 36, n. 5, p. 1466-1475, 2012. <http://dx.doi.org/10.1590/S0100-06832012000500010>

TOSSELL, J. A. H₂CO₃ (s): A new candidate for CO₂ capture and sequestration. **Environmental science & technology**, v. 43, n. 7, p. 2575-2580, 2009. <https://doi.org/10.1021/es802393s>.

UNFCCC, 2013. Views on Land Use, Land-use Change and Forestry Issues Referred to in Decision 2/CMP.7, Paragraphs 5e7. Submissions from Parties and Admitted Observer Organizations 12e18 (SBSTA, UNFCCC, 2013). Available at:<http://go.nature.com/hLAtTN> (accessed 15.06.20.).

VESTERDAL, L.; SCHMIDT, I. K.; CALLESEN, I.; NILSSON, L. O.; GUNDERSEN, P. Carbon and nitrogen in forest floor and mineral soil under six common European tree species. **Forest Ecology and Management**, v. 255, n. 1, p. 35-48, 2008. <https://doi.org/10.1016/j.foreco.2007.08.015>.

WALLENSTEIN, M.; ALLISON, S. D.; ERNAKOVICH, J.; STEINWEG, J. M.; SINSABAUGH, R. Controls on the temperature sensitivity of soil enzymes: a key driver of in situ enzyme activity rates. In: **Soil enzymology**. Springer, Berlin, Heidelberg, 2010. p. 245-258. https://doi.org/10.1007/978-3-642-14225-3_13.

WICK, A. F.; PHILLIPS, R. L.; LIEBIG, M. A.; WEST, M.; DANIELS, W. L. Linkages between soil micro-site properties and CO₂ and N₂O emissions during a simulated thaw for a northern prairie Mollisol. **Soil Biology and Biochemistry**, v. 50, p. 118-125, 2012. <https://doi.org/10.1016/j.soilbio.2012.03.010>.

WILLMOTT, C. J. 1981. On the validation of models. *Phys. Geogr.* 2, 184–194.

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. <https://doi.org/10.1016/j.eswa.2010.11.027>.

ZENDEHBOUDI, S.; REZAEI, N.; LOHI, A. Applications of hybrid models in chemical, petroleum, and energy systems: A systematic review. **Applied energy**, v. 228, p. 2539-2566, 2018. <https://doi.org/10.1016/j.apenergy.2018.06.051>.