

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
CÂMPUS DE CHAPADÃO DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA

PAULO ROBERTO NUNES VIANA

**ANÁLISE DOS MACRONUTRIENTES DO SOLO SOB ENFOQUE
MULTIESPECTRAL.**

CHAPADÃO DO SUL – MS

2023

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
CÂMPUS DE CHAPADÃO DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA

PAULO ROBERTO NUNES VIANA

**ANÁLISE DOS MACRONUTRIENTES DO SOLO SOB ENFOQUE
MULTIESPECTRAL.**

Orientador: Prof (a). Dr. Rafael Felipe Ratke

Dissertação apresentada à Universidade Federal de Mato Grosso do Sul, como requisito para obtenção do título de Mestre em Agronomia, área de concentração: Produção Vegetal.

CHAPADÃO DO SUL – MS

2023



Serviço Público Federal
Ministério da Educação
Fundação Universidade Federal de Mato Grosso do Sul



PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA

CERTIFICADO DE APROVAÇÃO

DISCENTE: Paulo Roberto Nunes Viana

ORIENTADOR: Dr. Rafael Felipe Ratke

TÍTULO: Análise dos macronutrientes do solo sob enfoque multiespectral

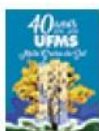
AVALIADORES:

Prof. Dr. Rafael Felipe Ratke

Profa. Dra. Larissa Pereira Ribeiro Teodoro

Prof. Dr. Alan Mario Zuffo

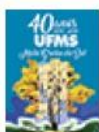
Chapadão do Sul, 17 de fevereiro de 2023.



Documento assinado eletronicamente por **Rafael Felipe Ratke, Professor do Magisterio Superior**, em 17/02/2023, às 10:15, conforme horário oficial de Mato Grosso do Sul, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Larissa Pereira Ribeiro, Professora do Magistério Superior**, em 17/02/2023, às 10:19, conforme horário oficial de Mato Grosso do Sul, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Alan Mario Zuffo, Usuário Externo**, em 17/02/2023, às 10:30, conforme horário oficial de Mato Grosso do Sul, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

AGRADECIMENTOS

Sem Deus nada seria.

Agradeço a todos os colaboradores que de uma vez ou outra me ajudaram na conclusão deste trabalho tão impressionante, são eles:

Meus pais, filho e família pelo apoio em momentos de dificuldade, seja financeira, seja emocional e suas motivações quando tudo parecia difícil.

Minha irmã pelo apoio em prosseguir quando tudo estava obscuro e sua atenta dedicação em corrigir os erros de meu trabalho me levando a outro nível de qualidade.

Aos amigos e colegas de faculdade que sempre se mostraram ajudantes quando precisei, seja nas áreas de confecção de figuras, tabelas, informática, conversas e ideias.

Ao conjunto de aportes recebidos da instituição UFMS para realização de disciplinas e construção de conhecimentos.

Em especial ao orientador professor Rafael pela dedicação e paciência em passar seus conhecimentos a um aluno que não é formado em sua área de atuação.

Ao professor Carlos amigo e ilustre conhecedor dos temas computacionais em abrir espaço para que eu observasse a amplitude do tema aberto em análise.

A professora Larissa pelas aulas de *Machine Learning* que sem elas não me permitiriam confeccionar as páginas escritas e nem pesquisar o tema proposto.

E todos aqueles que não foram citados, mas que direta ou indiretamente, contribuíram para este trabalho saísse do plano de projetos para realidade.

RESUMO

Solos analisados levam a discussões em meios acadêmicos devido a sua possível contaminação ao meio ambiente e ausência de sustentabilidade. Para sanar esses entraves estudaram-se os resultados obtidos pelos meios tradicionais de análise de macronutrientes do solo pela quimiometria e as comparou pela análise do solo através do uso de sensor multiespectral (550 a 800 nm), além disso, também fez a classificação dos teores de macronutrientes do solo por inteligência computacional. No capítulo 1 os estudos de correspondência preditiva dos resultados foram realizados usando a inteligência computacional e os algoritmos de decisão Random Forest, M5P, Regressão Linear, Multilayer Perceptron, Reptree e Random Tree. No capítulo 2 foi acrescentado os algoritmos de decisão J48 e Regressão Logística, retirados o M5P e Regressão Linear, além de predizer a classificação de teores quantitativos comparados em tabela de análise dos solos em baixo,adequado,médio,alto e sua granulometria em argiloso, arenoso e médio e medir a harmonia dos dados apresentados verificando a precisão dos experimentos. Foi possível observar que para análise do enxofre obteve-se um coeficiente de correlação positivo de 66.82% e erro médio absoluto de 8.22%, porém para outros macronutrientes a precisão de predição dos resultados obtidos por sensor multiespectral comparados a análise quimiométrica obtiveram valores preditivos pouco correlacionados. Também se observou que para avaliação de teores de fósforo no solo com a adição complementar de variáveis de entrada físicas de areia, silte e argila nos classificadores possibilitou precisão de 90% usando o classificador Random Forest ao passo que outros classificadores apresentaram diferenças estatísticas relevantes mostrando importar a técnica utilizada tanto para predizer quanto para classificar.

Palavras-chave: Análises. Algoritmos. Sustentabilidade. Solo.

ABSTRACT

Analyzed soils lead to discussions in academic circles due to their possible contamination to the environment and lack of sustainability. To remedy these obstacles, the results obtained by traditional means of soil macronutrient analysis by chemometrics were studied and compared by soil analysis using a multispectral sensor (550 to 800 nm), in addition, the levels were classified of soil macronutrients by computational intelligence. In chapter 1, studies of predictive correspondence of results were performed using computational intelligence and decision algorithms Random Forest, M5P, Linear Regression, Multilayer Perceptron, Reptree and Random Tree. In chapter 2, the decision algorithms J48 and Logistic Regression were added, removing the M5P and Linear Regression, in addition to predicting the classification of quantitative grades compared in a soil analysis table in low, adequate, medium, high and its granulometry in clayey, sandy and medium and measure the harmony of the presented data verifying the precision of the experiments. It was possible to observe that for the analysis of sulfur a positive correlation coefficient of 66.82% and an absolute mean error of 8.22% were obtained, however for other macronutrients the prediction accuracy of the results obtained by the multispectral sensor compared to the chemometric analysis obtained predictive values little correlated . It was also observed that for the evaluation of phosphorus levels in the soil with the complementary addition of physical input variables of sand, silt and clay in the classifiers, it allowed an accuracy of 90% using the Random Forest classifier, while other classifiers presented relevant statistical differences showing import the technique used for both prediction and classification.

Keywords: Analytics. Algorithms. Sustainability. Soil.

LISTA DE FIGURAS

CAPÍTULO 1

FIGURA 1. Correlação em porcentagem(%) entre valores observados e preditos usando o classificador Random Forest em modo padrão(default).....16

FIGURA 2. Índices de MAE (Erro Absoluto Médio), variando de 0 a 100 %, do classificador de *machine learning* Random Forest, com nutrientes comparados.....16

FIGURA 3. Boxplots para coeficiente de correlação de Pearson(r) e erro absoluto médio (MAE) para diferentes modelos de *machine learning* relacionado ao Enxofre 18

FIGURA 4. Boxplots para coeficiente de correlação de Pearson(r) e erro absoluto médio (MAE) para diferentes modelos de *machine learning* relacionado ao Magnésio 19

FIGURA 5. Boxplots para coeficiente de correlação de Pearson(r) e erro absoluto médio (MAE) para diferentes modelos de Machine Learning relacionado ao Potássio20

FIGURA 6. Boxplots para coeficiente de correlação de Pearson(r) e erro absoluto médio (MAE) para diferentes modelos de *machine learning* relacionado ao Cálcio.21

FIGURA 7. Comparação de cada macronutriente, variando matematicamente de 0 a 100 %, com o modelo de Regressão Linear em mod padrão(default)..... 22

CAPÍTULO 2

FIGURA 8. Análise em tabela comparativa do conjunto de 912 amostras de Cálcio em

porcentagem, variando de 0 a 100%33

FIGURA 9. Boxplots para classificações corretas (CC) e F-measure para diferentes modelos de *machine learning* relacionado ao Cálcio.34

FIGURA 10. Análise em tabela comparativa do conjunto de 358 amostras de Enxofre em porcentagem, variando de 0 a 100%35

FIGURA 11. Boxplots para classificações corretas (CC) e F-measure para diferentes modelos de *machine learning* relacionado ao Enxofre..... 36

FIGURA 12. Análise em tabela comparativa do conjunto de 688 amostras de Magnésio em porcentagem, variando de 0 a 100%37

FIGURA 13. Boxplots para classificações corretas (CC) e F-measure para diferentes modelos de *machine learning* relacionado ao Magnésio. 38

FIGURA 14. Análise em tabela comparativa do conjunto de 690 amostras de Potássio em porcentagem, variando de 0 a 100%39

FIGURA 15. Boxplots para classificações corretas (CC) e F-measure para diferentes modelos de Machine Learning relacionado ao Potássio. 40

FIGURA 16. Análise em tabela comparativa do conjunto de 105 amostras de Fósforo e sua relação com granulometria em porcentagem, variando de 0 a 100%..... 41

FIGURA 17. Boxplots para classificações corretas (CC) e F-measure para diferentes modelos de *machine learning* relacionado ao solo e sua granulometria.42

FIGURA 18. Análise em tabela comparativa do conjunto de 78 amostras de Fósforo em porcentagem, variando de 0 a 100%43

FIGURA 19. Boxplots para classificações corretas (CC) e F-measure para diferentes modelos de *machine learning* relacionado ao Fósforo obtido pelo método resina..... 44

SUMÁRIO

CAPÍTULO 1

PREDIÇÃO DOS TEORES DE NUTRIENTES NO SOLO UTILIZANDO SENSOR MULTISPECTRAL E APRENDIZAGEM DE MÁQUINA	10
RESUMO	10
ABSTRACT.....	11
INTRODUÇÃO	11
MATERIAL E MÉTODOS.....	13
RESULTADOS E DISCUSSÃO.	15
CONCLUSÕES	23
REFERÊNCIAS BIBLIOGRÁFICAS	23

CAPÍTULO 2

CLASSIFICAÇÃO DOS TEORES DE MACRONUTRIENTES E GRANULOMETRIA DO SOLO UTILIZANDO SENSOR MULTISPECTRAL E APRENDIZAGEM DE MÁQUINA.....	28
RESUMO	28
ABSTRACT.....	28
INTRODUÇÃO.....	29
MATERIAL E MÉTODOS.....	31
RESULTADOS E DISCUSSÃO.	33
CONCLUSÕES	45
REFERÊNCIAS BIBLIOGRÁFICAS	45

CAPÍTULO 1

PREDIÇÃO DOS TEORES DE NUTRIENTES NO SOLO UTILIZANDO SENSOR MULTISPECTRAL E APRENDIZAGEM DE MÁQUINA

RESUMO: As análises dos solos provocam discussões em meios acadêmicos devido a sua possível contaminação ao meio ambiente, ausência de sustentabilidade e a forma pela qual é executado as análises em laboratórios. Para sanar esses entraves estudaram-se os resultados obtidos pelos meios tradicionais de análise de macronutrientes do solo pela quimiometria e as comparou pela análise do solo através do uso de sensor multiespectral que vai do verde a borda do vermelho (550 a 800 nm). Para prever os teores utilizou-se os algoritmos de decisão Random Forest, M5P, Regressão Linear, Multilayer Percetron, Random Tree e RepTree e técnicas de inteligência computacional para obtenção de gráficos estatísticos. Após isso foi possível observar os resultados e que para análise do enxofre obteve-se um coeficiente de correlação de 66.82% e erro médio absoluto de 8.22%. Para outros macronutrientes houve resultados positivamente correlacionados, necessitando, contudo, de maiores estudos para definir a melhor banda espectral preditiva para cada macronutriente.

Palavras-chave: Análises. Algoritmos. Sustentabilidade. Solo.

ABSTRACT: Soil analyzes provoke discussions in academic circles due to their possible contamination of the environment, lack of sustainability and the way in which the analyzes are carried out in laboratories. To remedy these obstacles, the results obtained by traditional means of soil macronutrient analysis by chemometrics were studied and compared by soil analysis using a multispectral sensor ranging from green to red border (550 to 800 nm). To predict grades, decision algorithms Random Forest, M5P, Linear Regression, Multilayer Perceptron, Random Tree and RepTree and computational intelligence techniques were used to obtain statistical graphs. After that, it was possible to observe the results and that for the analysis of sulfur a correlation coefficient of 66.82% and an absolute mean error of 8.22% were obtained. For other macronutrients there were positively correlated results, requiring, however, further studies to define the best predictive spectral band for each macronutrient.

Keywords: Analytics. Algorithms. Sustainability. Soil.

INTRODUÇÃO

A análise geral do solo engloba o conjunto de processos químicos e indica o quanto aquele solo pode fornecer de determinado nutriente conforme orienta Raij (2001). Para tanto, é preciso a coleta de amostra do solo, a avaliação química, física, biológica e o estudo constitucional do histórico da área para avaliar a nutrição das culturas utilizadas. Com isso o resultado é a quantificação dos elementos necessários para o solo em estudo e a necessidade de adubação das culturas levando em conta sua fertilidade.

O método utilizado pela maioria dos cientistas agrônômicos é a extração de uma solução química conjugada com extratores que simulam a absorção de nutrientes pelas plantas. Ocorre que esse meio utilizado para análises tradicionais do solo causa alguns entraves ambientais devido ao descarte de resíduos gerados na análise. Segundo Guven e Akinci (2011) além de algumas análises serem contrárias à química verde, ou seja, danosas ao técnico de laboratório e ao meio circulante do resíduo químico, outras análises químicas dos solos têm alto custo dos reagentes químicos e demandam tempo.

A alternativa para análise dos solos advém do uso de novas tecnologias, a exemplo da espectroscopia de refletância com modelos de predição para variáveis químicas e físicas dos solos. Após análises espectrais, utilizam-se o PCA (Análise dos Componentes Principais), métodos de calibração e previsão (Regressão Linear múltipla, Regressão em Componentes Principais-PCR, Mínimos Quadrados Principais-PLS) além de métodos de inteligência artificial como redes neurais para correlação e predição acurada de variáveis (SENA et. al., 1999) que, por sua vez, permitem as correlações da capacidade de uso do solo, da fertilidade, redução dos custos de produção, manejo e conservação do solo.

Além disso, o sensor proximal com o espectrômetro, fluorescência dos raios X portáteis (PXRF) pode ser usado para medir metais pesados, cálcio do solo e pH do solo conforme analisa Vasques(2017) em seu estudo sobre fluorescência de raios X nas medidas químicas. Assim é possível a obtenção da matéria orgânica do solo com modelos matemáticos, colorimetria no sistema Munsell de cores, classificação de solos pelo sensor Landsat e a promissora transformação dos laboratórios químicos em laboratórios por sensores.

Já se tem várias vertentes futuristas do uso da inteligência computacional relacionada aos solos, como exemplo o uso em predição de dados meteorológicos no Rio de Janeiro usando Redes Neurais Artificiais (RNAs) e Regressão Linear Múltipla (RLM), obtendo

correlação positiva em seus dados de 91 a 96 % relatado de acordo com Coutinho, et. al.(2015).

Um estudo recente em Portugal mostrou a utilização de *machine learning* para classificação dos solos aptos à irrigação. Para tanto, usou-se os classificadores de decisão Suporte Vetor Machine (SVM), Árvore de decisão, Naive Bayes e Regressão logística para classificar a gleba em apta ou inapta à irrigação, obtendo precisão de 94.44% conforme relatam Torres et al. (2022).

Outro estudo mostrou a possibilidade de mapear digitalmente o solo usando *machine learning* e lógica Fuzzy na bacia do Ribeirão Inhaúma em Iconha/ES e foram obtidas as correlações positivas entre dados preditivos e observados entre as seguintes classes de solos: afloramentos rochosos com e sem vegetação, Cambissolo Háplico, Gleissolo Háplico, Latossolo e Neossolo Litólico. Usou-se o Random Forest, Ridge Regression, Cubist, Partial Least Squares(PLS), PCR, Adaptive Forward Backward Greedy Algorithm(FOBA), Generalized Boosted Regression Models(GBM) e Gradient Boosting with component wise linear models(GLMBOOST) obtendo-se correlações acima de 90% (CAMPBELL,2017).

A predição de dados agronômicos causam inúmeras discussões relacionadas a veracidade de seus resultados em temas como a deficiência de nutrientes pela medição de CTC (Capacidade de Troca Catiônica) e saturação por bases, toxicidade por alumínio, fixação do fósforo pela atração eletromagnética, deficiência de água e oxigênio no solo, impedimento à mecanização, índice climático, profundidade efetiva, erodibilidade do solo, rochosidade e/ou pedregosidade e nível de manejo na área em primitivo, intermediário e desenvolvido conforme se depreende de Dechen e Nachtigall (2007).

Segundo Damolin et al. (2005), a utilização da refletância espectral visando obter informações na identificação e quantificação de características do solo de maneira rápida e não invasiva, tanto em nível laboratorial quanto em nível orbital, tem ocorrido principalmente em países desenvolvidos. Essa tendência vem crescendo no Brasil, no entanto, considerando que se trata de uma linha de pesquisa recente, há a necessidade de suporte para melhor entendimento da interação da energia eletromagnética entre diferentes componentes do solo.

Um dos motivos para o crescimento dessa tendência pode ser relacionado à comprovação científica dos danos ambientais gerados pelas análises físico-químicas em laboratórios. De acordo com Penatti (2011), os subprodutos gerados pela maioria das análises físico-químicas, das quais ocorre a diluição de substâncias-teste com algum tipo de solvente

geram resíduos classificados como perigosos de acordo com as classes de perigo (NBR 10004), elaborada com base no Regulamento Técnico Federal Norte Americano-*Code of Federal Regulation (CFR)-title 40-Protection of environmental-Part 260-265-Hazardous waste management*. (ABNT, 2004a).

Além da necessidade de diminuição de resíduos gerados pelas análises químicas surge outro problema de grande impacto, conforme estuda Queirós (2014), a melhor maneira de prevenir e atender as necessidades alimentares da projeção de nove bilhões de pessoas da terra em 2050, é por meio da combinação de agrotecnologias com tecnologias de informação e da comunicação. Diante desses desafios é importante que se façam estudos relacionados ao tema para começar a surgir luzes em tema ainda obscuro pela ciência do solo haja vista as descobertas recentes relacionadas à análise espectral dos solos.

Nesse ínterim os objetivos do trabalho foram: analisar as associações entre variáveis espectrais e físico-químicas do solo e prever teores de atributos físico-químicos do solo utilizando variáveis espectrais como entrada nos modelos de aprendizagem de máquina.

MATERIAIS E MÉTODOS

Seleção de Amostras e obtenção de dados físico-químicos

Foram selecionadas 1000 amostras de solo em áreas agrícolas de 0 a 20 cm coletadas na região de Chapadão do Sul-MS-Brasil (18°46'26''S 52°37'28''W e altitude média de 810 m). O clima da região é classificado como tropical úmido, com estação chuvosa no verão e seca no inverno, com precipitação anual média de 1.850mm e temperatura anual média de 20,5 e variação de 7,5°C. Os solos da região dos Chapadões são classificados em sua maioria como Latossolo vermelho distrófico. (Santos, et. al. 2018).

As amostras de solo foram peneiradas em malha de 2 mm e secas ao ar. As análises químicas e físicas dos solos foram realizadas seguindo a metodologia descrita por Teixeira et al. (2017). Os atributos do solo avaliados foram: Ca, Mg, P (resina), K, S, Argila, Silte e Areia. Posteriormente essas amostras foram submetidas a avaliações multiespectrais.

Uma alíquota de 20 g de cada amostra de solo peneirada, seco e homogeneizado foi adicionada a placa de Petri para realização de medidas espectrais. A placa Petri foi colocada em uma bancada plana e o sensor foi instalado a 8 cm da superfície do solo. A área de incidência do feixe espectral foi de 3 cm². Duas lâmpadas halógenas externas de 50 W foram posicionadas a 35 cm da placa de Petri em um ângulo zenital de 30° formando um ângulo de 90° entre si seguindo o método descrito por Franceschini et. al.(2013).

Os espectros de refletância foram obtidos pelo CROP CIRCLE ACS-470 (Holland Scientific Inc, Lincoln, NE, EUA). As seis bandas espectrais usadas foram: verde (532 -550 nm), vermelho (670 -700 nm) e borda do vermelho (730-760 nm). O sensor foi calibrado usando FieldCal SC-1. As bandas espectrais foram aplicadas na superfície das amostras de solo. Foram realizadas 100 réplicas para cada banda. As refletâncias foram registradas em planilhas e as médias de refletância foram calculadas para cada banda espectral.

Análise dos dados físico-químicos pela inteligência computacional

As informações obtidas foram objeto de observação e comparação no software WEKA (Waikato Environment for Knowledge Analysis) versão 3.9.3(c) 1999-2018, acessada por um computador com processador AMD Phenom™ IIx4 B97 processador 3.20 GHz, memória instalada RAM 4GB, sistema operacional de 32 bits, Windows 7, usando a validação cruzada de 10 folds(K=10) e 10 repetições em uma análise espectral de 1000 amostras com dados de comprimento de onda obtidos como valores de entrada e macronutrientes como valores de saída a serem preditos.

Os parâmetros utilizados foram testados com alteração, mas optou-se em manter a configuração padrão do software (default) caso do classificador Multilayer Perceptron, usado para resolver problemas de aprendizagem de várias entradas.

Para o M5P utilizado para fornecer mais informações de equações matemáticas e lidar com variáveis categóricas e contínuas e com valores ausentes, os parâmetros também foram alterados, mas optou-se o padrão (default) do software.

Para o Random Forest (combinação de várias árvores de decisão aleatórias com previsões de variáveis de decisões aplicado a problemas de grande escala com sua facilidade de amostragem e acesso fácil a ideias principais), optou-se pela configuração padrão(default) do software.

Para o Random Tree utilizado para construir árvore de decisão com um conjunto aleatório de dados por meio da divisão de nós, optou-se por manter o padrão(default) do próprio software.

Para o Reptree que utiliza a lógica de regressão em múltiplas árvores em diferentes interações para selecionar a melhor árvore, optou-se pelo padrão (default) do software.

Depois disso escolheram-se as melhores redes, árvores e regressão usando sempre o default do software para chegar a melhor acurácia. Compararam-se os resultados de classificadores de Regressão Linear, Random Forest, M5P, Reptree, Multilayer Percetron, e

Random Tree usando como parâmetro de análise o coeficiente de correlação (r) e Erro Médio Absoluto (MAE) dos dados obtidos para chegar à metodologia mais acurada para prever os macronutrientes do solo em análise.

A acurácia dos modelos de predição foi avaliada pela correlação entre valores preditos e observados (r) e erro médio absoluto (MAE). As médias dos parâmetros de acurácia para cada um dos modelos testados foram submetidos à análise de variância para verificar a existência de diferenças entre os modelos de aprendizagem de máquina. Subsequentemente, foram gerados boxplots com as médias de r e MAE para cada modelo no caso da predição de macronutrientes.

As médias dos parâmetros de acurácia foram agrupadas pelo teste de Scott-Knott a 5% de probabilidade. Os boxplots e o agrupamento de médias foram gerados pelos pacotes ggplot2 and ExpDes.pt do software R.

O critério para verificação do coeficiente de correlação de Pearson (r) foi de Figueiredo Filho e Silva Júnior (2009) o qual considera que deve ser medido em fraco, médio e forte, sendo que se considera fraco quando o r estiver em torno de 0,10 a 0,30, moderado quando o r estiver entre 0,40 a 0,60 e forte quando o r estiver de 0,70 a 1.

RESULTADOS E DISCUSSÃO

As análises de aprendizagem de máquina permitiram a associação de cada macronutriente do solo com as bandas espectrais analisadas. Conforme se compara os resultados, diferentes conclusões se depreendem, como os retirados da figura 1. Percebe-se que o nutriente Enxofre obteve maior índice de correlação (r) levando a hipótese de que para os comprimentos de onda analisados variando de 500 a 800 nm, este é mais bem detectado por um sensor espectral comum, diferente dos outros nutrientes que podem precisar de sensores multiespectrais para aumentar suas correlações positivas com o software de aprendizagem.

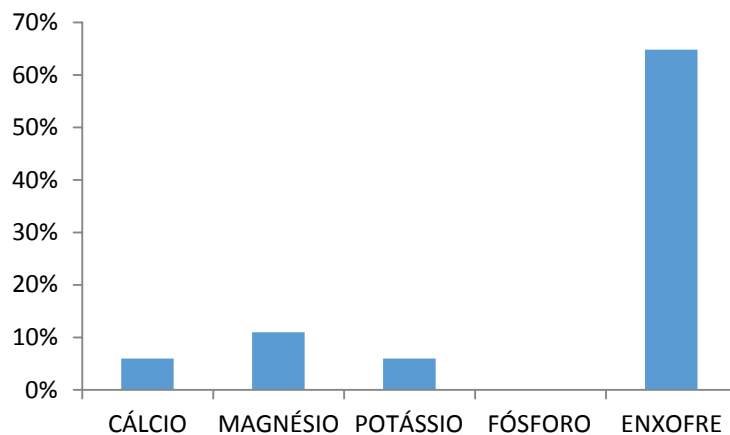


Figura 1: Correlação em porcentagem (%) entre valores observados e preditos usando o classificador Random Forest em seu modo padrão (default).

O cálcio, magnésio e potássio por sua vez apresentaram valores parecidos de correlação (r) mostrando que provavelmente nessa faixa de visibilidade esses nutrientes não são detectados pelo sensor espectral CROP CIRCLE ACS-470.

Contudo para uma mensuração correta não é só o coeficiente de correlação (r) que deve ser analisado, mas o grau de erro detectado pelo software de aprendizagem de máquina, conforme apontado pela figura 2.

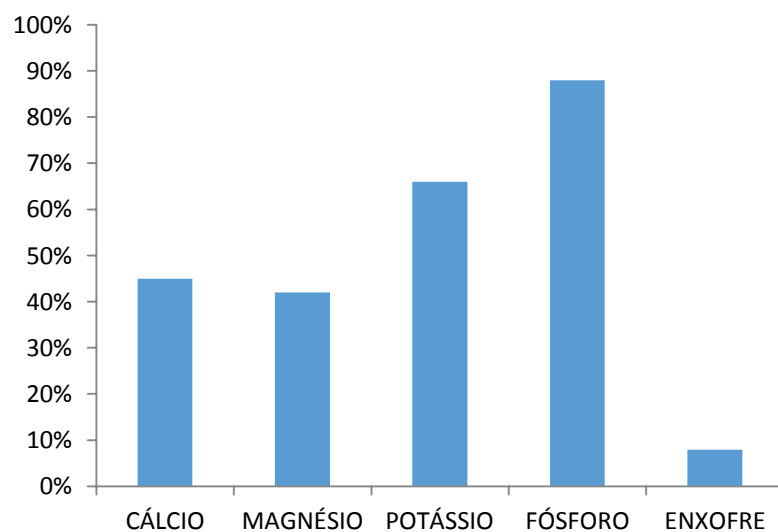


Figura 2: Índices de MAE (Erro Absoluto Médio), variando de 0 a 100 %, do classificador de *machine learning* Random Forest, com nutrientes comparados.

Os menores erros estatísticos comparativos foram encontrados com o magnésio e enxofre, contudo só isso não indica ser o Random Forest a única técnica disponível para classificações, já que de acordo com Martins et al. (2016) o MAE (Erro Médio Absoluto) não

leva em conta se o erro foi subestimado ou sobrestimado, caracterizando-se por ser a média dos erros cometidos pelo modelo de previsão durante uma série de execuções.

Desse modo, para se obter um modelo de predição acurado, é necessário que tanto o MAE seja baixo quanto o r seja alto, o que não se observou com o magnésio, possuindo apenas o MAE baixo e o coeficiente de correlação também baixo. Estatisticamente observa-se uma correlação imperfeita, já que de acordo com Filho e Junior (2009) uma correlação perfeita indica que o escore de uma variável pode ser determinado exatamente ao saber o escore da outra.

Os dados do magnésio tanto com mudança de parâmetros quanto sem mudança de parâmetros apontam pouca correlação entre variáveis, isso demonstrando possivelmente que a maneira de medição de correlação não foi a mais adequada já que os atuais estudos trazem excelentes resultados de correlação.

Demattê (1995) ao avaliar as relações entre dados espectrais e características físicas, químicas e mineralógicas de amostras de solos do Estado do Paraná por rochas básicas e ácidas para prever as cinco classes de solos do Estado pelos dados espectrais por meio de observação das curvas espectrais e da análise das feições de absorção e intensidade de refletância (956 a 1082 nm, 1500 a 1802 nm, 1997 a 2109 nm e 2246 a 2329nm) obteve 100 % de acerto usando dados espectrais para separar classes de solos pela análise discriminante e 100% para dados analíticos do solo.

Este trabalho se diferencia de Demattê (1995) por este ter utilizado o sensor IRIS (Infra Red Intelligent Spectroradiometer) em laboratório com medidas de 400 a 2500 nm (do violeta ao infravermelho próximo). Entretanto, utilizou-se nesta pesquisa comprimentos de onda entre 500 e 800 nm (do verde a borda do vermelho apenas) demonstrando que um intervalo de comprimento de onda maior pode favorecer resultados mais coerentes para uso de espectrometria para análise do solo.

Outra maneira de avaliar, comparar e analisar os resultados é pelo teste de média Scott Knott, de acordo com Jelihoschi et. al. (2014), um algoritmo de agrupamento hierárquico usado como ferramenta exploratória de dados, projetado pelos desenvolvedores da ANOVA, para comparar médias de tratamentos e encontrar grupos homogêneos no teste F significativo, conforme apresentado pela figura 3.

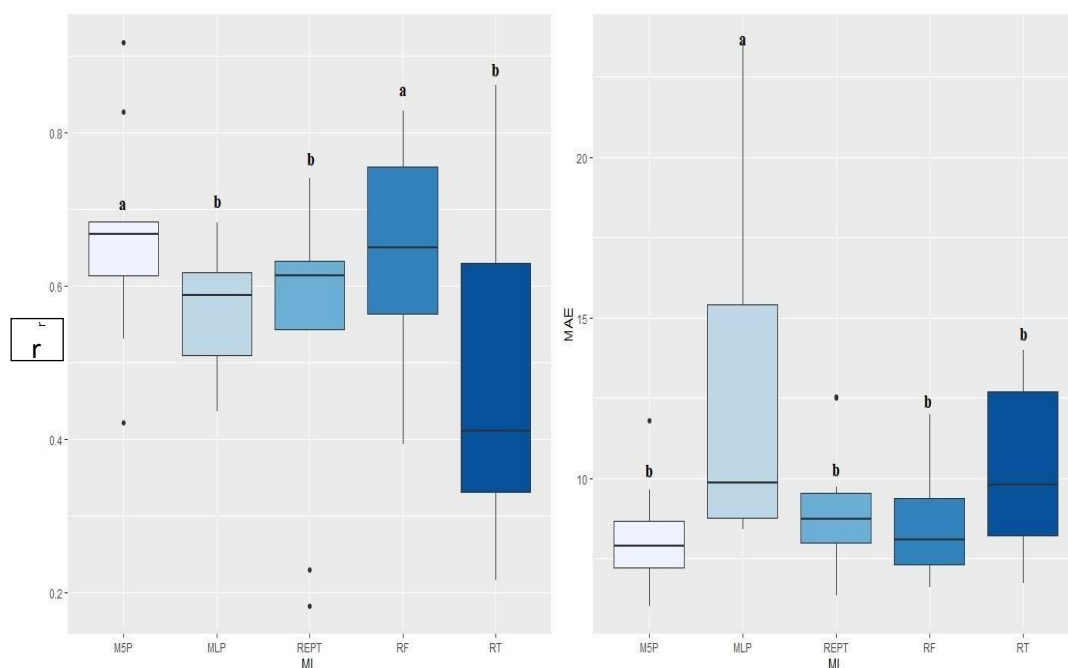


Figura 3. Boxplots para coeficiente de correlação de Pearson(r) e erro absoluto médio (MAE) para diferentes modelos de *machine learning* relacionado ao enxofre.

Dessa forma, com o uso do COP CIRCLE ACS-470 e o Weka, os resultados apresentados revelam que a técnica mais acurada pela análise do Enxofre foi o Random Forest, quando comparada a Reptree, M5P, Random Tree e Multilayer Perceptron. Além disso, evidenciaram-se diferenças estatísticas entre as análises.

Com relação ao MAE, o Multilayer Perceptron apresentou diferença estatística quando comparado a outras técnicas de *machine learning*, sendo neste algoritmo o erro maior que os demais. Apesar de o menor erro estar no classificador M5P, o Random Forest apresenta erro praticamente igual ao M5P, evidenciando-se coerência de análise.

Cientificamente o Random Forest vem obtendo excelentes resultados de predição. De acordo com Devetyarov e Nouretdinov (2010) a maioria dos algoritmos de aprendizado de máquina conhecidos podem ser usados como um algoritmo em um preditor estabelecido, mas eles dependem do tipo de dados analisados e a precisão do algoritmo utilizado.

Esses dados coincidem com o que foi constatado pela análise do enxofre com Random Forest, obtendo-se acurácia de 64,77%. Apesar de esses resultados serem promissores outras redes mostraram resultados parecidos o que levou a necessidade de observar, alterar e analisar um padrão de redes para caracterizar os bons resultados.

O início das mudanças do Random Forest foi com a profundidade da rede e após os testes de várias opções, a melhor rede foi com profundidade ilimitada (0). Posteriormente

foram alterados os tópicos para construção do conjunto notando-se o melhor resultado em (1). Além disso, os atributos aleatórios e a quantidade de sementes foram alterados, mas o padrão Random Forest permaneceu melhor com as indicações do próprio software, levando a crer na generalidade dos resultados da rede melhor que a rede específica tentada.

O classificador M5P do WEKA apresentou resultados parecidos ao Random Forest, mostrando que a parte do algoritmo utilizado se assemelha, sendo a correlação de Pearson (r) de 66,82%. Ao passo que os outros classificadores em uma mesma análise de enxofre tiveram acurácia inferior a 50%, demonstrando que a rede classificadora importa e diferencia a análise e também que para esses dados uma versão generalista de mineração é mais adequada.

Da mesma forma que o enxofre, os outros nutrientes também foram avaliados pelo Scott Knott, conforme se observa na figura 4.

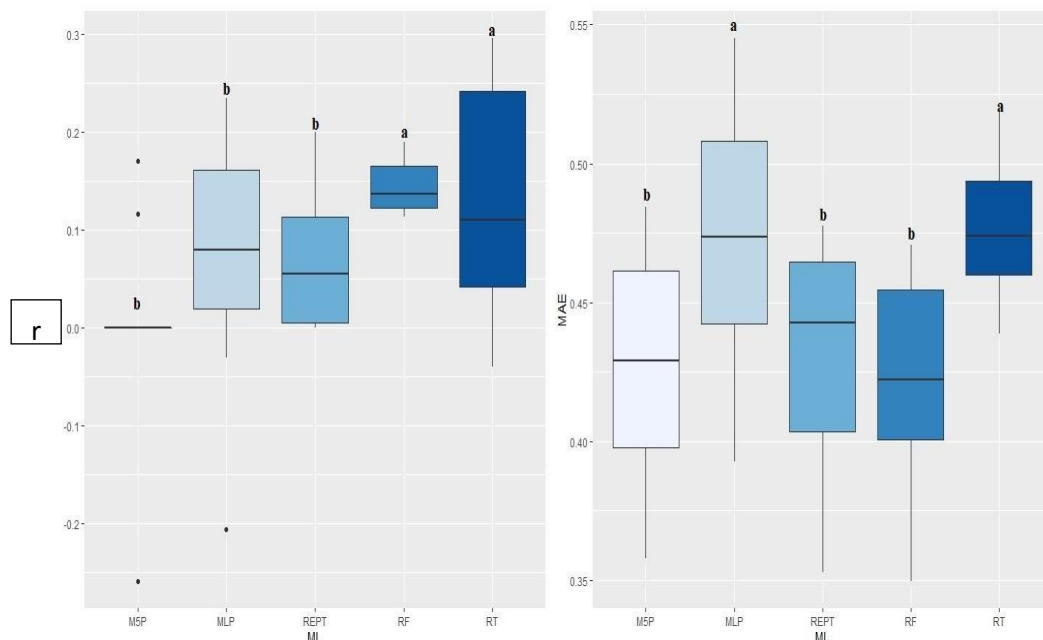


Figura 4. Boxplots para coeficiente de correlação de Pearson(r) e erro absoluto médio (MAE) para diferentes modelos de *machine learning* relacionado ao magnésio.

Com relação ao magnésio, observa-se que o Random Forest e o Random Tree apresentaram diferenças estatísticas relacionadas ao r , sendo melhor acurácia obtida pelo classificador Random Tree. O erro médio absoluto foi maior em Multilayer Perceptron não muito distante do Random Tree, evidenciando-se que para esses dados o Random Tree se mostrou mais adequado.

Ao analisar o potássio percebe-se que não houve diferenças estatísticas entre as técnicas de aprendizagem, sendo a de maior acurácia o Multilayer Perceptron. (Figura 5). O

maior erro absoluto foi obtido da mesma forma pelo Multilayer Perceptron, mostrando coerência da análise, da mesma forma que em relação ao erro não houve diferenças estatísticas entre as técnicas.

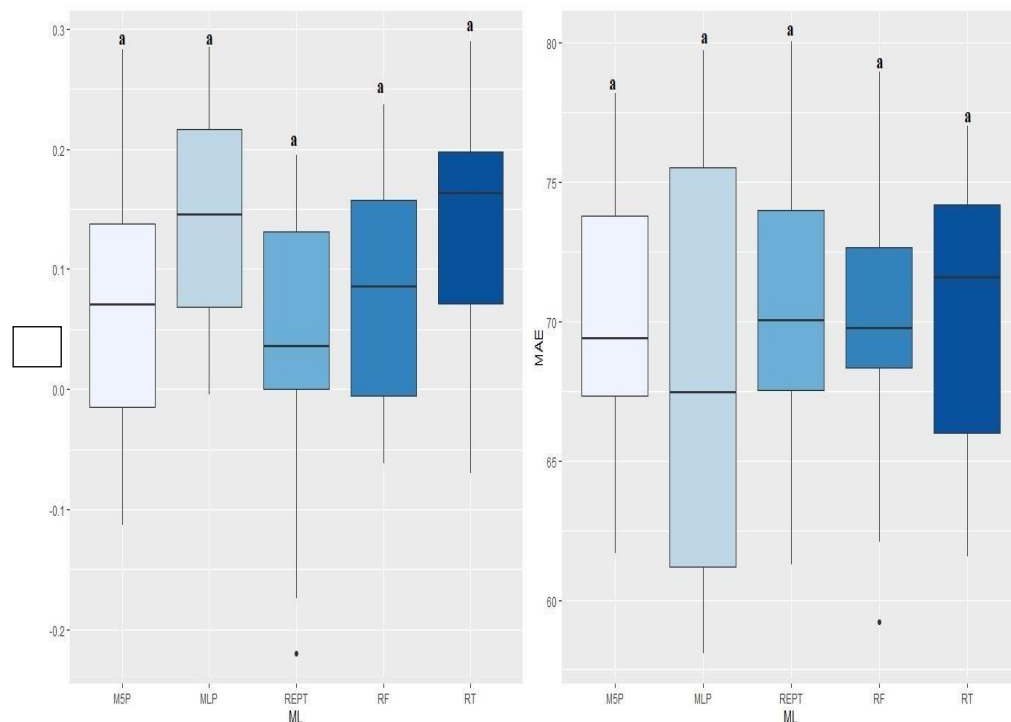


Figura 5. Boxplots para coeficiente de correlação de Pearson(r) e erro absoluto médio (MAE) para diferentes modelos de *machine learning* relacionado ao potássio.

Os dados de cálcio apresentaram diferença estatística em M5P relacionados ao r , sendo a maior correlação conseguida por este (Figura 6). Em relação ao erro médio absoluto, o maior erro foi obtido pelo classificador Random Tree ao passo que o menor erro foi obtido pela técnica M5P, confirmando a hipótese que este classificador é mais adequado aos dados de cálcio.

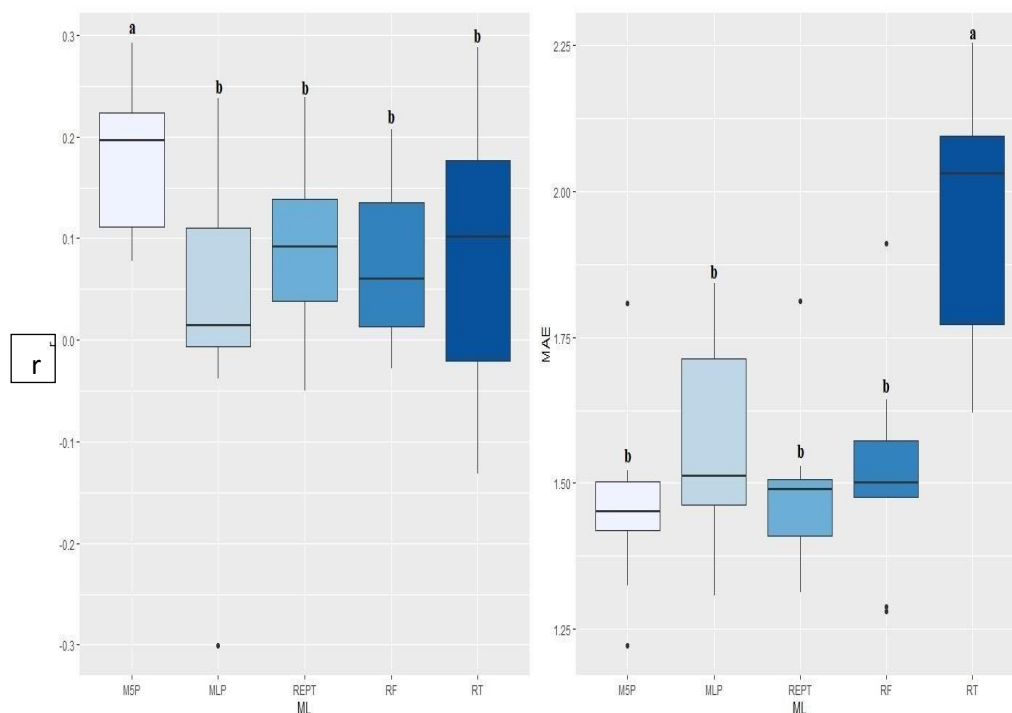


Figura 6. Boxplots para coeficiente de correlação de Pearson(r) e erro absoluto médio (MAE) para diferentes modelos de *machine learning* relacionado ao cálcio.

Da mesma forma que o enxofre, o magnésio, cálcio e potássio apontaram correlação positiva usando as redes neurais, principalmente o Random Tree e M5P, obtendo acurácia de 8,73 a 15,55% para esses nutrientes. Em um primeiro momento nota-se uma correlação pouco expressiva, contudo de acordo com Singhal e Jena (2013) os dados nem sempre levam a conclusões verdadeiras, ou seja, um resultado negativo ou pouco expressivo pode se dar por duplicidade de dados, dados sem qualidade, sem integração ou dados ruidosos.

No caso do potássio não houve diferenças estatísticas entre os algoritmos de aprendizagem. Para o cálcio e magnésio a melhor técnica foi o Random Tree.

Dados de qualidade proporcionam precisão, integridade, consistência, pontualidade, credibilidade e interpretabilidade. Conforme literatura científica de Soares et al. (2014), os macronutrientes já obtiveram correlações positivas acima 65%, chegando a 95 % com uso das medidas espectrais de NIR(Espectropia do infravermelho próximo) e do VIS(Espectropia na região do visível) com uma calibração correta e validação de modelos.

Nesse contexto não é exato afirmar que devido a pouca correlação preditiva fica demonstrado que os teores dos macronutrientes estão baixos ou inadequados.

Outra forma de análise comparativa está relacionada ao comportamento de cada variável de entrada relacionada a variável de saída equacionada matematicamente com base em valores conhecidos, conforme visto na figura 7.

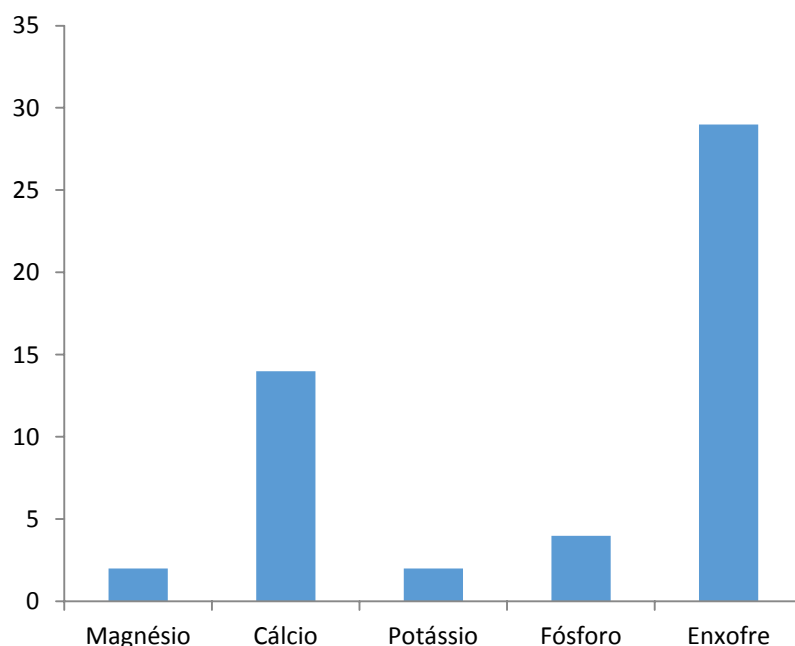


Figura 7: Comparação de cada macronutriente em porcentagem (%) com o modelo de Regressão Linear em modo padrão (default) utilizado pelo software Weka.

A análise da Regressão Linear dos macronutrientes levou a mesma conclusão das outras comparações em que o enxofre apresenta melhor índice de correlação a índices de comprimento de onda variando de 500 a 700 nm. Contudo, em número, apresenta índices muito inferiores que as árvores demonstrando que para esses dados as técnicas de *machine learning* foram melhores quando utilizado os algoritmos classificadores de árvores.

Rodrigues (2018) relata muitas vezes que o algoritmo de Regressão Linear não abarca todos os subconjuntos de treinamento e validação necessários para resultados positivos, daí a necessidade de estudar algoritmos evolutivos com modelos mais complexos. Em seu trabalho ele auferiu a predição de alumínio trocável em amostras de solos conseguindo $R^2=0,61$ e $RMSE=1,24$ no treinamento e $R^2=0,50$ e $RMSE=5,96$ na validação usando a técnica evolutiva SSA (Seleção de Subconjunto de atributos) de seleção de bandas espectrais e metaheurística PSO (Otimização por enxame de partículas) em conjunto com algoritmo de Regressão Linear Múltipla.

Esse autor Rodrigues (2018) diz em seu trabalho que resultados de baixa correlação podem ser aprimorados por meio de desenvolvimento de novas metaheurísticas,

transformação matemática de dados e estudo de novos atributos de interesse do algoritmo para predição do modelo de problema multi-objetivo.

Pelos dados observou-se uma correlação de *Pearson* fraca positiva, ou seja, quando o índice de acurácia está entre 0,1 e 0,5 ($0,1 < r < 0,5$). Nesse caso, de acordo com Matos (1995) o intervalo de confiança dos parâmetros e suas predições são baixos.

CONCLUSÃO

O uso do sensor multiespectral COP CIRCLE ACS-470 nas bandas espectrais de 532, 550, 670, 700, 730 e 760 nm (do verde a borda do vermelho) apresenta correlação positiva com macronutrientes do solo, em destaque ao enxofre que mostrou índices de acurácia acima de 65% com e sem mudança de parâmetros.

Para os macronutrientes P, K, Ca e Mg a precisão de predição dos resultados obtidos pelo mesmo sensor multiespectral comparados à análise quimiométrica obtiveram valores muito baixos, mostrando a necessidade de desenvolvimento de modelos mais acurados para aprimoramento da correlação.

REFERÊNCIAS

- ALVES, Marcelo Rodrigo. **Múltiplas Técnicas no mapeamento digital de solos**. 2008.160p. Tese. (Escola Superior de Agricultura Luís de Queiroz)-Universidade de São Paulo. Piracicaba, 2008.
- BANZATTO, David Arioaldo. **Experimentação agrícola**. 4 ed. Jaboticabal: Funep, 2006.
- BARBOSA, Mauro Cezar et al. **Desempenho agrônômico e componentes da produção de cultivares de soja em duas épocas de semeadura no arenito caiuá**. Semina: Ciências Agrárias, v. 34, n. 3, p. 945–960, 2013.
- BOOF, Leonardo. **Sustentabilidade: o que é e o que não é**. Editora Vozes; 5ª edição, 200 p, 2016.
- BORGES, Wander Luis Barbosa et al. **Plantas de cobertura para o noroeste do estado de São Paulo**. Ciência Rural, v. 45, n. 5, p. 799–805, 2015.
- BRANDT, Eduardo André et al. **Desempenho agrônômico de soja em função da sucessão de culturas em sistema de plantio direto**. Ciência e Agrotecnologia, v. 30, n. 5, p. 869–874, 2008.
- BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília, DF: Centro Gráfico, 1988.
- CAMPBELL, Patrícia Morais da Matta. **Mapeamento Digital de solos e predição de atributos utilizando Machine Learning e lógica fuzzy na bacia do Ribeirão Inhaúma**,

Iconha/ES.2017.Tese.227p.Programa de Pós graduação em Ciencias Ambientais e Florestais-Instituto Florestas. Universidade Federal Rural do Rio de Janeiro-UFRRJ.Rio de Janeiro,2017.

CASTRO, Paulo R.C. e VIEIRA,E.L.**Bioreguladores e bioestimulantes na cultura do milho. Milho: estratégias de manejo para alta produtividade.**Esalq/USP/LPV,p 99-115,Piracicaba,2003.

CONAB, C. N. DE A. **Acompanhamento da safra brasileira de grãos.** Disponível em: <http://www.conab.gov.br> Acesso em: 18 nov. 2021.

CORREIA, J.R. Et. Al.**Cerrado: Correção do solo e adubação.** Brasília-DF: EMBRAPA,2 ed, Informação Tecnológica,416p,2004.

COUTINHO, Eluã Ramos et al. **Utilização de Técnicas de inteligência computacional na predição de dados metereológicos.** Revista Brasileira de Meteorologia, v.31, n.1,24-36, Rio de Janeiro, 2016.

COUTO, Rafael Da Rosa et al. **Phosphorus fractions in soil with a long history of organic waste and mineral fertilizer addition.** Bragantia, v. 76, n. 1, p. 155–166, Rio Grande do Sul,2017.

CRUZ, André Camargo et al. **Fertilizantes Organominerais De Resíduos Do Agronegócio : Avaliação Do Potencial Econômico Brasileiro.** Indústria Química: BNDES Setorial, v. 45, p. 137–187,Rio de Janeiro,2017.

DAMOLIN, Ricardo Simão et. al. **Relação entre constituintes do solo e seu comportamento espectral.** Ciência Rural, Santa Maria, v.35, n.2, p.481-489, mar-abr,2005.

DEMATTE, José Alexandre Melo. **Relações entre dados espectrais e características físico, químicas e mineralógicas de solos desenvolvidos de rochas eruptivas.** Tese.265p.Esalq Piracicaba, 1995.

DEVETYAROV, Dmitry e NOURETDINOV, Ilia. **Artificial Intelligence Applications and Inovations: Prediction with confidence based on a Random Forest Classifier.** Lárnaca-Cyprus.403p. 2010.

FERNANDES, Adão B. da Luz et al. **Agrominerais para o Brasil.** CETEM/MCT,Rio de Janeiro,2010.

FILHO, Dalson Brito Figueiredo e JÚNIOR, José Alexandre Silva. **Desvendando os mistérios do Coeficiente de correlação de Pearson (r).** Revista Política Hoje, Vol. 18, n.1, Pernambuco, 2009.

GUVEN, Duyusen e Akinci, Gorkem. **Comparison of acid digestion techniques to determine heavy metals in sediment and soil samples**. Grazi University Journal of Science. 24: 29-34. Buca, Izmir, Turkey, 2011.

HORST, W. J. et al. **Agronomic measures for increasing P availability to crops**. Plant and Soil, v. 237, n. 1, p. 211–223, Netherlands, 2001.

KAMH, M. et al. **Mobilization of soil and fertilizer phosphate by cover crops**. Plant and Soil, v. 211, n. 1, p. 19–27, Netherlands, 1999.

MACEDO, Manoel Claudio Motta. **Integração lavoura e pecuária : o estado da arte e inovações tecnológicas**. Revista Brasileira de Zootecnia, Viçosa-MG, v. 38, p. 133–146, 2009.

MAROCO, João. **Análise Estatística- Como utilização do SPSS**. 2 ed. Edições Sílabo. Lisboa, 2003.

MARTINS Adalvan Manoel et al. **Relação entre índice SPAD, teores de clorofila extraível e nitrogênio na folha da batata**. Universidade Federal de Viçosa. Viçosa-MG, 2010.

MARTINS, R.S. O. et al. **Recomendação de parâmetros para COIN-OR Branch and Cut**. Simpósio Brasileiro de Pesquisa Operacional. Anais do XLVIII SBPO. Vitoria-Es, 27-30 setembro de 2016.

MATOS, Manoel Antonio. **Manual Operacional para Regressão Linear**. Faculdade de Engenharia da Universidade do Porto, Portugal, 1995.

MATTAR NETO, João Augusto. **Metodologia científica na era da informática**. 2 ed. São Paulo: Saraiva, 2005.

MOLIN, José Paulo. **Agricultura de precisão**. 1ed. Oficina de textos, São Paulo, 2015.

NASCIMENTO, Yasmim Cordeiro. **Proteção Legal Nacional do Solo e dos Lençóis freáticos**. JurisWay: revista eletrônica on line. 2015.

NOVAIS, Roberto Ferreira et al. **Fertilidade do solo**. Viçosa, MG; Sociedade Brasileira de Ciência do Solo, Minas Gerais, 2007.

PACHECO, Leonardo Pereira et al. **Ciclagem de nutrientes por plantas de cobertura e produtividade de soja e arroz em plantio direto**. Pesquisa Agropecuária Brasileira, v. 48, n. 9, p. 1228–1236, Goiás, 2008.

PASSOS, Alexandre Martins Abdao et al. **Cultivares de soja em sucessão ao trigo nos sistemas convencional e plantio direto**. Agrarian, v. 8, n. 27, p. 30–38, 2015. Disponível em: <https://ojs.ufgd.edu.br/index.php/agrarian/article/view/3015>. Acesso em: 8 mar. 2023.

PAVINATTO, Paulo Sergio et al. **Acúmulo de fosforo no solo em áreas agrícolas no Brasil: diagnóstico atual e potencialidades futuras**. Informações Agronômicas: Nutrição de Plantas Ciência e Tecnologia. Número 9, Março, Piracicaba-SP, 2021.

PEEL, M. C.; FINLAYSON, B. L.; MCMAHON, T. A. **Updated world map of the Köppen-Geiger climate classification**. Hydrology and Earth System Sciences, v. 11, n. 5, p. 1633–1644, Victoria-Australia, 11 out. 2007.

PENATTI, Fabio Eduardo et al. **Avaliação dos riscos ambientais causados pela disposição incorreta de resíduos de laboratórios**. Geografia: Ensino e Pesquisa, v.15, n.1, p. 43-52, jan/abr. Santa Maria-RS, 2011.

QUEIROS, Leonardo Ribeiro et al. **Análise das possibilidades e tendências do uso das tecnologias da informação e comunicação em Agricultura de Precisão**. Ed. Agricultura de precisão: resultados de um novo olhar. Brasília, DF: EMBRAPA, 2014.

QUEVEDO, Claudia Maria Gomes de. **As atividades do homem e a evolução do fosforo no meio ambiente**. Dissertação de mestrado. Universidade de São Paulo. 229p. São Paulo-SP, 2009.

RAIJ, Bernardo Van; ANDRADE, J.C. De; CANTARELLA, H; QUAGGIO, J.A. **Análise Química para avaliação da fertilidade de solos tropicais**. Instituto Agrônomo, Campinas-SP. 285p. 2001.

RATKE, Rafael Felipe. et al. **Agricultural potential and soil use based on the pedogenetic properties of soils from the cerrado-caatinga transition**. Semina-Ciências Agrárias, v. 41, p. 1119-1134, 2020p. 1119–1134, Londrina-PR, 2020.

RATKE, Rafael Felipe. et al. **Different limestone particle sizes for soil acidity correction, Ca and Mg supply and corn yield**. Comunicata Scientiae, v. 9, n. 2, p. 175–184, Piauí, 2018.

RIGO, Amanda Zolet et al. **Phosphorus fractions in soil with organic and mineral fertilization in integrated crop-livestock system**. Revista Brasileira de Ciencia do Solo, v. 43, p. 1–15, Viçosa-MG, 2019.

RODRIGUES, Giancarlo. **Seleção de bandas espectrais apoiada pela metaheurística PSO para predição do teor de alumínio trocável de amostras de solo**. Dissertação de mestrado. 70p. UEPG. Ponta Grossa-PR, 2018.

SASAKI, Yutaka. **The truth of the F-measure**. School of Computer Science. University of Manchester, octuber, Manchester, 2007.

SILVA, Ygor Jacques Agra Bezerra da. **Métodos de digestão de solos e fertilizantes para análises de metais pesados**. Dissertação. 41p. Universidade Federal Rural de Pernambuco-

PE.Recife-PE, 2012.

SOARES, Valdeir Moraes et al. **Calibração e validação de modelos para estimativa de atributos químicos dos solos com uso da espectroscopia**. Aprosoja: Agro cientista. Mato Grosso, 2014.

SOUSA, D. M. G.; LOBATO, E. **Cerrado: correção do solo e adubação**. 2^a ed. Brasília: Embrapa Informação Tecnológica, 2004.

SWASTI, Singhal e MONIKA Jena. **A study on weka tool for data preprocessing, classification and clustering**. International journal of innovative technology and exploring engineering. Volume 2, Issue 6. May,2013.

TEIXEIRA, Paulo Cezar et al. **Manual de métodos de análises de solos**. 3^a ed. Brasília: Embrapa, 2017.

TEODORO, Larissa Pereira Ribeiro et al. **A Random Forest ranking approach to predict yield in maize with uav-based vegetation spectral indices**. Elsevier-article.Vol 178.Nov. 2020.

TEODORO, Larissa Pereira Ribeiro et al. **Predicting days to maturity, plant height, and grain yield in soybean: a machine and deep learning approach using multispectral data**. Remote sensing-article.Out.2021.

TORRES, Pedro et al. **Aplicação de Técnicas de Machine Learning para classificação da aptidão dos solos para o regadio**. Instituto Politécnico de Castelo Branco-Escola Superior Agrária. Portugal. Castelo Branco, 2022.

VASQUES, Gustavo M. et al.**Medição in situ e predição de elementos químicos do solo usando espectrômetro de fluorescência de raios-X**, Congresso Brasileiro de Geoquímica,16,artigo em anais do Congresso, Niterói- RJ, 2017.

VOLPATO, Gilson Luiz. **Guia prático para redação científica**.286p.Botucatu: Best Writing,2015.

ZHAOYU, Zhai et al. **Decision support systems for agriculture 4.0: Survey and challenges**. Elsevier, Computers and electronics in Agriculture, March, 2020.

ZUFFO, Alan Mario et al. **Analysis of growth and agronomic characteristics of off-season corn grown insuccession with soybean and submitted to nitrogen doses**. Revista em Agronegócio e Meio Ambiente,Maringá-PR,v. 14, n. 1, p. e007960, 2021.

Capítulo 2

CLASSIFICAÇÃO DOS TEORES DE MACRONUTRIENTES E GRANULOMETRIA DO SOLO UTILIZANDO SENSOR MULTISPECTRAL E APRENDIZAGEM DE MÁQUINA.

RESUMO: A necessidade de aumento da produtividade agrícola aliado à manutenção do nível de fertilidade do solo em regiões muito demandadas vem trazendo novas oportunidades e desafios à nova era da informatização tecnológica na agricultura. A primeira inovação mais simples e já confirmada por outros pesquisadores se refere à substituição dos meios tradicionais de análises físico-químicas para sistemas de comparação de resultados por meio da inteligência computacional. A segunda oportunidade mais inovadora foi o objeto deste estudo: desenvolver um modelo de classificação e avaliação da fertilidade do solo por meio de algoritmos de decisão Random Forest, M5P, J48, Regressão Logística, Multilayer Percetron, Random Tree e Reptree e técnicas de inteligência computacional para obtenção de gráficos estatísticos sem a necessidade de uso de métodos tradicionais potencialmente danosos ao meio ambiente. Através de análises físico-químicas tradicionais e do estudo de bandas espectrais (do verde a borda do vermelho) usadas para a predição de macronutrientes, teores e granulometria foi obtido 90% de classificações corretas a partir de dados multiespectrais, areia, silte, argila e fósforo pelo método mehlich como variável avaliada e granulometria (argiloso, médio e arenoso) como variável a ser predita.

Palavras-chaves: Machine Learning. Fertilidade. Macronutrientes. Classificação.

ABSTRACT: The need to increase agricultural productivity combined with maintaining the soil fertility level in highly demanded regions has brought new opportunities and challenges to the new era of technological computerization in agriculture. The first, simpler innovation, already confirmed by other researchers, refers to the substitution of traditional means of physical-chemical analysis for systems for comparing results through computational intelligence. The second most innovative opportunity was the object of this study: to develop a classification and evaluation model of soil fertility through decision algorithms Random Forest, M5P, J48, Logistic Regression, Multilayer Percetron, Random Tree and RepTree and computational intelligence techniques to obtain statistical graphs without the need to use traditional methods potentially harmful to the environment. Through traditional physical-chemical analyzes and the study of spectral bands (from green to red border) used for the

prediction of macronutrients, contents and granulometry, 90% of correct classifications were obtained from multispectral data, sand, silt, clay and phosphorus by the mehlich method as the evaluated variable and granulometry (clayey, medium and sandy) as the variable to be predicted.

KEYWORDS: Machine Learning. Fertility. Macronutrients. Classification.

INTRODUÇÃO

A produção agropecuária impulsiona a economia do Brasil. Nos próximos 10 anos há uma tendência de aumento na produtividade agrícola em 2,6%, de 3,4 t ha⁻¹ para 4,1 t ha⁻¹ (MAPA, 2017). O aumento da produtividade está inerente à adoção de tecnologias desenvolvidas para a produção agropecuária. O bioma Cerrado ocupa 4% das regiões tropicais do mundo, no Brasil ocupa 207 milhões de hectares (FIGUEIREDO; RESCK; CARNEIRO, 2010). O cerrado está em uma zona intertropical com clima caracterizado pelo seu aspecto sazonal, com ocorrência de duas estações bem definidas: seca e úmida (BRANNSTROM et al., 2008).

A soja é a oleaginosa mais cultivada no mundo e a principal commodities agrícolas do Brasil, ocupando cerca de 57% da área total cultivada no país são com esta cultura (CONAB, 2022). Sendo o milho a segunda cultura mais importante, principalmente pela área cultivada na safra 2021/2022 foi de 17,25 milhões de hectares cultivados durante a primeira colheita (primavera-verão), que é semeada de setembro a dezembro e colhida de janeiro a abril, e a segunda colheita (outono-inverno), que é semeada de janeiro a março e colhida de maio a agosto (CONAB, 2022).

O Brasil é o quarto maior consumidor de fertilizantes do mundo, consumindo em 2021, 40 milhões de toneladas de fertilizantes, aumento de 14% no consumo em relação a 2020 (ANDA, 2021). A alta dependência da importação de fertilizantes traz o risco de, no futuro, o país enfrentar escassez desse insumo, devido países produtores de fertilizantes, como a China e EUA, serem também grandes consumidores (VEGRO, 2018). Portanto, é indispensável o uso de tecnologias adequadas para reduzir o uso de fertilizantes ou utilizar fertilizantes de maior eficiência.

A análise de solos e plantas nos sistemas de cultivos engloba o conjunto de processos químicos complexos (POPPIEL; PAIVA; DEMATTÊ, 2022). Atualmente o método tradicional utilizado pela maioria dos cientistas agrônômicos é a extração de uma solução

química conjugada e depois procedimentos analíticos em equipamentos especializados (DEMATTE et al., 2017)

A alternativa para análise de solos e plantas advém do uso de novas tecnologias, a exemplo da espectroscopia de refletância com modelos de predição para variáveis químicas e físicas dos solos (RAMIREZ-LOPEZ et al., 2013). Após análises espectrométricas utilizam-se o PCA (análise dos componentes principais), métodos de calibração e previsão (Regressão Linear múltipla, Regressão em Componentes Principais-PCR, Mínimos Quadrados Principais-PLS e métodos de inteligência artificial como redes neurais para correção e predição corretas de resultados (POPPIEL; PAIVA; DEMATTÊ, 2022)).

Os agricultores na região de cerrados utilizam doses de fertilizantes superiores aos preconizados em boletins de correção e adubação do solo utilizados na região (SOUSA; LOBATO, 2004), em virtude de que a não utilização destas quantidades podem não proporcionar produtividades compensatórias, principalmente quando o objetivo é o cultivo de culturas mais exigentes (DONAGEMMA et al., 2016).

As recomendações de calagem e adubação na agricultura são realizadas tradicionalmente através de consultas com os dados das análises de solos e tecidos foliares a tabelas de interpretações existentes. Esse método de recomendação apresenta algumas limitações, dentre as quais podemos destacar a necessidade de conhecimento especializado para consulta das tabelas e a idade dos estudos de origem das tabelas, que em alguns casos estão desatualizadas frente aos avanços alcançados nos últimos anos.

Diante disto, se fazem necessários estudos para atualização das tabelas de interpretação e adubação e desenvolvimento de novos métodos ou ferramentas que possibilitem levar estes conhecimentos aos produtores nos seus diferentes níveis tecnológicos, permitindo a todos um uso mais eficiente dos corretivos e fertilizantes. Uma alternativa ao sistema tradicional de interpretações e recomendações é a automatização desse processo por ferramentas computacionais, que além da possibilidade do uso das tabelas de interpretação e recomendação já existentes, permite também a integração de novos dados, armazenar e gerenciar estes dados permitindo maior eficiência no uso de fertilizantes para promover a produtividade das culturas.

Para sanar esses entraves ao avanço tecnológico esse trabalho teve como objetivos: analisar as associações entre variáveis espectrais e físico-químicas do solo e classificar teores e granulometria do solo por meio de modelos de aprendizagem de máquina e técnicas de inteligência computacional.

MATERIAIS E MÉTODOS

Seleção de Amostras e obtenção de dados físico-químicos

Foram selecionadas 1000 amostras de solo de 0 a 20 cm, coletadas na região de Chapadão do Sul-MS-Brasil (18°46'26''S 52°37'28''W e altitude média de 810 m). O clima da região é classificado como tropical úmido, com estação chuvosa no verão e seca no inverno, com precipitação anual média de 1.850mm e temperatura anual média de 20,5 e variação de 7,5°C. Os solos da região dos Chapadões são classificados em sua maioria como Latossolo vermelho distrófico. (Santos, et. al. 2018).

As amostras de solo foram peneiradas em malha de 2 mm e secas ao ar. As análises químicas e físicas dos solos foram realizadas seguindo a metodologia descrita por Teixeira et al. (2017). Os atributos do solo avaliados foram: Ca, Mg, P (resina), K, S, Argila, Silte e Areia. Posteriormente essas amostras foram submetidas a avaliações multiespectrais.

Uma alíquota de 20 g de cada amostra de solo peneirada, seco e homogeneizado foi adicionada a placa de Petri para realização de medidas espectrais. A placa Petri foi colocada em uma bancada plana e o sensor foi instalado a 8 cm da superfície do solo. A área de incidência do feixe espectral foi de 3 cm². Duas lâmpadas halógenas externas de 50 W foram posicionadas a 35 cm da placa de Petri em um ângulo zenital de 30° formando um ângulo de 90° entre si seguindo o método descrito por Franceschini et. al.(2013).

Os espectros de refletância foram obtidos pelo CROP CIRCLE ACS-470 (Holland Scientific Inc, Lincoln, NE, EUA). As seis bandas espectrais usadas foram: verde (532 -550 nm),vermelho (670 -700 nm) e borda do vermelho(730-760 nm). O sensor foi calibrado usando FieldCal SC-1. As bandas espectrais foram aplicadas na superfície das amostras de solo. Foram realizadas 100 réplicas para cada banda. As refletâncias foram registradas em planilhas e as médias de refletância foram calculadas para cada banda espectral.

Análise dos dados físico-químicos pela inteligência computacional

As informações obtidas foram objeto de observação e comparação no software WEKA (Waikato Environment for Knowledge Analysis) versão 3.9.3(c) 1999-2018, acessada por um computador com processador AMD Phenom™ IIx4 B97 processador 3.20 GHz, memória

instalada RAM 4GB, sistema operacional de 32 bits, Windows 7, usando a validação cruzada de 10 folds ($K=10$) e 10 repetições em uma análise espectral de 1000 amostras com dados de comprimento de onda obtidos como valores de entrada e macronutrientes como valores de saída a serem preditos. Além disso, usaram-se dados espectrais de entrada e teores como valores de saída para analisar as classificações corretas.

Os parâmetros utilizados foram testados com alteração, mas optou-se em manter a configuração padrão do software (default) do Weka no caso do classificador Multilayer Perceptron, usado para resolver problemas de aprendizagem de várias entradas.

Para o M5P utilizado para fornecer mais informações de equações matemáticas e lidar com variáveis categóricas e contínuas e com valores ausentes, os parâmetros também foram alterados, mas optou-se o padrão (default) do software Weka.

Para o Random Forest (combinação de várias árvores de decisão aleatórias com previsões de variáveis de decisões aplicado a problemas de grande escala com sua facilidade de amostragem e acesso fácil a ideias principais), optou-se pela configuração padrão (default) do software Weka.

Para o Random Tree utilizado para construir árvore de decisão com um conjunto aleatório de dados por meio da divisão de nós, optou-se por manter o padrão (default) do próprio software Weka.

Para o Reptree que utiliza a lógica de regressão em múltiplas árvores em diferentes interações para selecionar a melhor árvore, optou-se pelo padrão (default) do software Weka.

Para o J48, algoritmo que cria modelos de decisão na própria árvore por indução e posterior classificação foi usado o (default) do próprio software Weka.

Para Regressão Logística foi usado o mesmo padrão (default) dos parâmetros do software Weka

Depois disso escolheram-se as melhores redes, árvores e regressão usando sempre o default do software para chegar a melhor acurácia. Modelos de classificações dos solos foram testados utilizando os resultados de classificadores de Regressão Logística, Random Forest, M5P, J48, Reptree, Multilayer Perceptron, e Random Tree usando como parâmetro de análise classificações corretas (CC) e F-measure (medida de precisão).

A acurácia dos modelos de classificação foi mensurada pelas métricas de porcentagem de classificações corretas (CC) e medida F (F-measure). As médias dos parâmetros de

acurácia para cada um dos modelos testados foram submetidos à análise de variância para verificar a existência de diferenças entre os modelos de aprendizagem de máquina.

Subsequentemente, foram gerados boxplots com as médias de CC e F-measure no caso das classificações dos solos.

As médias dos parâmetros de acurácia foram agrupadas pelo teste de Scott-Knott a 5% de probabilidade. Os boxplots e o agrupamento de médias foram gerados pelos pacotes ggplot2 and ExpDes.pt do software R.

RESULTADOS E DISCUSSÃO

Os teores do solo foram classificados em baixo, médio, adequado e alto (Figura 8). De um conjunto de 912 amostras viáveis, 40% apresentaram níveis adequados de Ca para cultivos de plantas na região dos cerrados. Isso demonstra que para essas amostras de acordo com Sousa e Lobato (2004) e dependendo da espécie botânica a ser utilizada, da acidez do solo, do nitrogênio, fósforo e potássio associado além de outros fatores ambientais a calagem se mostra desnecessária.

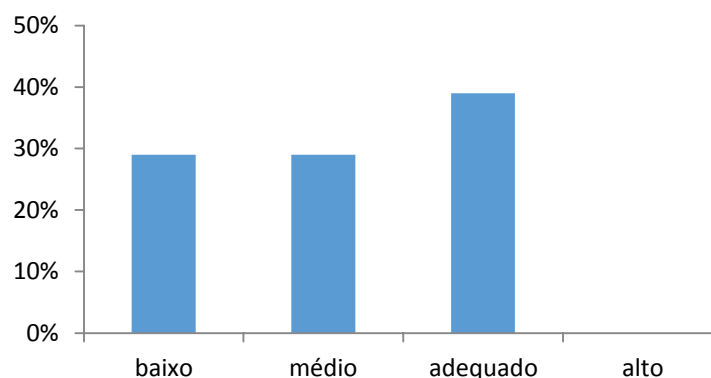


Figura 8: Análise em tabela comparativa do conjunto de 912 amostras de cálcio em porcentagem, variando de 0 a 100%.

Em seu artigo Albuquerque et al. (2003), relataram que além dos fatores citados aqui anteriormente para aplicação de calcário, outros itens devem ser levados em consideração como a dependência das estruturas da matéria orgânica, estabilidade dos agregados, condutividade elétrica, efeito salino e composição química do solo.

Albuquerque et al. (2003) avaliaram amostras de latossolos em parcelas de plantio direto e subparcelas usando calcário e concluíram também que a calagem em solos com alto tamponamento interfere negativamente na floculação das partículas, mas não compromete a estabilidade estrutural do solo. Assim para se caracterizar como adequado ou não para o cultivo, todos esses fatores devem ser levados em consideração e aqui os estudos

comparativos diferem quanto ao teor adequado para cada espécie, dependendo do caso, um ou outro teor se mostra apto ou não ao cultivo.

Para chegar ao nível adequado de calcário para fertilidade do solo em cultivo de cana-de-açúcar os autores Rossetto et al.(2004) analisaram que em solos ácidos o nível necessário de calcário aumenta mas também observaram que as cultivares plantadas sofrem adaptação e os nutrientes se reorganizam de maneiras diferentes, no caso do estudo da interação calcário-potássio o acréscimo de potássio não aumentou a produtividade da cana-de-açúcar mesmo em solo rico em cálcio para a maioria dos blocos estudados.

Em estudo de produção de bananas realizado por Ana Borges e Edilson Carvalho (2014), por exemplo, no estado do Pará, foi observado que para uma boa produção o alumínio no solo deve permanecer inferior a 10 % e o cálcio elevar a saturação a 50% dependendo da camada do solo a ser manejada.

Para o cálcio a maior porcentagem de classificações corretas foi de 41% usando Regressão Logística do software em modo default e que os dados estão em mais de 50% harmônicos de acordo com medida F, conforme se observa na figura 9.

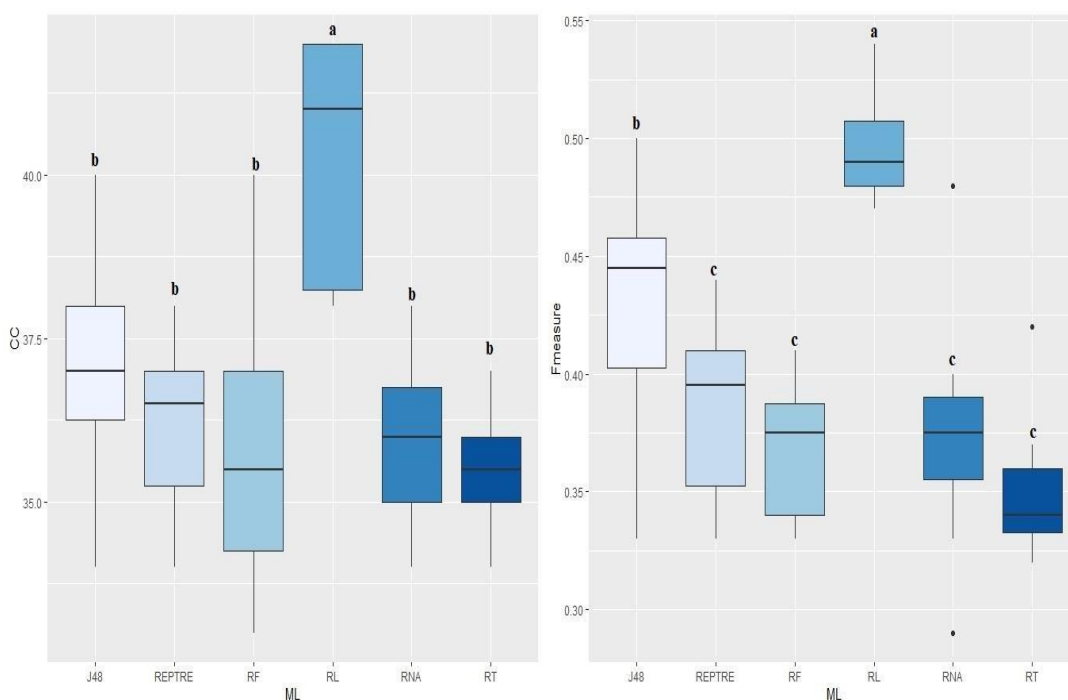


Figura 9: Boxplots para classificações corretas (CC) e F-measure para diferentes modelos de *machine learning* relacionado ao cálcio.

A figura permite depreender que o cálcio foi classificado corretamente em maior número usando o classificador de Regressão Logística. Assim, em tese, em um futuro

próximo e com os instrumentos adequados seria possível prever o nutriente e automaticamente indicar seu teor classificando o solo em apto ou inapto dependendo do uso do nutriente e da cultura a ser cultivada.

Observa-se também que o algoritmo de Regressão Logística difere estatisticamente dos outros classificadores para serem corretas. Outra forma de analisar estatisticamente a análise é por meio do F-measure (ou medida de verdade), a qual Sasaki (2007) define como a média harmônica entre precisão e recuperação de dados. Para F-measure houve diferenças entre os classificadores Regressão Logística, J48 e as demais técnicas de classificação, contudo seus resultados mostraram muita similaridade com a Regressão Logística de classificações corretas para esses dados.

Da mesma forma que o cálcio, também foi medido físico-quimicamente o Enxofre por meio das análises tradicionais obtendo 65% de teores baixos e posteriormente foi feita a comparação com técnicas de *machine learning* e seus teores conforme figura 10 e 11.

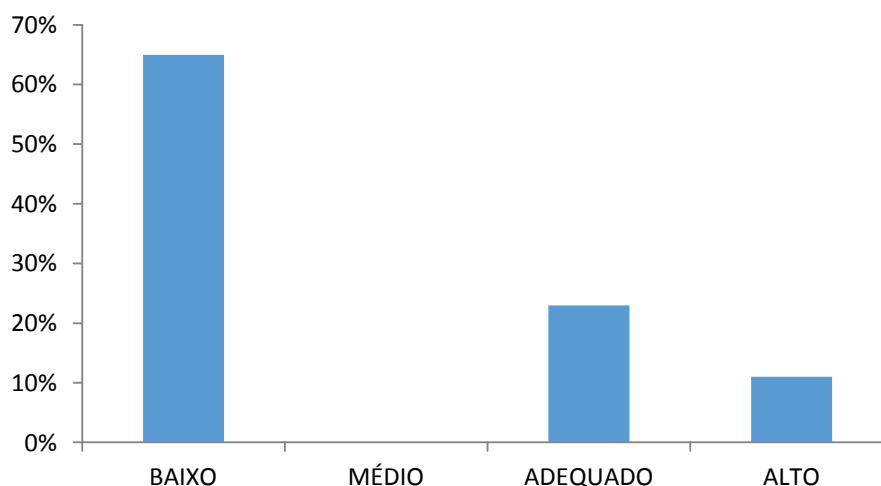


Figura 10: Análise em tabela comparativa do conjunto de 358 amostras de enxofre em porcentagem, variando de 0 a 100%.

Segundo Sousa e Rein (2003), o enxofre apresenta semelhanças com nitrogênio, sua maior parte está na forma orgânica em solos não adubados e assim como o nitrogênio a recomendação de adubação é complexa, devido à dinâmica de transformações desses nutrientes no solo, sua mobilidade e aproveitamento pelas plantas.

Nos estudos de Broch et. al.(2011) constaram que a produtividade de soja no cerrado é influenciada diretamente pelo enxofre. Em três anos de estudos na região de Maracaju-Ms usando superfosfato simples, MAP sulfurado+sulfurgran, Sulfurgran, enxofre elementar,

gesso granulado, Fosmag 509M6, gesso agrícola a soja teve maior produtividade média de grãos, chegando a valores superiores a 3100 kg/ha.

Assim para essas amostras o nível de enxofre está em sua maior parte baixo, levando a hipótese de que, em uma situação de plantio, será preciso uma adubação deste nutriente de acordo com a necessidade de cada cultivar, ambiente do momento da semeadura e os outros itens para correto manejo da área de cultivo.

Os resultados do enxofre foram em termos de médias de classificações corretas de mais de 70 % de acerto e mais de 80% de médias de harmonia de dados por F-measure conforme se observa na figura 11.

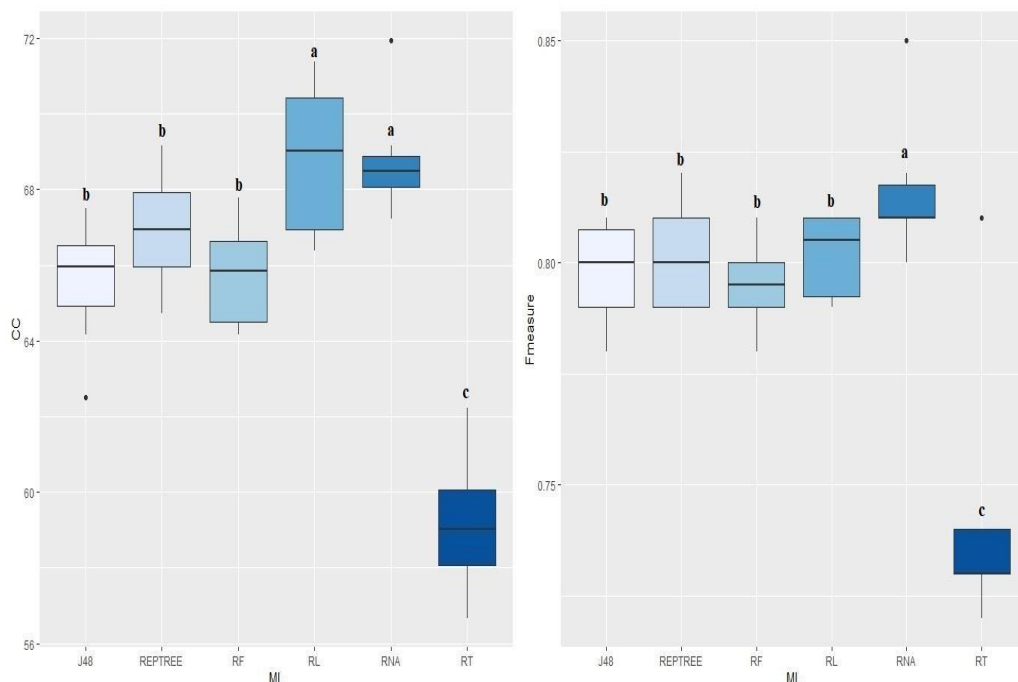


Figura 11: Boxplots para classificações corretas (CC) e F-measure para diferentes modelos de *machine learning* relacionado ao enxofre.

Percebe-se no caso comparativo das classificações corretas que a Regressão Logística obteve novamente maiores índices de acertos chegando a 70% mostrando ser possível classificar teores por meio do *machine learning*, apresentando, contudo, maiores variações de amplitude do teste Scott Knott. Estatisticamente obteve-se três diferenças entre classificadores, uma da Regressão Logística, outra de Random Forest, Reptree, J48 e outra do Random Tree.

Caleiro (2019) faz um estudo por meio de imagens de VANT (veículos aéreos não tripulados) das áreas de plantações de uva em Portugal estudando o comportamento da área,

detecção de seca, localização de pragas, estimativa de produtividade, modelação hidrológica e mapeamento de culturas por radiações eletromagnéticas (azul ao vermelho) e índices de vegetação.

Ainda Caleiro (2019) conclui por meio de estudos espectrométricos que com a adubação foliar os teores de enxofre acumulado aumentaram nas folhas assim como o vigor das uvas. De maneira complementar a detecção do nível de enxofre por meio do *machine learning*, outra aposta seria complementar a análise físico-química por estudos de índices de vegetação.

Nos dados deste estudo obtidos pela inteligência computacional o F-measure obteve melhores resultados com o classificador Multilayer Percetron (RNA) e assim como a outra análise do cálcio os piores resultados foram com o classificador Random Tree, mostrando que nesse caso uma correta ou incorreta escolha do algoritmo classificador pode alterar os resultados dos dados comparativos.

Foi analisado também o magnésio físico-quimicamente obtendo-se 70% de teores adequados para cultivo de solos do cerrado e posteriormente comparados com dados de inteligência computacional e os resultados obtidos podem ser observados na figura 12 e 13.

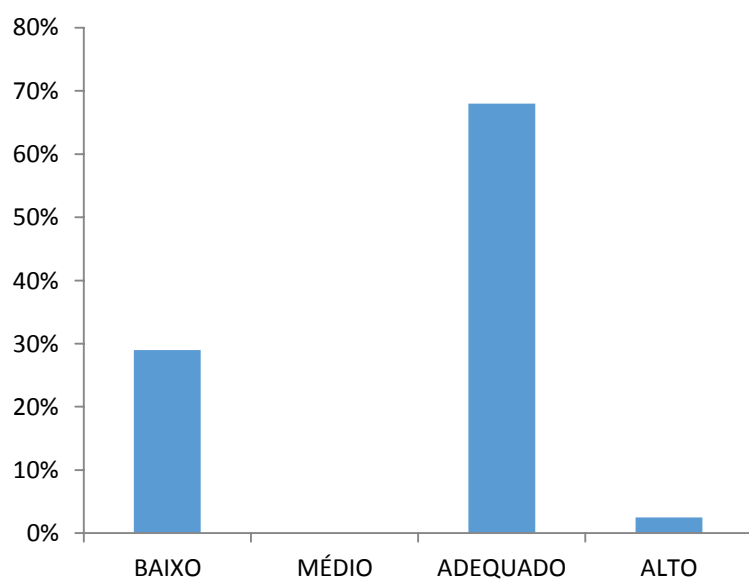


Figura 12: Análise em tabela comparativa do conjunto de 688 amostras de magnésio em porcentagem, variando de 0 a 100%.

Para essas amostras, os níveis de magnésio estão adequados em sua maior parte, assim em um possível uso de calcário, o acréscimo deste nutriente pode ajudar a planta em seu desenvolvimento.

Em estudo de Frazão (2008) foi observado que embora neossolos da região do cerrado possuam limitações para o cultivo, em razão de sua baixa fertilidade natural, se adequadamente manejado, este solo apresenta potencial para cultivo agrícola, ele ainda conclui que baixas concentrações de cálcio e magnésio, no sistema nativo do Cerrado é uma resposta esperada para esse tipo de solo e a correlação entre cálcio e magnésio é altamente significativa com coeficientes de correlação de 0,68 e 0,96 tanto para pastagem, plantio convencional e plantio direto.

Da mesma forma que análises anteriores, foram feitas comparações de resultados físico-químicas e posteriormente avaliados por meio do *machine learning* obtendo-se 68% em média das classificações corretas pelos classificadores utilizados, contudo, 34% em média de harmonia de dados pelo F-measure, como observado pela figura 13.

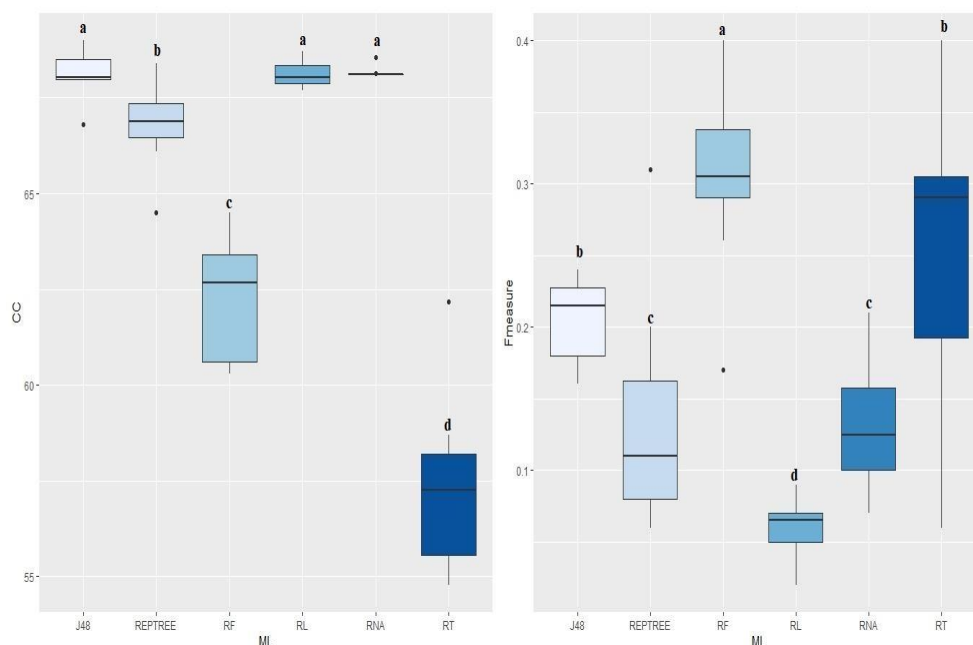


Figura 13: Boxplots para classificações corretas (CC) e F-measure para diferentes modelos de *machine learning* relacionado ao magnésio.

Resultados promissores foram conseguidos por meio das classificações corretas com RNA (Multilayer Perceptron), Regressão Logística e J48, ficando um pouco abaixo do previsto os classificadores Random Tree e Random Forest.

Para F-measure, contudo houve muitas variações entre as técnicas e resultados distantes da normalidade, o que pode gerar desconfiança quanto aos dados analisados, porém ao analisar o classificador F-measure isso se torna superado uma vez que Sasaki (2007) relata que este mede uma média harmônica dos dados e não sua veracidade.

Em estudo de Santos (2019) apontou que o magnésio possui correlação expressiva na faixa VIS-NIR de 2300nm, 2380nm, 2445nm e os valores de correlação variam de 0,35 a 0,34 e também Dematê et al.(2017) utilizou modelos de predição por Regressão Linear e modelos espectrais e obteve da mesma forma resultados satisfatórios com potássio e magnésio.

O potássio foi analisado físico-quimicamente com 60% de teores altos como se observa na figura 14, posteriormente foi comparado com software para analisar as classificações corretas.

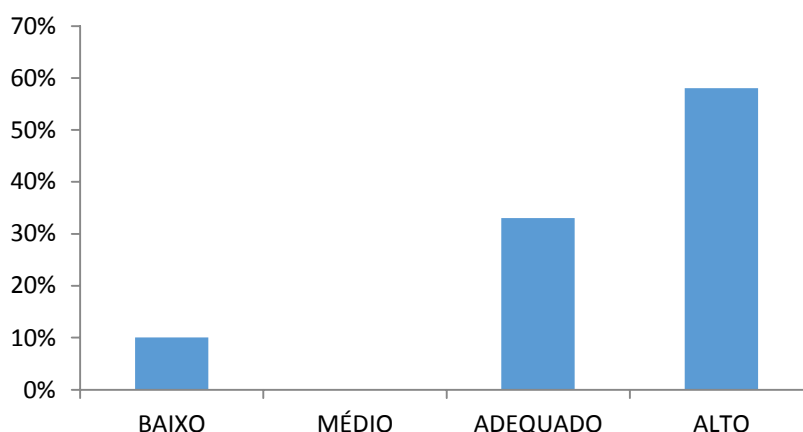


Figura 14: Análise em tabela comparativa do conjunto de 690 amostras de potássio em porcentagem, variando de 0 a 100%.

Observa-se de acordo com tabela comparativa que para as amostras analisadas mais da metade se mostraram com altos teores de potássio. Contudo conforme literatura de Grimme e Nemeth (1979) e Ritchey (1982) o índice de K extraível é objeto de muita controvérsia sobre sua disponibilidade no solo e sua absorção pela planta, já que em quase todas as estruturas da planta o potássio é encontrado, além de que cada cultura absorve esse macronutriente de uma forma diferente.

Assim não é correto dizer que este solo não precisa de adubação potássica e nem que este solo deve ser manejado com fertilizantes, sendo necessária a confluência de outros fatores para bem caracterizar uma área para apta ou não ao cultivo.

De acordo com Silva (2013) em estudo sobre avaliação de curvas espectrais de dois solos em função de atributos físico e químicos na região de Jaciara-MT mostrou que a absorbância diminui com aumento do teor de fósforo, tanto em leituras de campo quanto em

bancada e com aumento do potássio a absorvência diminuiu nas leituras de campo e aumentou nas leituras de bancada.

Comparando os resultados físicos e químicos do potássio com a técnica de *machine learning* foi classificado corretamente em média de 57,5% dos dados de cálcio, para harmonia de dados os resultados foram muito baixos, não chegando a 15% de média mostrando serem necessários ajustes objetivando melhora de F-measure, conforme se observa na figura 15.

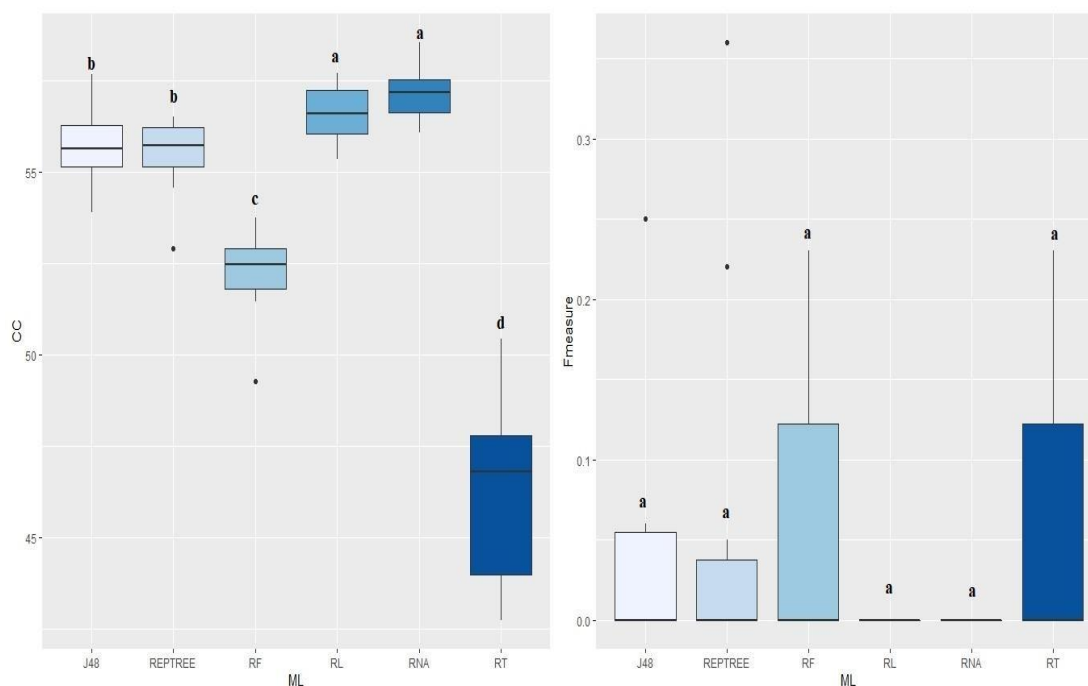


Figura 15: Boxplots para classificações corretas (CC) e F-measure para diferentes modelos de *machine learning* relacionado ao potássio.

Os resultados de correlação expressivamente positiva se deram com RNA (Multilayer Perceptron) e Regressão Logística muito próxima disso, seguido de J48 e Reptree, resultados com correlação fraca se observou em Random Forest e Random Tree. Observam-se poucas variações de resultados dentro de cada técnica, exceto no classificador Random Tree.

Para F-measure, não houve diferença estatísticas entre os classificadores e percebe-se que os dados de entrada não estão harmônicos, porém conforme elucidado anteriormente isso não mostra serem os dados incorretos ou não verdadeiros.

Para obter resultados mais harmônicos o ajuste do bloco de médias ou otimização do limite ideal dos dados da medida F como disserta Lipton et al.(2014) ou acréscimo de outras variáveis de entrada pode levar a índices mais promissores no caso do potássio.

De acordo com teores de areia, silte e argila analisada mais de 50% dos solos estudados são arenosos. (Figura 16).

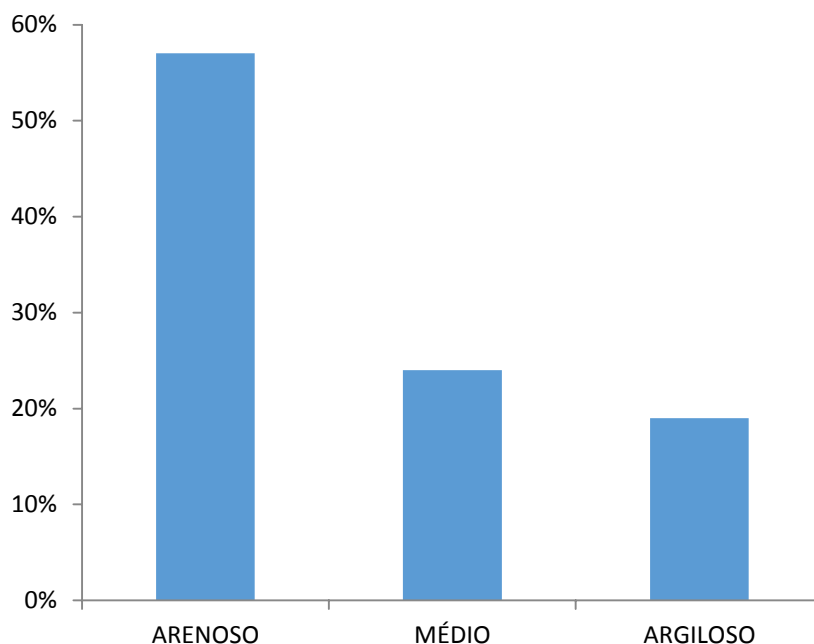


Figura 16: Análise em tabela comparativa do conjunto de 105 amostras de solo e sua relação com granulometria em porcentagem, variando de 0 a 100%.

Foi classificado as amostras em sua maioria como de solo arenoso e de acordo com Centeno (2017) isso não mostra ser inapto ao plantio uma vez que a granulometria isoladamente não está associada a produtividade do solo e conclui em seu estudo que solos arenosos podem ser utilizados para cultivo desde que seja realizada a implantação de sistemas de produção com arranjo de combinação de espécies.

Uma resposta mais rápida dessa análise pode levar economia de tempo e recursos na agricultura, para isso foi feito a comparação com a inteligência computacional para se tentar chegar aos mesmos resultados por meio de um software como mostrado na figura 17.

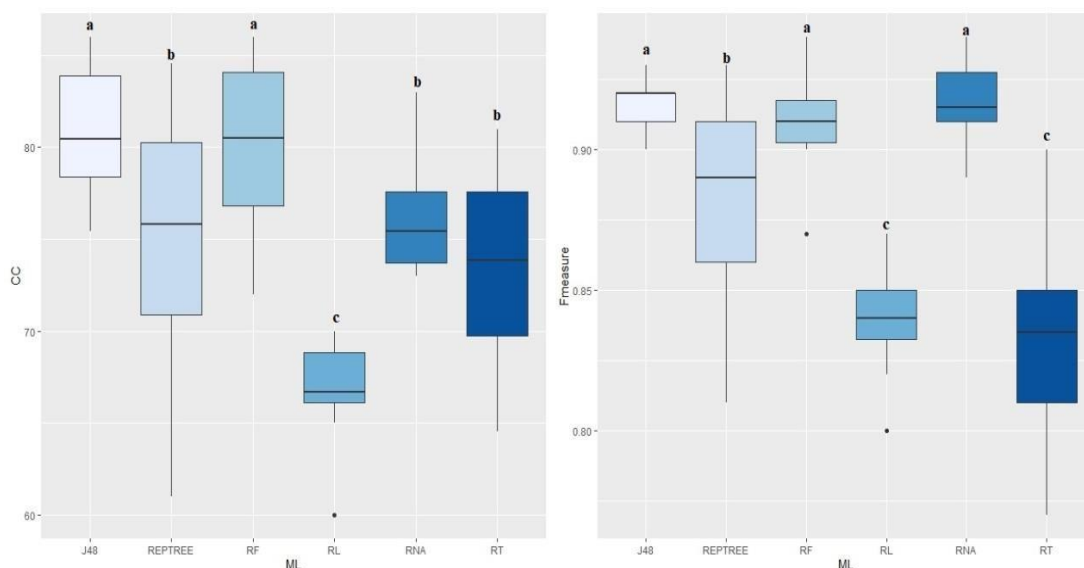


Figura 17: Boxplots para classificações corretas (CC) e F-measure para diferentes modelos de *machine learning* relacionado ao solo e sua granulometria.

A partir de informações numéricas de entrada com comprimentos de onda, areia, silte e argila e o software classificou corretamente o solo em mais de 80% das amostras usando os classificadores Random Forest e J48.

Além disso, o F-measure mostrou serem os dados harmônicos e precisos em mais de 90% dos casos estudados. Desse modo, percebe-se que para classes de solos e sua granulometria a inclusão de dados de areia, silte e argila como variáveis de entrada permite maior precisão de resultados, por esses fatores serem diretamente correlacionados ao fósforo no solo e sua disponibilidade.

Silva et al.(2010) estudaram a variabilidade espacial do fósforo e das frações granulométricas de um latossolo vermelho amarelo e concluíram por uma correlação significativa positiva entre as frações de areia e argila indicando que áreas com maiores concentrações de argila apresentam menor disponibilidade desse nutriente.

A análise do solo de acordo com Sousa et al.(2014) baseia-se na relação existente entre teores de nutrientes no solo e rendimento das culturas, assim como na relação entre doses de fósforo aplicadas e o rendimento em solos com diferentes teores de fósforo para analisar seu maior retorno econômico.

É possível, portanto classificar corretamente a granulometria do solo quando utilizados dados espectrométricos e dados físicos granulares de maneira complementar mantidas as condições de normalidade experimental.

Também foi estudado o fósforo analisado com resina trocadora de íons, conforme se observa na figura 18.

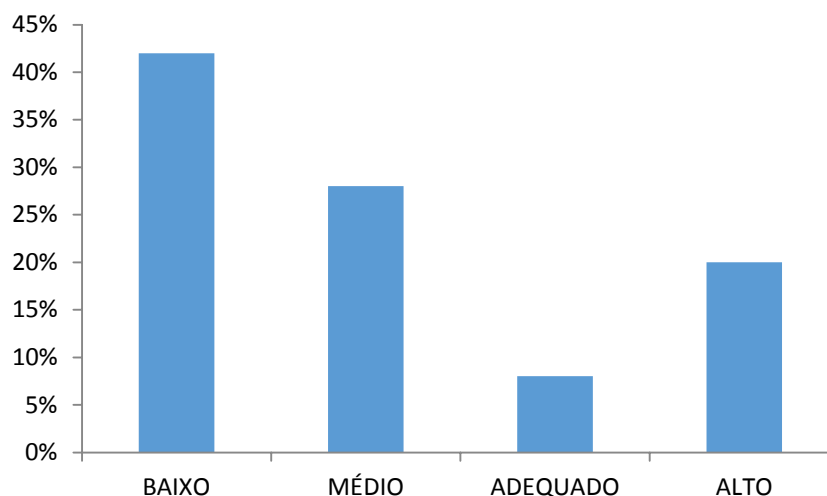


Figura 18: Análise em tabela comparativa do conjunto de 78 amostras de fósforo em porcentagem, variando de 0 a 100%.

Nota-se que o teor de fósforo para essas amostras está em sua maioria baixo, mostrando de acordo com Santos et. al.(2008) a necessidade de aplicação de fertilizantes fosfatados.

Desse modo como disserta Santos et.al.(2008) além de analisar os teores é preciso que se estudem as frações de fósforo no solo, a produtividade das culturas, as formas de aplicação, as alternativas de plantio direto, as relações do fósforo por camadas (0-20 cm), a presença de alumínio trocável e transferência do fósforo para camadas subaquáticas.

Conforme ministrado por Silva et. al.(2013) a resina possui maior tendência a variar em função do perfil do solo, mostrando maior veracidade na estimativa do fósforo disponível para respostas da planta independente da fonte utilizada ou tipo de solo, podendo ser utilizada tanto para solos ácidos como alcalinos.

Em estudo de Silva e Raij (1999) pela busca do melhor método para descobrir o teor de fósforo no solo o método resina apresenta maior coeficiente de determinação para correlação entre P absorvido pelas plantas e P no solo, além de ser considerado superior nas 72 pesquisas realizadas em solos ácidos e alcalinos por esses autores.

Portanto, o nível baixo de fósforo encontrado pelo método resina pode mostrar a realidade de um solo intemperizado e compactado e como as outras classificações mostraram amostras arenosas, este elemento pode ter sido perdido por lixiviação ou difusão, além de

poder apresentar óxido de ferro e alumínio e argilas do grupo caulinita, as quais retêm em sua superfície vários ânions, fenômeno conhecido como adsorção específica. (VAN RAIJ, 1991).

Em estudo de Abatti e Bayer (2019), estes autores estudaram o comportamento do fósforo a partir da aplicação de sete fontes de fósforo (solúveis e solubilidade moderada) e duas estratégias de adubação (de cultura e de sistema) em sistema de plantio direto por três anos em amostras de 0-20 cm, secas ao ar e peneiradas em malha de 2 mm, analisado por P resina, muito semelhante a metodologia deste trabalho e chegaram a conclusão que a utilização de fosfatos solúveis associada a adubação de cultura é preferível em situação de baixo nível de fósforo visando o manejo e reconstrução da fertilidade do solo.

Essa última análise difere da anterior em número de amostras, devido à variabilidade de amostras aptas a serem utilizadas para estudo e nesse caso foi feita a classificação de teor levando em consideração a tabela de Sousa e Lobato (2004) que simula esse nutriente no solo. Ou seja, nesse caso não foram analisados os dados de areia, silte e argila.

Também foi feita a análise do fósforo pelo método resina comparado com o resultado do *machine learning* do software Weka com 42% de classificações corretas e quase 60% para medida F, conforme figura 19.

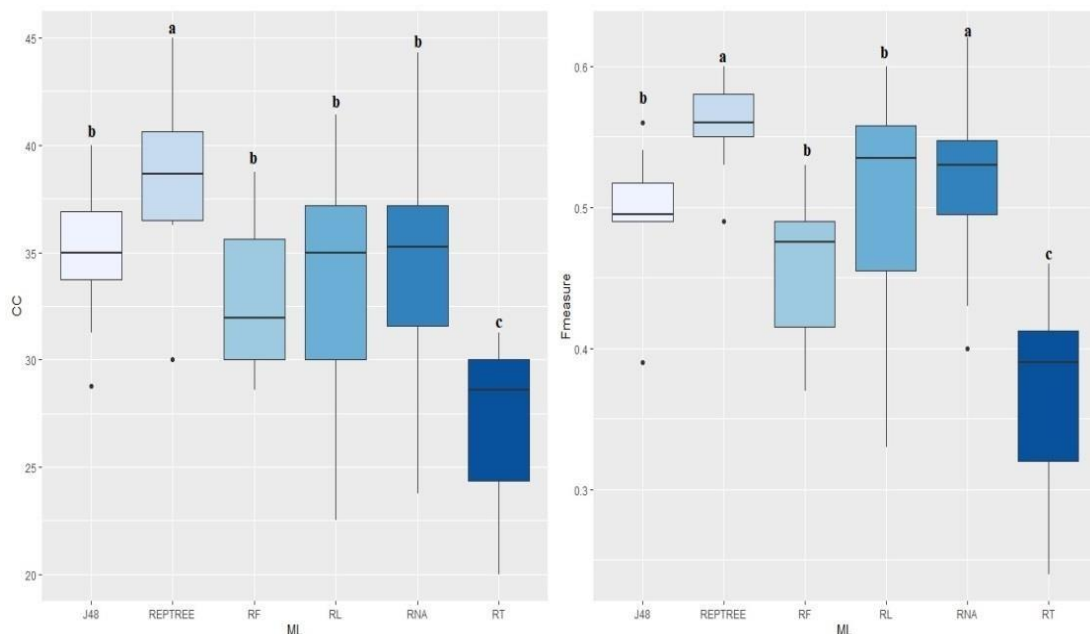


Figura 19: Boxplots para classificações corretas (CC) e F-measure para diferentes modelos de *machine learning* relacionado ao fósforo obtido pelo método resina.

Observa-se que o melhor resultado de classificações corretas foi obtido com o classificador Reptree para esse conjunto de dados, com três diferenças estatísticas entre classificadores.

Para F-measure, os dados não foram inaproveitáveis devido a sua baixa variação no caso do classificador J48 e Reptree. No entanto, com resultados inferiores ao fósforo pelo método mehlich com mais variáveis de entrada, mostrando estarmos caminhando na direção correta em relação à definição de quais variáveis de entrada devem ser utilizadas para chegar a 100 % de correlação positiva do nutriente e seus comprimentos de onda. Em um futuro que poderá trazer resultados exatos independentemente do software utilizado.

A título de exemplo, em recente estudo de Mayrink et al. (2022) buscando correlações entre fósforo e bandas espectrais, estes autores observaram que o fósforo pelo método resina tem forte relação com os comprimentos de onda (410, 460, 560, 645, 705nm) obtendo um R^2 de 0,97. Para tanto, em busca de resultados de precisão e desempenho, eles concluem que a estratégia correta de obtenção de dados do fósforo aliado a um modelo preciso de *machine learning* permitem resultados precisos na predição e análise de fósforo no solo e sua fertilidade.

CONCLUSÃO

O uso do sensor multiespectral COP CIRCLE ACS-470 nas bandas espectrais de 532, 550, 670, 700, 730 e 760 nm (do verde a borda do vermelho) apresenta correlação positiva com macronutrientes do solo.

Os melhores classificadores preditivos foram o Random Forest para classificações corretas e F-measure em estudo da granulometria seguido na média dos macronutrientes e seus teores pelo algoritmo Regressão Logística.

Também se conclui que acrescentando de maneira complementar os dados físicos de areia, silte, argila e comprimentos de onda como variável avaliada, para classificação da granulometria dos solos em médio, argiloso e arenoso como variável a ser predita usando o classificador Random Forest proporcionam precisão e harmonia nos resultados em mais de 90% das análises estudadas.

REFERÊNCIAS

ALBUQUERQUE, Jackson Adriano et al. **Aplicação de calcário e fósforo e estabilidade da estrutura de um solo ácido**. Revista Brasileira de ciência do Solo, Viçosa. 27, p. 799-806, 2003.

ALVES, Marcelo Rodrigo. **Múltiplas Técnicas no mapeamento digital de solos**. 2008.160p. Tese. (Escola Superior de Agricultura Luís de Queiroz)-Universidade de São Paulo. Piracicaba, 2008.

ANDA, A. N. **Principais Indicadores Do Setor De Fertilizantes**. Disponível em:https://anda.org.br/wpcontent/uploads/2022/05/Principais_Indicadores_2022.pdf. Acesso em 09/03/2023.

BANZATTO, David Arioaldo. **Experimentação agrícola**. 4 ed. Jaboticabal: Funep, 2006.

BARBOSA, Mauro Cezar et al. **Desempenho agrônômico e componentes da produção de cultivares de soja em duas épocas de semeadura no arenito caiua**. Semina:Ciencias Agrarias, v. 34, n. 3, p. 945–960, 2013.

BOOF, Leonardo. **Sustentabilidade: o que é e o que não é**. Editora Vozes; 5ª edição, 200 p, 2016.

BORGES, Wander Luis Barbosa et al. **Plantas de cobertura para o noroeste do estado de São Paulo**. Ciência Rural, v. 45, n. 5, p. 799–805, 2015.

BORGES, Wander Luis Barbosa et al. **Produção de soja e milho cultivados sobre diferentes coberturas**. Revista Ciência Agronômica, Fortaleza-CE,v. 46, n. 1, p. 89–98,jan-mar. 2015.

BRANDT, Eduardo André et al. **Desempenho agrônômico de soja em função da sucessão de culturas em sistema de plantio direto**. Ciência e Agrotecnologia, v. 30, n. 5, p. 869–874, 2008.

BRANNSTROM, C. et al. **Land change in the Brazilian Savanna (Cerrado), 1986-2002: Comparative analysis and implications for land-use policy**. Land Use Policy, v. 25, n. 04, p. 579–595, 2008.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília, DF: Centro Gráfico, 1988.

BRESSAN, Wellington e VASCONCELLOS, Carlos Alberto. **Alterações morfológicas no sistema radicular do milho induzidas por fungos micorrízicos e fósforo**. Pesquisa Agropecuaria Brasileira, v. 37, n. 4, p. 509–517, 2002.

BRESSAN, Wellington e VASCONCELLOS, Carlos Alberto. **Alterações morfológicas no sistema radicular do milho induzidas por fungos micorrízicos e fósforo**. Pesquisa Agropecuaria Brasileira,Brasília, v. 37, n. 4, p. 509–517, 2002.

CALERO, João Cravidão. **Monitoração do ciclo produtivo de uvas Fernão Pires biofortificadas em zinco com recurso de imagens multiespectrais-Estudo de caso.**Dissertação.68p. Faculdade de Ciências e Tecnologias da Universidade NOVA de Lisboa.Portugal. 2019.

CAMPBELL, Patrícia Morais da Matta. **Mapeamento Digital de solos e predição de atributos utilizando Machine Learning e lógica fuzzy na bacia do Ribeirão Inhaúma, Iconha/ES.**2017.Tese.227p.Programa de Pós graduação em Ciencias Ambientais e Florestais-Instituto Florestas. Universidade Federal Rural do Rio de Janeiro-UFRRJ.Rio de Janeiro,2017.

CARNEIRO, Franciele M. et al. **Correlations among vegetation indices and peanut traits during different crop development stages.** Engenharia Agrícola,Jaboticabal, v. 39, n. spe, p. 33–40, sep. 2019.

CASTRO, Paulo R.C. e VIEIRA,E.L.**Bioreguladores e bioestimulantes na cultura do milho. Milho: estratégias de manejo para alta produtividade.**Esalq/USP/LPV,p 99-115,Piracicaba,2003.

CENTENO,Luana Nunes. et. al. **Textura do solo: conceitos e aplicações em solos arenosos.**Revista brasileira de Engenharia e sustentabilidade..Pelotas-Rs.v.4,n.1,p 31-37,jul.2017.

CHIODEROLI, Carlos Alessandro et al. **Consórcio de Urochloas com milho em sistema plantio direto.** Ciência Rural, Santa Maria, v. 42, n. 10, p. 1804–1810,out. 2012.

CONAB, C. N. DE A. **Acompanhamento da safra brasileira de grãos.** Disponível em: <http://www.conab.gov.br> Acesso em: 18 nov. 2021.

CORREIA, J.R. Et. Al.**Cerrado: Correção do solo e adubação.** Brasília-DF: EMBRAPA,2 ed, Informação Tecnológica,416p,2004.

COUTINHO, Eluã Ramos et al. **Utilização de Técnicas de inteligência computacional na predição de dados meteorológicos.** Revista Brasileira de Meteorologia, v.31, n.1,24-36, Rio de Janeiro, 2016.

COUTO, Rafael Da Rosa et al. **Phosphorus fractions in soil with a long history of organic waste and mineral fertilizer addition.** Bragantia, v. 76, n. 1, p. 155–166, Rio Grande do Sul,2017.

CRUZ, André Camargo et al. **Fertilizantes Organominerais De Resíduos Do Agronegócio : Avaliação Do Potencial Econômico Brasileiro**. Indústria Química: BNDES Setorial, v. 45, p. 137–187, Rio de Janeiro, 2017.

DAMOLIN, Ricardo Simão et. al. **Relação entre constituintes do solo e seu comportamento espectral**. Ciência Rural, Santa Maria, v.35, n.2, p.481-489, mar-abr, 2005.

DEMATTÊ, José Alexandre Melo. **Relações entre dados espectrais e características físicas, químicas e mineralógicas de solos desenvolvidos de rochas eruptivas**. Tese. 265p. Esalq Piracicaba, 1995.

DEVETYAROV, Dmitry e NOURETDINOV, Ilia. **Artificial Intelligence Applications and Inovations: Prediction with confidence based on a Random Forest Classifier**. Lárnaca-Cyprus. 403p. 2010.

DONAGEMMA, Guilherme Kangussu et al. **Characterization, agricultural potential, and perspectives for the management of light soils in Brazil**. Pesquisa Agropecuária Brasileira, Brasília v. 51, n. 9, p. 1003–1020, set. 2016.

FERNANDES, Adão B. da Luz et al. **Agrominerais para o Brasil**. CETEM/MCT, Rio de Janeiro, 2010.

FIGUEIREDO, C. C. DE; RESCK, D. V. S.; CARNEIRO, M. A. C. **Labile and stable fractions of soil organic matter under management systems and native cerrado**. Revista Brasileira de Ciência do Solo, v. 34, n. 3, p. 907–916, 2010.

FILHO, Dalson Brito Figueiredo e JÚNIOR, José Alexandre Silva. **Desvendando os mistérios do Coeficiente de correlação de Pearson (r)**. Revista Política Hoje, Vol. 18, n.1, Pernambuco, 2009.

GUVEN, Duyusen e Akinci, Gorkem. **Comparison of acid digestion techniques to determine heavy metals in sediment and soil samples**. Grazi University Journal of Science. 24: 29-34. Buca, Izmir, Turkey, 2011.

HORST, W. J. et al. **Agronomic measures for increasing P availability to crops**. Plant and Soil, v. 237, n. 1, p. 211–223, Netherlands, 2001.

KAMH, M. et al. **Mobilization of soil and fertilizer phosphate by cover crops**. Plant and Soil, v. 211, n. 1, p. 19–27, Netherlands, 1999.

KLIEMANN, H. J. e LIMA, D. V. **Eficiência agronômica de fosfatos naturais e sua influência no fósforo disponível em dois solos de cerrado**. Pesquisa Agropecuária Tropical, Goiania-GO, v. 31, n. 2, p. 111–119, 2007.

LOUREMBAM, D.; LASKAR, S. e MUKHERJEE, S. **Framework for an optical sensor system for monitoring of soil nitrogen and tailoring soil pH**. Journal of Optics (India), v. 47, n. 2, p. 180–194, 2018.

MACEDO, Manoel Claudio Motta. **Integração lavoura e pecuária : o estado da arte e inovações tecnológicas**. Revista Brasileira de Zootecnia, Viçosa-MG, v. 38, p. 133–146, 2009.

MAROCO, João. **Análise Estatística- Como utilização do SPSS**. 2 ed. Edições Sílabo. Lisboa, 2003.

MARTINS Adalvan Manoel et al. **Relação entre índice SPAD, teores de clorofila extraível e nitrogênio na folha da batata**. Universidade Federal de Viçosa. Viçosa-MG, 2010.

MARTINS, R.S. O. et al. **Recomendação de parâmetros para COIN-OR Branch and Cut**. Simpósio Brasileiro de Pesquisa Operacional. Anais do XLVIII SBPO. Vitoria-Es, 27-30 setembro de 2016.

MATOS, Manoel Antonio. **Manual Operacional para Regressão Linear**. Faculdade de Engenharia da Universidade do Porto, Portugal, 1995.

MATTAR NETO, João Augusto. **Metodologia científica na era da informática**. 2 ed. São Paulo: Saraiva, 2005.

MAYRINK, G.O. et al. **Soil phosphorus test using a low cost spectrophotometer and machine learning**. Revista Engenharia Agrícola. V.42, n.6, Universidade Federal de Viçosa, Viçosa-MG, Brasil. 2022.

MOLIN, José Paulo. **Agricultura de precisão**. 1ed. Oficina de textos, São Paulo, 2015.

NASCIMENTO, Yasmim Cordeiro. **Proteção Legal Nacional do Solo e dos Lençóis freáticos**. JurisWay: revista eletrônica on line. 2015.

NOVAIS, Roberto Ferreira et al. **Fertilidade do solo**. Viçosa, MG; Sociedade Brasileira de Ciência do Solo, Minas Gerais, 2007.

NUNES, Rafael De Sousa et al. **Sistemas de manejo e os estoques de carbono e nitrogênio em latossolo de cerrado com a sucessão soja-milho**. Revista Brasileira de Ciência do Solo, Viçosa, v. 35, n. 4, p. 1407–1419, 2011.

ONG, C. C. H.; CUDAHY, T. J. **Mapping contaminated soils: Using remotely-sensed hyperspectral data to predict pH**. European Journal of Soil Science, v. 65, n. 6, p. 897–906, 2014.

OSCO, L.P. et al. **Leaf Nitrogen Concentration and plant height prediction for maize using UAV-based multispectral imagery and machine learning techniques.** Remote sensing-article. Out. 2020.

PACHECO, Leonardo Pereira et al. **Ciclagem de nutrientes por plantas de cobertura e produtividade de soja e arroz em plantio direto.** Pesquisa Agropecuária Brasileira, v. 48, n. 9, p. 1228–1236, Goiás, 2008.

PASSOS, Alexandre Martins Abdao et al. **Cultivares de soja em sucessão ao trigo nos sistemas convencional e plantio direto.** Agrarian, v. 8, n. 27, p. 30–38, 2015. Disponível em: <https://ojs.ufgd.edu.br/index.php/agrarian/article/view/3015>. Acesso em: 8 mar. 2023.

PAVINATTO, Paulo Sergio et al. **Acúmulo de fósforo no solo em áreas agrícolas no Brasil: diagnóstico atual e potencialidades futuras.** Informações Agronômicas: Nutrição de Plantas Ciência e Tecnologia. Número 9, Março, Piracicaba-SP, 2021.

PEEL, M. C.; FINLAYSON, B. L.; MCMAHON, T. A. **Updated world map of the Köppen-Geiger climate classification.** Hydrology and Earth System Sciences, v. 11, n. 5, p. 1633–1644, Victoria-Australia, 11 out. 2007.

PENATTI, Fabio Eduardo et al. **Avaliação dos riscos ambientais causados pela disposição incorreta de resíduos de laboratórios.** Geografia: Ensino e Pesquisa, v.15, n.1, p. 43-52, jan/abr. Santa Maria-RS, 2011.

PETTER, F. A. et al. **Doses e épocas de aplicação de potássio no desempenho agrônomo do milho no cerrado piauiense.** Comunicata Scientiae, v. 7, n. 3, p. 372–382, 2016.

PETTER, F. A. et al. **Productivity and Yield Components of Soybeans under Dose and Potassium Application Period in Piauí Savannah.** Journal of Agricultural Science, v. 4, n. 8, p. 173–181, 2012.

POPPIEL, R. R.; PAIVA, A. F. DA S.; DEMATTÊ, J. A. M. **Bridging the gap between soil spectroscopy and traditional laboratory: Insights for routine implementation.** Geoderma, v. 425, p. 116029, nov. 2022.

QUEIROS, Leonardo Ribeiro et al. **Análise das possibilidades e tendências do uso das tecnologias da informação e comunicação em Agricultura de Precisão.** Ed. Agricultura de precisão: resultados de um novo olhar. Brasília, DF: EMBRAPA, 2014.

QUEVEDO, Claudia Maria Gomes de. **As atividades do homem e a evolução do fósforo no meio ambiente.** Dissertação de mestrado. Universidade de São Paulo. 229p. São Paulo-SP, 2009.

RAIJ, Bernardo Van; ANDRADE, J.C. De; CANTARELLA, H; QUAGGIO, J.A. **Análise Química para avaliação da fertilidade de solos tropicais**. Instituto Agrônomo, Campinas-SP. 285p. 2001.

RAMIREZ-LOPEZ, L. et al. **The spectrum-based learner: A new local approach for modeling soil vis-NIR spectra of complex datasets**. Geoderma, v. 195–196, p. 268–279, 2013.

RATKE, Rafael Felipe. et al. **Agricultural potential and soil use based on the pedogenetic properties of soils from the cerrado-caatinga transition**. Semina-Ciências Agrárias, v. 41, p. 1119–1134, 2020p. 1119–1134, Londrina-PR, 2020.

RATKE, Rafael Felipe. et al. **Different limestone particle sizes for soil acidity correction, Ca and Mg supply and corn yield**. Comunicata Scientiae, v. 9, n. 2, p. 175–184, Piauí, 2018.

RIGO, Amanda Zolet et al. **Phosphorus fractions in soil with organic and mineral fertilization in integrated crop-livestock system**. Revista Brasileira de Ciencia do Solo, v. 43, p. 1–15, Viçosa-MG, 2019.

RODIGHERI, G. **Modelagem da produtividade primaria líquida utilizando dados coletados de sensores remotos: avaliação de impactos e perdas de área agrícola**. Dissertação (mestrado). 87 p. Universidade Federal do Rio Grande do Sul, Porto Alegre-RS, 2020.

RODRIGUES, Giancarlo. **Seleção de bandas espectrais apoiada pela metaheurística PSO para predição do teor de alumínio trocável de amostras de solo**. Dissertação de mestrado. 70p. UEPG. Ponta Grossa-PR, 2018.

ROSSETTO, Raffaella et al. **Fertilidade do solo e nutrição de plantas: calagem paracana-de-açúcar e sua interação com adubação potássica açúcar e sua interação com adubação potássica**. Bragantia, Campinas, v. 63, n. 1, p. 105–119, 2004.

SANTOS, D.R. et al. **Fatores que afetam a disponibilidade do fósforo e o manejo da adubação fosfatada em solos sob plantio direto**. Revista Ciencia Rural, Santa Maria, v. 38, n. 2, p. 576–586, mar-abr, 2008.

SANTOS, Jhonathans Jesus. **Modelagem de atributos químicos e espectrais dos solos de Feira de Santana-BA**. Dissertação. 121p. -Universidade Estadual de Feira de Santana-UEFS. Bahia. 2019.

SASAKI, Yutaka. **The truth of the F-measure**. School of Computer Science. University of Manchester, october, Manchester, 2007.

SCHENFERT, T. A. et al. **Lime and gypsum in the implantation no-till system promote the nutrition and yield of soybean**. Journal of Plant Nutrition, v. 43, n. 4, p. 1–14, 23 dez. 2019.

SHARIFIFAR, A. et al. **Evaluating a low-cost portable <scp>NIR</scp> spectrometer for the prediction of soil organic and total carbon using different calibration models**. Soil Use and Management, v. 35, n. 4, p. 607–616, dez. 2019.

SILVA, Ygor Jacques Agra Bezerra da. **Métodos de digestão de solos e fertilizantes para análises de metais pesados**, Dissertação. 41p. Universidade Federal Rural de Pernambuco-PE. Recife-PE, 2012.

SOARES, Valdeir Moraes et al. **Calibração e validação de modelos para estimativa de atributos químicos dos solos com uso da espectroscopia**. Aprosoja: Agro cientista. Mato Grosso, 2014.

SOUSA, D. M. G.; LOBATO, E. **Cerrado: correção do solo e adubação**. 2^a ed. Brasília: Embrapa Informação Tecnológica, 2004.

SOUZA, H. B.; BAIO, F. H. R.; NEVES, D. C. **Using passive and active multispectral sensors on the correlation with the phenological indices of cotton**. Engenharia Agrícola, Jaboticabal, v. 37, p. 782–789, 2017.

SWASTI, Singhal e MONIKA Jena. **A study on weka tool for data preprocessing, classification and clustering**. International journal of innovative technology and exploring engineering. Volume 2, Issue 6. May, 2013.

TEIXEIRA, Paulo Cezar et al. **Manual de métodos de análises de solos**. 3^a ed. Brasília: Embrapa, 2017.

TEODORO, Larissa Pereira Ribeiro et al. **A Random Forest ranking approach to predict yield in maize with uav-based vegetation spectral indices**. Elsevier-article. Vol 178. Nov. 2020.

TEODORO, Larissa Pereira Ribeiro et al. **Predicting days to maturity, plant height, and grain yield in soybean: a machine and deep learning approach using multispectral data**. Remote sensing-article. Out. 2021.

TERRA, F. S. et al. **Spectral libraries for quantitative analyses of tropical Brazilian soils: Comparing vis-NIR and mid-IR reflectance data.** Geoderma, v. 255–256, p. 81–93, 2015.

TIECHER, T. et al. **Cover crops affecting soil phosphorus dynamics in Brazilian highly weathered soils.** Cover Crops: Cultivation, Management and Benefits, v. 55, n. 55, p. 23–52, 2015.

TORRES, Pedro et al. **Aplicação de Técnicas de Machine Learning para classificação da aptidão dos solos para o regadio.** Instituto Politécnico de Castelo Branco-Escola Superior Agrária. Portugal. Castelo Branco, 2022.

URANO, E. O. M. et al. **Determinação de teores ótimos de nutrientes em soja pelos métodos chance matemática, sistema integrado de diagnose e recomendação e diagnose da composição nutricional.** Revista Brasileira de Ciência do Solo, v. 31, n. 1, p. 63–72, 2007.

VASQUES, Gustavo M. et al. **Medição in situ e predição de elementos químicos do solo usando espectrômetro de fluorescência de raios-X,** Congresso Brasileiro de Geoquímica, 16, artigo em anais do Congresso, Niterói- RJ, 2017.

VEGRO, C. L. R. **Mercado de fertilizantes.** Disponível em: <www.iaea.gov.br/ftp/iea/AIA/AIA-16-2018.pdf>. Acesso em: 4 mar. 2019.

VELHO, S. R. K. et al. **Nível de Maturidade Tecnológica: uma sistemática para ordenar tecnologias.** Parcerias Estratégicas, Brasília-DF, v. 22, n. 45, p. 119–140, 2017.

VOLPATO, Gilson Luiz. **Guia prático para redação científica.** 286p. Botucatu: Best Writing, 2015.

ZANDONÁ, Renan Ricardo et al. **Gesso e calcário aumentam a produtividade e amenizam o efeito do déficit hídrico em milho e soja.** Pesquisa Agropecuária Tropical, v. 45, n. 2, p. 128–137, 2015.

ZHAOYU, Zhai et al. **Decision support systems for agriculture 4.0: Survey and challenges.** Elsevier, Computers and electronics in Agriculture, March, 2020.

ZHAOYU, Zhai et al. **Decision support systems for agriculture 4.0: Survey and challenges.** Computers and Electronics in Agriculture: Elsevier, vol. 170, 2020.

ZUFFO, Alan Mario et al. **Analysis of growth and agronomic characteristics of off-season**

corn grown insuccession with soybean and submitted to nitrogen doses. Revista em Agronegócio e Meio Ambiente, Maringá-PR, v. 14, n. 1, p. e007960, 2021.