



MATHEUS STERZO NILSSON

**ESPECTRÔMETRO PORTÁTIL DE
FLUORESCÊNCIA DE RAIOS-X (pXRF) NA PREDIÇÃO E
MAPEAMENTO DIGITAL DE ATRIBUTOS DO SOLO**

**LAVRAS
2019**

Matheus Sterzo Nilsson

**Espectrômetro portátil de fluorescência de raios-X
(pXRF) na predição e mapeamento digital de atributos do
solo**

Monografia apresentada à
Universidade Federal de Lavras,
como parte das exigências do
Curso de Agronomia, para a
obtenção do título de Bacharel.

Prof. Dr. Sérgio Henrique Godinho Silva
Orientador

Prof. Dr. Michele Duarte de Menezes
Coorientadora

**LAVRAS MG
2019**

MATHEUS STERZO NILSSON

**ESPECTRÔMETRO PORTÁTIL DE FLUORESCÊNCIA DE RAIOS-X (pXRF)
NA PREDIÇÃO E MAPEAMENTO DIGITAL DE ATRIBUTOS DO SOLO**

**PORTABLE X-RAY FLUORESCENCE SPECTROMETER (pXRF) FOR THE
PREDICTION AND DIGITAL MAPPING OF SOIL PROPERTIES**

Monografia apresentada à
Universidade Federal de Lavras,
como parte das exigências do
Programa de Graduação em
Agronomia, para a obtenção do
título de Bacharel

APROVADA em 21 de maio de 2019

Dr. Sérgio Henrique Godinho Silva
Dra. Michele Duarte de Menezes
M.Sc. Renata Andrade
M.Sc. Marcelo Henrique Procópio Pelegrino

UFLA
UFLA
UFLA
UFLA

Prof. Dr. Sérgio Henrique Godinho Silva
(Orientador)

Prof. Dra. Michele Duarte de Menezes
(Coorientadora)

**Lavras MG
2019**

Este trabalho é dedicado à minha família,
minha namorada e principalmente à
minha avó Iracema, pelo amor, confiança
e apoio de todos depositados em mim.

Agradecimentos:

Agradeço a Deus por todas as graças concedidas durante toda minha vida, a minha saúde e principalmente a oportunidade de realizar este curso.

Agradeço a minha família por toda a confiança e apoio depositados em mim. Principalmente ao meu pai Gustavo, minha mãe Elaine, meus irmãos Vinícius e Bianca. Como também minha amada avó que não está mais conosco, por todos seus sábios conselhos.

Agradeço a minha namorada Rafaela por toda a ternura e companheirismo que me propiciou ao longo desse tempo juntos.

Agradeço ao meu orientador Sérgio e pela oportunidade de elaborar e concretizar este trabalho, contando com toda sua experiência e conhecimentos.

Agradeço ao núcleo de estudos G-Óleo principalmente aos professores Pedro e Fraga, e aos colegas de trabalho Camilo, Vitor, Pedro e Vinícius por toda a experiência que pude passar que me ajudaram em muito em ser um profissional melhor.

Agradeço ao núcleo de estudos NECS pelas oportunidades que me ajudaram a melhorar ainda mais quem eu sou. Principalmente aos colegas Maria Vitória, Mariany, Bruna por todo esse momento muito bom vivido. Além dos tutores.

Agradeço a todos os professores pelos quais tive contato, pela excelente oportunidade de aprender e me tornar um profissional com um ótimo conhecimento.

A todos meus sinceros agradecimentos!

“Não é possível libertar um povo, sem antes livrar-se da escravidão de si mesmo.

Sem esta, qualquer outra será insignificante, efêmera e ilusória, quando não um retrocesso.

Cada pessoa tem sua caminhada própria.

Faça o melhor que puder.

Seja o melhor que puder.

O resultado virá na proporção do seu esforço.

Compreenda que, se não veio, cumpre a você (a mim e a todos) modificar suas (nossas) técnicas, visões, verdades, etc.

Nossa caminhada somente termina no túmulo.

Ou até mesmo além...”

Mahatma Gandhi

RESUMO

A determinação dos atributos do solo são importantes indicadores das características e qualidade do solo. A acurada mensuração dos atributos do solo permite uma correta classificação de classes do solo, além de auxiliar no manejo agrônômico ao fornecer subsídios para recomendações de corretivos e fertilizantes. Nesse sentido, o equipamento portátil de fluorescência de raios-X (pXRF) surge como ferramenta apta a proporcionar inferências sobre os elementos bem como suas quantidades presentes nas amostras do solo, com baixo custo e sem destruição das amostras. Contudo, os estudos utilizando esta ferramenta para esta finalidade ainda são escassos no Brasil. Desta forma, são necessários estudos sobre o melhor modo de utilização do equipamento bem como a construção de modelos preditivos cada vez mais confiáveis. Portanto, o presente trabalho teve como objetivo realizar a predição espacial de pH, fósforo remanescente (P-rem), matéria orgânica (MO) e soma de bases (SB) a partir dos resultados do pXRF, através de regressão linear múltipla *Stepwise* (SMLR) e Random Forest (RF). E produzir mapas a partir desses atributos. Foram coletadas 90 amostras simples do horizonte A no campus da Universidade Federal de Lavras MG, das quais foram obtidos resultados analíticos de pH, P-rem, MO e SB como também a leitura dos teores dos elementos químicos pelo pXRF. Com base nos teores determinados pelo pXRF, resultados de laboratório foram preditos a partir do ajuste dos modelos de SMRL e RF. O método que apresentou resultado mais acurado, para cada atributo, foi escolhido para a confecção do mapa. As equações de validação da SMLR apresentaram maior coeficiente de correlação (R^2) para os atributos de SB e P-Rem, sendo estes de 0,41, 0,15 respectivamente e menores para MO 0,13 e pH 0,07. Para o RF os resultados os melhores resultados de R^2 foram para P-Rem 0,66 e SB 0,55 e menores para MO e pH o quais foram de 0,47 e 0,26 respectivamente. A SB e pH tiveram a variável CaO como sendo a de maior importância. Os atributos que tiveram maior R^2 na predição foram os que resultaram em mapas de maior R^2 , sendo eles: P-rem com 0,66 e SB com 0,54, já os menores valores de validação foram para matéria orgânica, com 0,47 seguida de pH, com 0,28. O RF apresentou melhores predições no que se diz respeito à construção dos modelos preditivos. Os dados do pXRF em conjunto com o RF pode ser utilizado para auxiliar na determinação dos atributos do solo, possibilitando, portanto, na confecção de mapas digitais.

Palavras chave: Variabilidade espacial. Sensor próximo. Predição espacial. Atributos do solo. Mapeamento digital.

ABSTRACT

The attributes of soil are important indicators of its quality. Their accurate measurement supports correct classifications of portions of land and may support the agricultural management, since it provides information for the recommendations of concealers and fertilizers. In this scenario, the portable equipment of X-ray fluorescence (pXRF) arises as a low cost and non-destructive tool able to provide inferences about the elements in the sample and their respective quantities. However, studies using this tool are still scarce in Brazil, so that investigations are necessary to elucidate the best use of such equipment as well as to build increasingly trustful predictive models. The present work aimed to predict pH, remaining phosphorous (P-rem), organic matter (OM) and sum of basis (SB) based on pXRF results by means of stepwise multiple linear regression (SMLR) and Random Forest (RF), as well as to produce maps of such attributes. Overall, 90 simple samples were collected from horizon A in the Federal University of Lavras MG, from which analytical results of pH, P-rem, OM, SB and pXRF readings were obtained. Results in the laboratory were predicted from adjustments of SMLR and RF models based on the contents determined by pXRF. The method with most accurate result for each attribute was chosen to produce the map. The SMLR validation equations showed a higher correlation coefficient (R^2) for the attributes of SB and P-Rem, which were 0.41, 0.15 respectively and lower for MO 0.13 and pH 0.07. For RF the best R^2 results were for P-Rem 0.66 and SB 0.55 and lower for MO and pH which were 0.47 and 0.26 respectively. For SB, CaO was the most important variable. Attributes with the greatest R^2 in the prediction resulted in maps with the greatest R^2 , namely: P-rem (0.66) and SB (0.54). On the other hand, the lowest validation values were recorded for MO (0.47) and pH (0.28). RF presented the best prediction. pXRF may be used with RF to help on determining the soil attributes and to produce digital maps.

Keywords: Spatial variability. Near sensor. Spatial prediction. Digital mapping. Attributes of soil.

Lista de figuras

Figura 1	Localização geográfica da área de estudo.....	22
Figura 2	Distribuição dos pontos amostrais de validação e modelagem na área de estudo.....	23
Figura 3	Gráficos de variáveis mais importantes obtidas para espectrômetro de fluorescência de raios X na predição de propriedades do solo pelo Random Forest.....	30
Figura 4	Gráficos de validação dos atributos do solo sendo que para os modelos de Random Forest os dados se apresentam à direita e de Regressões Múltiplas à esquerda.....	31
Figura 5	Mapas preditos para as propriedades do solo na camada de 0-10 cm, pelo método de Random Forest. Sendo que MO: matéria orgânica, SB: soma de bases e P-rem: fósforo remanescente.....	34

Lista de tabelas

Tabela 1	Estatística descritiva das propriedades do solo da validação e modelagem.	26
Tabela 2	Análise descritiva dos dados (ppm) obtidos pelo pXRF.....	27
Tabela 3	Equações de regressão múltiplas obtidas pelo método stepwise.....	28
Tabela 4	Erro do quadrado médio (EQM) e Variância explicada (Var exp %) do modelo Randon forest.....	29

Sumário:

1	Introdução	14
2	Referencial Teórico.....	16
2.1	Espacialização de atributos químicos no contexto do mapeamento digital.....	16
2.2	Tipos de sensores utilizados no mapeamento digital de solos.	17
2.3	Espectrômetro de fluorescência de raios X: Aplicações e conceitos.....	18
2.4	Métodos estatísticos usados na construção dos modelos de predição.....	19
2.5	Considerações gerais.....	20
3.	Materiais e métodos.....	21
3.1	Descrição da área de estudo.....	21
3.2	Coleta e análise de solo.....	22
3.3	Preparo e análise com o equipamento pXRF.....	22
3.4	Análise dos dados.....	23
3.4.1	Análise por Regressões lineares múltiplas.....	23
3.4.2	Análises pelo Random Forest.....	24
3.4.3	Acurácia entre os modelos.....	24
3.5	Espacialização das propriedades do solo utilizando dados do pXRF.....	25
4	Resultados e discussões.....	26
4.1	Estatística descritiva.....	26
4.2	Modelagem das propriedades através de Regressões múltiplas.....	28

4.3	Modelagem das propriedades através do Random Forest e importância das variáveis.....	30
4.4	Validação dos modelos gerados por regressão linear múltipla por stepwise e pelo método random forest.....	30
4.5	Mapeamento dos atributos e validação dos mapas.....	32
5	Conclusão.....	34
6	Referencial Bibliográfico.....	36

Abreviações

CV	Coefficiente de Variação
GPS	Sistemas de Posicionamento Global
MDS	Mapeamento Digital de Solos
ME	Erro Médio (do inglês Mean Error)
MO	Matéria Orgânica
MSE	Erro do Quadrado Médio (do inglês Mean Square Error)
P-rem	Fósforo remanescente
pXRF	Espectrômetro portátil de fluorescência de raios-X
R ²	Coefficiente de Correlação
RF	Floresta Aleatória (do inglês Random Forest)
RML	Regressão Linear Múltipla
RMSE	Raiz Quadrada do Erro Médio (do inglês Root Mean Square Error)
SB	Soma de Bases
SIG	Sistema de Informação Geográfica
SMLR	Regressão Linear Múltipla Stepwise (do inglês Stepwise Multiple Linear Regression)
TFSA	Terra Fina Seca ao Ar

1. Introdução

O crescente investimento em novas tecnologias, como também sua utilização pelos produtores rurais, foi um dos fatores que proporcionaram a consolidação do Brasil como um dos grandes exportadores de produtos agropecuários (RESENDE et al., 2016). Para esse progresso decisivo nos sistemas de produções agrícolas, foram necessários conhecimentos gerados por instituições de pesquisa e universidades visando à sustentabilidade e caracterização dos solos.

As vantagens em se aprender a respeito do solo são inúmeras, já que este é o principal substrato para produção de alimentos (RESENDE, 2007). Neste contexto, se tornam de extrema importância estudos sobre a variabilidade espacial dos atributos das diferentes classes de solos, como também o mapeamento de tais atributos, utilizando-se para este último ferramentas de sistemas de informações geográficas (SIG) como softwares e técnicas de geoprocessamento. O desenvolvimento de tais tecnologias possibilitou o trabalho com a aplicação de adubos a taxas variáveis, a qual tem como intuito a correta alocação de doses de corretivos e fertilizantes usados na produção agrícola, proporcionando assim a racionalização dos insumos e menores impactos ambientais, além de auxiliar na condução da cultura e aumentar a produtividade.

Uma análise de solo com resultados confiáveis é um dos aspectos mais importantes ligados ao manejo da fertilidade do solo (LOPES E GUILHERME, 2007), incluindo a correta recomendação e utilização eficiente de corretivos e fertilizantes. Essas análises são instrumentos básicos e insubstituíveis para o sucesso da lavoura, indispensáveis nas tomadas de decisões (ALVES, 2008). No entanto, os métodos utilizados para análises químicas e físicas de rotina em laboratórios são trabalhosos, demorados e necessitam de descarte apropriado de resíduos, caso o contrário, serão poluentes do meio ambiente (VISCARRA e McBRATNEY, 1998).

Neste cenário, com o propósito de se estimar resultados de atributos de solos em maior escala e em menor tempo, a área de ciência do solo conta com diferentes tecnologias. Dentre elas, podemos citar a utilização de sensores proximais de infravermelho-próximo (NIR), sensores utilizados no sensoriamento remoto e mais recentemente a espectrometria portátil de fluorescência de raios-X (pXRF). Estas técnicas proporcionam maior agilidade e eficiência, além de não modificarem as propriedades das amostras, que poderão ser utilizadas em outras análises (SILVA et al, 2017).

Os estudos sobre a utilização do pXRF são ainda poucos para solos brasileiros (SILVA et al, 2018), havendo deste modo necessidade de mais conhecimentos sobre a eficiência desta

técnica e a determinação de parâmetros para seu uso, visando o reconhecimento de padrões, e consequentemente, permitir a utilização destes dados por futuros usuários.

Deste modo, o presente trabalho teve por finalidade realizar a predição espacial de pH, P-rem, MO e SB a partir dos resultados do pXRF, através de Regressão Linear Múltipla *stepwise* e Randon Forest.

2. Revisão de literatura

2.1 A predição espacial de atributos do solo no contexto do mapeamento digital de solos (MDS)

A predição espacial de atributos do solo é amplamente utilizada como auxílio na diferenciação e delimitação de glebas, assim como utilizado na agricultura de precisão para melhor alocação de recursos agrícolas e estabelecimento de zonas de manejo. Oliveira et al. (2008) compararam a recomendação de calagem e adubação utilizando a técnica convencional de amostragem do solo (zigue-zague) versus a análise espacial, e deste modo, puderam identificar zonas de déficit ou excesso de calagem e adubação na área, que não poderiam ser definidas pelo método convencional de amostragem.

O método convencional de levantamento de solos utiliza etapas de amostragem que se caracteriza principalmente por ter caráter mental e qualitativo. Neste método, o pedólogo utiliza de pesquisas bibliográficas, análise de mapas, exame de fotos, imagens e/ou modelos digitais de elevação. A partir destas ferramentas, em conjunto com experiências passadas ao trabalhar em regiões semelhantes e análises dos perfis de solo, o profissional identifica o solo dominante de cada área. Porém, esse modelo apresenta algumas limitações, como: dificuldades em reproduzir os modelos mentais para outros pedólogos (principalmente os iniciantes) e alto despendido de tempo e recursos (LEPSCH, 2017).

As limitações do método convencional de levantamento de solos podem ser diminuídas de diferentes formas, entre elas, o desenvolvimento de projetos com auxílio de SIG. Os SIG trabalham com combinação de dados espaciais que auxiliam na descrição e análises de interações, de modo a fazer previsões através de modelos prospectivos empíricos e fornecer apoio para a definição de classes, unidades ou sítios de interesse no solo (NOLASCO-CARVALHO, C. C.; FRANCA-ROCHA, W.; UCHA, 2009). O conjunto destas técnicas e dados ajudam na elaboração de mapas digitais do solo, ou seja, a junção de diferentes dados por meio de modelagem matemática de diversos atributos e classes de solo (EMBRAPA, 2018).

Os mapas digitais do solo são uma realidade cada vez mais presente no cenário agrícola e ambiental brasileiro. A utilização desses mapas conjuntamente com outras técnicas propiciou o surgimento de uma nova área chamada de Pedometria, que foi definida por Lagacherie e McBratney (2007) como a criação e a população de sistemas de informação espacial de solos por meio de modelos numéricos, visando inferir as variações espaciais e temporais de classes e propriedades do solo, a partir de observações, conhecimento e dados de covariáveis ambientais

relacionados. Machado et al. (2007) observaram, o mapeamento digital de determinados atributos do solo só se tornou factível graças ao avanço dos computadores, do surgimento dos sistemas de posicionamento geográfico (GPS) e programas geoestatísticos.

A técnica de MDS é relativamente nova no Brasil, já que uma das primeiras publicações foi realizada em 2006 (GIASSON et al., 2006). Em trabalho de revisão, Ten Caten et al. (2012) descreveram que a utilização mais tardia do MDS estaria ligada a diversos fatores como: (i) à demorada disponibilização de softwares e hardwares no país; (ii) ao conservadorismo de muitos pedólogos que relutam em utilizar sistemas automatizados capazes de contribuir para o mapeamento de solos; (iii) à carência de pessoal qualificado para o emprego da tecnologia da informação na ciência do solo; (iv) à popularização tardia no Brasil no uso de tecnologias como sistema de posicionamento global e sensoriamento remoto.

2.2 Tipos de sensores utilizados no mapeamento digital de solos

Os princípios físicos que propiciam a mensuração das características químicas e físicas do solo são classificados em elétricos e eletromagnéticos, ópticos e radiométricos, acústicos, pneumáticos, mecânicos e eletroquímicos (EITELWEIN, 2017). Os espectrômetros são sensores que utilizam dos princípios ópticos e eletromagnéticos, funcionando basicamente através da emissão e absorção de certos comprimentos de onda da luz. Ocorre do seguinte modo: a intensidade de luz refletida e espalhada por um corpo atingido por uma certa radiação eletromagnética faz com que ocorra uma resposta (reflectância) diferenciada de acordo com o objeto analisado (CLARK, 1999; HUNK, 1980; VISCARRA ROSSEL et al., 2006). Portanto dependendo das características do solo, como, por exemplo, tamanho das partículas ou material de origem, as amostras analisadas podem ter a caracterização simultânea de outros elementos do solo (EITELWEIN, 2017).

A espectroscopia de reflectância com finalidade de determinação remota das propriedades químicas, físicas e de composição mineralógica dos solos, teve início em 1970, período em que os pesquisadores Hunt e Salisur deram início à formulação de um banco de dados espectrais (Chabrilat et al, 2013). Atualmente diversos são os trabalhos realizados com a tecnologia de espectroscopia de reflectância, principalmente com relação a determinações de textura do solo e de matéria orgânica das quais, Junior et al. (2008) revelaram uma variação da reflectância do solo devido aos diferentes atributos, como areia, argila, matéria orgânica e cor, mostrando ser possível o mapeamento de tais variações. Já Demattê et al. (2018), dispendo de um banco espectral de uma área já conhecida, criou um modelo de quantificação de textura do solo com erro menor de 30% para uma área externa, mostrando possibilidade de extrapolação de dados para uma região geográfica diferente.

Uma outra classe de sensores que auxiliam na determinação de atributos do solo são os eletrodos íons-seletivos que utilizam membranas para detectar a atividade de íons como hidrogênio, potássio e nitrato (GEBBERS E ADAMCHUCK, 2010). Os mecanismos de funcionamento dos eletrodos íons-seletivos são baseados na medição da diferença de potencial entre a membrana seletiva e a referência, assim essa diferença é convertida para atividade íon específico da solução (TALIBUDEEN, 1991). Viscarra Rossel et al. (2005) reporta ser necessário um determinado tempo para que o sensor atinja o equilíbrio entre o solo e a solução do solo. Adamchuck et al. (2005) descreveram esse tempo como sendo de 5 a 15 segundos.

Com a finalidade de diminuir o custo e tornar mais acessível a criação de uma densa malha amostral de uma área, além de mapear o pH para a correção de solo, Adamchuck et al. (1999) criaram o primeiro equipamento de mensuração de pH em movimento. O equipamento consiste em uma plataforma que coleta amostras a uma profundidade de 0,05 a 0,20 m. Esse equipamento foi um avanço, pois as leituras eram realizadas em média a cada dez segundos diretamente do solo por dois eletrodos íon-seletivos, sem a necessidade das amostras estarem preparadas, obtendo como resultados uma correlação linear de 0,83 entre os valores observados pelo equipamento e obtidos em laboratório.

Atualmente diversas linhas de pesquisas têm sido desenvolvidas utilizando o mesmo princípio dos eletrodos íons positivos. Sasaki et al. (2014), dispoendo de eletrodos de íons seletivos, determinaram o pH e teores de potássio, gerando mapas, e compararam com mapas gerados através de amostras de laboratório, chegando a um coeficiente de correlação de 0,56 para pH e 0,61 para potássio, classificando esses resultados como satisfatórios. Porém, Silva (2013), também mapeando pH, comparou a eficiência de diferentes eletrodos íons-seletivos com resultados de laboratório. Após fazer as análises estatísticas, verificou erros maiores que os determinados como limite e concluiu que a umidade do solo interferiu nos resultados. Assim, a pesquisadora sugeriu a utilização de um sensor de umidade em conjunto, como também o uso de ferramentas de padronização e homogeneização das amostras.

2.3 Espectrômetro de fluorescência de raios X: Aplicações e conceitos

O espectrômetro de fluorescência de raios-X também é um equipamento que recorre a energia eletromagnética com a intenção de fazer análises sobre amostras analisadas. A identificação do elemento, bem como sua quantificação (ppm) na amostra, ocorre pelo princípio de indução de fluorescência dos átomos da unidade analisada. Este princípio funciona da seguinte maneira: um feixe de raios-X é expedido sobre os átomos, fazendo com que elétrons de orbitais mais internos se movam para orbitais mais externos. Em seguida, os elétrons desse átomo retornam ao orbital inicial emitindo energia. Essa energia é conhecida como fluorescência e é

mensurada pelo aparelho através de um espectro eletromagnético (SHARMA et al., 2014; WEINDORF et al 2014). O comprimento de onda é diretamente dependente das características de cada átomo, permitindo, deste modo, sua classificação e quantificação da maioria dos elementos presentes da amostra (KALNICK e SINGHVI, 2001). Essa análise tem como característica ser rápida, além de proporcionar várias informações e necessitar de pouco preparo das amostras (ZHU et al., 2011).

Diversas pesquisas têm sido desenvolvidas visando o aproveitamento das qualidades do pXRF em várias áreas do conhecimento, como: mineração, química, metalurgia, arqueologia, ciência do solo e agronomia (WEINDORF et al. 2014). Suas utilizações na área da agricultura são relativamente novas e demonstram uma grande capacidade para expansão pois tem sido utilizadas para estimativa de vários atributos do solo como pH (SHARMA et al., 2014), capacidade de troca de cátions (SHARMA et al., 2015), teores de ferro e manganês (TEIXEIRA; VASQUES, 2017), além de sulfato de cálcio (WEINDORF et al. 2013). O' Rouker et al. (2016) testaram de maneira conjunta a integração de dados do pXRF e reflectância vis-NIR na predições de atributos do solo, obtendo uma melhora no desempenho das predições. Experimentos com o equipamento em solos brasileiro são novos como também técnicas de preparo das amostras para análises do pXRF. Nesse sentido, Silva et al. (2018), comparando os teores de Fe_2O_3 , Al_2O_3 , SiO_2 , TiO_2 e P_2O_5 obtidos pelo pXRF com o da análise de digestão por ácido sulfúrico, testaram também diferentes métodos de preparo das amostras, chegando à conclusão que diferentes preparos ocasionam diferenças nos resultados analisados pelo pXRF.

As pesquisas com pXRF vão além de aplicações com o solo. Ribeiro et al. (2017), em artigo de revisão, verificaram trabalhos relacionados com a medição da composição elementar em tecidos vegetais e grãos usando pXRF (McLAREN; GUPPY; TIGHE, 2012; PALTRIDGE et al., 2012). Nesse sentido, Towet, Shephred e Drake (2015), estudando a composição de folhas, encontraram alta correlação para os elementos Mg, P, S, K, Ca e Mn com os obtidos pelo método tradicional de digestão ácida obtidos pela técnica de ICP-OES.

2.4 Métodos estatísticos usados na construção dos modelos de predição

Alguns tipos de pesquisas necessitam de análises de modelos multivariados, as quais envolvem a relação entre múltiplas variáveis explicativas e, em alguns casos, múltiplas variáveis dependentes. A regressão linear múltipla (RLM) surge como um conjunto de técnicas capazes de analisar o efeito exercido por duas ou mais variáveis independentes sobre uma variável dependente (TABACHNICK e FIDELL, 1996). O resultado da RMSE é uma equação, a qual representa o melhor resultado das somas das variáveis preditoras para uma variável a ser predita (ABBAD e TORRES 2002).

Com o objetivo de analisar resultados estimados através de Fe_2O_3 utilizando dados de reflectância de imagem do satélite TM-Landsat 5 e de sensor IRIS, Fiorro et al. (2003) usaram regressões múltiplas e compararam esses resultados com os resultados de análises de laboratório, chegando a resultados de R^2 de 0,92 para o resultado do sensor radiométrico e de 0,93 para o global. Já Eitelwein, Trevisan e Colaço, (2015), conseguiram coeficientes de correlação para predição de argila e silte de respectivamente 0,87 e 0,85 também usando equações de regressão obtidas através de reflectância espectral (Vis-NIR).

Elaborada por Breiman (2001), como extensão do programa CART (Classification and Regression Trees) os modelos gerados pelo algoritmo RF é uma técnica não paramétrica, que visa buscar o melhor desempenho de predição do modelos. Se diferem da RMSE por se basearem na construção de grandes quantidades de árvores de decisões colocados para fora do subconjuntos de dados, utilizando um treinamento único e definido (BASGALUPP, 2010). No caso utilizando RF para regressão, o resultado final consiste da média dos resultados de todas as árvores (Breiman, 2001; Cutler et al., 2007).

Esse treinamento é realizado usando um meta-algoritmo com o intuito de melhorar a classificação e a regressão de modelos de acordo com a estabilidade e a precisão da classificação, pelo método conhecido como *bagging*. O método *bagging* consiste na geração de amostragens aleatórias dos dados para cada árvore de decisão, que são chamadas de amostras bootstrap. O bootstrap é uma técnica de amostragem com reposição, a partir do conjunto de treinamento inicial, no qual são selecionadas amostras aleatoriamente para um novo subconjunto de treinamento (OLIVEIRA, 2015).

As árvores de decisões são construídas através de casos extraídos a partir do treinamento dos conjuntos de dados originais, esses conjuntos de dados então construídos serão utilizados na elaboração de uma árvore que compõe o RF.

Se tratando do uso do modelo Radom Forest no mapeamento digital de solos, uma das aplicações foi utilizada por Pinheiro, (2015), que em trabalho de tese, utilizou o RF no mapeamento de classes e atributos dos solos em uma bacia hidrográfica, a qual apresenta grande variação de condições de paisagem, no Estado do Rio de Janeiro. A autora comparou a predição da composição da textura, na camada superficial do solo através de regressões lineares múltiplas e árvores de regressão. As análises indicaram desempenho superior do algoritmo de árvores de regressão, para todos os atributos testados (areia, silte, argila), utilizando dados harmonizados ou originais.

2.5 Considerações gerais

Os resultados de pesquisas têm demonstrado que a utilização de diversos sensores para a realização de mapeamento digital de solo é promissora, em várias regiões do mundo sob diversas situações. Na criação de modelos de predição dos atributos do solo cada método estatístico possui suas vantagens e desvantagens, interferindo diretamente na qualidade das predições e conseqüentemente na acurácia dos mapas. Sendo assim, surgem perspectivas quanto à utilização dos sensores na substituição das análises de solo para auxiliar no mapeamento digital, sendo necessário estudos mais aprofundados sobre o melhor modo de utilização dos sensores bem como a construção de modelos preditivos cada vez mais confiáveis.

3. Materiais e métodos

3.1 Descrição da área de estudo

O estudo foi desenvolvido no campus da Universidade Federal de Lavras, localizado no município de Lavras, em Minas Gerais, Brasil, situada entre as latitudes de $21^{\circ}13'07''\text{S}$ e $21^{\circ}14'27''$ longitudes $44^{\circ}59'24''\text{W}$ e $44^{\circ}57'29''\text{O}$, possuindo uma área de 314,5 hectares. Tem como as classes de solo: Latossolo Vermelho, Latossolo Vermelho-Amarelo, Cambissolo Háplico, Argissolo Amarelo, Argissolo Vermelho, Argissolo Vermelho-Amarelo, e Nitossolo Vermelho, (CURI et al., 2017).

Segundo a classificação de Köppen, o clima predominante é o Cwa com temperatura média anual de $20,4^{\circ}\text{C}$ e precipitação média anual de 1460mm estando em uma faixa de altitude de 841 e 980 m (DANTA et al.,2007). A área de estudo é caracterizada por um clima tropical de altitude com a ocorrência de verões quentes e úmidos e invernos frescos e secos.

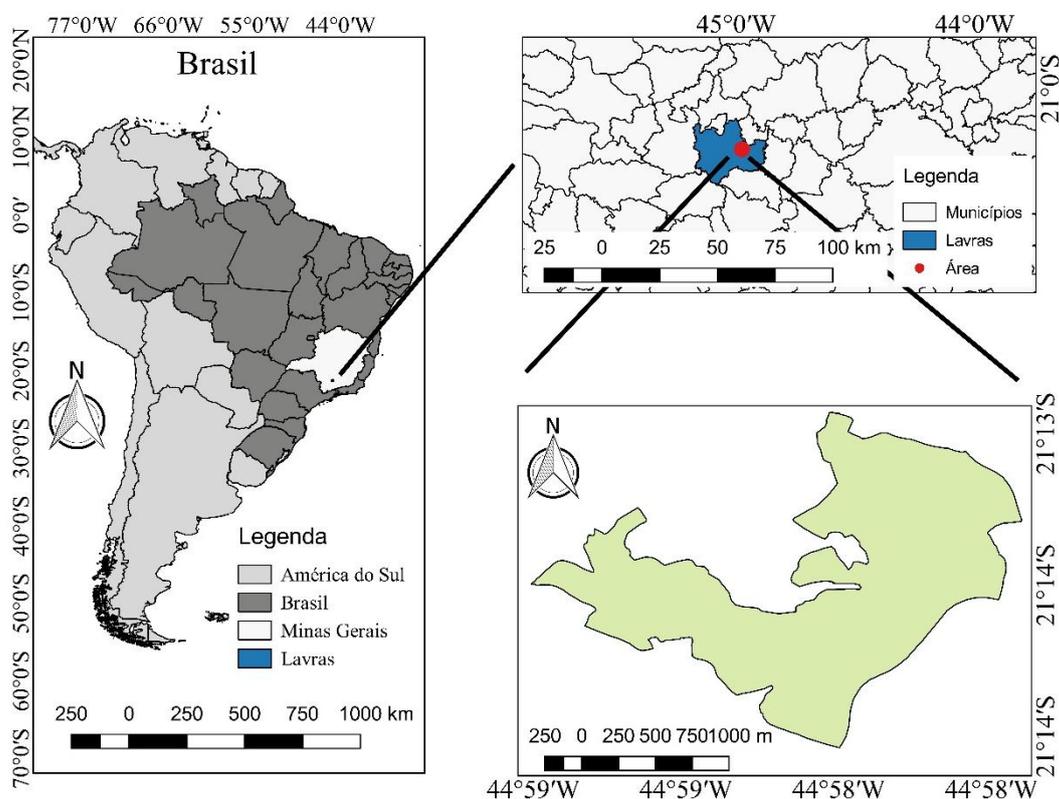


Figura 1: Localização geográfica da área de estudo.

Fonte: Do autor (2019)

3.2 Coleta e análise de solo:

Para a realização deste experimento foram coletadas 90 amostras simples do horizonte A dos solos da área, formando uma malha regular. Essas amostras foram secas em estufa com circulação forçada e logo após foram homogeneizadas e peneiradas em peneira de malha 2 mm (terra fina seca ao ar).

Em seguida, as amostras foram submetidas a análises laboratoriais para a determinação dos seguintes atributos do solo: teor de MO (WALKLEY, BLACK, 1934), pH em água, P-rem (ALVAREZ, FONSECA, 1990), teores trocáveis de Ca^{2+} , Mg^{2+} e Na^+ (McLEAN et al., 1958) e K disponível extraído com Mehlich-1.

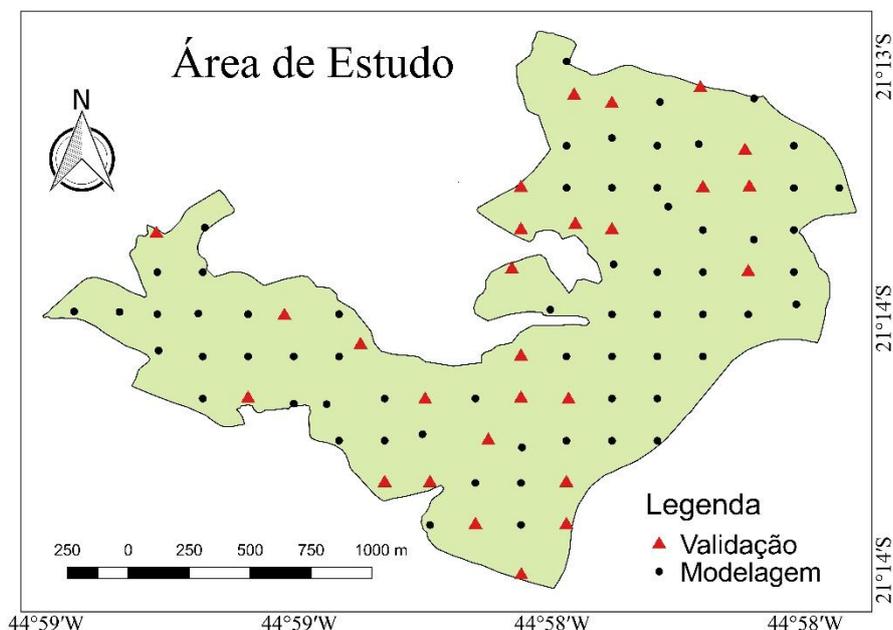


Figura 2: Pontos amostrais de modelagem e validação da área.

Fonte: Do autor (2019)

3.3 Análises com o pXRF

O aparelho utilizado foi um pXRF Bruker modelo S1 Titan LE foi utilizado para a determinação dos teores dos elementos químicos nas amostras de TFSA, usadas para a predição dos resultados em laboratório. Esse equipamento apresenta tubos de Rh com raios-X de 50 kV e 100 μA . Além disso, tem como vantagem ser portátil, podendo ser utilizado em campo, uma vez

que pesa cerca de 1,5 kg, tem dimensões de 25cm x 28cm x 9 cm e utiliza baterias de lítio recarregáveis como suprimento de energia. O total de 20 elementos ficaram acima do limite de detecção: Al, Si, P, Cl, K, Ca, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, As, Rb, Zr, Sr, Nb, Pb, que foram utilizados na criação dos modelos. Todas as leituras pelo equipamento foram realizadas em triplicata. O software utilizado pelo equipamento se trata do GeoChem, na configuração dual soil (Trace), durante 60 segundos, utilizando dois feixes de raios-X.

Para aferir a acurácia do equipamento, foi realizada a avaliação de dois materiais certificados pelo Instituto Nacional de Padrões e Tecnologia (NIST) como referência: 2710a e 2711a, e também foi realizada uma análise de amostra padrão (check sample – CS) fornecida pela fabricante do equipamento. Desta forma, foi calculado o valor de recuperação do pXRF para cada elemento através da fórmula: valor de recuperação = teor obtido pelo pXRF/teor certificado. Os valores de recuperação (2710a/2711a/CS) para os elementos usados neste estudo foram (valores iguais a zero indicam ou que o equipamento não detectou o elemento ou que o elemento não apresenta teor certificado nas amostras padrão): Al₂O₃ (0,80/0,71/0,91), As (0,86/0,68/0), CaO (0,39/0,46/0), Cl (0/0/0), Cr (0/1,11/0), Cu (0,82/0,85/0,89), Fe (0,74/0,80/0,87), K₂O (0,55/0,51/0,88), Mn (0,70/0,68/0,83), Ni (0/1,15/0,96), P₂O₅ (3,85/5,4/0), Pb (1,18/1,05/1,06), Rb (0,92/0,89/0), SiO₂ (0,64/0,51/0,87), Sr (1,11/0,83/0), Ti (0,78/0,75/0), V (0,78/1,15/0), Y (0/0/0), Zn (0,95/0,78/0), Zr (1,12/0/0).

3.4 Análises estatísticas dos dados

As análises de dados foram separadas em quatro etapas: a) estatística descritiva dos dados; b) ajuste de modelos para predições dos atributos estudados, empregando para este, dois métodos estatísticos: RMSE *stepwise* e outro com base em RF; c) avaliação dos resultados de predição e comparação com os resultados de laboratório, escolhendo assim o melhor método para realização dos mapas de atributos; d) predição espacial e validação com os resultados de laboratório. Os resultados das análises laboratoriais e do pXRF foram submetidos à estatística descritiva para verificação da posição (mínimo, máxima e média) e a dispersão (desvio padrão e coeficiente de variação) dos dados.

As análises das amostras foram divididas, de forma aleatória, em amostras a serem utilizadas para construção dos modelos de predição e para validação dos modelos conforme Figura 2. Foram utilizadas 63 amostras (70%) na construção do modelo de predição dos resultados e 27 amostras (30%) para validação.

Para as determinações de SB, P-rem, pH e MO do presente estudo por RMSE, utilizou-se o método *backward*, pelo software Sigma Plot. Este método faz com que todas as variáveis independentes sejam adicionadas de uma só vez, e depois retiradas uma a uma, de forma que o ajuste do modelo (R²) apresente o melhor resultado.

Para a predição de atributos utilizando o RF, foi utilizado o software R Studio pacote randomForest (LIAW e WIENER, 2015). Utilizando 1000 como o número de árvores (ntrees), 5 como o número de variáveis em cada nó (nodesize) e um terço do número total de amostras usadas para validação, conforme sugerido por Liaw e Wiener (2002).

Os ajustes do RF resulta na média quadrada dos resíduos (MSEoob), na porcentagem da variância explicada pelo modelo e a importância de todas as variáveis do modelo na predição dos dados. O MSEoob é calculado quando, para cada interação, apenas algumas variáveis preditoras são usadas para gerar uma árvore. O MSEoob é calculado pela equação 1. A importância das variáveis é o resultado da média da redução da precisão na previsão, pois quando uma variável é deixada de fora outras são incluídas. Assim, se uma variável é removida, mais o erro da previsão aumenta, o que significa que a precisão da predição diminui, mais importante a variável é para o ajuste do modelo. (BREIMAN, 2001; LIAW; WIENER, 2002).

$$(1) \text{MSEoob} = \frac{1}{n} \sum_{i=1}^n [y_i - y_i^{\text{oob}}]^2$$

Os parâmetros estatísticos utilizados para validação dos modelos de RML e RF foram RMSE calculado pela equação 2, EM calculado pela equação 3, além do coeficiente de determinação R².

$$(2) \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_i - m_i)^2} \quad (3) \text{ME} = \frac{1}{n} \sum_{i=1}^n (e_i - m_i)^2$$

n= número de observações, e_i= valores estimados pelo modelo, m_i= valores observados pelos resultados de laboratório.

Para esses cálculos foram utilizados dados independentes que não estavam compondo as equações de modelagem, compondo 30% do número total de amostras. O modelo com melhor performance para cada atributo de solo a ser predito foi escolhido de acordo com: maior R² e menores RMSE e EM.

3.5 Predição espacial dos atributos do solo

Após estabelecido o melhor modelo de predição dos atributos do solo, esses foram preditos para toda a área de estudo. Para tanto, os dados dos elementos químicos obtidos pelo pXRF foram espacializados para toda a área de estudo, utilizando o método *Inverse Distance Weighting* (IDW), no software ArcGis 10.3. Isso foi necessário pelo fato dos teores dos elementos obtidos pelo pXRF só estarem disponíveis para o local em que as amostras foram coletadas, enquanto que, para a criação dos mapas, tais informações são necessárias para toda a área de estudo.

Assim como os dados de modelagem, os mapas preditos da área também foram submetidos à validação através do cálculo do R^2 , RMSE e EM utilizando as amostras de validação.

4. Resultados e discussões

4.1 Estatística descritiva

Os resultados descritivos das análises tanto para modelagem quanto para validação dos resultados de laboratório estão inseridos na Tabela 1. Os valores de SB, MO e P Remanescente apresentaram coeficiente de variação (CV) considerado alto na classificação de Pimentel Gomes (1985) para ensaios agrícolas, já para o pH, o CV de uma maneira geral apresenta dispersão média.

Tabela 1: Estatística descritiva de atributos do solo pH, soma de bases (SB)(cmol dm^{-3}), matéria orgânica (MO) (dag kg^{-1}) e P-remanescente (P-rem) (mg dm^{-3}) nos conjuntos de modelagem e validação.

Parâmetro	Modelagem				Validação			
	SB	M.O	P-Rem	pH	SB	M.O	P-Rem	pH
Min	1,4	2,1	3,9	4,6	1,8	3,1	3,2	4,0
Max	26,6	14,1	35,1	7,2	13,3	11,2	36,1	7,0
Méd	6,2	6,2	22,8	5,8	6,9	6,3	22,7	5,8
Desv p	3,6	2,4	7,0	0,6	3,2	2,2	7,9	0,7
CV(%)	58,4	38,3	30,9	10,7	46,4	35,6	34,7	11,7

Min: Valor mínimo, Max: valor máximo, Méd: média dos valores, Desv p: desvio padrão da média e CV(%): Coeficiente de variação.

A área de estudo apresenta diferenças no com relação ao uso e ocupação do solo, tendo áreas agrícolas caracterizadas com culturas anuais, pastagens, cafezais, pomar, plantio de eucaliptos, candeias e pinus. Além de áreas de vegetação natural como floresta e mata ciliar, além de áreas de solo exposto (FERREIRA, 2013). Essa diferença no uso do solo explicam os altos coeficientes de variação de SB, pois em determinadas áreas se faz correção do solo com uso da calagem aumentando portanto os valores de Ca e adubação com fertilizantes diversos.

Os valores do pH são considerados bons segundo o manual de recomendação para uso de corretivos e fertilizantes em Minas Gerais (RIBEIRO; GUIMARÃES; ALVAREZ V, 1999), devido ao fato da correção do solo ser realizado em algumas áreas, cujos valores máximos e mínimos denotam a constante entre as áreas com pH corrigidos e não corrigidos. Os teores de P-rem variam de 3,2 e 36,1 mg/dm³ e têm média de 22,7 mg. Essas diferenças são ocasionadas devido a área possuir diversas classes de solos interferindo portanto na capacidade de adsorção dos fosfatos, já que essa adsorção é dependente da quantidade de matéria orgânica, da textura e da mineralogia da fração argila (NOVAIS; SMYTH, 1999). A classificação da matéria orgânica é considerada como boa segundo o manual de recomendação para uso de corretivos e fertilizantes em Minas Gerais (RIBEIRO; GUIMARÃES; ALVAREZ V, 1999).

O pXRF determinou 20 elementos ao todo em todas as amostras, cujas análises descritivas estão presentes na Tabela 2. Os elementos que apresentaram as maiores concentrações foram Al, Si, P, Cl e K. Segundo Lutgens e Tarbuck (2000), o elemento que está em maior quantidade na crosta terrestre é o oxigênio (46,6%), seguido por silício (27,7%), alumínio (8,1%), ferro (5%) e cálcio (3,6%). Dessa maneira se explica o fato do alumínio seguido do silício terem aparecido em maior quantidade.

Tabela 2 - Análise descritiva dos dados (ppm) obtidos pelo pXRF.

Atributo	Desvio padrão	Mínimo	Máximo	Média
Al ₂ O ₃	32240	103111	245797	178631
SiO ₂	74750	119532	414752	267958
P ₂ O ₅	510	0	2671	708
Cl	150	95	935	523
K ₂ O	4421	801	34101	3217
CaO	7538	219	66013	5481
Ti	3654	2428	18308	9583
V	112	0	462	114
Cr	881	0	5910	583
Mn	397	129	2153	617
Fe	44844	15902	187076	80872
Ni	118	0	750	715
Cu	24	10	140	36

Zn	21	15	123	41
As	2	0	9	1
Rb	17	0	138	11
Sr	19	0	158	15
Zr	49	98	332	216
Nb	7	0	25	7
Pb	15	0	53	16

4.2 Modelagem das propriedades do solo

As equações referentes aos atributos pH, MO e SB foram as que apresentaram maior número de variáveis sendo 8, 6 e 4 respectivamente (Tabela 3). A menor equação é a referente ao atributo P-rem, com 3 variáveis. Todas as equações apresentaram em sua composição a variável Ca, com exceção a equação do pH. A equação de pH não possui cálcio em sua fórmula, mesmo este sendo um atributo importante na relação com a elevação ou redução de seus valores (GARGANTINI; MELLO; ARZOLLA, 1982). Já na equação para SB, a variável CaO foi utilizada, o que era esperado uma vez que o teor de Ca^{2+} no solo é utilizado no cálculo de SB. Sharma et al (2014) encontraram um conjunto de variáveis diferentes para a determinação de pH, dentre eles o Ca, além de variações diferentes nas quantidades de variáveis, o que mostra que as variáveis podem mudar de acordo com o conjunto de dados e com a região de estudo.

Tabela 3: Equações de regressão linear múltiplas obtidas pelo método *stepwise*. Com SB (cmol dm³), M.O (dag kg⁻¹) e P-rem (mg dm³).

Equação	R ²
SB= 2,58000 + 0,000731CaO + 0,00397 V + 0,0156Ni - 0,0324Cu	0,95
P-rem= 3,338 + 0,0000601SiO ₂ + 0,001CaO - 0,0208Ni	0,85
M.O= 20,931 - 0,0000212 Al ₂ O ₃ - 0,0000307SiO ₂ + 0,000228CaO + 0,00854V - 0,0000644Fe + 0,00819Ni	0,75
pH= 3,063 + 0,00000717SiO ₂ - 0,000206K ₂ O + 0,00314V - 0,00029Cr - 0,0011Mn + 0,0000121Fe + 0,0705Rb + 0,0283Sr	0,67

Com relação aos valores de R² o P-rem e SB, ambos tiveram valores de R² acima de 0,80. Já para pH e MO, os valores foram de 0,75 e 0,67. Esses resultados de modelos de regressão linear múltipla *stepwise* são melhores que os encontrados por Silva et al. (2017) em áreas do cerrado brasileiro para os mesmos atributos e com o mesmo método de análise.

Os valores de RMSE da modelagem estimados pelo RF estão apresentados na Tabela 4. O RMSE é um estimador que tem a finalidade de informar a qualidade de um modelo preditivo, pois ele calcula a diferença dos quadrados médios dos valores estimados e os valores reais (LEHMANN, CASELHA, 1998). Observando os dados da tabela, verifica-se que os modelos que apresentaram melhor ajuste (menor erro e maior variância explicada) foram para pH, MO, seguidos de SB e P-Rem.

Tabela 4: Tendência do MSE e variância explicada (Var %) do modelo Random forest.

Parâmetro	SB	MO	P-Rem	pH
MSE	6,2	4	20,7	0,3
Variância explicada (%)	52,7	26,7	57,7	19,5

A escolha de uma variável importante para cada atributo através do RF se deve ao fato do algoritmo construir diversas árvores de decisões através dos dados originais, a partir cada árvore dessas é possível o programa contabilizar as variáveis de maior importância. (LOPES; GOEDTEL, ALESSANDRO, 2017).

Ao analisar a importância de cada variável para os modelos de predição dos atributos, como pode ser visto na Figura 3. Percebe-se que SB teve a variável CaO como sendo a de maior importância, de forma similar à sua importância pelo método de regressão múltipla stepwise, o que reforça o fato dessa variável ser importante na predição de SB independente do método utilizado. Para o pH, a variável CaO não esteve presente na equação de regressão múltipla stepwise, mas foi a de maior importância no modelo RF.

Para a matéria orgânica, a variável mais importante foi Al_2O_3 logo após aparece o Sr. Já para o P-rem, a de maior importância foi SiO_2 seguida também de Sr, tendo esses dois parâmetros a variável Zn como de menor importância.

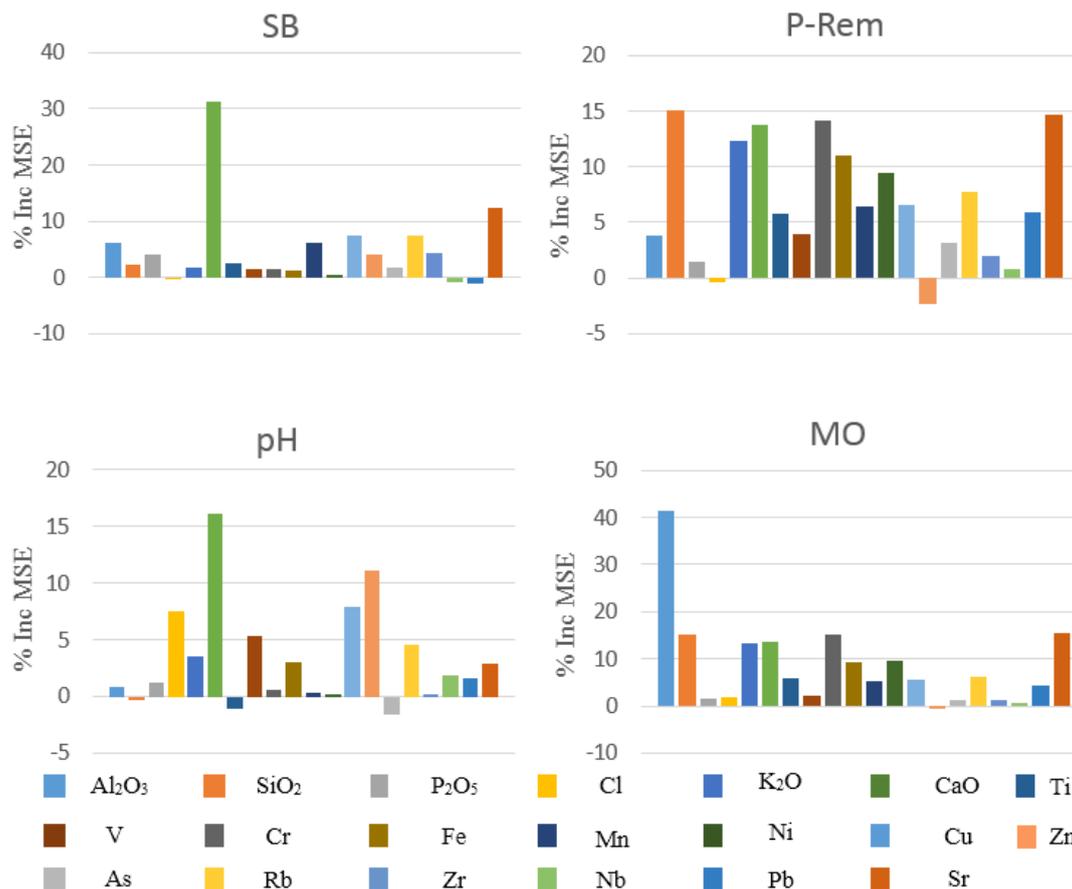


Figura 3: Índice de incremento no erro das variáveis obtidas pelo espectrômetro portátil de fluorescência de raios-X na predição de propriedades do solo pelo método de Random Forest.

Fonte: Do autor (2019)

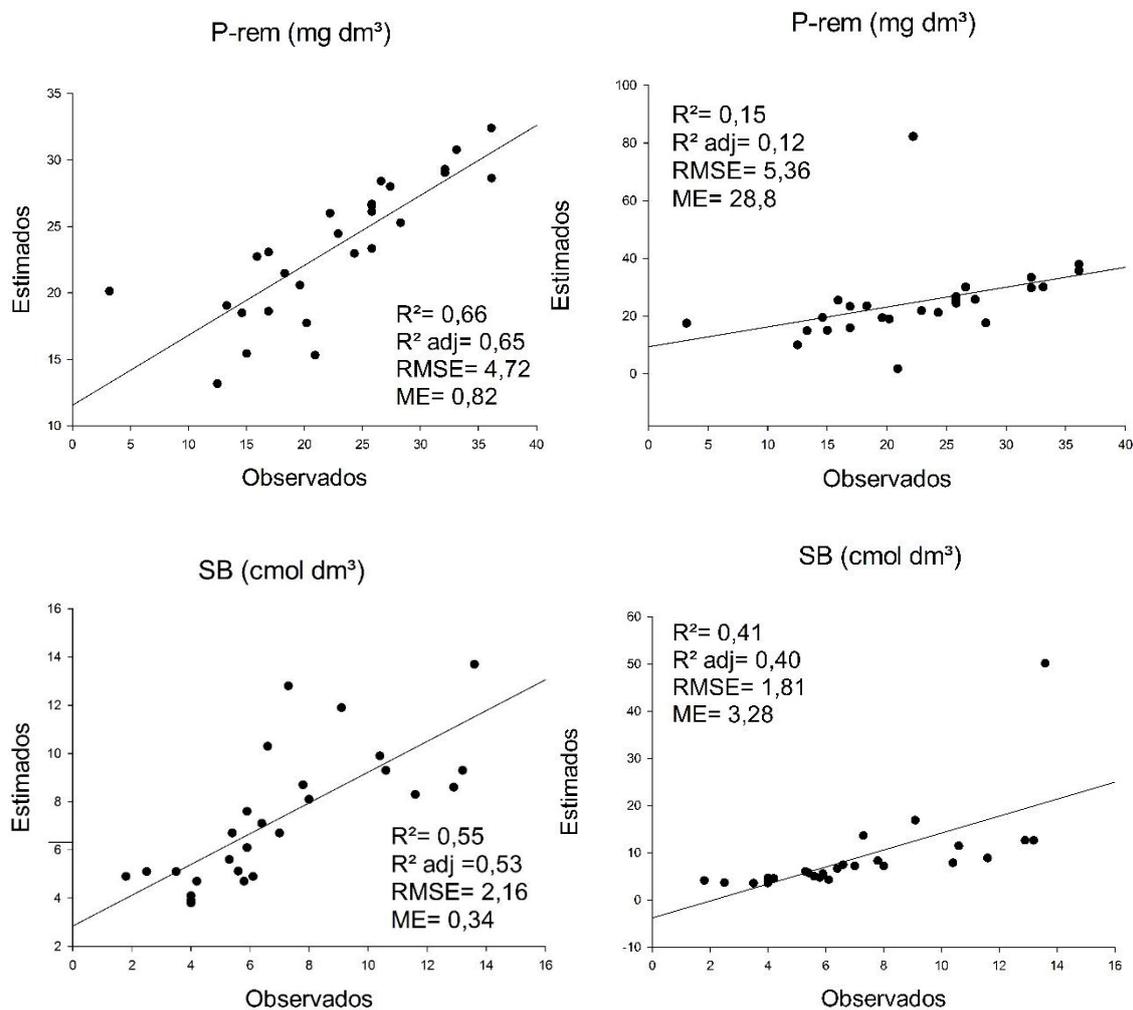
4.4 Validação dos modelos gerados por regressão linear múltipla *stepwise* e pelo RF

As análises de R^2 , R^2 ajustado, RMSE e ME das comparações entre os dados reais e preditos pelos métodos estatísticos estão apresentadas na Figura 5. Houve diferenças nos resultados obtidos pelos diferentes métodos, possibilitando a conclusão de que os parâmetros modelados pelo método RF foram os que tiveram o melhor desempenho. Esses resultados indicam que os modelos por RF têm melhor potencial em utilizar os dados dos elementos do pXRF para predições de atributos do solo que a regressão linear múltipla *stepwise*. Esse resultado se deve ao

fato das árvores de decisões trabalharem com modelagem não aditiva além de relações não lineares, (GREVE et al., 2010).

Chagas et al. (2018), a partir de dados de sensoriamento remoto e utilizando o RF, fez a predição de resultados da CTC e encontrou um R^2 de 0,47. Bhering et al (2016) analisaram a eficácia do modelo gerado por RF no mapeamento de areia, argila e carbono orgânico utilizando espectroscopia de reflectância difusa sob diferentes resoluções espaciais, obtendo resultados de R^2 de 0,4; 0,44 e 0,33, respectivamente. Ambos os trabalhos chegaram à conclusão de que a abordagem por RF apresenta potencial na estimativa de atributos do solo, concordando, deste modo, com os resultados deste experimento.

As predições que tiveram o melhor desempenho tanto para R^2 como para as demais análises foram o P-rem e a SB, com R^2 de 0,66 e 0,55 respectivamente. MO e pH tiveram os menores valores sendo de 0,47 e 0,26, respectivamente. SILVA et al. (2017), utilizando RF a partir de dados do pXRF, teve resultados gerais de 0,19 para P-rem e 0,51 para pH, reforçando o fato da qualidade dessas predições mudar conforme a região de estudo.



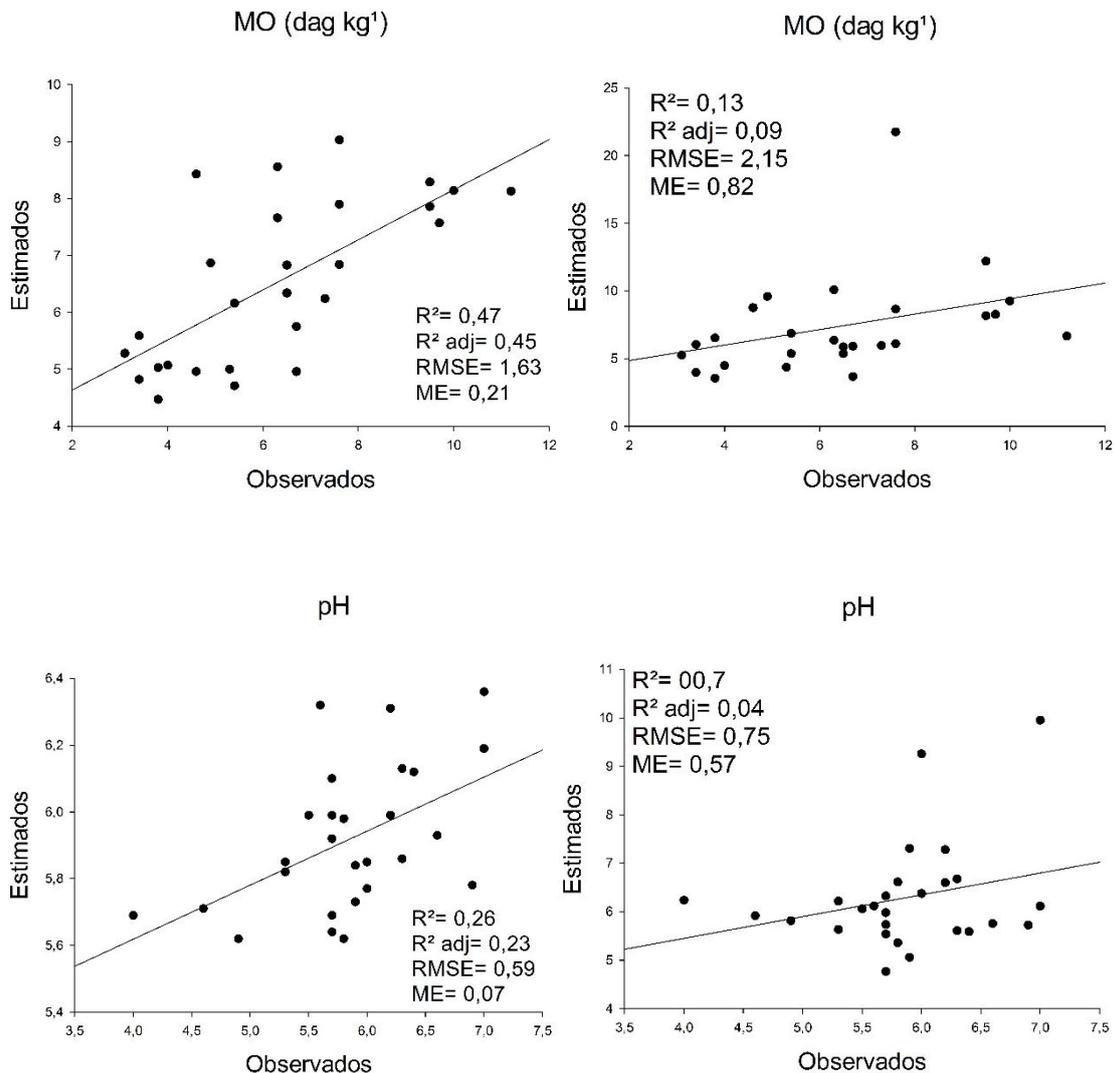


Figura 4: Validação das previsões dos atributos do solo. Para os modelos de *Random Forest* os dados se apresentam à esquerda e de regressão linear múltipla *Stepwise*, à direita.

Fonte: Do autor (2019)

4.5. Mapeamento dos atributos e validação dos mapas

Os modelos de maior acurácia foram utilizados na previsão espacial dos atributos, deste modo, os mapas estão apresentados na Figura 6. Ao analisar o mapa de MO, notamos que seus menores valores encontram-se em áreas de pastagens e áreas com solo descoberto. Isso se deve a essas regiões terem pouco acúmulo de massa vegetal, com menor ciclagem de nutrientes, diferente das regiões de eucaliptos ou mata nativa, já que diferentes sistemas de manejo interferem no acúmulo do carbono orgânico, e conseqüentemente na quantidade de matéria orgânica no solo (CONCEIÇÃO et al, 2005).

Valores mais baixos de P-rem foram encontrados em pastagens e cultivo de café, tendo valores relativamente maiores em áreas nativas e Pinus. A SB teve menores concentrações nas

regiões de pastagens, como também nas áreas de florestas nativas e pinus, aumentando gradualmente nas regiões de cultivo anual e tendo maior concentração nas áreas de plantio de café. Essa variação conforme o uso da terra ocorre pelo fato de haver intervenção com correções periódicas de teores de nutrientes no solo nas áreas cultivadas, elevando assim os teores da SB.

Assim como a SB, o pH apresentou menor valor nas áreas nativas e de cultivo florestal, aumentando seu valor para regiões de cultivo, devido também à ação antrópica. Para o pH o método IDW proporcionou a formação de “*bull eyes*” que são manchas em torno dos pontos amostrais de áreas que se diferenciam da suavização geral da variável, dificultando a interpretação dos resultados (JAKOB e YOUNG, 2016)

Os valores de R^2 , R^2 adj, o RMSE e ME gerados estão apresentados na Tabela 5. Observa-se que os atributos que tiveram maior R^2 na validação pontual foram os que resultaram também em mapas com validação de maiores R^2 . Dentre eles, os de melhor desempenho foram P-rem, com 0,66, e SB, com 0,55. Já os menores valores de validação foram de matéria orgânica com 0,47 seguida de pH com 0,28.

O pXRF é um aparelho que faz a leitura da concentração total dos elementos na amostra e não nos teores disponíveis para as plantas ou em outros atributos de características do solo que são medidos de maneira indireta como pH, MO, SB e P-rem. Porém os presentes resultados relatam que apesar do método apresentar erros acumulados da modelagem e da validação dos mapas, as predições resultaram em uma precisão satisfatória.

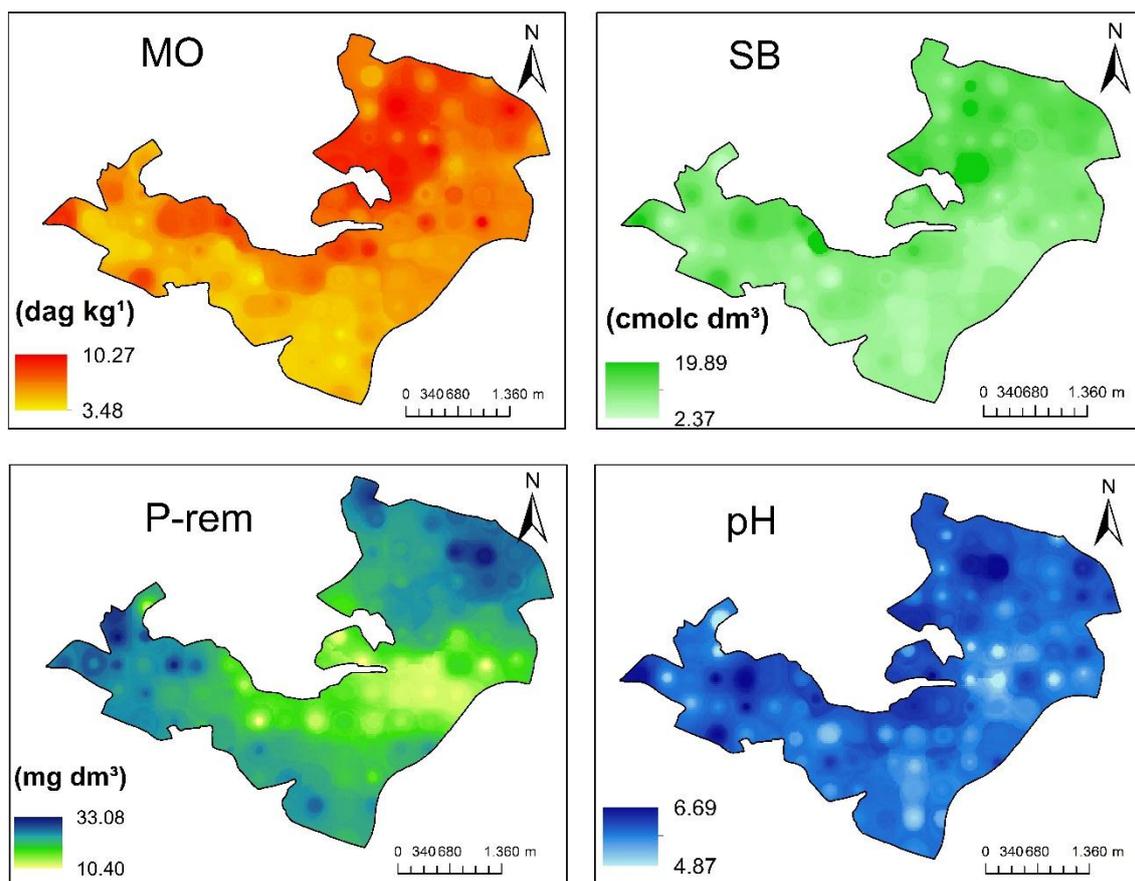


Figura 5. Mapas preditos para os atributos do solo na camada de 0-20 cm, obtidos através de modelos gerados pelo algoritmo Random Forest. MO: matéria orgânica, SB: soma de bases, P-rem: fósforo remanescente e pH.

Fonte: Do autor (2019)

Tabela 5. Resultados das análises de validação para os mapas da área. SB (cmol dm^3), M.O (dag kg^{-1}) e P-rem (mg dm^3).

Parâmetro	SB	MO	P-Rem	pH
R ²	0,54	0,47	0,66	0,28
R ² adj	0,52	0,45	0,65	0,24
ME	2,74	2,10	19,85	0,31
RMSE	1,66	1,45	4,45	0,55

5 Conclusão:

A variável CaO foi classificada como sendo de maior importância na determinação de SB e pH pelo modelo RF. Para a RMSE a variável CaO estava presente nas equações dos para atributos de SB, P-rem e MO.

Os atributos que tiveram maior R^2 na predição foram os que resultaram em mapas de maior R^2 , sendo eles: P-rem com 0,66 e SB com 0,54, já os menores valores de validação foram para matéria orgânica, com 0,47 seguida de pH, com 0,28.

Os modelos gerados por Random Forest apresentaram melhores resultados no que se diz respeito da predição de atributos do solo em relação ao método de regressão linear múltipla stepwise.

Ao levar em consideração todos os fatores que são utilizados na classificação de mapa do uso do solo, as análises realizadas pelo pXRF em conjunto com a modelagem do RF, foram suficientes para auxiliar na delimitação das glebas de diferentes uso do solo. Mostrando que o pXRF pode ser utilizado no mapeamento digital de solos tropicais em áreas com grande contraste de uso do solo.

Os dados extraídos a partir do pXRF podem ser utilizados na construção de modelos preditivos para atributos do solo de forma pontual ou espacial, neste último caso, realizando o mapeamento de atributos relacionados à fertilidade do solo local.

Referências:

ABBAD, G. S.; TORRES, C. V. **Regressão múltipla stepwise e hierárquica em Psicologia Organizacional: aplicações, problemas e soluções.** 2002.

ADAMCHUK, V.I.; MORGAN, M.T.; ESS, D.R. An automated sampling system for measuring soil pH. **Transactions of the ASAE**, St. Joseph, v. 42, n.4, p. 885-891, 1999.

ALVAREZ V., V. H.; RIBEIRO, A. C. Calagem. In: RIBEIRO, A. C.; GUIMARÃES, P. T. G.; ALVAREZ V., V. H. (Eds.). **Recomendações para o uso de corretivos e fertilizantes em Minas Gerais - 5º Aproximação.** Viçosa: CFSEMG, 1999. p.43-60.

ALVES, M. R. **Múltiplas técnicas no mapeamento digital de solos.** [s.l: s.n.]

BASGALUPP, M.P. LEGAL-Tree: Um algoritmo genético multi-objetivo lexicográfico para indução de árvores de decisão. Tese de Doutorado, ICMC-USP, São Carlos. 2010.

BHERING, S. B. et al. Mapeamento digital de areia, argila e carbono orgânico por modelos Random Forest sob diferentes resoluções espaciais. **Pesquisa Agropecuária Brasileira**, v. 51, n. 9, p. 1359-1370, 2016.

BREIMAN, L. Random Forests. **Machine Learnings**, v. 45, p. 5-32, 2001.

CHABRILLAT, S.; BEN-DOR, E.; VISCARRA ROSSEL, R. A. V.; DEMATTÊ, J. A. M. **Quantitative Soil Spectroscopy (Editorial).** Applied and Environmental Soil Science, v. 2013, p. 1-3, 2013.

CHAGAS, C. da S.; CARVALHO JÚNIOR, W. de; PINHEIRO, H. S. K.; XAVIER, P. A. M.; BHERING, S. B.; PEREIRA, N. R.; CALDERANO FILHO, B. Mapping Soil Cation Exchange Capacity in a Semiarid Region through Predictive Models and Covariates from Remote Sensing Data. **Revista Brasileira de Ciência do Solo**, v. 42, n. 0, p. 1–12, 2018.

CLARK, R, N. Spectroscopy of Rocks and Minerals, and Principles of Spectroscopy. In: EITELWEIR, M. T. **Sensoriamento proximal de solo para a quantificação de atributos químicos e físicos.** 2017. 124 p. Tese (Doutorado em Engenharia de Sistemas Agrícolas)-Universidade de São Paulo Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, 2017.

COMISSÃO DE FERTILIDADE DO SOLO DO ESTADO DE MINAS GERAIS. **Recomendações para o uso de corretivos e fertilizantes em Minas Gerais.** (5a aproximação). RIBEIRO, A.C.; GUIMARÃES, P.T.G.; ALVAREZ V., V.H. (Eds.). Viçosa: CFSEMG, 1999, 359p.

CONCEIÇÃO, P. C. et al. Qualidade do solo em sistemas de manejo avaliada pela dinâmica da matéria orgânica e atributos relacionados. **Revista brasileira de ciencia do solo.** Campinas. Vol. 29, n. 5 (2005), p. 777-788, 2005.

CURI, N. et al. **Mapeamento de solos e magnetismo no campus da UFLA como traçadores ambientais.** 1ª Edição. Lavras: Ed. UFLA, 2018.

CURI, N. et al. Geomorfologia, física, química e mineralogia dos principais solos da região de Lavras (MG). **Ciência e Prática**, 14(3):297-307, 1990.

CUTLER, R. D. et al. Random Forests for classification in Ecology. **Ecology**, v. 88, p.

2783-2792, 2007.

DEMATTÊ, J. A. M.; GUIMARÃES, C. C. B.; FONGARO, C. T.; VIDOY, E. L. F.; SAYÃO, V. M.; DOTTO, A. C.; SANTOS, N. V. dos. Satellite Spectral Data on the Quantification of Soil Particle Size from Different Geographic Regions. **Revista Brasileira de Ciência do Solo**, v. 42, n. 0, p. 1–14, 2018.

EITELWEIN, M. T.; TREVISAN, R. G.; COLAÇO, A. F. Predição da textura do solo através da reflectância espectral (Vis-NIR). **Journal of Hydrology**, v. 522, p. 475–487, 2015. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S002216941401066X>>.

EITELWEIR, M. T. **Sensoriamento proximal de solo para a quantificação de atributos químicos e físicos**. 2017. 124 p. Tese (Doutorado em Engenharia de Sistemas Agrícolas)-Universidade de São Paulo Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, 2017.

EMBRAPA SOLOS. **Mapeamento Digital de Solos**. Rio de Janeiro, 2010.

FERREIRA, E. et al. Dinâmica do uso e cobertura da terra no campus da Universidade Federal de Lavras de 1964 a 2009. **CERNE**, Lavras, v. 19, n. 1, p. 35-42, Mar. 2013.

FIORIO, P. R et al. Regressões lineares múltiplas utilizando-se dados espectrais obtidos por sensores em laboratório e orbital, na estimativa de atributos do solo. **SIMPÓSIO BRASILEIRO DE SENSORIAMENTO**, v. 11, p. 85-92, 2003.

GARGANTINI, H.; MELLO, F de A.F. de; ARZOLLA, S.. Efeitos da calagem no pH de perfis de solos de cerrado. **An. Esc. Super. Agric. Luiz de Queiroz, Piracicaba** , v. 39, n. 2, p. 1141-1158, 1982 .

GEBBERS, R.; ADAMCHUCK, V. I. **Precision agriculture and food security**. Science, v 327, p. 828-831, 2010.

GIASSON, E. et al. Digital soil mapping using multiple logistic regression on terrain parameters in southern Brazil. **Science Agricola**, Piracicaba, v. 63, n. 3, maio/jun. 2006.

GREVE, M. H. et al. Comparing decision tree modeling and indicator kriging for mapping the extent of organic soils in denmark. **Digital Soil Mapping: Bridging research, environmental application, and operation**. Progress in Soil Science, 2010, v.2, p.267-280.

HUNK, G. R. Electromagnetic radiation: the communication link in remote sensing. In: EITELWEIR, M. T. **Sensoriamento proximal de solo para a quantificação de atributos químicos e físicos**. 2017. 124 p. Tese (Doutorado em Engenharia de Sistemas Agrícolas)-Universidade de São Paulo Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, 2017.

JAKOB, A. A. E., & YOUNG, A. F. (2016). O uso de métodos de interpolação espacial de dados nas análises sociodemográficas. *Anais*, 1-22.

LAGACHERIE, P.; MCBRATNEY, A. B. Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping. In: TEN CATEN, A.; DALMOLIN, R. S. D.; MENDONÇA-SANTOS, M. de L.; GIASSON, E. Mapeamento digital de classes de solos: características da abordagem brasileira. **Ciência Rural**, v. 42, n. 11, p. 1989–1997, 2012.

LEHMANN, E.L; CASELHA, G., Teoria da Estimação Pontual (2ª ed.). **Nova Iorque**: Springer. (1998). ISBN 0-387-98502-6 . MR 1639875

LIAW, A.; WIENER, M. Classification and regression by random forest. **R News**, 2(December):18-22, 2002.

LIAW, A.; WIENER, M. **Package “randomForest.”** 2015.

LOPES, A. S.; GUILHERME, L. R. G. **I – Fertilidade Do Solo e Produtividade Agrícola.** [s.l.] Departamento de Ciência do Solo, Universidade Federal de Lavras – UFLA. Caixa, 2007.

LOPES, T. D.; GOEDEL, ALESSANDRO, R. H. C. P. W. **Aplicação Do Algoritmo Random Forest Como Classificador De Padroes De Falhas Em Rolamentos De Motores De Indução.** 2017.

LUTGENS, F. K.; TARBUCK, E. J., **Essentials of Geology**, 7th Ed., Prentice Hall, 2000.

MACHADO, L. de O; LANA, Â. M. Q.; LANA, R. M. Q.; GUIMARÃES, E. C.; FERREIRA, C. V. Variabilidade espacial de atributos químicos do solo em áreas sob sistema plantio convencional. **Revista Brasileira de Ciência do Solo**, v. 31, n. 3, p. 591–599, 2007.

MCLAREN, T. I.; GUPPY, C. N.; TIGHE, M. K. A rapid and nondestructive plant nutrient analysis using portable X-ray fluorescence. *Soil Science Society of America Journal*, 76:1446-1453, 2012. In: RIBEIRO, B. T.; SILVA, S. H. G.; SILVA, E. A.; GUILHERME, L. R. G. Aplicações da fluorescência de raios-X portátil (pXRF) na Ciência do Solo tropical. **Ciência e Agrotecnologia**, v. 41, n. 3, p. 245–254, 2017.

MCLEAN, E. O. et al. Aluminium in soils: I. Extraction methods and magnitud clays in Ohio soils. **Soil Science Society of America Proceedings**, 22(5):382-387, 1958.

NOLASCO-CARVALHO, C. C.; FRANCA-ROCHA, W.; UCHA, J. M. Mapa digital de solos: uma proposta metodológica usando inferência fuzzy. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 13, n. n. 1, p. 46–55, 2009.

NOVAIS, R.F; SMYTH, T.J. **Fósforo em solo e planta em condições tropicais.** Viçosa, MG, Universidade Federal de Viçosa, 1999. 399p.

OLIVEIRA, R. B. de et al. **Comparação entre métodos de amostragem do solo para recomendação de calagem e adubação do cafeeiro conilon.** *Engenharia Agrícola*, p. 176-186, 2008.

OLIVEIRA, M.F; **Mapeamento digital de solos da quadrícula de Ribeirão Preto – SP pelo método Random Forest.** 2016. 85 p. Dissertação (Mestre em Ciência do Solo) - Universidade Estadual Paulista Júlio de Mesquita Filho Faculdade de Ciências Agrárias e Veterinárias Câmpus de Jaboticabal, Jaboticabal. 2015

PALTRIDGE, N. G. et al. Energy-dispersive x-ray fluorescence spectrometry as a tool for zinc, iron and selenium analysis in whole grain wheat. *Plant Soil*, 361:261-269, 2012. In: RIBEIRO, B. T.; SILVA, S. H. G.; SILVA, E. A.; GUILHERME, L. R. G. Aplicações da fluorescência de raios-X portátil (pXRF) na Ciência do Solo tropical. **Ciência e Agrotecnologia**, v. 41, n. 3, p. 245–254, 2017.

PIMENTEL GOMES, F. **Curso de Estatística Experimental.** São Paulo: Nobel, 1985.

PINHEIRO, H. S. K.; **Métodos de mapeamento digital aplicados na predição de classes e atributos dos solos a bacia hidrográfica do rio Guapi Macacu, RJ.** 2015. 180 p. Tese (Doutorado em Ciência do Solo) – Universidade Federal Rural do Rio de Janeiro. Rio de Janeiro, 2015.

RESENDE, A. V. de (Embrapa); FONTOURA, S. M. V. (Fapa); BORGHI, E. (Embrapa); SANTIOS, F. C. do. (Embrapa); KAPPES, C. (Fundação M.; MOREIRA, S. G.; OLIVEIRA JUNIOR, A. de; BORIN, A. L. D. C. Solos de Fertilidade Construída: Características, Funcionamento e Manejo. **Informações Agronômicas**, v. 19, n. 156, p. 1–38, 2016.

RESENDE, M. et al. **Pedologia: Base para Distinção de Ambientes**. 5. ed. Lavras: Editora UFLA, 2007.

SASSAKI, M.H.; MACHADO, T.M.; SHIRATSUCHI, L.S.; DETERMINAÇÃO DE PH E POTÁSSIO EM SOLOS DO MEIO NORTE DE MATO GROSSO UTILIZANDO ELETRODOS SELETIVOS DE ÍONS. In: XLIII Congresso Brasileiro de Engenharia Agrícola, 42., 2014, Campo Grande. **Anais...** Campo Grande: 2014. p. 0-3.

SHARMA, A.; WEINDORF, D. C.; MAN, T.; ALDABAA, A. A. A.; CHAKRABORTY, S. Characterizing soils via portable X-ray fluorescence spectrometer: 3. Soil reaction (pH). **Geoderma**, v. 232–234, p. 141–147, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.geoderma.2014.05.005>>.

SHARMA, A.; WEINDORF, D. C.; WANG, D. D.; CHAKRABORTY, S. Characterizing soils via portable X-ray fluorescence spectrometer: 4. Cation exchange capacity (CEC). **Geoderma**, v. 239, p. 130–134, 2015. Disponível em: <<http://dx.doi.org/10.1016/j.geoderma.2014.10.001>>.

SILVA, F.C. S.; **Mensuração direta do pH do solo em campo por meio de eletrodos íon-seletivos**. 2013. p 102. Tese (Mestrado em Engenharia de Sistemas Agrícolas). Universidade de São Paulo Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, 2013.

SILVA, S. H. G.; SILVA, E. A.; POGGERE, G. C.; GUILHERME, L. R. G.; CURTI, N. Tropical soils characterization at low cost and time using portable X-ray fluorescence spectrometer (pXRF): effects of different sample preparation methods. **Ciência e Agrotecnologia**, v. 42, n. 1, p. 1–13, 2018.

SILVA, S. H. G.; TEIXEIRA, A. F. dos S.; MENEZES, M. D. de; GUILHERME, L. R. G.; MOREIRA, F. M. de S.; CURTI, N. Multiple linear regression and random forest to predict and map soil properties using data from portable X-ray fluorescence spectrometer (pXRF). **Ciência e Agrotecnologia**, v. 41, n. 6, p. 648–664, 2017.

SILVA, S. H. G.; SILVA, E. A.; GUILHERME, L. R. G. Aplicações da fluorescência de raios-X portátil (pXRF) na Ciência do Solo tropical. **Ciência e Agrotecnologia**, v. 41, n. 3, p. 245–254, 2017.

SOUSA JUNIOR, J. G. de A.; DEMATTE, J. M.; GENU, A. M. **Comportamento espectral dos solos na paisagem a partir de dados coletados por sensores terrestre e orbital**. Rev. Bras. Ciênc. Solo, Viçosa, v. 32, n. 2, p. 727-738, Apr. 2008. Available from <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-06832008000200027&lng=en&nrm=iso>. access on 11 Nov. 2018. <http://dx.doi.org/10.1590/S0100-06832008000200027>.

TABACHNICK, B. G.; FIDELL, L. S. **Using multivariate statistics**. Northridge. Cal.: Harper Collins, 1996.

TALIBUDEEN, O. **Ion-selective electrodes**. In: Smith, K.A (Ed.), Soil Analysis. Modern Instrumental Techniques, vol. 3, 2nd ed. Marcel Dekker, Inc, New York, NY, 1991. pp. 111-182.

TEIXEIRA, W. G.; VASQUES, G. M. **Uso de espectrômetro de fluorescência de raios x portátil (pxrf) para avaliação de teores de ferro e manganês em amostras de solo.** [s.d.]

TEN CATEN, A.; DALMOLIN, R. S. D.; MENDONÇA-SANTOS, M. de L.; GIASSON, E. Mapeamento digital de classes de solos: características da abordagem brasileira. **Ciência Rural**, v. 42, n. 11, p. 1989–1997, 2012.

TOWET, E. K.; SHEPHRED, K. D.; DRAKE, B. L. **Plant elemental composition and portable X-ray fluorescence (pXRF) spectroscopy: Quantification under different analytical parameters.** X-Ray Spectrometry, 45:117-124, 2015. In: RIBEIRO, B. T.;

UNIVERSIDADE FEDERAL DE LAVRAS. Biblioteca Universitária. **Manual de normalização e estrutura de trabalhos acadêmicos: TCCs, monografias, dissertações e teses.** 2. ed. rev., atual. e ampl. Lavras, 2016.

VISCARRA ROSSEL, R. A. & MCBRATNEY, A. B. **Laboratory Evaluation of a Proximal Sensing Technique for Simultaneous Measurement of Soil Clay and Water Content.** Geoderma, 85, 19–39, 1998.

VISCARRA ROSSEL, R. A; WALVOORT, D. J. J.; MCBRATNEY, A. B.; JANIK, L. J.; SKJEMSTA, J. O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. In: EITELWEIR, M. T. **Sensoriamento proximal de solo para a quantificação de atributos químicos e físicos.** 2017. 124 p. Tese (Doutorado em Engenharia de Sistemas Agrícolas) -Universidade de São Paulo Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, 2017.

VISCARRA ROSSEL, R.A.; GILBERTSON, M.; THYLEN, L.; HANSEN, O.; MCVEY, S.; MCBRATNEY, A.B. 2005. Field measurements of soil pH and lime requirement using an on-the-go soil pH and lime requirement measurement system. In: STAFFORD, J.V. (Ed.). **European Conference on Precision Agriculture**, 5., 2005. Uppsala, Sweden. **Proceedings**... Wageningen The Netherlands: Academic Publishers, 2005. p. 511 – 519.

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