

VIVIANE FLAVIANA CONDÉ

**PREDIÇÃO DE UMIDADE E ATRIBUTOS FÍSICOS DO SOLO COM USO DE
GPR (*GROUND PENETRATING RADAR*)**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Solos e Nutrição de Plantas, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2018

**Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa**

T

C745p
2018
Condé, Viviane Flaviana, 1991-
Predição de umidade e atributos físicos do solo com uso de
GPR (*Ground Penetrating Radar*) / Viviane Flaviana Condé. –
Viçosa, MG, 2018.
viii, 50f. : il. (algumas color.) ; 29 cm.

Orientador: Elpídio Inácio Fernandes Filho.
Dissertação (mestrado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Solos - Umidade - Medição. 2. Solos - Propriedades
elétricas. 3. Radar de penetração no solo. I. Universidade Federal
de Viçosa. Departamento de Solos. Programa de Pós-Graduação
em Solos e Nutrição de Plantas. II. Título.

CDD 22. ed. 631.4

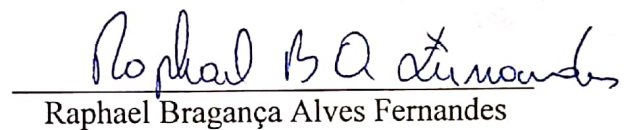
VIVIANE FLAVIANA CONDÉ

**PREDIÇÃO DE UMIDADE E ATRIBUTOS FÍSICOS DO SOLO COM USO DE
GPR (*GROUND PENETRATING RADAR*)**

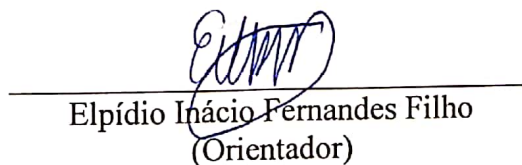
Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Solos e Nutrição de Plantas, para obtenção do título de *Magister Scientiae*.

APROVADA: 13 de julho de 2018.


Márcio Rocha Francelino
(Coorientador)


Raphael Bragança Alves Fernandes


Cibele Cláuver de Aguiar


Elpídio Inácio Fernandes Filho
(Orientador)

AGRADECIMENTOS

Agradeço a Deus, pela vida e todas as bênçãos ao longo dela.

À Universidade Federal de Viçosa (UFV) e ao Programa de Pós-Graduação em Solos e Nutrição de Plantas – UFV. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Ao Prof. Elpídio Inácio Fernandes Filho, pela orientação, apoio, confiança e amizade.

Ao Prof. Márcio Rocha Francelino, pela coorientação, apoio e sugestões.

A todos os membros do Labgeo, por todo apoio na execução deste trabalho.

Ao Adalberto Rodrigues Lira Gomes por todo apoio nas atividades de campo e laboratório.

Ao Gustavo Vieira Veloso pela paciência e disposição em colaborar na execução deste trabalho.

Aos membros da banca, pela contribuição e correção deste trabalho.

A minha família, em especial aos meus pais, José e Francisca, e as minhas irmãs, Fabiana e Lidiana, pelo incentivo, força, amor e carinho.

Aos meus amigos, pela amizade e presença em todos os momentos.

A todos aqueles que sabem da sua importância nesta e nas demais conquistas.

“Agradeço a todas as dificuldades que enfrentei; não fosse por elas, eu não teria saído do lugar (Chico Xavier)”

SUMÁRIO

RESUMO	v
ABSTRACT	vii
INTRODUÇÃO GERAL	1
REFERÊNCIAS	3
CAPÍTULO I	4
RESUMO	5
ABSTRACT	6
1 INTRODUÇÃO	7
2 MATERIAL E MÉTODOS	9
2.1. Caracterização da área de estudo	9
2.2. Coleta de dados com GPR	10
2.3. Coleta das amostras de solo	10
2.4. Processamento dos radargramas	10
2.5. Estimativa dos valores de umidade volumétrica a partir dos radargramas	14
3 RESULTADOS E DISCUSSÃO	16
3.1 Coleta dos valores de constante dielétrica	16
3.2 Umidade volumétrica - método padrão	18
3.3 Umidade volumétrica a partir dos radargramas	18
3.4 Proposição do modelo	19
4 CONCLUSÕES	22
REFERÊNCIAS	23
CAPÍTULO II	26
RESUMO	27
ABSTRACT	28
1 INTRODUÇÃO	29
2 MATERIAL E MÉTODOS	30
2.1. Caracterização da área de estudo	30
2.2. Coleta de dados	30
2.2.1. Coleta e análise de amostras de solo	30
2.2.2. Coleta de dados com GPR	31

2.2.3. Processamento dos radargramas	32
2.2.4. Aquisição das covariáveis de morfometria	32
2.2.5. Análises estatísticas.....	33
3 RESULTADOS E DISCUSSÃO	37
3.1. Textura do solo e porosidade	37
3.2. Constante dielétrica aparente do solo.....	37
3.3. Atributos físicos do solo a partir dos radargramas.....	40
3.3.1. Tratamento 01	40
3.3.2. Tratamento 02	40
3.3.3. Tratamento 03	41
3.3.4. Tratamento 04	41
3.3.5. Tratamento 05	42
4 CONCLUSÕES	45
REFERÊNCIAS.....	46
CONCLUSÕES GERAIS	50

RESUMO

CONDÉ, Viviane Flaviana, M.Sc., Universidade Federal de Viçosa, julho de 2018. **Predição de umidade e atributos físicos do solo com uso de GPR (*Ground Penetrating Radar*)**. Orientador: Elpídio Inácio Fernandes Filho. Coorientador: Márcio Rocha Francelino

O GPR (*Ground Penetrating Radar*) é um método não destrutivo e não invasivo que permite extrair informações ao longo do perfil de solo. É uma técnica indireta para investigação da subsuperfície do solo, que vêm sendo estudada como uma possível alternativa para determinação da umidade e atributos físicos do solo. O trabalho foi desenvolvido em uma área com cobertura por pastagem em relevo fortemente ondulado no município de Viçosa, Minas Gerais, Brasil, onde foram realizadas duas varreduras em três transectos distintos, com 70, 100 e 130 m de comprimento, no sentido do topo para o vale em relevo montanhoso. O GPR utilizado foi o SRI-3000 da GSSI®. Em cada uma das três áreas foram realizados transectos com antenas blindadas de 200, 400 e 900 MHz para geração de radargramas, os quais foram processados no software RADAN 7®. Nos radargramas foram realizados diversos processos como aplicação de filtros, ganhos e migração, todos visando a melhoria do sinal eletromagnético e determinação da constante dielétrica. A presente dissertação é dividida em dois capítulos. O primeiro capítulo objetivou avaliar a aplicabilidade do uso do *Ground Penetrating Radar* na determinação da umidade volumétrica em transectos comparando equações e propor modelo de calibração. Foram realizadas coletas em 26 pontos de amostras deformadas nas profundidades de 0,00-0,20; 0,20-0,40; 0,40-0,60; 0,60-0,80 e 0,80-1,00 m e realizada a análise laboratorial padrão de umidade. Para estimar a umidade volumétrica dos solos utilizando dados do GPR, a constante dielétrica foi obtida por meio dos processamentos dos radargramas e a umidade calculada conforme as equações de Topp et al. (1980), Ledieu et al. (1986), Roth et al. (1990) e Malicki et al. (1996). Foi utilizado o software estatístico R® para ajustar equações de regressão entre umidade do solo medida em laboratório e dados de constante dielétrica obtidos nos radargramas por GPR. Gerou-se uma equação de calibração local onde a umidade é função da constante dielétrica, da profundidade de coleta das amostras e da interação entre os fatores. Os valores de umidade volumétrica obtidos por GPR mostraram-se satisfatórios, o que indica possibilidade do uso da metodologia na espacialização da variável umidade do solo, otimizando estudos que necessitam de valores de umidade como uma variável com

distribuição contínua. No segundo capítulo, o objetivo foi elaborar modelos de predição de variáveis texturais (areia grossa, areia fina, silte, argila, porosidade total, macroporosidade e microporosidade) em solos do município de Viçosa, Minas Gerais, utilizando dados de *Ground Penetrating Radar*, variáveis morfométricas de terreno e técnicas de *Machine Learning*. Para as análises de textura e porosidade do solo foram realizadas coletas em 26 pontos de amostras indeformadas nas profundidades de 0,00-0,10; 0,10-0,20; 0,20-0,30; 0,30-0,40 e 0,40-0,50 m e realizada a análise laboratorial padrão. Para predição de textura e porosidade dos solos utilizando dados do GPR, a constante dielétrica foi obtida por meio dos processamentos dos radargramas e, juntamente aos resultados das análises laboratoriais, foram utilizados cinco algoritmos de predição: *Linear Models* (LM), *Generalized Linear Models* (GLM), *Cubist*, *Support Vector Machine Radial Sigma* (SVM) e *Random Forest* (RF), para verificar qual modelo se adequava melhor as cinco bases de dados testadas. Foi utilizado o software estatístico R[®] para ajustar os modelos de regressão. A verificação do desempenho dos algoritmos foi realizada por método de validação cruzada do tipo *leave one out*. O desempenho dos algoritmos em cada treinamento e validação foi avaliado pelo coeficiente de determinação (R^2) e pela raiz do erro médio quadrático (RMSE), para cada variável. Testou-se combinações distintas de bases de dados para avaliar quais variáveis são mais importantes na elaboração de modelos preditivos. Os modelos SVM, RF, e *Cubist* apresentaram maior desempenho quando comparados aos modelos lineares. Inserir variáveis de terreno possibilita modelagens de maior precisão, com valores de R^2 de 0,84 e 0,79 para variável argila, como os modelos SVM e RF, respectivamente. Os valores de RMSE foram extremamente baixos para todos os modelos (inferiores a 0,04). A predição das variáveis texturais gerou resultados superiores à predição de valores de porosidade.

ABSTRACT

CONDÉ, Viviane Flaviana, M.Sc., Universidade Federal de Viçosa, July, 2018. **Prediction of soil moisture and physical attributes using GPR (*Ground Penetrating Radar*)**. Adviser: Elpídio Inácio Fernandes Filho. Co-adviser: Márcio Rocha Francelino.

GPR (Ground Penetrating Radar) is a non-destructive, non-invasive method that allows information to be extracted along the soil profile. It is an indirect technique for soil subsurface research, which has been studied as a possible alternative for the determination of soil moisture and physical attributes. The work was carried out in a heavily corrugated grassland area in the municipality of Viçosa, Minas Gerais, Brazil, where two sweeps were performed on three distinct transects, 70, 100 and 130 m the valley in mountainous relief. The GPR used was the SRI-3000 from GSSI®. Transects with 200, 400 and 900 MHz shielded antennas were generated in each of three areas to generate radargrams, which were processed in RADAN 7® software. In the radargrams, several processes were applied, such as filters, gains and migration, all of them aimed at improving the electromagnetic signal and determining the dielectric constant. This dissertation is divided into two chapters. The first chapter aimed to evaluate the applicability of the use of the Ground Penetrating Radar in the determination of volumetric moisture in transects comparing equations and propose a calibration model. Samples were collected at 26 points of deformed samples at depths of 0.00-0.20; 0.20-0.40; 0.40-0.60; 0.60-0.80 and 0.80-1.00 m and performed the standard laboratory analysis of moisture. In order to estimate the volumetric moisture of the soils using GPR data, the dielectric constant was obtained by means of the radargrams and humidity calculated according to the equations of Topp et al. (1980), Ledieu et al. (1986), Roth et al. (1990) and Malicki et al. (1996). The R® statistical software was used to adjust regression equations between laboratory-measured soil moisture and dielectric constant data obtained on GPR-radargrams. It was generated a local calibration equation where humidity is a function of the dielectric constant, the depth of sample collection and the interaction between the factors. The values of volumetric moisture obtained by GPR were satisfactory, which indicates the possibility of using the methodology in the spatialization of the soil moisture variable, optimizing studies that require moisture values as a variable with continuous distribution. In the second chapter, the objective was to elaborate prediction models of textural variables (coarse sand, fine sand, silt, clay, total porosity, macroporosity and

microporosity) in soils of the municipality of Viçosa, Minas Gerais, using data from Ground Penetrating Radar, morphometric and Machine Learning techniques. For soil texture and porosity analysis, samples were collected at 26 points of undisturbed samples at depths of 0.00-0.10; 0.10-0.20; 0.20-0.30; 0.30-0.40 and 0.40-0.50 m and performed the standard laboratory analysis. In order to predict soil texture and porosity using GPR data, the dielectric constant was obtained by means of radargrams and, together with the results of the laboratory analysis, five predictive algorithms were used: Linear Models (LM), Generalized Linear Models (GLM), Cubist, Support Vector Machine Radial Sigma (SVM) and Random Forest (RF), to verify which model best suited the five databases tested. The R® statistical software was used to adjust the regression models. The performance verification of algorithms was performed by a one-off cross-validation method. The performance of algorithms in each training and validation was evaluated by the coefficient of determination (R^2) and root mean square error (RMSE), for each variable. Different combinations of databases have been tested to assess which variables are most important in the design of predictive models. The SVM, RF, and Cubist models presented higher performance when compared to linear models. Inserting field variables allows for more accurate modeling, with R^2 values of 0.84 and 0.79 for clay, such as the SVM and RF models, respectively. RMSE values were extremely low for all models (less than 0.04). The prediction of textural variables generated results superior to prediction of porosity values.

INTRODUÇÃO GERAL

Os métodos convencionais para determinação da umidade e atributos físicos do solo são trabalhosos e devem ser executados com cautela, pois estão sujeitos a sérios erros nas amostragens e procedimentos analíticos. Assim, torna-se necessário desenvolver metodologias alternativas para a avaliação dos atributos do solo, visando tornar o trabalho mais rápido e operacional (CEZAR et al., 2010). Neste sentido, técnicas de sensoriamento proximal vêm sendo estudadas como uma possível alternativa, dentre elas o uso do *Ground Penetrating Radar*.

O GPR (*Ground Penetrating Radar*) é um método não-invasivo e não destrutivo que permite extrair informações ao longo do perfil de solo. É uma técnica indireta para investigação da subsuperfície do solo (CEZAR et al., 2010).

O princípio de funcionamento do GPR consiste na emissão e recepção de ondas eletromagnéticas, por antenas transmissoras e receptoras, e armazenamento dos dados em forma de traços, que originam os radargramas. As ondas eletromagnéticas emitidas estão no comprimento de micro-ondas. A radiação em micro-ondas é muito utilizada para detectar a distância de objetos, calculada por meio da relação entre a velocidade de propagação da onda e o tempo entre a emissão e recepção da radiação eletromagnética. O tempo de viagem da onda é uma medida indireta que está relacionada com a constante dielétrica do material presente no solo que pode refletir ou atenuar a propagação da energia eletromagnética (CEZAR et al., 2010; WIJewardana et al., 2017).

A resolução e profundidade de penetração do sinal GPR são determinados pela frequência da antena e as propriedades elétricas dos materiais. Quanto maior a frequência da antena, menor o comprimento de onda e, conseqüentemente, menor a profundidade de penetração do sinal e maior a riqueza de detalhes nos radargramas gerados (TRAN et al., 2015).

O radargrama gera um perfil em subsuperfície em um transecto com grande volume de amostra, reduzindo o tempo de amostragem, o custo e o trabalho nos levantamentos de solos. Constitui-se em uma potencial ferramenta nos trabalhos pedológicos, entretanto, é pouco utilizado. Há relatos na literatura do uso de GPR para determinar espessuras de depósitos orgânicos (COLLINS et al., 1986); espessura dos horizontes do solo, fragmentos grosseiros, concreções, textura (DOOLITTLE; COLLINS, 1995); umidade e outras propriedades que auxiliem na classificação pedológica do solo, contudo, estudos relacionando atributos do solo a dados obtidos com

GPR, são necessários para otimizar a interpretação dos radargramas (CEZAR et al., 2010).

Embora os sensores proximais possam produzir resultados não tão precisos como análises laboratoriais convencionais, a facilidade de gerar grande quantidade de dados de forma, simples e menos trabalhosas e menor custo torna à metodologia atrativa (ROSSEL, 2013).

O processamento dos dados gerados com os radargramas podem envolver um grande esforço informático. O Machine Learning (aprendizado de máquina) surge como alternativa dentro da ciência do solo, por simplificar o ajuste de modelos e otimizar os parâmetros, possibilitando a interpretação de variáveis com maior nível de detalhamento e precisão. Os métodos de Machine Learning são amplamente utilizados na elaboração de modelos preditivos de diversos atributos do solo (KIRKWOOD et al., 2016; HENGL et al. 2017).

Diante do exposto, a presente dissertação é dividida em dois capítulos. O primeiro capítulo objetivou avaliar a aplicabilidade do uso do *Ground Penetrating Radar* na determinação da umidade volumétrica em transectos comparando equações e propor modelo de calibração. No segundo capítulo, o objetivo foi elaborar modelos de predição de variáveis texturais (areia grossa, areia fina, silte, argila, porosidade total, macroporosidade e microporosidade) em solos do município de Viçosa, Minas Gerais, utilizando dados de *Ground Penetrating Radar*, variáveis morfométricas de terreno e técnicas de *Machine Learning*.

REFERÊNCIAS

CEZAR, E.; NANNI, M. R.; CHICATI, M. L.; FABRIO, F. D.; HATA, F. T.; OLIVEIRA, R. B. Uso de sistema GPR (*Ground Penetrating Radar*) na avaliação de atributos de um solo sob plantio de cana-de-açúcar. *Revista Brasileira de Ciência de Solo*, Viçosa, MG, v. 34, n. 2, p. 291-297, 2010.

COLLINS, M.E. et al. Using ground penetrating radar to study changes in soil map unit composition in selected histosols. *Soil Science Society American Journal*, 50:408-412, 1986.

DOOLITTLE, J.A.; COLLINS, M.E. Use of soil information to determine application of ground penetrating radar. *Journal of Applied Geophysics*, 33:101-108, 1995.

HENGL, T. et al. SoilGrids250m: Global gridded soil information based on machine learning. *PloS one*, v. 12, n. 2, p. e0169748, 2017.

KIRKWOOD, C., CAVE, M., BEAMISH, D., GREBBY, S., FERREIRA, A. A Machine Learning approach to geochemical mapping. *Journal of Geochemical Exploration*, v. 167, p. 49-61, 2016.

ROSSEL, R. A. V.; MCBRATNEY, A. B.; MINASNY, B. (Ed.). *Proximal soil sensing*. Springer Science & Business Media, 2010.

TRAN, A. P., BOGAERT, P., WIAUX, F., VANCLOOSTER, M., LAMBOT, S. High-resolution space-time quantification of soil moisture along a hillslope using joint analysis of ground penetrating radar and frequency domain reflectometry data. *Journal of Hydrology*, v. 523, p. 252–261, 2015.

WIJEWARDANA, Y.N.S. et al. Ground penetrating radar (GPR) responses for sub-surface salt contamination and solid waste: modeling and controlled lysimeter studies. *Environ Monit Assess*, 189:1-15, 2017.

CAPÍTULO I

ESTIMATIVA DE UMIDADE VOLUMÉTRICA DO SOLO POR *Ground Penetrating Radar*

RESUMO

O estudo da dinâmica e distribuição espacial da umidade do solo é fundamental em especial na tomada de decisões quanto ao manejo do solo. O uso de sensores proximais ainda é incipiente na ciência do solo, apesar desses apresentarem um grande potencial de aplicações, como a possibilidade de mapeamento de propriedades do solo de forma rápida e não invasiva. Neste contexto, o objetivo foi avaliar a aplicabilidade do *Ground Penetrating Radar* (GPR) na determinação da umidade volumétrica o solo. O trabalho foi desenvolvido em uma área com cobertura por pastagem em relevo fortemente ondulado no município de Viçosa, Minas Gerais, Brasil, onde foram realizadas duas varreduras em três transectos distintos, com 70, 100 e 130 m de comprimento, no sentido do topo para o vale em relevo montanhoso. O GPR utilizado foi o SRI-3000 da GSSI®. Em cada uma das três áreas foram realizados transectos com antenas blindadas de 200, 400 e 900 MHz para geração de radargramas, os quais foram processados no software RADAN 7®. Nos radargramas foram realizados diversos processos como aplicação de filtros, ganhos e migração, todos visando a melhoria do sinal eletromagnético e determinação da constante dielétrica. Para a análise de umidade do solo foram realizadas coletas em 26 pontos de amostras deformadas nas profundidades de 0,00-0,20; 0,20-0,40; 0,40-0,60; 0,60-0,80 e 0,80-1,00 m e realizada a análise laboratorial padrão de umidade. Para estimar a umidade volumétrica dos solos utilizando dados do GPR, a constante dielétrica foi obtida por meio dos processamentos dos radargramas e a umidade calculada conforme as equações de Topp et al. (1980), Ledieu et al. (1986), Roth et al. (1990) e Malicki et al. (1996). Foi utilizado o software estatístico R® para ajustar equações de regressão entre umidade do solo medida em laboratório e dados de constante dielétrica obtidos nos radargramas por GPR. Gerou-se uma equação de calibração local onde a umidade é função da constante dielétrica, da profundidade de coleta das amostras e da interação entre os fatores. Os valores de umidade volumétrica obtidos por GPR mostraram-se satisfatórios, o que indica possibilidade do uso da metodologia na espacialização da variável umidade do solo, otimizando estudos que necessitam de valores de umidade como uma variável com distribuição contínua.

Palavras-chave: GPR, constante dielétrica do solo, umidade do solo.

ABSTRACT

The study of dynamics and spatial distribution of soil moisture is essential especially in making decisions as to soil management. The use of proximal sensors is still incipient in soil science, although these have a great potential for applications, such as the possibility of mapping soil properties in a fast and non-invasive way. In this context, the objective was to evaluate the applicability of Ground Penetrating Radar (GPR) in determining soil volumetric moisture. The work was carried out in a heavily corrugated grassland area in the municipality of Viçosa, Minas Gerais, Brazil, where two sweeps were performed on three distinct transects, 70, 100 and 130 m the valley in mountainous relief. The GPR used was the SRI-3000 from GSSI®. Transects with 200, 400 and 900 MHz shielded antennas were generated in each of the three areas to generate radargrams, which were processed in RADAN 7® software. In the radargrams, several processes were applied, such as filters, gains and migration, all of them aimed at improving the electromagnetic signal and determining the dielectric constant. For the analysis of soil moisture, samples were collected at 26 points of deformed samples at depths of 0.00-0.20; 0.20-0.40; 0.40-0.60; 0.60-0.80 and 0.80-1.00 m and performed the standard laboratory analysis of moisture. In order to estimate the volumetric moisture of soils using GPR data, the dielectric constant was obtained by means of radargrams and humidity calculated according to the equations of Topp et al. (1980), Ledieu et al. (1986), Roth et al. (1990) and Malicki et al. (1996). The statistical software R® was used to adjust regression equations between soil moisture measured in laboratory and data of dielectric constant obtained in the radargrams by GPR. A local calibration equation was generated where the humidity is a function of the dielectric constant, the depth of sample collection and the interaction between the factors. The values of volumetric moisture obtained by GPR were satisfactory, which indicates the possibility of using the methodology in the spatialization of the soil moisture variable, optimizing studies that require moisture values as a variable with continuous distribution.

Key words: GPR, soil dielectric constant, soil moisture.

1 INTRODUÇÃO

O uso de sensores proximais ainda é incipiente na ciência do solo. Poucos sensores são destinados a investigações do solo em subsuperfície. E equipamentos como *Ground Penetrating Radar* (GPR) possibilitam o levantamento de propriedades do solo de forma rápida e não invasiva. O GPR vem se destacando por ser um método não destrutivo de inspeção do solo (FRANCELINO et al., 2011). As medições são feitas em condições de campo e os dados adquiridos na superfície, o que permite produção de informações praticamente instantâneas (ROSSEL et al., 2011).

O GPR utiliza transmissão e reflexão de pulsos de alta frequência, na faixa das micro-ondas de 10 a 1500 MHz e pode ser operado em modo estacionário ou móvel. Trata-se de sensor ativo, composto por duas antenas (transmissora e receptora) e uma unidade de controle. Antenas de diferentes frequências produzem radargramas com resoluções e profundidades diferentes. Sabe-se que frequências mais elevadas aumentam a resolução da imagem em detrimento da profundidade de penetração do sinal eletromagnético. A propagação das ondas é dependente das propriedades elétricas, magnéticas e dielétricas dos materiais e estão sujeitas a reflexão, espalhamento, deflexão e refração (WIJEWARDANA et al., 2017).

A interação do sinal emitido pelo equipamento com o material depende da frequência utilizada e das propriedades do substrato estudado, com ênfase para a constante dielétrica, salinidade, composição mineralógica e textura. A constante dielétrica (K) determina a velocidade de propagação da onda eletromagnética (PAIXÃO et al., 2006) em um determinado material e no solo refere-se a média entre os valores de constante dielétrica de seus constituintes. Cezar et al. (2010) avaliou a utilização do GPR na determinação da umidade volumétrica e atributos físicos do solo em área de cultivo com cana-de-açúcar, e obteve resultados indicando que os valores de constante dielétrica aparente do solo apresentaram diferenças significativas em função da variação da umidade do solo. Oliveira Júnior e Medeiros (2008) observaram que o aumento do conteúdo de água gera aumento da constante dielétrica e redução na velocidade de propagação do sinal de GPR em amostras de areia. Wijewardana et al. (2017) verificaram que a intensidade do sinal da onda refletida diminui com o aumento dos níveis de condutividade elétrica. Os estudos de Dam et al. (2002) indicam que os óxidos de ferro podem diminuir a velocidade de propagação das ondas eletromagnéticas e,

consequentemente, apresentar influência na reflexão de ondas do GPR, o que gera dificuldades na interpretação do radargrama.

Topp et al. (1980) destacam que as propriedades eletromagnéticas do solo são controladas principalmente pelo teor de água. A água apresenta valor da constante dielétrica próximo a 80, o ar aproximadamente 1 e do material mineral e orgânico entre 3 e 7, portanto, pequenas variações na umidade volumétrica do solo (θ) podem resultar em grande mudança no valor de constante dielétrica do solo (PAIXÃO et al., 2006).

Devido ao grande contraste entre valores de constantes dielétrica da água, ar e minerais, o GPR pode ser usado para medir variações no teor de água do solo (LAMBOT et al., 2004). Há indícios que o conteúdo de óxidos de ferro correlaciona-se com o teor volumétrico de água. Essas variações do volume de água controlam a permissividade de penetração dos sinais de GPR, o que justifica a variação observada na velocidade da onda eletromagnética (DAM et al., 2002).

Para estimar a umidade volumétrica do solo com GPR, determina-se a constante dielétrica a partir das leituras nos radargramas (PAIXÃO et al., 2006). A maioria das equações que correlacionam valores de umidade e constante dielétrica foram estabelecidas para calibração de sondas TDR (*Time Domain Reflectometry*), em condições de laboratório utilizando solos de regiões temperadas. O modelo de Topp et al. (1980) e Ledieu et al. (1986), por exemplo, são de caráter empírico e não levam em consideração os efeitos dos componentes da constante dielétrica aparente do solo e das propriedades físicas do solo. O modelo de Roth et al. (1990) é de natureza semi-empírica ou física, em que a constante dielétrica aparente é fragmentada em seus componentes da matriz do solo, do ar e da água, que pode ser livre e adsorvida. Outros modelos, como o de Malicki et al. (1996), envolve o parâmetro físico densidade do solo.

O conhecimento do teor de água do solo é fundamental a qualquer estudo que envolva as propriedades do sistema solo-água-planta, em especial na tomada de decisões quanto ao manejo do solo e da irrigação (COELHO, 2006). A capacidade do solo em armazenar água varia em função de diversas características, tais como textura, tipo de argila, estrutura e teor de matéria orgânica (LEPSCH, 2016). Informações sobre umidade do solo são obtidas em geral de maneira pontual, o que torna oneroso o estudo da distribuição espacial da umidade. Neste contexto, o objetivo deste estudo foi avaliar a aplicabilidade do uso do *Ground Penetrating Radar* na determinação da umidade volumétrica em transectos comparando equações e propor modelo de calibração.

2 MATERIAL E MÉTODOS

2.1. Caracterização da área de estudo

A área de estudo é parte da bacia da Ribeirão São Bartolomeu, no local denominado Equideocultura, no *Campus* da Universidade Federal de Viçosa, município de Viçosa, na região da Zona da Mata de Minas Gerais (20°44' S, 42°51' W) (Figura 1). A área total da microbacia é de 112 ha, com uso predominante de pastagem formada por *Brachiaria spp.* A região é caracterizada por um relevo forte ondulado, com altitude média de 650 m e precipitação pluvial média anual de 1.203 mm (NASCIMENTO et al., 2006). O clima, de acordo com o sistema de classificação de Köppen, é do tipo Cwb com inverno seco e verão chuvoso. A temperatura do mês mais frio é inferior a 18 °C e do mês mais quente, superior a 22 °C (NASCIMENTO et al., 2006).

De maneira geral, há predomínio de Latossolos Vermelho-Amarelos, ocupando os topos convexos da paisagem; de Cambissolos, presentes nas áreas mais íngremes e côncavas, em que o horizonte C está mais próximo à superfície e de Argissolos Vermelho-Amarelos, nos terraços (NASCIMENTO et al., 2006).

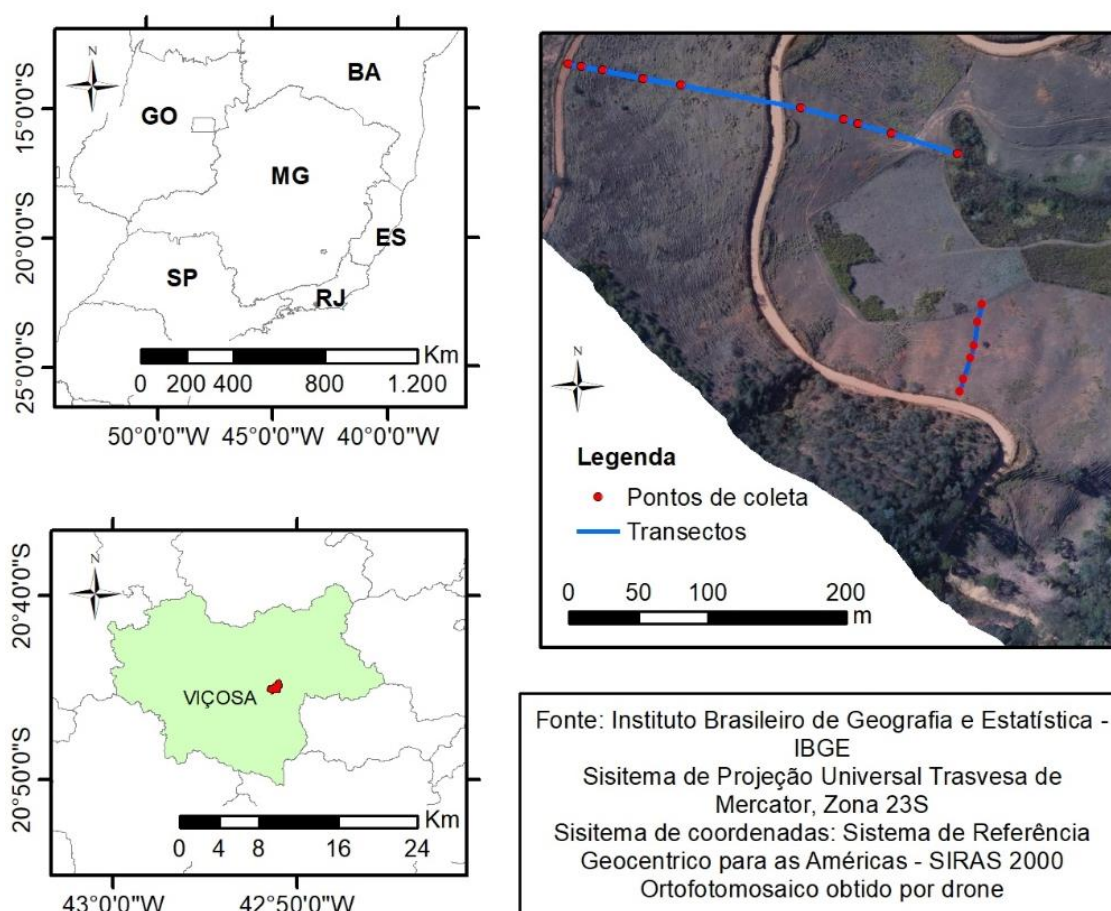


Figura 1. Localização da área de estudo, transectos e pontos de coleta.

2.2. Coleta de dados com GPR

Foram realizadas duas varreduras no período seco e chuvoso (setembro e fevereiro) em três transectos distintos na área de estudo (Figura 1), com 70, 100 e 130 metros de comprimento, no sentido do topo para o vale em relevo montanhoso. O equipamento utilizado foi o radar de penetração do solo da marca GSSI® modelo SIR-3000. Em cada área foram feitas varreduras com antenas blindadas de 200, 400 e 900 MHz. O aparelho foi configurado para valor de constante dielétrica de 13, com velocidade de 83 mm ns⁻¹, conforme recomendado no manual de instruções (GSSI, 2008). Os parâmetros utilizados na aquisição de dados foram: 50.000 varreduras por metro linear, 512 traços e 16 bits por amostra. A técnica de aquisição adotada foi a do afastamento constante (*common offset*), que consiste em passar o GPR em linha reta, utilizando odômetro para medir a distância percorrida.

2.3. Coleta das amostras de solo

Para a análise de umidade foram coletadas amostras deformadas nas profundidades de 0,00-0,20; 0,20-0,40; 0,40-0,60; 0,60-0,80 e 0,80-1,00 m.

A determinação da umidade do solo foi realizada por meio da diferença entre a massa fresca de solo e massa seca, pelo método padrão de determinação de umidade do solo. A amostra foi coletada, armazenada em saco plástico vedado para reduzir as perdas de água por evaporação e, logo em seguida, os valores de massa foram aferidos. As amostras foram levadas a estufa com circulação de ar a 105 °C, por 48 horas. Após estabilização do peso, foi realizado o cálculo da umidade conforme equação 1, descrito em Teixeira et al. (2017).

$$U = \frac{m_{um} - m_{sec}}{m_{sec}}$$

Onde:

U = umidade volumétrica do solo;

m_{um} = massa da amostra úmida;

m_{sec} = massa da amostra seca.

2.4. Processamento dos radargramas

Os radargramas (Figura 2) foram processados por meio do software RADAN 7®, disponibilizado pelo fabricante. Foram realizados os seguintes processamentos para melhoria do sinal eletromagnético nos radargramas, conforme sugerido por *Geophysical Survey Systems, Inc.* (GSSI) (2008):

1. Correção do tempo zero: remoção da faixa aérea entre a antena e a superfície (Figura 3);

2. Remoção do plano de fundo (*Background*): filtro aplicado para remover interferências de alta frequência. O ruído de fundo horizontal degrada a resolução dos dados GPR e está associado às interferências de ondas eletromagnéticas de alta frequência. Foi realizada a conversão de tempo em profundidade, utilizando modelos de velocidade média de propagação da onda através dos meios (Figura 4);

3. Filtro passa banda: filtro horizontal para remoção de ruídos e/ou reflexões múltiplas. O processamento de deconvolução elimina ruídos de baixa frequência melhorando a resolução dos dados GPR (Figura 5);

4. FIR (*Finite Impulse Response*): filtro vertical para redução de baixa e alta frequência. É um filtro digital linear caracterizado pela duração finita de seu sinal de entrada. O método retangular utilizado, *boxcar*, baseia-se em uma função de média simples determinada por meio da zona de corte de frequências desejada, chamada passa banda, e aplicada a janelas amostrais retangulares; e

5. Ganhos: adição de sinal para amplificar a resolução da imagem, possibilitando a identificação de refletores de baixa intensidade (Figura 6).

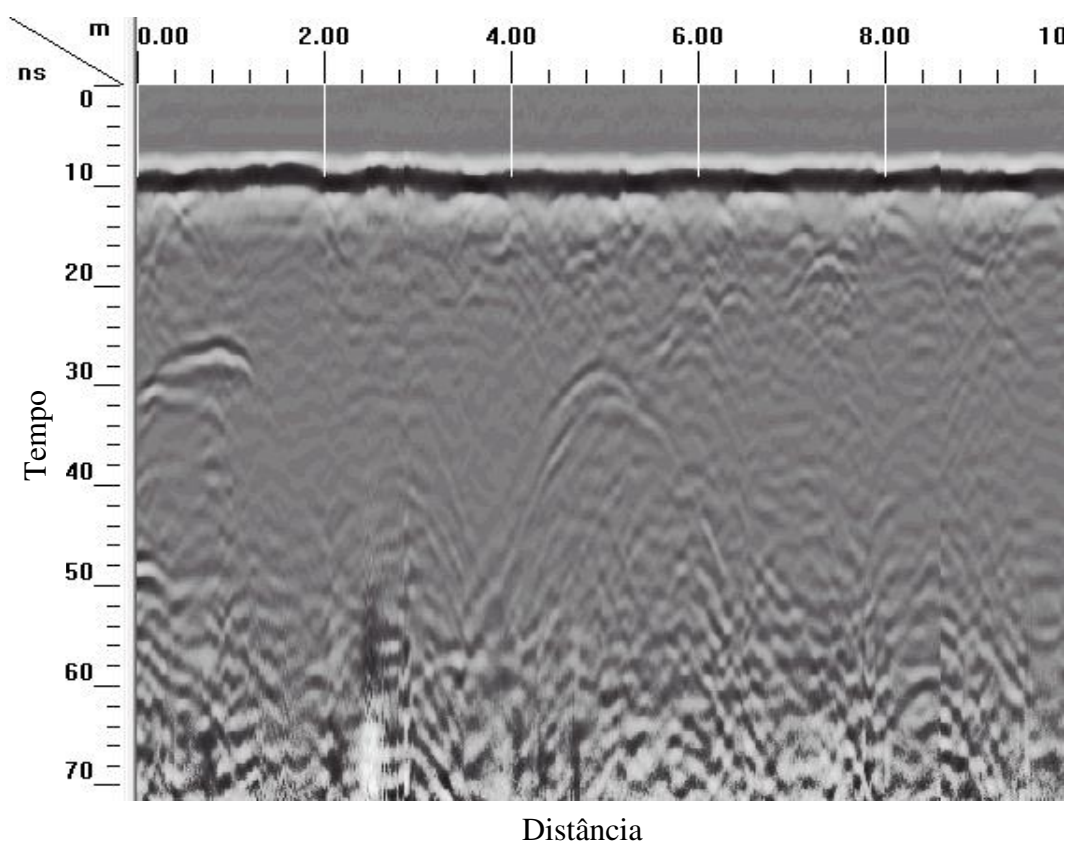


Figura 2. Radargrama bruto obtido com antena de 400 MHz.

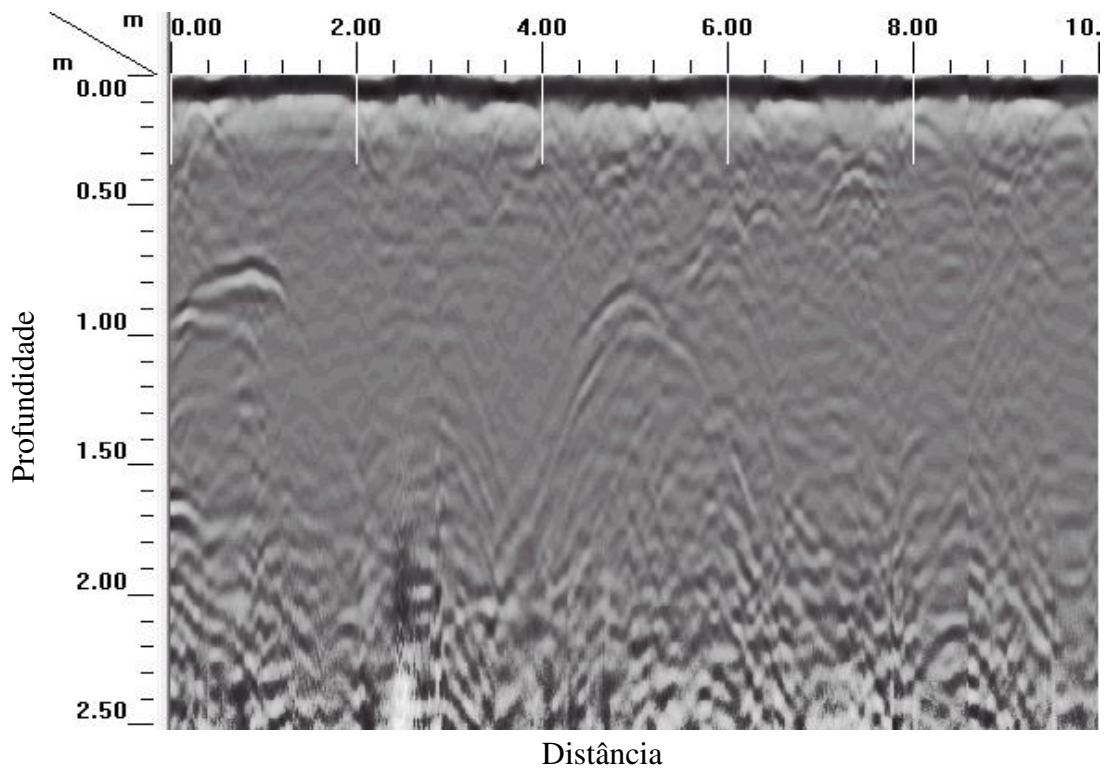


Figura 3. Radargrama em processamento – após remoção do tempo zero.

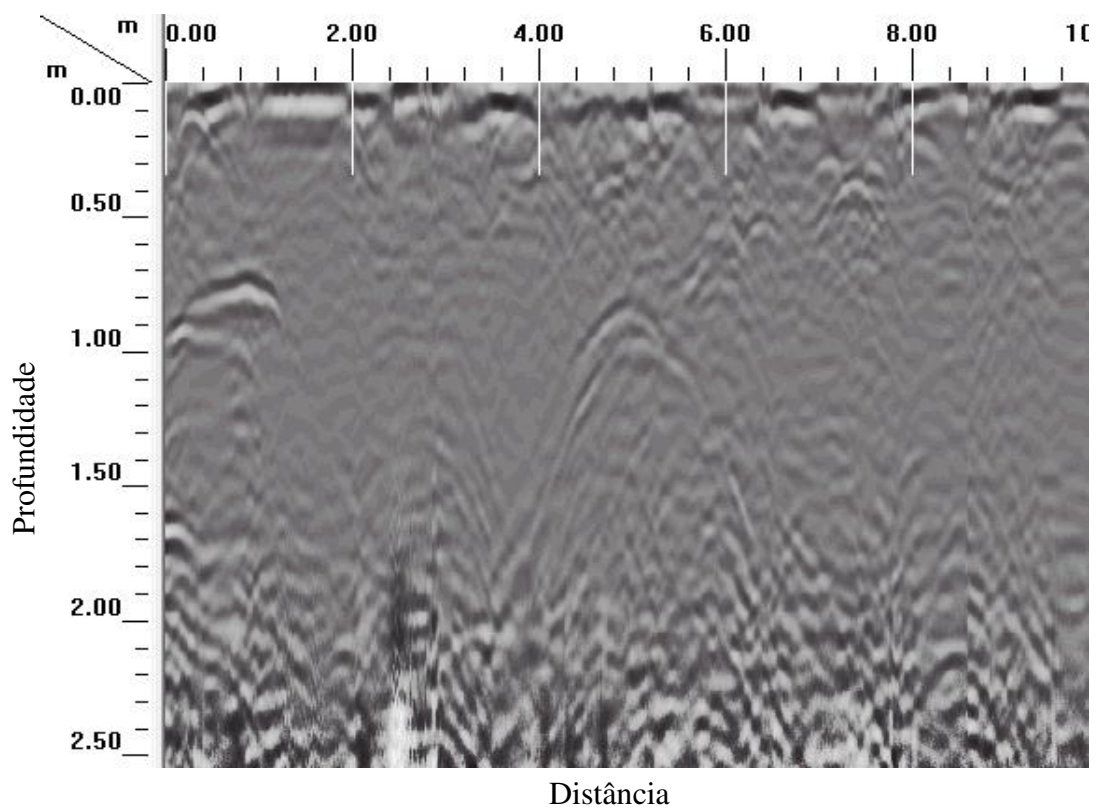


Figura 4. Radargrama em processamento – após remoção do plano de fundo.

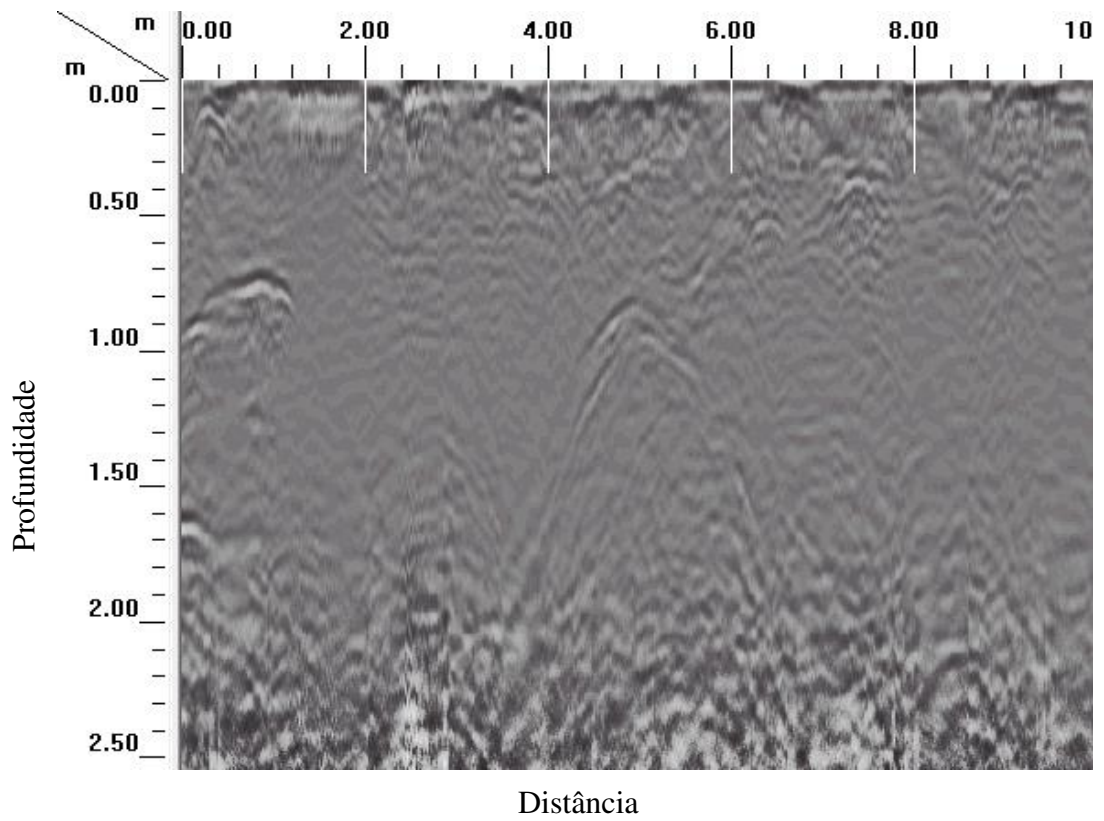


Figura 5. Radargrama em processamento – após remoção de ruído.

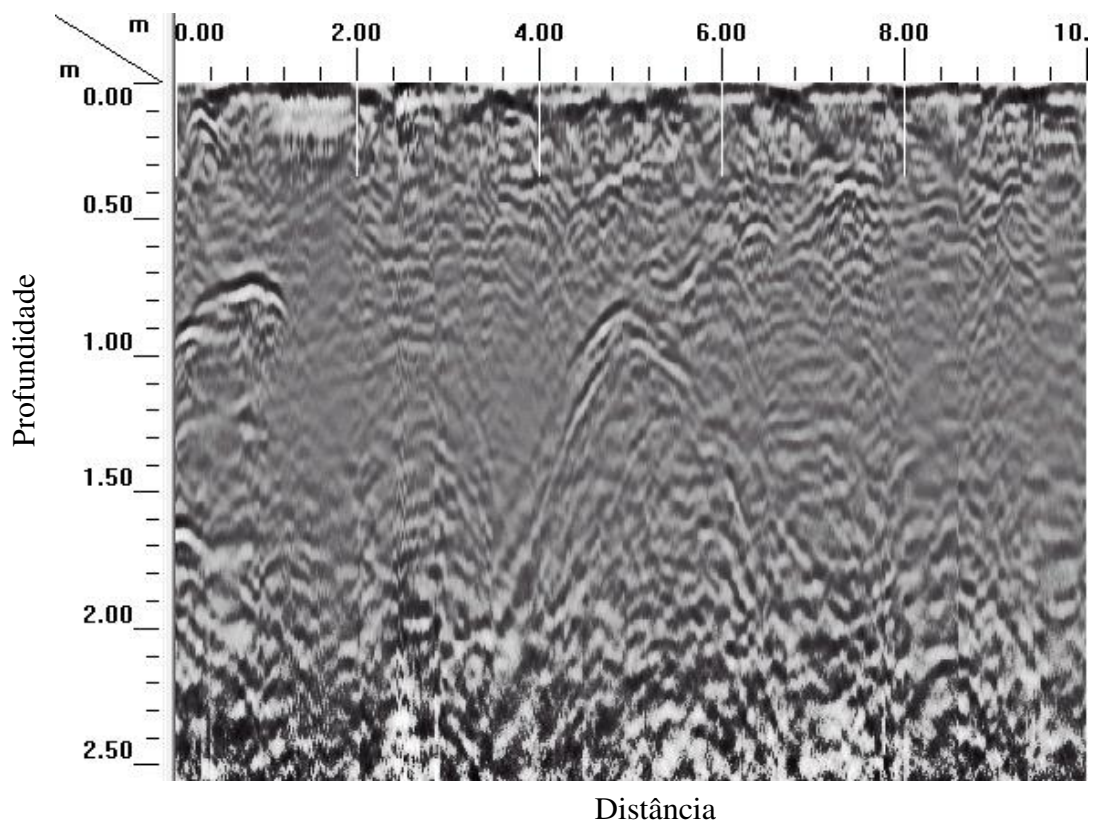


Figura 6. Radargrama em processamento – após aumento de ganhos.

2.5. Estimativa dos valores de umidade volumétrica a partir dos radargramas

Nos mesmos pontos e profundidades onde foram coletadas as amostras de solos para determinação da umidade em laboratório, foram selecionadas nos radargramas hipérboles representativas visando estimar o valor da constante dielétrica (K) do ponto amostral por meio de leitura direta nos radargramas, no software RADAN 7[®], (Figura 7).

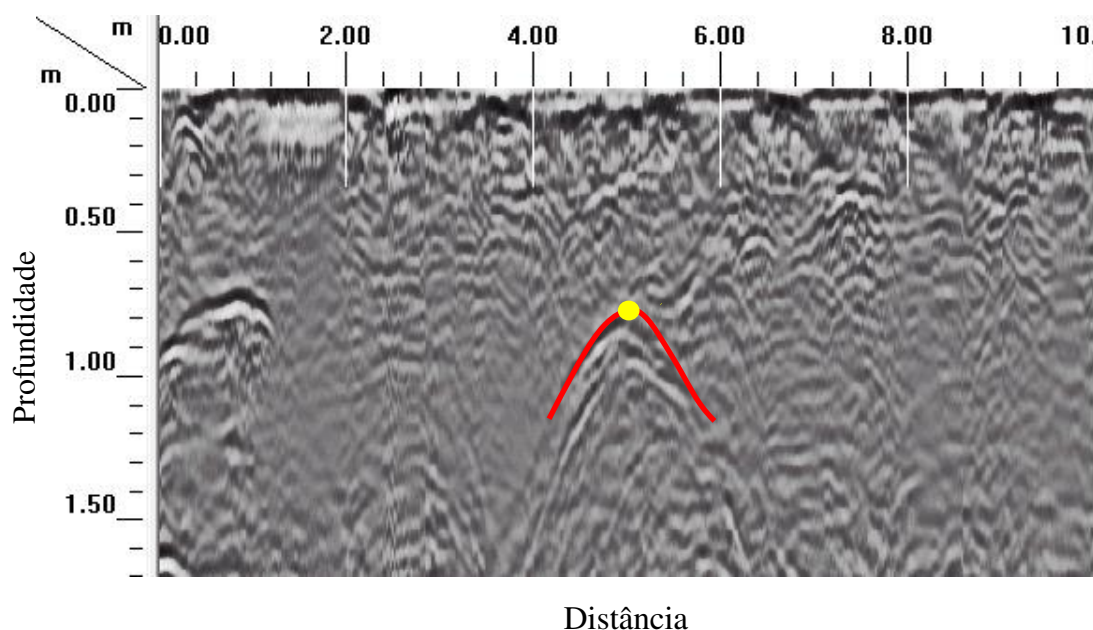


Figura 7. Coleta da hipérbole de reflexão. (O traço vermelho representa a hipérbole de reflexão e o ponto em amarelo, o ponto de coleta do valor de constante dielétrica)

Na estimativa da umidade volumétrica dos solos (U), a partir dos dados do GPR, a constante dielétrica obtida por meio dos radargramas foi usada conforme as seguintes equações:

- **Equação de Topp:** estabelece uma relação entre valores de constante dielétrica e umidade do solo. Proposta por Topp et al. (1980), é uma relação empírica obtida em laboratório por meio de dados experimentais com materiais de diferentes texturas. Os autores observaram que a constante dielétrica dos solos não era afetada significativamente por alterações na temperatura, textura, densidade e conteúdo de sais. O modelo consiste em uma equação polinomial do terceiro grau descrita matematicamente por:

$$U = -5,3 \times 10^{-2} + 2,92 \times 10^{-2}K + 4,3 \times 10^{-6}K^3$$

Onde:

U = umidade volumétrica do solo;

K = constante dielétrica observada no ponto de coleta.

- **Equação de Ledieu:** a equação proposta por Ledieu et al. (1986) estabelece uma relação linear entre umidade e a raiz quadrada da constante dielétrica, descrita por:

$$U = 0,1138\sqrt{K} - 0,1756$$

Onde:

U = umidade volumétrica do solo;

K = constante dielétrica observada no ponto de coleta.

- **Equação de Roth:** o estudo de Roth et al. (1990) visou estabelecer a existência de uma função de calibração universal. No estudo, foram utilizadas amostras de onze solos minerais e sete orgânicos com diferentes produtos e características físicas, químicas e propriedades magnéticas. Como resultado, obteve-se a equação:

$$U = -7,8 \times 10^{-2} + 4,48 \times 10^{-2}K - 1,95 \times 10^{-3}K^2 + 3,61 \times 10^{-5}K^3$$

Onde:

U = umidade volumétrica do solo;

K = constante dielétrica observada no ponto de coleta.

- **Equação de Malicki:** Malicki et al. (1996) propuseram um modelo relacionando umidade do solo e constante dielétrica, com inclusão do parâmetro físico densidade do solo (ρ). O modelo é descrito matematicamente por:

$$U = \frac{\sqrt{K} - 0,819\rho - 0,159\rho^2}{7,17 + 1,18\rho}$$

Onde:

U = umidade volumétrica do solo;

K = constante dielétrica observada no ponto de coleta; e

ρ = densidade do solo.

As equações estabelecem uma relação entre a constante dielétrica e umidade do solo. Os valores de constante dielétrica obtidos foram utilizados no ajuste de um modelo de regressão linear simples com interação para os dados locais (modelo local). Todas as equações foram utilizadas para comparar e calibrar o modelo local. Os resultados das análises de umidade em laboratório e os valores obtidos nos radargramas foram submetidos à análise de regressão. Foi utilizado o software R[®] (R CORE TEAM, 2017) a fim de obter o ajuste de equação de regressão entre umidade do solo e dados de constante dielétrica obtidos nos radargramas.

3 RESULTADOS E DISCUSSÃO

3.1 Coleta dos valores de constante dielétrica

Os estudos de Rossel et al. (2011) apontam que ondas de rádio com frequências mais elevadas aumentam a resolução da imagem (radargrama) e, em contrapartida, há redução da profundidade de penetração do sinal eletromagnético. Ao avaliar o alcance e a eficiência das três antenas, observou-se que a antena de 200 MHz atingiu em média 3,00 m de profundidade, enquanto as antenas de 400 MHz e 900 MHz alcançaram 2,00 e 0,60 m, respectivamente (Figura 8). O alcance da frequência de 200 MHz foi menor do que o esperado, o que é explicado pelo fato da onda do radar ser atenuada devido as perdas à medida que avança em profundidade em materiais condutores, como é o caso de solos argilosos da área de estudo. Nesses casos, a taxa de atenuação é maior e a onda pode penetrar a menor distância (GSSI, 2008).

Independente da frequência das antenas, as metodologias utilizando GPR são incapazes de distinguir mudanças sutis nas camadas do solo (DOOLITTLE; BUTNOR, 2009). A umidade do solo varia em função de diversos fatores tais como textura, atividade biológica, topografia e vegetação (DIONISIO, 2010; LEPSCH, 2016), e esses fatores dificultam a exatidão de metodologias de campo na determinação da umidade do solo.

Quanto maior a frequência da antena, menor o comprimento de onda, maior a riqueza de detalhes no radargrama e menor a profundidade de alcance do sinal (TRAN et al., 2015). A antena de 900 MHz gerou radargramas mais detalhados quando comparada às antenas de 200 MHz e 400 MHz, entretanto, apresentaram muitos ruídos mesmo após processamentos para a sua remoção, o que dificultou o uso das imagens para a coleta de hipérboles de reflexão. Os radargramas gerados com a antena de 200 MHz apresentaram menor número de hipérboles de reflexão próximas aos pontos de amostragem, o que limitou sua utilização nos cálculos de umidade.

Em decorrência dos resultados acima mencionados, foram realizados testes coletando-se 25 hipérboles em cada frequência. Os cálculos de umidade, utilizando-se os valores de K, foram efetuados e observou-se elevada dispersão dos dados coletados com as antenas de 200 MHz e 900 MHz. Diante disto, a metodologia para o cálculo de umidade foi empregada aos radargramas adquiridos somente com a antena de 400 MHz, onde os eventos de ondas aéreas e onda direta no solo foram facilmente distinguidos e a presença de hipérboles de reflexão nos radargramas mais evidente.

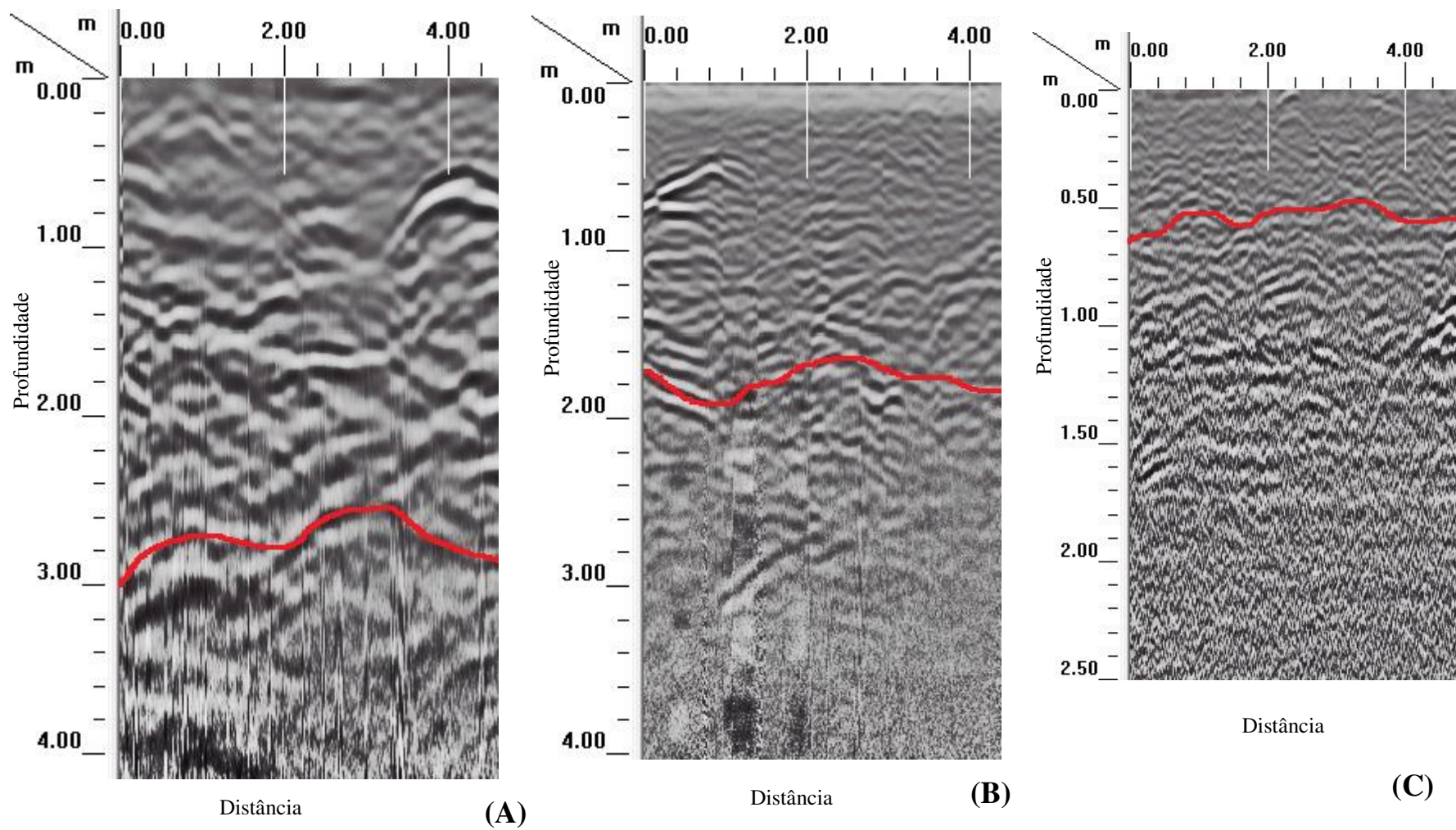


Figura 8. Comparação da profundidade de penetração do sinal eletromagnético em radargramas obtidos com antena distintas. Linha vermelha representa a profundidade máxima do sinal. Radargrama obtido com antena de 200 MHz (A), 400 MHz (B) e 900 MHz (C).

3.2 Umidade volumétrica - método padrão

A umidade volumétrica do solo aumentou em profundidade ao longo do perfil nos dois períodos de coleta. Observou-se incremento na umidade de 0,17 kg kg⁻¹ até 0,32 kg kg⁻¹ em profundidade, para as coletas no período de fevereiro. Para as coletas no período de setembro, houve incremento de 0,10 kg kg⁻¹ até 0,28 kg kg⁻¹ (Tabela 1). A umidade média nos dois períodos em todas as profundidades foi de 0,23 kg kg⁻¹ (Tabela 2).

Na camada de 0,00-0,20 m foram observados os menores valores de umidade nos dois períodos de coleta. A camada superficial tende a apresentar menor umidade por estar sujeita a maior atividade de absorção por raízes, maior evaporação e perdas nas trocas com a atmosfera (VAN LIER, 2010). Os maiores valores de umidade foram observados nas coletas mais profundas, entre 0,60-1,00 m.

3.3 Umidade volumétrica a partir dos radargramas

Na profundidade 0,4 e 0,60 m, os valores de umidade encontrados utilizando os modelos são mais similares do que nas demais, quando comparado ao método padrão (Tabela 1). Na camada superficial, todos os modelos superestimaram os valores de umidade, o que pode ser justificado pela dificuldade de obtenção de hipérbolos de reflexão nos primeiros centímetros da imagem. Este resultado está de acordo com o trabalho de Cezar et al. (2012) que, em um estudo utilizando o GPR para avaliar propriedades físicas do solo, não obteve resposta satisfatória nas investigações rasas, seja pela baixa frequência da antena ou textura muito argilosa. Grande quantidade de argila atenua a propagação das ondas eletromagnéticas devido a sua condutividade elevada, o que reduz a profundidade de investigação (DAVIS; ANNAN, 1989; CEZAR et al., 2010; CEZAR et al. 2012).

Barros Júnior (2011) destaca que para estimar a umidade volumétrica do solo com o uso do GPR fatores como classe textural do solo, aquisição de dados no campo e de seu processamento em laboratório são de fundamental importância na qualidade do resultado final.

Os valores de constante dielétrica obtidos nos radargramas foram utilizados para o cálculo da umidade volumétrica do solo, empregando as equações de Topp et al. (1980), Malicki et al. (1986), Roth et al. (1990) e Ledieu et al. (1996) e para proposição de uma nova equação.

O modelo Topp et al. (1980) apresentou desempenho similar ao método padrão (Tabela 1). Esses resultados estão coerentes com o trabalho de Paixão et al. (2006) que, utilizando antenas de 50 MHz, 100 MHz e 200 MHz, obteve valores de umidade próximos

aos obtidos com o método laboratoriais. Resultados semelhantes foram observados por Shamir et al. (2016), que obteve precisão dos resultados de umidade do solo entre 1% e 5%, em uma análise espacial até 1,5 m de profundidade em solo de textura franco argilosa (com cerca de 30% de argila, 30% de silte e 40% de areia) e textura areia franca (com cerca de 10% de argila, 5% de silte e 85% de areia) comuns em áreas rurais do sul de Israel.

3.4 Proposição do modelo

Gerou-se um modelo de regressão linear simples com interação para os dados locais afim de obter a equação de calibração. A equação de calibração proposta considera a umidade função da constante dielétrica e da profundidade de coleta das amostras, e a interação entre esses fatores. O modelo obtido é descrito por:

$$U = -7,637 \times 10^{-5} + 1,849 \times 10^{-2}K + 2,019 \times 10^{-3}P - 1,349 \times 10^{-4}KP$$

Onde:

U = umidade volumétrica do solo;

K = constante dielétrica observada no ponto de coleta; e

P = profundidade média de coleta da amostra de solo (cm).

Com o uso da equação proposta obteve-se o coeficiente de determinação (R^2) de 0,59 (Tabela 2). O desempenho superior do modelo indica a necessidade de estabelecer metodologias de calibração para as condições locais. Shamir et al. (2016) destacam que técnicas utilizando GPR atualmente necessitam de calibração, testes em laboratório e no campo para adequada compreensão das relações entre as propriedades físicas, permissividade dielétrica e umidade do solo.

Tabela 1: Valores mínimos, máximos e médios de umidade obtidos pelo método padrão (kg kg^{-1}), constante dielétrica (K) e umidade (kg kg^{-1}) obtidos por meio das equações de Topp et al. (1980), Ledieu et al. (1986), Roth et al. (1990) e Malicki et al. (1996), por período de coleta (meses de Fevereiro e Setembro) em profundidade

Fevereiro																		
Profundi- dade (m)	Umidade (kg . kg^{-1})			K			Topp et al. (1980)			Ledieu et al. (1986)			Roth et al. (1990)			Malicki et al. (1996)		
	Mín.	Máx.	Média	Mín.	Máx.	Média	Mín.	Máx.	Média	Mín.	Máx.	Média	Mín.	Máx.	Média	Mín.	Máx.	Média
0,00-0,20	0,17	0,32	0,24	8,39	15,95	12,35	0,16	0,29	0,23	0,15	0,28	0,22	0,18	0,29	0,24	0,18	0,30	0,25
0,20-0,40	0,16	0,31	0,23	8,65	14,11	11,91	0,16	0,26	0,22	0,16	0,25	0,22	0,19	0,27	0,24	0,18	0,28	0,24
0,40-0,60	0,17	0,28	0,24	10,52	15,95	12,87	0,20	0,29	0,24	0,19	0,28	0,23	0,22	0,29	0,25	0,22	0,30	0,26
0,60-0,80	0,20	0,28	0,26	9,51	16,64	13,47	0,18	0,30	0,25	0,18	0,29	0,24	0,20	0,29	0,26	0,20	0,31	0,27
0,80-1,00	0,25	0,32	0,27	12,12	18,17	14,38	0,23	0,32	0,27	0,22	0,31	0,26	0,24	0,31	0,27	0,25	0,34	0,28
Setembro																		
0,00-0,20	0,10	0,22	0,16	5,06	12,2	9,5	0,08	0,23	0,18	0,08	0,22	0,17	0,10	0,24	0,20	0,11	0,25	0,20
0,20-0,40	0,13	0,25	0,19	6,45	13,57	10,34	0,11	0,25	0,19	0,11	0,24	0,19	0,14	0,26	0,21	0,14	0,27	0,21
0,40-0,60	0,2	0,26	0,23	9,33	16,22	12,37	0,18	0,29	0,23	0,17	0,28	0,22	0,20	0,29	0,24	0,20	0,31	0,25
0,60-0,80	0,17	0,28	0,24	11,27	18,17	13,76	0,21	0,32	0,25	0,21	0,31	0,25	0,23	0,31	0,26	0,23	0,34	0,27
0,80-1,00	0,23	0,27	0,25	11,43	17,38	14,53	0,22	0,31	0,27	0,21	0,30	0,26	0,23	0,30	0,27	0,23	0,32	0,28

Tabela 2: Coeficiente de determinação e estatística descritiva dos valores de umidade obtidos pelo método padrão (kg kg^{-1}), constante dielétrica (K) e valores de umidade (kg kg^{-1}) obtidos por meio das equações de Topp et al. (1980), Ledieu et al. (1986), Roth et al. (1990) e Malicki et al. (1996).

	Mínimo	Máximo	Média	Desvio padrão	Coeficiente de variação	R ²
Umidade (kg.kg^{-1})	0,10	0,32	0,23	0,04	0,18	-
K	5,06	18,17	12,55	2,26	0,18	-
Topp et al. (1980)	0,08	0,32	0,23	0,04	0,17	0,5330
Ledieu et al. (1986)	0,08	0,30	0,22	0,03	0,16	0,5322
Roth et al. (1990)	0,10	0,30	0,24	0,03	0,13	0,5389
Malicki et al. (1996)	0,10	0,33	0,25	0,03	0,15	0,5322
Modelo local	0,11	0,26	0,22	0,037	0,17	0,5988

4 CONCLUSÕES

Os valores de constante dielétrica obtidos com a antena de 400 MHz, quando utilizados no cálculo da umidade do solo, apresentaram resultados similares aos da umidade volumétrica padrão quando comparados com os obtidos com as antenas de 200 e 900 MHz.

A equação desenvolvida neste trabalho apresenta melhor desempenho na estimativa da umidade do solo quando comparado aos demais modelos.

Os valores de constantes dielétricas determinados por meio da análise dos radargramas permitem estimar satisfatoriamente a umidade volumétrica em solo intemperizado e argiloso, otimizando estudos que necessitam de valores de umidade como uma variável com distribuição contínua.

REFERÊNCIAS

- BARROS JÚNIOR., V.L.M. **Utilização do radar de penetração em dois solos do norte fluminense.** Universidade Estadual do Norte Fluminense Darcy Ribeiro - UENF: Dissertação (mestrado em Produção Vegetal). 72p. 2011.
- BRANDÃO, V. S.; CECÍLIO, R. A.; PRUSKY, F. F.; SILVA, D. D. **Infiltração da água no solo.** Editora UFV. 3ª edição, 2006
- CEZAR, E., NANNI, M. R., CHICATI, M. L., DERBOCIO FABRIO, F., TERUHIKO HATA, F., BERTY DE OLIVEIRA, R.. Uso de sistema GPR (*Ground Penetrating Radar*) na avaliação de atributos de um solo sob plantio de cana-de-açúcar. **Revista Brasileira de Ciência do Solo**, v. 34, n. 2, 2010.
- CEZAR, E.; NANNI, M. R.; CHICATI, M. L.; DE OLIVEIRA, R. B. Emprego de GPR no estudo de solos e sua correlação com métodos laboratoriais Employment of GPR in the study of soils and its correlation with laboratory methods. *Semina: Ciências Agrárias*, v. 33, n. 3, p. 979-987, 2012.
- COELHO, E. F., VELLAME, L. D. M., COELHO FILHO, M. A., LEDO, C. D. S. Desempenho de modelos de calibração de guias de onda acopladas a TDR e a multiplexadores em três tipos de solos. **Revista Brasileira de Ciência do Solo**, v. 30, n. 1, p. 23-30, 2006.
- DAM, R. L. V., SCHLAGER, W., DEKKERS, M. J., HUISMAN, J. A.. Iron oxides as a cause of GPR reflections. **Geophysics**, v. 67, n. 2, p. 536-545, 2002.
- DAVIS, J. L.; ANNAN, A. P. Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy. **Geophysical Prospecting**, Netherlands, v. 37, n. 5, p. 531-551, 1989
- DIONISIO, H. A. F. Erosão hídrica: Suscetibilidade do solo. **Revista Eletrônica Thesis**, n. 13, p. 15-25, 2010.
- TEIXEIRA, P. C.; DONAGEMMA, G. K.; FONTANA, A.; TEIXEIRA, W. G. Manual de métodos de análise de solo 3ª ed. rev. e ampl. Brasília, DF: Embrapa, 574 p., 2017.
- DOOLITTLE, J. A.; BUTNOR, J. R. **Soils, peatlands, and biomonitoring.** In: JOL, H. M. (Org.). *Ground Penetrating Radar: Theory and Applications*. 1. ed. Amsterdam: Elsevier Science, p. 179-202, 2009.
- FRANCELINO, M. R.; MOURA, P. A.; SCHAEFER, C. E. G. R.; MENDONÇA, B. A. F. **Estudo do permafrost e solos com padrões com radar de penetração no solo**

(GPR) na ilha Livingston, Antártica Marítima. In: XXXIII Congresso Brasileiro de Ciência do Solo, 2011, Uberlândia. Anais do XXXIII Congresso Brasileiro de Ciência do Solo. Viçosa: SBCS, 2011

GSSI. Radan Software: **User's Manual.** Salem, USA: GSSI, 139 p.,2008.

LAMBOT, S., RHEBERGEN, J., VAN DEN BOSCH, I., SLOB, E. C., VANCLOOSTER, M. Measuring the soil water content profile of a sandy soil with an off-ground monostatic ground penetrating radar. **Vadose Zone Journal**, v. 3, n. 4, p. 1063-1071, 2004.

LANDEIRO, V. L. Coordenação de Pesquisas em Ecologia. Introdução ao uso do programa R. **Manaus: Instituto Nacional de Pesquisas da Amazônia**, 2011.

LEDIEU, J.; DE RIDDER, P.; DE CLERCK, P. & DAUTREBANDE, S. A method for measuring soil water moisture by time-domain reflectometry. **Journal of Hydrology**, v. 88, n. 3-4, p. 319- 328, 1986.

LEPSCH, I. F. **Formação e Conservação dos Solos.** Oficina de Textos. São Paulo. 2016.

MALICKI, M.A.; PLAGGE, R.; ROTH, C.H. Improving the calibration of dielectric TDR soil moisture determination taking into account the solid soil. **European Journal of Soil Science**, v.47, n. 3, p.357- 366, 1996.

NASCIMENTO, M. C., RIVA, R. D., CHAGAS, C. D. S., OLIVEIRA, H. D., DIAS, L. E., FERNANDES FILHO, E. I., SOARES, V. P. Uso de imagens do sensor ASTER na identificação de níveis de degradação em pastagens. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 10, n. 1, p. 196-202, 2006.

OLIVEIRA JÚNIOR, J. G.; MEDEIROS, W. E. Influência do conteúdo de água, granulometria e minerais pesados sobre as reflexões de GPR em corpos arenosos inconsolidados. **Revista Brasileira de Geofísica**, v. 26, n. 4, p. 437-449, 2008.

PAIXÃO, M. S. G.; PRADO, R. L.; DIOGO, L. A. Análise do emprego do GPR para estimar o teor de umidade do solo a partir de um estudo na cidade de São Paulo. **Revista Brasileira de Geofísica**, v. 24, n. 2, p. 189-198, 2006.

R CORE TEAM. **R: a language and environment for statistical computing.** Vienna, Austria: R Foundation for Statistical Computing; 2017. 2017.

ROSSEL, R. V., ADAMCHUK, V. I., SUDDUTH, K. A., MCKENZIE, N. J., LOBSEY, C. Proximal soil sensing: an effective approach for soil measurements in space and time. **Advances in Agronomy**, Academic Press, v. 113, p. 237-282, 2011.

ROTH, K.; SCHULIN, R.; FLUHLER, H.; ATTINGER, W. Calibration of time-domain reflectometry for water content measurement using composite dielectric approach. **Water Resources Research**, v. 26, n. 10, p. 2267-2273, 1990.

SHAMIR, O., GOLDSHLEGER, N., BASSON, U., RESHEF, M.. Mapping spatial moisture content of unsaturated agricultural soils with Ground-Penetrating Radar. **The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 41, p. 1279, 2016.

TOPP, G. C.; DAVIS, J. L.; ANNAN, A. P. Electromagnetic of soil water content: measurements in coaxial transmission lines. **Water Resources Research**, v. 16, n. 3, p. 574-582, 1980.

TRAN, A. P., BOGAERT, P., WIAUX, F., VANCLOOSTER, M., LAMBOT, S. High-resolution space-time quantification of soil moisture along a hillslope using joint analysis of ground penetrating radar and frequency domain reflectometry data. **Journal of Hydrology**, v. 523, p. 252–261, 2015.

VAN LIER, Q. J. Física do solo. Viçosa, MG: Sociedade Brasileira de Ciência do Solo, 2010.

WIJewardana, Y. N. S., SHILPADI, A. T., MOWJOOD, M. I. M., KAWAMOTO, K., GALAGEDARA, L. W. Ground penetrating radar (GPR) responses for sub-surface salt contamination and solid waste: modeling and controlled lysimeter studies.

CAPÍTULO II

PREDIÇÃO DE TEXTURA E POROSIDADE DO SOLO POR *Ground Penetratin* *Radar E Machine Learning*

RESUMO

O sensor proximal GPR (*Ground Penetrating Radar*) é um método não-invasivo e não destrutivo, que permite extrair informações sobre várias propriedades do solo. Para isso o processamento dos dados gerados podem envolver um grande esforço informático. O *Machine Learning* (aprendizado de máquina) surge como alternativa dentro da ciência do solo, uma ferramenta que simplifica o ajuste de modelos por otimizar os parâmetros, possibilitando a interpretação de variáveis com maior nível de detalhamento e precisão. Visto que com o GPR é possível determinar a constante dielétrica do solo e a partir dela estimar a umidade do solo, e como essa propriedade tem estreita relação com a textura do solo, o objetivo deste estudo foi elaborar modelos de predição de variáveis texturais utilizando dados de GPR e técnicas de *Machine Learning*. O trabalho foi desenvolvido em uma área de pastagem em relevo fortemente ondulado, no município de Viçosa, Minas Gerais, onde foram realizadas duas varreduras em três transectos distintos, com 70, 100 e 130 m de comprimento, no sentido do topo para o vale em relevo montanhoso. Em cada uma das três áreas foram realizados transectos com antena blindada de 400 MHz para geração de radargramas, os quais foram processados no software RADAN 7[®]. Nos radargramas foram realizados diversos processos como aplicação de filtros, ganhos e migração, todos visando a melhoria do sinal eletromagnético e determinação da constante dielétrica. Para as análises de textura e porosidade do solo foram realizadas coletas em 26 pontos de amostras indeformadas nas profundidades de 0,00-0,10; 0,10-0,20; 0,20-0,30; 0,30-0,40 e 0,40-0,50 m e realizada a análise laboratorial padrão. Para predição de textura e porosidade dos solos utilizando dados do GPR, a constante dielétrica foi obtida por meio dos processamentos dos radargramas e, juntamente aos resultados das análises laboratoriais, foram utilizados cinco algoritmos de predição: *Linear Models* (LM), *Generalized Linear Models* (GLM), *Cubist*, *Support Vector Machine Radial Sigma* (SVM) e *Random Forest* (RF), para verificar qual modelo se adequava melhor as cinco bases de dados testadas. Foi utilizado o software estatístico R[®] para ajustar os modelos de regressão. A verificação do desempenho dos algoritmos foi realizada por método de validação cruzada do tipo *leave one out*. O desempenho dos algoritmos em cada treinamento e validação foi avaliado pelo coeficiente de determinação (R^2) e pela raiz do erro médio quadrático (RMSE), para cada variável. Testou-se combinações distintas de bases de dados para avaliar quais variáveis são mais importantes na elaboração de modelos preditivos. Os modelos SVM, RF, e *Cubist* apresentaram maior desempenho quando comparados aos modelos lineares. Inserir variáveis de terreno possibilita modelagens de maior precisão, com valores de R^2 de 0,84 e 0,79 para variável argila, como os modelos SVM e RF, respectivamente. Os valores de RMSE foram extremamente baixos para todos os modelos (inferiores a 0,04). A predição das variáveis texturais gerou resultados superiores à predição de valores de porosidade.

Palavras-chave: aprendizado de máquina, GPR, modelagem textural do solo.

ABSTRACT

GPR (Ground Penetrating Radar) proximal sensor is a non-invasive, non-destructive method that allows you to extract information about various soil properties. For this the processing of the generated data can involve a great effort in computer science. Machine Learning emerges as an alternative within soil science, a tool that simplifies the adjustment of models by optimizing the parameters, allowing the interpretation of variables with greater level of detail and precision. Since with GPR it is possible to determine the dielectric constant of the soil and from it to estimate the soil moisture, and since this property has a close relation with the soil texture, the objective of this study was to elaborate models of prediction of textural variables using data GPR and Machine Learning techniques. The work was carried out in a heavily undulating grassland in the municipality of Viçosa, Minas Gerais, where two sweeps were carried out on three distinct transects, with 70, 100 and 130 m in length, from the top to the valley in mountainous relief. In each of three areas were transects with 400 MHz shielded antenna for generation of radargrams, which were processed in RADAN 7® software. In the radargrams, several processes were applied, such as filters, gains and migration, all of them aimed at improving the electromagnetic signal and determining the dielectric constant. For soil texture and porosity analysis, samples were collected at 26 points of undisturbed samples at depths of 0.00-0.10; 0.10-0.20; 0.20-0.30; 0.30-0.40 and 0.40-0.50 m and performed the standard laboratory analysis. In order to predict soil texture and porosity using GPR data, the dielectric constant was obtained by means of the radargrams and, together with the results of the laboratory analysis, five predictive algorithms were used: Linear Models (LM), Generalized Linear Models (GLM), Cubist, Support Vector Machine Radial Sigma (SVM) and Random Forest (RF), to verify which model best suited the five databases tested. The R® statistical software was used to adjust the regression models. The performance verification of the algorithms was performed by a one-off cross-validation method. The performance of the algorithms in each training and validation was evaluated by the coefficient of determination (R^2) and root mean square error (RMSE), for each variable. Different combinations of databases have been tested to assess which variables are most important in the design of predictive models. The SVM, RF, and Cubist models presented higher performance when compared to linear models. Inserting field variables allows for more accurate modeling, with R^2 values of 0.84 and 0.79 for clay, such as the SVM and RF models, respectively. RMSE values were extremely low for all models (less than 0.04). The prediction of textural variables generated results superior to prediction of porosity values.

Key words: machine learning, GPR, soil textural modeling

1 INTRODUÇÃO

O emprego de novas tecnologias na ciência do solo pode diminuir os custos de obtenção de amostras em campo e produzir informações mais detalhadas sobre as propriedades físicas do solo (PAIXÃO et al. 2006), otimizando a aquisição e processamento de dados, como também melhorar o detalhamento das informações (BARROS JÚNIOR, 2011). O uso de técnicas geofísicas de inspeção do solo pode contribuir para o avanço e otimização de estudos pedológicos (PAIXÃO et al. 2006; HARTEMINK; MINASNY, 2014; ABREU, 2017), porém não são devidamente aplicadas na ciência do solo. Uma das ferramentas mais promissora é o sensor proximal *Ground Penetrating Radar* (GPR). Trata-se de um método não-invasivo e não destrutivo, que permite extrair informações ao longo do perfil do solo; constituindo uma técnica indireta para investigação da subsuperfície do solo (CEZAR et al., 2010).

O estudo da dinâmica da distribuição espacial de atributos físicos do solo é fundamental para determinação de práticas de manejo, classificação do solo e estudos pedológicos em geral (ZANZARINI, 2013).

O *Machine Learning* (aprendizado de máquina) surge dentro da ciência do solo como uma ferramenta que simplifica o ajuste de modelos por otimizar os parâmetros, possibilitando a interpretação de variáveis com maior nível de detalhamento e precisão (KIRKWOOD et al. 2016; HENGL et al. 2017). O aprendizado de máquina permite mapear centenas de variáveis do solo em paralelo de forma rápida e eficiente, exigindo pouca interação do operador (HENGL et al., 2017). De acordo com Kuhn e Johnson (2013), o grupo de modelos de alto desempenho inclui *Support Vector Machines* (SVMs), *Boosted Trees*, *Random Forests* e *Cubist*. É interessante utilizar modelos não paramétricos em comparação a modelos lineares em estudos nas diversas áreas do conhecimento.

Os atributos do terreno, que são as variáveis morfométricas obtidas com o processamento do modelo digital de elevação, são amplamente utilizados como preditores de diversas características do solo devido à alta correlação com os processos de formação do solo (BEHRENS et al., 2010). Declividade, face de exposição à radiação solar, curvatura do terreno, índice de rugosidade do terreno, entre outros, são exemplos de variáveis morfométricas utilizadas como preditores de características do solo (GALLANT; HUTCHINSON, 1997; SCHÜNEMANN et al., 2018).

Neste contexto, o objetivo desse estudo foi elaborar modelos de predição de variáveis texturais (areia grossa, areia fina, silte, argila, porosidade total, macroporosidade e microporosidade) em solos do município de Viçosa, Minas Gerais, utilizando dados de *Ground Penetrating Radar*, variáveis morfométricas de terreno e técnicas de *Machine Learning*.

2 MATERIAL E MÉTODOS

2.1. Caracterização da área de estudo

A área de estudo localiza-se no *Campus* da Universidade Federal de Viçosa, município de Viçosa, na região da Zona da Mata de Minas Gerais (20°44' S, 42°51' W). A área total da microbacia é de 112 ha, com uso predominante de pastagem formada por *Brachiaria spp.* A região é caracterizada por um relevo forte ondulado, com altitude média de 650 m e precipitação pluvial média anual de 1.203 mm (NASCIMENTO et al., 2006). De modo geral, há predomínio de Latossolos Vermelho-Amarelos, ocupando os topos convexos da paisagem; de Cambissolos nas áreas íngremes e côncavas e Argissolos Vermelho-Amarelos nos terraços (NASCIMENTO et al., 2006).

2.2. Coleta de dados

2.2.1. Coleta e análise de amostras de solo

Foram definidos 13 pontos de coleta de amostras por período de coleta (meses de fevereiro e setembro) totalizando 26 pontos de coleta. Para a determinação da porosidade do solo foram coletadas cinco amostras indeformadas, em cada ponto de coleta, nas profundidades de 0,00-0,10; 0,10-0,20; 0,20-0,30; 0,30-0,40 e 0,40-0,50 m. O mesmo solo foi utilizado para a análise granulométrica.

As análises das características físicas das amostras de solo foram realizadas conforme Teixeira et al. (2017). A análise granulométrica foi realizada pelo método da tamisagem e da pipeta. Amostras de solo foram dispersas quimicamente em solução de hidróxido de sódio (NaOH 10 mmol L⁻¹). A dispersão física realizada com agitação lenta, em agitador rotatório tipo Wagner, a 50 rotações por minuto, durante 16 h. Foram quantificados os teores de areia fina, areia grossa, silte e argila.

A determinação da densidade foi realizada pelo método do anel volumétrico conforme Teixeira et al. (2017). Amostras de solo com estrutura indeformada foram coletadas por meio de um anel de aço (Kopecky) de bordas cortantes e volume interno de aproximadamente 50 cm³. O material foi submetido a secagem em estufa a 105 °C e após

estabilização da temperatura, os valores de massa foram aferidos. A densidade corresponde ao quociente entre os valores de massa seca e o volume do anel.

A densidade de partículas foi analisada pelo método de determinação do volume de álcool necessário para completar capacidade de um balão volumétrico, contendo solo seco em estufa, conforme Teixeira et al. (2017). Amostras de solo secas em estufa foram transferidas para balão aferido de 50 mL. Adicionou-se álcool etílico, agitando bem o balão para eliminar as bolhas de ar que se formam e completou-se o volume do balão. A densidade de partícula corresponde ao quociente entre o valor de massa seca de solo e o volume de álcool.

A porosidade determina o volume de poros totais do solo. A porosidade total foi calculada conforme equação:

$$PT = 1 - (DS/DP)$$

Onde:

PT = Porosidade total

DS = Densidade do solo

DP = Densidade de partícula

A microporosidade foi determinada pelo método da mesa de tensão a -6 kPa e a macroporosidade pela diferença entre porosidade total e microporosidade, conforme equação (Teixeira et al., 2017):

$$Ma = PT - Mi$$

Onde:

Ma = Macroporosidade do solo

PT = Porosidade total

Mi = Microporosidade do solo

2.2.2. Coleta de dados com GPR

Foram realizadas duas varreduras nos meses de fevereiro e setembro, em três transectos distintos na área de estudo com 70, 100 e 130 m de comprimento, no sentido do topo para o vale em relevo montanhoso. O equipamento utilizado foi o radar de penetração do solo da marca GSSI® modelo SIR-3000. Em cada área foram feitas varreduras com antena blindada de 400 MHz. O aparelho foi configurado para valor de constante dielétrica de 13, com velocidade de 83 mm ns⁻¹, correspondendo à área de pastagem, conforme recomendado no manual de instruções do equipamento (GSSI, 2008). Os parâmetros utilizados na aquisição de dados foram: 50.000 varreduras por metro linear, 512 traços e 16 Bits por amostra. O método de varredura adotado foi o

common offset que consiste em passar o GPR em linha reta, utilizando odômetro para medir a distância percorrida.

2.2.3. Processamento dos radargramas

Os radargramas foram gerados por meio do software RADAN 7[®]. Foram realizados processamentos para melhoria do sinal eletromagnético nos radargramas conforme *Geophysical Survey Systems, Inc.* - (GSSI) (2008). Os processamentos básicos foram a correção do tempo zero e remoção do plano de fundo (*Background*). Para remoção de ruídos e/ou reflexões múltiplas foi utilizado o filtro horizontal. O filtro vertical FIR (*Finite Impulse Response*) foi utilizado para redução de baixa e alta frequência. Aplicou-se o filtro de ganhos para amplificar a resolução da imagem, possibilitando a identificação de refletores de baixa intensidade.

A partir da imagem do radargrama, nos mesmos pontos e profundidades onde foram coletadas as amostras de solos para determinação da granulometria e porosidade, foram selecionadas hipérboles representativas da área a fim de obter o valor da constante dielétrica (*K*) do ponto amostral.

2.2.4. Aquisição das covariáveis de morfometria

O Modelo Digital de Elevação (MDE) já disponível no Laboratório de Geoprocessamento (dados não publicados) foi obtido utilizando o equipamento Laser scanner terrestre modelo VZ-1000, as Riegel. O equipamento foi configurado para obter alta densidade de pontos, com alcance de 1.400 m. Os pontos de instalação do equipamento foram definidos de forma a permitir no mínimo 60% de sobreposição entre visadas, objetivando o melhor ajuste da nuvem de pontos. A nuvem de ponto gerada foi processada no programa Riscan-Pro.

Foram geradas características de relevo a partir do MDE utilizando o pacote “RSAGA” (BRENNING, 2008) e a metodologia descrita por Schünemann et al. (2018), que de forma automatizada gerou as variáveis morfométricas listadas na Tabela 1.

Tabela 1: Covariáveis morfométricas do terreno utilizadas para análise de importância.

Variável explicativa	Descrição breve
Classificação da curvatura	Divide as curvaturas em oito classes
Convexidade da superfície terrestre	É calculado como a razão entre o número de células que têm curvatura positiva (células convexas) para o número de todas as células válidas dentro de um raio de pesquisa específico
Curvatura da linha de fluxo	
Curvatura tangencial	Descreve o primeiro mecanismo de acumulação
<i>Landforms</i>	Formas do relevo

<i>Module Topographic Wetness Index (TWI)</i>	Descreve a tendência de cada célula em acumular água em função do relevo
<i>Normalized height</i>	Distância vertical entre a base e o cume do declive normalizada
<i>Slope height</i>	Distância vertical entre a base e o cume do declive
<i>Solrad_diffuse2</i>	Radiação solar difusa incidente no mês de junho
<i>Solrad_dur1</i>	Duração da radiação solar no mês de janeiro
<i>Solrad_dur2</i>	Duração da radiação solar no mês de junho
<i>Textura da superfície terrestre</i>	Divide a textura da superfície em 8, 12 ou 16 classes
<i>Topographic position index</i>	Diferença vertical entre a base e o cume do declive padronizada

2.2.5. Análises estatísticas

As covariáveis físicas e morfométricas foram agrupadas e então realizada a normalização com o método z-score, utilizando a função *preProcess* presente na biblioteca “caret” (KUNH, 2008). Utilizou-se a função *high_correlation* do pacote “caret” para indicar as covariáveis altamente correlacionadas. O *high_correlation* determina a correlação entre todas as covariáveis entre si e exclui as covariáveis em que apresentam alta correlação acima de um determinado limite, sendo este definido pelo operador. Nesse estudo foi estabelecido o limite de 99% de correlação. Após a remoção das covariáveis altamente correlacionadas foi realizado o treinamento dos algoritmos.

A escolha dos algoritmos para o cálculo da predição das características texturais e porosidade foi realizada considerando modelos de uso mais comum na literatura para regressões, de forma a abranger uma gama variada, desde os modelos lineares mais simples, árvores de decisão e *Support Vector Machine*. Foram utilizados cinco algoritmos de predição: *Linear Models* (LM), *Generalized Linear Models* (GLM), *Cubist*, *Support Vector Machine Radial Sigma* (SVM) e *Random Forest* (RF). A verificação do desempenho dos algoritmos foi realizada por método de *leave one out*.

Regressões lineares são utilizadas para estimar valores reais com base em variáveis contínuas, estabelecendo uma relação entre variáveis dependentes e independentes. Esta relação é representada por uma equação linear. Os modelos lineares (LM) são os mais simples e permitem replicar a equação gerada de forma prática sem a necessidade de grande demanda computacional, dependendo do número de variáveis utilizadas. Os modelos lineares generalizados (GLM) estendem os modelos de regressão linear para as distribuições de resposta não normal. Formulados com o objetivo de unificar vários outros modelos estatísticos, incluindo regressão linear, GLM são modelos flexíveis que permitem que as variáveis de resposta apresentem modelos de distribuição de erro diferentes de uma distribuição normal (GOLDBURD et al., 2016).

Se a relação entre variável dependente e independente for próxima de um modelo linear, a regressão linear supera os resultados dos modelos baseados em árvore de decisão. Contudo, se houver não linearidade, dados não paramétricos e relação complexa entre as variáveis, os modelos de árvore de decisão, de modo geral, superarão os métodos clássicos de regressão (KUHN; JOHNSON, 2013).

Modelos de árvores de decisão, como *Random forest* e *Cubist*, podem apresentar resultados superiores a modelos lineares, por melhor representarem relações não-lineares complexas em grandes conjuntos de dados (KUHN; JOHNSON, 2013; HENGL et al. 2017). Os métodos baseados em árvore possibilitam gerar modelos preditivos com alta precisão e facilidade de interpretação. Árvores de decisão têm a vantagem de serem metodologias não paramétricas, não há uma hipótese sobre a distribuição do espaço e a estrutura do classificador, o que permite adequar os modelos a diversas bases de dados (BREIMAN, 2001; KUHN; JOHNSON, 2013).

O modelo *Random Forest* (RF) é um método de classificação e regressão, desenvolvido por Breiman (2001). É um método versátil capaz de executar tarefas de regressão e de classificação, sendo muito utilizado na ciência do solo (BREIMAN, 2001; HENGL et al. 2017). Uma das vantagens do *Random forest* é lidar com grande conjunto de dados, identificar as variáveis mais significativas e determinar a importância das variáveis. São classificadores do tipo árvores de decisões composto por várias árvores. Cada árvore é construída a partir da seleção aleatória de covariáveis preditoras e pontos de treinamento. Não há duas árvores exatamente iguais, contudo, todas começam com uma estrutura semelhante e estão relacionadas umas às outras. As árvores são construídas usando um subconjunto aleatório dos principais preditores em cada divisão da árvore (KUHN; JOHNSON, 2013). Em RF, todas as árvores são criadas independentes e cada árvore contribui igualmente para o modelo final. É vantajoso por reduzir o tempo de processamento, pois as árvores são independentes e podem ser facilmente processadas paralelamente (KUHN; JOHNSON, 2013). *Random Forest* é um algoritmo robusto, insensível a ruídos nos valores das variáveis e ao *overfitting* (super ajuste). Destaca-se também pela capacidade de fornecer medidas de importância das covariáveis utilizadas na predição (BREIMAN, 2001).

Já o algoritmo *Cubist* também é baseado em estrutura de árvore de regressão, gerando um modelo que contém várias regressões. No processo de criação das árvores, o algoritmo estabelece modelos de regressões lineares nas folhas terminais. Além disso, existem modelos lineares intermediários em cada etapa da árvore. As árvores são criadas em sequência, ou seja, a partir da primeira árvore, as demais são criadas usando versões

ajustadas para o resultado do conjunto de treinamento. Portanto, o resultado da regressão da folha final da primeira árvore é utilizado para ajustar o nó inicial da segunda árvore e assim sucessivamente. A previsão final é uma média simples das previsões geradas em cada árvore do modelo (MINASNY et al., 2013, KUHN et al., 2016).

O *Support Vector Machines* (SVM) é um método de aprendizagem de máquina usado para classificação e regressão. O objetivo do algoritmo é encontrar uma função que aproxima bem os pontos de treinamento pela minimização dos erros de predição. A vantagem de utilizar SVM é extrair padrões complexos de modelagem dos dados sem a necessidade de conhecer o comportamento prévio dos dados. SVM identifica padrões nos dados e cria uma estrutura para a classificação, um hiperplano (ou linha) que separa os dados em duas categorias. Quando os dados apresentam distribuição linear ou próxima a linear, as SVMs serão eficazes na classificação, já que os dados serão linearmente separáveis (LORENA; CARVALHO, 2007).

Os cinco algoritmos foram testados em cinco combinações (cinco tratamentos) distintas da base de dados afim de determinar modelos de melhor explicação e fácil obtenção de covariáveis. As covariáveis de interesse foram inseridas na base de dados de forma a possibilitar a elaboração de modelos de complexidade variada e melhor ajuste. As bases de dados utilizadas estão descritas na Tabela 2.

Tabela 2: Descrição das variáveis explicativas (X) utilizadas nas bases de dados

Tratamentos	Variáveis explicativas (X)	Variável resposta (Y)
01	Constante dielétrica (K)	
02	Constante dielétrica (K) Período de coleta	Areia grossa
03	Covariáveis morfométricas	Areia fina
04	Constante dielétrica (K) Período de coleta Covariáveis morfométricas	Silte Argila Macroporosidade
05	Constante dielétrica (K) Período de coleta Covariáveis morfométricas Remoção das amostras da camada superficial	Microporosidade Porosidade total

O desempenho dos algoritmos de predição foi verificado por separação aleatória entre amostras do conjunto de treinamento e validação, sendo os algoritmos avaliados pelo método *leave one out* modificado. Esse método foi aplicado considerando-se como ponto amostral cada ponto de coleta, com as respectivas cinco profundidades, totalizando 26 pontos. Foram utilizados 25 pontos treinamento e um ponto validação. Cada modelo foi analisado 26 vezes de forma a permitir que todos os pontos participem do processo de

treinamento e validação. Ao término das 26 repetições de treinamento e validação, foi realizado o cálculo das médias de coeficiente de determinação (R^2) e raiz do erro médio quadrático (RMSE) para cada variável: areia fina, areia grossa, silte, argila, porosidade total, macroporosidade e microporosidade.

O R^2 varia entre 0 e 1, indicando quanto o modelo consegue explicar os valores observados. Já a RMSE é comumente usada para expressar a acurácia dos resultados numéricos. A grande vantagem da RMSE é que os valores apresentados estão nas mesmas unidades da variável analisada (KUHN; JOHNSON, 2013).

As análises e gráficos gerados nesse trabalho foram desenvolvidos com o software R[®] (R Core Team, 2017) e suas bibliotecas *caret*, *corrplot*, *Cubist*, *doParallel*, *lattice*, *dplyr*, *e1071*, *janitor*, *hydroGOF*, *kernlab*, *scales* (KUHN, 2008; SARKAR, 2008; WICKHAM, 2010; ANALYTICS; WESTON, 2014; WEI et al., 2017; WICKHAM, 2017; ZAMBRANO-BIGIARINI, 2017; WICKHAM et al., 2018; FIRKE et al., 2018).

3 RESULTADOS E DISCUSSÃO

3.1. Textura do solo e porosidade

Nas análises laboratoriais de textura (Tabela 3), porosidade total, macroporosidade, microporosidade (Tabela 4) não foram observadas diferenças marcantes entre as amostras de solos.

3.2. Constante dielétrica aparente do solo

A metodologia para obtenção dos valores de constante dielétrica (K) foi empregada aos radargramas adquiridos com a antena de 400 MHz, onde os eventos de ondas aéreas e onda direta no solo puderam ser facilmente distinguidos e a presença de hipérboles de reflexão nos radargramas foi mais evidente. Os valores de constante dielétrica no mês de fevereiro variaram entre 8,39 e 16,64 e no mês de setembro entre 5,21 e 18,17. Nos dois períodos, os valores menores de constante dielétrica foram observados na camada superficial (Tabela 3).

Nos primeiros 10 cm da imagem do radargramas foram encontradas poucas hipérboles de reflexão representativas. Este resultado está de acordo com o trabalho de Cezar et al. (2012) que, em um estudo utilizando o georadar para avaliar propriedades físicas do solo, não obteve resposta satisfatória nas investigações rasas, seja pela baixa frequência da antena ou textura muito argilosa. Grande quantidade de argila atenua a propagação das ondas eletromagnéticas devido a sua condutividade elevada, o que reduz a profundidade de investigação (DAVIS; ANNAN, 1989; CEZAR et al., 2010; CEZAR et al., 2012).

Tabela 3: Valores mínimos, máximos e médios de constante dielétrica aparente do solo (K) e resultados da análise textural por período de coleta (meses de Fevereiro e Setembro) em profundidade.

Fevereiro												
Profundi- dade (m)	Areia grossa (kg.kg ⁻¹)			Areia fina (kg.kg ⁻¹)			Silte (kg.kg ⁻¹)			Argila (kg.kg ⁻¹)		
	Mín.	Máx.	Méd.	Mín.	Máx.	Méd.	Mín.	Máx.	Méd.	Mín.	Máx.	Méd.
0,00-0,10	0,15	0,25	0,20	0,11	0,23	0,16	0,03	0,18	0,10	0,44	0,66	0,55
0,10-0,20	0,13	0,31	0,21	0,10	0,22	0,16	0,03	0,14	0,08	0,42	0,69	0,56
0,20-0,30	0,12	0,25	0,18	0,08	0,23	0,15	0,03	0,17	0,08	0,44	0,70	0,58
0,30-0,40	0,11	0,22	0,16	0,09	0,22	0,14	0,04	0,12	0,08	0,45	0,71	0,62
0,40-0,50	0,10	0,21	0,15	0,09	0,22	0,13	0,04	0,12	0,08	0,46	0,73	0,64
Setembro												
0,00-0,10	0,15	0,32	0,22	0,11	0,18	0,14	0,03	0,16	0,09	0,40	0,66	0,55
0,10-0,20	0,13	0,31	0,23	0,10	0,18	0,13	0,03	0,13	0,07	0,42	0,69	0,56
0,20-0,30	0,12	0,32	0,21	0,08	0,19	0,13	0,03	0,13	0,08	0,42	0,70	0,58
0,30-0,40	0,11	0,30	0,18	0,09	0,16	0,12	0,04	0,11	0,08	0,46	0,71	0,62
0,40-0,50	0,10	0,30	0,17	0,09	0,14	0,11	0,04	0,10	0,07	0,50	0,73	0,64

Tabela 4: Valores mínimos, máximos e médios de Constante dielétrica aparente do solo (K), Porosidade total ($\text{m}^3.\text{m}^{-3}$), Macroporosidade ($\text{m}^3.\text{m}^{-3}$), Microporosidade ($\text{m}^3.\text{m}^{-3}$), por período de coleta (meses de Fevereiro e Setembro) em profundidade

Profundi- dade (m)	Fevereiro											
	K			Porosidade total ($\text{m}^3.\text{m}^{-3}$)			Macroporosidade ($\text{m}^3.\text{m}^{-3}$)			Microporosidade ($\text{m}^3.\text{m}^{-3}$)		
	Mín.	Máx.	Méd.	Mín.	Máx.	Méd.	Mín.	Máx.	Méd.	Mín.	Máx.	Méd.
0,00-0,10	8,39	15,70	12,50	0,40	0,56	0,47	0,01	0,16	0,06	0,37	0,45	0,41
0,10-0,20	8,65	14,50	11,88	0,40	0,57	0,48	0,02	0,24	0,10	0,33	0,44	0,38
0,20-0,30	9,83	15,95	12,48	0,39	0,58	0,47	0,01	0,21	0,09	0,34	0,46	0,38
0,30-0,40	10,52	15,50	12,81	0,45	0,58	0,50	0,01	0,20	0,11	0,34	0,45	0,39
0,40-0,50	9,51	16,64	13,34	0,46	0,57	0,52	0,05	0,19	0,13	0,33	0,43	0,39
Setembro												
0,00-0,10	5,21	12,20	9,52	0,42	0,52	0,47	0,03	0,14	0,08	0,37	0,43	0,39
0,10-0,20	6,55	13,36	10,36	0,44	0,55	0,47	0,06	0,13	0,08	0,35	0,41	0,39
0,20-0,30	6,45	15,43	11,36	0,45	0,57	0,49	0,04	0,18	0,12	0,31	0,41	0,37
0,30-0,40	9,33	16,22	12,34	0,42	0,60	0,51	0,01	0,21	0,13	0,30	0,41	0,37
0,40-0,50	12,12	18,17	13,89	0,43	0,58	0,50	0,03	0,21	0,10	0,31	0,53	0,40