

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS RURAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DO SOLO**

**MODELOS DIGITAIS DE ELEVAÇÃO E PREDIÇÃO
DO CARBONO ORGÂNICO DO SOLO NO PLANALTO
DO ESTADO DO RIO GRANDE DO SUL**

DISSERTAÇÃO DE MESTRADO

Jean Michel Moura Bueno

**Santa Maria, RS, Brasil
2014**

MODELOS DIGITAIS DE ELEVAÇÃO E PREDIÇÃO DO CARBONO ORGÂNICO DO SOLO NO PLANALTO DO ESTADO DO RIO GRANDE DO SUL

Jean Michel Moura Bueno

Dissertação apresentada ao Curso de Mestrado do Programa de Pós-Graduação em Ciência do Solo, Área de Concentração em Processos físicos e morfogenéticos do solo, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de Mestre em Ciência do Solo.

Orientador: Prof. Dr. Ricardo Simão Diniz Dalmolin

Santa Maria, RS, Brasil
2014

Ficha catalográfica elaborada através do Programa de Geração Automática da Biblioteca Central da UFSM, com os dados fornecidos pelo(a) autor(a).

Moura-Bueno, Jean Michel
MODELOS DIGITAIS DE ELEVAÇÃO E PREDIÇÃO DO CARBONO
ORGÂNICO DO SOLO NO PLANALTO DO ESTADO DO RIO GRANDE DO
SUL / Jean Michel Moura-Bueno.-2014.
110 p.; 30cm

Orientador: Ricardo Simão Diniz Dalmolin
Dissertação (mestrado) - Universidade Federal de Santa
Maria, Centro de Ciências Rurais, Programa de Pós-
Graduação em Ciência do Solo, RS, 2014

1. Modelos digitais de elevação 2. Atributos do
terreno 3. Precisão 4. Carbono orgânico do solo 5.
Mapeamento digital de solos I. Dalmolin, Ricardo Simão
Diniz II. Título.

Universidade Federal de Santa Maria
Centro de Ciências da Saúde
Programa de Pós-Graduação em Ciência do Solo

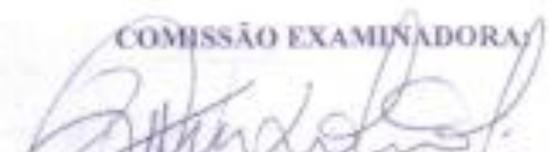
A comissão Examinadora, abaixo assinada,
aprova a Dissertação de Mestrado

**MODELOS DIGITAIS DE ELEVAÇÃO E PREDIÇÃO DO CARBONO
ORGÂNICO DO SOLO NO PLANALTO DO ESTADO DO RIO GRANDE DO
SUL**

elaborada por
Jean Michel Moura Bueno

como requisito parcial para obtenção do grau de
Mestre em Ciência do Solo

COMISSÃO EXAMINADORA



Ricardo Simão Diniz Dalmolin, Dr.
(Presidente/Orientador)



Alexandre ten Caten Dr. (UFSC)



Pablo Miguel Dr. (UFPeI)

Santa Maria, 08 de agosto de 2014.

**A todos aqueles que contribuíram para o meu crescimento como
pessoa, dedico.**

AGRADECIMENTOS

À Deus, que sempre esteve em minha vida, guiando meus passos e meu caminho, e me dando forças para que chegasse até o fim.

Aos meus pais Adair Sidnei Bueno e Heidi Selent, por me ensinarem que o estudo é sempre o melhor caminho. Obrigada por acreditarem em mim.

Ao meu irmão Charles Bueno por sempre estar ao mau lado e dividir as alegrias, tristezas e sonhos.

A minha noiva Fernanda dos Santos Pascotini, pelo companheirismo, pela paciência, pelo apoio e por toda ajuda.

Ao meu orientador e principalmente amigo, professor Ricardo Dalmolin, pela confiança depositada em mim durante os 7 anos de convivência no Setor de Pedologia do Departamento de Solos da UFSM. Tenha certeza que você contribui muito para que eu chegasse até aqui.

Aos professores da banca Prof. Pablo Miguel e Prof. Alexandre ten Caten pela disponibilidade, pelos momentos de discussão e pelo auxílio sempre que precisei.

Ao professor Fabrício de Araújo Pedron pela amizade e conselhos durante minha trajetória acadêmica.

Aos colegas e amigos Pablo Miguel e Alessandro Samuel-Rosa pela confiança, conselhos e momentos de descontração durante minha caminhada acadêmica.

Agradeço à ajuda, dos meus colegas e amigos do setor de Pedologia, André, Priscila, Sara Chagas, Ismael, Vicente, Ricardo Lopes, Cassiane Rodrigues, Gabriel Deobald, Ândrea Franco, e outros, pela ajuda, convivência e companheirismo.

Agradeço à UFSM e ao Programa de Pós-Graduação em Ciência do Solo pela oportunidade de realização do mestrado.

Ao amigo Carlos Alberto Flores pelos ensinamentos, troca de ideias e conselhos.

Ao proprietário de área de estudo, Darcy Urhy, pela disponibilidade da área para a realização desta pesquisa.

Agradeço aos amigos da turma 80 da Agronomia da UFSM por estarem ao meu lado nos momentos de descontração.

Agradeço a CAPES pela concessão da bolsa.

E agradeço a todas as pessoas que, de uma maneira ou de outra, contribuíram para a realização desse trabalho, muito obrigado.

“Por vezes sentimos que aquilo que fazemos não é senão uma gota de água no mar. Mas o mar seria menor se lhe faltasse uma gota”.

Madre Teresa de Calcutá

RESUMO

Dissertação de Mestrado
Programa de Pós-Graduação em Ciência do Solo
Universidade Federal de Santa Maria – Rio Grande do Sul

MODELOS DIGITAIS DE ELEVAÇÃO E PREDIÇÃO DO CARBONO ORGÂNICO DO SOLO NO PLANALTO DO ESTADO DO RIO GRANDE DO SUL

AUTOR: JEAN MICHEL MOURA BUENO
ORIENTADOR: RICARDO SIMÃO DINIZ DALMOLIN
Data e Local da Defesa: Santa Maria, 08 de agosto de 2014.

A demanda por informações quantitativas de solos em nível detalhado de bacias hidrográficas vêm aumentando em decorrência dos cenários globais. O mapeamento digital de solos (MDS) visa gerar dados compatíveis com essas necessidades por meio da geração de mapas funcionais de solos em multi-escalas. O objetivo desse trabalho foi avaliar a qualidade altimétrica e limitações de MDE com a finalidade de auxiliar na escolha do MDE mais adequados para derivar atributos do terreno (AT) para desenvolver funções de predição espacial para serem aplicadas ao mapeamento digital do carbono orgânico do solo em escala de propriedade rural no Planalto do Estado do Rio Grande do Sul (RS). O estudo foi realizado em uma área de 937 ha localizada no município de Giruá, RS, Brasil. Foram coletados 243 pontos amostrais na camada de 0-5 cm e realizado um levantamento planialtimétrico (LP) considerado neste estudo a verdade do terreno dos valores de altitude. Os MDE avaliados foram: MDE-LP (gerado a partir do LP), MDE-CARTA (gerado pela interpolação das curvas de nível da carta topográfica), MDE-ASTER, MDE-SRTM e MDE-TOPODATA. Os MDE foram avaliados quanto à precisão altimétrica por meio de teste estatísticos, valor da raiz quadrada do erro médio quadrático (RMSE) e aplicação da Norma Brasileira de Cartografia para definição da escala de cada MDE com base na precisão da altitude. Os AT derivados de cada MDE foram confrontados com os AT derivados do MDE-LP. Os resultados mostraram que o MDE-LP apresentou a melhor qualidade dos dados de altitude (RMSE = 1,93 m), seguido dos MDE-SRTM (RMSE = 5,95 m), MDE-CARTA (RMSE = 8,28 m), MDE-TOPODATA (RMSE = 9,78 m) e MDE-ASTER (RMSE = 15,57 m). O tamanho da área e relevo suave ondulado foram os principais fatores que influenciaram nos resultados. O MDE-LP se adequou na escala 1:10.000 classe D, enquanto os MDE-CARTA e MDE-SRTM se adequaram na escala 1:50.000 classe B, o MDE-TOPODATA a escala 1:50.000 classe D e o MDE-ASTER escala 1:100.000 classe B. Em relação aos AT, os MDE- SRTM e MDE-TOPODATA apresentam resultados mais próximos do MDE-LP e os menores valores de RMSE para cada AT avaliado. A função de predição construída a partir dos AT derivados do MDE-LP e índice de vegetação obtido de imagens Landsat-7 explicou apenas 16% da variância do COS na área. A resolução espacial pequena do MDE-LP e das imagens associado ao uso apenas dessas variáveis preditoras foram os principais fatores que influenciaram nos resultados. Com base nesses resultados, os MDE- SRTM e MDE-TOPODATA podem ser utilizados no MDS semidetalhado de classes de solos. No caso do MDS do COS sugere-se o uso desses MDE associado com pontos de controle de campo para verificação da precisão altimétrica e a inclusão de variáveis relacionadas a práticas de manejo do solo.

Palavras-Chave: modelos digitais de elevação, atributos do terreno, precisão, carbono orgânico do solo, mapeamento digital de solo.

ABSTRACT

Master Dissertation
Graduate Program in Soil Science
Federal University of Santa Maria – Rio Grande do Sul

DIGITAL ELEVATION MODELS AND PREDICTION OF SOIL ORGANIC CARBON IN PLATEAU STATE OF RIO GRANDE DO SUL

AUTHOR: JEAN MICHEL MOURA BUENO

ADVISOR: RICARDO SIMÃO DINIZ DALMOLIN

Date and Local of the Defense: Santa Maria, August 8th 2014.

The requirement for quantitative soil information has increased as a consequence of the global scenarios. The digital soil mapping (DSM) seeks to produce consistent data with the current needs through the generation of functional soil maps in multi-scales. The aim of this study was to evaluate the altimetry quality and limitations of digital elevation model (DEM) in order to assist in choosing the most suitable DEM to derive terrain attributes (TA) to develop spatial prediction functions to be applied to digital mapping of soil organic carbon (SOC) of farm scale in the Plateau of the state of Rio Grande do Sul (RS). The study was conducted on a 937 ha area located in the municipality of Giruá, RS, Brazil. They collected 243 sampling points in the 0-5 cm layer and an accomplished planialtimetric survey (PS) considered in this study the truth of the ground of altitude values. DEM evaluated were: DEM-PS (generated from the PS), DEM-LETTER (generated by interpolating the level of topographic map curves), DEM-ASTER, DEM-SRTM and DEM-TOPODATA. The DEM were evaluated for precision altimetry through statistical test value of the square root of the mean square error (RMSE) and application of the Brazilian Cartographic Standard for defining the scale of each DEM based on the accuracy of the altitude. TA derived from each DEM were faced with the AT derived from the DEM-PS. The results showed that the DEM-PS presented the best quality of elevation data (RMSE = 1.93 m), followed by DEM-SRTM (RMSE = 5.95 m), DEM- (RMSE = 8.28 m), DEM-TOPODATA (RMSE = 9.78 m) and DEM-ASTER (RMSE = 15.57 m). The size of the area and gently rolling relief were the main factors that influenced the results. The DEM-PS is suited in scale 1: 10,000 Class D, while DEM-LETTER and DEM-SRTM are suited in scale 1: 50,000 class B, the DEM-TOPODATA the scale 1: 50,000 class D and the DEM-ASTER scale 1: 100,000 Class B. With regard to TA, the DEM-SRTM and DEM-TOPODATA present results closer to the DEM-PS and smaller RMSE values for each TA assessed. The prediction function constructed from the DEM-PS derived from the TA and vegetation index Landsat-7 obtained images explained only 16% of the variance in SOC area. The small spatial resolution of DEM-PS and images associated with the use only of these predictors were the main factors that influenced the results. Based on these results, the DEM-SRTM and DEM-TOPODATA can be used in DSM semi-detailed soil classes. In the case of the SOC DSM suggest the use of these DEM associated with field control points to verify the precision altimetry and the inclusion of variables related to soil management practices.

Keywords: digital elevation models, terrain attributes, precision, soil organic carbon, digital mapping soil.

SUMÁRIO

1 INTRODUÇÃO	09
2 REVISÃO DE LITERATURA	12
2.1 Mapeamento digital de solos	12
2.2 Funções de predição espacial de propriedades do solo	16
2.3 Predição espacial do carbono orgânico do solo (COS)	21
2.4 Modelos digitais de elevação	24
3 HIPÓTESE	31
4 OBJETIVOS	32
4.1 Geral	32
4.2 Específicos	32
5 CAPÍTULO 1 QUALIDADE DE MODELOS DIGITAIS DE ELEVAÇÃO PARA SEREM UTILIZADOS NO MAPEAMENTO DIGITAL DE SOLOS	33
5.1 Introdução	33
5.2 Materiais e Métodos	35
5.3 Resultados e Discussão	42
5.4 Conclusões	51
6 CAPÍTULO 2 FUNÇÕES DE PREDIÇÃO ESPACIAL DO CARBONO ORGÂNICO DO SOLO	52
6.1 Introdução	52
6.2 Materiais e Métodos	54
6.3 Resultados e Discussão	60
6.4 Conclusão	74
7 CONSIDERAÇÕES FINAIS	75
8 REFERÊNCIAS	76
9 ANEXOS	90

1 INTRODUÇÃO

O homem sempre manteve relações próximas com o recurso natural solo. Varias civilizações deixaram de ser nômades em razão de fatores relacionados ao potencial dos solos em produzir alimentos para o sustento do homem. Há evidências que os primeiros relatos a respeito do solo como meio para produção de alimentos foi quando o homem deixou de ser nômade e passou a compor pequenas comunidades a 9000 anos na Mesopotâmia.

Nesse contexto, o solo é um componente fundamental para os ecossistemas terrestres, pois está envolvido desde a produção de alimentos até a regulação do clima. Em virtude do acelerado crescimento populacional, em 2050 a Terra será habitada por cerca de 9 bilhões de pessoas segundo estimativas da Organização das Nações Unidas (ONU). Diante desse cenário, a produção de alimentos terá que aumentar entre 60 e 90%, podendo causar a utilização não racional dos recursos naturais como, por exemplo, o solo, se não forem tomadas medidas adequadas de uso e conservação desse recurso.

Para minimizar esses impactos, nas últimas décadas o homem tem buscado no meio científico, soluções mais consistentes para alguns problemas que existem no planeta em que habitamos. Entre esses enigmas, está os cenários da mudança climática, necessidade de aumentar a produção de alimentos e energia, preocupação com as formas de uso da terra, conteúdo e sequestro de carbono orgânico (CO) no solo. Nesse contexto, o solo esta em entre os cinco principais reservatórios de carbono (C) no ecossistema terrestre, onde os solos agrícolas podem atuar como dreno ou fonte de C para a atmosfera dependendo do uso e manejo a que forem submetidos. O carbono orgânico do solo (COS) é considerado um indicador de qualidade do solo e responsável por uma série de funções no ambiente.

Portanto, conhecer o conteúdo e a distribuição de COS na paisagem é extremamente importante tendo em vista os cenários globais, porém é uma tarefa difícil e que ainda necessita de avanços na pesquisa. Para alcançar esse objetivo, o conhecimento e a compreensão dos tipos de solos e suas propriedades e a inter-relação com variáveis ambientais responsáveis pela dinâmica do COS em escala de propriedade rural, bacia hidrográfica ou município é de fundamental importância para um planejamento racional do recurso natural solo.

Realizar mapeamento do conteúdo de COS nas escalas citadas acima é geralmente dispendioso e demanda tempo devido ao elevado números de pontos que precisam ser

amostrados na paisagem. Sendo isso consequência do grande número de variáveis ambientais e ações antrópicas responsáveis pela variabilidade do COS na pedosfera, onde cada variável apresenta uma função e porcentagem de contribuição nessa variação. Por isso, é necessário o uso de novas técnicas de amostragem, processamento de dados e a construção de modelos que representem as relações entre o solo e o ambiente em que o mesmo ocorre.

Isso tem sido facilitado nos últimos anos com os avanços da informática, softwares estatísticos e os Sistemas de Informações Geográficas (SIG), que têm possibilitado o processamento de grandes volumes de informações e o desenvolvimento de modelos matemáticos capazes de estimar propriedades do solo a partir de covariáveis ambientais. Segundo ten Caten e Dalmolin (2014) o geoprocessamento é uma ferramenta muito promissora para uso no mapeamento de solos na área de Pedologia.

Essa nova linha de estudo constitui o escopo do mapeamento digital de solos (MDS), o qual teve avanço nas pesquisas a partir do trabalho publicado em 2003 por Alex McBratney e seus colaboradores, intitulado “On digital soil mapping”. O MDS foi definido por Lagacherie e McBratney (2006) como “a criação e população de sistemas de informação espacial de solos através do uso de métodos observacionais de campo e laboratório, acoplados a sistemas de inferência espacial e não-espacial de solos”.

Para a construção desses sistemas de informações espacial de solo, utiliza-se a pedometria, disciplina da ciência do solo que estuda a “aplicação de métodos matemáticos e estatísticos para o estudo da distribuição e gênese dos solos” (Heuvelink, 2003). As variáveis mais utilizadas para gerar esses modelos são as derivadas de modelos digitais de elevação (MDE) em virtude da alta correlação com algumas propriedades do solo e a facilidade de aquisição desses dados. Esses modelos são denominados de Funções de Predição Espacial de Solos (FPES), as quais constituem em métodos de ajuste empírico entre as relações quantitativas existentes entre o solo e o ambiente em que o mesmo ocorre.

Neste contexto, pode-se dizer que a pesquisa em MDS é recente, principalmente no Brasil. Pesquisas regionalizadas precisam ser realizadas para aumentar a base de conhecimento e desenvolver estratégias e metodologias para o MDS, em especial a predição do COS na paisagem, e possíveis extrapolações para outras regiões com condições ambientais semelhantes.

O Planalto do estado do Rio Grande do Sul, Brasil, é considerado uma região com alto potencial agrícola e com uma diversidade de uso e manejo do solo. Portanto deve-se ter consciência que para utilizar o solo necessita-se de um planejamento e exploração controlada, a fim de preservar a qualidade do mesmo para as gerações futuras.

Diante do exposto, o presente trabalho buscou aumentar o conhecimento a respeito do desempenho das funções de predição espacial do COS construídas através de covariáveis ambientais. Este trabalho concentrou-se em avaliar a qualidade das informações contidas nos MDE e construir uma função de predição para o COS em escala de propriedade rural.

2 REVISÃO BIBLIOGRÁFICA

2.1 Mapeamento digital de solos

Estudos pioneiros a respeito do solo como corpos naturais organizados e a sua gênese foram difundidos pelo russo Dokuchaev em 1883, posteriormente por seus discípulos na metade do século 19. Com base nesses estudos, Hans Jenny em 1941 formalizou as bases conceituais sobre a gênese e a distribuição dos solos na paisagem. Jenny propôs um modelo clássico abordando os fatores relacionados com a formação do solo. Para Jenny, o solo (S) é função do clima (cl), dos organismos (o) e do relevo (r), agindo sobre o material parental (p) durante um período de tempo (t) e fatores desconhecidos (...), conforme representado pela equação 1:

$$S=f(cl,o,r,p,t,...)$$

O modelo determinístico de Jenny representou uma maneira complexa de prever os solos, em função das interações entre os fatores de formação (VILLELA, 2013). Dessa forma, se o solo é função dos fatores de formação e se a distribuição espacial desses fatores for conhecida, conclui-se que o solo e suas propriedades podem ser inferidos a partir da interpretação de informações de campo e construção de um modelo conceitual (relação solo-paisagem) (JENNY, 1941; HUDSON, 1992). Esse modelo busca compreender as relações entre a paisagem e demais variáveis ambientais que caracterizam o solo, sendo esse aplicado para inferir a distribuição espacial dos solos e suas propriedades em áreas análogas (HUDSON, 1992).

Em um ambiente no qual os fatores de formação e os processos de pedogênese sejam iguais, tem-se grande probabilidade dos solos apresentarem características muito semelhantes. Adotando essa ideia como padrão, é possível fazer inferências sobre os tipos de solos encontrados em variados locais na paisagem, podendo ser mais ou menos explícito de acordo com a extensão da área (MENDONÇA-SANTOS e SANTOS, 2003).

A partir da introdução do modelo de Jenny na ciência do solo, mais especificamente na área da Pedologia, vários levantamentos e mapeamentos de solos foram elaborados em

escala mundial, conforme as necessidades e demandas da sociedade em geral de conhecer os tipos de solo e sua distribuição na superfície terrestre. Inicialmente o levantamento pedológico era realizado com descrições dos solos, classificação, determinação de limites entre as classes e, com a realização de inferências sobre o comportamento dos solos perante o uso e manejo (MENDONÇA-SANTOS e SANTOS, 2003).

No mapeamento do solo utilizando métodos convencionais, o pedólogo adquire conhecimentos tácitos com o decorrer do tempo a respeito da ocorrência das classes de solos na paisagem, sendo de fundamental importância a experiência do pedólogo adquirida no decorrer dos muitos trabalhos de campo (HUNSON, 1992). No entanto, existem poucos pedólogos treinados e capazes de descrever um modelo mental generalizado que atenda as diversas situações que ocorrem no ambiente.

Nos últimos anos a demanda por informações de qualidade sobre o solo tem aumentado devido a estreita relação desse recurso natural com os ciclos hidrológicos, dinâmica do clima em escala global, ciclagem de nutrientes e produção de alimentos e energia (HARTEMINK, 2008; SANCHEZ et al., 2009; PACHECO et al., 2011; GRUNWALD et al., 2012). A gestão e o planejamento para fins específicos, com alto nível de detalhe, requerem mapas funcionais de solos (ZHU et al., 2013), por exemplo, mapa de aptidão agrícolas das terras e área de risco de erosão (VAN ZIJL et al., 2014).

Nesse contexto, uma das limitações é a indisponibilidade de mapas de solo em escalas adequadas para cada propósito. Segundo Sanchez et al. (2009), apenas 109 países possuem mapas coropléticos de solos nas escalas 1:1.000.000 ou maiores, representando somente 31% da superfície da Terra. O restante dispõe apenas do mapa de solos na escala 1:5.000.000 elaborado pela FAO - *Food and Agriculture Organization of the United Nations*, em escala incompatível para planejamento de uso do solo necessário para atender as demandas atuais.

Ainda, os levantamentos e mapeamentos de solos realizados com base no “modelo mental” desenvolvido pelo pedólogo experiente, representados em mapas coropléticos, estabelecem limites abruptos arbitrários entre as classes de solo (TEN CATEN et al., 2011b) e as informações não ficam documentadas, dificultando a reprodução e atualização desses mapas no futuro. Esses são alguns dos motivos pelo qual o mapeamento convencional vem recebendo críticas a respeito de suas limitações para atender as necessidades da sociedade sobre informações quantitativas e com precisão sobre a distribuição espacial do solo e suas propriedades (BASHER, 1997; HARTEMINK e MCBRATNEY, 2008; ARRUDA et al., 2013; CARVALHO JÚNIOR et al., 2014).

Diante disso, surge a necessidade de compreender melhor os processos que ocorrem na pedosfera em conjunto e quantificar com precisão a distribuição espacial do solo e suas propriedades físicas, químicas e biológicas, as quais exercem funções no ambiente. Porém, essa é uma tarefa difícil de ser realizada, uma vez que, o solo é um sistema heterogêneo e contínuo na paisagem, onde seus processos de formação e dinâmica de funcionamento são de difícil compreensão. Além disso, as coletas de solos para formação de bancos de dados de solos apresentam um custo elevado e demanda de tempo (OLIVEIRA et al., 2012).

O desenvolvimento de novas técnicas e também o uso mais frequente pelos pesquisadores de tecnologias digitais, têm contribuído para reduzir esses problemas, como é o caso da técnica de MDS. O MDS foi definido por Lagacherie e McBratney (2006) como “a criação e população de sistemas de informação espacial de solos através do uso de métodos observacionais de campo e laboratório, acoplados a sistemas de inferência espacial e não-espacial de solos”. No MDS as informações extraídas de amostras de solo e associação com variáveis ambientais são empregadas em modelos de predição que possibilitam gerar informação espacial a respeito do solo com as incertezas e erros associados (MINASNY et al., 2008) e que permitem análises e interpretações em qualquer época. No MDS são utilizadas tecnologias tais como, sensoriamento remoto, informática, gerenciamento de dados espaciais em SIG, visualização e integração em conjunto do solo e as variáveis ambientais correlacionadas, métodos quantitativos para prever os processos que ocorrem no solo e suas propriedades (MCBRATNEY et al., 2003; GRUNWALD, 2009; TEN CATEN et al., 2012).

A utilidade dos levantamentos de solos não se restringe somente a gerar dados de inventários e distribuição geográfica de classes de solos, mas, também, prover a distribuição espacial quantitativa de propriedades dos solos. Nesse contexto, o MDS pode atender a essas necessidades, através da criação em 2004 do Grupo de Trabalho de Mapeamento Digital de Solo da União Internacional de Ciências do Solo (IUSS) e do consórcio GlobalSoilMap.net (GSM). O objetivo do consórcio é gerar um mapa digital em escala global das propriedades funcionais do solo como, por exemplo, a distribuição de partículas (areia, silte e argila), teor de carbono orgânico, pH, capacidade de troca de cátions, condutividade elétrica e densidade do solo. Este novo mapa global do solo visa auxiliar a tomada de decisões em uma série de questões globais, como a produção de alimentos e a erradicação da fome, mudanças climáticas e degradação ambiental, sendo esse tipo de informação necessária para atender a demanda dos usuários (SANCHEZ et al., 2009).

O reconhecimento do MDS pela comunidade científica internacional como uma ferramenta promissora para fornecer informações de qualidade e precisão a respeito do solo se

deu a partir da criação desse grupo de trabalho em MDS, sendo isso um indicativo que as pesquisas nessa área são recentes. No Brasil as pesquisas em MDS são mais recentes ainda, com o primeiro trabalho publico por Giasson et al. (2006). Segundo ten Caten et al. (2012) somente após o 2º Workshop Global em Mapeamento Digital de Solos realizado no Rio de Janeiro em 2006, os pesquisadores brasileiros da área da ciência do solo deram importância para os métodos empregados no MDS.

As primeiras pesquisas em MDS realizados no Brasil deram destaque para o mapeamento de classes de solos, como os trabalhos de Giasson et al. (2006), Crivelenti et al. (2009) e ten Caten (2011b). Com relação a predição de propriedades do solo, existem vários trabalhos, porém, utilizam como variáveis preditoras apenas o solo (*fator s*) e as coordenadas do local (*fator n*) e como técnica de predição a geoestatística pelo método da krigagem. São poucos os trabalhos que utilizam a abordagem do MDS apresentada no trabalho de McBratney et al. (2003), para o mapeamento de propriedades do solo. Nessa linha de estudo, pode-se citar os trabalhos de Bernoux et al. (2006) e Mendonca-Santos et al. (2010), ambos realizam o mapeamento dos estoques de carbono no solo e de Samuel-Rosa et al. (2013) que trabalhou com predição da granulometria do solo.

Além disso, os trabalhos de predição de classes de solos estão muito ligados aos sistemas taxonômicos de classificação e poucos utilizam dados atualizados de campo. Esse último fator é um aspecto que se deve tomar cuidado, pois segundo Minasny e McBratney (2007), a qualidade das predições é mais dependente da qualidade dos dados de solos utilizados do que da complexidade dos modelos matemáticos utilizados para construir as FPES.

A comunidade científica brasileira que desenvolve suas pesquisas na área de solo, em específico na área da Pedologia, tem um longo caminho a ser percorrido com relação ao MDS. Ênfase maior deve ser dada ao MDS de propriedades do solo, com a confecção de mapas funcionais de solos, utilizando a abordagem quantitativa preconizada pelo MDS nas diferentes regiões do país e em multi-escalas.

Para isso, é necessário o uso de técnicas mais robustas de coletas de dados, métodos estatísticos, geoprocessamento e processamento de dados. No entanto, para que isso ocorra de forma mais rápida e consistente é necessário mais integração entre grupos de pesquisas das diferentes áreas do conhecimento com o objetivo de aumentar a interdisciplinaridade e assim se consiga alcançar os objetivos propostos pelo grupo de trabalho em MDS.

2.2 Funções de predição espacial de propriedades do solo

O termo pedometria tem significado em um neologismo derivado do grego *pedos* (solo) e *metron* (medida). A pedometria busca solucionar os problemas relacionados com as incertezas inerentes ao método convencional de mapeamento de solos, com investigações sobre novas abordagens de modelagem quantitativa, tendo como objetivo aprofundar o conhecimento da variabilidade espacial dos solos, da qualidade e precisão da informação (WEBSTER, 1984; HEUVELINK, 2003. MCBRATNEY et al., 2003).

Para atender esse objetivo, a pedometria lança mão do uso algumas ferramentas como a geoinformação, as informações do solo e a estatística (modelos lineares generalizados, redes neurais, árvores de decisão, regressão, componentes principais) (Figura 1). Para Alex McBratney, a pedometria está relacionada à parte de mensuração de informações de solos, sendo restrita aos métodos quantitativos matemáticos e estatísticos. Segundo ten Caten et al. (2012), o avanço dos estudos em pedometria é também resultado de avanços tecnológicos, técnicas de sensoriamento remoto, posicionamento global de satélites e ao acesso à informação através da rede mundial de computadores.



Figura 1 – Diagrama de interdisciplinaridade da pedometria. Adaptado de Hengl (2003).

A linha de pesquisa em pedometria é reconhecida mundialmente através das publicações, conferências e workshops organizados pelo Grupo Internacional de Trabalho sobre Pedometria, da União Internacional de Ciência do Solo (IUSS). No Brasil foi recentemente criada, pela Sociedade Brasileira de Ciência do Solo (SBCS), a Comissão Especializada em Pedometria dentro da divisão Solo no Espaço e no Tempo. Isso significa um grande avanço para a ciência do solo no país, onde a pedometria passa a fazer parte da estrutura de ensino e pesquisa.

O MDS utiliza os fundamentos básicos da pedometria através do uso das funções de predição espacial de solos (FPES). Tais funções constituem um método de ajuste empírico das relações quantitativas existentes entre o solo e o ambiente em que o mesmo ocorre (MCBRATNEY et al., 2003). A equação que representa as FPES constitui o modelo $S=f(\textit{scorpan})$, nas quais o S representa a classe ou propriedade do solo a ser predita em função de s = informação prévia do solo, c = clima, o = organismos, r = relevo, p = material de origem, a = período decorrido, e n = posição espacial (MCBRATNEY et al., 2003). Este modelo apresenta fatores semelhantes ao modelo de Jenny (1941). Porém o modelo *scorpan* acrescenta em sua composição os fatores s e n , que representam a relação solo e ambiente em determinado espaço geográfico, possibilitando inferir a variabilidade espacial.

Para a construção de uma FPES necessitamos de um conjunto de dados a partir do qual serão estabelecidas relações entre as propriedades dos solos e as variáveis ambientais predictoras. As informações de solos são provenientes de amostragens que podem ser realizadas em toda a área de estudo (MCBRATNEY et al., 2003) ou em uma pequena área representativa, a qual constitui uma área de referência (LAGACHERIE et al., 1995; TEN CATEN et al., 2011b). As duas formas de amostragem mais utilizadas em trabalhos de pesquisa na ciência do solo são: 1) amostragem aleatória, na qual toda população de solos têm a mesma probabilidade de ser escolhida para a amostragem (BRUS e GRUIJTER, 1997); 2) amostragem sistemática em malha regular, onde o plano de amostragem é dividido em intervalos regulares de mesmo tamanho (LARK e BISHOP, 2007), sendo esta bastante comum em estudo relacionado à agricultura de precisão (CHERUBIN et al., 2014).

Os principais fatores limitantes e que influenciam fortemente na escolha de uma metodologia de amostragem, é a restrição financeira, mão-de-obra e difícil acesso aos locais de amostragem. Portanto, mais recentemente, a metodologia de amostragem recomendada para contornar essas dificuldades, é a utilização da amostragem em uma área de referência. Segundo Lagacherie et al. (1995) duas paisagens na qual ocorrem padrões de geologia, topografia e geomorfologia tem grandes probabilidades de apresentar o mesmo padrão de

ocorrência de solo. Assim, é de fundamental importância, verificar a representatividade da área de referência escolhida em umas das paisagens, para que a FPES consiga capturar as relações entre solo e outras variáveis ambientais (LAGACHERIE et al., 2001), sendo possível assim a extrapolação dessas relações para áreas vizinhas com as mesmas características.

A partir do conjunto de informações da área de referência é possível estabelecer relações entre o solo e as variáveis preditoras para construção das FPES para serem aplicadas na predição de propriedades do solo de uma área adjacente, sem que toda a população desses solos tenha sido amostrada, como mostra o trabalho de ten Caten et al., 2011b.

Da mesma forma, Cambule et al. (2013) realizou o mapeamento do COS em uma área de em Moçambique utilizando métodos geoestatísticos a partir de uma área de referência. O trabalho enfatizou o uso limitado de informações de solos e variáveis preditoras de uma área acessível para calibrar um modelo para ser utilizado na estimativa do COS para locais de difícil acesso. Os resultados observados mostram o potencial dessa metodologia, porém as relações estabelecidas entre as variáveis só podem ser extrapoladas para áreas com características semelhantes.

Dentre as variáveis preditoras, as mais utilizados são os atributos do terreno (MCBRATNEY et al., 2003; GRUNWALD, 2009), consequência da estreita relação entre o solo e os atributos do terreno. As variáveis do terreno caracterizam os fluxos de água na paisagem e consegue identificar a variabilidade espacial das propriedades do solo, uma vez que, a água é o fator responsável pela erosão em uma bacia hidrográfica e, portanto, têm influência nas características do terreno e consequentemente nas propriedades do solo ao longo de uma topossequência de solos (MOORE et al., 1993). Além disso, as variáveis do terreno são de mais fácil obtenção (TEN CATEN et al., 2011a), sendo adquiridas de sensores remotos, modelos digitais de elevação (MDE), digitalização de curvas de nível de cartas topográficas ou de levantamentos topográficos. Em geral, essas fontes de dados estão disponíveis em resolução espacial média a alta.

As FPES são então construídas utilizando modelos estatísticos (modelos lineares, modelos lineares generalizados, árvores de decisão, entre outros), geoestatísticos (krigagem e co-krigagem) ou uma combinação de ambos (métodos híbridos) (MCBRATNEY et al., 2003; BISHOP e MINASNY, 2006; GRUNWALD, 2009). Neste contexto, Moore et al. (1993) usando modelos de regressão linear para estimar o teor de silte a argila do solo conseguiram explicar 66% da variâncias dessa propriedade na paisagem. McKenzie e Ryan (1999) conseguiram explicar 54% da variância do teor de COS utilizando modelos lineares

generalizados. Sumfleth e Duttmann (2008) não conseguiram explicar mais de 41% da variância do teor de COS em solos da China.

Da mesma forma, Kheir et al. (2010) realizam predição a do COS em solos da Dinamarca utilizando método de árvore de decisão (AD), o modelo gerado conseguiu explicar 71% da variabilidade do COS na paisagem. Samuel-Rosa et al. (2013) trabalhado em uma em bacia hidrográfica na região sul do Brasil com relevo e geologia complexa, construí FPES para a composição granulometria solo, a quais explicaram 55, 67 e 75% da variação para o conteúdo de argila, silte e areia, respectivamente. Taghizadeh-Mehrjardi et al. (2014) realizaram a modelagem da salinidade em um área de 72.000 ha na região de Ardakan no centro do Irã. Os autores utilizaram atributos do terreno, imagens Landsat-7 ETM e condutividade elétrica aparente para gerar as FPES por meio de árvore de regressão. As funções conseguiram explicar 78% da variabilidade nas camadas superficiais e apenas 11% na camada de 60 -100 cm.

Estudos tem mostrado a necessidade de desenvolver tecnologias para caracterizar de forma rápida e precisa todo o perfil de solo (BEN-DOR et al., 2008; VISCARRA ROSSEL et al., 2011; STOCKMANN et al, 2014). Para anteder essa demanda, recente foi criado o grupo de trabalho em Mapeamento Digital de Morfometria do solo da IUSS. Em trabalho publicado por Hartemink e Minasny (2014), os autores apresentam uma revisão de técnicas com potencial para serem utilizadas para medir e quantificar algumas características do perfil de solo como, horizontes, textura, cor, estrutura, umidade, mosqueados, consistência, carbonatos, fragmentos de rocha, poros e raízes. Os autores denominaram essa linha de estudo de “morfometria digital de solos”, sendo considerado um grande avanço para área de pedologia.

O desempenho de FPES é muito viável, sendo influenciados pelas características geomorfológicas da área em estudo, da propriedade do solo a ser predita e do conjunto de variáveis ambientais preditoras, informações de solo pré-existentes e da resolução espacial dessas informações (GRUNWALD, 2006; CAVAZZI et al., 2013).

A respeito da resolução de MDE utilizados na predição de solos, estudo realizado por Cavazzi et al. (2013), em três áreas diferentes em termos de geomorfologia na Irlanda mostra que nem sempre o MDE de maior resolução espacial é o mais adequado para realizar e predição de solos. Thompson et al. (2001) já havia relatado isso, salientado que MDE de alta resolução podem não ser necessário para a construção de modelos de relação-solo-paisagem.

Uma linha de pesquisa na área de MDS é o uso de múltiplas escalas para análise do terreno e sua influência nas predições de solos. Estudo elaborado por Behrens et al. (2010) mostra uma melhoria na precisão das estimativas de solos utilizando essa estratégia. Da

mesma forma, Behrens et al. (2014) mostraram o uso de mineração de dados para determinar medidas estatísticas de variáveis do terreno para mapear estruturas geomorfológicas, relacionadas com a formação do solo, além de identificar locais onde determinada medida estatística apresenta mais influencia. Os autores encontraram vales de R^2 de 0,63 para predição de argila para uma área de 300 km² em Piracicaba, Brasil. No mesmo estudo, em uma área de 1.150 km² na Alemanha, o modelo explicou 68% da variação do conteúdo de silte no solo. Essa nova técnica pode auxiliar na predição de processos mais complexos de formação do solo na paisagem. Segundo os autores, em uma região onde se tem apenas um MDE e com resolução baixa, o uso da técnica de multi-escala pode solucionar o problema da falta de variáveis ambientais e assim, fornecer informações sobre geomorfológica da área.

Existem alguns trabalhos avaliando as incertezas e erros no MDS, sendo as causas relacionadas às variáveis preditoras usadas na construção das FPES (BISHOP et al., 2006; BEHRENS et al., 2008; GRIMM e BEHRENS, 2010). Estudo de Grimm e Behrens (2010) buscou estimar e visualizar o efeito de erros de localização de amostras de solo coletadas em tempos remotos. A partir de métodos estatísticos algumas amostras foram removidas e outras movidas na paisagem, sendo observado influência na predição do COS, com redução do valor de RMSE dos modelos obtidos a partir dos dados originais, movidos e removidos, respectivamente. Essa estratégia deve ser adotada no MDS, pois grande parte dos trabalhos utilizada para a calibração dos modelos dados de solo pré-existent, consequência do custo de coleta e análises dos mesmos (CARRÉ et al., 2007). Porém, alguns desses dados foram adquiridos em momentos anteriores ao uso do GPS, assim provavelmente apresentam erros de posicionamento na paisagem.

Diante dos resultados apresentados pelos trabalhos citados anteriormente, onde mostram diferentes desempenhos das FPES em prever algumas propriedades do solo, sendo as causas atribuídas as características intrínsecas dos locais de estudo. Grunwald (2009) mostra em seu estudo que é impossível construir uma FPES que possa ser usada em todas as superfícies geomorfológicas, sendo indispensável à realização de estudos e construção de FPES para as distintas condições ambientais (MCKENZIE et al., 2000). Nesse sentido, o Brasil carece de pesquisas na área de MDS com ênfase na construção de FPES para as diferentes regiões do país e em diferentes escalas (bacia hidrográfica, propriedade rural, entre outras).

2.3 Predição espacial do carbono orgânico do solo (COS)

Para entender a dinâmica do carbono dos ecossistemas terrestres, é preciso entender os processos de formação dos solos e a variabilidade espacial do COS ao longo da paisagem (BATJES e SOMBROEK, 1997). Vários mapeamentos tem sido realizado com essa finalidade em escala global (ESWARAN et al, 1995), continental (BAJTES, 2000), nacional (YU et al., 2007) e regional (MENDONÇA-SANTOS et al., 2007; GRIMM et al., 2008; KHEIR et al., 2010; PEI et al., 2010). Esses trabalhos utilizam diferentes técnicas de interpolação para realizar o mapeamento do COS, sendo requerido para isso uma malha de amostragem bastante densa em cada local de estudo, acarretando em alto custo e demanda de tempo (OLIVEIRA et al., 2012), consequência da alta variabilidade que essa propriedade do solo apresenta na paisagem. Ainda, em alguns casos o volume de dados é limitado devido a esses fatores refletindo em mapeamentos inconsistentes.

Portanto, realizar inventários do conteúdo de COS em escala de paisagem como, em bacias hidrográficas, demanda de um alto custo e tempo devido ao elevado números de pontos que precisam ser amostrados na paisagem. Logo, pesquisas regionalizadas precisam ser realizadas para aumentar a base de conhecimento sobre a variabilidade do COS em escala de paisagem, desenvolver estratégias e metodologias para o mapeamento do COS e possíveis extrapolações para outras regiões com condições ambientais semelhantes (PHACHOMPHON et al., 2010). Tendo em vista isso, o grupo de trabalho em MDS da IUSS através do consórcio GlobalSoilMap.net (GlobalSoilMap.net, 2011) estabeleceu algumas propriedades do solo a serem mapeadas, entre elas o COS, e especificações que regem alguns aspectos que devem ser estudados e seguidos nos trabalhos de MDS.

No MDS a associação de informações extraídas de amostras de solo, sensoriamento remoto e variáveis ambientais derivadas de MDE, é empregada em funções de predição que possibilitam gerar informação espacial quantitativa a respeito do solo, além de incluir as incertezas e erros a respeito dessas informações (MCBRATNEY et al., 2000). Como já visto anteriormente, para prever classes de solo e propriedades do solo McBratney et al. (2003) desenvolveram o modelo *scorpan*. Esse modelo tem dado suporte para que pesquisas relacionadas a predição de propriedades do solo sejam desenvolvidas, onde o objetivo dos estudos é gerar informações para a compreensão dos processos envolvidos na redução das emissões de CO₂ para a atmosfera (GRUNWALD, 2009).

Estudos relacionados ao teor de COS encontram-se em destaque por ser uma propriedade do solo fundamental, uma vez que, o solo está entre os cinco principais reservatórios de carbono no ecossistema terrestre (BATJES e SOMBROEK, 1997), podendo o solo agrícola atuar como dreno ou fonte de C para a atmosfera dependendo do uso e manejo a que forem submetidos (MBAH e IDIKE, 2011).

O COS pode apresentar correlação significativa com as variáveis do terreno. Diante disso, vários trabalhos têm utilizado atributos do terreno derivadas a partir de MDE e associadas a informações extraídas de amostras de solos e incorporados à métodos geoestatísticos e funções de predição espacial para o mapeamento e modelagem digital do COS (KHEIR et al., 2010; MENDONÇA-SANTOS et al., 2010; PEI et al., 2010; RAZAKAMANARIVO et al., 2011; KEMPEN et al., 2011; CHENG-ZHI et al., 2011; BALLABIO et al., 2012; CROFT et al., 2012; KERRY et al., 2012).

Nessa linha de pesquisa, Kheir et al. (2010) mostrou que mapas digitais de COS podem ser desenvolvidos a partir de avaliações de amostras de solo e atributos do terreno como, por exemplo, índice de vegetação por diferença normalizada (NDVI), índice de umidade por diferença normalizada (NDWI), índice de cor do solo (SCI), índice de umidade topográfica (IUT), entre outros, sendo essas variáveis utilizadas como preditoras na construção de modelos preditivos de COS a partir do método de AD.

Mendonça-Santos et al. (2007) testando técnicas de MDS para o mapeamento do estoque de COS, no estado do Rio de Janeiro, encontraram desempenho satisfatório para modelos de análise multilinear construídos a partir das variáveis preditoras mapas de solos, mapas de geologia, bandas de imagem Landsat ETM, NDVI e atributos do terreno derivados do MDE. Pei et al. (2010) utilizando o método de krigagem ordinária para realizar o mapeamento da MOS, encontraram correlações significativas entre a MOS e índice de umidade topográfica (IUT), em uma área de 5.176 ha, com pouca diferença de altitude e declividade média de 5%.

Grimm et al. (2008) avaliado o desempenho de modelos de Árvore de Decisão (AD) do tipo *Random Forest* para a predição do COS na Ilha do Barro Colorado, encontraram resultados satisfatórios, utilizando como variáveis preditoras os atributos do terreno, formação geológica, mapas de solos e histórico da vegetação. Kempen et al. (2011) em estudando a distribuição tridimensional do conteúdo de matéria orgânica do solo (MOS), na província de Drenthe no nordeste da Holanda em uma área de 125 km², constataram que a modelagem geoestatística associada a variáveis preditoras é uma ferramenta promissora para o MDS.

Razakamanarivo et al. (2011) realizou o mapeamento do estoque de COS na camada de 0 - 30 cm, em uma área de 1.590 ha em Madagascar, onde gerou FPES a partir de árvores de regressão, regressão linear simples e regressão linear múltipla, utilizando como variáveis preditoras parâmetros biofísicas da vegetação, tipo de solo, altitude e declividade. O melhor desempenho foi observado no modelo de árvores de regressão, explicando 61% da variação do estoque de COS. Ballabio et al. (2012) também utilizaram parâmetros da vegetação extraídos do sensoriamento remoto e atributos do terreno para realizar o mapeamento do COS por meio de regressão de krigagem. Os autores observaram um aumento no desempenho do modelo em explicar a variação do COS de 32% para 66% quando as variáveis de vegetação foram inseridas no modelo.

A respeito do uso de técnicas de sensoriamento remoto (TSR) para monitoramento espacial do carbono orgânico em sistemas agrícolas, Croft et al. (2012) em revisão sobre o assunto, os autores apontam que as TSR são ferramentas promissoras para cumprir muitas exigências de dados de pesquisa em ciências do solo, onde os dados podem ser criados, armazenado e manipulado digitalmente. Essa técnica vem sendo utilizada para estimar algumas propriedades do solo, dentre elas se destaca o COS e textura (DALMOLIN et al., 2005; NOCITA et al., 2014; RAMIREZ-LOPEZ et al., 2014).

Segundo Viscarra-Rossel e McBratney (2008) a aplicação de TSR proximal em solos pode reduzir os custos de amostragem e melhorar ou substituir os métodos tradicionais de laboratórios de análise de solo. Dotto (2012) utilizou a técnica de espectroscopia de reflectância difusa na predição do teor de argila no solo de 412 amostras no estado do Rio Grande do Sul, para enquadramento nos tipos texturais segundo a Normativa n. 2 do Ministério da Agricultura Pecuária e Abastecimento (MAPA). Os resultados mostraram que 80% das amostras não precisariam ser analisadas pelo método tradicional de determinação de argila, sendo apenas necessária a leitura das amostras em um espectro radiômetro em laboratório.

Em relação aos trabalhos publicados no Brasil, observa-se que a maioria dos trabalhos sobre o mapeamento de propriedades do solo no Brasil, com destaque para o COS, não usam os fundamentos do MDS baseado no modelo SCORPAN discutidos por McBratney et al. (2003), com exceção do trabalho de Mendonça-Santos et al. (2007). A maioria dos trabalhos utiliza a abordagem geoestatística para realizar a espacialização do COS na paisagem, sendo requerido para isso um grande número de observações.

Os cientistas do solo brasileiros tem um longo caminho a ser percorrido com relação à pesquisa em MDS de propriedades do solo, sendo necessários estudos com ênfase em

protocolos para gerar mapas de COS em diferentes escalas. Nesse sentido, é possível realizar a predição do COS utilizando como variáveis preditoras apenas atributos do terreno derivados de MDE, já que essas informações são de fácil acesso.

Essas informações são exigidas para compreender os processos relacionados com a mitigação de gases de efeito estufa e qualidade ambiental. Segundo Grunwald (2009) e Sanchez et al. (2009), é esse tipo de informação que os gestores, tomadores de decisão e a sociedade como um todo necessitam atualmente para o planejamento e conservação dos recursos naturais.

2.4 Modelos digitais de elevação (MDE)

Uma das formas de representar numericamente uma característica ou fenômeno espacial ambiental é através do Modelo Digital do Terreno (MDT). O conceito de MDT foi introduzido por MILLER e LAFLAMME (1958) como sendo “uma representação estatística da superfície contínua do terreno por um grande número de pontos selecionados com coordenadas conhecidas X , Y e Z em um campo arbitrário de coordenadas”. Em resumo, o MDT representa de forma quantitativa a variação contínua do relevo na paisagem (Moore et al., 1993). Exemplos de fenômenos ambientais que os MDT podem representar são dados de altitude, variáveis geoquímicas e geofísicas, declividade, entre outras. O MDE é um exemplo de MDT, que representa matematicamente a superfície do terreno por um grande número de pontos com coordenadas planimétrica (X e Y) e um valor de intensidade do pixel, que corresponde à elevação (Z) (PARANHOS FILHO et al., 2008).

Segundo Hutchinson e Gallant (2000), um MDE pode ser confeccionado a partir de três fontes de dados diferentes: pontos cotados ou digitalização de curvas de nível, a partir de dados de elevação de sensoriamento remoto aéreo e orbital ou através de dados obtidos por laser scanner. Pontos cotados da superfície do terreno, incluindo pontos altos e baixos. Esses dados podem ser obtidos através de levantamentos topográficos e manualmente através de estéreo-modelos fotogramétricos.

Com o advento do GPS, os levantamentos planialtimétricos em campo com pontos de maior precisão foram facilitados. Atualmente, são utilizados três métodos para obtenção de dados altimétricos via GPS: o estático, o cinemático ou semi-cinemático. No primeiro método, os pontos no terreno são coletados com GPS na forma estática (parado por alguns

segundos). No segundo, o terreno é percorrido (caminhamento) com o GPS onde é coletado um grande número de pontos a cada intervalo pré-definido. O terceiro método é uma combinação dos dois anteriores. Os dados coletados são submetidos ao pós-processamento com o objetivo de fazer correções e calcular a precisão das observações. A coleta de pontos no campo é aplicável apenas em pequenas extensões de área devido ao elevado tempo de execução e custo.

Um mapa de isolinhas é a representação de uma superfície por meio de curvas de nível que apresentam os valores de altitudes. As curvas de nível ainda são a fonte de dados altimétricos mais comum tratando-se de grandes áreas. Estes dados são digitalizados a partir de mapas topográficos existentes que são a única fonte de dados altimétricos para algumas partes do mundo. No Brasil um exemplo do uso deste tipo de dados é a digitalização das curvas de nível das cartas topográficas do Exército. As curvas nível também podem ser extraídas automaticamente de estéreo-modelos fotogramétricos, embora este método esteja sujeito a erros (HUTCHINSON e GALLANT, 2000). Santos (2001) considera a digitalização de documentos cartográficos como a técnica de extração de atributos menos precisa existente. Isto devido aos originais de impressão que trazem incorporados em si erros inerentes aos procedimentos operacionais de confecção da folha topográfica no formato analógico.

A aquisição de dados altimétricos é cada vez mais estimulada com o objetivo de geração de MDE de alta resolução e com cobertura global. Para isso é preciso estudos a respeito da precisão das técnicas de interpolação para a geração de MDE em relação aos tipos de terreno com diferentes geomorfologias e quantidade de dados (DOBOS et al., 2000; CHAPLOT et al., 2006). Produtos de sensoriamento remoto podem fornecer larga cobertura espacial e com grande volume de dados, mas com algumas limitações. Nenhum dos sensores pode medir, de maneira confiável, a altimetria em área com alta cobertura de dossel vegetativo. Mesmo com a ausência de cobertura vegetal no terreno todos os métodos medem altitudes com erros significativos associados às limitações inerentes aos instrumentos de observação, como também ao declive e a rugosidade da superfície (HUTCHINSON e GALLANT, 2000). Os métodos também necessitam de pontos de controle em campo precisamente localizados, nem sempre fáceis de ser coletados, devido ao custo e demanda de tempo, especialmente em regiões de relevo acidentado (HUTCHINSON e GALLANT, 2000).

Atualmente a interferometria de radar que usa a técnica de Radar de Abertura Sintética (*Spaceborne Synthetic Aperture Radar - SAR*) é um método alternativo ao método estereoscópico tradicional de extração de dados altimétricos. O SAR é um sensor ativo que atua na região de microondas do espectro eletromagnético, o qual não sofre a influencia de

fatores ambientais prejudiciais para a qualidade os dados de altimétricos, como nuvens e condições climáticas adversas (Dixon, 1995). Os dados obtidos da missão SRTM utilizam essa técnica, onde os MDE gerados podem apresentar uma precisão de 16m e em áreas de extensão limitada no globo podem apresentar erro padrão de altitude que variam de 1 a 3m (HUTCHINSON e GALLANT, 2000).

Ao longo dos anos as pesquisas têm mostrado o potencial do uso de informações derivadas dos MDE, as quais fornecem dados sobre o relevo da superfície terrestre para serem utilizados em levantamentos de recursos naturais como, o solo, e em pesquisas de cunho ambiental (TEN CATEN e DALMOLIN, 2014). Com o avanço das técnicas computacionais e análises de dados espaciais têm contribuído para o avanço e maior rapidez de execução dos levantamentos de solos no mundo. Assim, os atributos do terreno relacionados aos fatores de formação do solo, vêm sendo usados para compreender a relação solo-paisagem, as relações espaciais e temporais entre as propriedades do solo e variáveis ambientais (MCBRATNEY et al., 2003; MINASNY e MCBRATNEY, 2007; SANCHEZ et al., 2009).

Dessa forma, os atributos do terreno derivados dos MDE são utilizados como variáveis auxiliares na predição espacial de classes de solo e propriedades, como composição granulométrica, COS, cor, umidade, entre outras (GESSLER et al., 1995; MENDONÇA-SANTOS et al., 2007; PEI et al., 2010; KEMPEN et al., 2011; SAMUEL-ROSA et al., 2013; BALLABIO et al., 2012; TESKE et al., 2014), devido alta relação desses atributos com a formação do solo e a ampla disponibilidade de MDE (MCBRATNEY et al., 2003).

Diante disso, o uso de MDE na modelagem ambiental aumentou em escala global. No entanto, algumas regiões o globo carecem de cartas topográficas em escalas grandes para aplicação em estudos de cunho ambiental. No território brasileiro, apenas 14% da área possui cartas topográficas na escala 1:50.000 e somente 1% na escala 1:25.000, sendo a maioria desses levantamentos realizados a mais de 30 anos (ARCHELA, 2008). Uma alternativa para o Brasil e demais regiões do globo que apresentam esse problema, seria o uso de MDE globais disponíveis gratuitamente como, o ASTER (*Advanced Spaceborne Thermal Emission and Reflection Radiometer*) obtido por meio da técnica de estereoscopia ótica, o SRTM (*Shuttle Radar Topographic Mission*) obtido por interferometria de SAR ou os dados do TOPODATA obtidos a partir do refinamento dos dados originais do SRTM para o Brasil (VALERIANO e ROSSETI, 2009). Há também a possibilidade de se obter MDE a partir de dados comerciais como, por exemplo, imagens dos satélites QuickBird, ALOS (*Advanced Land Observing Satellite*) e SPOT (*Satellite Pour l'Observation de la Terre*), onde as imagens adquiridas apresentam alta resolução espacial (2,5 a 5 metros), podendo proporcionar melhor

qualidade altimétrica, mas conseqüentemente, apresentam elevado custo de aquisição (PINHEIRO et al., 2013) e nem sempre apresentam qualidade altimétrica superior aos MDE disponibilizados gratuitamente dependendo da região (IORIO et al., 2012).

As fontes de dados para obtenção dos MDE mais utilizadas internacionalmente são ASTER, o SRTM, Cartas Topográficas e em estudos realizados no Brasil observa-se o uso do MDE TOPODATA. No Brasil pesquisas em modelagem ambiental e MDS têm utilizado dados obtidos de MDE gerados através das curvas de nível obtidas das cartas topográficas como, por exemplo, os trabalhos de Giasson et al. (2006), Figueiredo et al. (2008), Samuel-Rosa et al. (2013) e ten Caten et al. (2013). Por outro lado, os MDE obtidos da missão SRTM possuem um diferencial que atraem muitos pesquisadores, como a sua disponibilidade gratuita em escala global e boa qualidade dos dados para fins de modelagem ambiental em escala pequena (GUTH, 2010; CARVALHO e CARVALHO, 2012).

Os MDE obtidos a partir de técnica de sensoriamento remoto orbital apresentam algumas vantagens como: cobertura territorial a nível mundial, resolução média à alta, apresentam pontos de controle no terreno, relatórios de possíveis erros de elevação e alta velocidade de processamento dos dados (DIXON, 1995; SUBRAMANIAN et al., 2005). Entretanto, estudos tem mostrado que as informações derivadas de um MDE podem apresentar limitações de uso em relação aos campos de aplicação, onde cada MDE apresenta suas potencialidades, estando isso relacionado com a sua resolução espacial e a magnitude dos seus erros (FELICÍSIMO, 1994; HUTCHINSON e GALLANT, 2000).

Nesse sentido, o sucesso na predição de classes e propriedades do solo depende em parte da escala em que ocorrem os processos pedogenéticos na paisagem e se a escala do MDE consegue representar as características no terreno (HUTCHINSON e GALLANT, 2000; MCBRATNEY et al., 2003). Este é um problema crítico, pois diferentes fatores pedogenéticos e processos paisagem operam em escalas espaciais distintas (FLORINSKY e KURYAKOVA, 2000) e, portanto, é preciso conhecer as limitações dos MDE em relação a escala. Ainda, a magnitude dos erros no MDE nem sempre estão disponíveis para os usuários e quando há essas informações, elas foram avaliadas para regiões específicas, diferentes do local onde se deseja aplicar os dados. Além disso, a precisão das técnicas de interpolação dos dados altimétricos é influenciada pela forma do terreno, densidade de dados e tamanho da escala (DOBOS et al., 2000; CHAPLOT et al., 2006), somada a falta de padronização das metodologias para geração dos MDE, assim como, pouca atenção é dada para a qualidade e limitações dos dados gerados e informações derivadas dos MDE para locais específicos,

sendo isso em muitos casos negligenciados (OKSANEM e SARJAKOSKI, 2005; LIN e OGUCHI, 2006).

Diante desse problema, alguns trabalhos têm sido desenvolvidos na tentativa de avaliar a qualidade dos MDE, afim de, apresentar dados com maior confiabilidade e limitações de uso conhecidas, para os usuários que utilizam essas informações. A abordagem mais comum utilizada para esse propósito é a comparação de pontos de elevação conhecidos e mais precisos (valores considerados verdade do terreno) com os valores do MDE desenvolvidos. O índice utilizado na avaliação da concordância entre as duas fontes de dados é a raiz quadrada do erro médio quadrático (RMSE) da elevação. As feições drenagem e curvas de nível também são utilizadas como critério de avaliação do MDE. Outros trabalhos têm utilizado a Norma Brasileira de Cartografia (NBC) que avalia a qualidade altimétrica dos MDE e define a escala e a classe de precisão do MDE (PINHEIRO et al., 2012; MOURA et al., 2014).

Pinheiro et al. (2012) estudando a qualidade de MDE com objetivo de serem utilizados no MDS em uma bacia hidrográfica de 125.078 ha, com variação de relevo de plano a escarpado, observaram melhor precisão dos dados SRTM em relação ao MDE gerado a partir da carta topográfica. Esses autores realizaram uma avaliação qualitativa com base nos aspectos da superfície (feições de drenagem e curvas de nível) e número de depressões espúrias. Nessa mesma linha de estudo, Chagas et al. (2010) avaliando MDE em uma bacia de 28.000 ha, com relevo plano e escarpado, encontraram valores de RMSE de 12,00 14,00 e 37,00 metros para os MDE SRTM, CARTA e ASTER, respectivamente. Iorio et al. (2012) avaliando o desempenho de MDE em uma extensão de área cobrindo três municípios no estado do Mato Grosso do Sul, com relevo suave ondulado, encontraram valores de RMSE de 19,83, 17,23 e 16,69 metros para o TOPODATA, ASTER e SRTM respectivamente.

Mukherjee et al. (2013) avaliando a precisão vertical de MDE, observaram o mesmo comportamento para os valores de altitude dos MDE obtidos pelo ASTER e SRTM em uma área na Índia. Segundo estes autores, isso está associado à presença de vegetação, artefatos na superfície e terreno plano, onde os MDE gerados a partir de dados de sensores remotos orbitais, o ASTER é o mais influenciado por esses fatores (GUTH, 2010). Slater et al. (2009), que avaliaram esse modelo em 20 locais diferentes, e constataram que, em quase todas as áreas, a média de elevação do ASTER foi menor que as altitudes medidas no campo. Da mesma forma, Guth (2010) estudando a qualidade de MDE obtidos de dados de sensores remotos, observou que o ASTER apresenta valores de altitude menor em relação ao SRTM.

Druzina (2007), que avaliou a qualidade dos MDE SRTM, CARTA e ASTER no município de Porto Alegre no estado do RS. Os autores observaram que em áreas com relevo

plano a suave ondulado, o MDE gerado a partir da CARTA apresentou maior RMSE (5,64 m) em relação à área com relevo misto (RMSE = 4,61 m). Para o MDE SRTM e ASTER os autores encontram RMSE de 3,40 e 4,62 m para áreas planas; 8,26 e 6,69 m para áreas mistas; 10,96 e 6,52 m para áreas de relevo acidentado, respectivamente.

Estudo realizado em Honduras por Jarvis et al. (2004) mostrou que o MDE, na resolução 90 metros, gerados a partir de dados do SRTM apresentam maior acurácia em relação a MDE obtido de dados extraídos de uma carta topográfica (escala 1:50.000). Nesse estudo os erros médios dos valores de altitude foram de 8m para o MDE SRTM e 20 m para a MDE CARTA. Em relação a MDE gerados a partir de dados do ASTER, Hirano et al. (2003) encontrou valores de RMSE entre 7 e 15m na avaliação de MDE comparando com pontos de controle de campo. Segundo os autores, valores menores podem ser alcançados quando utilizados pontos de controle no campo e imagens de boa qualidade. Além disso, os dados do ASTER tem potencial para uso na modelagem ambiental em escalas de 1:50.000 a 1:100.000 e intervalos das curvas de nível de 40m ou mais.

Estudos também tem avaliado a influência da incerteza dos preditores ambientais derivados dos MDE nos modelos de predição de solos (BISHOP et al., 2006). Giasson et al. (2013) estudaram o uso de três diferentes MDE e de cinco algoritmos árvores de decisão no MDS de classes de solos. Os autores constaram que o modelo preditivo gerado pelo MDE ASTER produziu mapas de solos com maior semelhança com o mapa de referência. Teske et al. (2014) avaliaram o desempenho de seis MDE com diferentes escala no MDS de classes de solos utilizando árvore de decisão. Os autores concluíram que os MDE com maior resolução espacial, apresentaram maiores oscilações nos valores da elevação e geraram correlações solo-paisagem menos precisas e menor quantidade de unidades de mapeamento preditas em relação ao MDE de menor resolução espacial. Em relação a resolução dos MDE, Cavazzi et al. (2013), demonstraram que em áreas com relevo menos declivoso, os MDE com menores resoluções espaciais (140 m) geram modelos de predição de solos com maior precisão.

Diante do exposto, nota-se que diferentes MDE apresentam características distintas, as quais podem ser negativas ou positivas para o usuário, refletindo no produto final gerado pelo uso dessas informações. Portanto, definir a precisão e escala do MDE é importante, uma vez que, existe protocolo para execução de levantamentos pedológicos baseados na escala da base cartográfica (EMBRAPA, 1995), sugestões de resolução dos MDE conforme a escala de mapeamento (MCBRATNEY et al., 2003) e uso de multi-escalas para análise digital do terreno em estudos de MDS. Estudo de Behrens et al. (2010) mostrou isso, onde cada classe de solo é melhor predita por distintas combinações de variáveis preditoras (elevação,

declividade, entre outras) filtradas em diferentes escalas (pixel de 30 m e 15 m, por exemplo) a partir da análise do terreno. de Ainda, segundo Dalmolin et al. (2004) são escassos os levantamentos de solos em escala compatível com as necessidades do Brasil, que desenvolvem programas de manejo e conservação do solo em escala de bacia hidrográfica e propriedade rural. Neste sentido, qual seria o MDE mais adequado para mapear solos e suas propriedades em escala detalhada (\geq a 1:20.000) e ultradetalhada (1:5.000, 1:2.000, 1:1.000 e 1:500) ? O mesmo MDE apresenta a mesma precisão altimétrica em diferentes escalas?

Diante disso, surge a necessidade de conhecer a precisão e a escala dos MDE para assim definir a amostragem de solo, nível de detalhamento dos levantamentos de solos e suas propriedades e aumentar a precisão do MDS. No entanto, há uma carência de estudos para áreas de pequena extensão territorial, além de não haver a observância da precisão altimétrica dos MDE com relação à escala conforme especificado pela NBC e a não utilização de dados de campo como referência para avaliação dos MDE.

3 HIPÓTESE

- Os dados altimétricos e atributos do terreno (AT) derivados da carta topográfica, modelos digitais de elevação globais gratuitos, apresentam a mesma qualidade nas informações em relação a um levantamento planialtimétrico (LP) em escala de propriedade rural com pouca variação de relevo,

- A partir dos AT podem-se estabelecer relações para a construção de funções de predição espacial para o carbono orgânico do solo (COS) em escala propriedade rural, com capacidade de explicar mais de 50% da variância.

4 OBJETIVOS

4.1 GERAL

Avaliar a qualidade altimétrica e limitações de MDE com a finalidade de auxiliar na escolha do MDE mais adequados para derivar covariáveis do terreno para desenvolver funções de predição espacial para serem aplicadas ao mapeamento digital do carbono orgânico do solo em escala de propriedade rural no Planalto do Estado do Rio Grande do Sul.

4.2 ESPECÍFICOS

- Avaliar a qualidade dos dados altimétricos e dos AT derivados da carta topográfica, modelos digitais de elevação globais gratuitos em relação ao LP para a área estudada,
- Avaliar se o mesmo MDE apresenta a mesma precisão altimétrica para diferentes escalas em área de pequena extensão e com pouca variação de relevo,
- Identificar o potencial dos modelos digitais de elevação globais gratuitos com resolução pequena (30 metros) em relação ao LP para serem utilizadas no MDS em áreas de pequena extensão e com pouca variação de relevo,
- Identificar os AT derivados no MDE que mais se correlacionam com o carbono orgânico do solo e sua distribuição na paisagem.

5 CAPÍTULO 1

QUALIDADE DE MODELOS DIGITAIS DE ELEVAÇÃO PARA SEREM UTILIZADOS NO MAPEAMENTO DIGITAL DE SOLOS

5.1 Introdução

A demanda por informações de propriedades do solo em resoluções finas para grandes áreas é crescente (HARTEMINK, 2008). Diante disso, AT derivados de modelos digitais de elevação (MDE) vêm sendo utilizados em uma frequência maior como covariáveis auxiliares na predição espacial de classes e propriedades do solo (ARRUDA et al., 2013; TEN CATEN et al., 2013; TESKE et al., 2014) devido alta relação desses atributos com a formação do solo e a ampla disponibilidade de MDE (MCBRATNEY et al., 2003).

O MDE representa de forma quantitativa a variação contínua do relevo na paisagem (MOORE et al, 1993). O MDE pode ser gerado a partir de dados topográficos de levantamentos de campo, interpolação de bases vetoriais (pontos cotados, curvas de nível e hidrografia) extraídos de cartas topográficas, pares de imagens estereoscópicas em aerofotos e por meio de imagens de satélites obtidas por sensores óticos, radares orbitais ou laser scanner (HUTCHINSON e GALLANT, 2000).

No território brasileiro, apenas 14% da área possui cartas topográficas na escala 1:50.000 e somente 1% na escala 1:25.000, sendo a maioria desses levantamentos realizados a mais de 30 anos (ARCHELA, 2008). Essa realidade também é comum em outras regiões do globo que carecem de cartas topográficas em escalas grandes. Uma alternativa para esses locais seria o uso de MDE globais disponíveis gratuitamente como, o ASTER obtido por meio da técnica de estereoscopia ótica, o SRTM obtido por interferometria ou os dados do TOPODATA obtidos a partir do refinamento dos dados originais do SRTM para o Brasil.

As informações derivadas de um MDE podem apresentar limitações de uso em relação aos campos de aplicação (VALERIANO, 2008). Cada MDE apresenta suas potencialidades, estando isso relacionado com a sua resolução espacial e a magnitude dos seus erros (HUTCHINSON e GALLANT, 2000). Isso porque o sucesso na predição de classes e

propriedades do solo depende em parte da escala em que ocorrem os processos pedogenéticos na paisagem e se a escala do MDE consegue representar as características no terreno (HUTCHINSON e GALLANT, 2000; MCBRATNEY et al., 2003). Ainda, a magnitude dos erros no MDE nem sempre estão disponíveis para os usuários e quando há essas informações, elas foram avaliadas para regiões específicas, diferentes do local onde se deseja aplicar os dados. Além disso, não existe uma padronização das metodologias de interpolação para geração dos MDE (CHAPLOT et al., 2006), assim como, a avaliação da qualidade e limitações das informações derivadas dos MDE para locais específicos, sendo isso em muitos casos negligenciados (LIN e OGUCHI, 2006).

Diante desse problema, trabalhos têm sido desenvolvidos na tentativa de avaliar a qualidade dos MDE, a fim de apresentar dados com maior confiabilidade e tornar suas limitações conhecidas para posterior aplicação no MDS (CHAGAS et al., 2010; PINHEIRO et al., 2012). A abordagem mais comum utilizada para esse propósito é a comparação de pontos de altitude conhecidos e mais acurados (valores considerados verdadeiros) com os seus pontos respectivos do MDE, como mostra os trabalhos de Chagas et al. (2010) e Pinheiro et al. (2012). No geral, o índice comumente utilizado na avaliação é a raiz quadrada do erro médio quadrático (RMSE) dos valores de altitude. As feições drenagem e curvas de nível também são utilizadas como critério de avaliação do MDE.

Outros trabalhos na área de cartografia (PINHEIRO, 2006; IORIO et al., 2012; MOURA et al., 2014), têm utilizado a Norma de Cartografia Brasileira (NCB) que avalia a qualidade altimétrica dos MDE e define a escala e a classe de precisão do MDE. Contudo uma limitação da maioria desses estudos é não realizar um levantamento topográfico para a coleta in loco de dados de referência (verdade de terreno) para avaliação dos MDE. Outra maneira de avaliação dos MDE é estudar o desempenho dos modelos preditivos de classes de solos gerados a partir dos AT de distintos MDE, como foi abordado nos trabalhos de Giasson et al. (2013) e Teske et al. (2014). No primeiro, os autores estudaram o uso de três diferentes MDE no MDS de classes de solos, e concluíram que o modelo preditivo gerado pelo MDE ASTER produziu mapas de solos com maior semelhança com o mapa de referência. No segundo, foi avaliado o desempenho de seis MDE com diferentes escala no MDS de classes de solos utilizando árvore de decisão. Foi observado que os MDE com maior resolução espacial, apresentaram maiores oscilações nos valores da elevação e geraram correlações solo-paisagem menos precisas.

Associado a isso, existem protocolos para execução de levantamentos pedológicos baseados na escala da base cartográfica (EMBRAPA, 1995), sugestões de resolução dos MDE

conforme a escala de mapeamento (MCBRATNEY et al., 2003) e uso de multi-escalas para análise digital do terreno em estudos de MDS (BEHRENS et al., 2010). Diante disso, surge a necessidade de conhecer a precisão e a escala dos MDE para assim definir a amostragem de solo, nível de detalhamento dos levantamentos de solos e suas propriedades e aumentar a precisão do MDS. Trabalhos de Chagas et al. (2010) e Pinheiro et al. (2012) avaliaram a qualidade dos MDE para áreas de grande extensão territorial que podem ser utilizadas em levantamentos de solos de escala pequena. No entanto, há uma carência de estudos para áreas de pequena extensão territorial, além de não haver a observância da precisão altimétrica dos MDE com relação à escala conforme especificado pela NBC e a não utilização de dados de campo como referência para avaliação dos MDE.

Nesse contexto, o objetivo do presente trabalho foi avaliar a qualidade altimétrica e limitações quanto à escala de quatro MDE globais disponíveis gratuitamente em comparação com um MDE por levantamento planialtimétrico (LP). Esse diagnóstico tem por finalidade auxiliar na escolha de MDE mais adequados para derivar covariáveis do terreno para serem utilizados em estudos de modelagem ambiental e no MDS em nível de bacia hidrográfica de pequena extensão e com pouca variação de relevo.

5.2 Material e Métodos

O estudo foi realizado em uma área de 937 ha localizada no município de Giruá, no estado Rio Grande do Sul (RS), Brasil (Figura 1). A área faz parte da carta topográfica Giruá, pertence à sub-região das Missões, região fisiográfica do Planalto do RS. A área apresenta altitudes que variam entre 340 e 420 m, com média de 380 m. O clima predominante é do tipo Cfa conforme a classificação de Köppen, caracterizado como clima subtropical úmido sem período estiagem definido (ALVARES et al., 2013). O relevo é suave ondulado a ondulado, com declividades mais acentuadas próximo das drenagens, sendo comuns colinas com declividades de 3 a 10 %. A área apresenta geologia derivada da Formação Serra Geral. As classes de solos predominante na área são os Latossolos Vermelhos Distroféricos típicos e Gleissolos Háplicos Tb Distrófico típicos. As coberturas da terra predominantes na área são: mata nativa, campo nativo e lavouras anuais em sistema plantio direto (SPD).

Foram utilizados neste estudo, cinco MDE (Figura 2) obtidos de formas distintas, os quais foram avaliados com relação à qualidade altimétrica e ao potencial para derivar atributos do terreno a serem utilizados no MDS.

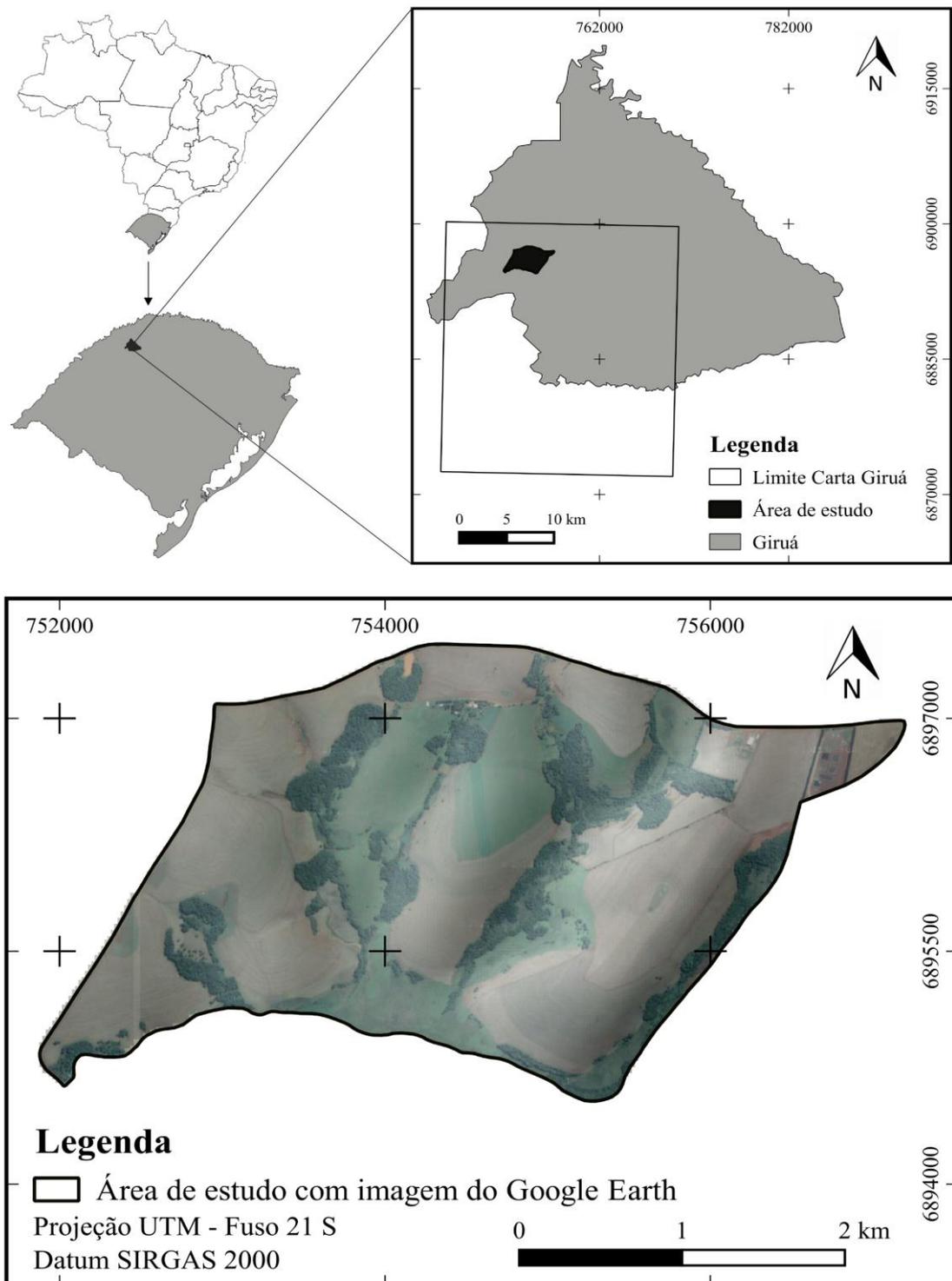


Figura 1 – Localização de área de estudo.

Os MDE submetidos à avaliação neste estudo foram:

1) MDE-LP gerado a partir de um levantamento planialtimétrico (LP) realizado na área de estudo. O LP foi realizado a partir da demarcação de pontos no terreno de modo estático, utilizando um receptor GPS - *Global Positioning System*, topográfico, modelo Leica GS20, com frequência L1, e com precisão de ± 10 cm de erro para posicionamento e ± 30 cm de erro de vertical. Esses pontos foram submetidos a correções (ajustamento de observações geodésicas) nas posições indicadas pelos satélites medidas no campo com o receptor GPS. Foram demarcados 663 pontos na área de estudo para gerar um MDE com resolução espacial de 30 metros.

2) MDE-CARTA, gerado a partir das curvas de nível extraídas da carta topográfica Giruá (SH.21-X-B-III/1/SE – escala 1:50.000 – curvas de nível de 20 metros), disponíveis na Base Cartográfica do Estado do Rio Grande do Sul (HASENACK e WEBER, 2010). No SAGA GIS 2.1 (SAGA GIS, 2010), a ferramenta *Grid - spline interpolation* foi utilizada para realizar a interpolação das curvas de nível e gerar o modelo na resolução de 30 m. Nesse procedimento foi utilizado o interpolador *Thin Plate Spline (TIN)*.

3) MDE-ASTER obtido da base de dados espaciais da *United States Geological Survey* (USGS), disponível gratuitamente na resolução de 30 metros para todo globo.

4) MDE-SRTM obtido da missão *Shuttle Radar Topography Mission* (SRTM) versão 4. Os dados do SRTM estão disponíveis gratuitamente pelo USGS para todo globo, na resolução espacial de 90 metros.

5) MDE-TOPODATA obtido da base de dados do Projeto TOPODATA, disponível gratuitamente pelo INPE na resolução de 30 metros (VALERIANO e ROSSETI, 2009).

Para viabilizar a comparação dos MDE, todos foram gerados com resolução espacial de 30 m, utilizando a ferramenta *Resampling* no SAGA GIS. Além disso, os MDE foram projetados para coordenadas *Universal Transversa de Mercator* – UTM da Zona 21 Sul e o Sistema Geodésico de Referência – datum SIRGAS 2000. Além disso, a altitude dos pontos do LP e dos MDE representam a altitude em relação a um elipsóide de referência, com exceção dos pontos do MDE-CARTA que representam a altitude ortométrica. Para tanto, foi calculada a ondulação geoidal de cada ponto, para posterior transformação da altitude elipsoidal para altitude ortométrica. Esses procedimentos foram realizado no *software* MapGeo (IBGE, 2010). Segundo o IBGE o erro médio na região para ondulação geoidal é de 0,20 metros. Todo o geoprocessamento e geração dos MDE foram realizadas no SAGA GIS e as análises estatísticas no ambiente R (R CORE TEAM, 2014).

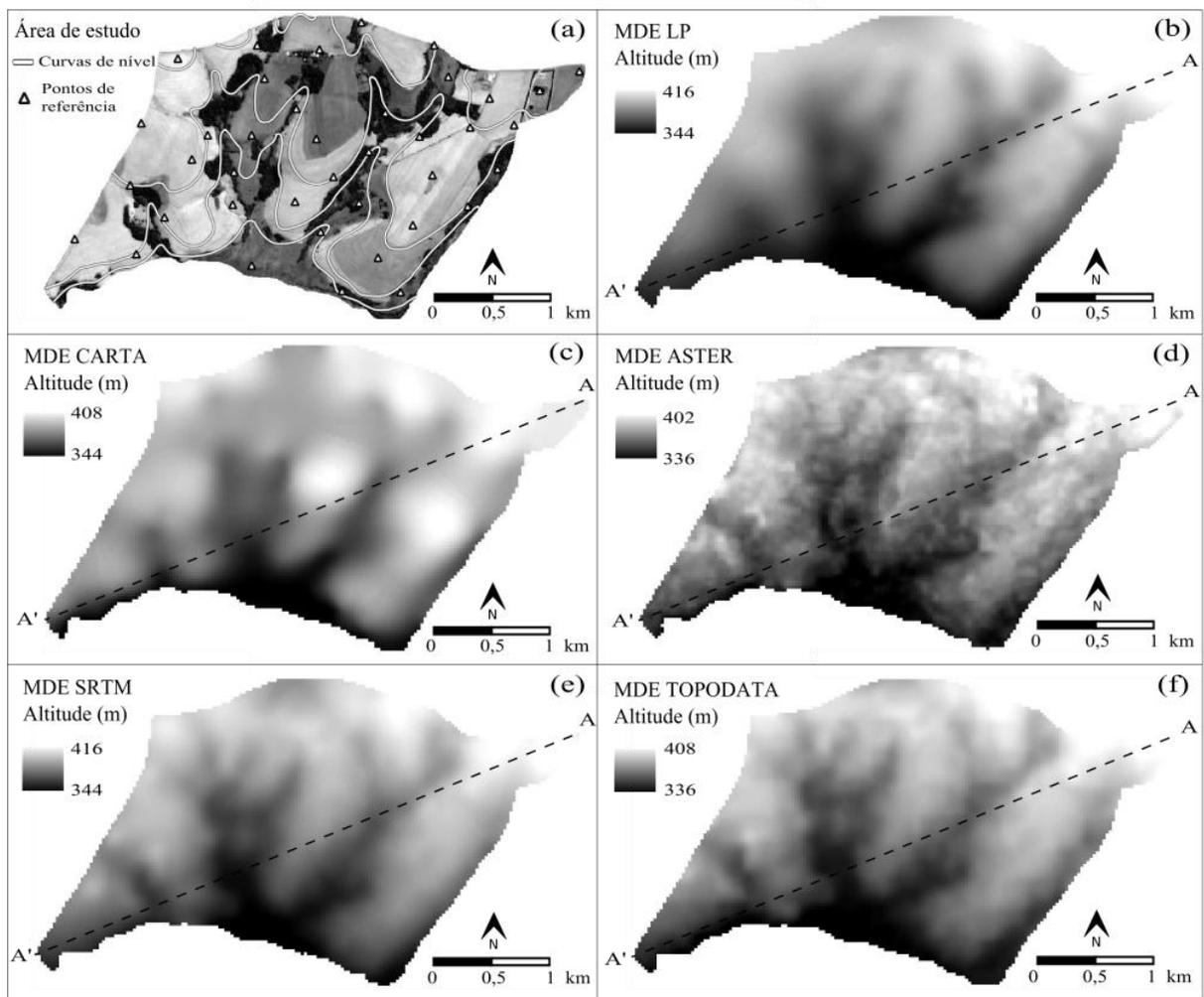


Figura 2 – (a) Área de estudo com a localização dos 41 pontos de referência (PR) e curvas de nível (20 m) da carta topográfica Giruá; (b) MDE-LP; (c) MDE-CARTA; (d) MDE-ASTER; (e) MDE-SRTM; (f) MDE-TOPODATA; a linha tracejada em cada MDE representa a transeção A-A'. Projeção UTM – Fuso 21 S / Datum SIRGAS 2000.

Os MDE foram avaliados quanto à qualidade dos dados altimétricos e dos seguintes atributos do terreno (AT) derivados de cada MDE: declividade, índice de umidade topográfica, fator LS e convergência. No entanto, os testes estatísticos foram aplicados apenas nos dados altimétricos, uma vez que, os AT são derivados a partir dos valores de altitude. Neste estudo, os valores dos AT de cada MDE foram comparados utilizando como referências os valores derivados do MDE-LP. Uma avaliação qualitativa foi realizada através da geração de mapas das diferenças de altitude de cada pixel para cada MDE, utilizando como referência o MDE-LP. Para isso foi utilizada a ferramenta *calculadora raster* do SAGA GIS 2.1.

O primeiro procedimento realizado na avaliação altimétrica foi calcular os erros de altitude (EH) de cada MDE a partir da diferença entre os valores da altitude dos pontos de

referência (H_{PR}) do valor de altitude dos MDE (H_{MDE}), a média dos erros (\overline{EH}) e o desvio-padrão dos erros (S_{EH}). Neste estudo, os valores de H_{PC} foram obtidos a partir de 41 PR distribuídos na área de estudo, demarcados utilizando a mesma metodologia da coleta dos pontos do LP. Com essas informações, foram calculados três parâmetros conforme as equações 1, 2 e 3 respectivamente, demonstradas a seguir:

$$EH = H_{PC} - H_{MDE} \quad (1)$$

$$\overline{EH} = \sum_{i=1}^n EH_i \quad (2)$$

$$S_{EH} = \sqrt{\frac{\sum_{i=1}^n (EH_i - \overline{EH})^2}{n-1}} \quad (3)$$

A partir do EH , \overline{EH} e S_{EH} obtidos, foram realizadas as seguintes análises estatísticas: 1) identificação de erros grosseiros na altitude; 2) cálculo da raiz quadrada do erro quadrático médio (RMSE); 3) teste de tendência; 4) teste de precisão e 5) aplicação na norma do Padrão de Exatidão Cartográfica dos Produtos Cartográficos Digitais (PEC-PCD) em cada MDE. Essas análises seguem as instruções da Norma Brasileira de Cartografia conforme as Especificações Técnicas para a Aquisição de Dados Geoespaciais Vetoriais (ET-ADGV) (BRASIL, 2011).

A identificação de erros grosseiros (EG) nos valores de altitude foi realizada utilizando equação 4:

$$EG = 3 * S_{EH} \quad (4)$$

Onde S_{EH} é o desvio-padrão dos valores de EH . Os valores de EH que forem maiores que o EG calculado, esse é considerado um ponto de altitude com erro grosseiro. Esse tipo de erro geralmente é causado por alguma falha no equipamento de coleta de dados.

A RMSE do valor do erro de altitude (EH_i) em relação aos 41 PR ($n=41$) foi calculada conforme a equação 5:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (EH_i^2)} \quad (5)$$

O Teste de Tendência tem como objetivo identificar os erros sistemáticos nos valores de altitude, por meio do teste de hipótese conforme mostra a expressão abaixo:

$$H_0: \overline{EH} = 0, \text{ não é tendencioso}$$

$$H_1: \overline{EH} \neq 0, \text{ é tendencioso}$$

Para testar essa hipótese deve-se calcular o valor de “ t_h ” conforme a equação (6):

$$t_h = \frac{\overline{EH}}{S_{EH}} * \sqrt{n} \quad (6)$$

Em que EH é a média dos erros e S_{EH} é o desvio padrão do conjunto dos erros e n é o número de pontos utilizados na avaliação. Para testar a hipótese utiliza-se o teste “ t ” de Student, a partir de valores tabelados, verificando se o valor de “ t_h ” está no intervalo de aceitação ou rejeição da hipótese nula. O intervalo de confiança do teste “ t ” é dado pela seguinte desigualdade:

$$|t_h| < t_{(n-1; \alpha/2)}$$

Dessa forma, se o valor do “ t_h ” for menor que o valor “ t ” tabelado para $n-1$ graus de liberdade e nível de significância α , os dados estão livre de tendências sistemáticas. Esse tipo de erro esta relacionado com fatores que podem influenciar na coleta de dados, como a copa das árvores, equipamentos desregulados, tecnologia de coletas de dados e fatores atmosféricos.

O Teste de Precisão é calculado comparando o desvio padrão dos EH com o erro padrão (EP) especificado pelo PEC-PCD para cada escala e classe, tentando a seguinte hipótese:

$$H_0: S_{EH}^2 = \sigma_H^2$$

$$H_1: S_{EH}^2 > \sigma_H^2$$

Onde σ_H é o erro padrão pré-estabelecido para cada classe e escala, e S_{EH} corresponde ao desvio-padrão dos valores de EH de cada MDE avaliado. Para testar essa hipótese, utiliza-

se o teste Qui-Quadrado (χ^2), com valores tabelados, para o nível de significância de 90%. O cálculo do desvio padrão esperado é realizado conforme a equação (7):

$$\chi^2_H = \frac{S_{EH}^2}{\sigma_H^2} * (n - 1) \quad (7)$$

Após, verifica-se se o valor de χ^2_{EH} se encontra dentro do intervalo de aceitação conforme a expressão abaixo:

$$\chi^2_H \leq \chi^2_{(n-1;\alpha)}$$

Se o valor não atender a expressão acima, rejeita-se a hipótese nula, e os valores de altitude do MDE não atendem a precisão pré-estabelecida.

Segundo a Norma Brasileira de controle de qualidade de produtos cartográficos (BRASIL, 2011), um produto cartográfico digital (PCD) para atender os critérios pré-estabelecidos pelo PEC-PCD, noventa por cento (90%) dos erros dos pontos coletados no PCD avaliado, quando comparado com as coordenadas demarcadas em campo com equipamento de alta precisão, devem apresentar os valores iguais ou inferiores aos previstos pelo PEC-PCD. Por exemplo, para um PEC-PCD de 1 metro (escala 1:5.000 – classe B), 90% dos pontos avaliados devem apresentar valores de EH menores que 1 m. Os valores do PEC-PCD para cada escala e classe estabelecidas pela Norma Brasileira de Cartografia (NBC) contempladas neste estudo são apresentados no Quadro 1.

Quadro 1 – Resumos das escalas contempladas no presente estudo, suas respectivas equidistâncias das curvas de nível (ECN), valores do erro padrão (EP) e PEC-PCD segundo a NB (BRASIL, 2011).

Escala	ECN	EP	PEC-PCD	EP	PEC-PCD	EP	PEC-PCD	EP	PEC-PCD
		Classe A ⁽¹⁾		Classe B		Classe C		Classe D	
----- m -----									
1:50.000	20 m	3,33	5,50	6,66	10,00	8,00	12,00	10,00	15,00
1:25.000	10 m	1,67	2,70	3,33	5,00	4,00	6,00	5,00	7,50
1:10.000	5 m	0,84	1,35	1,67	2,50	2,00	3,00	2,50	3,75
1:5.000	2 m	0,34	0,54	0,66	1,00	0,80	1,20	1,00	1,50

⁽¹⁾ Classes especificadas conforme o valor do PEC-PCD e erro padrão admitido. A rigidez nos valores de PEC-PCD e erros diminuem de A para D.

5.3 Resultados e Discussão

Na Figura 2 (a e b) são apresentados os perfis topográficos de cada MDE, a distribuição dos valores de altitude dos pontos de referência (PR) e as altitudes derivadas de cada MDE, respectivamente. Observa-se que a representação do relevo na área estudada foi diferenciada para cada MDE (Figura 2a), apresentando diferentes valores mínimos e máximos de altitude. Isso é um indicativo que as diferentes fontes de dados para obtenção de informações do terreno nessa área de estudo apresentam diferentes desempenho. Isso vai de encontro aos trabalhos de Chagas et al. (2010) e Teske et al. (2014), onde nesses trabalhos foi constatado que o relevo é representado de forma distinta por cada MDE.

Teske et al. (2014) constataram em seu estudo que os MDE que apresentam elevada variação nos valores de altitude geram modelos de predição de classes de solos menos acurados e com menor número de unidades de mapeamento preditas. Os resultados desses trabalhos e os observados na Figura 2 confirmam a necessidade de um diagnóstico dos MDE, visando selecionar o mais adequado para cada local e finalidade.

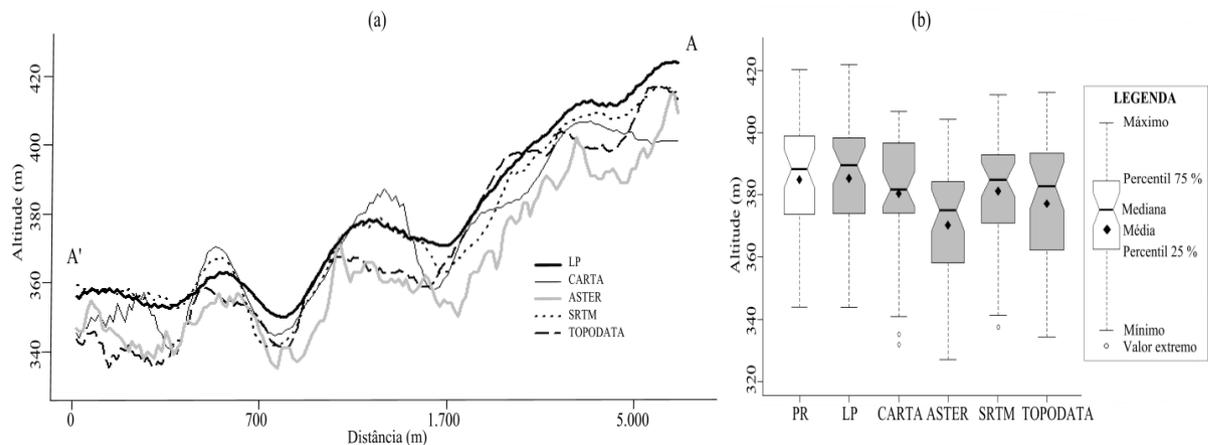


Figura 2 – perfis topográficos gerados a partir de transeção A-A' dos cinco MDE (a); distribuição das altitudes dos pontos de referência (PR) e dos cinco MDE estudados (b).

Não houve diferença entre os valores médios de altitude dos PR, MDE-LP, MDE-SRTM e MDE-TOPOTADA que apresentam valores médios de 385,13; 385,17; 381,10 e 380,54 metros respectivamente (Figura 2b). Isso é um indicativo que os valores de altitude desses MDE condizem com os valores de altitude dos PR, considerados a verdade de campo

neste estudo. Já os MDE-CARTA e MDE-ASTER apresentam valores médios de 370,37 e 376,95 metros, respectivamente, sendo esses valores diferentes em relação aos PR e aos demais MDE.

Os valores mais altos de altitude foram observados nos MDE-LP, MDE-SRTM e MDE-TOPODATA; e os mais baixos, nos MDE-CARTA e MDE-ASTER. O MDE-TOPODATA foi o modelo que apresentou maior amplitude, demonstrando maior variação nos valores de altitude. O MDE-ASTER foi o modelo que apresentou a maior diferença nos valores de altitude em relação aos PR e aos demais MDE (Figura 2b). Observa-se que os valores de altitude dos MDE-CARTA, MDE-ASTER, MDE-SRTM e MDE-TOPODATA são subestimados em relação aos PR e ao MDE-LP, com destaque para o MDE-ASTER.

Mukherjee et al. (2013) avaliando a precisão vertical de MDE, observaram o mesmo comportamento para os valores de altitude dos MDE obtidos pelo ASTER e SRTM em uma área na Índia. Segundo estes autores, isso está associado à presença de vegetação, artefatos na superfície e terreno plano, onde entre os MDE gerados a partir de dados de sensores remotos orbitais, o ASTER é o mais influenciado por esses fatores (GUTH, 2010).

Para o MDE-TOPODATA, segundo Valeriano e Rossetti (2012) em áreas com pouca variação de altitude, os dados altimétricos do TOPODATA podem ter sido afetados pela técnica de refinamento em relação a terrenos com relevo forte ondulado, refletindo na subestimação dos valores de altitude em relação aos PR como observado na Figura 2.

Observando a Figura 3, nota-se que o MDE-LP apresentou a melhor qualidade altimétrica, com valor médio do EH próximo a zero, confirmando a pequena diferença entre as altitudes dos PR e do MDE-LP, ou seja, os valores condizem com a verdade de campo nesta área de estudo. Esse resultado é confirmado pelo menor valor de RMSE (Figura 3) e amplitude dos valores de EH do MDE-LP em relação aos demais MDE avaliados neste estudo.

O MDE-SRTM apresentou qualidade superior comparado aos MDE CARTA, ASTER e TOPODATA, conforme os valores de EH e RMSE (Figuras 3 e 4c). Observa que os valores de altitude do MDE-SRTM foram pouco subestimados e superestimados em relação aos MDE da CARTA, ASTER e TOPODATA, sendo esses valores observados em locais com relevo forte ondulado e presença de vegetação arbórea (Figura 4c). Pinheiro et al. (2012) estudando a qualidade de MDE, através de uma avaliação qualitativa com base nos aspectos da superfície e número de depressões espúrias, com objetivo de serem utilizados no MDS em uma bacia hidrográfica de 125.078 ha, com variação de relevo de plano a escarpado, observaram melhor precisão dos dados SRTM em relação ao MDE gerado a partir da carta topográfica.

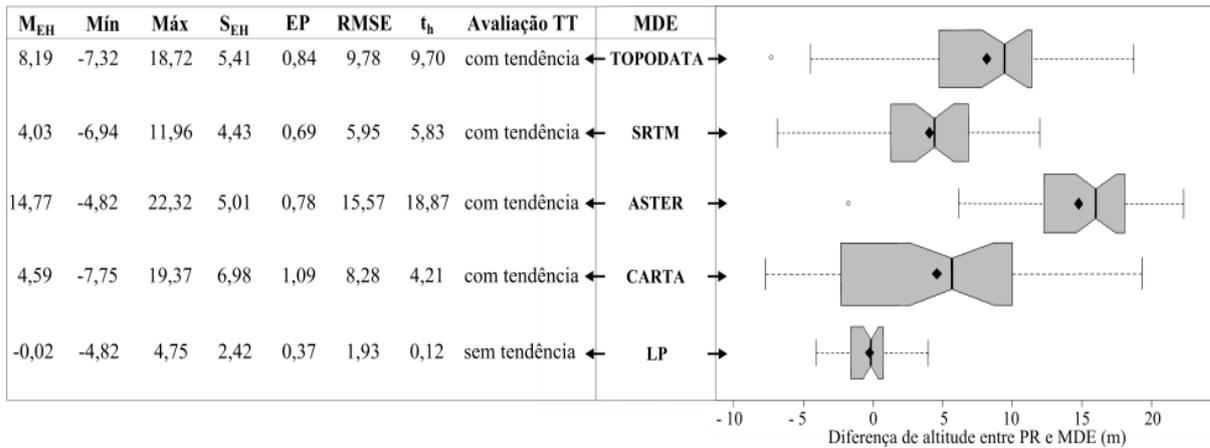


Figura 3 – Estatística descritiva dos valores do erro de altitude (EH), valor da raiz quadrada do erro médio quadrático (RMSE) e resultado do Teste de Tendência (TT) para os cinco MDE estudados.

Nessa mesma linha de estudo, Chagas et al. (2010) avaliando MDE em uma bacia de 28.000 ha, com relevo plano a escarpado, encontraram valores de RMSE de 12,00 14,00 e 37,00 metros para os MDE SRTM, CARTA e ASTER, respectivamente. No entanto, na avaliação geral a partir de indicadores qualitativos, como distribuição da rede de drenagem, curvas de nível e presença de depressões espúrias, o MDE gerado a partir dos dados da carta topográfica apresentou os melhores resultados.

Os resultados de RMSE encontrados para o MDE-TOPODATA (Figura 3) podem estar relacionados ao processo e refinamento dos dados do SRTM original que é a fonte de dados desse MDE. Em áreas com relevo suave ondulado o processo de refinamento da resolução de 90 para 30 metros, pode não ter alterado os valores de altitudes do pixel, conseqüentemente, os novos valores gerados pelo refinamento não se apresentam próximos dos valores do terreno (VALERIANO, 2008). Neste caso, como a área apresenta relevo com pouca variação, isso pode ter influenciado na qualidade desse modelo. Na Figura 4d observa-se que as maiores diferenças de altitude entre os MDE TOPODATA e MDE-LP ocorrem em locais de relevo plano a suave ondulado, onde os valores de elevação foram subestimados (valores positivos). Já em locais onde ocorre relevo ondulado, os valores foram superestimados (valores negativos).

Iorio et al. (2012) avaliando o desempenho de MDE em uma extensão de área cobrindo três municípios no estado do Mato Grosso do Sul, com relevo suave ondulado, encontraram valores de RMSE de 19,83 - 17,23 - 16,69 metros para o TOPODATA, ASTER e SRTM respectivamente. No entanto, comparando esses MDE gratuitos com resolução de

30m com MDE gerados a partir de imagens do sensor PRISM do satélite ALOS com resolução de 2,5 metros, no qual se esperava os melhores resultados devido a sua resolução, os autores concluíram que os MDE gratuitos apresentam melhor qualidade altimétrica.

O MDE-CARTA apresentou menor valor RMSE em comparação aos MDE ASTER e TOPODATA. Isso é confirmado pelos valores médios da altitude e do EH estarem próximos da altitude dos PR e do EH do MDE-LP respectivamente (Figura 2 e 3). Porém, apresentou precisão inferior ao MDE-LP e SRTM. Isso pode ser devido ao fato que as cartas topográficas foram geradas pela técnica de fotogrametria em 1975 sendo que, conforme, Subramanian et al. (2005) podem apresentar limitações em relação ao uso o GPS e técnicas de radar para coleta de dados para obtenção de MDE.

No MDE-CARTA observa-se elevada amplitude (27,13 m) no conjunto de valores do EH (Figura 3). Essa alta amplitude dos valores de EH indica que em alguns pontos do terreno os valores de altitude são subestimados e em outros são superestimados como mostra a Figura 4a. Fazendo uma análise em cada valor de altitude extraído desse MDE e comparando com os valores de altitudes dos PR, foi constatado que em 66% dos pontos os valores de altitude foram subestimados, refletindo no sinal positivo do valor médio dos EH (Figura 3). Pinheiro et al. (2006) constatou que os valores de altitude do MDE derivados das cartas topográficas podem apresentar tendência negativa em relevo plano, indo ao encontro do nosso estudo onde foi observado essa mesma tendência.

Essa baixa qualidade do MDE-CARTA está relacionada com a reduzida amostragem do terreno pelas curvas de nível com equidistâncias de 20 metros na área de estudo, devido ao relevo pouco ondulado e tamanho da área. Observando os pixels constatou-se que os maiores EH ocorrem nos pixels localizados no centro da distância entre duas curvas de nível (Figura 4a). Esse problema também foi constatado no estudo de Druzina (2007), que avaliou a qualidade dos MDE SRTM, CARTA e ASTER no município de Porto Alegre no estado do RS. Os autores observaram que em áreas com relevo plano a suave ondulado, o MDE gerado a partir da CARTA apresentou maior RMSE (5,64 m) em relação à área com relevo misto (RMSE = 4,61 m). Para o MDE SRTM e ASTER os autores encontram RMSE de 3,40 e 4,62 m para áreas planas; 8,26 e 6,69 m para áreas mistas; 10,96 e 6,52 m para áreas de relevo acidentado, respectivamente.

O MDE-ASTER apresentou os maiores valores de EH e RMSE, indicando a baixa qualidade altimétrica desse modelo para a área de estudo (Figura 3). O perfil topográfico do ASTER (Figura 2a) e o mapa de diferenças entre o MDE-LP e MDE-ASTER (Figura 4b) mostra que os valores de altitude foram subestimados. Esses dados são corroborados pelo

trabalho de Slater et al. (2009), que avaliaram esse modelo em 20 locais diferentes, e constataram que, em quase todas as áreas, a média de elevação do ASTER foi menor que as altitudes medidas no campo, ou seja, o EH apresentou tendência positiva, como observado na Figura 3 e Figura 4b. Da mesma forma, Guth (2010) estudando a qualidade de MDE obtidos de dados de sensores remotos, observou que o ASTER apresenta valores de altitude menor em relação ao SRTM.

Em relação ao Teste de Tendência (TT), todos os MDE apresentaram tendência a erros sistemáticos nos valores de altitude (Figura 3), com exceção do MDE-LP. Nesse MDE o t_h calculado foi menor que t_{tabelado} . Erros sistemáticos indicam a influencia de fatores externos na aquisição dos dados do terreno, como deslocamento de imagens, a presença de nuvens e erros nos métodos de interpolação (DIXON, 1995; DOBOS et al., 2000; CHAPLOT et al., 2006) que podem comprometer a qualidade dos dados. Em virtude da forma de aquisição de dados dos MDE-CARTA, MDE-ASTER, MDE-SRTM e MDE-TOPODATA, esse MDE são mais suscetíveis a erros sistemáticos em relação ao MDE gerado a partir de um LP. Portanto, o diagnóstico dos dados de altitude de um MDE por meio desse teste, pode auxiliar na escolha do MDE mais adequado para cada estudo e da necessidade de uma avaliação mais robusta da precisão altimétrica do MDE.

A qualidade dos MDE-CARTA, MDE-SRTM e MDE-TOPODATA encontrados nesse estudo (Figura 3), estão de acordo com alguns trabalhos encontrados na literatura, onde MDE gerados a partir destas fontes apresentam melhor qualidade em relação ao MDE obtidos com dados do ASTER (CHAGAS et al., 2010; MOURA et al., 2014). A geração de MDE que usam a técnica de estereoscopia, como o ASTER apresenta algumas desvantagens, como por exemplo, influencia da cobertura de nuvens, copa das árvores e absorção atmosférica do comprimento de onda, em relação a técnica de interferometria de radar com abertura sintética utilizada por exemplo na missão SRTM, considerada a técnica mais avançada para gerar informações do terreno. Segundo Guth (2010), a fonte de dados do MDE-ASTER apresenta maior quantidade de artefatos que são quantificados como valores de altitude, sendo isso intensificado em locais de relevo mais suave, prejudicando a caracterização do relevo.

Observando a Figura 4, pode-se fazer uma avaliação qualitativa por meio da identificação visual dos locais do terreno onde ocorrem os pixels com as maiores diferenças de altitude entre os MDE, considerando o MDE-LP como referencial, onde os dados mostram que existe uma diferença significativa entre os MDE. As causas dessas imperfeições nos MDE podem estar associadas aos métodos de interpolação utilizados para obtenção dos modelos (DOBOS et al., 2000; CHAPLOT et al., 2006). Segundo estudo de Hengl et al. (2004)

discrepâncias nos MDE obtidos de sensores remotos é comum em decorrência de erros grosseiros advindos do método de coleta de dados dos sensores, com destaque para o ASTER. Os MDE SRTM e CARTA apresentaram menor EH em relação aos MDE ASTER e TOPODATA, mostrando qualidade mediana para esta área de estudo.

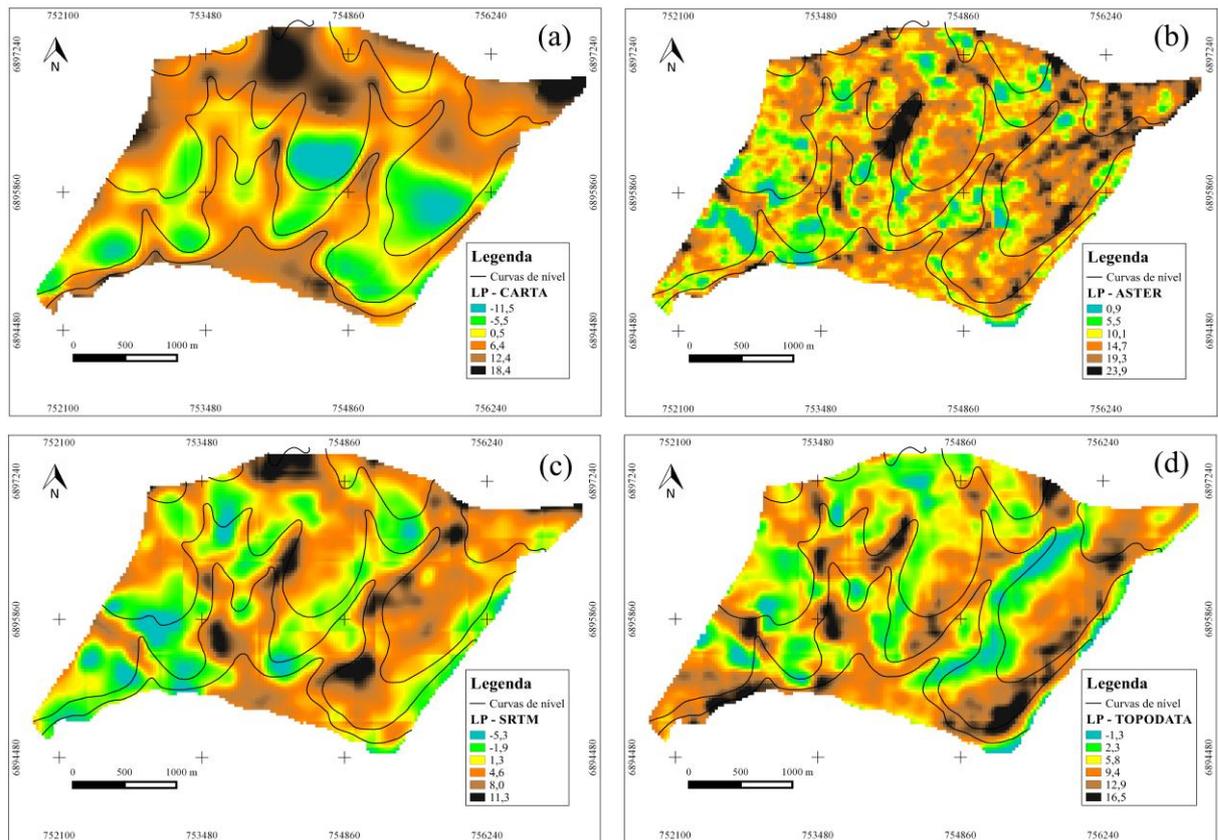


Figura 4 – Mapa de diferenças dos valores de altitude dos MDE CARTA (a), ASTER (b), SRTM (c) e TOPODATA (d) em relação ao MDE-LP. Projeção UTM – Fuso 21 S / Datum SIRGAS 2000.

Esses resultados tem implicação direta na caracterização dos aspectos de superfície do terreno e na direção de fluxo, os quais definem as relações hidrológicas na paisagem (Wine, 2000). Neste estudo, o terreno apresenta relevo com pouca variação e associado com a resolução espacial dos MDE de 30m, a presença de EH em vários locais da paisagem pode refletir significativa na caracterização do relevo da área e, conseqüentemente, na relação-solo paisagem e nos modelos de predição espacial de classes e propriedades do solo ajustados a partir de AT derivados desses MDE.

A comparação dos MDE segundo o Teste de Precisão (TP) tem como pré-requisito a definição de uma escala e classe conforme a Quadro 1, sendo definida de acordo com os valores do desvio-padrão (DP) de cada MDE investigado. Nesse estudo, para todos os MDE avaliados, optou-se por iniciar os teste a partir da escala 1:50.000 – Classe A (Quadro 1), em decorrência da utilização da carta topográfica nessa mesma escala. A partir disso, a avaliação da escala e classe mais adequada para cada produto cartográfico digital testado, seguiu até a maior escala possível em acordo com a NBC conforme o PEC-PCD (BRASIL, 2011).

O TP indicou que o MDE-LP atendeu as especificações para a escala 1:50.000 nas quatro classes (A, B, C e D). Na escala 1:25.000 atendeu apenas as classes B, C e D. Na escala 1:10.000 atendeu apenas a classe D. A escala 1:5.000 não apresentou precisão adequada em nenhuma das classes. O MDE-CARTA apresenta precisão apenas na escala 1:50.000 nas classes B, C e D. O MDE-SRTM atendeu ao TP na escala 1:50.000 nas classes B, C e D e na escala 1:25.000 na classe D. O MDE-TOPODATA atendeu ao TP apenas para a escala 1:50.000 na classe D.

Segundo a norma PEC-PCD a máxima escala que pode ser utilizada para a área de estudo desse trabalho com o MDE-LP é 1:10.000 – Classe D, onde é aceito um erro padrão dos valores de altitude menor ou igual a 2 metros. Para o MDE-CARTA a escala máxima foi 1:50.000 – classe B. Essa escala apresenta baixo detalhamento do terreno, não sendo recomendado para uso no MDS em áreas de pequena extensão territorial e com baixa variação de relevo. Esses resultados encontrados para o MDE-CARTA se referem a um modelo gerado a partir das informações derivadas da carta topográfica, onde no processo de geração do MDE podem ocorrer erros que influenciam no PEC-PCD. No entanto, para trabalhos em áreas maiores (bacias hidrográficas grandes e municípios) o MDE-CARTA na escala 1:50.000 pode ser utilizado como base cartográfica para projetos de levantamentos de solos semidetalhados como, por exemplo, o trabalho de Flores et al. (2007). O MDE-SRTM pode ser utilizado na escala 1:50.000 nas classe B,C e D e o MDE-TOPODATA apenas na classe D, em levantamentos de solos semidetalhados, com restrições maiores em relação a precisão.

Trabalho de Iorio et al. (2012) avaliando o PEC-PCD de MDE gratuitos com resolução 30m e MDE gerados a partir de imagens do sensor ALOS com resolução de 2,5 m, mostrou que esses MDE se enquadraram na escala 1:100.000 nas classe B e C, com exceção do MDE gerado do SRTM que atendeu a classe A nessa mesma escala. Isso mostra que diferentes MDE podem apresentar diferenças na precisão altimétrica, independentes da resolução espacial, sendo as características do terreno e a qualidade dos dados originais utilizados para gerar o MDE os fatores que mais influenciam na qualidade dos MDE.

Esse tipo de informação norteia estudos futuros de MDS em regiões com essas características. Esse enquadramento dos MDE em relação à escala, a qual é definida pela precisão dos dados altimétricos, possibilita a elaboração de estudos a posteriores como, por exemplo, mapeamento digital de propriedades com maior qualidade e confiabilidade das informações geradas. Isso vai de encontro aos estudos de Breimer et al. (1986) e McBratney et al. (2003), onde os autores relatam a existência de uma relação direta entre precisão, escala e objetivo do levantamento de solos. Conforme Embrapa (1995), a qualidade e escala da base cartográfica utilizada são fatores de extrema importância no mapeamento de solos.

Na Tabela 2 é apresentada a estatística descritiva e o RMSE dos AT derivados dos cinco MDE. A comparação neste caso foi realizada a partir dos valores dos atributos do MDE-LP (considerados valores de referência) em relação aos demais MDE.

Tabela 2 – Resultados da estatística descritiva dos atributos do terreno derivados de cada MDE.

MDE	Média	Mín ⁽¹⁾	Máx ⁽²⁾	DV ⁽³⁾	EP ⁽⁴⁾	RMSE ⁽⁵⁾	CV ⁽⁶⁾
Declividade (graus^0,5)							
LP	0,22	0,08	0,37	0,18	0,04	-	0,72
CARTA	0,24	0,05	0,48	0,22	0,05	0,20	0,88
ASTER	0,28	0,03	0,48	0,21	0,05	0,22	0,82
SRTM	0,22	0,04	0,35	0,16	0,03	0,17	0,71
TOPODATA	0,24	0,04	0,42	0,17	0,04	0,17	0,71
Índice de umidade topográfica (adimensional)							
LP	8,14	5,82	17,13	1,80	0,14	-	0,22
CARTA	8,15	6,05	17,54	1,71	0,13	2,05	0,14
ASTER	7,79	4,67	14,52	2,00	0,15	2,24	0,26
SRTM	8,11	5,62	14,34	1,44	0,11	1,63	0,18
TOPODATA	8,11	5,94	15,35	1,69	0,13	1,86	0,21
Fator LS (adimensional)							
LP	1,44	0,05	4,64	0,93	0,08	-	0,64
CARTA	1,69	0,06	8,29	1,57	0,12	1,37	0,93
ASTER	2,04	0,025	6,94	1,40	0,11	1,57	0,69
SRTM	0,59	0,012	1,70	0,36	0,003	1,06	0,61
TOPODATA	1,45	0,02	0,97	0,97	0,007	1,20	0,66
Convergência (adimensional)							
LP	0,34	-35,69	23,07	7,81	0,60	-	22,9
CARTA	1,10	-32,35	24,90	7,26	0,56	9,94	6,6
ASTER	0,88	-32,00	42,45	10,34	0,79	10,11	11,75
SRTM	0,35	-16,59	24,34	6,81	0,52	8,35	19,45
TOPODATA	0,32	-29,93	47,71	9,01	0,69	8,37	28,15

⁽¹⁾ Mín: valor mínimo; ⁽²⁾ Máx: valor máximo; ⁽³⁾ DV: desvio padrão; ⁽⁴⁾ EP: erro padrão; ⁽⁵⁾ RMSE: raiz quadrada do erro médio quadrático; ⁽⁶⁾ CV: coeficiente de variação.

Observando os parâmetros estatísticos dos quatro AT, os MDE-LP, MDE-CARTA, MDE-SRTM e MDE-TOPODATA apresentam resultados semelhantes, diferindo do MDE-ASTER. Analisando os valores do RMSE para cada atributo, nota-se que os menores erros em relação ao MDE-LP são observados nos MDE-SRTM e MDE-TOPODATA, ou seja, apresentam valores mais próximos do MDE-LP, considerado a verdade de campo. O MDE-CARTA apresentou valores intermediários e o MDE-ASTER os valores mais altos de RMSE. Os resultados observados para o MDE-CARTA estão relacionados a baixa amplitude do relevo na área de estudo, onde a análise do terreno pelas curvas de nível é severamente agravada (VALERIANO, 2008). Em relação MDE-ASTER, os resultados são consequência da menor precisão altimétrica do ASTER neste estudo.

Esses AT estão diretamente relacionados com a velocidade dos fluxos superficiais no terreno, refletido na taxa de erosão, conteúdo de água no solo e nos processos de formação do solo (WILSON e GALLANT, 2000). Em trabalhos de MDS esses AT são utilizados na construção de modelos de predição de classes e propriedades do solo, neste caso, valores que não representam adequadamente o terreno, podem comprometer o desempenho dos modelos (McBRATNEY et al., 2003). Além disso, a escala do MDE interfere na análise morfométrica de bacias hidrográficas de pequena extensão (ZANATA et al., 2011).

Esses resultados indicam que não se pode aumentar a escala desses modelos a fim de obter maior detalhamento no terreno, uma vez que, os valores dos AT não seriam precisos o suficiente para representar o terreno corretamente. Assim, trabalhos de modelagem ambiental e MDS que utilizam os MDE devem levar em consideração a qualidade do MDE e sua escala para minimizar a propagação dos erros nas informações secundárias geradas.

Nesse sentido, os MDE obtidos por sensoriamento remoto, com baixa resolução, disponíveis gratuitamente, apresentam limitações para serem utilizados em áreas de pequena extensão com pouca variação no relevo, comparados ao MDE-LP. Porém, os MDE gratuitos (MDE-SRTM e MDE-TOPODATA) apresentaram resultados próximos o MDE-LP e enquadram-se na escala 1:50.000. Cabe ressaltar que esses MDE foram desenvolvidos para fornecer informações do terreno em escala global, ou seja, para aplicação em área de grande extensão com escalas entre 1:100.000 e 1:250.000 (VALERIANO, 2008).

Portanto, esses MDE gratuitos apresentam potencial para uso do MDS de pequenas bacias hidrográficas mediante uma pré-análise dos dados para condições de terreno específicas, além de haver a necessidade de execução de levantamentos planialtimétricos (LP) para diagnosticar o potencial e limitações dos MDE em relação a verdade de terreno. O limitante é a demanda de tempo e custo elevado para obtenção dos LP. Uma alternativa é o

uso de Veículos Aéreos Não Tripulado (VANT) para obtenção de dados altimétricos do terreno, sendo essa fonte de dados uma linha de pesquisa ainda pouco estudada no Brasil.

5.4 Conclusões

O MDE-LP apresentou, no geral, os melhores resultados, podendo ser utilizado para derivar AT na modelagem ambiental e no MDS para a área de estudo. O MDE-SRTM foi entre os MDE disponíveis gratuitamente o que apresentou resultados mais próximos do MDE-LP, seguido do MDE-TOPODATA. O MDE-CARTA apresentou comportamento intermediário e o MDE-ASTER o pior desempenho para a área de estudo.

Os MDE disponíveis gratuitamente apresentaram limitações quanto à qualidade dos dados altimétricos para áreas de pequena extensão e com pouca variação no relevo, porém podem ser utilizados para o MDS semidetalhado de classes de solos nessas áreas mediante um diagnóstico da precisão desses MDE utilizando pontos de referências demarcados *in loco*.

É necessário definir as escalas e classes compatíveis com a precisão altimétrica dos MDE a partir dos parâmetros estatísticos da Norma Brasileira de Cartografia para regiões específicas, com objetivo de melhorar a qualidade dos trabalhos de modelagem ambiental e MDS.

6 CAPÍTULO 2

FUNÇÕES DE PREDIÇÃO ESPACIAL DO CARBONO ORGÂNICO DO SOLO

6.1. Introdução

O solo está diretamente relacionado com fluxos de CO_2 no ambiente, uma vez que está entre os cinco maiores reservatórios de carbono do ecossistema terrestre, sendo responsável por 75% do acúmulo de carbono (BATJES e SOMBROEK, 1997). Este estudo de mostrou que os solos do Brasil, na camada superficial dos primeiros 30 cm, representam 5% do estoque mundial de carbono no solo. Nesse contexto, os solos podem atuar como dreno ou fonte de carbono (C) para atmosfera, dependendo do uso e manejo a que forem submetidos (MBAH e IDIKE 2011).

Para o melhor entendimento do processo de sequestro de carbono no solo e avanço da modelagem do fluxo líquido de carbono nos ecossistemas terrestres, é necessário de um entendimento dos processos de formação dos solos e da variabilidade espacial de carbono orgânico do solo (COS) na paisagem. O COS é uma das propriedades do solo funcionais do solo consideradas de fundamental importância de se ter inventários, uma vez que, pode ser utilizado para avaliar impactos ao solo, fornecer dados para a modelagem do ciclo do carbono e simulação de cenários futuros (CARRÉ et al., 2007; SANCHEZ et al., 2009).

No entanto, a variabilidade espacial na paisagem dessa propriedade do solo é complexa devida as interações com vários fatores ambientais, sendo isso um os principais problemas que limitam a modelagem mais precisa do conteúdo de COS (BERNOUX et al., 2002). Além disso, o elevado número de amostragens, associado ao custo e demanda de trabalho para quantificação do COS para posterior espacialização é um dos limitantes para mapeamento em áreas extensas e de difícil acesso.

Para contornar esse cenário, pesquisas têm mostrado o uso técnicas quantitativas digitais, a partir da linha de pesquisa em Mapeamento Digital de Solos (MDS). O MDS foi definido por Lagacherie e McBratney (2006) como “a criação e a população de sistemas de

informação espacial de solos por meio de modelos numéricos para inferir as variações espaciais e temporais de classes de solos e suas propriedades, a partir de observações, conhecimento e de dados de variáveis ambientais relacionados”. Assim, vários trabalhos têm utilizado métodos geoestatísticos e ajustes de funções de predição associados a variáveis preditoras como, os atributos do terreno derivados dos modelos digitais de elevação (MDE) e informações de solo e vegetação, para o mapeamento digital do conteúdo do COS (MENDONÇA-SANTOS et al., 2007; GRIMM et al. 2008; KHEIR et al., 2010; PEI et al., 2010; KEMPEN et al., 2011).

Nessa linha de pesquisa, Mendonça-Santos et al. (2007) testando técnicas de MDS para o mapeamento do estoque de COS, no estado do Rio de Janeiro, encontraram desempenho satisfatório para modelos de análise multilinear construídos a partir das variáveis preditoras mapas de solos, mapas de geologia e atributos do terreno derivados do MDE. Grimm et al. (2008) avaliado o desempenho de modelos de Árvore de Decisão (AD) do tipo *Random Forest* para a predição do COS na Ilha do Barro Colorado, encontraram resultados satisfatórios, utilizando como variáveis preditoras os atributos do terreno, formação geológica, mapas de solos e histórico da vegetação.

Kheir et al. (2010) mostrou que mapas digitais de COS podem ser desenvolvidos a partir de avaliações de amostras de solo e atributos do terreno como, por exemplo, índice de vegetação por diferença normalizada (NDVI), índice de umidade por diferença normalizada (NDWI), índice de cor do solo (SCI), índice de umidade topográfica (IUT), entre outros, sendo essas variáveis utilizadas como preditoras na construção de modelos preditivos de COS a partir do método de AD.

Kempen et al. (2011) estudando a distribuição tridimensional do conteúdo de matéria orgânica do solo (MOS), na província de Drenthe no nordeste da Holanda em uma área de 125 km², constataram que a modelagem geoestatística associada a variáveis preditoras é uma ferramenta promissora para o MDS. Pei et al. (2010) utilizando o método de krigagem ordinária para realizar o mapeamento da MOS, encontraram correlações significativas entre a MOS e índice de umidade topográfica (IUT), em uma área de 5.176 ha, com pouca diferença de altitude e declividade média de 5%.

Diante do exposto, estudos regionalizados e em áreas de atividade agrícola intensa são importantes para compreender os processos de distribuição do COS na paisagem, gerar informações para o planejamento e manejo conservacionista do solo e garantir uma escala apropriada de estudo da distribuição espacial do COS. Dessa forma, é possível construir funções de predição para COS em áreas pequenas e com pouca variação de relevo utilizando

como variáveis preditoras apenas os atributos do terreno derivados de um MDE com resolução espacial de 30 metros. Para buscar essa informação, o objetivo desse estudo foi construir funções de predição espacial para o COS, utilizando regressão linear múltipla em escala de propriedade rural, utilizando como variáveis preditoras apenas atributos do terreno.

6.2 Material e Métodos

O estudo foi realizado em uma área de 937 há (Figura 1), localizada no município de Giruá (RS), Brasil. A área apresenta altitudes que variam entre 340 e 420 m, com média de 380 m. O clima predominante é do tipo Cfa conforme a classificação de Köeppen, caracterizado como clima subtropical úmido sem período estiagem definido, com temperatura média anual de 19,5°C (ALVARES et al., 2013). O relevo é suave ondulado a ondulado, com declividades mais acentuadas próximo das drenagens, sendo comuns colinas com declividades de 3 a 10 %.

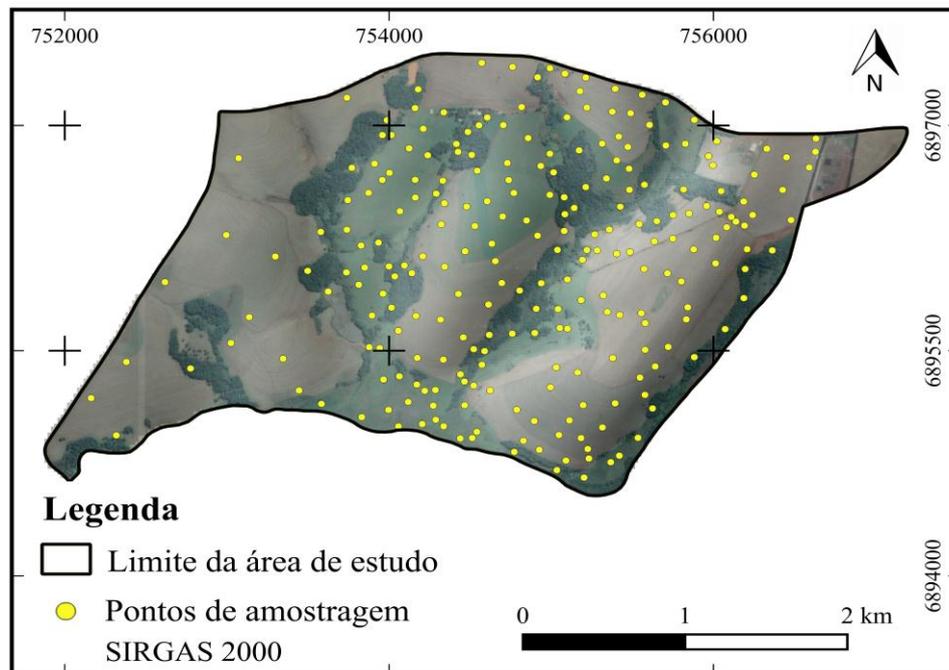


Figura 1 – Área de estudo e localização dos 243 pontos amostrados de solo na área.

A área apresenta geologia derivada da Formação Serra Geral, com rochas denominadas de basalto. As classes de solos predominante na área são os Latossolos Vermelhos Distroféricos típicos e Gleissolos Háplicos Tb Distrófico típicos. As coberturas da terra predominantes na área são: mata nativas, lavouras anuais em sistema plantio diretas (SPD) a mais de 10 anos, campo nativo transformado em lavoura anual com SPD e campo nativo com sistema de pecuária de corte.

Para caracterização do solo e suas propriedades foram coletadas 243 amostras (Figura 1) na camada superficial do solo na profundidade 0 – 5 cm. A definição dos pontos de coleta no campo foi realizada de duas formas: (1) foi confeccionada uma malha de amostragem irregular, no SAGA GIS 2.1, contendo 400 pontos, com o objetivo de coletar toda a variabilidade ambiental existente na área; (2) a partir dos 400 pontos foram selecionados 243 pontos com base em dois critérios, primeiro o conhecimento do pedológico a respeito da área de estudo, uso da terra, classe de solo e unidades geomorfológica. O segundo critério para escolha dos locais de coleta, foi o acesso fácil aos locais de amostragem. Em cada ponto amostral foi coleta três amostras e a partir disso foi obtida uma amostra composta que foi utilizada para as análises laboratoriais. Os pontos amostrais foram georreferenciados no campo utilizando um aparelho de GPS. Também foram coletadas informações ambientais sobre a condição de drenagem, presença de erosão, pedregosidade e demais informações pertinentes, para auxiliar na discussão dos resultados.

Em laboratório as amostras foram secas ao ar, destorroadas e passadas em peneira com malha de 2 mm, e submetidas à análises químicas em triplicata. O teor de carbono orgânico do solo (COS) foi determinado via combustão úmida (0,5 g de solo + 10 ml $K_2Cr_2O_7$ 0,067 mol L^{-1}) com aquecimento externo, segundo Yeomans e Bremner (1988). A titulação dos extratos do COT foi realizada com $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ 0,5 mol L^{-1} , utilizando como indicador a fenantrolina.

A partir do MDE-LP pré-selecionado no estudo 1, o qual apresentou a melhor qualidade dos dados altimétricos, foram derivados os atributos do terreno utilizados como variáveis preditoras na construção da FPES para o teor de COS. A derivação dos atributos do terreno foi realizada conforme Wilson e Gallant (2000) e executada no SAGA-GIS. As bandas da imagem LANDSAT-7 também foram utilizadas para gerar as variáveis relacionadas ao índice de vegetação que buscam combinar, sob a forma de razões, as informações contidas na radiância refletida pelos dosséis da vegetação, nas regiões do vermelho e infravermelho próximo do espectro eletromagnético.

As variáveis preditoras derivadas do MDE-LP e LANDSAT-7 foram: posição relativa do declive (RSP); profundidade de vale (VD); distância vertical da rede de drenagem (VDCN); elevação (ELEV); área de contribuição (CA); declividade média da área de contribuição (DMAC); declividade (DECL); aspecto (ASP); nível na base da rede de drenagem (CNBL); curvatura planar (PLAN); curvatura de perfil (PROF); índice de umidade topográfica (TWI); convergência (CONV); índice de rugosidade do terreno (TRI); índice potencial de escoamento (IPE); fator LS (LS); curvatura (CURV); refletância do infravermelho próximo (REFL-IVP); índice de vegetação por diferença normalizada (NDVI) e índice de vegetação ajustado ao solo (SAVI).

As informações contidas em cada variável preditora foi amostrada para os 243 pontos contendo as informações do teor de COS. Essa amostragem foi realizada no SAGA GIS, utilizando o algoritmo do vizinho mais próximo (*nearest neighbor*). A partir disso, foi gerada uma tabela contendo o valor de cada variável para os 243 pontos. Para gerar a FPES o banco de dados (n = 243) foi particionado aleatoriamente em dois conjuntos de dados. Um conjunto de calibração da FPES, composto por 170 amostras (70%) e outro de validação com 73 amostras (30%).

Antes de gerar a FPES foi realizada por meio da matriz de correlação linear entre as variáveis a identificação da multicolinearidade. A alta correlação entre duas variáveis é prejudicial para a construção de FPES, podendo causar desempenho inadequado da FPES. Nesse estudo foi estabelecido como valor máximo de correlação o valor de $r = 0,75$. Dessa forma, as variáveis preditoras que apresentam $r \geq 0,75$ foram excluídas do conjunto de dados.

A partir das variáveis pré-selecionadas foi construída a FPES para o teor de COS na camada de 0 – 5 cm, utilizando as 170 amostras por meio do modelo de regressão linear múltipla (RLM). A metodologia utilizada para a seleção das variáveis preditoras inseridas na RLM foi a estimação *stepwise*, onde o modelo é ajustado seguindo o Critério de Informação de Akaike (AIC). Esse método seleciona o melhor preditor para a variável dependente e as demais variáveis preditoras são incluídas no modelo à medida que essas aumentam o poder explicativo do modelo, resultando na maior minimização do valor de AIC.

A avaliação da qualidade do ajuste do modelo de RLM construído foi por meio de gráficos elaborados a partir dos seguintes parâmetros estatísticos:

Os resíduos (*resi*),

$$resi_i = \hat{y}_i - y_i \quad (1)$$

onde $resi_i$ são os resíduos, \hat{y}_i é o valor predito de COS na i -ésima observação e o y_i é o valor medido de COS na i -ésima e $i = 1, 2, \dots, n$ (HAIR et al., 2009).

Os resíduos padronizados ($resip_i$),

$$resip_i = \frac{resi_i}{Dp_r} \quad (2)$$

onde $resip_i$ são os resíduos padronizados, $resi_i$ é o resíduo da i -ésima observação e Dp_r é o desvio padrão dos resíduos e $i = 1, 2, \dots, n$. Para conjuntos de amostras com $n > 50$ os $resip_i$ seguem a distribuição t de Student, sendo assim, observações com valores de $resip_i > 1,96$ podem ser consideradas atípicas (outliers) (HAIR et al., 2009).

A distância de Cook ($dist_c$),

$$dist_c = \frac{\sum_{j=1}^n (\hat{y}_j - \hat{y}_{j(i)})^2}{p * QM_{erro}} \quad (3)$$

onde $dist_c$ é a distância de Cook, \hat{y}_j é o valor predito na j -ésima observação, $\hat{y}_{j(i)}$ é o valor predito na j -ésima observação quando a i -ésima observação é eliminada do ajuste do modelo, p é o número de graus de liberdade do modelo e QM_{erro} é o erro quadrático médio. Esse parâmetro visa avaliar a influência do peso da i -ésima observação sobre os demais n valores preditos pelo modelo, sendo utilizada para identificar *outliers*. Dessa forma, valores de $dist_c > 1$ podem influenciar no ajuste do modelo, sendo necessária a remoção da observação neste caso (HAIR et al., 2009).

Alavancagem (*alav*), resultado da diagonal do produto da matriz das variáveis independentes e indica o efeito que uma observação possui sobre os resultados do modelo ajustado devido às suas diferenças em relação às outras observações (HAIR et al., 2009). Seu valor varia entre 0 e 1, e o valor médio é igual a k/n , onde k é o número de variáveis preditoras no modelo e n é o número de observações utilizadas no ajuste. Para modelos de regressão ajustados com menos de dez parâmetros Hair et al. (2009) sugerem que seja utilizado como limite máximo aceitável de alavancagem o valor de três vezes a média ($3k/n$).

A utilização desses parâmetros mencionados acima tem por objetivo identificar as observações individualmente visando encontrar aquelas que estão fora dos padrões gerais do conjunto de dados ou que influenciam fortemente nos resultados do modelo. Dependendo da magnitude do impacto de uma observação no modelo, esta pode ser removida.

Por fim, foi aplicada a análise de variância (ANOVA) para avaliar se o ajuste do modelo de regressão para o conjunto de dados foi adequado. Essa análise permite identificar a contribuição de cada variável preditora, neste caso, os atributos do terreno utilizados na construção do modelo de predição na explicação da variância (EV) do COS na área de estudo. A EV pode ser calculada da seguinte forma:

$$EV(\%) = \frac{SQ_{total}}{SQ_{extra(i)}} * 100 \dots\dots\dots (4)$$

onde EV(%) é percentual da variância explicada pela variável preditora p , SQ_{total} é a soma de quadrados total, a qual estima a variância, e $SQ_{extra(i)}$ é a soma de quadrados extra associado a i -ésima variável preditora e $i = 1, 2, \dots, p$ (HAIR et al., 2009).

Para a validação da FPES construída foi utilizado o conjunto de amostra de validação ($n = 73$) e avaliada a qualidade do modelo pelos seguintes parâmetros estatísticos:

Erro médio (EM) ou diferença sistemática entre os valores preditos e medidos (BIAS)

$$EM = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \dots\dots\dots (5)$$

onde EM é o erro médio, \hat{y}_i é o valor predito de carbono na i -ésima observação e y_i é o valor medido na i -ésima observação e $i = 1, 2, \dots, n$. Valores positivos indicam superestimação e valores negativo indicam subestimação.

A Raiz quadrada do erro quadrático médio (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \dots\dots\dots (6)$$

onde RMSE é a raiz quadrada do erro quadrático médio, \hat{y}_i é o valor predito de carbono na i -ésima observação e y_i é o valor medido na i -ésima observação e $i = 1, 2, \dots, n$.

Relação de desempenho do desvio (RPD),

$$RPD = \frac{Dp}{RMSE} \quad (7)$$

onde RPD é a relação de desempenho do desvio, Dp é o desvio padrão dos valores de COS medidos calculado pela equação 8 e RMSE é a raiz quadrada do erro quadrático médio calculado pela equação 6. De acordo com Chang et al. (2001), valores de RPD acima de 2,0 pode ser considerado como modelos excelentes, de 1,4 à 2,0 modelos aceitáveis e menor que 1,4 modelos não confiáveis.

Desvio padrão (Dp),

$$Dp = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (8)$$

onde Dp é desvio padrão dos valores medidos de COS, y_i é o valor medido de COS na i -ésima observação, \bar{y}_i é a média dos n valores medidos de COS e $i = 1, 2, \dots, n$.

Coefficiente de determinação ajustado ($R^2_{ajustado}$),

$$R^2_{aj} = 1 - \frac{n-1}{n-(p+1)} (1 - R^2) \quad (9)$$

onde R^2_{aj} é o coeficiente de determinação ajustado, n é o número de observações utilizadas no ajuste da função de regressão, p é número de variáveis utilizadas no ajuste e R^2 é o coeficiente de determinação (HAIR et al., 2009).

A análise estatística dos dados foi realizada utilizando o pacote estatístico R Core Team (2014). O geoprocessamento dos dados foi realizado nos *softwares* QGIS 2.0 e SAGA GIS 2.1.

6.3 Resultados e Discussão

Na Tabela 1 é apresentada a estatística descritiva para os 17 atributos do terreno derivados do MDE e as três variáveis relacionadas com a vegetação, extraídas da imagem Landsat-7. Esse conjunto de dados constituiu as variáveis preditoras utilizadas para iniciar o processo de construção das FPES para os COS no presente estudo. Neste estudo para avaliar a distribuição dos dados em relação a média, a moda e a mediana foi adotado o valor de coeficiente de assimetria $\geq 0,50$, sendo neste caso, dados com valores superiores a este considerados dados que não apresentam uma distribuição próxima a curva normal.

Tabela 1 – Estatística descritiva do conjunto de das variáveis preditoras.

Variável preditora ⁽¹⁾	Mínimo	Máximo	Média	Desvio padrão	Coefficiente de assimetria ⁽²⁾
RSP - metro ⁻¹	0,00	0,99	0,50	0,34	- 0,15
VD - metro	0,00	47,24	16,68	13,72	0,33
VDCN - metro	0,00	52,70	20,26	14,95	0,25
ELEV - metro	332,98	413,44	377,33	20,53	- 0,29
CA - metro²	1.296,85	1.402.120,00	45.651,73	14.3979,5	5,69
DMAC - graus	0,00	0,08	0,04	0,01	- 0,15
DECL - graus	0,13	6,77	3,08	1,42	0,21
ASP - graus	53,68	338,67	194,29	59,39	0,15
CNBL - metro	335,36	378,91	357,08	8,88	- 0,47
PLAN - metro⁻¹	0,56	0,61	0,58	0,00	0,66
PROF - metro ⁻¹	0,21	0,63	0,40	0,07	- 0,08
TWI - adimensional	5,63	16,65	8,15	1,61	2,23
CONV - porcentagem	0,26	0,61	0,49	0,04	- 0,53
TRI - adimensional	0,06	2,54	1,17	0,53	0,26
IPE - adimensional	- 2,00	6,37	2,71	1,20	- 0,22
LS - adimensional	0,00	1,70	0,61	0,36	0,38
NDVI - adimensional	0,00	1,00	0,50	0,31	0,36
REFL-IVP - porcentagem	0,00	0,55	0,30	0,12	- 0,69
SAVI - adimensional	0,02	0,94	0,48	0,27	0,17
CURV - metro ⁻¹	0,15	0,69	0,38	0,08	- 0,03

⁽¹⁾ simbologia adotada para cada variável preditora seguida da unidade de medida; RSP: posição relativa do declive; VD: profundidade de vale; VDCN: distância vertical da rede de drenagem; ELEV: elevação; CA: área de contribuição; DMAC: declividade média da área de contribuição; DECL: declividade; ASP: aspecto; CNBL: nível na base da rede de drenagem; PLAN: curvatura planar; PROF: curvatura de perfil; TWI: índice de umidade topográfica; CONV: convergência; TRI: índice de rugosidade do terreno; IPE: índice potencial de escoamento; LS: fator LS; NDVI: índice de vegetação por diferença normalizada REFL-IVP: refletância do infravermelho próximo; SAVI: índice de vegetação ajustado para o solo; CURV: curvatura. ⁽²⁾ variáveis preditoras com coeficiente de assimetria $\geq 0,50$ estão grifados em negrito.

A Tabela 2 mostra as variáveis que não apresentaram distribuição próxima da normal conforme o valor de assimetria apresentado na Tabela 1, e a respectiva transformação dos

dados para logaritmo natural ou raiz quadrada com objetivo de obter uma distribuição mais próxima da curva normal para as variáveis.

Tabela 2 – Variáveis predictoras, coeficientes de assimetria e transformação utilizada.

Variável predictora ⁽¹⁾	Transformação utilizada	
	Logaritmo natural	Raiz quadrada
CA	X	
PLAN		X
TWI	X	
CONV		X
REFL-IVP		X

⁽¹⁾ CA: área de contribuição; PLAN: curvatura planar; TWI: índice de umidade topográfica; CONV: convergência; REFL-IVP: refletância do infravermelho próximo.

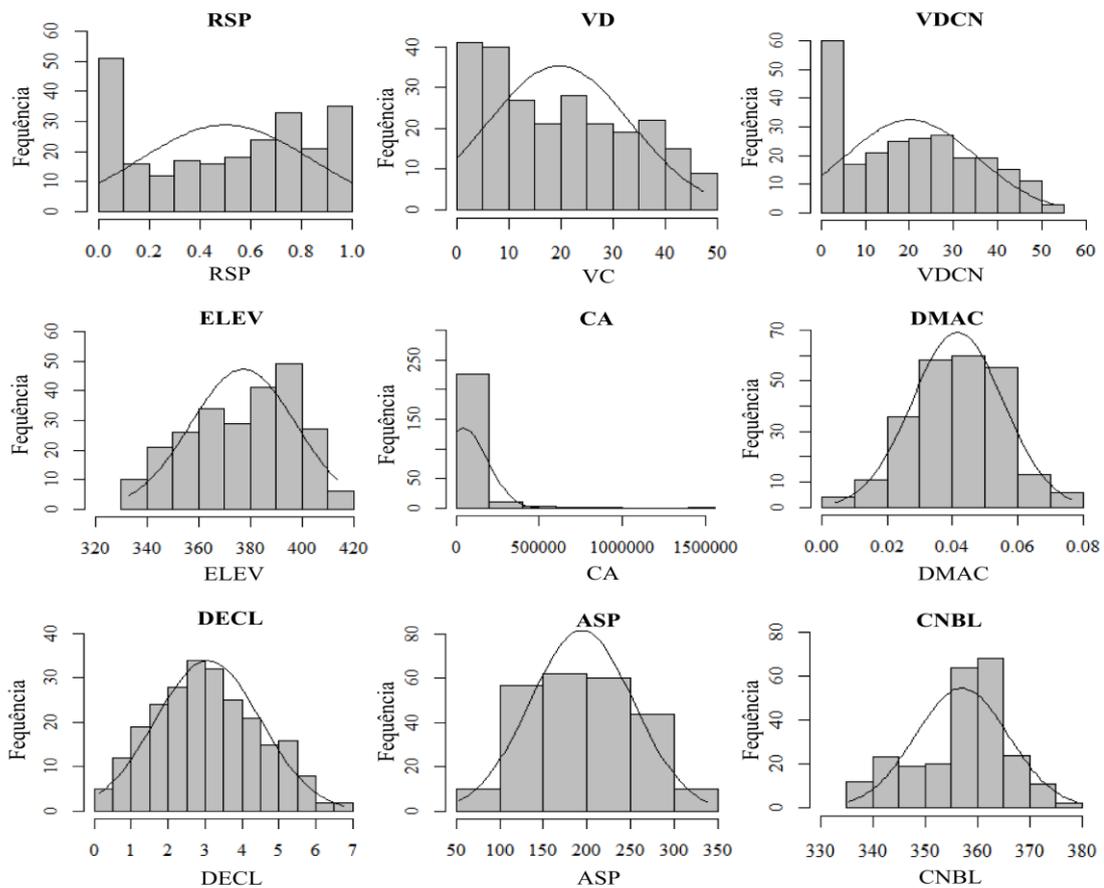


Figura 1 – Histogramas de frequência com a curva de distribuição normal das variáveis predictoras. Continua...

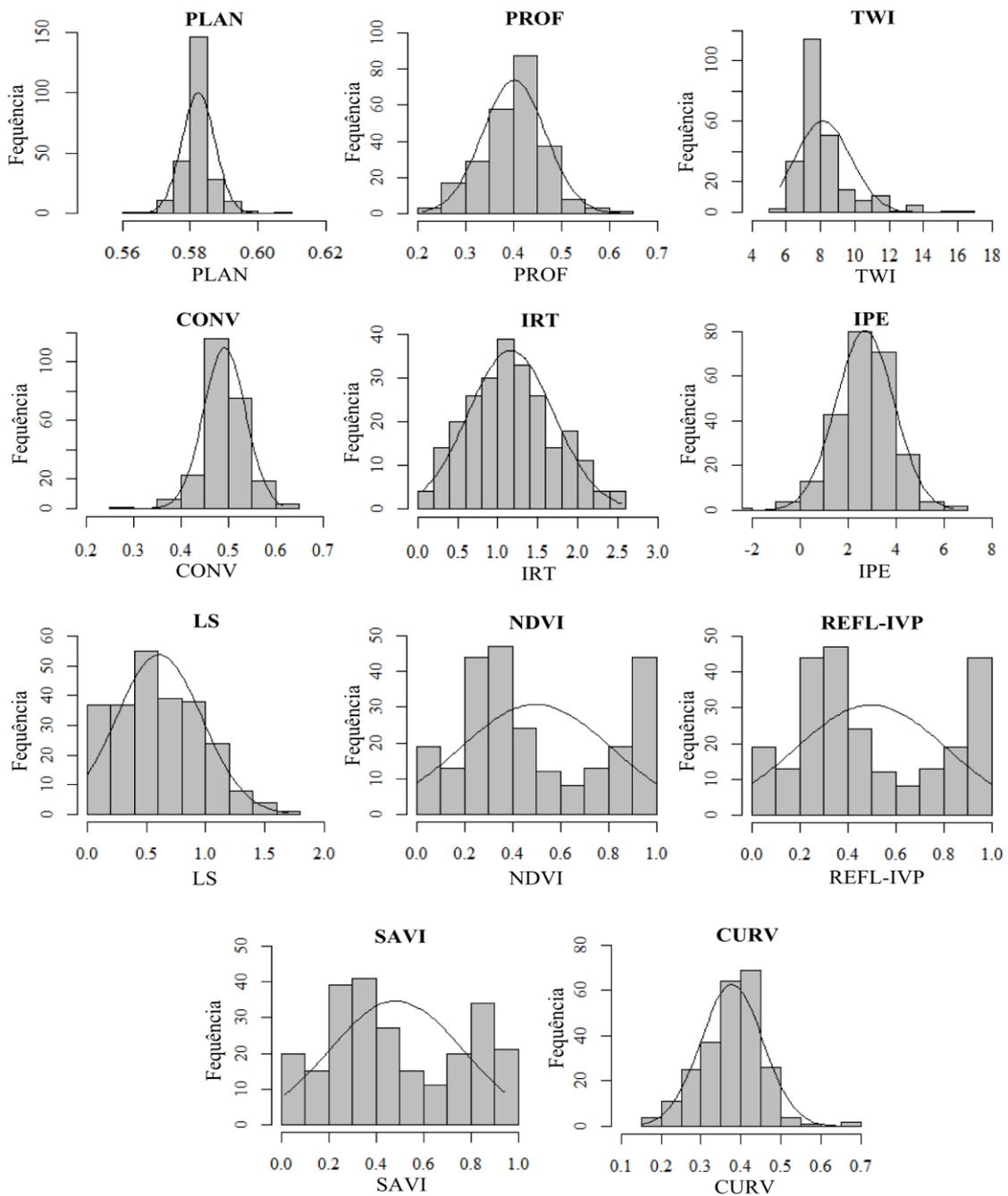


Figura 1 – Continuação.

Notação: RSP - posição relativa do declive; VD - profundidade de vale; VDCN - distância vertical da rede de drenagem; ELEV - elevação; CA - área de contribuição; DMAC - declividade média da área de contribuição; DECL - declividade; ASP - aspecto; CNBL - nível na base da rede de drenagem; PLAN - curvatura planar; PROF - curvatura de perfil; TWI - índice de umidade topográfica; CONV - convergência; TRI - índice de rugosidade do terreno; IPE - índice potencial de escoamento; LS - fator LS; NDVI - índice de vegetação por diferença normalizada REFL-IVP - refletância do infravermelho próximo; SAVI - índice de vegetação ajustado para o solo; CURV - curvatura.

A Figura 1 mostra os histogramas de frequência das 20 variáveis preditoras após a transformação dos dados exigida ou não para cada uma delas. Esse conjunto de variáveis foi extraído de cada plano de informação, na resolução espacial de 30 metros, em cada um dos 243 pontos amostrados na área de estudo. Os histogramas mostram que esse conjunto de variáveis representa bem a variabilidade dos atributos do terreno na área, uma vez que, grande parte das variáveis apresenta à curva de distribuição da variação dos atributos próximo a curva normal.

Os coeficientes da análise de correlação linear de Pearson aplicada nas variáveis preditoras com objetivo de identificar a multicolinearidade entre as variáveis são apresentados na Tabela 3. As variáveis com $r \geq 0,75$ foram eliminadas do conjunto de atributos do terreno utilizadas na construção da FPES. As variáveis que apresentaram alta correlação estão destacadas em negrito na Tabela 3. Observa-se que as variáveis com $r \geq 0,75$ foram: VD e RSP ($r = -0,98$), VDCN e RSP ($r = 0,95$), VDCN e VD ($r = -0,90$), ELEV e RSP ($r = 0,91$), ELEV e VD ($r = -0,90$), ELEV e VDCN ($r = 0,92$), CNBL e ELV ($r = 0,76$), CONV e PLAN ($r = 0,92$), LS e DECL ($r = 0,93$), LS e TRI ($r = 0,93$), SAVI e NDVI ($r = 0,99$) e CURV e PROF ($r = 0,86$). Com base nesses resultados algumas variáveis foram eliminadas conforme o grau de importância para a área de estudo e informações contempladas por outras variáveis. As variáveis eliminadas foram: VD, VDCN, ELEV, PLAN, LS, NDVI e PROF.

Tabela 3 – Coeficientes de correlação linear de Pearson entre as variáveis preditoras. Continua...

	RSP	VD	VDCN	ELEV	CA	DMAC	DECL	ASP	CNBL	PLAN
RSP	1,00									
VD	-0,98	1,00								
VDCN	0,95	-0,90	1,00							
ELEV	0,91	-0,90	0,92	1,00						
CA	-0,62	0,64	-0,56	-0,51	1,00					
DMAC	-0,67	0,61	-0,62	-0,60	0,34	1,00				
DECL	-0,25	0,20	-0,24	-0,21	-0,11	0,52	1,00			
ASP	-0,07	0,08	-0,07	-0,05	0,05	-0,07	0,03	1,00		
CNBL	0,50	-0,55	0,44	0,76	-0,24	-0,35	-0,09	0,00	1,00	
PLAN	0,32	-0,31	0,28	0,22	-0,63	-0,22	-0,18	0,02	0,04	1,00
PROF	0,49	-0,50	0,43	0,36	-0,46	-0,24	0,00	-0,02	0,11	0,38
TWI	-0,42	0,45	-0,37	-0,34	0,90	0,09	-0,43	0,00	-0,16	-0,56
CONV	0,30	-0,30	0,25	0,19	-0,65	-0,17	-0,08	0,02	0,02	0,92
TRI	-0,25	0,20	-0,24	-0,21	-0,10	0,52	1,00	0,03	-0,09	-0,18
IPE	-0,64	0,63	-0,58	-0,53	0,59	0,60	0,29	-0,08	-0,25	-0,35
LS	-0,40	0,35	-0,37	-0,31	0,17	0,58	0,93	0,03	-0,10	-0,34
NDVI	0,24	-0,27	0,10	0,09	-0,28	-0,35	-0,11	0,14	0,04	0,10
REFL-IVP	-0,14	0,12	-0,25	-0,28	-0,01	-0,04	0,05	0,05	-0,22	-0,05
SAVI	0,22	-0,25	0,08	0,06	-0,27	-0,33	-0,09	0,17	0,01	0,10
CURV	0,45	-0,45	0,39	0,30	-0,65	-0,20	0,00	0,01	0,04	0,65

Tabela 3 – Continuação.

	PROF	TWI	CONV	TRI	IPE	LS	NDVI	REFL-IVP	SAVI	CURV
RSP										
VD										
VDCN										
ELEV										
CA										
DMAC										
DECL										
ASP										
CNBL										
PLAN										
PROF	1,00									
TWI	-0,41	1,00								
CONV	0,42	-0,63	1,00							
TRI	-0,01	-0,43	-0,09	1,00						
IPE	-0,28	0,38	-0,34	0,29	1,00					
LS	-0,16	-0,15	-0,27	0,93	0,42	1,00				
NDVI	0,20	-0,20	0,14	-0,11	-0,40	-0,19	1,00			
REFL-IVP	0,04	-0,03	0,03	0,04	-0,01	0,03	0,71	1,00		
SAVI	0,19	-0,21	0,14	-0,10	-0,37	-0,18	0,99	0,74	1,00	
CURV	0,86	-0,62	0,71	0,00	-0,34	-0,25	0,20	0,04	0,19	1,00

⁽¹⁾ simbologia adotada para cada variável preditora seguida da unidade de medida; RSP: posição relativa do declive; VD: profundidade de vale; VDCN: distância vertical da rede de drenagem; ELEV: elevação; CA: área de contribuição; DMAC: declividade média da área de contribuição; DECL: declividade; ASP: aspecto; CNBL: nível na base da rede de drenagem; PLAN: curvatura planar; PROF: curvatura de perfil; TWI: índice de umidade topográfica; CONV: convergência; TRI: índice de rugosidade do terreno; IPE: índice potencial de escoamento; LS: fator LS; NDVI: índice de vegetação por diferença normalizada REFL-IVP: refletância do infravermelho próximo; SAVI: índice de vegetação ajustado para o solo; CURV: curvatura.

Outra técnica que foi utilizada para identificar a ocorrência de multicolinearidade entre as variáveis preditoras foi o teste estatístico de Bartlett. O resultado deste teste mostra que após a remoção de algumas variáveis preditoras pela análise de correlação linear de Pearson (Tabela 3) ainda existe correlação entre as variáveis ($p\text{-valor} = 0,00$) (Tabela 4).

Neste caso foi aplicada análise de componentes principais (ACP) com o objetivo de identificar as variáveis preditoras correlacionadas e assim reduzir o conjunto de variáveis preditoras para a construção das FPES. Essa técnica é comumente utilizada para seleção de variáveis do terreno em estudos de MDS como mostra os trabalhos de ten Caten et al. (2011a) e Samuel-Rosa et al. (2013).

Antes da aplicação da ACP o conjunto de dados foi submetido ao teste de adequação amostral KMO (Kaiser-Meyer-Olkin) e MAS (Measure of Sample Adequacy) para verificar se o conjunto de variáveis preditoras é adequado para a aplicação da ACP. Os resultados da Tabela 5 mostram que o conjunto de variáveis preditoras são correlacionadas, indicando que

são adequadas para aplicação da ACP, com exceção das variáveis ASP e REFL-IVP, a quais foram removidas do conjunto de dados.

Tabela 4 – Resultado do teste estatístico de esfericidade de Bartlett das variáveis preditoras selecionadas pra construção das FPES.

Estatística do teste χ^2	Valor p	Graus de liberdade
3535,72	0,00	78

Tabela 5 – Estatística do teste de adequação amostral KMO e MAS das variáveis preditoras selecionadas pra construção das FPES.

Teste de adequação amostral KMO													
											a) ¹ 0,71 (valor desejável entre 0,5 e 1,0)		
											b) ² 0,74 (valor desejável entre 0,5 e 1,0)		
Teste de adequação amostral MSA													
	DECL	ASP	SAVI	CNBL	RSP	TRI	DMAC	IPE	TWI	CA	CONV	REFL-IVP	CURV
a)	0,59	0,21	0,58	0,75	0,8	0,58	0,83	0,88	0,71	0,73	0,81	0,45	0,83
b)	0,59	-	0,59	0,59	0,59	0,59	0,59	0,59	0,59	0,59	0,59	-	0,83

⁽¹⁾ Teste estatístico realizado com as 13 variáveis preditoras; ⁽²⁾ Teste estatístico realizado após a remoção das variáveis preditoras ASP e REFL-IVP, pois apresentaram valor $MSA \leq 0,50$. DECL: declividade; ASP: aspecto; SAVI: índice de vegetação ajustado para o solo; CNBL: nível na base da rede de drenagem; RSP: posição relativa do declive; TRI: índice de rugosidade do terreno; DMAC: declividade média da área de contribuição; IPE: índice potencial de escoamento; TWI: índice de umidade topográfica; CA: área de contribuição; CONV: convergência; REFL-IVP: refletância do infravermelho próximo; CURV: curvatura.

Após a verificação e adequação do conjunto das variáveis preditoras, foi aplicada a ACP. Na Figura 2 é apresentada a relação entre os autovalores de cada componente principal para o conjunto das 11 variáveis preditoras. Apenas três componentes apresentam autovalor maior que 1, os quais explicam 76% da variância presente nos dados, como pode ser observado na Tabela 6. Nestas componentes, as variáveis que apresentaram pesos mais elevados foram a AC, DECL e CNBL. No entanto, os pares de variáveis DECL e TRI, TWI e AC, apresentaram pesos elevados na mesma componente principal (Tabela 7), indicando que esses pares são correlacionados, podendo influenciar na construção das FPES. Para minimizar esse problema, os atributos TRI e AC foram removidos do conjunto de dados.

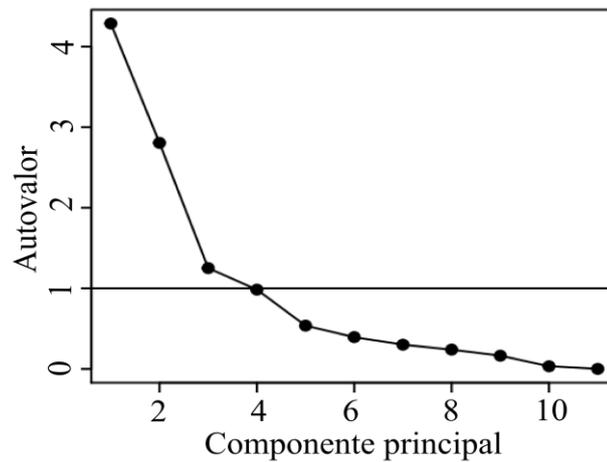


Figura 2 – Autovalor correspondente a cada componente principal obtida a partir da matriz de correlação das 11 variáveis predictoras.

Tabela 6 – Autovalores e proporção da variância explicada por cada componente principal.

Componente principal	Autovalor	Proporção da variância explicada	Proporção cumulativa da variância explicada
1	2,071	0,390	0,390
2	1,675	0,255	0,645
3	1,119	0,114	0,759
4	0,992	0,089	0,848
5	0,733	0,049	0,897
6	0,628	0,036	0,933
7	0,549	0,027	0,960
8	0,489	0,022	0,982
9	0,406	0,015	0,997
10	0,184	0,003	1,000
11	0,031	0,000	1,000

A projeção da correlação entre as variáveis predictoras e os escores obtidos na análise de componentes principais é apresentado na Figura 3. Na projeção da primeira componente com a segunda, observa-se a formação de cinco grupos, sendo o primeiro formado pelos atributos CURV e CONV que aparece na primeira e segunda projeção. Da mesma forma, ocorre com os atributos DECL e TRI, formando o segundo grupo. O terceiro grupo é formado pelos atributos DMAC e IPE. O quarto agrupamento é formado pelos atributos RSP, SAVI e CNBL, onde os dois últimos aparecem bem próximos no primeiro círculo, porém o atributo CNBL aparece bem destacado dos demais no segundo círculo. O quinto grupo é formado pelos atributos TWI e AC.

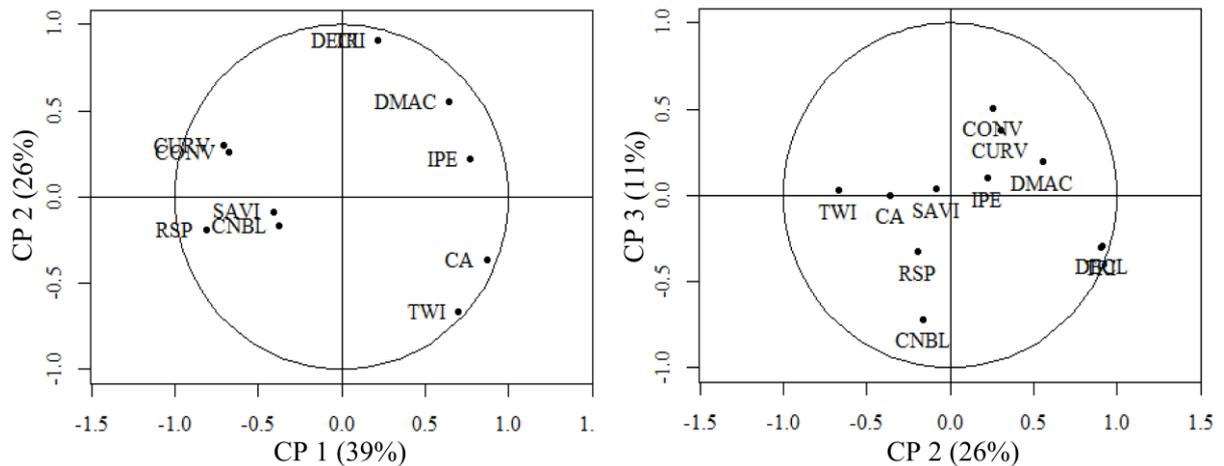


Figura 3 – Projeção da correlação entre as 11 variáveis predictoras e os escores das componentes principais na primeira (CP 1), segunda (CP 2) e terceira (CP 3) dimensões.

Notação: RSP - posição relativa do declive; CA - área de contribuição; DMAC - declividade média da área de contribuição; DECL - declividade; CNBL - nível na base da rede de drenagem; TWI - índice de umidade topográfica; CONV - convergência; TRI - índice de rugosidade do terreno; IPE - índice potencial de escoamento; SAVI - índice de vegetação ajustado para o solo; CURV - curvatura.

Com base nos resultados apresentados na Tabela 7 e Figura 3, foram removidos do conjunto de variáveis predictoras os seguintes atributos do terreno: TRI, CA, CURV e IPE. Na eliminação de cada atributo em detrimento de outro foi considerada sua importância para a predição do COS na área de estudo. As variáveis predictoras selecionadas para a construção das FPES foram: DECL, DMAC, CNBL, RSP, CONV, TWI e SAVI.

Tabela 7 – Pesos das 11 variáveis predictoras em cada uma das componentes principais.

Variável predictoradora ⁽¹⁾	Componente principal										
	1	2	3	4	5	6	7	8	9	10	11
DECL	0,10	0,54	-0,26	-0,13	-0,08	0,12	-0,04	-0,23	0,12	-0,15	0,71
SAVI	-0,20	-0,05	0,03	-0,84	0,48	-0,01	0,05	0,10	0,09	-0,01	-0,01
CNBL	-0,18	-0,10	-0,64	0,33	0,57	-0,25	0,12	-0,16	-0,11	-0,02	0,00
RSP	-0,39	-0,12	-0,29	0,11	-0,16	0,42	0,16	0,37	0,61	0,05	0,00
TRI	0,11	0,54	-0,27	-0,13	-0,08	0,13	-0,05	-0,25	0,12	-0,09	-0,71
DMAC	0,31	0,33	0,17	0,11	0,16	-0,34	0,64	0,40	0,20	0,01	-0,01
IPE	0,37	0,13	0,09	0,21	0,51	0,41	-0,46	0,40	0,00	-0,05	0,00
TWI	0,34	-0,40	0,02	-0,02	0,04	0,15	0,20	-0,26	0,23	-0,74	-0,03
CA	0,42	-0,22	0,00	-0,03	0,13	0,20	0,15	-0,43	0,31	0,64	0,03
CONV	-0,33	0,15	0,46	0,23	0,24	-0,30	-0,28	-0,30	0,53	-0,08	0,00
CURV	-0,34	0,18	0,34	0,17	0,23	0,54	0,44	-0,24	-0,33	-0,02	0,00

⁽¹⁾ DECL: declividade; SAVI: índice de vegetação ajustado para o solo; CNBL: nível na base da rede de drenagem; RSP: posição relativa do declive; TRI: índice de rugosidade do terreno; DMAC: declividade média da área de contribuição; IPE: índice potencial de escoamento; TWI: índice de umidade topográfica; CA: área de contribuição; CONV: convergência; CURV: curvatura.

Na Tabela 8 é apresentada a correlação linear de Pearson entre as variáveis preditoras pré-selecionadas para a construção das FPES para o COS. As correlações foram baixas para todos os atributos, sendo observado os maiores valores para as variáveis RSP, CNBL e SAVI.

Tabela 8 – Coeficientes de correlação entre as variáveis preditoras e o teor de carbono orgânico do solo (COS) na camada de 0 – 5 cm.

Variáveis preditoras ⁽¹⁾	COS (0 – 5 cm)
DECL	-0,04 (p= 0,61) ⁽²⁾
DMAC	0,05 (p = 0,51)
CONV	-0,04 (p = 0,57)
CNBL	-0,21 (p = 0,01)
RPS	-0,29 (p = 0,00)
TWI	0,12 (p = 0,13)
SAVI	-0,19 (p = 0,01)

⁽¹⁾ DECL: declividade; DMAC: declividade média da área de contribuição; CONV: convergência; CNBL: nível na base da rede de drenagem; RSP: posição relativa do declive; TWI: índice de umidade topográfica; SAVI: índice de vegetação ajustado para o solo; ⁽²⁾ p = probabilidade a 5% de significância (n = 243).

Na Tabela 3 é apresentada a análise descritiva dos teores de carbono orgânico do solo (COS) para os conjuntos de calibração e validação. Como os dados não apresentam distribuição normal conforme o valor de assimetria, esse foram transformados para logaritmo do valor de COS. Observa-se a semelhança entre os conjuntos de dados de calibração e validação conforme os valores de média, mediana, desvio padrão e variância amostral. Isso indicada que os conjuntos de validação são representativos do conjunto de calibração.

Tabela 9 – Estatísticas descritivas do COS para os conjuntos de calibração e validação.

	COS – C ⁽¹⁾	COS - V
	(g kg ⁻¹)	
Média	29,65	29,70
Erro padrão	0,61	0,82
Mediana	28,35	29,21
Desvio Padrão	7,99	7,00
Variância	63,97	49,08
Assimetria	1,27	0,66
Mínimo	14,21	17,4
Máximo	60,28	50,63

⁽¹⁾ COS – C: conjunto de calibração; COS – V: conjunto de validação.

A FPES para o COS na camada de 0 – 5 cm, construída a partir das variáveis preditoras pré-selecionadas, utilizando como modelo a RLM em conjunto com o método stepwise, para o conjunto dos 170 pontos foi:

$$\ln(COS) = 62,323924 + 3,842303*SAVI - 13,925191*RSP - 227,325020*DMAC + 7,991431*TWI$$

onde SAVI é o índice de vegetação ajustado para o solo, RSP é a posição relativa do declive, DMAC é a declividade média da área de contribuição e TWI é o índice de umidade topográfica.

Na Tabela 10 são apresentados os coeficientes de FPES, assim como o valor do R^2 , R^2_{aj} e valor de AIC. A função consegue explicar apenas 18 % (R^2 : 0,18) da variância do COS na área de estudo. Ainda, o valor de AIC neste estudo é considerado elevado. Os coeficientes da função para cada variável preditora indicam que o teor de COS diminui com o aumento do RSP e DMAC. O inverso ocorre para o SAVI e TWI, onde valores maiores proporcionam o aumento do COS. Isso está de acordo com a distribuição do COS na área de estudo, sendo observado os maiores teores de COS nas áreas de baixada, onde o relevo é mais plano e os valores de RSP e DMAC são menores e o TWI é mais elevado. Essas áreas são consideradas locais de deposição de sedimentos oriundos do processo erosivo, e conseqüentemente, áreas de acúmulo de COS. As áreas com maior SAVI são locais de maior aporte de resíduos vegetais, como as áreas de floresta nativa e campo nativo, onde os valores de COS também foram maiores.

Tabela 10 – Coeficientes da função de predição espacial do carbono orgânico para área de estudo.

ln (cos)	Estimativa	Erro padrão	Valor t	Pr(> t)
Intercepto	62,3239	10,456	5,960	1,480e-08 *** ⁽²⁾
SAVI ⁽¹⁾	3,8423	1,241	3,097	2,300e-03 **
RSP	-13,9251	2,610	-5,336	3,100e-07 ***
DMAC	-227,3250	58,019	-3,918	1,300e-04 ***
TWI	7,9914	4,096	1,951	5,277e-02 "

R^2 : 0,18; R^2_{aj} : 0,16
 Critério de Informação de Akaike (AIC): 681,8646

⁽¹⁾ SAVI: índice de vegetação ajustado para o solo; RSP: posição relativa do declive; DMAC: declividade média da área de contribuição; TWI: índice de umidade topográfica; ⁽²⁾ *** 0,000; ** 0,001; " 0,05.

A Tabela 11 mostra os resultados de análise da variância, onde a partir dos valores da soma dos quadrados (SQ) pode-se calcular a contribuição de cada variável preditora na variância dos resultados por meio da explicação da variância (EV%). Neste estudo, a EV das variáveis SAVI, RSP, DMAC e TWI foram 3%, 7%, 6% e 2%, respectivamente. Apesar dos valores baixos de EV, atributos relacionados ao processo erosivo, como o RSP e DMAC apresentaram os valores maiores de EV em relação aos demais. A baixa contribuição do SAVI pode estar relacionada com resolução pequena da imagem LANDSAT-7 para a área de estudo. Esses resultados mostram que apenas o uso dos atributos do terreno e índices de vegetação não foi suficiente para estabelecer uma alta correlação com o conteúdo de COS para a construção de uma FPES com potencial para a predição do COS na área de estudo.

Tabela 11 – Análise da variância da função de predição espacial do COS para área de estudo.

Fonte de variação	Graus de liberdade	Soma dos Quadrados	Quadrado médio	F	Pr(>F)
Regressão	4	1.963,8	490,9	9,155	1,052e-06
SAVI ⁽¹⁾	1	372,4	372,4	6,944	0,009208** ⁽²⁾
RSP	1	740,1	740,1	13,802	0,000278***
DMAC	1	647,2	647,2	12,069	0,000655***
TWI	1	204,1	204,1	3,806	0,052770 ^{''}
Resíduos	165	8848,0	53,6	-	-
Total	169	10.811,8	63,9	-	-

⁽¹⁾ SAVI: índice de vegetação ajustado para o solo; RSP: posição relativa do declive; DMAC: declividade média da área de contribuição; TWI: índice de umidade topográfica; ⁽²⁾ *** 0,000 ; ** 0,001; ^{''} 0,05.

A análise da Figura 4 a, b e d permite identificar varias observações com valores de $resi > 2$, $resip > 1,96$ e valores de $alav > 0,09$ ($3 \times 5 / 170 = 0,09$), respectivamente. As observações que apresentaram os maiores valores de $resi$ foram 51, 60, 151, 121, 92 e 73, onde o valor de COS predito foi superestimado nestas observações ($resi > 0$). Da mesma forma, o mesmo conjunto de observações aparece nos $resip$ na Figura 4b. Verifica-se também que várias observações foram subestimadas ($resi < 0$), com valores de $resi > -10$ (Figura 4a).

Segundo McDonald (2002), para que uma observação tenha influencia no ajuste do modelo de predição quando $n > 15$ (mais de 15 amostras no conjunto de dados) e $p > 3$ (mais de três variáveis preditoras) o valor de $dist_c$ deve ser $> 0,85$. A Figura 4c mostra que as observações com maiores valores de $dist_c$ foram 51, 60, 124 e 15, no entanto, nenhuma

observação apresentou valor de $dist_c > 0,85$, portanto, neste estudo não foi removida nenhuma observação durante o ajuste do modelo.

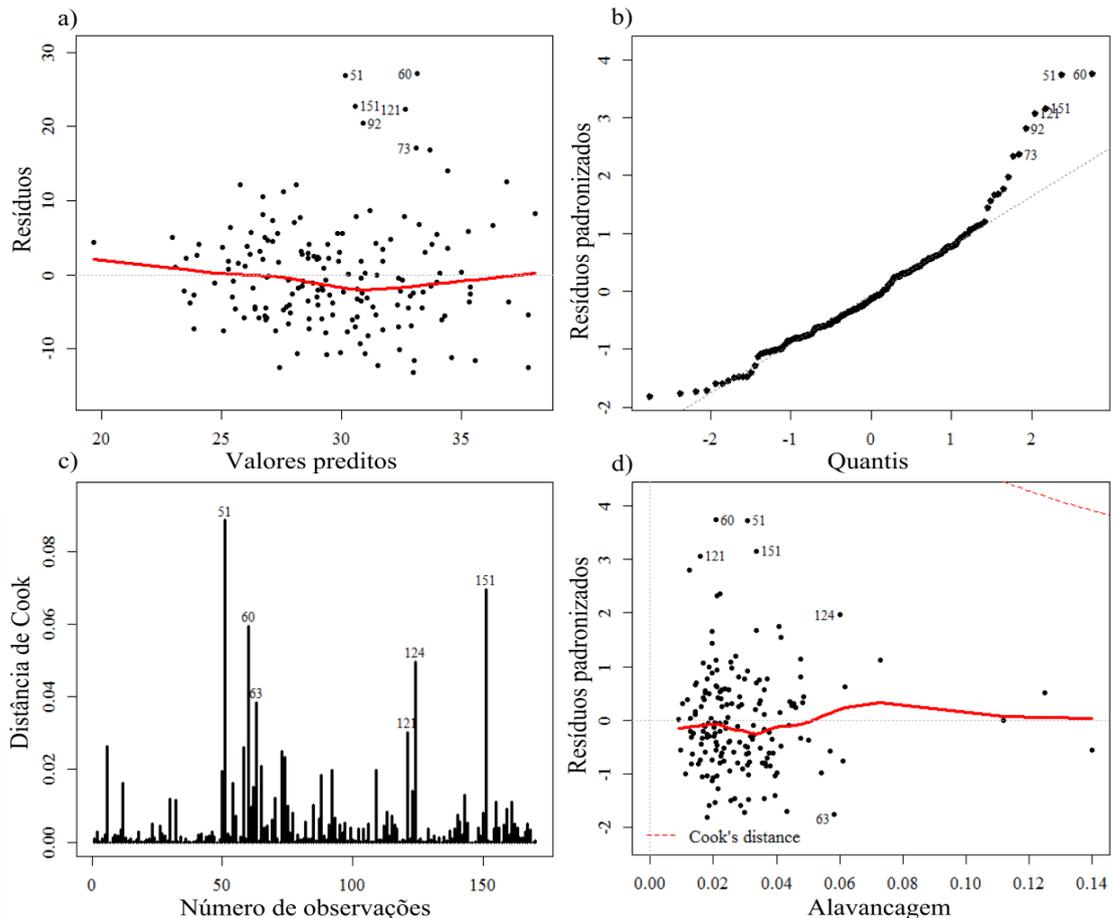


Figura 4 – Análise dos resíduos (*resi*) do ajuste da FPES para o COS: a) resíduos dos valores preditos; b) resíduos preditos padronizados; c) distâncias de Cook associada a cada observação; d) alavancagem.

Os valores de *resi* elevados nas observações 51 e 60 podem estar relacionados com sua posição na paisagem, sendo encontradas em áreas mal drenadas, onde predomina solos mal drenados (Gleissolos Háplicos) e a vegetação é campo nativo. Já as observações 73 e 92 foram amostras em área de floresta nativa sobre Latossolo Vermelho (LV). E as observações 121 e 151 são representativas de áreas de lavoura sobre LV. Vale ressaltar que as informações classe de solo e tipo de uso da terra não foram inseridas como variáveis preditoras na FPES para o COS, podendo isso ter influenciado nestes resultados. Uma alternativa para melhorar a

FPES seria separar o conjunto de observações por tipo de solo e uso da terra, no entanto, isso não foi realizado devido ao tamanho do conjunto de observações, apenas 170, disponíveis para construção e ajuste da FPES neste estudo.

Outra forma seria inserir essas informações como variáveis preditoras na construção da FPES para o COS como realizado no trabalho de Mendonça-Santos et al. (2007). Neste trabalho os autores construíram modelos de predição com bom desempenho para estoque de carbono no estado do Rio de Janeiro, utilizando variáveis preditoras que representavam os fatores: relevo, organismos, material de origem e o solo. Porém, esta estratégia também não foi realizada no presente estudo, uma vez que, este estudo buscou desenvolver uma FPES para COS utilizando apenas variáveis preditoras relacionadas com o terreno (fator relevo) e índices de vegetação (fator organismos).

Na Tabela 12 é apresentado os valores dos parâmetros estatísticos para os conjuntos de calibração e validação. Valores de EM positivos indicam superestimação dos teores de COS para os dois conjuntos. Já para o conjunto de validação o valor do EM foi negativo, indicando subestimação dos teores de COS. Os valores de RMSE são considerados altos, pois esses valores de erro de predição em escala de propriedade rural são significativos. Os valores de RPD são baixos ($< 1,4$), mostrando que a FPES para os COS não é confiável.

Tabela 12 – Estatísticas da calibração e validação da função de predição espacial do COS para área de estudo.

Conjunto de dados ⁽¹⁾	Parâmetros estatísticos ⁽²⁾			
	EM (g kg ⁻¹)	RMSE (g kg ⁻¹)	RPD	R ² _{aj}
Calibração	5,88	7,21	1,10	0,16
Validação	-0,45	6,07	1,11	0,19

⁽¹⁾ conjunto de calibração (n = 170) e conjunto de validação (n = 73); ⁽²⁾ EM: erro médio; RMSE: raiz quadrada do erro quadrático médio; RMSE_n: raiz quadrada do erro quadrático médio normalizado; RPD: relação do desempenho do desvio; R²_{aj}: coeficiente de determinação ajustado.

Na Figura 5 são apresentados os teores de COS preditos confrontado com os valores medidos de COS para o conjunto de calibração e validação. Observa-se uma grande dispersão dos pontos em relação a linha 1:1, principalmente para o conjunto de calibração quando os valores de COS medidos são maiores a 40 g kg⁻¹.

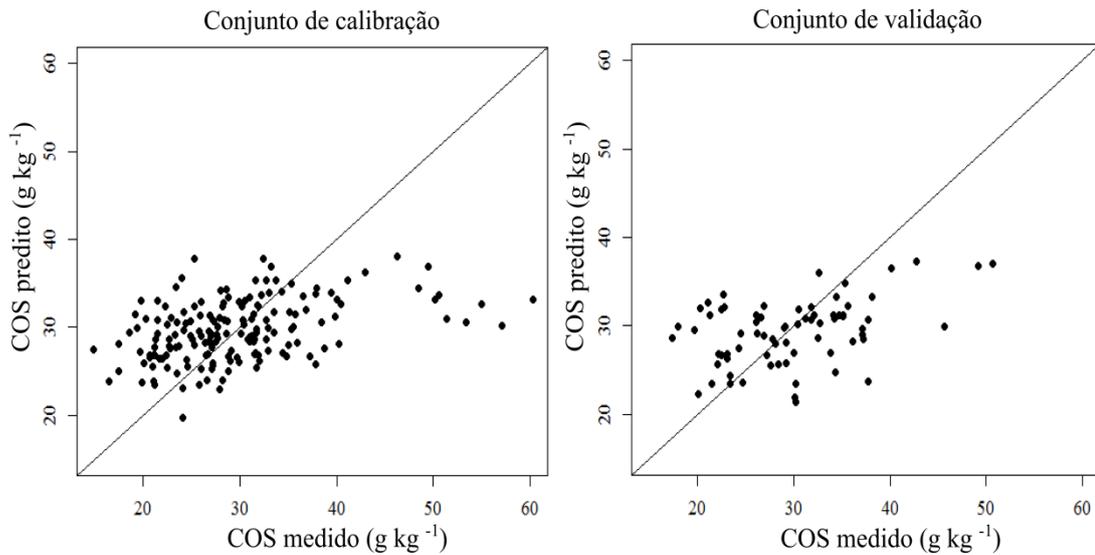


Figura 5 – Teor de carbono orgânico do solo (COS) predito versus medido para os conjuntos de calibração (n = 170) e validação (n = 73).

O valor de R^2 encontrado neste estudo é baixo comparado a outros trabalhos encontrados na literatura. Razakamanarivo et al. (2011) uma área de 1.590 ha encontraram valores de $R^2 = 0,61$ nas predições do COS, porém utilizando em conjunto atributos do terreno e de vegetação. Kempen et al. (2011) encontraram para uma área de 152.500 ha, valores de $EM=4,5$, $RMSE=9,9$ e $R^2=0,75$ em predições de COS usando co-krigagem, e como variáveis, utilizou atributos do MDE, mapa de classes de solo, mapa de uso da terra, mapa de geomorfologia. Kheir et al. (2010) utilizando e variáveis do terreno, sensoriamento remoto, tipo de solo e geologia, para modelar o COS por meio da técnica árvore de decisão, constaram 69,5% a acerto comparando os mapas preditos versus observações de campo.

No entanto, alguns trabalhos também têm encontrados valores baixos de ajuste do R^2 de modelos de predição do COS. Wheeler et al. (2012) encontraram $R^2 = 0,18$ mapeando o estoque de carbono em escala de propriedade rural, utilizando como covariáveis atributos do solo, tipos de uso da terra e atributos do terreno. Da mesma forma, Pei et al. (2010) utilizaram o método de co-krigagem e como variáveis predito o índice de umidade topográfica para realizar a modelagem de COS, atingindo valores de R^2 de 0,23.

O motivo dos baixos de R^2 encontrados neste trabalho pode estar relacionado principalmente ao uso apenas de atributos do terreno e índices de vegetação, uma vez que, essas variáveis preditoras não foram capazes de identificar a variações no conteúdo de COS

na paisagem, consequência da área de estudo ser relativamente pequena e homogênea em termos de relevo, onde nesta área, as variações no COS na área estariam mais relacionadas a variáveis práticas de manejo do solo, tipo de uso da terra e tipo de solo.

A utilização de um MDE de maior resolução espacial e associado com informações espectrais do solo e mapas de cobertura da terra com alta resolução poderia ser uma alternativa para melhorar a predição de COS em áreas pequenas e com características homogêneas. Além disso, a distribuição do COS na paisagem é influenciada por uma gama de fatores, onde a associação de um conjunto maior de variáveis preditoras podem melhorar as predições de COS em áreas pequenas e homogêneas, no entanto, a complexidade das funções de predição torna-se maior. Ainda, o uso de técnicas geoestatísticas como co-krigagem e regressão de krigagem podem apresentar desempenho melhor nessas áreas, no entanto, necessitam de uma malha de amostragem densa que represente a variabilidade do COS na área.

6.4 Conclusões

A função de predição espacial construída a partir dos atributos do terreno e índices de vegetação não apresentou desempenho satisfatório conforme o ajuste do modelo e o valor de R^2 .

As variáveis preditoras utilizadas conseguiram explicar apenas 16% da variância do COS na área de estudo. Esses resultados estão relacionados ao tamanho da área e as características do terreno, onde os atributos MDE-LP não conseguiu identificar diferenças no terreno relacionadas com variabilidade COS para essa área.

Para melhorar a predição desses modelos é necessária a inclusão de variáveis relacionadas ao manejo do solo e mapas de cobertura do solo. O uso da técnica de espectroscopia de reflectância difusa pode ser uma alternativa para melhorar a predição do COS em escala de propriedade rural nessas áreas.

7 CONSIDERAÇÕES FINAIS

Os modelos digitais de elevação globais disponíveis gratuitamente apresentaram diferenças quanto à qualidade altimétrica e atributos do terreno para área de estudo. O MDE-LP apresentou o melhor desempenho entre os MDE. Isso confirma a necessidade de avaliação dessas informações antes de utilizá-las no Mapeamento Digital de Solos, visando minimizar as incertezas e a propagação de erros.

Os MDE disponíveis gratuitamente globais podem apresentar potencial para uso no MDS de classes de solos em áreas de média e pequena extensão territorial mediante um diagnóstico da qualidade dos dados altimétricos com base em dados de controle de campo (*in loco*) e definição da acurácia de cada MDE em relação a escala de trabalho.

O uso apenas das variáveis preditoras atributos do terreno e índices de vegetação derivados de MDE com resolução espacial pequena, não possibilitaram estabelecer correlações com desempenho satisfatório para um bom ajuste de FPES para explicar a variação do conteúdo de carbono orgânico do solo na área de estudo.

Sugere-se em áreas pequenas e com características homogêneas de terreno o uso de outros métodos estatísticos para construção de funções de predição espacial, MDE com maior resolução espacial e um conjunto maior de variáveis ambientais preditoras, principalmente relacionadas ao uso da terra e práticas de manejo.

8 REFERÊNCIAS

ALVARES, C.A.; STAPE, J.L.; SENTELHAS, P.C.; GONÇALVES, J.L.M.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v.22, n.6, 711-728, 2013.

ARCHELA, R.; ARCHELA, E. Síntese cronológica da cartografia no Brasil. **Portal da Cartografia**, Londrina, 2008. p.93-110.

ARRUDA, G.P.; DEMATTÊ, J.A.M; CHAGAS, C.S. Mapeamento digital de solos por redes neurais artificiais com base na relação solo-paisagem. **Revista Brasileira de Ciência do Solo**, v.37, p.327-338, 2013.

BALLABIO, C.; FAVA, F.; ROSENMUND, A. A plant ecology approach to digital soil mapping, improving the prediction of soil organic carbon content in alpine grasslands. **Geoderma**, v.187-188, p.102-116, 2012.

BASHER, L.R. Is pedology dead and buried? **Australian Journal of Soil Research**, v.35, p.979-994, 1997.

BATJES, N. H.; SOMBROEK, W. G. Possibilities for carbon sequestration in tropical and subtropical soils. **Global Change Biology**. v.3, n.2, p.161-173, 1997.

BAJTES, N.H. Effects of mapped variation in soil conditions on estimates of soil carbon and nitrogen stocks for South America. **Geoderma**, v.97, p.135-144, 2000.

BEHRENS, T.; SCHMIDT, K.; SCHOLTEN, T. An approach to removing uncertainties in nominal environmental covariates and soil class maps. In: HARTEMINK, A.; MCBRATNEY, A.; MENDOÇA-SANTOS, M.L. (Eds.), **Digital Soil Mapping with Limited Data**. Springer. 2008. p.213-224.

BEHRENS, T.; ZHU, A.; SCHMIDT, K.; SCHOLTEN, T. Multi-scale digital terrain analysis and features selection for digital soil mapping. **Geoderma**, v.155, p.175-185, 2010.

BEHRENS, T.; SCHMIDT, K.; RAMIREZ-LOPEZ, L.; GALLANT, J.; ZHU, A-X.; SCHOLTEN, T. Hyper-scale digital soil mapping and soil formation analysis. **Geoderma**, v. 213, p.578-588, 2014.

BEN-DOR, E.; HELLER, D.; CHUDNOVSKY, A. A novel method of classifying soil profiles in the field using optical means. **Soil Science Society American Journal**, v.72, p.1113-1123, 2008.

BERNOUX, M.; CARVALHO, M.C.S.; VOLKOFF, B.; CERRI, C.C. Brazil's soil carbon stocks. **Soil Science of America Journal**, Madison, v.66, p.888-896, 2002.

BERNOUX, M.; ARROUAYS, D.; CERRI, C.; CERRI, C. Regional organic carbon storage maps of the western brazilian amazon based on prior soil maps and geostatistical interpolation. In: LAGACHERIE, P.; MCBRATNEY, A.B.; VOLTZ, M. (Eds.). **Digital soil mapping - an introductory perspective**. Amsterdam: Elsevier, 2006. p.497-506.

BISHOP, T.F.A.; MINASNY, B.; MCBRATNEY, A.B. Uncertainty analysis for soil-terrain models. **International Journal of Geographical Information Science**, v.20, n.2, p.117-134, 2006.

BRASIL. **Especificação Técnica para a Aquisição de Dados Geoespaciais Vetoriais (ET-ADGV) Versão 2.1.3**. Diretoria de Serviço Geográfico do Exército Brasileiro. 2011. 254p. Disponível em <<http://www.geoportal.eb.mil.br/index.php/inde2?id=140>>. Acesso em: 20 jan. 2014.

BREIMER F.R.; van KEKEN, A.J.; van REULER, H. Guidelines for soil survey and land evaluation in ecological research. Paris: **UNESCO**; 1986. 124p.

BRUS, D.; DE GRUIJTER, J. Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). **Geoderma**, v.80, p.1-44, 1997.

CAMBULE, A.H.; ROSSITER, D.G.; STOORVOGEL, J.J. A methodology for digital soil mapping in poorly-accessible areas. **Geoderma**, v.192, p.341-353, 2013.

CARRÉ, F.; MCBRATNEY, A.B.; MAYR, T.; MONTANARELLA, L. Digital soil assessments: beyond DSM. **Geoderma**, v.142, p.69-79, 2007.

CARVALHO JÚNIOR, W.; CHAGAS, C.S.; MUSELLI, A.; PINHEIRO, H.S.K.; PEREIRA, N.R.; BHERING, S.B. Método do hipercubo latino condicionado para a amostragem de solos na presença de covariáveis ambientais visando o mapeamento digital de solos. **Revista Brasileira de Ciência do Solo**, v.38, p.386-396, 2014.

CARVALHO, T.M.; CARVALHO, C.M. Sistemas de informações geográficas aplicadas à descrição de habitats. **Acta Scientiarum**. Human and Social Sciences Maringá, v.34, n.1, p.79-90, 2012.

CAVAZZI, S.; CORSTANJE, R.; MAYR, T.; HANNAM, J.; FEALY, R. Are fine resolution digital elevation models always the best choice in digital soil mapping? **Geoderma**, v.195–196, p.111–121, 2013.

CHAGAS, C.S.; FERNANDES FILHO, E.; ROCHA, M.F.; CARVALHO JÚNIOR, W.; SOUZA NETO, N.C. de. Avaliação de modelos digitais de elevação para aplicação em um mapeamento digital de solos. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.14, p.218-226, 2010.

CHANG, C.; LAIRD, D.A.; MAUSBACH, M.J.; HURBURG JUNIOR, C.R. Near infrared reflectance spectroscopy – principal components regression analyses of soil properties. **Soil Science Society of American Journal**, v.25, n.2, p.480-490, 2001.

CHAPLOT, V., DARBOUX, F., BOURENNANE, H., LEGUÉDOIS, S., SILVERA, N., PHACHOMPON, K. Accuracy of interpolation techniques for the derivation of digital elevation models in relation to landform types and data density. **Geomorphology**, v.77, p.126-141, 2006.

CHENG-ZHI, Q.; A-XING, Z.; WEI-LI, Q.; YAN-JUN, L.; BAO-LIN, L.; PEI, T. Mapping soil organic matter in small low-relief catchments using fuzzy slope position information. **Geoderma**, v.171-172, p. 64-74, 2012.

CHERUBIN, M.R.; SANTI, A.L.; EITELWEIN, M.T. MENEGOL, D.R.; DA ROS, C.O.; PIAS, O.H. de C.; BERGHETTI, J. Eficiência de malhas amostrais utilizadas na caracterização da variabilidade espacial de fósforo e potássio. **Ciência Rural**, v.44, p.425-432, 2014.

CRIVELENTI, R.C.; COELHO, R.M.; ADAMI, S.F.; OLIVEIRA, S.R.M. Mineração de dados para inferência de relações solo-paisagem em mapeamentos digitais de solo. **Pesquisa Agropecuária Brasileira, Brasília**, v.44, n.12, p.1707-1715, 2009.

CROFT, H.; KUHN, N. J.; ANDERSON, K. On the use of remote sensing techniques for monitoring spatio-temporal soil organic carbon dynamics in agricultural systems. **Catena**, v.94, p.64-74, 2012.

DALMOLIN, R. S. D.; KLAMT, E.; PEDRON, F. A.; AZEVEDO, A. C. Relação entre as características e o uso das informações de levantamentos de solos de diferentes escalas. **Ciência Rural**, v.34, p.1479-1486, 2004.

DALMOLIN, R.S.D.; GONCALVES, C.N.; KLAMT, E.; DICK, D.P. Relação entre os constituintes do solo e seu comportamento espectral. **Ciência Rural**, v.35, p.481-489, 2005.

DIXON, T.H. **SAR Interferometry and surface change detection**. Miami: University of Miami, Rosenstiel School of Marine and Atmospheric Sciences, RASMAS TR 95-003, 1995. 97p.

DOBOS, E.; MICHELI, E.; BAUMGARDNER, M.F.; BIEHL, L.; HELT, T. Use of combined digital elevation model and satellite data regional soil mapping. **Geoderma**, v.97, p.367-391, 2000.

DOTTO, A.C. **Funções de pedotransferência do solo: estimativa por radiometria**. 2012. 68f. Dissertação (Mestrado em Ciência do Solo) - Universidade Federal de Santa Maria, Santa Maria, RS. 2012.

DRUZINA, A.G.S. **Integração de dados altimétricos obtidos através de diferentes técnicas para geração de um novo modelo digital de elevação**. 2007. 69f. Dissertação (Mestrado em Sensoriamento Remoto). Universidade Federal do Rio Grande do Sul, Porto Alegre, RS. 2007.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA - EMBRAPA. Centro Nacional de Pesquisa de Solo. **Procedimentos Normativos de Levantamentos Pedológicos**. Brasília, Embrapa, 1995.

ESWARAN, H.; VAN DEN BERG, E.; REICH, P.; KIMBLE, J. Global soil carbon resources. In: R. Lal et al. (ed.) **Soils and global change**. Advances in Soil Science, Boca Raton, 1995. p. 27-43.

FELICÍSIMO, A. M. **Modelos digitais del terreno: Introducción aplicaciones en las ciencias ambientales**. Oviedo: Universidad de Oviedo, 1994. Disponível em <<http://www.etsimo.uniovi.es/~feli>>. Acesso em: 20 jan. 2014.

FIGUEIREDO, S.R. et al. Uso de regressões logísticas múltiplas para mapeamento digital de solos no Planalto Médio do RS. **Revista Brasileira de Ciência do Solo**, v.32, p.2779-2785, 2008.

FLORES, C.A.; PÖTER, R.O.; FASOLO, P.J.; HASENACK, H.; WEBER, E. **Mapa semidetalhado de solos – Folha Palomas**. Porto Alegre: Editora da Universidade, 2007. 1 mapa, 1 CD-ROM.

FLORINSKY, I.V.; KURYAKOVA, G.A. Determination of grid size for digital terrain modeling in landscape investigations-exemplified by soil moisture distribution at a micro-scale. **International Journal of Geographical Information Science**, v.14, p.815-832, 2000.

GESSLER, P.E.; MOORE, I.D.; MCKENZIE, N.J.; RYAN, P.J. Soil-landscape modelling and spatial prediction of soil attributes. **International Journal of Geographical Information Systems**, v.9, p.421-432, 1995.

GIASSON, E.; CLARKE, R.T.; INDA JUNIOR, A.V.; MERTEN, G.H.; TORNQUIST, C.G. Digital soil mapping using multiple logistic regression on terrain parameters in Southern Brazil. **Scientia Agricola**, v.63, p.262- 268, 2006.

GIASSON, E.; HARTEMINK, A.E.; TORNQUIST, C.G.; TESKE, R. & BAGATINI, T. Avaliação de cinco algoritmos de árvores de decisão e três tipos de modelos digitais de elevação para mapeamento digital de solos a nível semidetalhado na Bacia do Lageado Grande, RS, Brasil. **Ciência Rural**, v.43, p.61-67, 2013.

GLOBALSOILMAP.NET. **Specifications Version 1 GlobalSoilMap.net products**, 2011. Disponível em: <http://www.globalsoilmap.net/system/files/GlobalSoilMap_net_specifications_v2_0_edited_draft_Sept_2011_RAM_V12.pdf>. Acesso em: 10 dez. 2012.

GRIMM, R.; BEHRENS, T.; MÄRKER, M.; ELSENBEER, H. Soil organic carbon concentrations and stocks on Barro Colorado Island – Digital soil mapping using Random Forests analysis. **Geoderma**, v.146, n.1-2, p.102-113, 2008.

GRIMM, R.; BEHRENS, T. Uncertainty analysis of sample locations within digital soil mapping approaches. **Geoderma**, v.155, p.154-163, 2010.

GRUNWALD, S. What do we really know about the space–time continuum of soil landscapes? In: GRUNWALD, S. (Ed.). **Environmental soil-landscape modeling geographic information technologies and pedometrics**. Boca Raton: Taylor and Francis, 2006. p.3-36.

GRUNWALD, S. Multi-criteria characterization of recent digital soil mapping and modeling approaches. **Geoderma**, v.152, p.195-207, 2009.

GRUNWALD, S.; MCBRATNEY, A.B.; THOMPSON, J.A.; MINASNY B.; BOETTINGER, J.L. Digital Soil Mapping in a Changing World. In: MINASNY, B.; MALONE, B.P.; MCBRATNEY, A.B. (Eds) **Digital Soil Assessments and Beyond: Proceedings of the 5th Global Workshop on Digital Soil Mapping**. Sydney, Australia. 2012. 482p.

GUTH, P.L. **Geomorphometric comparison of ASTER GDEM and SRTM**. A special joint symposium of ISPRS Technical Commission IV & AutoCarto in conjunction with ASPRS/CaGIS 2010. Disponível em: <<http://www.isprs.org/proceedings/XXXVIII/part4/files/Guth.pdf>>. Acesso em: 03 de mai. 2013.

HAIR, J.F.; BLACK, W. C.; BABIN, B.; ANDERSON, R.E.; TATHAM, R.L. **Análise multivariada de dados**. Tradução: Adonai Schlup Sant'Anna. – 6 ed. – Porto Alegre: Bookman, 2009. 688p.

HARTEMINK, A.E. Soils are back on the global agenda. **Soil Use Manage**, v.24, p.327-330, 2008.

HARTEMINK, A.E.; MCBRATNEY, A.B. A soil science renaissance. **Geoderma**, v.148, p.123-129, 2008.

HARTEMINK, A.E.; MINASNY, B. Towards digital soil morphometrics. **Geoderma**, v. 230-231, p.305-317, 2014.

HASENACK, H.; WEBER, E. (Org.). **Base cartográfica vetorial contínua do Rio Grande do Sul - escala 1:50.000**. Porto Alegre: UFRGS-IB-Centro de Ecologia, 2010. 1 DVD-ROM. (Série Geoprocessamento, 3).

HENGL T. **Pedometric mapping: bridging the gaps between conventional and pedometric approaches**. PhD thesis, University of Wageningen, Enschede, 2003. 214p.

HENGL, T.; GRUBER, S.; SHRESTHA, D.P. Reduction of errors in digital terrain parameters used in soil-landscape modeling. **International Journal of Applied Earth Observation and Geoinformation**, v.5, p.97-112, 2004.

HEUVELINK, G. The definition of pedometrics. **Pedometron**, v.15, p.11-12, 2003.

HIRANO, A.; WELCH, R.; LANG, H. **Mapping from ASTER stereo image data: DEM validation and accuracy assessment**. ISPRS Journal of Photogrammetry & Remote Sensing, v.57, p.356-370, 2003.

HUDSON, B.D. The soil survey as paradigm-based science. **Soil Science Society of America Journal**, v.56, p.836-841, 1992.

HUTCHINSON, M.F.; GALLANT, J.C. Digital elevation models and representation of terrain shape. In: WILSON, D.J.; GALLANT, J.C. (Eds.). **Terrain analysis: principles and applications**. New York: John Wiley and Sons, 2000. p.29-50.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. **MAPGEO: Sistema de Interpolação de Ondulação Geoidal**. Versão 1.0, Brasil: IBGE, 2010.

IORIO, M.M.; LASTORIA, G.; MIOTO, C.L.; ALBREZ, E.A.; PARANHOS FILHO, A.C. Avaliação de modelos digitais de elevação extraídos de imagens ALOS/PRISM e comparação com os modelos disponibilizados gratuitamente na web. São Paulo, UNESP, **Geociências**, v.31, p.650-664, 2012.

JARVIS, A.; RUBIANO, J.; NELSON, A.; FARROW, A.; MULLIGAN, M. **Practical use of SRTM data in the tropics: comparisons with digital elevation models generated from cartographic data**. Cali: Centro Internacional de Agricultura Tropical (CIAT), 2004. 32p. Working document n.198

JENNY, H. **Factors of soil formation - a system of quantitative pedology**. New York: Dover Publications, 1941. 281p.

KEMPEN, B.; BRUS, D.J.; STOORVOGEL, J.J. Three-dimensional mapping of soil organic matter content using soil type-specific depth functions. **Geoderma**, v.162, p.107-123, 2011.

KERRY, R.; GOOVAERTS, P.; RAWLINS, B.G.; MARCHANT, B.P. Disaggregation of legacy soil data using area to point kriging for mapping soil organic carbon at the regional scale. **Geoderma**, v.170, p.347-358, 2012.

KHEIR, R.B.; GREVE, M.H.; BØCHER, P.K.; GREVE, M.B.; LARSEN, R.; MCCLOY, K. Predictive mapping of soil organic carbon in wet cultivated lands using classification-tree based models: The case study of Denmark. **Journal of Environmental Management**, v.91, p.1150-1160, 2010.

LAGACHERIE, P.; LEGROS, J.; BURFOUGH, P. A soil survey procedure using the knowledge of soil pattern established on a previously mapped reference area. **Geoderma**, v.65, p.283-301, 1995.

LAGACHERIE, P.; ROBBEZ-MASSON, J.; NGUYEN-THE, N.; BARTHÈS, J. Mapping of reference area representatively using a mathematical soil scape distance. **Geoderma**, v.101, p.105-118, 2001.

LAGACHERIE, P.; MCBRATNEY, A.B.; VOLTZ, M. (Eds.). **Digital soil mapping - an introductory perspective**. Amsterdam: Elsevier, 2006. p.3-22.

LARK, R.; BISHOP, T. Cokriging particle size fractions of the soil. **European Journal of Soil Science**, v.58, p.763-774, 2007.

LIN, Z.; OGUCHI, T. DEM analysis on longitudinal and transverse profiles of steep mountainous watersheds. **Geomorphology**, v.78, p.77-89, 2006.

MBAH, C.N.; IDIKE, F.I. et al. Carbon Storage in Tropical Agricultural Soils of South Eastern Nigeria under different Management Practices. **International Research Journal of Agricultural Science and Soil Science**, v.1, n.2, p.053-057, 2011.

MCBRATNEY, A.B.; ODEH, I.O.; BISHOP, T.F.; DUNBAR, M.S.; SHATAR, T.M. An overview of pedometric techniques for use in soil survey. **Geoderma**, v.97, p.293-327, 2000.

MCBRATNEY, A.B.; MENDONÇA-SANTOS, M.; MINASNY, B. On digital soil mapping. **Geoderma**, v.117, p.3-52, 2003.

MCDONALD, B. A teaching note on Cook's distance – a guideline. **Research Letters in the Information and Mathematical Sciences**, v.3, p.127-128, 2002.

MCKENZIE, N.J.; RYAN, P.J. Spatial prediction of soil properties using environmental correlation. **Geoderma**, v.89, p.67-94, 1999.

MCKENZIE, N.; GESSLER, P.E.; RYAN, P.J.; O'CONNELL, D. The role of terrain analysis in soil mapping. In: WILSON, J.; GALLANT, J. (Eds.). **Terrain analysis: principles and applications**, New York: John Wiley and Sons, 2000. p.245-266.

MCKENZIE, N.; GRUNDY, M.; WEBSTER, R.; RINGROSE-VOASE, A. **Guidelines for survey in soil and land resources**. 2.ed. Melbourne: CSIRO Publishing, 2008. 576p.

MENDONÇA-SANTOS, M.L.; SANTOS, H.G. **Mapeamento Digital de Classes e Atributos de Solos**. Embrapa. 2003.17p.

MENDONÇA-SANTOS, M.L.; DART, R.; SANTOS, H.; COELHO, M.; BERBARA, R.; LUMBRERAS, J. Digital soil mapping of top soil organic carbon content. Rio de Janeiro State, Brazil. In: BOETTINGER, J.L.; HOWELL, D.W.; MOORE, A.C.; HARTEMINK, A.E.; KIENAST-BROWN, S. (Eds.). **Digital soil mapping: bridging research, environmental application, and operation**, New York: Springer, 2010. p.255-265.

MENDONÇA-SANTOS, M.L.; SANTOS, H.G.; DART, R.G.; PARES, J.G. **Modelagem e Mapeamento Digital de Estoque de Carbono Orgânico na Camada Superficial dos Solos (0-10 cm) do Estado do Rio de Janeiro**. Embrapa, Boletim de pesquisa e desenvolvimento. 2007. 23p.

MILLER, C.L.; LAFLAMME, R.A. The Digital Terrain Model – Theory and Application. **Photogrammetric Engineering**, v.24, n.4, p.433-442, 1958.

MINASNY, B.; MCBRATNEY, A.B. Spatial prediction of soil properties using EBLUP with the Matérn covariance function. **Geoderma**, v.140, p.324-336, 2007.

MINASNY, B.; MCBRATNEY, A.B. Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. **Chemometrics and Intelligent Laboratory Systems**, v.94, n.1, p.72-79, 2008.

MINASNY, B.; MCBRATNEY, A.B.; SALVADOR-BLANES, S. Quantitative models for pedogenesis — a review. **Geoderma**, v.144, p.140-157, 2008.

MOORE, I.D.; GESSLER, P.E.; NIELSEN, G.A.; PETERSON, G.A. Soil attribute prediction using terrain analysis. **Soil Science Society of America Journal**, v.57, p.443-452, 1993.

MOURA, L. Z.; BIAS, E.S.; BRITES, R. Avaliação da acurácia vertical de modelos digitais de elevação (mdes) nas bacias do Paranoá e São Bartolomeu. **Revista Brasileira de Cartografia**, v.66, p.1-14, 2014.

MUKHERJEE S.; JOSHI, P.K.; MUKHERJEE S., GHOSH, A.; GARG, R.D.; MUKHOPADHY, A.Y.A. Evaluation of vertical accuracy of open source Digital Elevation Model (DEM). **International Journal of Applied Earth Observation and Geoinformation**, v.21, p.205-217, 2013.

NOCITA, M.; STEVENS, A.; TOTH, G.; PANAGOS, P.; WESEMAEL, B.V.; MONTANARELLA, L. Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least square regression approach. **Soil Biology and Biochemistry**, v.68, p.337-347, 2014.

OLIVEIRA, A. H.; SILVA, M. A.; SILVA, M. L. N.; AVANZI, J. C.; CURI, N.; LIMA, G. C.; PEREIRA, P. H. Caracterização ambiental e predição dos teores de matéria orgânica do solo na Sub-Bacia do Salto, Extrema, MG. **Ciências Agrárias**, Londrina, v.33, n.1, p.143-154, 2012.

OKSANEN, J.; SARJAKOSKI, T. Error propagation of DEM-based surface derivatives. **Computer and Geoscience**, v.31, p.1015-1027, 2005.

PACHECO, L.P.; BARBOSA, J.M.; LEANDRO, W.M.; MACHADO, P.L.O.A.; ASSIS, R.L.; MADARI, B.E.; PETTER, F.A. Produção e ciclagem de nutrientes por plantas de cobertura nas culturas de arroz de terras altas e de soja. **Revista Brasileira de Ciência do Solo**, v.35, p.1787-1799, 2011.

PARANHOS FILHO, A.C.; LASTORIA, G.; TORRES, T.G. **Sensoriamento Remoto Ambiental Aplicado**. Campo Grande: Editora UFMS, 2008.

PEI, T.; QIN, C.; ZHU, A.; YANG, L.; LUO, M.; LI, B.; ZHOU, C. Mapping soil organic matter using the topographic wetness index: A comparative study based on different flow-direction algorithms and kriging methods. **Ecological Indicators**, v.10, p.610-619, 2010.

PHACHOMPHON, K.; DLAMINI, P.; CHAPLOT, V. Estimating carbon stocks at a regional level using soil information and easily accessible auxiliary variables. **Geoderma**, v.155, p.372-380. 2010.

PINHEIRO, E.S. Comparação entre dados altimétricos Shuttle Radar Topography Mission, cartas topográficas e GPS: numa área com relevo escarpado. **Revista Brasileira de Cartografia**, v.58, n.1, p.1-9, 2006.

PINHEIRO, H.S.K.; CHAGAS, C.S.; CARVALHO JÚNIOR, V.; ANJOS, L.H.C. Modelos de elevação para obtenção de atributos topográficos utilizados em mapeamento digital de solos. C. **Pesquisa Agropecuária Brasileira**, v.47, p.1384-1394, 2012.

PINHEIRO, E.C.C.; BUSCIOLI, A.; MILANEZI, C.H.S.; SILVA, E.F.C.; STEFANES, M. **Disponibilidade de imagens de satélites gratuitas para gestão ambiental territorial dos municípios**. In: XVI Simpósio Brasileiro de Sensoriamento Remoto; Abril 2013; Foz do Iguaçu, PR, Brasil, p.4242-4249, 2013.

R Core Team (2014). R: **A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

RAZAKAMANARIVO, R.H.; GRINAND, C.; RAZAFINDRAKOTO, M.A.; BERNOUX, M.; ALBRECHT, A. Mapping organic carbon stocks in eucalyptus plantations of the central highlands of Madagascar: A multiple regression approach. **Geoderma**, v.162, p.335-346, 2011.

SAGA DEVELOPMENT TEAM. **SAGA GIS**. Hamburg, Alemanha, 2010. Versão 2.0.6. Disponível em: <<http://www.saga-gis.org/en/index.html>>.

SAMUEL-ROSA, A; DALMOLIN, R.S.D.; MIGUEL, P. Building predictive models of soil particle-size distribution. **Revista Brasileira de Ciência do Solo**, v.37, p.422-430, 2013.

SANCHEZ, P.A.; AHAMED, S.; CARRE, F.; HARTEMINK, A.E.; HEMPEL, J.; HUISING, J.; LAGACHERIE, P.; MCBRATNEY, A. B.; MCKENZIE, N. J.; MENDONCA SANTOS, M.L.; MINASNY, B.; MONTANARELLA, L.; OKOTH, P.; PALM, C.A.; SACHS, J.D.; SHEPHERD, K.D.; VAGEN, T.-G.; VANLAUWE, B.; WALSH, M.G.; WINOWIECKI, L.A.; ZHANG, G.-L. Digital soil map of the world. **Science**, v.325, p.680-681, 2009.

SARMENTO, E.C.; FLORES, C.A.; WEBER, E.; HASENACK, H. REINALDO OSCAR PÖTTER, R.O. Sistema de informação geográfica como apoio ao levantamento detalhado de solos do Vale dos Vinhedos. **Revista Brasileira de Ciência do Solo**, v.32, p.2795-2803, 2008.

SLATER, J.A.; HEADY, B.; KROENUNG, G.; CURTIS, W.; HAASE, J.; HOEGEMANN, D.; SHOCKLEY, C.; KEVIN, T. **Evaluation of the New ASTER Global Digital Elevation Model**. 2009. 27p. Disponível em: <<http://earth-info.nga.mil/GandG/elevation/>>. Acesso em: 18 jan. 2014.

STOCKMANN, U.; MINASNY, B.; MCBRATNEY, A.B. How fast does soil grow? **Geoderma**, v.216, p.48-61, 2014.

SUBRAMANIAN, K.S.S.; SINGH, A.; SUDHAKAR, M. **Evaluation of digital elevation models created from different satellite images**. 2005. 8p. Disponível em:<<http://www.gisdevelopment.net/technology/rs/mi03079pf.htm>>. Acesso em: 21 jan. 2014.

TAGHIZADEH-MEHRJARDI, R.; MINASNY, B.; SARMADIAN, F; MALONE, B.P. Digital mapping of soil salinity in Ardakan region, central Iran. **Geoderma**, v.213, p.15-28, 2014.

TEN CATEN, A.; DALMOLIN, R.S.D. **Geoprocessamento como aliado da pedologia**. In Boletim Informativo da Sociedade Brasileira de Ciência do Solo – Em debate, a utilização dos mapeamentos de solos no Brasil. v.39, n.1, p.0-36, 2014.

TEN CATEN, A.; DALMOLIN, R.S.D.; PEDRON, F.A.; RUIZ, L.F.C.; SILVA, C.A. An appropriate data set size for digital soil mapping in Erechim, Rio Grande do Sul, Brazil. **Revista Brasileira de Ciência do Solo**, v.37, p.359-366, 2013.

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

TEN CATEN, A.; DALMOLIN, R.S.D.; PEDRON, F.A.; MENDONÇA-SANTOS, M.L. Estatística multivariada aplicada à diminuição do número de preditores no mapeamento digital do solo. **Pesquisa Agropecuária Brasileira**, v.46, p.554-562, 2011a.

TEN CATEN, A.; DALMOLIN, R.S.D.; PEDRON, F.A.; MENDONÇA-SANTOS, M.L. Regressões logísticas múltiplas: fatores que influenciam sua aplicação na predição de classes de solos. **Revista Brasileira de Ciência do Solo**, v.35, p.53-62, 2011b.

TESKE, R.; GIASSON E.; BAGATINI, T. Comparação do uso de modelos digitais de elevação em mapeamento digital de solos em Dois Irmãos, RS, **Brasil**. **Revista Brasileira de Ciência do Solo**, v.38, p.1367-1376, 2014.

THOMPSON, A.J.; BELL, J.C.; BUTLER, C.A. Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling. **Geoderma**, v.100, p.67-89, 2001.

VALERIANO, M.M. **TOPODATA: guia de utilização de dados geomorfométricos locais**. São José dos Campos: INPE, 2008. 44p.

VALERIANO, M.M.; ROSSETTI, D.F. **TOPODATA: Seleção de coeficientes geoestatísticos para o refinamento unificado de dados SRTM**. São José dos Campos: INPE, 2009.

VALERIANO, M.M.; ROSSETTI, D.F. TOPODATA: Brazilian full coverage refinement of SRTM data. **Applied Geography**, v.32, p.300-309, 2012.

VAN ZIJL, G.M.; BOUWER, D.; VAN TOL, J.J.; LE ROUX, P.A. L. Functional digital soil mapping: A case study from Namarroi, Mozambique. **Geoderma**, v.219-220, p.155-161, 2014.

VILLELA, A.L.O. **Mapeamento Digital de Solos da Formação Solimões Sob Floresta Tropical Amazônica**. 2013. 114 f. Tese (Doutorado em Ciências, no Curso de Pós-Graduação em Agronomia, Área de Concentração em Ciência do Solo). Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ. 2013.

VISCARRA ROSSEL, R.A.; 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**, v.113, p.243-291, 2011.

VISCARRA-ROSSEL, R.A.; MCBRATNEY, A.B. Diffuse reflectance spectroscopy as a tool for digital soil mapping. In: HARTEMINK, A. A.M.L.M.-S. (Ed.), **Digital Soil Mapping with Limited Data**. Springer. 2008. p.165-172.

WEBSTER, R. **Quantitative spatial analysis of soil in the field**. Advances in Soil Science, New York, v.3, p.1-70, 1984.

WHEELER, I.; MCBRATNEY, A.B.; MINASNY, B.; GRUIJTER, J.J. Digital Soil Mapping to inform design-based sampling strategies for estimating total organic carbon stocks at the farm scale. In: MINASNY, B.; MALONE, B.P.; MCBRATNEY, A.B. **Proceedings of The 5th Global Workshop on Digital Soil Mapping**. Sydney, Australia. "Digital Soil Assessments and Beyond". 2012. 482p.

WILSON, J.P.; GALLANT, J.C. **Terrain Analysis: Principles and Applications**. New York: John Wiley & Sons. 2000. 485p.

WISE, S. Assessing the quality for hydrological applications of digital elevation models derived from contours. **Hydrological Processes**, v.14, p.1909-1929, 2000

YEOMANS, J.C.; BREMNER, J.M. A rapid and precise method routine determination of organic carbon in soil. **Communications in Soil Science and Plant Analysis**, v.19, p.1467-1474.

YU, D.; SHI, X.; WANG, H.; SUN, W.; CHEN, J. LIU, Q.; ZHAO, Y. Regional of soil organic carbon stocks in China. **Journal Environmental Management**, v.85, p.680-689, 2007.

ZANATA, M.; PISSARRA, T.C.T.; ARRAES, C.L.; RODRIGUES, F.M.; CAMPOS, S. Influência da escala na análise morfométrica de microbacias hidrográficas. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.15, n.10, p.1062-1067, 2010.

ZHU, Q.; LIN, H.S.; DOOLITTLE, J.A. Functional soil mapping for site specific soil moisture and crop yield management. **Geoderma**, v. 200-201, p.45-54, 2013.

9 ANEXO

Anexo – Rotina das análises estatísticas realizadas no pacote estatístico R

1 - Análise dos dados de altitude dos modelos digitais de elevação (MDE)

```
# Carregar o arquivo de dados para o ambiente R
dados_altitude=read.table("dados_altitude.csv",sep=";",head=T)

# Vincular objeto à área de trabalho
attach(dados_altitude)

# Verificar os nomes das colunas do objeto fixado
colnames(dados_altitude)
vetor1 <-c(pc)
vetor2 <-c(lp)
vetor3<-c(cartas)
vetor4 <-c(aster)
vetor5 <-c(srtm)
vetor6 <-c(topodata)

# Instalar os pacotes estatísticos requeridos
install.packages("MASS")
library(MASS)
install.packages("fBasics")
library(fBasics)
install.packages("fitdistrplus")
library(fitdistrplus)

# Análise estatística descritiva dos dados

summary(dados_altitude)
basicStats(pc, ci=0.95)
```

```
descdist(pc)
basicStats(lp, ci=0.95)
descdist(lp)
basicStats(carta, ci=0.95)
descdist(carta)
basicStats(aster, ci=0.95)
descdist(aster)
basicStats(srtm, ci=0.95)
descdist(srtm)
basicStats(topodata, ci=0.95)
descdist(topodata)
```

Teste de normalidade do conjunto dos dados

```
attach(dados_altitude)
vetor1 <-c(pc)
vetor2 <-c(lp)
vetor3<-c(carta)
vetor4 <-c(aster)
vetor5 <-c(srtm)
vetor6 <-c(topodata)
```

Teste de shapiro-wilk

```
shapiro.test(pc)
shapiro.test(lp)
shapiro.test(carta)
shapiro.test(aster)
shapiro.test(srtm)
shapiro.test(topodata)
```

Gerar histograma de frequência e curva de distribuição do conjunto de dados

```
histPlot(as.timeSeries(pc))
```

```
hist(pc, main="Histograma de PC", ylab="Frequência", font=6, font.lab=6, font.main=6,  
col="gray")
```

```
qqnorm(pc) # gera a linha de distribuição das amostras em relação a NORMAL
```

```
qqnormPlot(pc)
```

```
histPlot(as.timeSeries(lp))
```

```
hist(lp, main="Histograma de LP", ylab="Frequência", font=6, font.lab=6, font.main=6,  
col="gray")
```

```
qqnorm(lp) # gera a linha de distribuição das amostras em relação a NORMAL
```

```
qqnormPlot(lp)
```

```
histPlot(as.timeSeries(carta))
```

```
hist(carta, main="Histograma de CARTA", ylab="Frequência", font=6, font.lab=6,  
font.main=6, col="gray")
```

```
qqnorm(carta) # gera a linha de distribuição das amostras em relação a NORMAL
```

```
qqnormPlot(carta)
```

```
histPlot(as.timeSeries(aster))
```

```
hist(aster, main="Histograma de CARTA", ylab="Frequência", font=6, font.lab=6,  
font.main=6, col="gray")
```

```
qqnorm(aster) # gera a linha de distribuição das amostras em relação a NORMAL
```

```
qqnormPlot(aster)
```

```
histPlot(as.timeSeries(srtm))
```

```
hist(srtm, main="Histograma de SRTM", ylab="Frequência", font=6, font.lab=6,  
font.main=6, col="gray")
```

```
qqnorm(srtm) # gera a linha de distribuição das amostras em relação a NORMAL
```

```
qqnormPlot(srtm)
```

```
histPlot(as.timeSeries(topodata))
```

```
hist(topodata, main="Histograma de topodata", ylab="Frequência", font=6, font.lab=6,  
font.main=6, col="gray")
```

```
qqnorm(topodata) # gera a linha de distribuição das amostras em relação a NORMAL
```

```
qqnormPlot(topodata)
```

Criar perfil topográfico dos valores de altitude para cada modelo digital de elevação

```
# Carregar o arquivo de dados para o ambiente R
```

```
dados_perfis_elevation =read.table("dados_perfis_elevation.csv",sep=";",head=T)
```

```
# Vincular objeto à área de trabalho
```

```
attach(dados_perfis_elevation)
```

```
distancia<-0:50
```

```
tri1<-c(lp)
```

```
tri2<-c(cartas)
```

```
tri3<-c(aster)
```

```
tri4<-c(srtm)
```

```
tri5<-c(topodata)
```

```
install.packages('MASS')
```

```
library(MASS)
```

```
plot(dados_perfis_elevation$topodata, type="l", main="Perfis topográficos dos MDE
avaliados", xlab="", ylab="Altitude", col=8, lty=5, lwd=2, ylim=c(330,430), pch=1)
```

```
lines(dados_perfis_elevation $carta, lty=1, lwd=2, col=8,type="l")
```

```
lines(dados_perfis_elevation $aster, col=1,lty=3,lwd=2,type="l")
```

```
lines(dados_perfis_elevation $srtm, col=1, lty=4, lwd=2, type="l")
```

```
lines(dados_perfis_elevation $lp, col=1, lty=1, lwd=2,type="l")
```

```
legend(120,370,c("LP","CARTA", "ASTER", "SRTM", "TOPODATA"), col=c(1,8,1,1,8),
lty=c(1,1,3,4,5), lwd=c(2,2,2,2,2), cex=(0.8), bty=("n"))
```

Elaboração dos gráficos boxplot

```
# Boxplot dos valores de altitude dos MDE
```

```
attach(dados_altitude)
```

```
boxplot(dados_altitude, notch=T , col = 8, ylab="Altitude (m)", ylim=c(320,420))
```

```
# Boxplot dos valores das diferenças de altitude dos MDE
```

```
attach(dados_diferença_altitude)
```

```
boxplot(dados_diferença_altitude, notch=T, col = 8, ylab="Altitude (m)", ylim=c(-20,20))
```

2 – Análises das variáveis preditoras e construção da função de predição para o COS

```
# Carregar arquivo de dados para o ambiente R
```

```
dados_variaveis_preditoras=read.table("dados_atributos_terreno.csv",sep=";",head=T)
```

```
# Vincular objeto à área de trabalho
```

```
attach(dados_variaveis_preditoras)
```

```
vetor1 <-c(RSP)
```

```
vetor2 <-c(VD)
```

```
vetor3 <-c(VDCN)
```

```
vetor4 <-c(ELEV)
```

```
vetor5 <-c(CA)
```

```
vetor6 <-c(DMAC)
```

```
vetor7 <-c(DECL)
```

```
vetor8 <-c(ASP)
```

```
vetor9 <-c(CNBL)
```

```
vetor10 <-c(PLAN)
```

```
vetor11 <-c(PROF)
```

```
vetor12 <-c(TWI)
```

```
vetor13 <-c(CONV)
```

```
vetor14 <-c(TRI)
```

```
vetor15 <-c(IPE)
```

```
vetor16 <-c(LS)
```

```
vetor17 <-c(NDVI)
```

```
vetor18 <-c(RFL_IVP)
```

```
vetor19 <-c(SAVI)
```

```
vetor20 <-c(CURV)
```

```
# Instalar os pacotes estatísticos requeridos
```

```
install.packages("MASS")
```

```
library(MASS)
```

```
install.packages("fBasics")
```

```
library(fBasics)
```

```
install.packages("fitdistrplus")
```

```
library(fitdistrplus)
```

```
# Análise estatística descritiva das variáveis preditoras
```

```
summary(dados_variaveis_preditoras)
```

```
basicStats(RSP, ci=0.95)
```

```
descdist(RSP)
```

```
basicStats(VD, ci=0.95)
```

```
descdist(VD)
```

```
basicStats(VDCN, ci=0.95)
```

```
descdist(VDCN)
```

```
basicStats(ELEV, ci=0.95)
```

```
descdist(ELEV)
```

```
basicStats(CA, ci=0.95)
```

```
descdist(CA)
```

```
basicStats(DMAC, ci=0.95)
```

```
descdist(DMAC)
```

```
basicStats(DECL, ci=0.95)
```

```
descdist(DECL)
```

```
basicStats(ASP, ci=0.95)
```

```
descdist(ASP)
```

```
basicStats(CNBL, ci=0.95)
```

```
descdist(CNBL)
```

```
basicStats(PLAN, ci=0.95)
```

```
descdist(PLAN)
```

```
basicStats(PROF, ci=0.95)
```

```
descdist(PROF)
basicStats(TWI, ci=0.95)
descdist(TWI)
basicStats(CONV, ci=0.95)
descdist(CONV)
basicStats(TRI, ci=0.95)
descdist(TRI)
basicStats(IPE, ci=0.95)
descdist(IPE)
basicStats(LS, ci=0.95)
descdist(LS)
basicStats(NDVI, ci=0.95)
descdist(NDVI)
basicStats(REFL-IVP, ci=0.95)
descdist(REFL-IVP)
basicStats(SAVI, ci=0.95)
descdist(SAVI)
basicStats(CURV, ci=0.95)
descdist(CURV)
```

Teste de normalidade do conjunto dos dados

```
# Teste de shapiro-wilk
```

```
shapiro.test(RSP)
shapiro.test(VC)
shapiro.test(VDCN)
shapiro.test(ELEV)
shapiro.test(CA)
shapiro.test(DMAC)
shapiro.test(DECL)
shapiro.test(ASP)
shapiro.test(CNBL)
shapiro.test(PLAN)
```

```

shapiro.test(PROF)
shapiro.test(TWI)
shapiro.test(CONV)
shapiro.test(TRI)
shapiro.test(IPE)
shapiro.test(LS)
shapiro.test(NDVI)
shapiro.test(REFL-IVP)
shapiro.test(SAVI)
shapiro.test(CURV)

```

Gerar histograma de frequência e curva de distribuição do conjunto de variáveis preditoras

```

x <-RSP

h<-hist(x, breaks=10, col="grey", xlab = expression(paste("RSP")),font.lab=6),
ylab="Frequência",

      xlim=c(0.0,1.1), ylim=c(0,60),font.lab=6, font.axis=6)

xfit<-seq(min(x),max(x),length=80)

yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))

yfit <- yfit*diff(h$mids[1:2])*length(x)

lines(xfit, yfit, col="black", lwd=1)

x <-VD

h<-hist(x, breaks=10, col="grey", xlab = expression(paste("VD")),font.lab=6),
ylab="Frequência",

      xlim=c(0.0,50), ylim=c(0,48),font.lab=6, font.axis=6)

xfit<-seq(min(x),max(x),length=100)

yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))

yfit <- yfit*diff(h$mids[1:2])*length(x)

lines(xfit, yfit, col="black", lwd=1)

```

```
x <-VDCN
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("VDCN")),
font.lab=6),ylab="Frequência",
```

```
    xlim=c(0,60), ylim=c(0,60),font.lab=6, font.axis=6)
```

```
xfit<-seq(min(x),max(x),length=100)
```

```
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
```

```
yfit <- yfit*diff(h$mids[1:2])*length(x)
```

```
lines(xfit, yfit, col="black", lwd=1)
```

```
x <-ELEV
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("ELEV")),font.lab=6),
ylab="Frequência",
```

```
    xlim=c(320,420), ylim=c(0,60),font.lab=6, font.axis=6)
```

```
xfit<-seq(min(x),max(x),length=100)
```

```
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
```

```
yfit <- yfit*diff(h$mids[1:2])*length(x)
```

```
lines(xfit, yfit, col="black", lwd=1)
```

```
x <-CA
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("CA")), font.lab=6),
ylab="Frequência",
```

```
    xlim=c(0,1500000), ylim=c(0,300),font.lab=6, font.axis=6)
```

```
xfit<-seq(min(x),max(x),length=100)
```

```
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
```

```
yfit <- yfit*diff(h$mids[1:2])*length(x)
```

```
lines(xfit, yfit, col="black", lwd=1)
```

```
x <-DMAC
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("DMAC")),font.lab=6),
ylab="Frequência",
```

```
    xlim=c(0,0.08), ylim=c(0,70),font.lab=6, font.axis=6)
```

```
xfit<-seq(min(x),max(x),length=100)
```

```
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
```

```
yfit <- yfit*diff(h$mids[1:2])*length(x)
```

```
lines(xfit, yfit, col="black", lwd=1)
```

```
x <-DECL
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("DECL")), font.lab=6),
ylab="Frequência",
```

```
    xlim=c(0,7), ylim=c(0,40),font.lab=6, font.axis=6)
```

```
xfit<-seq(min(x),max(x),length=40)
```

```
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
```

```
yfit <- yfit*diff(h$mids[1:2])*length(x)
```

```
lines(xfit, yfit, col="black", lwd=1)
```

```
x <-ASP
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("ASP")),font.lab=6),
ylab="Frequência",
```

```
    xlim=c(50,350), ylim=c(0,80),font.lab=6, font.axis=6)
```

```
xfit<-seq(min(x),max(x),length=100)
```

```
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
```

```
yfit <- yfit*diff(h$mids[1:2])*length(x)
```

```
lines(xfit, yfit, col="black", lwd=1)
```

```
x <-CNBL
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("CNBL")),font.lab=6),
ylab="Frequência",
```

```

xlim=c(330,380), ylim=c(0,90),font.lab=6, font.axis=6)
xfit<-seq(min(x),max(x),length=80)
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
yfit <- yfit*diff(h$mids[1:2])*length(x)
lines(xfit, yfit, col="black", lwd=1)

```

```
x <-PLAN
```

```

h<-hist(x, breaks=10, col="grey", xlab = expression(paste("PLAN")),font.lab=6),
ylab="Frequência",

```

```

xlim=c(0.55,0.63), ylim=c(0,150),font.lab=6, font.axis=6)
xfit<-seq(min(x),max(x),length=100)
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
yfit <- yfit*diff(h$mids[1:2])*length(x)
lines(xfit, yfit, col="black", lwd=1)

```

```
x <-PROF
```

```

h<-hist(x, breaks=10, col="grey", xlab = expression(paste("PROF")),font.lab=6),
ylab="Frequência",

```

```

xlim=c(0.2,0.7), ylim=c(0,100),font.lab=6, font.axis=6)
xfit<-seq(min(x),max(x),length=100)
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
yfit <- yfit*diff(h$mids[1:2])*length(x)
lines(xfit, yfit, col="black", lwd=1)

```

```
x <-TWI
```

```

h<-hist(x, breaks=10, col="grey", xlab = expression(paste("TWI")),font.lab=6),
ylab="Frequência",

```

```

xlim=c(4,18), ylim=c(0,130),font.lab=6, font.axis=6)

```

```
xfit<-seq(min(x),max(x),length=40)
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
yfit <- yfit*diff(h$mids[1:2])*length(x)
lines(xfit, yfit, col="black", lwd=1)
```

```
x <-CONV
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("CONV")),font.lab=6),
ylab="Frequência",
      xlim=c(0.2,0.7), ylim=c(0,130),font.lab=6, font.axis=6)
```

```
xfit<-seq(min(x),max(x),length=80)
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
yfit <- yfit*diff(h$mids[1:2])*length(x)
lines(xfit, yfit, col="black", lwd=1)
```

```
x <-TRI
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("TRI")),font.lab=6),
ylab="Frequência",
      xlim=c(0,3), ylim=c(0,45),font.lab=6, font.axis=6)
```

```
xfit<-seq(min(x),max(x),length=100)
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
yfit <- yfit*diff(h$mids[1:2])*length(x)
lines(xfit, yfit, col="black", lwd=1)
```

```
x <-IPE
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("IPE")),font.lab=6),
ylab="Frequência",
      xlim=c(-2,8), ylim=c(0,80),font.lab=6, font.axis=6)
```

```
xfit<-seq(min(x),max(x),length=100)
```

```
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
```

```
yfit <- yfit*diff(h$mids[1:2])*length(x)
```

```
lines(xfit, yfit, col="black", lwd=1)
```

```
x <-LS
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("LS")),
font.lab=6),ylab="Frequência",
```

```
      xlim=c(0,2), ylim=c(0,60),font.lab=6, font.axis=6)
```

```
xfit<-seq(min(x),max(x),length=80)
```

```
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
```

```
yfit <- yfit*diff(h$mids[1:2])*length(x)
```

```
lines(xfit, yfit, col="black", lwd=1)
```

```
x <-NDVI
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("NDVI")),font.lab=6),
ylab="Frequência",
```

```
      xlim=c(0,1), ylim=c(0,50),font.lab=6, font.axis=6)
```

```
xfit<-seq(min(x),max(x),length=100)
```

```
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
```

```
yfit <- yfit*diff(h$mids[1:2])*length(x)
```

```
lines(xfit, yfit, col="black", lwd=1)
```

```
x <-REFF-IVP
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("REFL-IVP")),font.lab=6),
ylab="Frequência",
```

```
      xlim=c(0,1), ylim=c(0,50),font.lab=6, font.axis=6)
```

```
xfit<-seq(min(x),max(x),length=100)
```

```
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
```

```
yfit <- yfit*diff(h$mids[1:2])*length(x)
```

```
lines(xfit, yfit, col="black", lwd=1)
```

```
x <-SAVI
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("SAVI")),font.lab=6),
ylab="Frequência",
```

```
      xlim=c(0,1), ylim=c(0,50),font.lab=6, font.axis=6)
```

```
xfit<-seq(min(x),max(x),length=100)
```

```
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
```

```
yfit <- yfit*diff(h$mids[1:2])*length(x)
```

```
lines(xfit, yfit, col="black", lwd=1)
```

```
x <-CURV
```

```
h<-hist(x, breaks=10, col="grey", xlab = expression(paste("CURV")),font.lab=6),
ylab="Frequência",
```

```
      xlim=c(0.1,0.7), ylim=c(0,80),font.lab=6, font.axis=6)
```

```
xfit<-seq(min(x),max(x),length=100)
```

```
yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
```

```
yfit <- yfit*diff(h$mids[1:2])*length(x)
```

```
lines(xfit, yfit, col="black", lwd=1)
```

Transformação do conjunto de variáveis preditoras

```
CA=log(CA)
```

```
TWI=log(TWI)
```

```
PLAN=sqrt(PLAN)
```

```
CONV=sqrt(CONV)
```

```
REFL-IVP=sqrt(REFL-IVP)
```

Gerar matriz de correlação linear das variáveis preditoras

```
# Carregar o arquivo de dados para o ambiente R
```

```
variaveis_preditoras =read.table("variaveis_preditoras.csv",sep=";",head=T)
```

```
# Vincular objeto à área de trabalho
```

```
attach(variaveis_preditoras)
```

```
matriz=cbind(RSP, VD, VDCN, ELEV, CA, DMAC, DECL, ASP, CNBL, PLAN, PROF,  
TWI, CONV, TRI, IPE, LS, NDVI, REFL-IVP, SAVI, CURV)
```

#Análise de Componentes Principais (ACP)

```
# Carregar o arquivo de dados para o ambiente R
```

```
acp =read.table("acp.csv",sep=";",head=T)
```

```
# Vincular objeto à área de trabalho
```

```
attach(acp)
```

```
# Obtenção da matriz de correlação
```

```
acp=cor(acp)
```

```
# Carregar o pacote necessário para o teste de esfericidade de Bartlett
```

```
install.packages("psych")
```

```
library(psych)
```

```
# Adequação do conjunto de dados para aplicação da ACP
```

```
# Realização do teste de esfericidade de Bartlett e KMO (Kaiser-Meyer-Olkin) e MSA  
(Measure of Sample Adequacy)
```

```
cortest.bartlett(acp,243)
```

```
KMO(acp)
```

```
# o atributo ASP e REFL-IVP apresentaram valor menor que 0,5 e devem ser removidos do conjunto de dados
```

```
acp$ASP=NULL
```

```
acp$REFL-IVP=NULL
```

```
View(acp) # mostra a tabela de dados sem os atributos ASP e REFL-IVP
```

```
KMO(acp)
```

```
# Análise de componentes principais
```

```
# Calcular autovalores e autovetores da matriz de correlação
```

```
auto1=eigen(cor(acp))
```

```
# Gerar scree-plot
```

```
plot(auto1$values,type="l",xlab="Componente principal",ylab="Autovalor",font.lab=6,font.axis=6)
```

```
points(auto1$values,pch=20)
```

```
abline(h=1)
```

```
# Realizar a análise de componentes principais
```

```
pca1=princomp(acp,cor=T)
```

```
summary(pca1)
```

```
print(pca1$loadings,digits=3,cutoff=NULL)
```

```
pca1$scores
```

```
# Calcular a correlação entre as variáveis originais e os escores das componentes principais
```

```
CP1=pca1$loadings[,1]*sqrt(auto1$values[1])
```

```
CP2=pca1$loadings[,2]*sqrt(auto1$values[2])
```

```
CP3=pca1$loadings[,3]*sqrt(auto1$values[3])
```

```

# Gerar os gráficos das projeções

library(plotrix)

par(ps=16)

proj1=cbind(CP1,CP2)

proj2=cbind(CP2,CP3)

plot(proj1,asp=1,xlim=c(-1,1),ylim=c(-1,1),xlab="CP 1 (39%)",ylab="CP 2
(26%)",pch=20,font.lab=6,font.axis=6)

draw.circle(0,0,1)

abline(v=0,h=0)

text(proj1,rownames(proj1),pos=2,font=6)

plot(proj2,asp=1,xlim=c(-1,1),ylim=c(-1,1),xlab="CP 2 (26%)",ylab="CP 3
(11%)",pch=20,font.lab=6,font.axis=6)

draw.circle(0,0,1)

abline(v=0,h=0)

text(proj2,rownames(proj2),pos=1,font=6)

```

Gerar grupo de amostras de calibração e validação de forma aleatória

```

# Carregar o arquivo de dados para o ambiente R
variaveis_pred_cot =read.table("variaveis_pred_cot.csv",sep=";",head=T)

# Vincular objeto à área de trabalho
attach(variaveis_pred_cot)

#170 – calibração do modelo
#73 – validação do modelo
sample.int(243, size = 73, replace = FALSE, prob = NULL)

# Construção da função de predição por regressão linear múltipla (RLM)

# Instalar os pacotes estatísticos requeridos

```

```
install.packages("MASS")
library(MASS)

# Carregar o arquivo de dados para o ambiente R
variaveis_predi_cot_170 =read.table("variaveis_predi_cot_170.csv",sep=";",head=T)

# Vincular objeto à área de trabalho
attach(variaveis_pred_cot_170)

fit<-lm(COS~DECL+SAVI+CNBL+RSP+TRI+DMAC+IPE+TWI+CA+CONV+CURV)
fit<-stepAIC(fit, direction="both",ci=0.05)
summary(fit)

# coeficientes do modelo
coefficients(fit)

# valores preditos
fitted(fit)

# resíduos
residuals(fit)

# tabela da anova
anova(fit)

# valor de AIC
extractAIC(fit)

# Avaliação do ajuste da função de predição
influence(fit)

# Avaliação da função de predição por meio de gráficos

par(ps=18)
```

```
par(font=6)
plot(fit,which=1,font=6,font.lab=6,font.axis=6,font.main=6,pch=20
     ,font.sub=6,lwd=3,id.n=6, labels.id=atributos.cot$id)
plot(fit,which=2,font=6,font.lab=6,font.axis=6,font.main=6,pch=20
     ,font.sub=6,lwd=3,id.n=6, labels.id=atributos.cot$id)
plot(fit3,which=4,font=6,font.lab=6,font.axis=6,font.main=6,pch=20
     ,font.sub=6,lwd=3,id.n=6, labels.id=atributos.cot$id)
plot(fit3,which=5,font=6,font.lab=6,font.axis=6,font.main=6,pch=20
     ,font.sub=6,lwd=3,id.n=6, labels.id=atributos.cot$id)

# Calcular a raiz quadrada do erro médio quadrático (RMSE)

install.packages("qpcR")
library(qpcR)
RMSE(fit,which=NULL)
```