

UFRRJ
INSTITUTO DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA
CIÊNCIA DO SOLO

TESE

**Mapeamento Digital de Estoque de Carbono e
Frações Granulométricas de Solos da Amazônia
Central e Predição de Densidade dos Solos do Brasil**

Andréa da Silva Gomes

2018



UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA -
CIÊNCIA DO SOLO

MAPEAMENTO DIGITAL DE ESTOQUE DE CARBONO E FRAÇÕES
GRANULOMÉTRICAS DE SOLOS DA AMAZÔNIA CENTRAL E
PREDIÇÃO DE DENSIDADE DOS SOLOS DO BRASIL

ANDRÉA DA SILVA GOMES

Sob a Orientação do Professor

Marcos Bacis Ceddia

e Co-orientação dos Professores

Lúcia Helena Cunha dos Anjos e

Gustavo de Mattos Vasques

Tese submetida como requisito parcial para obtenção do grau de **Doutora**, no Programa de Pós-graduação em Agronomia, Área de Concentração em Ciência do Solo.

Seropédica, RJ
Fevereiro de 2018

Universidade Federal Rural do Rio de Janeiro Biblioteca
Central / Seção de Processamento Técnico

Ficha catalográfica elaborada com os dados fornecidos
pelo (a) autor (a)

G633m Gomes, Andréa da Silva, 1988-
Mapeamento Digital de Estoque de Carbono e Frações
Granulométricas de Solos da Amazônia Central e Predição de
Densidade dos Solos do Brasil / Andréa da Silva Gomes. -
2018.
56 f.: il.

Orientador: Marcos Bacis Ceddia.

Coorientador: Gustavo de Mattos Vasques.

Tese(Doutorado). -- Universidade Federal Rural do Rio
de Janeiro, Pós-Graduação em Agronomia - Ciência do Solo,
2018.

1. Sensoriamento remoto. 2. Regressão linear múltipla.
3. Funções de pedotransferência. I. Ceddia, Marcos Bacis,
1968-, orient. II. Vasques, Gustavo de Mattos, 1979-
coorient. III Universidade Federal Rural do Rio de Janeiro.
Pós-Graduação em Agronomia - Ciência do Solo. IV.
Título.

É permitida a cópia parcial ou total desta Tese, desde que seja citada a fonte.

UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA – CIÊNCIA DO SOLO

ANDRÉA DA SILVA GOMES

Dissertação/Tese submetida como requisito parcial para obtenção do grau de **Doutora**, no Programa de Pós-Graduação em Agronomia, área de Concentração em Ciência do Solo.

TESE APROVADA EM 28/02/2018

Marcos Bacis Ceddia. Dr. UFRRJ
(Orientador)

Marcos Gervasio Pereira. Dr. UFRRJ

Renato Nunes Pereira. Dr. UFRRJ

Gustavo Souza Valladares. Dr. UFPI

Guilherme Kangussu Donagemma. Dr. Embrapa Solos

*"Eu sou de uma terra que o povo padece
Mas não esmorece e procura vencer.
Da terra querida, que a linda cabocla
De riso na boca zomba no sofrer
Não nego meu sangue, não nego meu nome
Olho para a fome, pergunto o que há?
Eu sou brasileiro, filho do Nordeste,
Sou cabra da Peste, sou do Ceará."*

Patativa do Assaré

*A minha família e amigos, pelo apoio, carinho e confiança,
Por serem meu porto seguro e o motivo pelo qual não me deixo vencer pelos obstáculos
encontrados no meio do caminho e sempre sigo em frente.*

Ofereço

*Á minha mãe, Antônia Cláudia, pela dedicação, preocupação, paciência e pelo amor
incondicional a mim dedicado.
Dedico*

AGRADECIMENTOS

A Deus que é conforto, guia, refúgio, me deu forças para chegar até aqui e concluir mais essa etapa de aprendizagem.

À minha família pelo apoio e por terem investido na minha educação por tantos anos, sempre compreensivos e atenciosos.

Ao professor Marcos Bacis Ceddia pela orientação nesse trabalho, pela atenção, paciência, amizade, pelo exemplo de profissional, competente, ético, responsável, por sempre ter uma boa idéia para sugerir, pelos ensinamentos, pelos risos, por ter confiado em minha capacidade e esforço.

Aos co-orientadores professores Lúcia Helena Anjos e Gustavo Vasques pelo apoio, ensinamentos e cooperação com o desenvolvimento da tese.

Ao grupo LASA, ao qual tenho o prazer e a alegria de fazer parte, por todos os bons frutos que irão surgir da união desse grupo.

Aos amigos de curso e de vida, em especial, Selenobaldo, Silvana, Carlos, Aurea, Mayan, Ana Carolina, Sandra Lima, Iara, Lílian, Mariana, Danilo, Andressa, Cícero, Camila, Carol, Amanda por todo o apoio, disponibilidade, parceria, carinho e horas agradáveis de descontração, que foram essenciais para o desencadeamento saudável deste trabalho.

Ao meu namorado, Jadér Lobato, pelo apoio, carinho, atenção, compreensão, por estar me acompanhando e incentivando meu crescimento tanto pessoal quanto profissional e a seus pais, dona Arlete e seu Renato, por terem me acolhido em sua família e confiado em meu esforço.

Ao professor Gustavo Valladares, por ter me dado a oportunidade de participar de uma iniciação científica, ter incentivado minha vinda para a UFRRJ e sempre ter depositado em mim confiança e acreditado em meu esforço.

Ao professor Mauro Antunes por ter contribuído com o conhecimento sobre sensoriamento remoto e o tratamento de imagem de satélite.

A todos os professores do PPGA-CS, pelo o conhecimento transmitido de forma tão competente, em especial aos professores Érika Pinheiro, Lúcia Helena e Marcos Gervasio.

Aos professores José Claudio Mura e Sidnei João Siqueira Sant'Anna pela ajuda, disponibilidade e atenção e por ter me possibilitado a oportunidade de participar do curso de Processamento de Imagens SAR ofertado pela Pós-Graduação do INPE, o que me conferiu maior entendimento sobre os fundamentos e processamento de imagens de radar.

Ao estagiário Michel Gomes por toda a ajuda e disponibilidade para a realização da organização do banco de dados do Brasil.

Aos funcionários da secretaria do CPGA CS, Marquinhos, Michele e Wagner pela atenção e gentileza.

Ao INPE, Ministério do Meio Ambiente e Universidade do Alasca pela disponibilidade das imagens ópticas e de radar utilizadas na tese.

A Universidade Federal Rural do Rio de Janeiro, uma instituição de tradição, com excelentes profissionais e ao CPGA CS pela oportunidade de ter estudado e conhecido todos os amigos e professores que com certeza vão fazer parte da minha história para sempre.

A CAPES pela concessão de bolsa, recurso indispensável para realização do trabalho de pesquisa.

BIOGRAFIA

Andréa da Silva Gomes, filha de Francisco Antônio Gomes e Antônia Cláudia da Silva Santos, nasceu em 10 de setembro de 1988 na cidade de Fortaleza, Estado do Ceará. Ingressou na Universidade Federal do Ceará em fevereiro de 2008 no curso de Zootecnia. Em 2010 iniciou trabalhos de iniciação científica na área de solos, atuando nos temas geoprocessamento e pedologia, sendo bolsista PIBIC/CNPq. Formou-se em 2011 e em março de 2012 iniciou o mestrado no Curso de Pós-graduação em Agronomia – Ciência do Solo pela Universidade Federal Rural do Rio de Janeiro, onde desenvolveu a dissertação na área de pedometria/mapeamento digital de solos na Amazônia Central. Em março de 2014 iniciou o doutorado no mesmo curso e mesma instituição de ensino, onde continua desenvolvendo trabalhos na área de pedometria/mapeamento digital de solos, incluindo sensoriamento remoto. Neste mesmo ano, em maio, iniciou o curso de Pós-Graduação *latu sensu* (Especialização) em Estatística Aplicada, concluindo em novembro de 2015.

RESUMO GERAL

GOMES, Andréa da Silva. **Mapeamento digital de estoque de carbono e frações granulométricas de solos da Amazônia Central e predição de densidade dos solos do Brasil**. 2018. 56f. Tese (Doutorado em Agronomia, Ciência do Solo). Instituto de Agronomia, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2018.

O conhecimento do solo e suas propriedades é essencial para o planejamento ambiental e estudo da dinâmica em agroecossistemas. Algumas regiões, como Amazônia brasileira, apresentam carência em informações detalhadas sobre o meio físico, devido a dificuldades referentes ao acesso e movimentação em mata densa típica dessas regiões. Atributos de difícil mensuração em campo como densidade do solo, não são determinados comumente em levantamentos de solos e são necessários para o cálculo de outras propriedades como estoque de carbono e porosidade total. Nesse contexto, modelos de predição e algoritmos de interpolação são alternativas para a estimativa desses atributos em áreas não amostradas. Os objetivos deste trabalho foram: avaliar a adição de covariáveis do relevo e derivadas de imagens multiespectrais e de radar na predição de estoque de carbono e textura do solo em área da Amazônia Central, comparando a performance de quatro métodos geoestatísticos; e desenvolver Funções de Pedotransferência (FPTs) para estimar densidade do solo (Ds) em solos brasileiros, utilizando outros atributos físicos e químicos. No Capítulo I foram utilizadas variáveis do relevo e derivadas de imagens multiespectrais (NDVI) e de radar (coeficiente de retroespalhamento) como covariáveis na predição dos atributos estoque de carbono e textura do solo. Em geral, os resultados mostraram que, mesmo sob floresta densa, o Índice de Vegetação de Diferença Normalizada (NDVI) e o coeficiente de retroespalhamento ALOS PALSAR melhoraram a acurácia das predições de estoque de carbono (EC) e argila subsuperficial. O NDVI derivado de imagem do sensor RapidEye melhorou a predição de EC usando cokrigagem isotópica, enquanto o NDVI derivado de imagem do sensor Landsat 8 e o coeficiente de retroespalhamento foram selecionados para predição de EC e argila subsuperficial usando o krigagem com regressão (KR). A melhoria relativa da aplicação de cokrigagem e KR sobre o krigagem ordinária foi inferior a 10%, indicando que são necessárias análises futuras para conectar covariáveis derivadas da vegetação e do relevo com atributos do solo. No Capítulo II foram desenvolvidas FPTs para estimar densidade dos solos do Brasil, a partir de atributos físicos e químicos disponíveis em levantamentos de solo. Foram desenvolvidos modelos de regressão linear múltipla para o Brasil e para as cinco regiões e comparado sua performance com o modelo referência publicado na literatura proposto por Benites et al. (2007). As FPTs geradas tiveram melhor performance comparada ao modelo proposto por Benites et al. (2007), exceto para a região Sul. Comparando-se a performance do modelo geral (Brasil) em relação aos modelos gerados por regiões, observou-se menores valores de MPE, MAE e RMSPE, para os modelos gerados por regiões, exceto para a região Sul. Quando se considera um modelo geral para prever a Ds, o modelo gerado para o Brasil, obteve melhor performance, aproximadamente 7% a mais de acurácia comparado ao modelo referência.

Palavras-chave: SAR. NDVI. Relevo. Sensoriamento remoto. Regressão linear múltipla. Funções de pedotransferência.

GENERAL ABSTRACT

GOMES, Andréa da Silva. **Digital mapping of carbon stock and particle size fractions of soils from Central Amazonia and prediction of bulk density of Brazilian soils.** 2018. 56p. Thesis (Doctor in Agronomy, Soil Science). Instituto de Agronomia, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2018.

The knowledge of the soil and its properties is essential for the environmental planning and study of the dynamics in agroecosystems. Some regions, such as the Brazilian Amazon, lack detailed information about the physical aspects due to difficulty of access and movement in the dense forest. Attributes complicated to measure in the field such as soil density commonly are not determined in soil surveys, although they are necessary for calculation of properties such as carbon stock and total porosity. In this context, prediction models and interpolation algorithms are alternatives for estimating these attributes in non-sampled areas. The objectives of this work were: to evaluate the addition of relief covariates, and those derived from multispectral and radar images, in the prediction of carbon stock and soil texture in a Central Amazonia area, comparing the performance of four geostatistical methods; and to develop Pedotransfer Functions (PTFs) to estimate the bulk density of Brazilian soils using physical and chemical soil attributes. In Chapter I, relief variables derived from multispectral images (NDVI) and radar (backscattering coefficient) were used as covariates in the prediction of carbon stock and soil texture attributes. In general, the results showed that, even under the dense forest, the Normalized Difference Vegetation Index (NDVI) and the ALOS PALSAR backscattering coefficient improved the accuracy of carbon stock (CS) and the clay predictions for the soil subsurface. The NDVI derived from RapidEye sensor improved the prediction of CS using isotopic cokriging, while the NDVI derived from Landsat 8 and backscattering coefficient were selected to predict clay content at the subsurface using regression kriging (RK). The relative improvement of applying cokriging and RK over ordinary kriging were lower than 10%, indicating that further analyses are necessary to connect soil proxies (vegetation and relief types) with soil attributes. In Chapter II, PTFs were developed to estimate bulk density in Brazilian soils, based on physical and chemical attributes available in soil surveys. Multiple linear regression models were developed for Brazil and for the five regions and compared their performance with the reference model published in the literature proposed by Benites et al. (2007). The generated PTFs performed better than the model proposed by Benites et al. (2007), except for the southern region. Comparing the performance of the general model (Brazil) in relation to the models generated by regions, we observed smaller values of MPE, MAE and RMSPE, for the models generated by regions, except for the South region. When considering a general model to predict bulk density, the model generated for Brazil obtained better performance, approximately 7% more accuracy than the reference model.

Key Words: SAR. NDVI. Relief. Remote sensing. Multiple linear regression. Pedotransfer functions.

SUMÁRIO

1	INTRODUÇÃO GERAL	1
2	REVISÃO DE LITERATURA	3
	2.1 Sensoriamento Remoto e Análise Ambiental.....	3
	2.2 Estoque de Carbono e Textura do Solo.....	4
	2.3 Métodos de Predição Geoestatísticos e Híbridos.....	6
	2.4 Funções de Pedotransferência (FPTs).....	8
3	CAPÍTULO I SOIL CARBON STOCK AND PARTICLE SIZE FRACTIONS IN THE CENTRAL AMAZON PREDICTED FROM REMOTELY SENSED RELIEF, MULTISPECTRAL AND RADAR DATA	11
	3.1 RESUMO	12
	3.2 ABSTRACT	13
	3.3 INTRODUCTION	14
	3.4 MATERIALS AND METHODS	16
	3.4.1 Study Area.....	16
	3.4.2 Soil Data	18
	3.4.3 Remote Sensing Covariate Data	18
	3.4.4 Prediction Methods	19
	3.5 RESULTS	21
	3.5.1 Descriptive Statistics and Linear Regression Models.....	21
	3.5.2 Geostatistical Analysis	23
	3.6 DISCUSSION	29
	3.6.1 Potential of Using Multispectral and Radar Data as Covariates.....	29
	3.6.2 Performance of the Different Prediction Methods.....	30
	3.7 CONCLUSIONS	32
	3.8 REFERENCES	33
4	CAPÍTULO II FUNÇÃO DE PEDOTRANSFERÊNCIA PARA ESTIMAR DENSIDADE DE SOLOS DO BRASIL	36
	4.1 RESUMO	37
	4.2 ABSTRACT	38
	4.3 INTRODUÇÃO	39
	4.4 MATERIAL E MÉTODOS	40
	4.4.1 Banco de Dados	40
	4.4.2 Desenvolvimento das FPTs.....	40
	4.4.3 Avaliação da Performance das FPTs	42
	4.5 RESULTADOS E DISCUSSÃO	43
	4.5.1 Estatística Descritiva do Banco de Dados	43
	4.5.2 FPTs Geradas	45
	4.5.3 Avaliação do Desempenho das FPTs	46
	4.6 CONCLUSÕES	50
5	CONCLUSÕES GERAIS	51
6	REFERÊNCIAS BIBLIOGRÁFICAS	52

1. INTRODUÇÃO GERAL

A região da Amazônia Legal do Brasil abrange uma grande diversidade de vegetação e solos. O território abrange uma área de 5.217.423 km² (59% do total do território brasileiro) e, apesar de sua importância estratégica, o conhecimento sobre seus solos é relativamente baixo. O único projeto sistemático para mapear a região da Amazônia Legal (O projeto RADAMBRASIL), foi gerado entre os anos de 1973 e 1984. Este projeto é a principal fonte de dados usada por pesquisadores para desenvolver modelos e mapas de atributos do solo. No entanto, de acordo com Ceddia et al. (2015), os mapas de solo foram publicados em escala muito pequena e os dados de solo disponíveis são muito espaçados ao longo da região Amazônica (~0,7-3,5 perfis de solo a cada 10.000 km²). A principal razão para a baixa densidade de dados do solo é a presença da Floresta Amazônica, uma barreira para os pedólogos e fonte de potenciais doenças tropicais, como leishmaniose, malária e febre amarela, e a falta de vias de acesso (baixa densidade de rotas, maior parte do território só é acessível por meio de barcos ou avião).

Informações detalhadas sobre atributos do solo são de extrema relevância para a gestão de recursos da terra, monitoramento de impactos ambientais, desenvolvimento de práticas de manejo sustentáveis na agricultura e implantação de sistemas agrícolas. A textura do solo é considerada um fator importante por influenciar muitas propriedades físicas e químicas e padrões observados no solo, tais como armazenamento de água, capacidade de troca de cátions, fertilidade e drenagem interna do perfil (Manrique et al., 1991; Liao et al., 2013). Diante do cenário de mudanças climáticas e a preocupação com a mitigação da emissão de gases de efeito estufa, o estudo da variabilidade espacial de estoque de carbono orgânico em solos tropicais se faz necessário e constitui uma alternativa para essa problemática.

Métodos de predição de propriedades do solo a partir de dados de sensoriamento remoto vêm sendo bastante empregados em ciência do solo. Os modelos de predição mais comumente utilizados na literatura geralmente envolvem métodos estatísticos (regressão linear múltipla, árvore de regressão), interpolação geoestatística (krigagem ordinária e co-krigagem) e técnicas híbridas (*regression kriging*) (Bishop & McBratney 2001; Lark & Bishop, 2007; Sumfleth & Duttmann 2008; Liao et al. 2013; Ahmed & Iqbal 2014). A abordagem de regressão é relativamente simples, mas não faz nenhuma tentativa de explicar a dependência espacial dos dados, uma vez que depende, exclusivamente da relação entre o atributo do solo de interesse e os atributos selecionados como preditores (Mueller & Pierce 2003). Krigagem é uma técnica geoestatística básica que fornece a melhor estimativa sem tendência para uma variável espacialmente dependente (Liao et al., 2013). Co-krigagem incorpora informações secundárias em estimativas espaciais de uma forma que explica a correlação espacial cruzada entre as variáveis primárias e secundárias (Goovaerts, 1997). *Regression kriging* abrange duas abordagens: primeiro utiliza a regressão para prever uma variável e depois usa krigagem simples para interpolar os resíduos do modelo de regressão (Hengl et al., 2004).

Dados de sensoriamento remoto e atributos do relevo são muitas vezes utilizados como variáveis auxiliares nesses modelos de predição, pois são dados de obtenção barata e podem ser expandidas para grandes áreas (Bishop & McBratney, 2001). Imagens de sensoriamento remoto tomadas no domínio óptico são informativos de propriedades da superfície da terra que podem estar relacionados às características do solo (Maselli et al. 2008). Esses dados são utilizados para derivar propriedades do solo, além de serem utilizados como uma fonte de dados de suporte de mapeamento digital de solos (Ben-Dor et al., 2008).

Ainda, facilitam o mapeamento de áreas inacessíveis, reduzindo a necessidade de levantamentos de campo caros e demorados (Mulder et al., 2011). Atributos do terreno (por exemplo, altitude, aspecto, inclinação, curvatura) podem ajudar na estimativa espacial das propriedades do solo porque o relevo tem grande influência nos processos de formação do solo (Jenny, 1941; Mueller & Pierce, 2003). Esses processos ocorrem em muitas paisagens em resposta à maneira como a água se move através da paisagem. Além disso, os atributos do terreno podem caracterizar esses caminhos de fluxo e as interações com os atributos do solo (Dobos et al., 2000). Dados de sensoriamento remoto são muito influenciados pela variabilidade do terreno, por isso tem que ser complementado com informações sobre o terreno que podem ser usadas para corrigir as distorções dos dados de satélites decorrentes das variações topográficas da paisagem e fornecer dados adicionais para a modelagem solo-paisagem (Dobos et al., 2000).

A combinação de dados espectrais e atributos do relevo na predição de propriedades do solo, como estoque de carbono e frações granulométricas, apresenta-se como alternativa viável e promissora, uma vez que, no caso da área de estudo (Amazônia Central), existe dificuldade de acesso e falta de informações detalhadas sobre o meio físico.

Ainda nesse contexto, diante da escassez de dados medidos de densidade do solo, principalmente nas regiões norte e nordeste, e de sua relevância para o cálculo de outros atributos, como estoque de carbono, Funções de Pedotransferência (FPTs) vem sendo desenvolvidas com o intuito de suprir essa escassez. Vários algoritmos têm sido desenvolvidos e os resultados mostram-se bastante promissores quando os modelos são utilizados de forma adequada. Sendo essa uma alternativa viável para situações onde não é possível a medição desse atributo em campo.

Diante disso, a hipótese deste trabalho é que o uso de dados de relevo e sensoriamento remoto (sensores multispectrais e radar) pode melhorar a acurácia da predição de atributos do solo sob Floresta Amazônica. É possível gerar FPTs para estimar densidade de solos do Brasil utilizando atributos físicos e químicos do solo.

Os objetivos específicos foram:

Capítulo I: Soil carbon stock and particle size fractions in the central amazon predicted from remotely sensed relief, multispectral and radar data.

- a) Avaliar a adição de co-variáveis do relevo e de vegetação derivadas de imagens multispectrais com distintas resoluções espaciais e espectrais (Landsat 8 e RapidEye) e radar banda L (ALOS PALSAR), para a predição de estoque de carbono e textura do solo;
- b) Avaliar o desempenho de quatro métodos geoestatísticos para mapear a variabilidade espacial destas propriedades do solo (Krigagem ordinária, co-krigagem isotópica e heterotópica, krigagem com regressão).

Capítulo II: Função de pedotransferência para estimar densidade de solos do Brasil.

- a) Desenvolver FPTs para estimar a densidade de solos do Brasil empregando atributos físicos e químicos disponíveis em levantamentos de solos;
- b) Analisar se a subdivisão do banco de dados de acordo com as cinco regiões (Norte, Nordeste, Centro-Oeste, Sudeste e Sul) gera modelos mais acurados;
- c) Comparar o desempenho dos modelos gerados com o proposto por Benites et al. (2007) que é referência na predição de Ds para o Brasil.

2. REVISÃO DE LITERATURA

2.1 Sensoriamento Remoto e Análise Ambiental

A preocupação com as questões ambientais vem crescendo muito nos últimos anos, o que impulsionou o desenvolvimento de tecnologias para estudo da dinâmica dos processos naturais e a recuperação dos danos causados pela ação antrópica. O sensoriamento remoto mostra-se uma ferramenta importante na avaliação dessas mudanças, principalmente quanto às variações nas propriedades edáficas devido à mudança do uso ao longo do tempo (Lopes et al., 2010). Como a atualização dos dados sobre o uso e cobertura do solo é frequente, possibilitada pelo monitoramento da superfície da terra por satélites, essa informação tem se destacado pela confiabilidade e precisão dos dados gerados (Tornquist et al., 2009).

Para a avaliação da evolução da vegetação, o sensoriamento remoto permite o uso de índices espectrais que facilitam o reconhecimento de feições da vegetação de acordo com a fase do desenvolvimento ou variações ao longo do ano onde esta vegetação está inserida.

Dentre os diversos índices de vegetação, o mais conhecido e utilizado na literatura é o NDVI (*Normalized Difference Vegetation Index*) ou Índice de Vegetação por Diferença Normalizada. O NDVI é expresso pela razão entre a diferença da medida da reflectância nos canais do infravermelho próximo e vermelho e a soma desses canais. Os valores de NDVI variam de -1 a +1, quanto mais próximo de 1, maior a densidade da cobertura vegetal, ou seja, ela apresenta-se em seu estágio denso, úmida e bem desenvolvida (Melo et al., 2011). Este índice é bastante recomendado quando se pretende fazer comparações ao longo do tempo de uma mesma área, pois é esperado que seja menos influenciado pelas variações das condições atmosféricas (Meneses & Almeida, 2012).

Outro índice bastante conhecido é o RVI (*Ratio Vegetation Index*), ele consiste na divisão da banda situada no infravermelho próximo (NIR), região que a vegetação mostra a mais intensa reflectância, pela banda situada no vermelho (R), onde a vegetação possui a mais alta absorção da luz solar visível devido à presença do pigmento verde clorofila, indicativo de uma vegetação sadia (Meneses & Almeida, 2012). Os valores absolutos de RVI são sempre maiores que zero e assumem valores positivos altos em áreas de densa cobertura vegetal. Quanto mais claro o tom de cinza, maior a densidade de vegetação. Tons de cinzas médios indicam pouca vegetação e tons de cinza totalmente escuros indicam ausência de vegetação, o que só pode ocorrer nas áreas de corpos d'água, de solo exposto, áreas cobertas por neve, ou nuvens.

O EVI (*Enhance Vegetation Index*) ou índice de vegetação realçado (EVI) é um índice mais sensível às variações na atividade fotossintética e estrutura do dossel vegetal, ao mesmo tempo em que é mais resistente à contaminação atmosférica e menos suscetível à saturação (Huete, 2006).

Diversos trabalhos foram desenvolvidos mostrando os benefícios do uso de sensoriamento remoto na evolução espaço-temporal do uso do solo e da vegetação. Leite et al. (2012) buscaram analisar ao longo das últimas duas décadas, através de imagens de satélite, a dinâmica no uso do solo da bacia do rio Pacuí. Demarchi et al. (2011) analisaram a evolução temporal do uso do solo no município de Santa Cruz do Rio Pardo – SP por meio da classificação supervisionada de imagens de satélite Landsat-5 TM e compararam os índices de vegetação NDVI e SAVI nas datas 26-05-2007, 07-01-2009 e 29-04-2009. Melo et al. (2011) apresentaram um método de análise da cobertura fitogeográfica da microbacia do Riacho dos Cavalos através do Índice de Vegetação por Diferença Normalizada (NDVI). Lopes et al. (2010) avaliaram, por meio de técnicas de sensoriamento remoto, as modificações na

cobertura do solo decorrentes de atividades antrópicas, em uma área piloto ao norte envolvendo parte da zona urbana de Vacaria, RS.

Além do uso de imagem de satélite (imagem óptica) para a análise ambiental, a imagem de radar também tem sido usada com esse propósito. Sensores ópticos possuem algumas limitações como: aquisição de imagens ópticas é dependente de ausência de cobertura de nuvens ou de fumaça durante a passagem do satélite; e a capacidade de penetração da radiação eletromagnética (REM) incidente é restrita à porção superior do dossel ou dos solos; a interpretação de dados ópticos precisa levar em consideração a complexa interação entre a REM e a cobertura vegetal e de solos, o sombreamento interno dos alvos e a interferência atmosférica (Moran et al., 2002; Sano et al., 2009).

O uso de imagem de radar de abertura sintética (SAR) corresponde a uma alternativa diante dessas limitações. Sensores de radar de abertura sintética (SAR) são sensores ativos, ou seja, possuem fonte própria de REM que operam na faixa espectral de microondas e que permitem discriminar alvos ou feições no terreno com base nas diferenças nas suas propriedades dielétricas (conteúdo de água nos solos e nas plantas) e geométricas (rugosidade do terreno e estrutura da vegetação) (Paradella et al., 2005; Sano et al., 2009). Possibilitam ainda a obtenção de imagens independentemente da presença de nuvens, portanto pode-se adquirir imagens em qualquer época do ano, tem capacidade de operação durante o dia ou a noite e independência em relação às condições de iluminação solar (Paradella et al., 2005). A principal desvantagem é a falta de um melhor entendimento sobre os processos físicos que efetivamente ocorrem quando os sinais de radar incidem num alvo geológico, agrícola ou com vegetação natural (Chaves et al., 2003).

Pesquisas tem sido desenvolvidas para abordar o uso de dados de sensores SAR na análise ambiental. Sano et al. (2009) avaliaram o potencial das imagens do satélite ALOS PALSAR (banda L, polarizações HH, HV e VV) para discriminar as classes de uso e cobertura da terra mais representativas do Distrito Federal e concluíram que o uso dessas imagens mostraram-se eficientes para a identificação de classes de uso e cobertura da terra representativas da região. Neta et al. (2010) avaliaram a discriminação do incremento de desflorestamento (área desflorestada por corte raso há no máximo um ano) na Amazônia Legal a partir de dados multipolarizados obtidos em banda L pelo radar ALOS/PALSAR e concluíram que o uso de imagens desse sensor foi eficiente na discriminação do incremento de desflorestamento. Freitas et al. (2014) analisaram o potencial das imagens ALOS/PALSAR (banda L, polarizações HH, HV e VV) para discriminar as diferentes classes de cobertura de terras representativas da região do Médio Taquari-MS e os resultados desse estudo demonstraram que é possível discriminar essas classes de cobertura vegetal natural e antrópica com base nas imagens de radar do satélite ALOS/PALSAR da estação seca.

2.2 Estoque de Carbono e Textura do Solo

Diante das mudanças climáticas, aquecimento global devido ao aumento da emissão de gases do efeito estufa (GEEs), estudos que desenvolvem métodos de mitigar essa emissão vem sendo cada vez mais necessários no cenário mundial. A ação antrópica comumente é relatada como fator relevante no aumento da emissão de GEEs, por exemplo, a agricultura introduzida após desmatamento tem fortes impactos, aumentando a emissão desses gases para a atmosfera, conforme relata Machado (2005).

Ainda segundo esse autor, o desmatamento aliado a introdução de sistemas agrícolas/pastagens aumenta a emissão de GEEs devido às práticas de manejo, em que se tem o desmatamento seguido da queimada e o preparo do solo por aração seguida de sucessivas gradagens niveladoras para preparar a superfície do solo para receber as sementes, o que também favorece o processo de erosão hídrica do solo.

O uso de geoestatística e sensoriamento remoto na elaboração de mapas, tanto em escala regional ou global, vem crescendo muito no meio científico e bastante informação tem sido gerada aplicando essas ferramentas. Esses estudos são importantes, pois permitem um melhor entendimento da dinâmica do carbono em agroecossistemas, além de possibilitar a previsão deste atributo ao longo dos anos.

Bernoux et al. (2002) geraram mapa de estoque de carbono (EC) através de associações entre vegetação e solo, em que o mapa final de EC foi gerado pela interseção dos mapas de solos e de vegetação. Cerri et al. (2007) elaboraram mapa de estimativa de estoque de carbono, usando o sistema de modelagem GEFSOC. Esse sistema de modelagem requer seis classes de dados básicos para construir o conjunto de dados necessário para uma simulação regional: vegetação nativa, histórico de manejo e uso do solo, clima, solos e latitude/longitude. Chig et al. (2008) utilizaram abordagem diferente das citadas, com avaliação de dependência espacial e krigagem ordinária. Os resultados obtidos por esses autores mostraram que houve dependência espacial para o atributo carbono orgânico, sendo assim, foi possível a interpolação de dados não amostrados pelo método da krigagem.

O sensoriamento remoto permite a observação, análise e estudos de áreas extensas de uso/cobertura de uma forma geral e com custos financeiros e menor tempo do que apenas com campanhas de campo (Nunes et al., 2012), que são otimizadas quando ocorre essa etapa como fase inicial dos estudos. Com a aquisição de imagens de satélites e aplicação de métodos de aritmética de bandas é possível a geração de índices de vegetação que auxiliam na análise de evolução da vegetação e, assim, pode-se inferir sobre áreas que potencialmente podem sequestrar mais carbono e áreas que estão sendo desmatadas ou queimadas, logo contribuindo para emissão de GEEs. Essas técnicas também vêm sendo bastante usadas em pesquisas e os resultados mostram a eficiência e êxito do seu uso para o planejamento ambiental.

Informações detalhadas sobre as propriedades do solo associadas a informações climáticas desempenham um importante papel no planejamento do uso da terra. A textura do solo é uma propriedade física fundamental e que influencia fortemente muitas outras propriedades do solo que são altamente relevantes para a produção agrícola, a saber, fertilidade e capacidade de retenção de água.

Modelos de predição de classe e atributos do solo, dentre eles, estoque de carbono e textura, vêm sendo amplamente utilizado em ciência do solo. Os métodos mais comumente utilizados na predição de propriedades do solo usando dados de sensoriamento remoto e covariáveis ambientais são regressão linear múltipla, regressão por quadrados mínimos parciais, métodos híbridos e geoestatísticos (Chagas et al., 2016).

O sensoriamento remoto permite determinar o comportamento espectral de alvos terrestres, como o solo por exemplo. Isso é possível graças a interação da radiação eletromagnética e o alvo que pode ser medida por meio de sensores (tanto terrestres quanto orbitais). A possibilidade de estimativa de atributos do solo utilizando sua curva espectral está baseada na intensidade da curva, assim como nas zonas de absorções de energia eletromagnética em pontos específicos ou em intervalos da curva, já que diferentes quantidades de certos atributos presentes no solo podem aumentar ou diminuir a reflectância espectral em diferentes comprimentos de onda do espectro eletromagnético (Cezar et al. 2013).

Viscarra-Rossel et al. (2006) fizeram predições simultâneas de diferentes propriedades do solo (pH, C orgânico, argila, silte, areia, CTC, Ca, P, K) em diferentes regiões do espectro eletromagnético, concluindo que a região do infravermelho médio foi a mais adequada para este tipo de análise, devido às bandas espectrais de incidência serem mais elevadas nesta região, bem como a intensidade mais elevada, além da especificidade do sinal.

Djuuna et al. (2011) fizeram a predição de algumas propriedades do solo (pH, C total, N total, C/N, NH₄-N, NO₃-N, P, K, argila, silte, areia e carbono da biomassa microbiana)

usando infravermelho médio (mid-IR) – método de mínimos quadrados parciais (MQP). Os autores concluíram que o método mid-IR – MQP pode ser usado na predição de propriedades do solo, baseado nas calibrações de valores mid-IR.

2.3 Métodos de Predição Geoestatísticos e Híbridos

A Geoestatística é definida como um tópico especial da estatística aplicada que trata de problemas referentes às variáveis regionalizadas, as quais têm um padrão espacial mostrando características intermediárias entre as variáveis verdadeiramente aleatórias (apresentam irregularidades e variação imprevisível de um ponto para outro) e as totalmente determinísticas (apresentam relações existentes entre os pontos no espaço motivadas pela sua gênese) (Landim, 2006). Portanto, a análise geoestatística se aplica quando uma determinada propriedade ou atributo varia de um local para outro com certo grau de organização ou continuidade, expresso pela dependência espacial (Vieira, 2000).

O uso da estatística clássica em experimentos da ciência do solo deve ser feita de forma cuidadosa, pois para alguns atributos a hipótese de independência estatística ou distribuição espacial aleatória não se aplica. Na maioria dos casos o que se tem observado é a existência de dependência espacial, ou seja, o atributo de um fenômeno varia com a distância e a direção e nesse caso uso de geoestatística é mais apropriado (Valencia et al., 2004).

A geoestatística se baseia na determinação e modelagem do semivariograma experimental, através do qual se quantifica a variabilidade espacial entre as amostras (Gonçalves et al. 2005). O semivariograma é definido pela equação 1:

$$\gamma(h) = (1/2N(h)) E \{ [Z(x_i) - Z(x_i+h)]^2 \} \quad \text{Eq. 1}$$

Em que:

- $\gamma(h)$ – é a semivariância em função do vetor h ;
- E – valor esperado;
- $Z(x_i)$ – valor da variável na posição x_i
- $Z(x_i+h)$ - valor da variável na posição x_i+h .
- N – Número de pares distantes em um vetor h

Os parâmetros numéricos do semivariograma (Figura 1) são importantes para o ajuste de um modelo teórico que melhor represente a distribuição espacial dos dados e para a etapa de krigagem (estimação de dados em locais não amostrados). O efeito pepita (C_0) revela a descontinuidade do semivariograma para distâncias menores do que a menor distância entre as amostras. O patamar ($C_0 + C_1$) é aproximadamente igual a variância dos dados. O alcance (a) é a distância na qual $\gamma(h)$ atinge o patamar e é a distância limite de dependência espacial, a partir da qual tem distribuição espacial aleatória. A contribuição (C_1) é a variância estrutural, ou seja, corresponde a variação modelada pelo semivariograma e é calculada pela diferença entre o patamar e o efeito pepita (Vieira, 2000; Valencia et al., 2004).

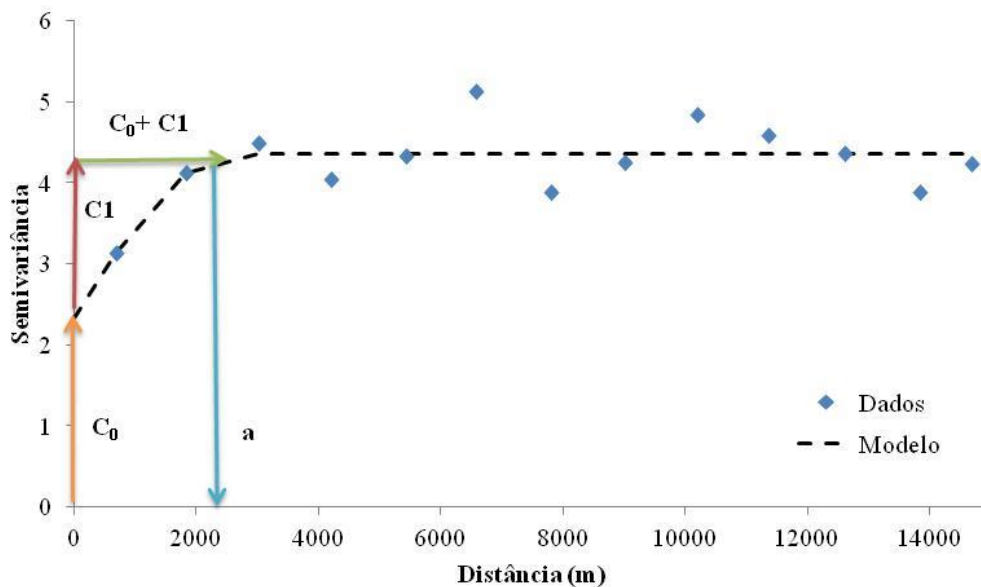


Figura 1. Parâmetros do semivariograma experimental (Autoria própria).

O semivariograma é necessário para a verificação da existência de dependência espacial, e quando detectada, é possível o ajuste de um modelo teórico que melhor represente a distribuição espacial dos dados (linear, esférico, exponencial, gaussiano). O ajuste de um modelo teórico ao semivariograma experimental é um dos pontos mais relevantes das aplicações da Teoria das Variáveis Regionalizadas e acaba sendo uma das maiores fontes de ambiguidade e polêmica nessas aplicações, pois se o modelo ajustado estiver errado todos os cálculos das próximas etapas da análise também estarão (Vieira, 2000).

Com os parâmetros do semivariograma com o modelo ajustado, é possível se estimar valores para locais não amostrados através do método de krigagem. Landim (2006) define krigagem (Eq. 2) como sendo o processo de estimativa de valores de variáveis distribuídas no espaço, e/ou no tempo, a partir de valores adjacentes enquanto considerados como interdependentes pelo semivariograma.

$$Z^*(x_0) = \sum_{i=1}^N \lambda_i z(x_i) \quad \text{Eq. 2}$$

Em que:

$z^*(x_0)$ – valor a ser estimado;

N – número de valores medidos, $z(x_i)$, envolvidos na estimativa; e.

λ_i – são os pesos associados a cada valor medido, $z(x_i)$.

Segundo Vieira (2000), no método da krigagem, os pesos são variáveis e dependem da variabilidade espacial expressa no semivariograma. Esse estimador nada mais é que uma média móvel ponderada, o que torna a krigagem um interpolador ótimo (sem tendência e com variância mínima), então, é a forma como os pesos são distribuídos. Esse método de interpolação vem sendo bastante aplicado em estudos de variabilidade espacial de atributos do solo.

A cokrigagem é outro método de interpolação utilizado na geoestatística, sendo bastante eficiente quando se pretende estimar uma variável que possui poucos dados, por ser de difícil ou laboriosa medição, através de outra variável com um maior número de informações. Para isso, essas variáveis devem apresentar correlação entre si e, caso exista dependência espacial com o semivariograma cruzado, é possível estimar valores para um atributo com dados de outros atributos correlacionados. A cokrigagem não requer que a

variável secundária esteja disponível em todos os pontos de amostragem. Quando as variáveis, principal e secundária, foram medidas em todos os pontos de amostragem, esse é o caso da cokrigagem isotópica. Quando as variáveis são medidas em pontos diferentes, a cokrigagem heterotópica pode ser usada.

O semivariograma cruzado é calculado pela Equação 3:

$$\gamma_{12} * (h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \{ [z_1(x_i + h) - z_1(x_i)] [z_2(x_i + h) - z_2(x_i)] \} \quad \text{Eq. 3}$$

O semivariograma cruzado ideal teria a mesma aparência mostrada na equação acima, porém, com significados diferentes pelo simples fato de envolver o produto das diferenças de duas variáveis distintas. O alcance neste caso representa o final ou a distância máxima de dependência espacial entre as variáveis. O patamar, caso exista, deve aproximar-se do valor da covariância entre as duas variáveis. Assim, quando as duas forem de correlação inversa o semivariograma cruzado será negativo (Vieira, 2000).

A krigagem com regressão (*Regression kriging*) é um método geoestatístico híbrido, pois abrange duas abordagens: primeiro utiliza a regressão para prever uma variável e depois usa krigagem simples para interpolar os resíduos do modelo de regressão (Hengl et al., 2004). Resumindo, o método de *regression kriging* é dado pela Equação 4:

$$z(s_0) = \sum_{k=0}^p \beta_k \cdot q_k(s_0) + \sum_{i=1}^n \lambda_i \cdot e(s_i) \quad \text{Eq. 4}$$

Em que:

β_k = coeficientes estimados na regressão

$q_k(s_0)$ = valores das variáveis independentes

λ_i = pesos da krigagem determinada pela estrutura da dependência espacial do resíduo

$e(s_i)$ = resíduo na posição s_i

Alguns tipos de regressão com krigagem são citados na literatura e comparados a outros métodos de predição. A regressão com krigagem tipo A se dá por uma regressão normal seguida pela krigagem dos valores preditos. A regressão com krigagem tipo B envolve um modelo de regressão normal seguido pela krigagem dos valores preditos e dos resíduos da regressão. A krigagem com regressão tipo C é similar a krigagem com regressão tipo B, no entanto, só é feita a krigagem dos resíduos da regressão (Odeh et al., 1994; Odeh et al., 1995).

Os métodos geoestatísticos híbridos têm a capacidade de superar as deficiências dos métodos univariados como a krigagem ordinária; por exemplo, e têm sido recomendados para prever as propriedades do solo em casos onde os processos de formação do solo são complexos e onde os impactos antropogênicos alteraram severamente a paisagem (Rivero et al., 2007). Alguns trabalhos têm mostrado que a krigagem com regressão tem melhor performance quando comparada a krigagem ordinária, cokrigagem e regressão linear múltipla (Odeh et al., 1995; Herbst et al., 2006; Sumfleth & Duttman, 2008). Vale ressaltar que o método de krigagem com regressão apresenta algumas limitações que estão relacionadas à qualidade dos dados; poucos dados de amostragem; extrapolação da predição para fora da área amostrada; preditores que não estão bem relacionados com a variável principal (Hengl et al., 2007).

No entanto, ainda são poucos os trabalhos que trazem a utilização destes métodos, havendo a necessidade do desenvolvimento de pesquisas que testem a acurácia destes métodos em comparação aos mais utilizados como a krigagem ordinária e co-krigagem.

2.4 Funções de Pedotransferência (FPTs)

A expressão Função de Pedotransferência foi criada por Bouma (1989) e significa “*transformando dados que possuímos em dados que necessitamos*”. Utilizando essas funções

é possível se estimar dados de difícil mensuração, laboriosa e de custo elevado, a partir de informações de menor custo e fáceis de serem obtidas, como dados morfológicos (através de árvores de regressão) e granulométricos. Segundo Minasny et al. (2003), quando a pedotransferência é expressa por meio de equações matemáticas, ela passa a ser designada por Funções de Pedotransferência – FPT (*Pedotransfer Functions*).

O desenvolvimento de novas FPTs é uma tarefa árdua (Minasny et al., 2003), pois necessita de ampla base de dados de solos contendo muitos dados morfológicos e determinações de laboratório. A princípio esses autores recomendam que se utilize FPTs já desenvolvidas, mas atentando para o fato de que a validade de uma determinada FPT não deve ser interpolada ou extrapolada além do ambiente pedológico ou tipo de solo em que foi desenvolvida, e que, a aplicação direta de uma FPT em solos com propriedades distintas só deve ser feita após o teste de seu desempenho, para que FPTs mais adequadas possam ser identificadas.

Com a evolução do uso e desenvolvimento de FPTs, surgiram e foram discutidos alguns problemas relacionados à aplicação das funções (Benites et al., 2007). A precisão das FPTs é limitada, apesar da utilização de preditores potencialmente úteis e uso de ferramentas sofisticadas de processamento de dados. A aplicabilidade das FPTs também é limitada, pois funções desenvolvidas para uma região ou com determinado banco de dados não podem ser aplicadas para condições bioclimáticas diferentes sem serem validadas (Williams et al., 1992; Tietje & Tapkenhinrichs, 1993; Wosten et al., 2001). Uma explicação hipotética para as limitações das FPTs, é que elas não descrevem a estrutura do espaço poroso e, portanto, não representam bem as relações entre estrutura e função do solo. Os preditores típicos nas FPTs, textura do solo, densidade do solo, teor de carbono orgânico, estão relacionados com a estrutura do solo em um sentido amplo, mas não são suficientes para caracterizar a porosidade do solo (tamanho e geometria).

Segundo Minasny et al. (2003), para evitar o mau uso do conceito de pedotransferência, foram definidos dois princípios de FPTs: eficiência e incerteza. O princípio da eficiência é “*Não prognosticar algo que é mais fácil e mais barato medir ou determinar*”. As FPTs têm como objetivo estimar propriedades que são difíceis, laboriosas ou de custo elevado para determinação, então, é necessário que os preditores sejam de obtenção mais fácil e barata. Caso contrário, não seria vantajosa a utilização das FPTs. O princípio da incerteza é “*Não usar FPTs, a menos que se possa avaliar a incerteza associada e, para um determinado problema, se um conjunto de FPTs alternativas está disponível, usar aquela com menor variância*”. Isso significa que é mais conveniente escolher a função que tenha o menor erro de variância ou que se ajuste melhor ao tipo de solo e que a incerteza das FPTs deve ser quantificada, assim, a estimativa das variáveis preditas será a mais realista possível.

A densidade do solo tem reconhecida aplicação para cálculo de atributos dos solos (porosidade total, estoque de carbono e armazenamento de água), bem como a execução de transformações de massa para volume e vice-versa. Portanto, é um preditor chave no prognóstico de retenção e movimento da água e estoque de carbono no solo, além disso, indica a existência de compactação/adensamento. No entanto, a densidade do solo não é frequentemente determinada durante a execução de levantamentos de solos para todos os tipos de horizontes e solos.

A amostragem de campo e a obtenção de amostras para determinar a densidade do solo (Ds) são etapas laboriosas e fatigantes, principalmente abaixo de 30 cm de profundidade (Benites et al., 2007) e para determinar o estoque de carbono, geralmente são necessários dados de densidade até 1 m de profundidade. Araújo et al. (2012) usaram equações lineares múltiplas que relacionaram teor de argila, C e pH, para estimar Ds nas camadas de 20–40 cm com intuito de calcular o estoque de carbono em solos na região leste do Acre. Fidalgo et al.

(2007) estimaram o estoque de carbono no solo no Brasil considerando os diferentes tipos de solo sob diferentes sistemas de uso e distribuídos nos diversos biomas brasileiros. Para isso, a D_s foi estimada a partir de uma FPT utilizando valores de argila, carbono e soma de bases.

Bernoux et al. (1998) e Tomasella & Hodnett (1998) publicaram os primeiros trabalhos de predição de densidade do solo a partir das propriedades dos solos na Amazônia brasileira. O modelo proposto por Bernoux et al. (1998) incluiu os atributos argila, carbono orgânico total, pH e areia, já o modelo proposto por Tomasella & Hodnett (1998) incluiu carbono orgânico total, silte e argila. Esses autores utilizaram como fonte de dados, as informações contidas no projeto RadamBrasil. Barros & Fearnside (2015) geraram modelos para estimar a densidade do solo, no estado do Amazonas, a partir de parâmetros facilmente medidos em campo, presentes em muitos inventários pedológicos. O modelo que melhor estimou a densidade nas condições desse estudo foi o que incluiu o conteúdo de argila e o pH em H_2O como variáveis independentes. Os resultados mostraram que o uso da equação calibrada localmente evidenciou melhor desempenho para estimar a densidade do solo para as florestas de terra-firme da região de Manaus. Mais recentemente, Gomes et al. (2017) desenvolveram FPTs para estimar densidade dos solos a partir de propriedades facilmente disponíveis de solos da formação Iça na região Amazônica Brasileira. O modelo melhor ajustado proposto pelos autores incluiu carbono orgânico, pH_{H_2O} , soma de bases, alumínio e argila.

Quando se considera um modelo geral, para estimar densidade dos solos do Brasil, tem-se como referência o modelo proposto por Benites et al. (2007). Os autores desenvolveram um procedimento estatístico para prever densidade dos solos brasileiros presentes na maior parte dos biomas a partir de propriedades do solo de fácil obtenção e compararam o modelo gerado com três modelos existentes na literatura, dois para o Brasil (Bernoux et al., 1998; Tomasella & Hodnett, 1998) e um para os EUA (Manrique & Jones, 1991). O modelo incluiu teor de argila, carbono orgânico total e soma de bases e a fonte de dados utilizadas na geração do modelo foram as informações contidas em levantamentos de solos da Embrapa Solos. Os autores concluíram que comparado aos três modelos existentes, o proposto apresentou menor viés, maior precisão e maior exatidão, ou seja, foi o mais preciso.

3. CAPÍTULO I
SOIL CARBON STOCK AND PARTICLE SIZE FRACTIONS IN THE
CENTRAL AMAZON PREDICTED FROM REMOTELY SENSED
RELIEF, MULTISPECTRAL AND RADAR DATA

Document published in “Remote Sensing”.

Ceddia, M. B., Gomes, A. S., Vasques, G. M., Pinheiro, E. F. M., Soil carbon stock and particle size fractions in the central amazon predicted from remotely sensed relief, multispectral and radar data. Remote Sensing, 2017.

3.1 RESUMO

Os solos das áreas remotas da Floresta Amazônica no Brasil são mal mapeados devido à presença de floresta densa e falta de vias de acesso. O uso de covariáveis derivadas de sensores remotos multiespectrais e de radar permite mapear grandes áreas e tem o potencial de melhorar a precisão dos mapas de atributos do solo. Os objetivos deste estudo foram: (a) avaliar a adição de covariáveis de relevo, e de vegetação derivadas de imagens multiespectrais com resoluções espaciais e espectrais distintas (Landsat 8 e RapidEye) e radar de banda L (ALOS PALSAR) para a predição de estoque de carbono orgânico do solo (EC) e frações granulométricas; e (b) avaliar o desempenho de quatro métodos geoestatísticos para mapear essas propriedades do solo. Em geral, os resultados mostram que, mesmo sob cobertura vegetal densa, o Índice de Vegetação por Diferença Normalizada (NDVI) e o coeficiente de retroespalhamento ALOS PALSAR melhoraram a precisão das estimativas de EC e argila subsuperficial. O NDVI derivado do sensor RapidEye melhorou a predição de EC usando cokrigagem isotópica, enquanto o NDVI derivado do Landsat 8 e o coeficiente de retroespalhamento foram selecionados para predizer o conteúdo de argila subsuperficial usando a krigagem com regressão (KR). A melhora relativa da aplicação de Cokrigagem e KR sobre a krigagem ordinária foi inferior a 10%, indicando que análises adicionais são necessárias para conectar proxies de solo (tipos de vegetação e relevo) com os atributos do solo.

Palavras-chave: Mapeamento Digital de Solo. Geoestatística. Regressão. krigagem. Cokrigagem.

3.2 ABSTRACT

Soils from the remote areas of the Amazon Rainforest in Brazil are poorly mapped due to the presence of dense forest and lack of access routes. The use of covariates derived from multispectral and radar remote sensors allows mapping large areas and has the potential to improve the accuracy of soil attribute maps. The objectives of this study were to: (a) evaluate the addition of relief, and vegetation covariates derived from multispectral images with distinct spatial and spectral resolutions (Landsat 8 and RapidEye) and L-band radar (ALOS PALSAR) for the prediction of soil organic carbon stock (CS) and particle size fractions; and (b) evaluate the performance of four geostatistical methods to map these soil properties. Overall, the results show that, even under forest coverage, the Normalized Difference Vegetation Index (NDVI) and ALOS PALSAR backscattering coefficient improved the accuracy of CS and subsurface clay content predictions. The NDVI derived from RapidEye sensor improved the prediction of CS using isotopic cokriging, while the NDVI derived from Landsat 8 and backscattering coefficient were selected to predict clay content at the subsurface using regression kriging (RK). The relative improvement of applying cokriging and RK over ordinary kriging were lower than 10%, indicating that further analyses are necessary to connect soil proxies (vegetation and relief types) with soil attributes.

Key words: Digital soil mapping. Geostatistics. Regression kriging. Cokriging.

3.3 INTRODUCTION

The Brazilian Legal Amazon covers 5,217,423 km² (59% of the Brazilian territory) and encompasses a great diversity of vegetation and soils. However, despite its strategic importance, the knowledge about its soils is relatively poor and outdated (Brasil, 1973). The RADAMBRASIL project is the main source of data used to develop models and maps of soil attributes in the Brazilian Amazon. However, the soil maps were published at a coarse scale and the soil data available are very sparse across the Amazon, with ~0.7–3.5 soil profiles per 10,000 km² (Ceddia et al., 2015). The main reason for the low density of soil data is the presence of the dense Amazon Rainforest, which acts as a barrier for the pedologists and source of many natural threats (tropical diseases, predators), combined with limited access routes, with most of the territory only accessible by boats or airplane.

Digital soil mapping allows for the creation of maps of soil attributes by exploring the potential of ancillary environmental data such as passive optical and active microwave instruments, and digital elevation models. The covariates derived from digital elevation models have been exhaustively explored in traditional soil surveys and digital soil mapping projects to derive soil–landscape correlations and prediction models, since the topographic maps can often be used as proxies of the spatial distribution of hydrologic, geomorphologic and biologic factors controlling soil formation (Moore et al., 1991). Topography controls the flow of water, solutes and sediments, and these processes affect the catenary soil development, allowing to characterize soils and their properties in relation to their topographic spatial patterns (Sumfleth & Duttmann, 2008).

Multispectral and especially radar remote sensing data are relatively less explored in digital soil mapping, due to costs, the more complex processing required, and to their less direct connection with soil formation processes (Mulder et al., 2011). These data can be used in various ways, including to: (a) support the segmentation of the landscape into more homogeneous soil–landscape units; (b) predict soil properties using physically-based or empirical prediction methods; and (c) facilitate mapping in large inaccessible areas, where extensive field surveys are impractical. In sparsely vegetated areas, some soil properties (mineralogy, particle size distribution, iron, moisture, organic carbon, carbonate and salinity) have been successfully predicted using spaceborne, airborne and in situ measurements from multispectral and microwave sensors (Mulder et al., 2011; Shabou et al., 2015). However, in densely vegetated areas, the use of remote sensing images for mapping soils suffers from the fact that the vegetation covers much of the soil making it necessary to search for indirect evidences of soil variation inferred from the vegetation. In these cases, the efficiency of using remote sensing to map soil properties is even more dependent on indirect relations between soil attributes and soil formation proxies (vegetation, topography and drainage patterns, for example) (McBratney et al., 2003; Mulder et al., 2011).

While visible and infrared sensors only measure surface characteristics, radar can provide spectral information beyond the vegetative cover and the soil surface (L-band or lower frequencies). In the case of synthetic aperture radar (SAR), the backscattering signal is used to retrieve target properties, including the aboveground biomass and soil moisture (Dubois et al., 1995). The backscattering coefficient is calculated based on radar sensor configuration (observation frequency, polarization and incident angle) and surface characteristics (roughness and dielectric properties). As the soil water content is the main factor influencing its dielectric constant, the soil moisture content is the most frequently retrieved attribute (Dubois et al., 1995). Singh & Kathpalia (2007), developed a modeling approach based on a genetic algorithm to retrieve simultaneously soil moisture, roughness and

texture from the dielectric constant derived from ERS-2 SAR backscatter data. Despite the agreement of the predictions with the field observations, there were problems with the retrieval of the input variables of the model.

The relationship between soil and vegetation has the potential to improve the prediction of soil attributes. However, more detailed information on the vegetation cover may be needed, such as plant functional types and Ellenberg indicator values (Mulder et al., 2011). Such an approach could be useful for the Amazon basin, which is commonly densely vegetated. Ceddia et al. (2015) presented a qualitative model explaining the relationships between soil, relief and vegetation types in the same study area of the current study, in the Urucu River basin, Central Amazon, Brazil. They found that regions with higher slope and lower compound topographic index (CTI) presented higher clay contents and soil organic carbon stocks (CS). Under the conditions of the study, and considering that there is no water deficit during the year, the main vegetation type classified as Upland Dense Forest and the input of carbon from the vegetation, including that from deep roots, explained the higher values for both clay and CS. On the other hand, regions with higher CTI and lower slope (river plains and flatlands) presented both lower clay contents and CS. In these regions the soils are imperfectly drained and the vegetation types are classified as Flooded Lowland Open Tropical Rainforest, and Upland Open Tropical Rainforest, with species that are more adapted to soil aeration restrictions. The CS are lower because the inputs from leaves, trunks and roots are lower and concentrated at the surface layers. The authors showed that the use of CTI improved the prediction of CS using heterotopic cokriging.

Given the potential of remote sensing data to aid digital soil mapping in relatively remote and densely vegetated regions, we hypothesize that the predictions of CS and particle size fractions (PSFs) in the Central Amazon can be improved by adding relief and remotely sensed multispectral and radar covariates. Thus, the objectives were to: (a) evaluate the addition of relief and vegetation covariates derived from multispectral images with distinct spatial and spectral resolutions (Landsat 8 and RapidEye) and L-band radar (ALOS PALSAR) for the prediction of soil CS and PSFs; and (b) evaluate the performance of four geostatistical methods to map these soil properties, including ordinary kriging (OK), isotopic cokriging (ICOK), heterotopic cokriging (HCOK), and regression kriging (RK).

3.4 MATERIALS AND METHODS

This study extends the previous study reported by Ceddia et al. (2015). It uses the same study site and soil data used by the referred authors, but extends the analysis by including: the prediction of PSFs (clay, silt and sand contents), besides CS; other geostatistical methods; and covariates derived from multispectral, and radar remote sensing imagery to help predicting soil CS and PSFs.

3.4.1 Study Area

The study site is located in the central region of the Amazonas state, in the Urucu River basin, in the municipality of Coari, Brazil (Figure 1). It has an elongated shape of about 50 km by 8 m buffering an oil duct where access was possible, and covers about 7967 ha. The site is located about 640 km from Manaus (the state capital) and can be accessed only by boat or airplane. The climate is equatorial (Af, according to Köppen classification), with the temperature of the coldest month higher than 20 °C, mean annual precipitation of 2500 mm, and no pronounced dry period. According to Brasil (1978), the soils of the region were formed from Tertiary and Quaternary sediments of the Içá Formation, which covers an area of 563,264 km² (36% of the Amazonas state). They are composed of very fine sandstone, claystone and siltstone. The Holocene alluvium of the Quaternary deposits are related to the current Amazonian drainage networks. In the study area, the main soils include Argissolos (Ultisols) and Cambissolos (Inceptisols) (Embrapa, 2013; Soil Survey Staff, 2014), most of them with low sum of bases, high aluminum and medium to high sand contents. Some soils in the area present aquic characteristics, especially those near the Urucu River floodplain.

According to Brasil (1978), the main vegetation types along the study site are Upland Dense Tropical Rainforest, Flooded Lowland Open Tropical Rainforest and Upland Open Tropical Rainforest. The Upland Dense Tropical Rainforest is commonly found in areas with relief forms called terra firme, which are upland areas on flat-topped terrain located in river interflaves. The soils in these relief forms are well drained, with no water deficits for plants and have higher clay contents. The Flooded Lowland Open Tropical Rainforests are observed in relief forms called várzeas, which are the river floodplains, while the Upland Open Tropical Rainforests are commonly found in imperfectly drained flatlands. The soils of the study site were surveyed between 2008 and 2009 and the main soil classes and their respective numbers of soil profiles (94 in total) are shown in Table 1.

Table 1. Number of soil profiles (N) and frequency of soil suborders in the visited sites.

SiBCS	Soil Taxonomy	N	Frequency (%)
Argissolos Amarelos	Oxyaquic Hapludults	26	28
Argissolos Vermelhos	Typic Hapludults	2	2
Argissolos Vermelho-Amarelos	Typic Hapludults	21	22
Argissolos Acinzentados	Typic Endoaquults	3	3
Cambissolos Háplicos	Typic Dystrudepts	38	41
Espodossolos Humilúvicos	Humods	1	1
Neossolos Quartzarênicos	Quartzipsamments	2	2
Planossolos Háplicos	Aquults	1	1
Total	-	94	100

SiBCS, Brazilian Soil Classification System.

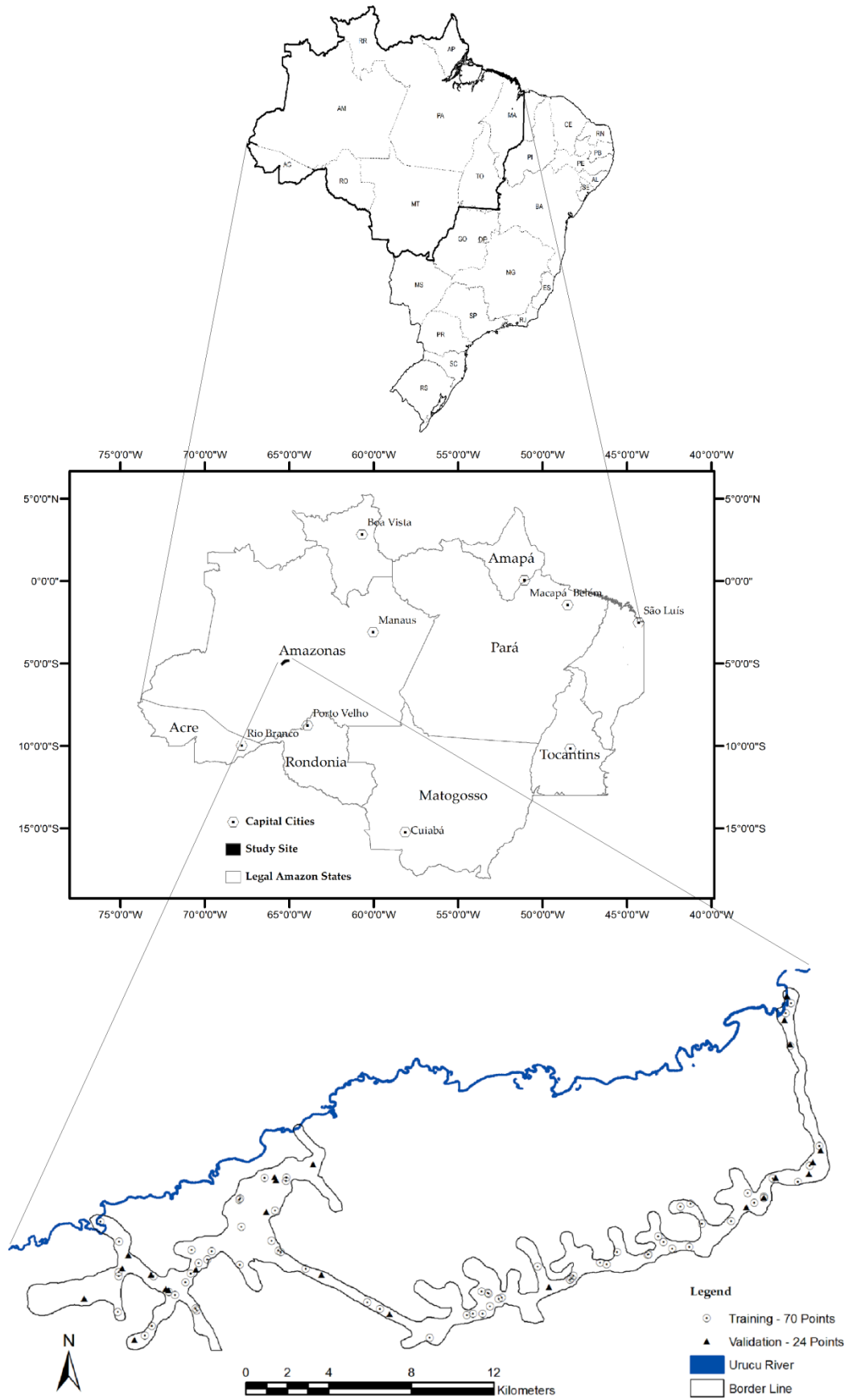


Figure 1. Location of the study area in the Central Amazon, Brazil (upper map), and the study site showing the location of the soil profiles (bottom map).

3.4.2 Soil Data

In each soil profile soil samples were collected following the soil horizon stratification, and analyzed for organic carbon (OC) by wet combustion (Walkley & Black, 1934), bulk density (BD) from Kopeck rings (4.2-cm high and 4.0-cm in diameter), and PSFs—sand, silt and clay contents—by the Pipette method (Embrapa, 1997). Soil CS were calculated for the 0-30 and 0-100-cm layers as the depth-weighted sum of the CS across all horizons with portions lying within the desired depth interval (Equation (1)). These two layers are the most used in soil OC studies (Batjes, 2000).

$$CS = \sum_{i=1}^n OC_i \times BD_i \times T_i \quad (1)$$

where CS is the OC stock ($\text{kg}\cdot\text{m}^{-2}$) at the desired layer (0–30 or 0–100 cm), OC_i is the OC content ($\text{g}\cdot\text{kg}^{-1}$) at horizon i , BD_i is the soil bulk density ($\text{Mg}\cdot\text{m}^{-3}$) at horizon i , T_i is the thickness (m) of the portion of horizon i that lies within the desired layer, and n is the number of horizons that have a portion within the desired layer.

Similarly, the sand, silt, and clay contents of the surface (Surf) and subsurface (Sub) layers were calculated as the depth-weighted average across the A and AB horizons for the Surf layer, and the BA and B horizons for the Sub layer (Equation (2)). The BC and C horizons were not included.

$$PSF_{Surf/Sub} = \sum_{i=1}^n PSF_i \times T_i / \sum_{i=1}^n T_i \quad (2)$$

where $PSF_{Surf/Sub}$ is the particle size fraction content ($\text{g}\cdot\text{kg}^{-1}$) at the desired layer (surface or subsurface), PSF_i is the PSF content ($\text{g}\cdot\text{kg}^{-1}$) at horizon i , T_i is the thickness (m) of the portion of horizon i that lies within the desired layer, and n is the number of horizons that have a portion within the desired layer.

3.4.3 Remote Sensing Covariate Data

Remotely sensed relief, multispectral, and radar imagery were assembled as independent variables, or covariates, for the prediction of soil CS at 0–30 and 0–100 cm, and sand, silt and clay contents at the surface and subsurface. Relief and multispectral data are commonly used in digital soil mapping studies, however, radar imagery is hardly ever included, although it has shown potential to predict soils and soil properties (McBratney et al., 2003; Metternicht & Zinck, 2003).

A topographic map with 2-m contour line intervals and a stream network map of the study area were derived at the 1:5000 scale from radar images (X-band, 0.5-m spatial resolution, HH polarization) using interferometry. A digital terrain model (DTM, 0.5-m resolution) with enforced drainage network and filled sinks was derived from the topographic map using the ANUDEM algorithm (Hutchinson, 1989), implemented in the Topo-to-Raster tool in ArcGIS version 9.3 (Esri, 2006). Four terrain attributes, namely slope gradient, slope aspect, curvatures, and CTI (Gessler, et al., 1995) were calculated from the DTM using the Spatial Analyst tools in ArcGIS.

Three orthorectified RapidEye (RE) satellite images from 16 June 2014, with 5-m resolution (resampled from the original 6.5-m resolution), were obtained from the Brazilian Ministry of Environment, mosaicked, and then used to derive vegetation indices from three bands, including band 3 at 0.630–0.685 μm (red), band 4 at 0.690–0.730 μm (red edge), and

band 5 at 0.760–0.850 μm (near infrared, NIR). In addition, a Landsat 8 Operational Land Imager (OLI) image from 31 October 2015 was obtained from the Brazilian National Institute for Space Research. Accordingly, two Landsat 8 OLI sensor bands were used, including band 4 at 0.636–0.673 μm (red), and band 5 at 0.851–0.879 μm (NIR), with 30-m spatial resolution. The RE images and the Landsat 8 OLI image were radiometrically and atmospherically corrected using the 6S model (Second Simulation of Satellite Signal in the Solar Spectrum), originally developed for the simulation of radiance at the satellite level by Vermote et al. (1997), and adapted for atmospheric correction by Antunes et al. (2012). As initial conditions, the tropical atmosphere, and the continental aerosol model were used.

Four vegetation indices were derived from the multispectral imagery, including the Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1973), the Enhance Vegetation Index (EVI) (Huete et al., 1994), the Soil-Adjusted Vegetation Index (SAVI) (Huete, 1988), and, for the RE images only, the Normalized Red Edge Vegetation Index (NDVI_Ed) (Equation (3)).

$$NDVI_{Ed} = (NIR - Red\ edge)/(NIR + Red\ edge) \quad (3)$$

where NIR is the near infrared band, and Red edge is the red edge band.

Along with multispectral images, three L-band microwave (23.6-cm wavelength) radar images from the Advanced Land Observing Satellite (ALOS) Phased Array type L-band Synthetic Aperture Radar (PALSAR) sensor from 6 July 2009 (two images), and 23 July 2009 (one image), with 12.5-m spatial resolution, were acquired and mosaicked. These images were downloaded from ASF DAAC (2016). The images were acquired in ascending orbit, with off-nadir angles of 38.8°, in Level 1.5 Fine Beam Dual (FBD) format, containing two bands with HH, and HV polarization, respectively. According to Shimada (2011), the ALOS PALSAR's long operating wavelength in L-band is suitable for tropical forest monitoring due to its high sensitivity to forest structure and moisture characteristics. The ALOS PALSAR images were used to derive the backscattering coefficients (σ°) of the two polarizations (HH and HV), which indicate the amount of microwave energy that is reflected by the target and returns to the sensor antenna per unit area, in decibels (dB). This was performed using MapReady version 3.2.1 (Alaska Satellite Facility, 2016).

All remotely sensed covariate rasters were assembled in a Geographic Information System and their values extracted to the field soil data in ArcGIS, thus deriving the database used to build the prediction models.

3.4.4 Prediction Methods

Four geostatistical methods were compared to predict soil CS and PSFs, including ordinary kriging (OK), isotopic cokriging (ICOK), heterotopic cokriging (HCOK), and regression kriging (RK). Ordinary kriging is a univariate method that uses the primary variable (CS or PSF) measured at sampled locations to predict the same primary variable at unsampled locations. In OK, the mean is taken as a constant but unknown value, and its stationarity is assumed only within a local neighborhood centered at the location being predicted. The OK predictor is a best linear unbiased predictor and is written as a linear combination, or weighted average, of the values at the sampled points, with weights adding up to 1 (Equation (4), (Goovaerts, 1997)). The unknown local mean $m(u)$ is filtered from the linear estimator by forcing the kriging weights to sum to 1.

$$Z_{OK}^*(u) = \sum_{\alpha=1}^{n_{\alpha}(u)} \lambda_{\alpha}^{OK}(u)[Z_1(u)], \text{ with } \sum_{\alpha=1}^{n_{\alpha}(u)} \lambda_{\alpha}^{OK} = 1 \quad (4)$$

Cokriging is a bivariate extension of kriging in which a secondary variable (that is, relief, multispectral or radar covariates— Z_v) is incorporated for the prediction of the target variable (CS or PSFs— Z_1) at unsampled locations by accounting for the spatial correlation between them (Equation (5), (Wackernagel, 2003)).

$$Z_{COK}^*(u) = \sum_{\alpha=1}^{n_z(u)} \lambda_{\alpha}^{COK}(u)[Z_1(u_{\alpha})] + \sum_{v=1}^{n_v(u)} \lambda_v^{COK}(u)[Z_v(u_v)], \text{ with } \sum_{\alpha=1}^{n_z(u)} \lambda_{\alpha}^{COK}(u) = 1 \text{ and } \sum_{v=1}^{n_v(u)} \lambda_v^{COK}(u) = 0 \quad (5)$$

Cokriging does not require that the secondary variable is available at all locations to make predictions. When the target and secondary variables have been measured at the same sampling locations, this is a case of isotopic cokriging. Otherwise, when they have been measured at different locations, heterotopic cokriging can be used (Wackernagel, 2003). Since the influence of the secondary variable on predicting the target one depends on: (i) the correlation between them; (ii) their spatial continuity; and (iii) their sampling density and spatial configuration (Simbahan et al., 2006), these two forms of cokriging (ICOK and HCOK) were compared.

Regression kriging combines regression (prediction) methods to model the global spatial trend, or external drift, with geostatistical methods to model (interpolate) the residuals from the regression (Grunwald, 2006). In RK, multiple linear regression (MLR) with stepwise variable selection ($p < 0.05$) models were first derived in R version 3.1.1 (R Development Core Team, 2014) for the target soil variables (CS and PSFs) as a function of relief covariates only, or relief, multispectral and radar covariates together (all variables). For each variable, the model with the highest adjusted R^2 was used to make predictions at all pixels in the study area. Then, the MLR model residuals, derived at the sampled locations, were interpolated across the study area using OK. The final prediction maps were derived as the sum of the MLR predictions and interpolated residuals across the study area. For all geostatistical methods, the auto- and cross-variograms were fitted manually using the spherical model.

In order to run and compare the accuracy of the prediction methods, the whole dataset containing 94 observations was randomly split into training (70 observations) and validation (24 observations) sets (Figure 1). The training observations were used for model development, while the validation observations were set aside for model validation and comparison using the mean error (ME) and root mean square error (RMSE) (Equations (6) and (7)). The ME is used to check the bias of the predictions, whereas the RMSE measures the accuracy of the predictions. The relative improvement (RI, in %; Equation (8)) in the RMSE was derived to compare the ICOK, HCOK and RK against OK.

$$ME = \frac{1}{n} \sum_{i=1}^n (O_i - P_i) \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2} \quad (7)$$

$$RI = \frac{RMSE_{OK} - RMSE_{COKorRK}}{RMSE_{OK}} \times 100 \quad (8)$$

where O_i and P_i are the n observed and predicted values, respectively, OK is ordinary kriging, COK is cokriging, and RK is regression kriging.

3.5 RESULTS

3.5.1 Descriptive Statistics and Linear Regression Models

The descriptive statistics of the whole data set (94 observations), as well as the training (70 observations) and validation (24 observations) data sets are presented in Table 2. The different data sets have similar descriptive statistics, with the mean values of all attributes statistically equal between the training and validation data sets, according to a Student's t-test at 0.05 significance. The main differences between data sets are observed in the minimum and maximum values of the attributes. Overall, the data sets are adequate for the development of predictive models and their validation, and represent the total set of data collected in the field.

Table 2. Descriptive statistics of target soil variables.

Variables	Set	N	Min	Max	Mean	Median	SD	Skewness	Kurtosis
Sand Surf (g·kg ⁻¹)	W	94	182	883	456	435	153	0.450	-0.324
Sand Surf (g·kg ⁻¹)	T	70	182	883	454 ^a	434	154	0.445	-0.077
Sand Surf (g·kg ⁻¹)	V	24	249	777	463 ^a	437	151	0.577	-0.508
Sand Sub (g·kg ⁻¹)	W	94	66	855	317	288	145	0.709	0.431
Sand Sub (g·kg ⁻¹)	T	70	66	855	317 ^a	290	147	0.805	1.060
Sand Sub (g·kg ⁻¹)	V	24	145	600	317 ^a	259	138	0.501	-1.279
Clay Surf (g·kg ⁻¹)	W	94	34	352	173	167	77	0.304	-0.538
Clay Surf (g·kg ⁻¹)	T	70	34	352	171 ^a	165	75	0.435	-0.191
Clay Surf (g·kg ⁻¹)	V	24	37	335	180 ^a	208	83	-0.035	-1.147
Clay Sub (g·kg ⁻¹)	W	94	13	531	327	334	104	-0.452	0.145
Clay Sub (g·kg ⁻¹)	T	70	13	531	321 ^a	332	102	-0.414	0.524
Clay Sub (g·kg ⁻¹)	V	24	121	478	346 ^a	388	106	-0.665	-0.472
Silt Surf (g·kg ⁻¹)	W	94	58	707	371	354	134	0.164	-0.376
Silt Surf (g·kg ⁻¹)	T	70	58	707	375 ^a	359	138	0.102	-0.217
Silt Surf (g·kg ⁻¹)	V	24	158	627	357 ^a	339	119	0.388	-0.473
Silt Sub (g·kg ⁻¹)	W	94	85	584	333	338	108	-0.087	-0.096
Silt Sub (g·kg ⁻¹)	T	70	85	584	336 ^a	342	112	-0.097	-0.122
Silt Sub (g·kg ⁻¹)	V	24	118	529	325 ^a	321	94	-0.158	-0.036
CS30 (kg·m ⁻²)	W	94	1.54	6.44	3.44	3.22	1.11	0.606	-0.071
CS30 (kg·m ⁻²)	T	70	1.54	6.44	3.47 ^a	3.21	1.18	0.582	-0.273
CS30 (kg·m ⁻²)	V	24	1.86	5.40	3.34 ^a	3.32	0.86	0.428	0.220
CS100 (kg·m ⁻²)	W	94	3.26	11.93	7.38	7.54	2.02	0.042	-0.408
CS100 (kg·m ⁻²)	T	70	3.64	11.93	7.46 ^a	7.68	2.09	0.059	-0.479
CS100 (kg·m ⁻²)	V	24	3.26	10.64	7.16 ^a	7.19	1.75	-0.213	-0.252

W, whole dataset; T, training dataset; V, validation dataset; Surf, surface layer; Sub, subsurface layer; CS30, carbon stock at 0–30 cm; CS100, carbon stock at 0–100 cm; N, number of observations; Min, minimum; Max, maximum; SD, standard deviation. ^a Equal means between the training and validation data sets for each soil attribute, respectively, according to a Student's t-test at 0.05 significance.

The soil CS at 0–100 cm (CS100) varied from 3.26 to 11.93 kg·m⁻², with an average value of 7.38 kg·m⁻² (Table 2). About 47% of this stock was in the 0–30-cm layer, with an average of 3.44 kg·m⁻². The CS obtained in this study, compared with those observed by different studies in the same region (Moraes et al., 1995; Batjes & Dijkshoorn, 1999; Cerri et al., 2000; Batjes, 2005), is 14% to 28% lower at 0–30 cm, and 5% to 17% lower at 0–100 cm. The differences occur because this study is constrained to a single geologic formation in the

Central Amazon, whereas the other studies are not. Moreover, the bulk density data used to calculate stocks also differed among the studies. The soil textural classification was loam at the surface, and clay loam at the subsurface, reflecting the constitution of the parent material of the study site, which is mainly composed of very fine sandstone and ferruginous claystone and siltstone. We call attention to the relatively high silt values of the region, both at the surface (mean: 371 g.kg⁻¹; maximum: 707 g.kg⁻¹) and subsurface (mean: 333 g.kg⁻¹; maximum: 584 g.kg⁻¹) layers, as these amounts of silt are not frequently found in Brazilian soils.

The regression models developed for soil CS and PSFs are presented in Table 3, except for sand and silt at the subsurface, which had no significant predictor selected among relief, multispectral and radar covariates. The regression models to predict silt at the subsurface, CS at 0–100 cm, and clay at the subsurface presented the highest adjusted R², though all regression models had low adjusted R², which means that the proportion of the variance in the soil variables that is predictable from the covariates is low. These poor correlations among variables also affected cokriging, as shown later. Nevertheless, the inclusion of covariates derived from Landsat 8 OLI and ALOS PALSAR increased the R², mainly for silt at the surface, clay at the subsurface, and CS at 0–30 cm (CS30).

Table 3. Linear regression models using relief only, and all remote sensing variables.

Target	Covariates	Regression Model	R ² _{adj}
Sand Surf	Relief	-	-
	All variables	924.43 + 5.83 × Slope – 604.31 × NDVI_L8	0.12
Sand Sub	Relief	-	-
	All variables	-	-
Silt Surf	Relief	445.91 – 9.40 × Slope	0.09
	All variables	130.01 + 490.94 × NDVI_L8 – 9.68 × Slope + 11.59 × BackHH – 18.58 × BackHV	0.22
Silt Sub	Relief	-	-
	All variables	-	-
Clay Surf	Relief	310.71 – 14.69 × CTI – 75.86 × CurvP	0.15
	All variables	157.64 + 173.17 × NDVI_L8 – 0.10 × Asp – 48.16 × CurvC – 18.56 × CTI – 61.89 × CurvP – 61.89 × BackHH	0.19
Clay Sub	Relief	431.57 + 0.17 × Asp – 15.75 × CTI	0.07
	All variables	50.85 + 337.20 × NDVI_L8 – 15.86 × CTI – 14.03 × BackHH	0.15
CS30	Relief	0.07 + 0.05 × Elev	0.09
	All variables	–0.68 + 0.05 × Elev – 0.12 × BackHH	0.11
CS100	Relief	4.14 – 0.32 × CTI + 0.10 × Elev	0.15
	All variables	3.02 – 0.38 × CTI + 0.09 × Elev – 0.25 × BackHH	0.19

Surf, surface layer; Sub, subsurface layer; CS30, carbon stock at 0–30 cm; CS100, carbon stock at 0–100 cm; NDVI_L8, Normalized Difference Vegetation Index derived from Landsat 8 Operational Land Imager; BackHH and BackHV, backscattering coefficients derived from ALOS PALSAR using HH and HV polarizations, respectively; CTI, compound topographic index; CurvP, profile curvature; Asp, aspect; CurvC, combined (standard) curvature; Elev, elevation.

The covariates derived from RapidEye, despite having a higher spatial resolution (5 m), were not selected in any model to predict CS or PSFs. This indicates that better resolution does not necessarily improve or grant significant soil–landscape correlations. The NDVI derived from Landsat 8 OLI (NDVI_L8) was selected in all significant PSF models, and the radar backscatter coefficient was selected in all significant models but for sand at the surface. The selection of relief covariates varied among target variables, with the slope selected to predict sand and silt at the surface, elevation selected only to predict CS, and the CTI selected to predict clay at both layers and CS100.

3.5.2 Geostatistical Analysis

The results of the spatial dependence analysis of the target soil variables, residues of the regression models, and remote sensing covariates individually, as well as of the spatial cross-dependence analysis between target soil and remote sensing variables are presented in Table 4 and Figures 2 and 3.

Table 4. Fitted auto- and cross-variogram parameters (spherical model).

Variables	C_0	C_1	Sill	Range (m)	$\frac{C_0}{C_0+C_1} \times 100$ (%)
Original variables					
Sand Surf	14.656	9314	23.970	2256	61.1
Silt Surf	5000	14.000	19.000	2000	26.3
Silt Sub	6971	6042	13.013	3256	53.6
Clay Sub	8200	3500	11.700	5500	70.1
CS100	3.70	1.30	5.0	8000	74.0
Regression residuals					
Sand Surf	15,000	6870	21,870	2000	68.6
Silt Surf	10,000	4100	14,100	3000	70.9
Clay Sub	7000	2000	9000	8000	77.8
CS100	3.20	2.70	5.90	8200	54.2
Silt Sub × secondary variables					
Elevation	15	35	50	3256	30.0
Slope	16	9	25	3256	64.0
CTI	2.5	0.8	3.0	3256	75.8
NDVI_RE	0.001	0.002	0.003	3256	20.0
NDVI_Ed	0.0012	0.0005	0.0017	3256	72.7
BackHH	2.5	2.2	4.7	3256	53.2
BackHV	3.7	0.9	4.6	3256	80.4
Silt Sub × Elevation	-50.0	-109.0	-159.0	3256	31.4
Silt Sub × Slope	0.0	-130.0	-130.0	3256	0.0
Silt Sub × CTI	0.0	41.0	41.0	3256	0.0
CS100 × secondary variables					
Elevation	20	35	55	8000	36.4
Slope	19	6	25	8000	76.0
CTI	2.5	0.8	3.3	8000	75.8
NDVI_RE	0.001	0.002	0.000	8000	33.3
NDVI_Ed	0.0012	0.0005	0.0000	8000	72.7
BackHH	3.0	1.8	4.8	8000	62.5
BackHV	3.8	0.9	4.7	8000	80.9
CS100 × Elevation	0.9	5.1	6.0	8000	15.0
CS100 × Slope	0.0	3.5	3.5	8000	0.0
CS100 × CTI	-0.80	-0.40	-1.20	8000	66.7
CS100 × NDVI_RE	0.0000	0.0098	0.0098	8000	0.0

C_0 , nugget variance; C_1 , sill variance; Surf, surface layer; Sub, subsurface layer; CS100, carbon stock at 0–100 cm; NDVI_RE and NDVI_Ed, Normalized Difference Vegetation Index and Normalized Red Edge Vegetation Index derived from RapidEye, respectively; BackHH and BackHV, backscattering coefficients derived from ALOS PALSAR using HH and HV polarizations, respectively; CTI, compound topographic index.

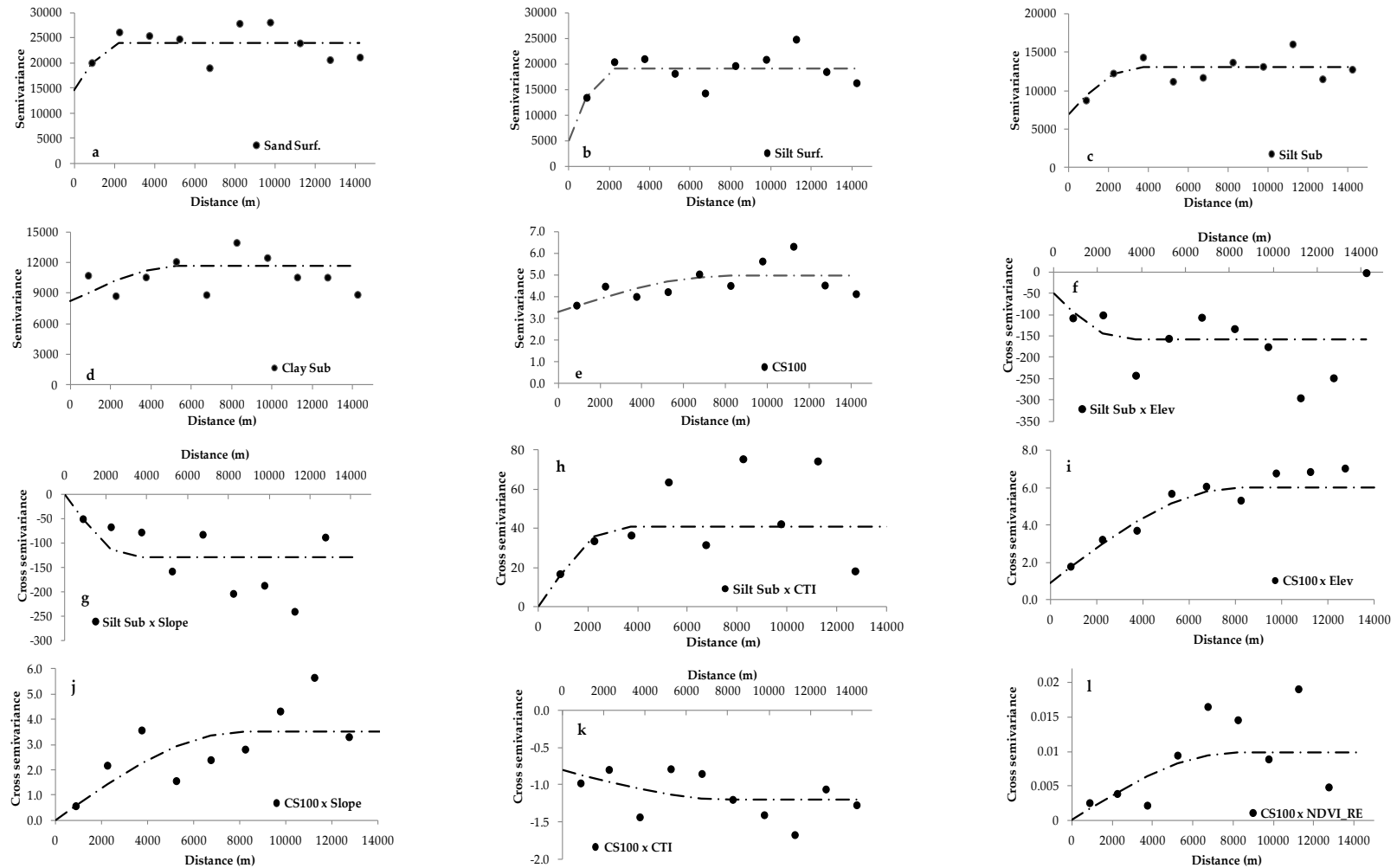


Figure 2. Experimental and fitted auto- and cross-variograms: (a) sand content at the surface; (b) silt content at the surface; (c) silt content at the subsurface; (d) clay content at the subsurface; (e) carbon stock at 0–100 cm (CS100); (f) silt content at the subsurface and elevation (Elev); (g) silt content at the subsurface and slope; (h) silt content at the subsurface and compound topographic index (CTI); (i) CS100 and elevation; (j) CS100 and slope; (k) CS100 and CTI; and (l) CS100 and Normalized Difference Vegetation Index derived from RapidEye (NDVI_RE).

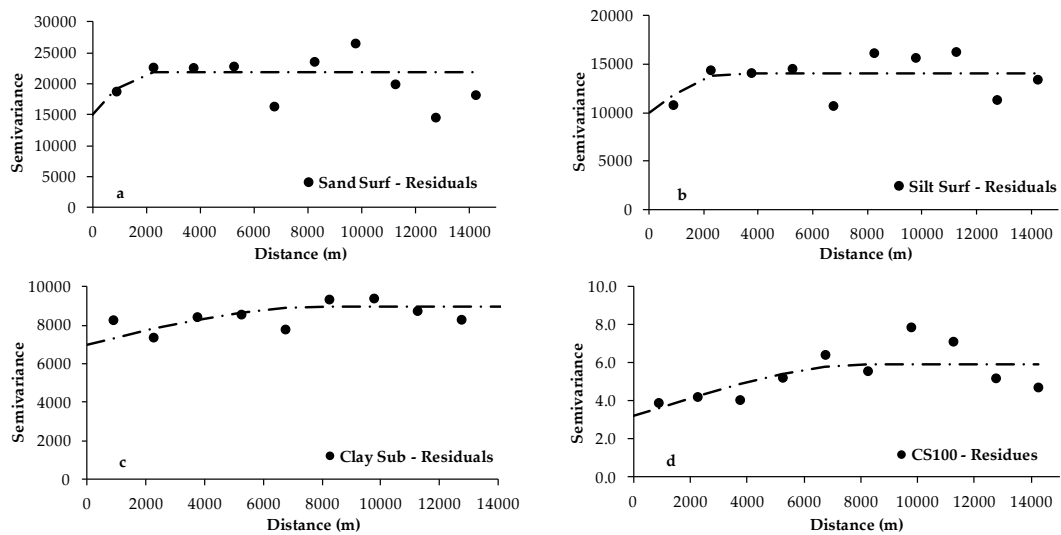


Figure 3. Experimental and fitted variograms of the regression residuals: (a) sand content at the surface; (b) silt content at the surface; (c) clay content at the subsurface; and (d) carbon stock at 0–100 cm.

The target soil variables that showed spatial autocorrelation were sand at the surface, silt at the surface and subsurface, clay at the subsurface, and CS100 (Table 4; Figure 2). The CS100 showed the highest range of spatial autocorrelation (8000 m), followed by the clay content at the subsurface (5500 m). Silt at the surface layer had the lowest range (2000 m) and the lowest proportion of the nugget variance over the total variance (that is, nugget-to-sill ratio) among the soil variables.

Ordinary cokriging (ICOK or HCOK) is more demanding than OK inasmuch as two auto-variograms (of the primary and secondary variable, respectively), as well as their cross-variogram, need to be known so that a model of co-regionalization can be defined. Considering these requisites, only two variables (silt at the subsurface and CS100) could be modeled by ICOK and HCOK. The candidate secondary variables that were spatially autocorrelated and could be used as covariates included elevation, slope, CTI, NDVI_RE, NDVI_Ed, BackHH and BackHV, and only a few of them were spatially correlated to silt or CS100 (Table 4; Figure 2). The cross-variograms show that silt at the subsurface is negatively correlated with elevation and slope, and positively correlated with CTI, whereas CS100 had the opposite behavior in relation to these covariates. Indeed, areas with higher elevation and slope presented higher CS and lower silt contents, whereas areas with higher CTI presented lower CS and higher silt contents in the study area. All cross-variograms presented lower nugget-to-sill ratio compared to the original soil variables, meaning that the spatial continuity is stronger when the secondary variables are considered for cokriging.

The NDVI_RE was the only multispectral index that could be used to map CS using cokriging. In this case, areas with higher NDVI_RE had higher CS. The lack of other multispectral indices that present cross spatial dependence with CS or PSFs was unexpected because the soil types are to some extent related to vegetation types in the study area (Ceddia et al., 2015). Moreover, the vegetation indices derived from RapidEye images did not cross-vary with the soil variables, which suggests that the higher spatial resolution of the indices did not help cokriging, except for CS100.

Regression kriging requires that the regression residuals are interpolated using OK. To do so, the regression residuals need to present spatial autocorrelation. From all soil attributes with fitted regression models, only sand and silt contents at the surface, clay content at the subsurface, and CS100 presented autocorrelated residuals (Table 4; Figure 3), which could be

interpolated and added back to the regression predictions. Clay at the surface and CS30 could not be predicted using RK.

In summary, considering the requirements of the kriging methods, most of the soil variables could be mapped by OK and RK, but only a couple could be cokriged. In OK, the variables not mapped included sand at the subsurface, clay at the surface and CS30, which did not show spatial autocorrelation. In RK, the variables not mapped included sand and silt at the subsurface, which lacked significant correlations with the covariates to derive a regression model, and clay at the surface and CS30, whose regression residuals were not spatially autocorrelated. Finally, only silt at the subsurface and CS100 could be mapped by ICOK and HCOK.

The final predicted maps of the soil attributes are shown in Figure 4. Among the prediction methods compared to map CS100, the best results (that is, with the lowest RMSE of external validation) were obtained by ICOK using slope as covariate, followed by ICOK using CTI, and then HCOK using NDVI from RapidEye (Table 5). For clay at the subsurface, RK outperformed OK. In this case, the regression model selected CTI, NDVI from Landsat 8 OLI, and the backscattering coefficient in HH polarization, contributing to improve both the ME and RMSE. These results show that cokriging (isotopic or heterotopic) has potential to improve the accuracy of soil attribute predictions using relief and multispectral remote sensing data. However, the gain in accuracy from applying these methods need to be evaluated from a cost-efficiency perspective, when compared to the simpler and cheaper OK method. Thus, the relative improvement (RI) of adding covariates to help in the predictions (as in ICOK, HCOK and RK) over OK was computed (Table 5). The RI shows that the RMSE of CS100 and clay at the subsurface were only modestly improved by up to 5% by ICOK, and 8% by RK. In comparison, OK derived the best predictions for sand at the surface, and silt at the surface and subsurface. Clay at the surface, sand at the subsurface, and CS30 were not mapped.

Table 5. External validation errors, with the best model indicated in bold for each target soil variable, respectively.

Variables	ME	RMSE	RI (%)
Ordinary kriging			
CS100 ($\text{kg}\cdot\text{m}^{-2}$)	0.08	1.79	-
Sand Surf ($\text{g}\cdot\text{kg}^{-1}$)	4.40	198.80	-
Silt Surf ($\text{g}\cdot\text{kg}^{-1}$)	6.26	139.61	-
Silt Sub ($\text{g}\cdot\text{kg}^{-1}$)	-27.38	80.02	-
Clay Sub ($\text{g}\cdot\text{kg}^{-1}$)	-33.43	127.96	-
Isotopic cokriging			
CS100/Slope ($\text{kg}\cdot\text{m}^{-2}$)	0.24	1.70	5.0
CS100/Elevation ($\text{kg}\cdot\text{m}^{-2}$)	0.24	1.75	2.2
CS100/NDVI_RE ($\text{kg}\cdot\text{m}^{-2}$)	0.20	1.78	0.5
CS100/CTI ($\text{kg}\cdot\text{m}^{-2}$)	0.26	1.73	3.4
Silt Sub/Slope ($\text{g}\cdot\text{kg}^{-1}$)	-17.22	83.40	-4.2
Silt Sub/Elevation ($\text{g}\cdot\text{kg}^{-1}$)	-16.62	82.15	-2.7
Silt Sub/NDVI_RE ($\text{g}\cdot\text{kg}^{-1}$)	-19.42	84.82	-6.0
Silt Sub/CTI ($\text{g}\cdot\text{kg}^{-1}$)	-19.16	81.39	-1.7
Heterotopic cokriging			
CS100/Slope ($\text{kg}\cdot\text{m}^{-2}$)	-0.03	2.11	-17.8
CS100/Elevation ($\text{kg}\cdot\text{m}^{-2}$)	0.27	1.91	-6.7
CS100/NDVI_RE ($\text{kg}\cdot\text{m}^{-2}$)	0.28	1.74	2.8
CS100/CTI ($\text{kg}\cdot\text{m}^{-2}$)	0.16	1.89	-5.6
Silt Sub/Slope ($\text{g}\cdot\text{kg}^{-1}$)	-7.40	98.78	-23.4
Silt Sub/Elevation ($\text{g}\cdot\text{kg}^{-1}$)	-19.32	81.59	-2.0
Silt Sub/NDVI_RE ($\text{g}\cdot\text{kg}^{-1}$)	-18.19	83.17	-3.9
Silt Sub/CTI ($\text{g}\cdot\text{kg}^{-1}$)	-29.34	83.72	-4.6
Regression kriging			
CS100 ($\text{kg}\cdot\text{m}^{-2}$)	0.51	2.52	-42
Sand Surf ($\text{g}\cdot\text{kg}^{-1}$)	-84.46	212.83	-7.1
Silt Surf ($\text{g}\cdot\text{kg}^{-1}$)	134.80	224.69	-61
Clay Sub ($\text{g}\cdot\text{kg}^{-1}$)	-32.32	117.54	8.2

ME, mean error; RMSE, root mean square error; RI, relative improvement.

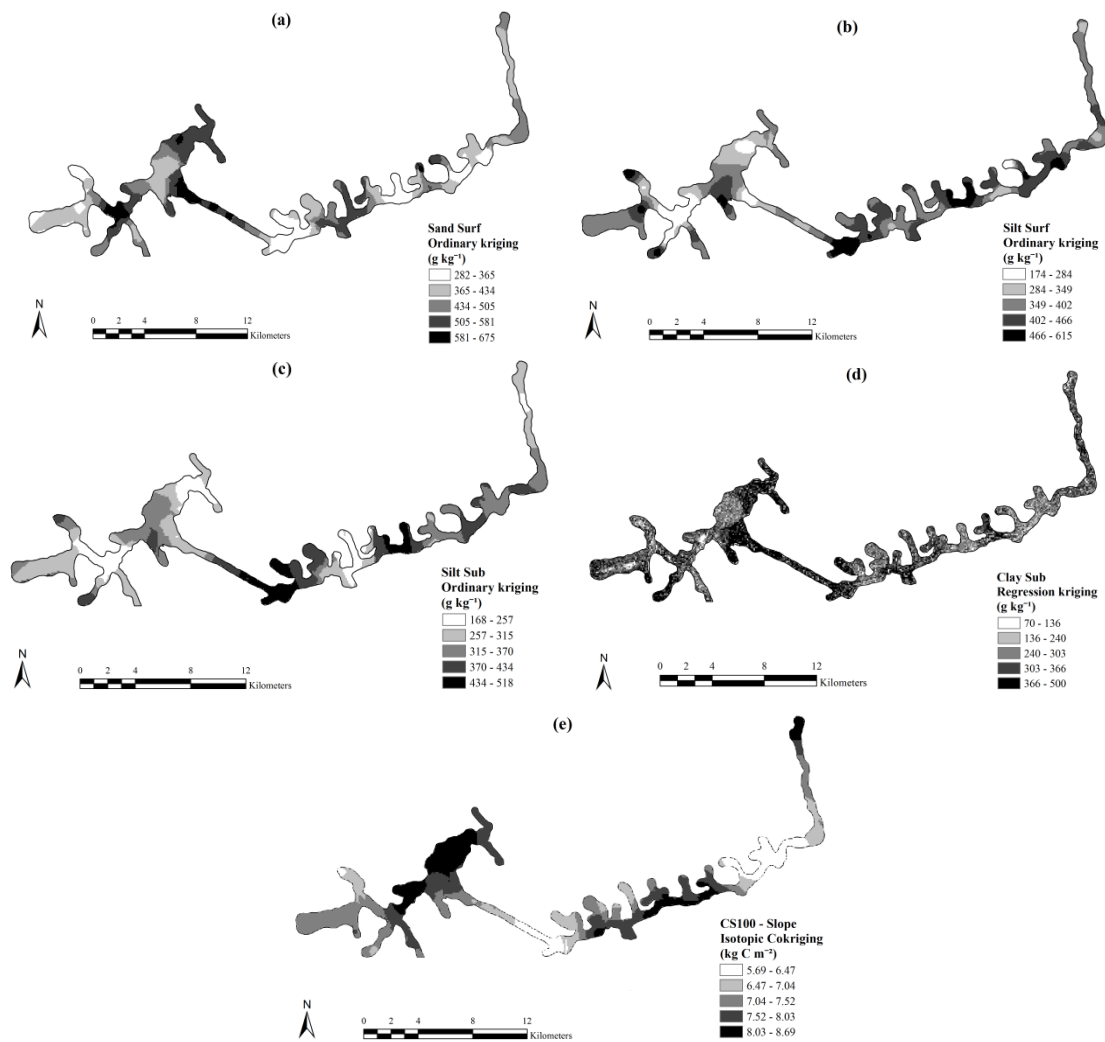


Figure 4. Final predicted maps of soil attributes derived from their best prediction methods: (a) sand content at the surface, (b) silt content at the surface, and (c) silt content at the subsurface predicted by ordinary kriging; (d) clay content at the subsurface predicted by regression kriging; and (e) carbon stock at 0–100 cm (CS100) predicted by isotopic cokriging using slope as covariate.

3.6 DISCUSSION

3.6.1 Potential of Using Multispectral and Radar Data as Covariates

This study offers new insights on the use of covariates derived from relief, multispectral and radar remote sensing data to improve the accuracy of soil attribute predictions from geostatistical methods under native Amazon forest. Considering the regression models selected to predict soil attributes, the covariates derived from relief, multispectral and radar data together explained only a low proportion (<25%) of the variance of the soil attributes, respectively. Nonetheless, the addition of NDVI (from Landsat 8 OLI) and backscattering coefficients (from ALOS PALSAR) increased substantially the prediction power of some models, compared to the others using only relief covariates (Table 3). This is the case of the models for silt at the surface (an increase from 9% to 22%) and clay at the subsurface (an increase from 7% to 15%). The clay content at the subsurface was better predicted by RK than OK, with the regression model selecting relief, multispectral, and radar variables. Although this was the only soil attribute better predicted by RK, the preference of OK over RK is not unusual, and has been reported elsewhere (Lozano-Garcia, 1991; Vasques et al., 2010; Zhu & Lin, 2010; Vasques et al., 2016). Thus, there is potential to improve soil attribute predictions by adding remote sensing covariates.

The clay content at the subsurface is higher when the value of NDVI is higher and both the backscattering coefficient and CTI are lower, respectively. The NDVI indicates the “greenness” of the land cover, or the vegetation health, and is related to the aboveground biomass. According to Ceddia et al. (2015), in the study site the densest vegetation type is Upland Dense Forest, whose NDVI are the highest in the study area. This vegetation type occurs more frequently in areas of low CTI on soils with relatively high clay contents. In the opposite direction, regions with the highest CTI and lowest clay contents occur closer to the river floodplain, where the plant species are more adapted to soil aeration restrictions. The vegetation types in these regions are classified as Flooded Lowland Open Tropical Rainforest and Upland Open Tropical Rainforest, which have lower density and smaller trunk diameters (thus, lower NDVI). Some authors have found similar results, showing that some soil attributes or types affect the NDVI. For example, Lozano-Garcia et al. (1991) derived NDVI values from the Advanced Very High Resolution Radiometer (AVHRR) sensor over 45 fields representing 8 soil associations in the state of Indiana (USA), finding higher NDVI values in forested areas with soils without restrictive layers, and higher clay content and water-holding capacity. Farrar et al. (1994) examined the extent to which differences in the rate of soil moisture generation, as a function of soil type or locality, accounts for the NDVI in semiarid Botswana. According to the authors, moisture availability is not the only soil characteristic determining the amount of vegetation growth. Instead, a group of soil attributes including moisture, porosity, nutrient availability, profile characteristics, and chemical attributes play a role on vegetation growth and rain-use efficiency. Although the latter example (Ferrar et al., 1994) refers to an environment completely different from the Amazon region, it emphasizes the complexity of factors directly or indirectly affecting the water dynamics in soils and the response of vegetation and NDVI.

The backscattering coefficient in HH polarization was selected in most regression models of PSFs and CS. The results found by Shimada (2011) can explain the predominance of the HH polarization in the regression models. The author developed a polarimetric-calibration method for ALOS PALSAR using time series imagery acquired over dense forest in the Brazilian Amazon (Rio Branco, Acre), and found that the HH signal of PALSAR

penetrates the forest canopy deeper and returns from the bottom of the forest stronger than the VV signal. The author argued that the total received power from the HH polarization could be decomposed into, or originated from, the double-bounce (or penetrating) signal, which accounts for about 20% of the signal, and the volume of the canopy, which accounts for the remaining 80%.

Although NDVI and radar backscattering coefficients showed higher potential to be used as covariates to improve the prediction of soil attributes, the prediction power of the models remained low. The causes may be related to the low sensitivity of these covariates to differentiate the soil patterns under native rainforest. In other words, the range of the NDVI values may not have been sufficient to discretize the different vegetation types and their associated soil attributes (CS and PSFs). Theoretical modeling results proved that large aboveground biomass in dense forests can decrease backscatter, due to the saturation of the free-space (or not attenuation) (Mermoz et al., 2015). This decrease in the backscatter can have strong implications for the use of L-band radar data for mapping, such as severe underprediction of soil water content and aboveground biomass (Dubois et al., 1995; Mermoz et al., 2015). This is especially important because the soil water content and dynamics play an important role to connect soil formation factors (vegetation and topography) with soil formation processes and soil attributes. However, according to Mermoz et al. (2015), the L-band does not entirely lose sensitivity at large biomass values, suggesting that much progress can be made by refining our understanding of radar backscattering behavior in the L-band.

The NDVI derived from RapidEye was selected as covariate to cokriging soil CS100 using HCOK (though the best CS100 predictor was ICOK using slope as covariate). Its higher spatial resolution (5 m) seems to be important to find the spatial cross-continuity between CS and NDVI. Rocchini (2007) points out that sensors with higher spatial resolution capture higher levels of detail, consequently better representing the spatial variation within a short distance. On the other hand, coarser resolution data (for example, Landsat 8 OLI with 30-m resolution) tend to have mixed-pixel problems and hence less sensitivity to spatial complexity at short distances. However, in the opposite direction, recent research has shown that higher spatial resolution in covariates does not necessarily improve prediction quality (Samuel-Rosa et al., 2015).

3.6.2 Performance of the Different Prediction Methods

The performance of the prediction methods applied in this study changed according to the soil attribute being mapped. Ordinary kriging presented the best performance to predict sand content at the surface and silt both at the surface and subsurface, whereas the clay content at the subsurface was better predicted using RK. The only soil variables that could be cokriged were CS100 and silt at the subsurface. This is unfortunate because CS100 was best predicted by ICOK and the cokriged silt maps were practically as accurate as the OK one. This suggests that the other soil variables could have been better predicted by cokriging but only a few covariates had the necessary spatial configurations to meet the assumptions of this method (for example, NDVI derived from RapidEye). These results are in accordance with other studies (Odeh et al., 1995; Hengl & Heuvelink, 2004) showing that there is no single best prediction method for all soil attributes. Cokriging and RK, despite having potential to improve the predictions, are more complex methods, and thus, the extra effort for their implementation must be taken into consideration. The RI (Table 5) gives an idea of how much improvement can be obtained from applying cokriging or RK, relative to OK, which takes the least effort. Kravchenko & Robertson (2007), mapping total soil carbon, evaluated the RI of applying RK compared to OK. The authors found that the combination of the primary variable having strong spatial continuity (nugget-to-sill ratio < 30%), and the primary and secondary

variables being relatively weakly correlated ($R^2 < 0.40$) appeared to be particularly unfavorable to obtain any improvement from RK. Besides, when the R^2 was smaller than 0.30, the maximum improvement from RK did not exceed 10%. The same reasoning from Kravchenko & Robertson (2007) can be considered to explain the poor performance of RK in this study ($RI < 10\%$), since the R^2 of all regression models were less than 0.30. The exception was the clay content at the subsurface, which had a R^2 of 0.15 and a nugget-to-sill ratio of 70.1%, but still was better predicted by RK than OK. Its high nugget-to-sill ratio can be caused by measurement errors and by the limited number and spatial configuration of the training observations, which might not have captured the structural variance at shorter distances. The weak correlation between the clay content at the subsurface and the covariates from the regression models can be a consequence of the reduction of the effect of topography and surface processes as the soil depth increases. For example, within a relatively short distance, some soils in the study site have textural gradients (intensive increase of clay content with depth), while others do not. The textural gradient depends on the translocation of clay, which, in turn, depends on both the water dynamics and the degree of clay flocculation or dispersion within the soil profile.

3.7 CONCLUSIONS

This work hypothesized that the use of relief, multispectral and radar remote sensing data can improve the accuracy of soil attribute predictions under Amazon native forest. Although the improvements were modest and limited to a couple of soil attributes, there is potential to use remotely sensed multispectral and radar data in the Amazon to improve the accuracy of soil attribute predictions. However, further research is necessary to better connect soil formation proxy variables derived from remote sensing to the spatial distribution of soil attributes in the Amazon, including the more commonly used vegetation indices and relief variables, but especially those still little explored like L-band radar data. The Amazon Rainforest offers harsh conditions for both soil data collection in the field (due to its remoteness, access limitations, risks, etc.), and data correlation between field soil and remote sensing data (due to the dense land cover that hinders direct soil observation from above, with the models relying on indirect correlations among remote sensing reflectance data, vegetation/land cover, and soil attributes).

The source and type (multispectral vs. radar) of the remote sensing data influenced which covariates were selected by the prediction models, with the NDVI derived from Landsat 8 OLI and the ALOS PALSAR backscattering coefficients most frequently selected. On the other hand, among multispectral and radar covariates, only RapidEye NDVI presented spatial cross-continuity to predict CS100 by cokriging. This may be related to the spatial resolution of the data, because only fine-resolution (multispectral and relief) covariates could be used for cokriging, contrary to the coarser-resolution radar (12.5 m) and Landsat 8 data (30 m). This is important because ICOK and HCOK presented lower RMSE to predict CS100, thus it may be worth testing other fine-resolution data in the search for a better improvement than those found in this study that reached a maximum of 8%.

3.8 REFERENCES

- ALASKA SATELLITE FACILITY. **MapReady 1.0**. Disponível em: <http://www.asf.alaska.edu/sardatacenter/softwaretools>. Acesso em: 15 mar 2016.
- ANTUNES, M. A. H.; DEBIASI, P.; COSTA, A. R.; GLERIANI, J. M. Correção atmosférica de imagens ALOS/AVNIR-2 utilizando o modelo 6S. **Revista Brasileira de Cartografia**, v. 64, p. 531-539, 2012.
- ASF DAAC. **Alaska Satellite Facility Distributed Active Archive Center**. Available online: <https://www.asf.alaska.edu/sar-data/palsar/download-data/> (accessed on 23 March 2016).
- BATJES, N. H.; DIJKSHOORN, J. A. Carbon and nitrogen stocks in the soils of the Amazon Region. **Geoderma**, v. 89, p. 273-286, 1999.
- BATJES, 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.
- BATJES, N. H. Organic carbon stocks in the soils of Brazil. **Soil Use and Management**, v. 21, p. 22-24, 2005.
- BRASIL. Departamento Nacional da Produção Mineral. **Projeto RADAMBRASIL**. Levantamento de Recursos Naturais (Geologia, Geomorfologia, Pedologia, Vegetação, Uso Potencial da Terra). Ministério das Minas e Energia (DNPM), Rio de Janeiro Brasil, 1973.
- BRASIL. Departamento Nacional da Produção Mineral. **Projeto RADAMBRASIL**. Folha SB. 20 Purus: geologia, geomorfologia, pedologia, vegetação e uso potencial da terra. Rio de Janeiro, 1978. 561p.,+ mapas. (Levantamento de Recursos Naturais, 17).
- CEDDIA, M. B.; VILLELA, A. L. O.; PINHEIRO E. F. M.; WENDROTH, O. Spatial variability of soil carbon stock in the Urucu river basin, Central Amazon-Brazil. **Science Total Environmental**, v. 526, p. 58–69, 2015.
- CERRI, C. C.; BERNOUX, M.; ARROUAYS, D.; FEIGL, B.; PICCOLO, M. C. **Carbon pools in soils of the Brazilian Amazon**. In Global Climate Change and Tropical Ecosystems. Lal, R., Ed.; CRC Press: Boca Raton, FL, USA, 2000; pp. 33–50.
- DUBOIS, P. C.; VAN ZYL, J.; ENGMAN, T. Measuring soil moisture with imaging radar. **IEEE Transactions Geoscience and Remote Sensing**, v. 33, p. 915–926, 1995.
- EMBRAPA. Centro Nacional de Pesquisa de Solos. **Sistema Brasileiro de Classificação de Solos**. 353p. 3ª edição. 2013.
- EMBRAPA. **Manual de métodos de análise de solo**. 2. ed. Rio de Janeiro, Centro Nacional de Pesquisa de Solos, 1997. 212p.
- ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE (ESRI). 2006. **ArcInfo 9.3**.
- FARRAR, J. F.; MINCHIN, P. E. H.; THORPE, M. R. Carbon import into barley roots: Stimulation by galactose. **Journal of Experimental Botany**, v. 45, p. 17–22, 1994.
- GESSLER, P. E.; MOORE, I. D.; MCKENZIE, N. J.; RYAN, P. J. Soil-landscape modeling and spatial prediction of soil attributes. **International Journal of Geographical Information Science**, v. 9, p. 421-432, 1995.
- GOOVAERTS, P. **Geostatistics for Natural Resource Evaluation**; Oxford University Press: New York, NY, USA, 1997; Volume 27, pp. 315–334.

- GRUNWALD, S. Environmental Soil-Landscape Modeling-Geographic Information Technologies and Pedometrics; CRC Press: New York, NY, USA, 2006; pp. 154–196.
- HENGL, T.; HEUVELINK, G. B. M.; STEIN, A. A generic framework for spatial prediction of soil variables based on regression-kriging. **Geoderma**, v. 120, p. 75-93, 2004.
- HUETE, A. A soil-adjusted vegetation index (SAVI). **Remote Sensing Environmental**, v. 25, p. 295–309, 1988.
- HUETE, A.; JUSTICE, C.; LIU, H. Development of vegetation and soil indices for MODIS-EOS. **Remote Sensing Environmental**, v. 29, p. 224-234, 1994.
- HUTCHINSON, M. F. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. **Journal of Hydrological**, v. 106, p. 211-232, 1989.
- LOZANO-GARCIA, D. F.; FERNANDEZ, R. N.; JOHANNSEN, C. J. Assessment of regional biomass-soil relationships using vegetation indexes. **IEEE Transactions Geoscience and Remote Sensing**, v. 29, p. 331–338, 1991.
- KRAVCHENKO, A. N.; ROBERTSON, G.P. Can topographical and yield data substantially improve total soil carbon mapping by regression kriging? **Agronomy Journal**, v. 99, p. 12–17, 2007.
- MCBRATNEY, A. B.; MENDONÇA SANTOS, M. L.; MINASNY, B. On digital soil mapping. **Geoderma**, v. 117, p. 3–52, 2003.
- MERMOZ, S.; REJOU-MECHAIN, M.; VILLARD, L.; TOAN, T. L.; ROSSI, V.; GOURLET-FLEURY, S. Decrease of L-band SAR backscatter with biomass of dense forests. **Remote Sensing Environmental**, v. 159, p. 307-317, 2015.
- METTERNICHT, G. I.; ZINCK, J. A. Remote sensing of soil salinity: Potentials and constraints. **Remote Sensing Environmental**, v. 85, p. 1–20, 2003.
- MOORE, I. D.; GRAYSON, B. R.; LADSON, A. R. Digital terrain modeling: A review of hydrological, geomorphological, and biological applications. **Hydrological Processes**, v. 5, p. 3–30, 1991.
- MORAES, J. L.; CERRI, C. C.; MELILLO, J. M.; KICKLIGHTER, D.; NEILL, C.; SKOLE, D. L.; STEUDLER, P. A. Soil carbon stocks of the Brazilian Amazon Basin. **Soil Science Society of America Journal**, Madison, v. 59, p. 244-247, 1995.
- MULDER, V. L.; BRUIN, S.; SCHAEPMAN, M. E.; MAYR, T. R. The use of remote sensing in soil and terrain mapping - A review. **Geoderma**, v. 162, p. 1–19, 2011.
- ODEH, I. O. A.; MCBRATNEY, A. B.; CHITTLEBOROUGH, D. J. Further results on prediction of soil properties from terrain attributes: heterotpic cokriging and regression-kriging. **Geoderma**, v. 67, p. 215-225, 1995.
- R Development Core Team. **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing, Vienna, Austria. Version 3.1.1. 2014. Available online: <http://www.R-project.org> (accessed on 24 January 2017).
- ROCCHINI, D. Effects of spatial and spectral resolution in estimating ecosystem alpha-diversity by satellite imagery. **Remote Sensing Environmental**, v. 111, p. 423-434, 2007.
- ROUSE, J. W.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W. **Monitoring vegetation systems in the great plains with ERTS**. In Earth Resources Technology Satellite-1 Symposium; NASA: Washington, DC, USA, 1973; Volume 1, pp. 309–317.

- SAMUEL-ROSA, A.; HEUVELINK, G. B. M.; VASQUES, G. M.; ANJOS, L. H. C. Do more detailed environmental covariates deliver more accurate soil maps? **Geoderma**, v. 243, p. 214-227, 2015.
- SHABOU, M.; MOUGENOT, B.; CHABAANE, Z. L.; WALTER, C.; BOULET, G.; AISSA, N. B.; ZRIBI, M. Soil clay content mapping using a time series of Landsat TM data in semi-arid lands. **Remote Sensing**, v. 7, p. 6059–6078, 2015.
- SHIMADA, M. Model-Based polarimetric SAR calibration method using forest and surface-scattering targets. **IEEE Transactions Geoscience and Remote Sensing**, v. 49, p. 1712-1733, 2011.
- SIMBAHAN, G. C.; DOBERMANN, A.; GOOVAERTS, P.; PING, J.; HADDIX, M. L. Fine-resolution mapping of soil organic carbon based on multivariate secondary data. **Geoderma**, v. 132, p. 471-489, 2006.
- SINGH, D.; KATHPALIA, A. An efficient modeling with GA approach to retrieve soil texture, moisture and roughness from ERS-2 SAR data. **Prog. Electromagn. Res.** v. 77, p. 121–136. 2007.
- SOIL SURVEY STAFF. **Keys to Soil Taxonomy**, 12th ed.; Natural Resources Conservation Service, United States Department of Agriculture: Washington, DC, USA, 2014.
- SUMFLETH, K.; DUTTMANN, R. Prediction of soil property distribution in paddy soil landscapes using terrain data and satellite information as indicators. **Ecological Indicators**, v. 8, p. 485–501, 2008.
- VASQUES, G. M.; COELHO, M. R.; DART, R. O.; OLIVEIRA, R. P.; TEIXEIRA, W. G. Mapping soil carbon, particle-size fractions, and water retention in tropical dry forest in Brazil. **Pesquisa Agropecuária Brasileira**, v. 51, p. 1371-1385, 2016.
- VASQUES, G. M.; GRUNWALD, S.; COMERFORD, N. B.; SICKMAN, J. O. Regional modelling of soil carbon at multiple depths within a subtropical watershed. **Geoderma**, v. 156, p. 326-336, 2010.
- VERMOTE, E. F. D.; TANRÉ, D.; DEUZÉ, J. L.; HERMAN, M.; MORCRETE, J.J. Second simulation of the satellite signal in the solar spectrum, 6S: An overview. **IEEE Transactions Geoscience and Remote Sensing**, v. 35, p. 675-686, 1997.
- WACKERNAGEL, H. **Multivariate Geostatistics: An Introduction with Applications**, 3rd ed.; Springer: New York, NY, USA, 2003; p. 387.
- WALKLEY, A.; BLACK, I. A. An examination of the Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. **Soil Science**, v. 37, p. 29–38, 1934.
- ZHU, Q.; LIN, H. S. Comparing ordinary kriging and regression kriging for soil properties in contrasting landscapes. **Pedosphere**, v. 20, p. 594-606, 2010.

CAPÍTULO II
FUNÇÃO DE PEDOTRANSFERÊNCIA PARA ESTIMAR DENSIDADE
DE SOLOS DO BRASIL

4.1 RESUMO

Funções de Pedotransferência (FPTs) são modelos de predição de atributos de difícil mensuração, utilizando como variáveis preditoras atributos facilmente disponíveis em levantamentos de solos. Essas funções são alternativas bastante difundidas e utilizadas na literatura, podendo minimizar o problema de escassez de valores de densidade do solo. Esse estudo teve por objetivos: i) desenvolver FPTs para estimar a densidade de solos do Brasil empregando atributos físicos e químicos disponíveis em levantamentos de solos; ii) analisar se a subdivisão do banco de dados de acordo com as cinco regiões (Norte, Nordeste, Centro-Oeste, Sudeste e Sul) gera modelos mais acurados; iii) Comparar a performance dos modelos gerados com o proposto por Benites et al. (2007) que é referência na predição de Ds para o Brasil. O banco de dados foi organizado a partir de dados de levantamentos de solos, relatórios, reuniões, teses, dissertações e artigos, disponíveis em bibliotecas e por meio virtual, realizados pelo território brasileiro. O banco de dados foi subdividido por regiões (Norte, Nordeste, Centro-Oeste, Sul e Sudeste). Foram desenvolvidos modelos de regressão linear múltipla para o Brasil e para as cinco regiões e comparado sua performance com o modelo referência publicado na literatura proposto por Benites et al. (2007), sendo aplicados os seguintes indicadores estatísticos: coeficiente de determinação (R^2), o erro médio de predição (MPE), a Raiz do Quadrado médio do erro de predição (RMSPE), o erro médio absoluto (MAE) e melhoria relativa (MR). As FPTs geradas tiveram melhor performance comparada ao modelo proposto por Benites et al. (2007), exceto para a região Sul. Todos os modelos gerados, exceto para a região Sul, apresentaram valores de MPE, MAE, RMSPE mais baixos e R^2 mais elevados quando comparados ao modelo proposto por Benites et al. (2007) aplicado aos dados de validação de cada sub-banco. Comparando-se a performance do modelo geral (Brasil) em relação aos modelos gerados por regiões, observou-se menores valores de MPE, MAE e RMSPE, para os modelos gerados por regiões, exceto para a região Sul. Quando se considera um modelo geral para predizer a Ds, o modelo gerado para o Brasil, obteve melhor performance, aproximadamente 7% a mais de acurácia comparado ao modelo referência proposto por Benites et al. (2007). Para a região Sul, a FPT gerada para o Brasil ou a proposta por Benites et al. (2007), apresentaram melhor desempenho, podendo ser utilizadas para predição de Ds nessa região.

Palavras-chave: Regressão linear múltipla. Solos minerais. Textura do solo.

4.2 ABSTRACT

Pedotransfer Functions (PTFs) are predictive models of attributes that are difficult to measure, using as attributes predictors attributes easily available in soil surveys. These functions are quite widespread alternatives and used in the literature, being able to minimize the problem of scarcity of values of bulk density. The objective of this study was to: i) develop PTFs to estimate the soil density of Brazil using physical and chemical attributes available in soil surveys; ii) analyze whether the subdivision of the database according to the five regions (North, Northeast, Central West, Southeast and South) generate more accurate models; iii) Compare the performance of the models generated with that proposed by Benites et al. (2007) which is a reference in the prediction of Ds for Brazil. The database was organized based on soil survey data, reports, meetings, theses, dissertations and articles, available in libraries and by virtual means, carried out by the Brazilian territory. The database was subdivided by regions (North, Northeast, Midwest, South and Southeast). Multiple linear regression models were developed for Brazil and for the five regions and compared their performance with the reference model published in the literature proposed by Benites et al. (2007). The following statistical indicators were applied: coefficient of determination (R^2), mean prediction error (MPE), mean square root of prediction error (RMSPE), mean absolute error (MAE) and relative improvement (MR). The generated PTFs performed better than the model proposed by Benites et al. (2007), except for the South region. All models generated, except for the South region, presented higher values of MPE, MAE, RMSPE and R^2 higher when compared to the model proposed by Benites et al. (2007) applied to the validation data of each sub-dataset. Comparing the performance of the general model (Brazil) in relation to the models generated by regions, smaller values of MPE, MAE and RMSPE were observed for the models generated by regions, except for the South region. When considering a general model to predict Ds, the model generated for Brazil, obtained better performance, approximately 7% more accuracy compared to the reference model proposed by Benites et al. (2007). For the South region, the PTF generated for Brazil or the proposal by Benites et al. (2007), obtained better performance and can be used to predict Ds in this region.

Key words: Mineral soils. Multiple Linear Regression. Soil texture.

4.3 INTRODUÇÃO

A densidade do solo (Ds) tem relevante aplicação para o cálculo de outros atributos do solo como estoque de carbono, porosidade total, conversão de massa para volume entre outros, além disso, é indicador de compactação/adensamento do solo (Tamminen & Starr, 1994; De Vos et al., 2005; Han et al., 2012). No entanto, sua amostragem em campo é considerada laboriosa e demorada, especialmente abaixo de 30 cm de profundidade (Kaur et al., 2002; Heuschek et al., 2005; Benites et al., 2007).

Funções de Pedotransferência (FPTs) são modelos de predição de atributos de difícil mensuração, utilizando como variáveis preditoras atributos facilmente disponíveis em levantamentos de solos (Bouma, 1989). Essas funções são alternativas bastante difundidas e utilizadas na literatura, podendo minimizar o problema de escassez de valores de densidade do solo, possibilitando o cálculo de estoque de carbono e estudo de sua dinâmica.

Bernoux et al. (1998) e Tomasella & Hodnett (1998) publicaram os primeiros trabalhos de predição de densidade do solo a partir das propriedades dos solos mais facilmente disponíveis na Amazônia brasileira. O modelo proposto por Bernoux et al. (1998) incluiu os atributos argila, areia, carbono orgânico total e pH_{H_2O} , já o modelo proposto por Tomasella & Hodnett (1998) incluiu carbono orgânico total, silte e argila. Barros & Fearnside (2015) geraram modelos para estimar a densidade do solo, na Amazônia Central, sendo que o modelo que melhor estimou a densidade nas condições desse estudo foi o que incluiu o conteúdo de argila e o pH_{H_2O} como variáveis independentes. Mais recentemente, Gomes et al. (2017) desenvolveram FPTs para estimar densidade dos solos a partir de propriedades facilmente disponíveis de solos da formação Iça na região Amazônica Brasileira. O modelo melhor ajustado proposto pelos autores incluiu carbono orgânico, pH_{H_2O} , soma de bases, alumínio e argila.

Quando se considera um modelo geral, para estimar densidade dos solos do Brasil, tem-se como pioneiro o modelo proposto por Benites et al. (2007). Esses autores desenvolveram FPTs para prever densidade dos solos brasileiros presentes na maior parte dos biomas a partir de propriedades do solo de fácil obtenção. O modelo incluiu teor de argila, carbono orgânico total e soma de bases. Essas variáveis independentes conseguiram explicar 66% da variação da densidade do solo, mesmo sendo este um modelo robusto que engloba a grande variabilidade dos solos distribuídos pelos biomas brasileiros.

Esse estudo teve por objetivos: i) desenvolver FPTs para estimar a densidade de solos do Brasil empregando atributos físicos e químicos disponíveis em levantamentos de solos; ii) analisar se a subdivisão do banco de dados de acordo com as cinco regiões (Norte, Nordeste, Centro-Oeste, Sudeste e Sul) gera modelos mais acurados; iii) Comparar o desempenho dos modelos gerados com o proposto por Benites et al. (2007) que é referência na predição de Ds para o Brasil.

4.4 MATERIAL E MÉTODOS

4.4.1 Banco de Dados

O banco de dados foi organizado a partir de dados de levantamentos de solos, relatórios, reuniões, teses, dissertações e artigos, disponíveis em bibliotecas e por meio virtual, realizados pelo território brasileiro. Dispunha de 10.181 perfis de solo (39.036 horizontes/camadas), com informações de profundidade superior e inferior dos horizontes/camadas, fração grosseira, areia, silte, argila, argila dispersa, grau de floculação, densidade do solo e da partícula, pH em água e em KCl, bases trocáveis (Ca^{2+} , Mg^{2+} , K^+ , Na^+), soma de bases, Al^{3+} , H^+ , CTC, saturação por bases, saturação por alumínio, carbono orgânico total (COT), nitrogênio, óxidos (SiO_2 , Al_2O_3 , Fe_2O_3), condutividade elétrica, água disponível e equivalente de umidade. No entanto, essas informações não estavam disponíveis para todos os perfis ou horizontes/camadas.

Foram selecionados os perfis que possuíam valores de densidade do solo medidos até $1,90 \text{ Mg.m}^{-3}$, carbono orgânico menor que 80 g.kg^{-1} e todos os valores das propriedades que iriam ser testadas como preditoras de Ds (Areia, silte, argila, soma de bases, $\text{pH}_{\text{H}_2\text{O}}$, Al^{3+} , H^+ e COT). Ao final, foram utilizados 769 perfis de solo para gerar as FPTs, cerca de 7,5% do banco original. O exposto reflete a escassez de informações de Ds medidas em levantamentos de solos convencionais e reforça a necessidade de desenvolvimento de técnicas de predição desse atributo.

4.4.2 Desenvolvimento das FPTs

Foram selecionados 769 perfis de solo que dispunham de valores medidos de densidade do solo e valores de carbono abaixo de 80%, ou seja, os solos orgânicos foram retirados do banco. Foi sorteado, de forma aleatória, 20% dos perfis para validação dos modelos gerados e 80% para treinamento do modelo. Dos 769 perfis (3255 horizontes/camadas), 615 (2591 horizontes/camadas) foram utilizados para treinamento e 154 (664 horizontes/camadas) para validação.

Posteriormente, o banco de dados geral (Brasil) foi dividido em sub-bancos por regiões – Norte, Nordeste, Centro-Oeste, Sudeste e Sul. A subdivisão por regiões seguiu a mesma proporção de separação de dados para treinamento e validação (80/20) que foi aplicado ao banco de dados geral (Tabela 1). A distribuição dos perfis por região está apresentada na Figura 1. Devido a alguns levantamentos serem muito antigos, alguns perfis não apresentavam coordenadas geográficas disponíveis, portanto, não foram representados na figura.

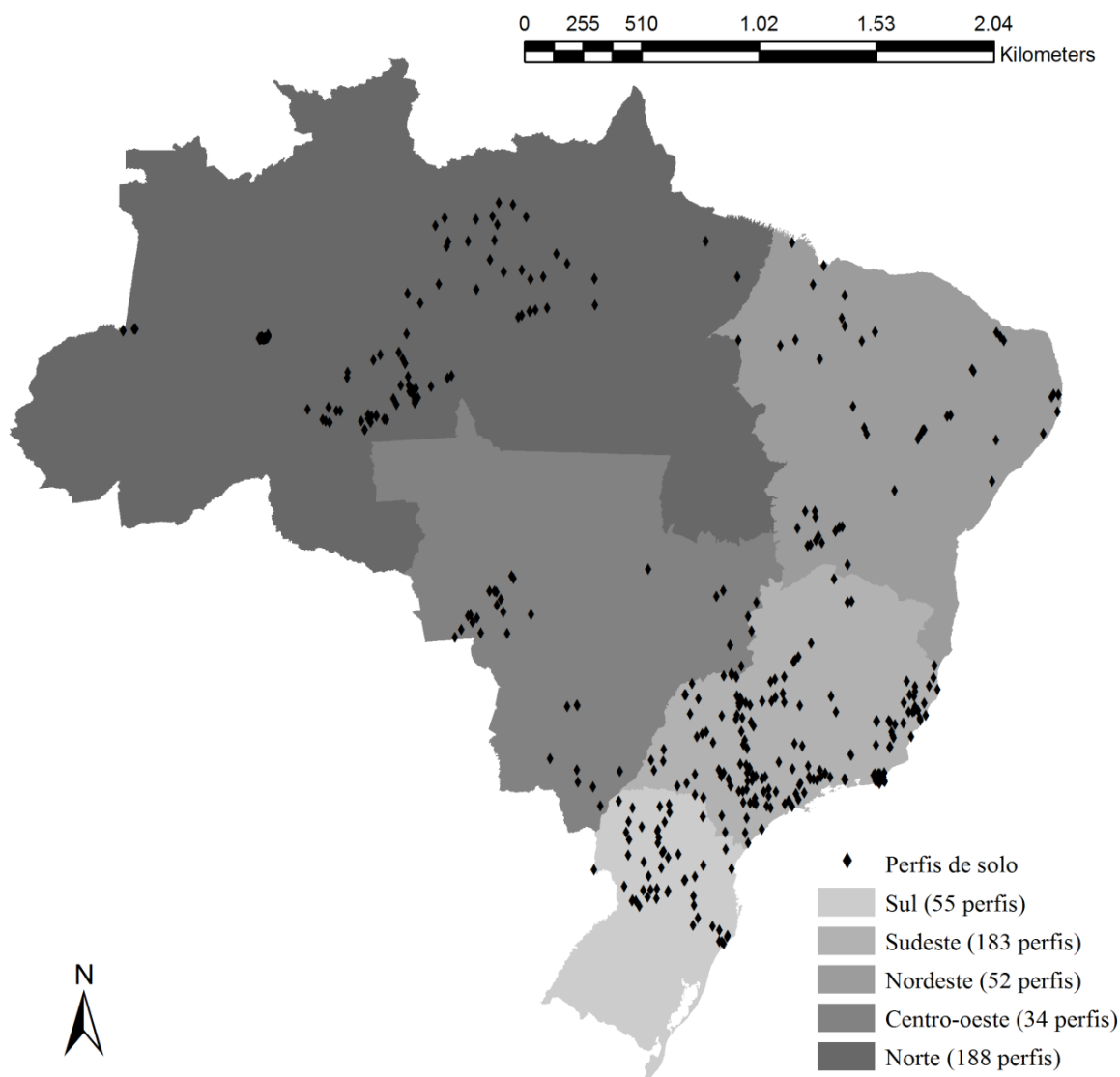


Figura 1. Distribuição dos perfis por regiões do Brasil.

Tabela 1. Proporção de dados de treinamento e validação para os sub-bancos.

Sub-bancos	Treinamento	Validação
Norte	162 perfis (740 horizontes/camadas)	41 perfis (192 horizontes/camadas)
Nordeste	84 perfis (339 horizontes/camadas)	21 perfis (90 horizontes/camadas)
Centro-Oeste	50 perfis (203 horizontes/camadas)	12 perfis (55 horizontes/camadas)
Sudeste	258 perfis (1123 horizontes/camadas)	64 perfis (272 horizontes/camadas)
Sul	62 perfis (192 horizontes/camadas)	15 perfis (49 horizontes/camadas)

Estatísticas descritivas do conjunto de dados foram realizadas para avaliar a existência de dados anômalos (*outliers*), normalidade, bem como, a possível necessidade de divisão do conjunto de dados para aumentar a acurácia da função.

O procedimento estatístico *stepwise* foi utilizado para seleção de variáveis preditoras e as FPTs foram geradas a partir de modelos de regressão linear múltipla. A Multicolinearidade foi minimizada pela remoção de variáveis com fator de inflação de variância > 4 . A normalidade dos resíduos de regressão foi analisada através do teste de Kolmogorov-Smirnov

(KS). O procedimento de *stepwise* foi realizado no programa R (R Development Core Team, versão 3.1.1).

4.4.3 Avaliação da Performance das FPTs

Na literatura, encontra-se publicado um modelo para estimar densidade dos solos do Brasil proposto por Benites et al. (2007):

$$D_s = 1,560 - 0,0005 * (\text{Argila}/\text{g kg}^{-1}) - 0,010 * (\text{COT}/\text{g kg}^{-1}) + 0,0075 * (\text{SB})$$

Portanto, esse modelo é o pioneiro para a predição de D_s no Brasil, logo, o desempenho dos modelos gerados será comparada a performance do modelo proposto por Benites et al. (2007) que será aplicado a todos os conjuntos de dados de validação analisados.

Os valores preditos de D_s (y_i) usando as FPTs geradas e a proposta por Benites et al., (2007) foram comparados aos valores observados dos bancos de validação. A diferença entre os valores preditos e os observados ($\hat{y}_i - y_i$) foi usada para calcular as seguintes medidas de erro: o erro médio de predição (MPE); a raiz do quadrado médio do erro de predição (RMSPE); o erro médio absoluto (MAE); coeficiente de determinação (R^2) e melhoria relativa (MR).

$$\text{MPE} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i) \quad (\text{Eq. 1})$$

$$\text{RMSPE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (\text{Eq. 2})$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (\text{Eq. 3})$$

$$R^2 = \frac{[\text{COV}(y_i, \hat{y}_i)]}{\text{Var}(y_i) \cdot \text{Var}(\hat{y}_i)} \quad (\text{Eq. 4})$$

$$\text{MR} = \frac{\text{RMSPE}_{\text{Benites et al. (2007)}} - \text{RMSPE}_{\text{modelos gerados}}}{\text{RMSPE}_{\text{Benites et al. (2007)}}} \times 100 \quad (\text{Eq. 5})$$

Em que:

y_i = densidade do solo observada da i -ésima amostra de solo;
 \hat{y}_i = densidade do solo predita para a i -ésima amostra de solo;
 n = número total de observações.

O MPE é usado para determinar o grau de enviesamento das estimativas, frequentemente referido como viés (Isaaks & Srivastava, 1989), revelando a tendência da FPT de superestimar os valores (quando positivo) ou de subestimar (quando negativo). De acordo com Willmott (1982), MAE e RMSPE estão entre as melhores medidas globais de desempenho de modelos, pois eles resumem a diferença média em unidades de valores observados e preditos. Em comparação com MAE, o RMSPE é mais sensível para valores extremos. A Melhora Relativa (MR, em%) no RMSPE foi derivada para comparar os modelos de predição.

4.5 RESULTADOS E DISCUSSÃO

4.5.1 Estatística Descritiva do Banco de Dados

A estatística descritiva das variáveis dos conjuntos de dados utilizadas na geração das FPTs está apresentada na Tabela 2. São apresentadas as estatísticas: média, valor mínimo, valor máximo, desvio padrão, coeficiente de variação e coeficiente de assimetria e curtose.

Considerando a estatística descritiva dos conjuntos de dados, pode-se observar um panorama geral dos atributos dos solos do Brasil e de suas regiões. Em relação às frações granulométricas do solo, há um predomínio das frações areia e argila em relação à silte. Quanto à composição química, observa-se que as médias dos valores de $\text{pH}_{\text{H}_2\text{O}}$ foram próximas, no entanto há uma grande variação nos atributos SB, Al^{3+} , H^+ e COT. A região Nordeste apresentou o valor médio de SB mais elevado (33,89 mmolc.kg^{-1}) e os mais baixos valores médios de Al^{3+} (4,95 mmolc.kg^{-1}) e COT (4,90 g.kg^{-1}), o que reflete os solos desenvolvidos nessa região, geralmente, pouco intemperizados e pobres em matéria orgânica. A região Norte apresentou o valor médio de SB mais baixo (13,03 mmolc.kg^{-1}) e valores médios de Al^{3+} e H^+ elevados, 41,62 e 34,41 mmolc.kg^{-1} , respectivamente. A região Sul apresentou os valores médios mais elevados de H^+ e COT, 59,82 mmolc.kg^{-1} e 11,36 g.kg^{-1} , respectivamente. Todos os atributos químicos apresentaram $\text{CV} > 45\%$, exceto $\text{pH}_{\text{H}_2\text{O}}$.

Os coeficientes de assimetria e curtose podem ser usados para inferir sobre a normalidade dos dados. Um valor zero para ambos os coeficientes sugere que os dados apresentam distribuição normal. Não foi observada distribuição normal para nenhum atributo utilizado na geração dos modelos de regressão.

A média da densidade do solo para o Brasil foi de 1,29 Mg.m^{-3} , com mínimo e máximo de 0,57 e 1,87 Mg.m^{-3} , respectivamente e CV de 16%. O valor médio de densidade do solo para as regiões apresentou médias próximas ao valor médio de densidade do solo para o Brasil. A região Nordeste foi a que apresentou o valor mais elevado e a região Sul apresentou o valor mais baixo de densidade do solo, 1,42 e 1,10 Mg.m^{-3} , respectivamente. A densidade do solo para todos os conjuntos de dados apresentou $\text{CV} > 25\%$, sendo o maior CV observado para a região Sul e o menor para a região Nordeste, 21 e 10%, respectivamente.

Tabela 2. Estatística descritiva das variáveis dos conjuntos de dados utilizadas para geração das FPTs.

Variável	Bancos	Média	Mín	Máx	DP	CV	Ass.	Curt.
Ds (Mg.m ⁻³)	B	1,29	0,57	1,87	0,21	16	-0,35	-0,13
Ds (Mg.m ⁻³)	N	1,28	0,75	1,83	0,17	13	-0,24	-0,20
Ds (Mg.m ⁻³)	NE	1,42	0,95	1,87	0,15	10	-0,09	0,42
Ds (Mg.m ⁻³)	CO	1,20	0,77	1,79	0,20	17	0,01	-0,74
Ds (Mg.m ⁻³)	SE	1,30	0,57	1,82	0,20	15	-0,37	0,04
Ds (Mg.m ⁻³)	S	1,10	0,66	1,72	0,23	21	0,36	-0,54
AT (g.kg ⁻¹)	B	432	0	989	263,62	61	0,14	-1,11
AT (g.kg ⁻¹)	N	361	15	974	228,20	63	0,50	-0,60
AT (g.kg ⁻¹)	NE	605	10	1000	218,09	36	-0,67	-0,04
AT (g.kg ⁻¹)	CO	451	20	900	256,57	57	-0,02	-1,33
AT (g.kg ⁻¹)	SE	477	0	989	263,46	55	0,13	-1,07
AT (g.kg ⁻¹)	S	216	20	920	242,06	111	1,37	0,61
Silte (g.kg ⁻¹)	B	181	0	659	130,84	72	0,88	-0,003
Silte (g.kg ⁻¹)	N	279	6	721	148,45	53	0,08	-0,72
Silte (g.kg ⁻¹)	NE	145	0	540	100,24	69	1,10	1,07
Silte (g.kg ⁻¹)	CO	128	30	430	68,81	57	1,14	1,95
Silte (g.kg ⁻¹)	SE	137	0	530	103,17	74	1,16	0,93
Silte (g.kg ⁻¹)	S	162	12	393	75,95	46	0,13	-0,08
Argila (g.kg ⁻¹)	B	384	2	960	228,64	59	0,38	-0,80
Argila (g.kg ⁻¹)	N	358	13	910	197,68	55	0,64	-0,06
Argila (g.kg ⁻¹)	NE	249	0	831	172,31	69	0,89	0,45
Argila (g.kg ⁻¹)	CO	420	40	890	239,22	57	0,20	-1,27
Argila (g.kg ⁻¹)	SE	385	4	930	226,95	59	0,13	-0,98
Argila (g.kg ⁻¹)	S	587	40	920	250,04	42	-0,87	-0,51
pH _{H2O} (adm)	B	5,05	3,20	8,40	0,69	14	0,86	1,26
pH _{H2O} (adm)	N	4,63	3,20	6,30	0,46	10	0,16	0,64
pH _{H2O} (adm)	NE	5,44	4,00	7,70	0,79	14	0,71	-0,22
pH _{H2O} (adm)	CO	5,32	3,80	7,30	0,70	13	0,70	0,36
pH _{H2O} (adm)	SE	5,22	3,50	7,90	0,67	13	0,78	0,72
pH _{H2O} (adm)	S	5,06	3,90	6,80	0,48	10	0,71	1,12
SB (mmolc.kg ⁻¹)	B	22,87	1,00	266,00	33,96	148	2,79	9,14
SB (mmolc.kg ⁻¹)	N	13,03	1,00	93,00	12,75	97	2,05	6,38
SB (mmolc.kg ⁻¹)	NE	33,89	1,00	258,00	46,69	138	2,17	5,02
SB (mmolc.kg ⁻¹)	CO	18,69	1,00	174,00	30,25	162	2,49	6,37
SB (mmolc.kg ⁻¹)	SE	26,81	0,00	266,00	36,43	135	2,32	6,21
SB (mmolc.kg ⁻¹)	S	22,71	1,00	189,00	33,71	148	2,44	6,53
Al ³⁺ (mmolc.kg ⁻¹)	B	23,98	0,00	890,00	60,18	251	9,83	122,66
Al ³⁺ (mmolc.kg ⁻¹)	N	41,62	0,00	252,00	40,93	98	2,00	5,20
Al ³⁺ (mmolc.kg ⁻¹)	NE	4,95	0,00	53,00	7,35	148	3,38	15,27
Al ³⁺ (mmolc.kg ⁻¹)	CO	6,03	0,00	50,00	7,66	127	2,04	6,36
Al ³⁺ (mmolc.kg ⁻¹)	SE	19,68	0,00	890,00	81,84	415	9,04	85,00

Al ³⁺ (mmolc.kg ⁻¹)	S	19,08	0,00	80,00	18,55	97	0,97	0,13
H ⁺ (mmolc.kg ⁻¹)	B	29,13	0,00	218,00	27,36	94	1,83	4,92
H ⁺ (mmolc.kg ⁻¹)	N	34,41	0,50	185,00	26,07	75	2,21	7,26
H ⁺ (mmolc.kg ⁻¹)	NE	22,56	0,00	128,00	19,39	86	2,11	6,02
H ⁺ (mmolc.kg ⁻¹)	CO	32,11	0,00	136,20	27,09	84	1,65	2,87
H ⁺ (mmolc.kg ⁻¹)	SE	20,44	0,00	180,00	24,03	117	1,95	5,33
H ⁺ (mmolc.kg ⁻¹)	S	59,82	2,00	189,00	35,62	59	1,08	0,85
COT (g.kg ⁻¹)	B	7,50	0,10	44,80	6,78	90	1,80	3,67
COT (g.kg ⁻¹)	N	6,45	0,10	29,30	5,02	77	1,54	2,35
COT (g.kg ⁻¹)	NE	4,90	0,20	31,50	4,45	90	2,07	5,65
COT (g.kg ⁻¹)	CO	8,45	0,50	32,80	6,57	78	1,47	2,05
COT (g.kg ⁻¹)	SE	7,62	0,20	44,80	6,89	90	1,77	3,70
COT (g.kg ⁻¹)	S	11,36	0,31	43,7	9,10	80	1,10	0,61

B = Banco de dados referente ao Brasil; N = Banco de dados referente a região Norte; NE = Banco de dados referente a região Nordeste; CO = Banco de dados referente a região Centro-Oeste; SE = Banco de dados referente a região Sudeste; S = Banco de dados referente a região Sul; Mín = mínimo; Max. = máximo; DP = desvio padrão; CV = coeficiente de variação (%); Ass = assimetria; Curt. = curtose; Ds = densidade do solo; AT = areia total; SB = soma de bases; COT = carbono orgânico total.

4.5.2 FPTs Geradas

O resultado da análise de *stepwise* para a seleção das variáveis preditoras, tanto para o modelo geral (Brasil), como para os sub-bancos (regiões) estão apresentados na Tabela 3. Todos os modelos gerados apresentaram normalidade dos resíduos de regressão (Teste Kolmogorov-Smirnov), a 5% de significância. Em relação a colinearidade entre as variáveis preditoras, observa-se que o valor de VIF foi menor que 4 para todas as variáveis preditoras em todos os modelos gerados, portanto, não há colinearidade entre as variáveis preditoras (Tabela 3).

As variáveis incluídas no modelo geral (Brasil) foram: carbono orgânico, soma de bases, alumínio trocável, hidrogênio, argila e silte. Para a região Norte, foram incluídas no modelo as variáveis: carbono orgânico, soma de bases, pH, areia total e alumínio trocável. Para a região Nordeste, as variáveis: hidrogênio, soma de bases, argila e silte. Para a região Centro-Oeste, as variáveis: carbono orgânico, soma de bases, alumínio trocável e areia total. Para a região Sudeste, as variáveis: carbono orgânico, soma de bases, pH, areia total e alumínio trocável. Para a região Sul, as variáveis: carbono orgânico, silte e areia total.

Para todos os modelos gerados, exceto para a região Nordeste, o carbono orgânico foi selecionado como variável preditora e foi a variável que mais contribuiu para a predição de Ds. O papel dominante da matéria orgânica em relação a predição de Ds é relatado em vários trabalhos na literatura (Adams, 1973; Heuschek et al., 2005; Han et al., 2012; Gomes et al., 2017). O conteúdo de carbono orgânico relaciona-se de forma negativa com a Ds, indicando decréscimo no valor de Ds com o aumento do conteúdo de carbono orgânico. A textura do solo é outra propriedade bastante relacionada à Ds, sendo rotineiramente encontrada como variável preditora desse atributo (Bernoux et al., 1998; Tomasella & Hodnett, 1998; Benites et al., 2007; Barros & Fearnside, 2015; Gomes et al., 2017). Bernoux et al., (1998) geraram FPTs para estimar Ds na Amazônia brasileira e observaram que no modelo que considerou todos as camadas do solo, o conteúdo de argila foi o melhor preditor de Ds, respondendo por 37% da variação. Barros & Fearnside (2015) geraram modelos para estimar a densidade do solo, em Latossolos Amarelos aluminicos, na Amazônia Central, observaram que o modelo que melhor estimou a densidade nas condições desse estudo foi o que incluiu o conteúdo de

argila e o $\text{pH}_{\text{H}_2\text{O}}$ como variáveis independentes. Sendo que o conteúdo de argila foi responsável por explicar cerca de 70% da variação da densidade do solo.

Na literatura, a relação entre Ds e atributos químicos ($\text{pH}_{\text{H}_2\text{O}}$, SB, Al^{3+} , H^+) é mal abordada e uma relação direta não está claramente apresentada. Bernoux et al., (1998) e Barros & Fearnside (2015), encontraram $\text{pH}_{\text{H}_2\text{O}}$ como variável preditora de Ds. Benites et al., (2007) usaram SB como variável preditora de Ds. Gomes et al., (2017) utilizaram SB, $\text{pH}_{\text{H}_2\text{O}}$ e Al^{3+} como variáveis predictoras de Ds. As concentrações de Al^{3+} , H^+ e SB estão relacionadas à formação de agregados, onde os cátions floculantes aproximam as partículas permitindo boa agregação e, conseqüentemente, maior porosidade total, o que implica em redução dos valores de densidade do solo. Ainda tem-se um efeito indireto já que com o aumento da fertilidade natural, tem-se uma melhoria para o desenvolvimento da vegetação e conseqüentemente mais carbono é adicionado ao solo.

Tabela 3. Modelos de regressão para estimar Ds considerando o Brasil e a divisão por regiões.

Modelo (Brasil)	Int.	COT	SB	Al^{3+}	H^+	Argila	Silte	R^2_{ajust}	KS
1	1,566	-0,0102	0,0009	0,0003	-0,0014	-0,0003	-0,0002	0,5240	0,3977
VIF	-	1,7110	1,7110	1,0555	1,7590	1,0779	1,0779		
Modelo (Sudeste)	Int.	COT	SB	$\text{pH}_{\text{H}_2\text{O}}$	AT	Al^{3+}		R^2_{ajust}	KS
2	1,325	-0,0118	0,0015	-0,0346	0,0004	0,0002		0,5332	0,7344
VIF	-	1,1448	1,6969	1,7567	1,1162	1,0610			
Modelo (Norte)	Int.	COT	SB	$\text{pH}_{\text{H}_2\text{O}}$	AT			R^2_{ajust}	KS
3	0,981	-0,0183	-0,0015	0,0775	0,0002			0,532	0,6639
VIF	-	1,4436	1,1708	1,3053	1,0246				
Modelo (Nordeste)	Int.	H^+	SB	Argila	Silte			R^2_{ajust}	KS
4	1,572	-0,0025	0,0004	-0,0002	-0,0002			0,2805	0,6611
VIF	-	1,1101	1,3412	1,1640	1,3046				
Modelo (Centro-Oeste)	Int.	COT	SB	Al^{3+}	AT			R^2_{ajust}	KS
5	0,965	-0,0102	0,0021	0,0044	0,0005			0,7698	0,3452
VIF	-	1,2646	1,1782	1,1866	1,2194				
Modelo (Sul)	Int.	COT	Silte	AT				R^2_{ajust}	KS
6	0,926	-0,0092	0,0008	0,0006				0,6453	0,5475
VIF	-	1,1921	1,1671	1,1033					

KS (p-valor) – valores acima de 0,05 significa que os resíduos da regressão são normalmente distribuídos (KS=Teste de Kolmogorov-Smirnov); SB = Soma de bases; VIF = Fator de Inflação da Variância; Comoreferência, quando o valor de VIF dos atributos do solo for menor que 4, significa que o respectivo atributo não apresenta colinearidade com as demais variáveis predictoras.

4.5.3 Avaliação do Desempenho das FPTs

A comparação do desempenho das FPTs foi feita aplicando-se os modelos gerados e o proposto por Benites et al. (2007) para prever a densidade do solo aos conjuntos de dados de

validação (20% selecionados para validação). Os gráficos de dispersão entre os valores medidos e estimados de Ds e as medidas de erro de estimativa para todos os modelos são apresentados na Figura 2 e Tabela 4, respectivamente.

As medidas de erro de estimativa aplicadas neste trabalho foram MPE, MAE, RMSPE, R^2 , MR, além da avaliação dos gráficos de dispersão. A FPT de melhor performance será aquela que apresentar valores de MPE, MAE, RMSPE mais próximos de 0 e o valor de R^2 mais próximo de 1.

Todos os modelos gerados, exceto para a região Sul, apresentaram valores de MPE, MAE, RMSPE mais baixos e R^2 mais elevados quando comparados ao modelo proposto por Benites et al. (2007) aplicado aos dados de validação de cada sub-banco. O modelo gerado para a região Nordeste apresentou os valores de MAE, RMSPE iguais aos valores encontrados para o modelo proposto por Benites et al. (2007), além de ter apresentado o valor mais baixo de R^2 .

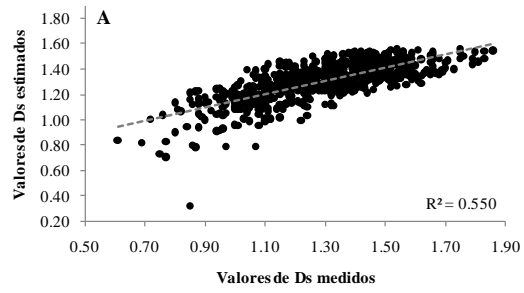
Comparando-se o desempenho do modelo geral (Brasil) em relação aos modelos gerados por regiões, observou-se menores valores de MPE, MAE e RMSPE, para os modelos gerados por regiões, exceto para a região Sul. Em relação ao MPE, os modelos gerados para as regiões Norte e Nordeste apresentaram o menor viés, pois alcançaram o valor referência (MPE = 0). Os modelos gerados para a região Sudeste e Sul tenderam a subestimar os valores de Ds (MPE = -0,01 e -0,03 Mg.m^{-3} , respectivamente), sendo o maior valor de MPE observado para o modelo gerado para a região Sul. Os valores de MAE variaram entre 0,03 e 0,12 Mg.m^{-3} , sendo o valor mais baixo observado para o modelo gerado para a região Sudeste (MAE = 0,03 Mg.m^{-3}) e o mais alto observado para a região Sul (MAE = 0,12 Mg.m^{-3}). O mesmo foi observado para o RMSPE, sendo o valor mais baixo observado para o modelo gerado para a região Sudeste (MAE = 0,07 Mg.m^{-3}) e o mais alto observado para a região Sul (MAE = 0,15 Mg.m^{-3}).

A melhora relativa (MR) nos permite ter ideia da melhora na acurácia dos modelos gerados em relação ao modelo proposto por Benites et al. (2007). O maior valor de MR foi observado para o modelo gerado para a região Centro-Oeste (MR = 35,7%), ou seja, esse modelo apresentou uma melhoria em acurácia cerca de 36% a mais que o modelo proposto por Benites et al. (2007). Também foi o modelo que apresentou o maior valor de R^2 ($R^2 = 0,829$). A região Norte teve o segundo maior valor de MR (MR = 29,4), ou seja, aproximadamente 29% a mais de acurácia em relação ao modelo proposto por Benites et al. (2007). Comparando-se o modelo gerado para o Brasil ao proposto por Benites et al. (2007), observou-se uma melhoria de apenas 6,6%, o valor de RM mais baixo encontrado. Para a região Sul não foi observado melhoria na acurácia em relação ao modelo proposto por Benites et al. (2007), sendo este 7,1% mais acurado.

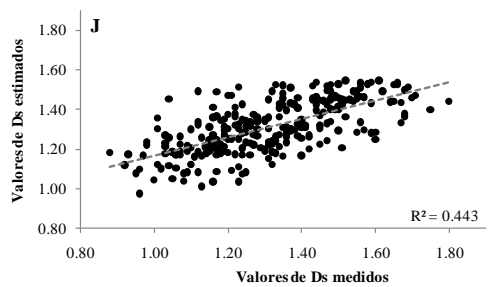
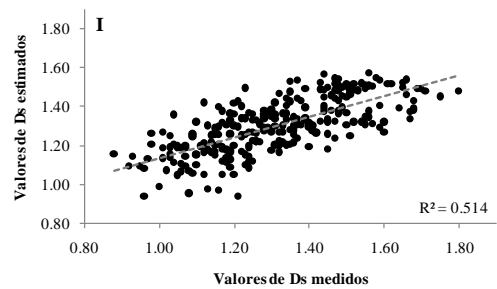
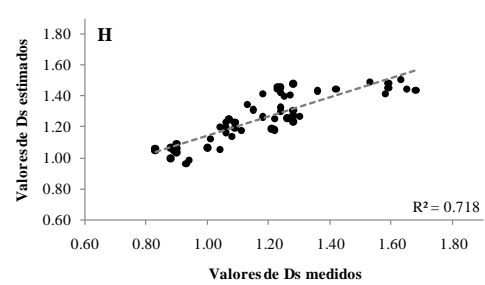
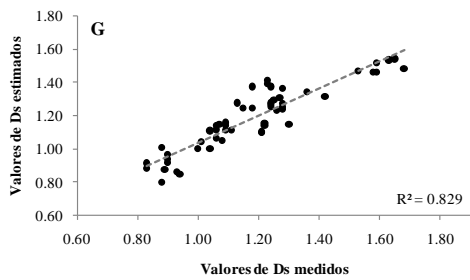
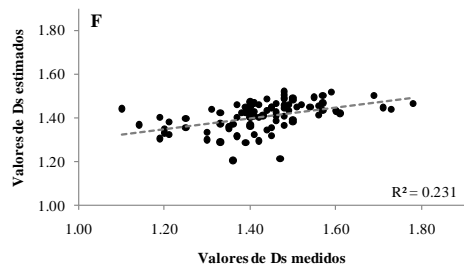
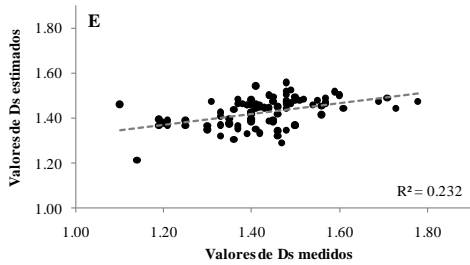
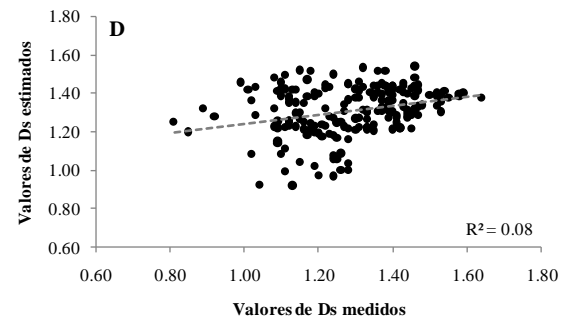
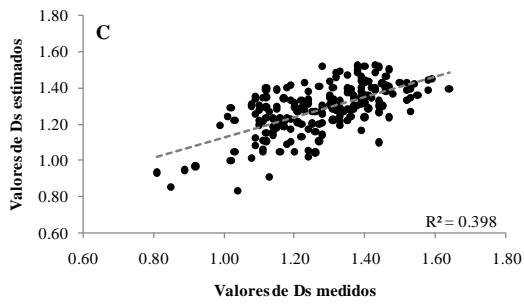
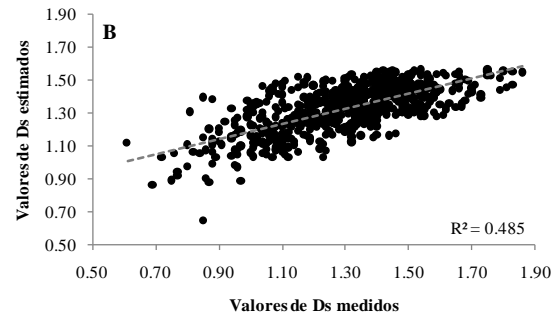
O número de amostras por região e a distribuição dos pontos na área amostrada influenciam nos resultados obtidos pelos modelos de regressão. A subdivisão por regiões ainda abrange muita heterogeneidade devido a particularidades e variações de solos dentro de cada região, no entanto, quando se considera um modelo para todo o Brasil, essas variações são ainda maiores. O que reflete na melhor acurácia dos modelos por regiões comparados ao modelo proposto por Benites et al. (2007).

Nesse contexto, se reforça a questão da necessidade de um maior número de pontos associado a uma distribuição mais uniforme destes pontos, de forma que a amostra represente a maior parte da variação presente na área amostrada. Chen et al. (2018) sugerem que uma estratégia de amostragem sistemática e densa pode ajudar muito numa modelagem mais robusta.

Modelos propuestos



Benites et al. (2007)



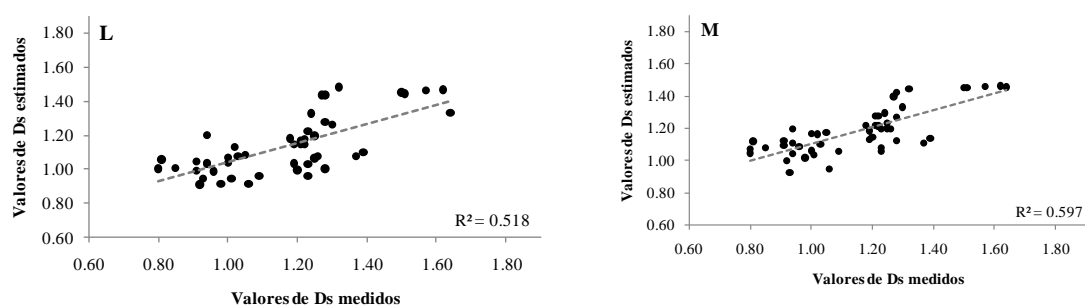


Figura 2. Valores medidos vs valores estimados de densidade do solo (Mg.m^{-3}) para os modelos gerados comparados ao proposto por Benites et al. (2007) aplicado aos conjuntos de validação. A) Modelo gerado para o Brasil; C) Modelo gerado para a região Norte; E) Modelo gerado para a região Nordeste; G) Modelo gerado para a região Centro-Oeste; I) Modelo gerado para a região Sudeste; L) Modelo gerado para a região Sul; B, D, F, H, J, M) Modelo proposto por Benites et al. (2007).

Tabela 4. Medidas de erro de predição para os modelos gerados e o proposto por Benites et al. (2007).

Modelos	MPE	MAE	RMSPE	R ²	MR (%)
Centro-Oeste	0,01	0,07	0,09	0,829	35,7
Benites et al. (2007)	0,08	0,12	0,14	0,718	-
Norte	0,00	0,10	0,12	0,398	29,4
Benites et al. (2007)	0,03	0,13	0,17	0,080	-
Nordeste	0,00	0,08	0,11	0,232	0
Benites et al. (2007)	-0,02	0,08	0,11	0,231	-
Sudeste	-0,01	0,03	0,07	0,514	12,5
Benites et al. (2007)	-0,01	0,04	0,08	0,443	-
Sul	-0,03	0,12	0,15	0,518	-7,1
Benites et al. (2007)	0,03	0,11	0,14	0,597	-
Brasil	0,01	0,11	0,14	0,550	6,6
Benites et al. (2007)	0,03	0,12	0,15	0,485	-
Valores de referência	0	0	0	1	

MPE – erro médio de predição; MAE – erro médio absoluto; RMSPE – raiz do quadrado médio do erro de predição; MR – melhora relativa.

4.6 CONCLUSÕES

Quando se considera um modelo geral para prever a Ds, o modelo gerado para o Brasil, apresentou melhor desempenho, aproximadamente 7% a mais de acurácia comparado ao modelo pioneiro proposto por Benites et al. (2007).

Foi observado um melhor desempenho na predição das FPTs quando se subdividiu o banco geral (considerando o Brasil) por regiões, com exceção do modelo gerado para a região Sul.

Todos os modelos gerados obtiveram melhor desempenho quando comparados ao modelo proposto por Benites et al. (2007), com exceção dos modelos gerados para as regiões Sul e Nordeste.

Para a região Sul, a FPT gerada para o Brasil ou a proposta por Benites et al. (2007), obtiveram melhor desempenho, podendo ser utilizadas para predição de Ds nessa região.

5 CONCLUSÕES GERAIS

Este trabalho investigou o uso de algumas covariáveis multiespectrais e de radar com diferentes resoluções (espacial e espectral) para mapear frações granulométricas e estoque de carbono na região norte e também o uso de FPTs para estimar densidade do solo para o Brasil e suas regiões.

Embora as melhorias tenham sido modestas e limitadas a alguns atributos do solo, existe potencial para o uso de dados multispectrais e de radar na Amazônia para melhorar a precisão das predições de atributos do solo. No entanto, pesquisas adicionais são necessárias para conectar melhor as variáveis de formação de solo e as derivadas de sensoriamento remoto com a distribuição espacial dos atributos do solo na Amazônia, incluindo os índices de vegetação e variáveis de relevo mais comumente usados, mas especialmente aqueles ainda pouco explorados como dados de radar da banda L.

Por outro lado, entre covariáveis multiespectrais e de radar, apenas o NDVI derivado de imagem RapidEye apresentou continuidade espacial cruzada para predição de CS100 por cokrigagem. Isso pode estar relacionado à resolução espacial dos dados, pois apenas covariáveis de resolução fina (RapidEye e de relevo) foram utilizadas para cokrigagem, contrariamente ao radar de resolução mais baixa (12,5 m) e aos dados Landsat 8 (30 m).

Isso é importante porque ICOK e HCOK apresentaram menor RMSE na predição de CS100, portanto, pode valer a pena testar outros dados de resolução mais alta na busca de melhorias mais expressivas que as encontradas neste estudo que atingiram um máximo de 8%.

Quando se considera um modelo geral para prever a Ds, o modelo gerado para o Brasil, obteve melhor desempenho, aproximadamente 7% a mais de acurácia comparado ao modelo pioneiro proposto por Benites et al. (2007).

Foi observado um melhor desempenho na predição das FPTs quando se subdividiu o banco geral (considerando o Brasil) por regiões, com exceção do modelo gerado para a região Sul.

Todos os modelos gerados obtiveram melhor desempenho quando comparados ao modelo proposto por Benites et al. (2007), com exceção dos modelos gerados para as regiões Sul e Nordeste.

Para a região Sul, a FPT gerada para o Brasil ou a proposta por Benites et al. (2007), obtiveram melhor performance, podendo ser utilizadas para predição de Ds nessa região.

Pesquisas futuras são relevantes no que se refere ao uso de outras fontes de covariáveis derivadas do sensoriamento remoto, uma vez que, mostrou-se existir potencial para o uso desse tipo de covariável na predição de alguns atributos do solo. Reforça-se também a necessidade de mais informações sobre as propriedades dos solos do Brasil, a fim de compor um banco para um melhor gerenciamento dos dados de solos no país.

6 REFERÊNCIAS BIBLIOGRÁFICAS

- ADAMS, W. A. The effect of organic matter on the bulk and true densities of some uncultivated podzolic soils. **Journal of Soil Science**, v. 24, p. 10-17, 1973.
- AHMED, Z.; IQBAL, J. Evaluation of Landsat TM5 multispectral data for automated mapping of surface soil texture and organic matter in GIS. **European Journal Remote Sensing**, v. 47, 557–573, 2014.
- ARAÚJO, E. A.; KER, J. C.; MENDONÇA, E. S.; SILVA, I. R.; OLIVEIRA, E. K. Impacto da conversão floresta - pastagem nos estoques e na dinâmica do carbono e substâncias húmicas do solo no bioma Amazônico. **Revista Acta Amazônica**, v. 41, p. 103 – 114, 2012.
- BARROS, H. S.; FEARNSIDE, P. M. Pedo-transfer functions for estimating soil bulk density in central Amazonia. **Revista Brasileira de Ciência Solo**, v. 39, p. 397-407, 2015.
- BEN-DOR, E.; TAYLOR, R. G.; HILL, J.; SOMMER, S. Imaging Spectrometry for Soil Applications. **Advances in Agronomy**, v. 97, p. 321-392, 2008.
- BENITES, V. M.; MACHADO, P. L. O. A.; FIDALGO, E. C. C.; COELHO, M. R.; MADARI, B. E. Pedotransfer functions for estimating soil bulk density from existing soil survey reports in Brazil. **Geoderma**, Amsterdam, v. 139, p. 90–97, 2007.
- BERNOUX, M.; CARVALHO, M. C. S.; VOLKOFF, B.; CERRI, C. C. Brazil's soil carbon stocks. **Soil Science Society America Journal**, v. 66, p. 888–896, 2002.
- BERNOUX, M.; ARROUAYS, D.; CERRI, C.; VOLKOFF, B.; JOLIVET, C. Bulk densities of Brazilian Amazon soils related to other soil properties. **Soil Science Society of America Journal**, Madison, v. 62, p. 743 – 749, 1998.
- BISHOP, T. F. A.; MCBRATNEY, A. B. A comparison of prediction methods for the creation offield-extent soil property maps. **Geoderma**, v. 103, p. 149–160, 2001.
- BOUMA, J. Using soil survey data for quantitative land evaluation. **Advances in Soil Science**, New York, v. 9, p. 177-213, 1989.
- CEDDIA, M. B.; VILLELA, A. L. O.; PINHEIRO E.F.M.; WENDROTH, O. Spatial variability of soil carbon stock in the Urucu river basin, Central Amazon-Brazil. **Science of the Total Environment**, v. 526, p. 58–69, 2015.
- CERRI, C. E. P.; EASTER, M.; PAUSTIAN, K.; KILLIAN, K.; COLEMAN, K.; BERNOUX, M.; FALLOON, P.; POWLSON, D. S.; BATJES, N. H.; MILNE, E.; CERRI, C.C. Predicted soil organic carbon stocks and changes in the Brazilian Amazon between 2000 and 2030. Agriculture, **Ecosystems and Environment**, v. 122, p. 58 – 72, 2007.
- CEZAR, E.; NANNI, M. R.; DEMATTE, J. A. M.; CHICATI, M. L.; OLIVEIRA, R. B. Estimativa de atributos do solo por meio de espectrorradiometria difusa. **Revista Brasileira de Ciência do Solo**, v. 37, p. 858 – 868, 2013.
- CHAGAS, C. S.; JUNIOR, W. C.; BHERING, S. B.; CALDERANO-FILHO, B. Spatial prediction of soil surface texture in a semiarid region using random forest and multiple linear regressions. **Catena**, v. 139, p. 232 – 240, 2016.
- CHAVES, J. M.; SANO, E. E.; GUIMARÃES, E. M.; SILVA, A. B.; MENESES, P. R. Sinergismo entre dados ópticos e de radar para o estudo geológico na região de Bezerra-Cabeceiras, Goiás. **Revista Brasileira de Geociências**, v. 33, n. 2 - Suplemento, p. 137-146, 2003.

- CHEN, S.; RICHER-DE-FORGES, A. G.; SABY, N. P. A.; MARTIN, M. P.; WALTER, C.; ARROUAYS, D. Building a pedotransfer function for soil bulk density on regional dataset and testing its validity over a larger area. **Geoderma**, v. 312, p. 52 – 63, 2018.
- CHIG, L. A.; COUTO, E. G.; NOVAES FILHO, J. P.; RODRIGUES, L. C. M.; JOHNSON, M. S.; WEBER, O. L. S. Distribuição espacial da granulometria, cor e carbono orgânico do solo ao longo de um transecto em microbacias na Amazônia meridional. **Acta Amazônica**, v. 38, não paginado, 2008.
- DEMARCHI, J. C.; PIROLI, E. L.; ZIMBACK, C. R. L. Análise temporal do uso do solo e comparação entre os índices de vegetação NDVI e SAVI no município de Santa Cruz do Rio Pardo – SP usando imagens Landsat-5. **RA E GA**, Curitiba, v. 21, p. 234-271, 2011.
- DE VOS, B.; MEIRVENN, M.V.; QUATAE, P.; DECKER, J.; MUYS, B. Predictive quality of pedotransfer functions for estimating bulk density of forest soils. **Soil Science Society of America Journal**, v. 69, p. 500-510, 2005.
- DJUUNA, I. A. F.; ABBOTT, L.; RUSSELL, C. Determination and Prediction of Some Soil Properties using Partial Least Square (PLS) Calibration and Mid-Infra Red (MIR) Spectroscopy Analysis. **Journal Tropical Soils**, v. 16, p. 93 – 98, 2011.
- DOBOS, E.; MICHELI, E.; BAUMGARDNER, M.F.; BIEHL, L.; HELT, T. Use of combined digital elevation model and satellite radiometric data for regional soil mapping. **Geoderma** v. 97, p. 367–391, 2000.
- FIDALGO, E. C. C.; BENITES, V. M.; MACHADO, P. L. O. A.; MADARI, B. E.; COELHO, M. R.; MOURA, I. B.; LIMA, C. X. **Estoque de carbono nos solos do Brasil. Rio de Janeiro: Boletim de Pesquisa e Desenvolvimento (Embrapa Solos)**, n.121, 2007.
- FREITAS, D. M.; SANO, E. E.; SOUZA, R. A. Potencial das imagens multipolarizadas do satélite ALOS PALSAR na discriminação de cobertura vegetal do bioma Pantanal: Estudo de caso na região do médio Taquari, MS. **Revista Brasileira de Cartografia**, v. 66, p. 209 – 221, 2014.
- GOMES, A. S.; FERREIRA, A. C. S.; PINHEIRO, E. F. M.; MENEZES, M. D.; CEDDIA, M. B. The use of pedotransfer functions and the estimation of carbon stock in the Central Amazon region, **Scientia Agricola**, v. 74, n. 6, p. 450-460, 2017.
- GONÇALVES, M. G.; FARIAS, L. N.; COSTA, A. D.; MOTA, L. S.; OLIVEIRA, B. L.; CEDDIA, M. B.; VIEIRA, S.R. Contribuição da geoestatística para o planejamento experimental de uma pastagem do sistema integrado de produção agroecológica (Fazendinha km 47). **Revista Universidade Rural**, Seropédica, v. 25, p. 1 – 5, 2005.
- GOOVAERTS, P. **Geostatistics for natural resource evaluation**. Oxford University Press, New York, 1997.
- HAN, G. Z.; ZHANG, G. L.; GONG, Z. T.; WANG, G. F. Pedotransfer functions for estimating soil bulk density in China. **Soil Science**, v. 177, p. 158-164, 2012.
- HENGL, T.; HEUVELINK, G. B. M.; ROSSITER, D. G. About regression-kriging: From equations to case studies. **Computers & Geosciences**, v. 33, p. 1301–1315, 2007.
- HENGL, T.; HEUVELINK, G. B. M.; STEIN, A. A generic framework for spatial prediction of soil variables based on regression-kriging. **Geoderma**, v. 120, p. 75 – 93, 2004.
- HERBST, M.; DIEKKRÜ, B.; VERECKEN, H. Geostatistical coregionalization of soil hydraulic properties in a micro-scale catchment using terrain attributes. **Geoderma**, v. 132, p. 206-221, 2006.

- HEUSCHER, S. A.; BRANDT, C. C.; JARDINE, M. P. Using soil physical and chemical properties to estimate bulk density. **Soil Science Society of America Journal**, v. 106, p. 52-62, 2005.
- HUETE, A. R.; DIDAN, K.; SHIMABUKURO, Y.; RATANA, P.; SALESKA, S. R.; HUTYRA, L. R.; YANG, W.; NEMANI, R. R.; MYNENI, R.. Amazon rainforests green-up with sunlight in dry season. **Geophysical Research Letters**, v.33, 2006.
- ISAACS, E. H. & SRIVASTAVA, R. M. **Applied Geostatistics**. New York: Oxford University Press, 1989. 600 p.
- JENNY, H. **Factors of Soil Formation**. McGraw-Hill, New York, 1941.
- KAUR, R.; KUMAR, S.; GURUNG, H. P. A pedo-transfer function (PTF) for estimating soil bulk density from basic soil data and its comparison with existing PTFs. **Australian Journal of Soil Research**, v. 40, p. 847-857, 2002.
- LANDIM, P. M. B. **Sobre Geoestatística e mapas**. Terra e Didática, v. 2, p. 19-33, 2006. Disponível em: <<http://www.ige.unicamp.br/terraedidatica/>> Acesso em 7 de janeiro de 2014.
- LARK, R. M.; BISHOP, T. F. A. Cokriging particle size fractions of the soil. **European Journal Soil Science**, v. 58, p. 763–774, 2007.
- LEITE, M. E.; ALMEIDA, M. I. S.; VELOSO, G. A.; FERREIRA, M. F. F. Sensoriamento remoto aplicado ao mapeamento da dinâmica do uso do solo na bacia do rio Pacuí, no norte de Minas Gerais, nos anos de 1989, 1999 e 2009. **Revista do Departamento de Geografia – USP**, v. 23, p. 217-231, 2012.
- LIAO, K.; XU, S.; WU, J.; ZHU, Q. Spatial estimation of surface soil texture using remote sensing data. **Soil Science Plant Nutrition**, v. 59, p. 488–500, 2013.
- LOPES, F.; MIELNICZUK, J.; OLIVEIRA, E. S. & TORNQUIST, C. G. Evolução do uso do solo em uma área piloto da região de Vacaria, RS. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.14, n.10, p.1038–1044, 2010.
- MACHADO, P. L. O. A. Carbono do solo e a mitigação da mudança climática global. **Revista Química Nova**, v. 28, p. 329-334, 2005.
- MASELLI, F.; GARDIN, L.; BOTTAI, L. Automatic mapping of soil texture through the integration of ground, satellite and ancillary data. **International Journal of Remote Sensing**, v. 29, p. 5555–5569, 2008.
- MANRIQUE, L. A.; JONES, C. A.; DYKE, P. T. Predicting cation exchange capacity from soil physical and chemical properties. **Soil Science Society of America Journal**, v. 55, p. 787–794, 1991.
- MELO, E. T.; SALES, M. C. L.; OLIVEIRA, J. G. B. Aplicação do índice de vegetação por diferença normalizada (NDVI) para a análise da degradação ambiental do microbacia hidrográfica do Riacho dos Cavalos, Crateús – CE. **RA'E GA**, v. 23, p. 520 – 533, 2011.
- MENESES, P. R.; ALMEIDA, T. Aritmética de Bandas. In: MENESES, P. R.; ALMEIDA, T. (eds). **Introdução ao processamento de imagens de sensoriamento remoto**. Brasília, 2012. Cap 9, p. 138 - 152.
- MINASNY, B.; McBRATNEY, A. B.; SANTOS, M. de L. M.; SANTOS, H. G. dos. **Revisão de pedotransferência (PTFs) e novos métodos de predição de classes e atributos do solo**. Rio de Janeiro: Embrapa Solos, 2003. 50p. (Documentos, 45).

- MORAN, M. S.; HYMER, D. C.; QI, J.; KERR, Y. Comparison of ERS-2 SAR and Landsat TM imagery for monitoring agricultural crop and soil conditions. **Remote Sensing of Environment**, v. 79, p. 243-252, 2002.
- MUELLER, T.G.; PIERCE, F.J. Soil carbon maps: enhancing spatial estimates with simple terrain attributes at multiple scales. **Soil Science Society American Journal**, v. 67, p. 258–267, 2003.
- MULDER, V. L.; BRUIN, S.; SCHAEPMAN, M. E.; MAYR, T. R. The use of remote sensing in soil and terrain mapping – A review. **Geoderma**, v. 162, p. 1 – 19, 2011.
- NETA, S. R. A.; FREITAS, C. C.; DUTRA, L. V. Uso de imagens ALOS PALSAR multipolarizadas para detecção de incremento de desflorestamento na Amazônia. **Revista Brasileira de Cartografia**, v. 62, p. 417 – 431, 2010.
- NUNES, G. M.; FILHO, C. R. S.; FERREIRA, L. G. Discriminação de fitofisionomias na Amazônia Central por meio de índices de vegetação de imagens com resolução espacial moderada. **Revista Geográfica Acadêmica**, v.6, p. 05-14, 2012.
- ODEH, I. O. A.; MCBRATNEY, A. B.; CHITTLEBOROUGH, D. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. **Geoderma**, v. 67, p. 215-226, 1995.
- ODEH, I. O. A.; MCBRATNEY, A. B.; CHITTLEBOROUGH, D. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. **Geoderma**, v. 63, p. 197-214, 1994.
- PARADELLA, W. R.; SANTOS, A. R.; VENEZIANI, P.; CUNHA, E. S. P. Radars imageadores nas Geociências: estado da arte e perspectivas. **Revista Brasileira de Cartografia**, v. 57, n. 1, p. 56-62, 2005.
- R Development Core Team, 2014. **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing, Vienna, Austria. Version 3.1.1. 3-900051-07-0 (URL<http://www.R-project.org>).
- RIVERO, R. G.; GRUNWALD, S.; BRULAND, G. L. Incorporation of spectral data into multivariate geostatistical models to map soil phosphorus variability in a Florida wetland. **Geoderma**, v 140, p. 428–443, 2007.
- SANO, E. E.; SANTOS, E. M.; MENESES, P. R. Análise de imagens do satélite Alos Palsar para o mapeamento de uso e cobertura da terra do Distrito Federal. **Geociências**, v. 28, p. 441 – 451, 2009.
- SUMFLETH, K.; DUTTMANN, R. Prediction of soil property distribution in paddy soil landscapes using terrain data and satellite information as indicators. **Ecological Indicators**. v. 8, p. 485–501, 2008.
- TAMMINEN, P.; STARR, M. Bulk density of forested mineral soils. **Silva Fennica**, v. 28, p. 53-60, 1994.
- TIETJE, O.; TAPKENHINRICHS, M. Evaluation of pedotransfer functions. **Soil Science Society of America Journal**, Madison, v.57, p.1088– 1095, 1993.
- TOMASELLA, J.; HODNETT, M.G. Estimating soil water retention characteristics from limited data in Brazilian Amazonia. **Soil Science**,v. 163, p. 190-202, 1998.

- TORNQUIST, C. G.; GASSMAN, P. W.; MIELNICZUK, J.; GIASSON, E.; CAMPBELL, T. Spatially explicit simulations of soil C dynamics in Southern Brazil: Integrating Century and GIS with iCentury. **Geoderma**, v.150. p. 404 - 414, 2009.
- VALENCIA, L. I. O.; MEIRELLES, M. S. P.; BETTINI, C. **Geoestatística Aplicada à Agricultura de Precisão**. In: MACHADO, P. L. O. A.; BERNARDI, A. C. C.; SILVA, C. A. Agricultura de Precisão para o Manejo da Fertilidade do Solo em Sistema Plantio Direto. Rio de Janeiro: Embrapa Solos, 2004. Cap. 02, p. 37 – 56.
- VIEIRA S. R. **Geoestatística em estudos de variabilidade espacial do solo**. In: NOVAIS, R.F.; ALVAREZ V. V. H. & SCHAEFER, C. E. G. R. (Eds). Tópicos em Ciência do Solo. Viçosa. Sociedade Brasileira de Ciência do Solo, v. 1, p. 1-53, 2000.
- VISCARRA-ROSSEL, R. A.; WALVOORT, D. J. J.; MCBRATNEY, A. B.; JANIK, L. J.; SKJEMSTAD, J. O. Visible, near-infrared, mid-infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. **Geoderma**, v. 131, p. 59 – 75, 2006.
- WILLMOTT, C. J. Some comments on the evaluation of model performance. **Bulletin American Meteorological Society**, v. 63, p. 1309 – 1317, 1982.
- WILLIWANS, J.; ROSS, P.; BRISTOW, K. **Prediction of the Campbell water retention function from texture, structure, and organic matter**. In: VAN GENUCHTEN, M. T. H., LEIJ, F. J., LUND, L.J. (Eds.), Proc. Int. Workshop on Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils. University of California, Riverside, 1992, p. 427–442.
- WÖSTEN, J.H.M.; PACHEPSKY, Y. A. A.; RAWLS, W. J. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. **Hydrology Journal**, v. 251, p.123–150, 2001.