

**TÂNIA MARIA DE CARVALHO**

**MODELAGEM DIGITAL DE ATRIBUTOS DE SOLO  
DA FAZENDA EDGÁRDIA - BOTUCATU-SP**

**Botucatu**

**2016**

**TÂNIA MARIA DE CARVALHO**

**MODELAGEM DIGITAL DE ATRIBUTOS DE SOLO  
DA FAZENDA EDGÁRDIA - BOTUCATU-SP**

Tese apresentada à Faculdade de Ciências Agronômicas da Unesp Câmpus de Botucatu, para obtenção do título de Doutor em Agronomia (Energia na Agricultura).

Orientadora: Prof<sup>a</sup> Dr<sup>a</sup> Célia Regina Lopes  
Zimback

**Botucatu**

**2016**

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉCNICA DE AQUISIÇÃO E TRATAMENTO DA INFORMAÇÃO - DIRETORIA TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - UNESP - FCA - LAGEADO - BOTUCATU (SP)

C331m Carvalho, Tânia Maria de, 1965-  
Modelagem digital de atributos de solo da Fazenda Edgárdia - Botucatu-SP / Tânia Maria de Carvalho. - Botucatu : [s.n.], 2016  
143 p. : ils. color., grafs., tabs.

Tese (Doutorado) - Universidade Estadual Paulista, Faculdade de Ciências Agronômicas, Botucatu, 2016  
Orientador: Célia Regina Lopes Zimback  
Inclui bibliografia

1. Geoprocessamento. 2. Modulação digital. 3. Solos. 4. Krigagen. 5. Cuestas. I. Zimback, Célia Regina Lopes. II. Universidade Estadual Paulista "Júlio de Mesquita Filho" (Câmpus de Botucatu). Faculdade de Ciências Agronômicas. III. Título.

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte"



## CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: MODELAGEM DIGITAL DE ATRIBUTOS DE SOLO DA FAZENDA EDGARDIA -  
BOTUCATU-SP

**AUTORA: TÂNIA MARIA DE CARVALHO**

**ORIENTADORA: CÉLIA REGINA LOPES ZIMBACK**

Aprovada como parte das exigências para obtenção do Título de Doutora em AGRONOMIA  
(ENERGIA NA AGRICULTURA) pela Comissão Examinadora:



Prof. Dr. CÉLIA REGINA LOPES ZIMBACK  
Depto de Solos e Recursos Ambientais / Faculdade de Ciências Agrônomicas - UNESP



Prof. Dr. PAULO MILTON BARBOSA LANDIM  
Depto de Solos e Recursos Ambientais / UNESP - Faculdade de Ciências Agrônomicas de Botucatu



Profa. Dra. ALESSANDRA FAGIOLI DA SILVA  
Departamento de Engenharia Rural e Desenvolvimento Agrário / Universidade Estadual do Norte do Paraná -  
Campus Luiz Meneguel



Dra. BRUNA SOARES XAVIER DE BARROS  
./ Botucatu/SP



Prof. Dr. DIEGO AUGUSTO DE CAMPOS MORAES  
Depto de Análise e Desenvolvimento de Sistemas / Faculdade Eduvale de Avaré

Botucatu, 19 de dezembro de 2016



Às vidas entrelaçadas à minha vida,  
que me inspiram sonhos e impulsionam às realizações:  
meus pais Aurélio (*in memoriam*) e Maria Célia,  
meus irmãos e sobrinhos,  
e especialmente à Débora e Marina,



## **AGRADECIMENTOS**

À minha família, pelo amor e apoio permanente.

À Profª Drª Célia Regina Lopes Zimback, pela orientação, confiança, apoio e paciência. Por ser uma verdadeira mestra, sobre as ciências e sobre a vida.

Ao Prof. Dr. Paulo Milton Barbosa Landim, pelas sugestões em todas as etapas do trabalho e inestimáveis ensinamentos.

Ao Prof. Dr. Luiz Alberto Blanco Jorge pelos dados, ensinamentos e especialmente as conversas sobre a Fazenda Edgárdia.

Aos professores e funcionários dos Departamentos de Solos e Recursos Ambientais e Engenharia Rural, pelos ensinamentos, auxílios práticos e convivência.

Aos funcionários da Biblioteca “Prof. Paulo de Carvalho Mattos”, pela convivência, presteza e pelo ambiente aconchegante dentro da FCA.

Ao GEPAG, por me prover estrutura física para o desenvolvimento do trabalho e pela rica e prazerosa convivência com cada um dos membros, numa diversidade enriquecedora de contribuições, em um ambiente de incentivo e amizade.

Aos colegas da pós-graduação, especialmente Roberto Filgueiras e Donizeti Nicolete, que muito me acompanharam, opinaram, auxiliaram e incentivaram em relação à pesquisa.

À profª Regina Célia de Carvalho, da FAMEV/ UFMT pelo incentivo e por compartilhar a sua experiência de vida e trabalho acadêmico.

À profª Virindiana Colet Bortolotto, minha colega de área, de sala e de Instituto na UFMT, que acompanhou à distância minha trajetória durante esse curso, pelas sugestões e auxílio.

Ao Prof. Dr. Frederico Terra de Almeida e à Laura Cristini, da UFMT, pelo apoio.

Ao programa de capacitação da Universidade Federal de Mato Grosso.

À FCA/Unesp, representando a universidade pública e gratuita.

Muito obrigada!



## RESUMO

O mapa de solos é uma ferramenta essencial para o planejamento de uso da terra e estudos que envolvem aspectos ambientais relativos a esse importante recurso natural. Técnicas quantitativas e ferramentas de geoprocessamento têm sido aliadas à interpretação dos processos pedogenéticos para possibilitar a elaboração de mapas mais precisos, obtidos por processo mais rápido e menos oneroso. Dentre os modelos aplicados, os denominados modelos híbridos empregam variáveis auxiliares preditoras e autocorrelação espacial, para viabilizar a predição de atributos de solo em locais não amostrados. A iniciativa para mapeamento digital do solo em escala mundial – *GlobalSoilMap.net* atua no sentido de disponibilizar representações globais de atributos de solo, elaboradas por meio da aplicação de modelo híbrido em dados legados de solos, realizando a prática do Mapeamento Digital de Solos (MDS). Com base nesse princípio, esse trabalho baseou-se na hipótese de que a aplicação da técnica híbrida regressão-krigagem, utilizando dados legados de levantamento de solo e covariáveis de relevo e sensoriamento remoto proveem mapa de atributos de solo representativos de uma área da Cuesta de Botucatu. O modelo foi aplicado localmente, a duas profundidades, para representação contínua do Índice de Avermelhamento (IAV), saturação de bases (V%), teor de areia, teor de argila, CTC e pH dos solos da Fazenda Experimental Edgárdia, para a qual são disponíveis dados de levantamento de solo. As covariáveis preditoras derivadas de um MDE e de imagem orbital foram uniformizadas a uma resolução espacial de 10 m, e os métodos foram selecionados de acordo com a verificação de correlação linear significativa entre atributos e covariáveis e autocorrelação espacial dos atributos ou dos resíduos de regressões lineares múltiplas (RLM). Os dados foram separados em subconjuntos de treinamento e validação. Os coeficientes de correlação entre atributos de solo e covariáveis foram significativos e variaram de -0,40 a 0,51. Os preditores mais correlacionados aos atributos foram Índice Topográfico de Umidade (ITU), Declividade (Decl), Aspecto (Aspc), Elevação (Elev) e índice de vegetação NDVI, sendo os quatro últimos os principais na estimação das frações texturais. Os valores de  $R^2$  ajustado das RLM, entre 0,10 e 0,36, foram considerados baixos. De modo geral, os mapas de predição expuseram padrões característicos da variação espacial observada nos mapas das covariáveis preditoras, usadas na calibração dos

modelos. Foi observado um incremento na acurácia entre as duas etapas do processo de RK, indicando que o mapa final é superior em relação à RLM. No entanto, os modelos apresentaram, de modo geral, um baixo desempenho quando avaliados por meio de validação externa, mesmo com a estratificação em duas áreas mais uniformes em termos de relevo. Os resultados indicaram a limitação do uso de amostragem para fins de levantamento em modelos de predição. Houve ainda dificuldade de aplicação dos modelos em função do contexto litológico complexo e da dinâmica local de formação de solos, que não puderam ser detectadas pelas covariáveis selecionadas. Apesar das limitações, os mapas de predição apresentaram coerência com o conhecimento relativo aos atributos, nas condições locais.

**Palavras-chave:** Predição, regressão linear múltipla, regressão-krigagem, Cuesta de Botucatu.

## ABSTRACT

The soil map is an essential tool for land use planning and studies related to environmental aspects of this important natural resource. Quantitative techniques and geoprocessing tools are currently combined with the interpretation of pedogenic processes to enable the development of more accurate maps obtained by faster and less costly process. Among the models applied to it, the hybrid models employ predictive auxiliary variables and spatial autocorrelation, to enable the prediction of soil attributes in unsampled locations. The digital soil mapping worldwide project – GlobalSoilMap.net acts in order to provide global representations of soil attributes developed through the application of hybrid model in legacy soil data, performing the practice of Digital Soil Mapping (MDS). This work was based on the assumption that the application of the hybrid technique of regression-kriging (RK), using legacy data of soil survey and covariates of relief and remote sensing provide representative map of soil attributes of an area in Cuesta of Botucatu. The goal was to apply locally, in two depths, prediction models and continuous representation of Soil Redness Index (IAV), base saturation index (V%), sand content and clay content, cation-exchange capacity (CTC) and pH of the soils in Edgardia Experimental Farm, for which are available soil survey data. The predictor covariates were derived from an Digital Elevation Model (MDE) and an orbital image. They were all standardized at spatial resolution of 10 m, the methods were selected by checking significant linear correlation between attributes and covariates and spatial autocorrelation of attributes or residues of multiple linear regressions (RLM). The data were separated into training and validation subsets. The correlation coefficients ( $r$ ) between soil attributes and covariates were significant and ranged from -0.40 to 0.51. The predictors more correlated to attributes were topographic wetness index (ITU), slope (Decl), aspect (Aspc), elevation (Elev) and vegetation index (NDVI), and the last four are key definers of granulometric fractions. The values of adjusted  $R^2$  of RLM were between 0.10 and 0.36, which is considered low. In general, the prediction maps exhibited characteristic patterns of spatial variation observed in the covariates maps, used in the calibration of the models. An increase in accuracy was observed between the two steps of the modeling process by RK, indicating that the final map is better than the RLM. However, the models showed generally low performance, and did not provide good results when evaluated by external validation and even if the area was stratified

in two smaller plots, with more homogeneous relief. The results indicated the restricted use of soil survey sampling in prediction models, and the difficulty of applying MDS in areas with complex lithology, especially where the correlation between local dynamics of soil genesis and selected covariates are not strong. Despite the limitations, the prediction maps were consistent with knowledge about soil properties in local conditions.

**Keywords:** Prediction, multiple linear regression, regression-kriging, Cuesta of Botucatu.

## LISTA DE ILUSTRAÇÕES

<b>Figura 1</b>	– Representação esquemática da modelagem espacial de uma variável ambiental por regressão-krigagem.....	41
<b>Figura 2</b>	– Localização da Fazenda Experimental Edgárdia e distribuição de pontos de amostras de solos.....	52
<b>Figura 3</b>	– Formações geológicas na área da Fazenda Edgárdia.....	53
<b>Figura 4</b>	– Mapa de solos da Fazenda Experimental Edgárdia.....	55
<b>Figura 5</b>	– Modelo digital de terreno da Fazenda Experimental Edgárdia.....	76
<b>Figura 6</b>	– Declividade (Decl) e Fator LS (FLS) da Fazenda Experimental Edgárdia.....	76
<b>Figura 7</b>	– Curvatura vertical (CPE) e horizontal (CPL) da Fazenda Experimental Edgárdia.....	77
<b>Figura 8</b>	– Aspecto (Aspc) e Índice Topográfico de Umidade (ITU) da Fazenda Experimental Edgárdia.....	78
<b>Figura 9</b>	– NDVI e SAVI da Fazenda Experimental Edgárdia.....	78
<b>Figura 10</b>	– Variogramas dos resíduos da RLM do atributo IAV.....	85
<b>Figura 11</b>	– Mapas de predição de IAV por RLM, por RK e da qualidade da predição – DP, da área total da Fazenda Edgárdia .....	87
<b>Figura 12</b>	– Variogramas dos resíduos da RLM do atributo V%.....	92
<b>Figura 13</b>	– Mapas de predição de V% por RLM, por RK e da qualidade de predição – DP, da área total da Fazenda Edgárdia.....	94
<b>Figura 14</b>	– Variogramas dos resíduos da RLM do atributo Areia.....	99
<b>Figura 15</b>	– Mapas de predição da Areia por RLM, por RK e da qualidade de predição – DP, da área total da Fazenda Edgárdia.....	101
<b>Figura 16</b>	– Variogramas dos resíduos da RLM do atributo Argila.....	104
<b>Figura 17</b>	– Mapas de predição da Argila por RLM, por RK e da qualidade de predição – DP, da área total da Fazenda Edgárdia.....	106
<b>Figura 18</b>	– Variogramas do atributo CTC .....	109
<b>Figura 19</b>	– Mapas de predição da CTC por KO e da qualidade de predição – DP, da área total da Fazenda Edgárdia.....	111
<b>Figura 20</b>	– Mapas de predição do pH por IDW da área total da Fazenda Edgárdia.....	113
<b>Figura 21</b>	– Representação gráfica da comparação entre valores estimados e observados nos pontos de validação para os atributos IAV, V, e Ar do solo da Fazenda Edgárdia.....	115
<b>Figura 22</b>	– Representação gráfica da comparação entre valores estimados e observados nos pontos de validação para os atributos Ar, Arg, CTC e pH do solo da Fazenda Edgárdia.....	116

<b>Figura 23</b> – Mapas de atributos IAV e V% do solo na parte alta da Fazenda Edgárdia.....	125
<b>Figura 24</b> – Mapas de atributos Areia e Argila do solo na parte alta da Fazenda Edgárdia.....	126
<b>Figura 25</b> – Mapas de atributos IAV e V% do solo na parte baixa da Fazenda Edgárdia.....	127
<b>Figura 26</b> – Mapas de atributos Areia e Argila do solo na parte baixa da Fazenda Edgárdia.....	128

## LISTA DE TABELAS

<b>Tabela 1</b>	– Variáveis ambientais intervenientes na pedogênese e na distribuição de solos em região tropical.....	44
<b>Tabela 2</b>	– Unidades de mapeamento de solos da Fazenda Experimental Edgárdia.....	56
<b>Tabela 3</b>	– Características da imagem SPOT sensor HRG2.....	58
<b>Tabela 4</b>	– Regra de decisão para seleção da função de MDS.....	67
<b>Tabela 5</b>	– Estatística descritiva dos atributos de solo.....	73
<b>Tabela 6</b>	– Matriz de correlação entre as covariáveis ambientais.....	79
<b>Tabela 7</b>	– Correlações entre atributos do solo e covariáveis.....	80
<b>Tabela 8</b>	– Resultado da aplicação das regras de seleção do modelo de predição para cada atributo de solo para a área total.....	82
<b>Tabela 9</b>	– Equações RLM para IAV.....	84
<b>Tabela 10</b>	– Modelos e parâmetros dos variogramas isotrópicos ajustados aos resíduos do ajuste de regressão linear para o atributo IAV.....	86
<b>Tabela 11</b>	– Estatísticas dos dados amostrais e dos modelos ajustados para o atributo IAV para a área total da Fazenda Edgárdia.....	89
<b>Tabela 12</b>	– Equações da RLM para V%.....	91
<b>Tabela 13</b>	– Modelos e parâmetros dos variogramas isotrópicos ajustados aos resíduos do ajuste de regressão linear para a saturação por bases.....	92
<b>Tabela 14</b>	– Estatísticas dos dados amostrais e dos modelos ajustados para o atributo V% para a área total da Fazenda Edgárdia.....	96
<b>Tabela 15</b>	– Equações da RLM para Areia.....	97
<b>Tabela 16</b>	– Modelos e parâmetros dos variogramas isotrópicos ajustados aos resíduos do ajuste de regressão linear para o atributo areia...	99
<b>Tabela 17</b>	– Estatísticas dos dados amostrais e dos modelos ajustados para o atributo Areia para a área total da Fazenda Edgárdia.....	102
<b>Tabela 18</b>	– Equações da RLM para Argila.....	103
<b>Tabela 19</b>	– Modelos e parâmetros dos variogramas isotrópicos ajustados aos resíduos do ajuste de regressão linear para o atributo argila..	104
<b>Tabela 20</b>	– Estatísticas dos dados amostrais e dos modelos ajustados para o atributo Argila para a área total da Fazenda Edgárdia.....	108
<b>Tabela 21</b>	– Modelos e parâmetros dos variogramas isotrópicos ajustados aos dados do atributo CTC.....	109
<b>Tabela 22</b>	– Estatísticas dos dados amostrais e dos modelos ajustados para o atributo CTC para a área total da Fazenda Edgárdia.....	112

<b>Tabela 23</b> – Estatísticas dos dados amostrais e dos modelos ajustados para o atributo pH para a área total da Fazenda Edgárdia.....	114
<b>Tabela 24</b> – Modelos de RLM para atributos de solo nas áreas alta e baixa da Fazenda Edgárdia.....	121
<b>Tabela 25</b> – Modelos e parâmetros dos variogramas isotrópicos ajustados aos resíduos da RLM para os atributos de solo nas partes alta e baixa da Fazenda Edgárdia.....	122
<b>Tabela 26</b> – Estatística descritiva das amostras da área total, parte alta e parte baixa da Fazenda Experimental Edgárdia.....	123

## LISTA DE ABREVIATURAS E SIGLAS

ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
CTC	Capacidade de Troca de Cátions
DAEE	Departamento de Águas e Energia Elétrica
EPP	Efeito Pepita Puro
EUPS	Equação Universal de Perda de Solo
FAO	Organização das Nações Unidas para Alimentação e Agricultura
GPS	Sistema de Posicionamento Global por Satélites
IAV	Índice de Avermelhamento do Solo
IBGE	Instituto Brasileiro de Geografia e Estatística
IDE	Índice de Dependência Espacial
IGC	Instituto Geográfico e Cartográfico
IPT	Instituto de Pesquisas Tecnológicas
KO	Krigagem ordinária
MDE	Modelo Digital de Elevação
MDS	Modelagem Digital de Solos
NDVI	Índice de Vegetação da Diferença Normalizada
RLM	Regressão Linear Múltipla
RK	Regressão krigagem
SIG	Sistema de Informação Geográfica
SR	Sensoriamento Remoto
SRTM	Shuttle Radar Topography Mission
WRB	World Reference Base



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>21</b>
<b>2</b>	<b>REVISÃO DE LITERATURA.....</b>	<b>23</b>
2.1	Demanda de informações sobre o solo.....	23
2.2	A formação e os atributos dos solos.....	24
2.3	Variabilidade espacial dos solos.....	26
2.4	O mapeamento de solos.....	29
2.5	O mapeamento digital de solos .....	32
2.5.1	O processo e a adoção do MDS.....	34
2.5.2	Técnicas de MDS.....	36
2.5.3	A regressão-krigagem.....	39
2.5.4	As covariáveis ambientais.....	43
2.6	Aplicações da regressão-krigagem.....	48
<b>3</b>	<b>MATERIAL E MÉTODOS.....</b>	<b>51</b>
3.1	Descrição da área de estudo.....	51
3.2	Material.....	54
3.2.1	Seleção da base cartográfica.....	54
3.2.2	Mapa de solos.....	55
3.2.3	Imagens orbitais.....	58
3.2.4	Banco de dados de atributos de solo.....	58
3.2.5	Programas.....	59
3.3	Métodos.....	59
3.3.1	Pré-processamento de material cartográfico.....	59
3.3.2	Organização dos dados.....	60
3.3.3	Atributos de solo .....	61
3.3.4	Seleção das covariáveis ambientais preditoras.....	62
3.3.4.1	Obtenção do MDE e variáveis morfométricas.....	62
3.3.4.2	Índices de vegetação.....	65
3.3.5	Análise exploratória dos dados.....	66
3.3.6	Seleção da predição espacial dos atributos de solo .....	67
3.3.7	Aplicação do modelo híbrido.....	68
3.3.7.1	Avaliação das covariáveis e RLM.....	68
3.3.7.2	Krigagem ordinária dos resíduos e mapa de predição.....	69

3.3.8	Validação.....	71
3.3.9	Subdivisão da área.....	72
4	<b>RESULTADOS E DISCUSSÃO.....</b>	<b>73</b>
4.1	Análise exploratória dos dados.....	73
4.2	Covariáveis ambientais.....	75
4.3	Covariáveis e correlações.....	79
4.4	Modelos de predição dos atributos de solo.....	82
4.4.1	Índice de avermelhamento do solo (IAV).....	83
4.4.2	Índice de saturação por bases (V%).....	90
4.4.3	Areia.....	96
4.4.4	Argila.....	102
4.4.5	Capacidade de troca de cátions (CTC).....	108
4.4.6	pH do solo.....	112
4.5	Representação da validação dos modelos.....	114
4.6	Considerações sobre a modelagem dos atributos para a Fazenda Edgárdia.....	117
4.7	Modelagem considerando duas áreas geomorfologicamente distintas.....	119
4.8	Considerações finais.....	129
5	<b>CONCLUSÕES.....</b>	<b>130</b>
	<b>REFERÊNCIAS .....</b>	<b>131</b>

## 1 INTRODUÇÃO

O solo é um dos principais recursos naturais que proporcionam suporte à vida na Terra. Sua utilização provê alimentos, fibras e biomassa para produção de energia. Para o uso sustentável do solo é necessário que haja conhecimento suficiente de suas características e propriedades, um planejamento do uso e ocupação da terra e que sejam adotadas práticas de manejo adequadas.

A 68ª Assembleia Geral das Nações Unidas declarou o ano de 2015 como o Ano Internacional dos Solos e segundo a *Food and Agriculture Organization of the United Nations* (FAO), o destaque para o recurso natural visou aumentar a consciência e entender a importância do solo para a segurança alimentar e funções essenciais nos ecossistemas. Um dos objetivos de fóruns e discussões referiu-se ao desenvolvimento de políticas públicas e efetiva adoção de manejo sustentável.

Para alcançar esses objetivos, entretanto, um requisito básico é o conhecimento de informações suficientes e precisas sobre o solo. O levantamento de dados em uma área, com a finalidade de realizar o diagnóstico, deve providenciar informações confiáveis que possam inclusive ser usadas para a estimativa de valores dos atributos em locais não amostrados. Além disso, deve possibilitar a representação gráfica das informações, na forma de mapas, que sejam úteis ao planejamento agrícola ou ambiental.

As informações de solos geralmente não estão disponíveis em escalas adequadas, para todo o território brasileiro e para alcançar maior detalhamento, é necessário levantamento de campo, custoso em termos de recursos financeiros e tempo. O nível de detalhamento geralmente não é grande, ou ainda não há qualquer representação de solos para a área de interesse.

O mapa de solos é uma ferramenta essencial para planejamento de uso e estudos que envolvem aspectos ambientais. Tradicionalmente os mapas de solos são elaborados baseados na sua classificação e são delimitadas unidades de mapeamento. O estudo é realizado por especialistas, adotando o modelo solo-paisagem, que considera as condições de formação do solo e as relações entre os solos e a paisagem. Atualmente tais procedimentos têm sido aliados a uma modelagem mais quantitativa, notadamente para mapear os atributos de solos.

Na Pedometria, o mapeamento digital de solos (MDS) adota metodologias usando ferramentas de geoprocessamento e técnicas quantitativas para a confecção

de representações espaciais de atributos dos solos, de forma que sejam mais aplicáveis a objetivos práticos, com interpretação imediata. Várias questões como adequação de técnicas, escalas de trabalho e resolução, amostragem, definição de variáveis intervenientes no processo, adequação de modelos, bem como acurácia nos produtos finais são exemplos dos temas investigados nessa área. Dentre as técnicas, modelos híbridos, que empregam variáveis auxiliares preditoras e autocorrelação espacial, têm apresentado bons resultados.

A aplicação global de técnicas MDS está se tornando realidade no projeto *GlobalSoilMap.net*, consórcio que visa elaborar o mapa digital de solos do planeta, para o qual foi definido um protocolo que inclui o uso de dados e mapas de levantamentos já realizados, valorizando recursos financeiros e tempo empregados na sua elaboração. Algumas aplicações regionais também foram pesquisadas.

Com base nessas considerações, as hipóteses básicas do trabalho foram:

- a) A aplicação de um modelo geoestatístico híbrido multivariado, baseado em dados legados e em variáveis ambientais auxiliares de relevo e índices de sensoriamento remoto é eficiente na elaboração de mapas de atributos de solos para a área da Fazenda Experimental Edgárdia.
- b) O padrão espacial dos atributos de solos está correlacionado com o relevo nas condições de ocorrência da Cuesta de Botucatu.

A proposta deste trabalho consistiu na modelagem digital da predição de atributos de solos, com quantificação dos erros associados, que poderão ser utilizados como suporte no planejamento de levantamentos em áreas com relevo típico de cuesta. Para tanto, foi utilizada técnica geoestatística híbrida e ferramentas de geoprocessamento, com dados legados de levantamento de solos e covariáveis ambientais preditoras na Fazenda Experimental Edgárdia, Faculdade de Ciências Agrônômicas (FCA), da Universidade Estadual Paulista (Unesp), no município de Botucatu-SP.

Os objetivos específicos foram:

- a) elaborar um banco de dados digital de solos e características físicas da Fazenda Experimental Edgárdia;
- b) definir e mapear variáveis preditoras a serem utilizadas na modelagem digital de atributos do solo;
- c) aplicar modelo geoestatístico e híbrido para modelagem digital dos atributos de solos e avaliar a incerteza dos modelos de predição.

## 2 REVISÃO DE LITERATURA

### 2.1 Demanda de informações sobre o solo

Informações precisas sobre o solo são necessárias nas mais diferentes escalas, para diversas aplicações, desde o estabelecimento de políticas públicas globais e nacionais até a adoção de práticas pontuais na agricultura de precisão (BARTLEY et al., 2013; CHIG et al., 2010; DOBOS et al., 2006).

Segundo Resende et al. (2014), em condições naturais, as características e propriedades do solo resultam de inter-relações de fatores ecológicos que compõem o ambiente, atingindo um estado de equilíbrio dinâmico e fazendo desse recurso natural um componente de importância fundamental no ecossistema. O seu uso e manejo acarretam alterações em suas propriedades, principalmente, na camada superficial, o que afeta a sua qualidade, como avaliado por Bertol et al. (2004) e Marques et al. (2010). O uso do solo deve ser pautado, então, por um manejo sustentável, baseado em um modelo holístico, de acordo com o ambiente e com seus atributos físicos, químicos e biológicos, regido por uma governança responsável da terra (CARTA ..., 2015).

O conhecimento dos atributos do solo é requisito para planejamento e manejo sustentável de áreas agrícolas. Mendonça-Santos e Santos (2003) colocaram a execução de mapeamentos de solos como uma demanda permanente por parte de várias instituições, na busca por informações do meio físico para o planejamento da ocupação racional das terras e adequada gestão ambiental, seguindo as premissas do desenvolvimento sustentável. O mapeamento de solos permite inferir sobre o seu comportamento no espaço, constituindo um material básico e essencial para planejamentos e previsões. Alguns tipos de levantamento, em escalas maiores e nível mais detalhado, possibilitam interpretações muito práticas, como ocorre, por exemplo, nos relatórios de levantamentos de solos dos Estados Unidos da América do Norte (LEPSCH, 2011).

Na ciência são muitos os modelos de previsão de cenários ambientais que demandam informações de solos para realizar prognósticos. Gray et al. (2015) ressaltaram as modelagens da segurança global dos solos, segurança alimentar,

sequestro de carbono, mudanças climáticas, sistemas hidrológicos, ecológicos, além de outros sistemas. Em escalas regional e local, modelos de perda de solos (CUOMO et al., 2015) risco de degradação e contaminação (ZEWDU et al., 2015), de deslizamentos de terras (KIM et al., 2015; NANDI ; SHAKOOR, 2009), dentre outros.

Adhikari e Hartemink (2016) avaliaram as inter-relações entre solos e serviços ecossistêmicos e atestaram uma deficiência no reconhecimento da contribuição do solo para o bem-estar humano. Os autores recomendaram que ferramentas para avaliação do serviço ecossistêmico do solo sejam desenvolvidas, de forma interdisciplinar, e que estudos avaliem e monitorem o recurso solo em um contexto de serviços ecossistêmicos.

Além da informação a respeito dos valores dos atributos do solo em si, e sua espacialização, o padrão de variação ao longo da área constitui importante informação adicional. Sob o aspecto da classificação de solos, o padrão de variação pode ser referência para identificação de relações solo-paisagem, que auxiliam na predição de comportamento e no uso do solo como uma referência para estratificar o ambiente (RESENDE et al., 2014), principalmente para fins de uso e manejo. Essa prática é possível devido ao padrão previsível de atributos e classes de solos, resultado de uma síntese dos processos e fatores de formação do solo específicos para um determinado local. O conhecimento dos padrões da variação, a partir de amostras em locais específicos em campo se dá pela aplicação de metodologias estatísticas espaciais e geoestatísticas, que viabilizam a explicação, inferências e previsões de fenômenos em diversas aplicações (ISAAKS; SRIVASTAVA, 1989; SRIVASTAVA, 2013; YAMAMOTO; LANDIM, 2013;).

## **2.2 A formação e os atributos dos solos**

Desde os primeiros estudos realizados por Vasily Dokuchaev, na Rússia, identificando as características diferenciais que ocorriam nos solos e o caráter evolutivo implícito no processo, foi constatado que tais diferenças se desenvolviam ao longo do tempo e que permitiam separar solo e rocha. Segundo Florinsky (2012b), Dokychaev postulou que o solo é sempre, e em qualquer lugar, uma função da rocha de origem, do clima, da vegetação, da idade e da topografia do terreno. A pesquisa relacionada à pedogênese buscou uma definição da relação entre os

fatores, além do conhecimento da influência deles na formação dos solos (BOCKHEIM et al., 2014). Os fatores interagem entre si e mudam ao longo do tempo, interferindo de maneira combinada em todo o processo, propiciando condições que favorecem determinados processos em detrimento de outros, resultando em uma condição particular que origina um determinado tipo de solo e define suas características e propriedades.

Posteriormente, Jenny (1941) sintetizou a atuação dos fatores na pedogênese como variáveis, na equação que estabelece a relação entre os atributos do solo (ou ainda, o tipo de solo) aos fatores do estado do sistema solo, considerando apenas o tempo como variável totalmente independente e as demais independentes, com forte interdependência entre elas (Equação 1):

$$S \text{ (ou } s) = f(Cl, o, r, p, t \dots) \quad (1)$$

Onde  $S$  = tipo de solo;  $s$  = atributos do solo;  $Cl$  = clima;  $o$  = organismos;  $r$  = relevo;  $p$  = material de origem;  $t$  = tempo.

Segundo White (2009), as reticências indicam que ainda outros fatores podem afetar a pedogênese, mesmo que não sejam preponderantes no processo.

Muitos cientistas do solo buscaram estabelecer as relações quantitativas que possibilitassem definir o resultado final em termos de  $s$  ou  $S$ , considerando que a combinação dos fatores de formação dos solos é única para cada condição local. Os pedólogos inclusive simplificaram, considerando apenas um dos estados como variável enquanto os demais fossem praticamente constantes, dando origem a relações de litossequências, climossequências, cronossequências, biossequências e topossequências de solos (RESENDE et al., 2014). Dessas, as topossequências são as mais estudadas, como abordados por Campos et al. (2012) e Rodolfo Júnior et al. (2015).

A modelagem matemática de características e propriedades do solo considera as condições iniciais e aquelas atuantes durante a formação do solo. Assim, White (2009) previu que a variabilidade dos atributos físicos e químicos, ou mesmo de tipos de solos resultantes, pode ser grande, mesmo em áreas restritas, como na abrangência de pequenas bacias hidrográficas ou propriedades rurais. De acordo com Mendonça-Santos e Santos (2003), a previsão é possível uma vez que

condições idênticas de fatores e processos de formação de solos originam solos semelhantes.

Estudos de pedogênese têm um caráter de conhecimento básico, mas segundo Lepsch (2011), é fundamental para outras áreas da Ciência do Solo e para o desenvolvimento de tecnologias e práticas de manejo dos solos. Do ponto de vista do mapeamento, permitem identificar as variáveis atuantes e a definição das propriedades e classes dos solos. Uma situação de equilíbrio dinâmico solo-paisagem resulta de um processo de interações geomorfológicas e pedológicas (QUEIROZ NETO, 2000).

Por fim, o conhecimento dos processos e fatores de formação do solo são base de sistemas de classificação de solos (BOCKHEIM et al., 2014; SANTOS et al., 2013), pelo qual os solos são agrupados, em diferentes níveis, pelo seu processo de formação.

### **2.3 Variabilidade espacial dos solos**

A variabilidade espacial do solo deve ser considerada na previsão e representação dos solos, uma vez que o conhecimento da estrutura de variação possibilita um mapeamento mais eficaz. Nesse caso, a quantificação do erro que está inserido no mapeamento passa a ser também informação valiosa, apesar da maior complexidade na análise (KRAVCHENKO; BULLOCK, 1999).

Heuvelink e Webster (2001) consideraram que as variações espaciais e temporais das propriedades do solo resultam em diversidade de habitats e riqueza de flora e também resiliência às flutuações de clima. Além disso, podem ser citados vários exemplos de aplicações práticas do conhecimento da variação espacial dos solos, como sua influência nos fluxos de água e solutos através do solo, nos processos erosivos em função das propriedades dos solos e predição de atributos (HEUVELINK; WEBSTER, 2002).

A variabilidade do solo no contexto espaço-temporal foi considerada por Malone et al. (2011) como a causa da impossibilidade de serem estabelecidos modelos precisos e acurados para fenômenos relativos aos solos. A imprevisibilidade da variação do solo associada à simplificação dos modelos incorre na incerteza, que deve então ser avaliada e minimizada, ou pelo menos, conhecida.

A representação gráfica dos solos exige predição, devido à sua inerente

variabilidade. Hengl (2009) considerou difícil explicar a variação de variáveis ambientais, principalmente, porque observações de campo são muito custosas e amostragens para fins de mapeamentos normalmente representam uma porcentagem muito abaixo de 1% do total da área a ser mapeada.

Segundo Oliver e Webster (2014), o valor observado para o atributo a partir de uma amostra resulta de uma combinação complexa de processos físicos, químicos e biológicos que implica em uma variação aparentemente aleatória, que geralmente não pode ser quantificada por meio de uma solução determinística. Os autores apresentaram as bases da formalização do tratamento matemático de uma variável aleatória, conforme o seguinte:

Um atributo assume um valor  $z$ , em um dado local  $X$ , definido pelas coordenadas  $(X_1, Y_1)$ , como uma realização, naquele local, do processo que poderia resultar em infinitos valores possíveis da variável aleatória  $Z(x)$ .

O processo aleatório é o conjunto de todos os valores aleatórios possíveis de ocorrer em todos os locais de uma região, também denominado  $Z(x)$ .

A variável aleatória tem correlação espacial em alguma escala.

A variabilidade de uma variável ambiental é constituída de três componentes, conforme a Equação 2:

$$Z(x) = Z'(x) + \mathcal{E}'(x) + \mathcal{E}'' \quad (2)$$

Onde  $Z(x)$  é a variável ambiental,  $Z'(x)$  é a componente determinística,  $\mathcal{E}'(x)$  é a componente aleatória correlacionada espacialmente e  $\mathcal{E}''$  é a variação de micro escala ou o erro da medida.

A parte determinística é a superfície de tendência e as demais são a superfície de erros (HENGL, 2009).

As componentes  $Z'(x)$  e  $\mathcal{E}'(x)$  representam a variação espacial natural, espacialmente dependente, que pode ser explicada, até certo nível, por modelos matemáticos, e  $\mathcal{E}''$  é o ruído espacialmente independente, geralmente inerente às medições (McBRATNEY et al., 2000).

A maneira como será feita a predição da variável abrange técnicas que estabelecem a relação entre os atributos de solos e os fatores que influenciam a sua variabilidade (CARVALHO JÚNIOR et al., 2014b).

Na abordagem geoestatística, a inferência espacial se baseia na concepção de um processo aleatório que caracteriza o conjunto de observações, considerando a continuidade espacial da variável aleatória, pelo estimador do variograma, realizando o ajuste de um modelo teórico a um modelo experimental (OLIVER; WEBSTER, 2014).

Com base em uma amostragem representativa, consistente e aleatória, Yamamoto e Landim (2013) mostraram as possibilidades de interpolar ou estimar o valor de uma variável em pontos não amostrados, por meio de funções matemáticas locais (pontos mais próximos aos amostrados) ou globais (todos os pontos amostrais). Se houver correlação espacial da variável, métodos geoestatísticos serão adotados e o principal processo de interpolação de dados esparsos para geração de mapas é a krigagem, com as características de não tendenciosidade e variância mínima associada ao valor estimado (MATHERON, 1963).

Oliver e Webster (2014) destacaram o grande uso de métodos geoestatísticos e a adoção da krigagem como o principal método de interpolação em pesquisas em ciências ambientais; relacionados principalmente pela introdução de ferramentas geoestatísticas em Sistemas de Informações Geográficas (SIG) e a disponibilidade de pacotes estatísticos com essas funções. Os autores afirmaram que corriqueiramente o principal objetivo nas pesquisas é a obtenção de um mapa da distribuição espacial de uma variável, obtido por um processo que começa com amostragem adequada, segue com análise exploratória, representação dos dados, transformação (se necessário), a estimação e modelagem de um ou mais variogramas, seguida da interpolação e, finalmente, a representação do mapa.

O nível da dependência espacial pode ser baseado na relação entre os valores da componente estrutural e patamar (CAMBARDELLA et al., 1994), ou ainda nessa relação atrelada ao modelo ajustado (SEIDEL; OLIVEIRA, 2014). Oliver e Webster (2014) atentaram para a necessidade de conhecimento adicional do campo e dos processos físicos, geológicos e geomórficos, ao se realizar inferências a partir da modelagem do variograma.

Na presença de tendência dos dados, ou seja, se um atributo de solo pode ser relacionado a um fator causa, a aplicação da krigagem envolve o grupo de técnicas denominadas híbridas por McBratney et al. (2000). Para esses casos são utilizados os métodos krigagem universal, krigagem com deriva externa ou regressão-

krigagem (RK), e o discernimento entre os métodos se baseia em quais as informações auxiliares que são usadas na predição espacial e como a análise é conduzida (HENGL et al., 2004). A escala e a resolução espacial refletem diretamente na capacidade de se compreender a variabilidade dos solos, segundo Behrens et al. (2010).

## **2.4 O mapeamento de solos**

No portal de solos da FAO, o levantamento de solos é definido como “um estudo sistemático do solo em uma área, que inclui a classificação e o mapeamento das propriedades e a distribuição das várias unidades de solos”.

O mapeamento torna possível a interpretação dos mapas para aplicação em estudos de ecossistemas e para manejo das terras. A partir desses mapas deve ser possível realizar predições precisas e úteis, obtendo o máximo de informação sobre o solo, na ausência de informação específica local. Para que esse propósito seja atingido, é necessário: determinar o padrão de solos na área, dividir esse padrão em unidades relativamente homogêneas, levando em conta critérios de classificação e finalmente mapear a distribuição dessas unidades, em uma dada escala, seguindo normas e procedimentos definidos (Empresa Brasileira de Pesquisa Agropecuária - Embrapa, 1995).

No Brasil, segundo Coelho e Giasson (2010), os mapas de solos com maior abrangência em termos de território nacional foram produzidos no âmbito do projeto Radam e, posteriormente, Radambrasil, constituindo material adequado para planejamentos estratégicos e abrangentes, uma vez que o nível do levantamento foi exploratório, na escala 1:1.000.000. A publicação de tais mapas e relatórios não alcançou todo o território brasileiro (Instituto Brasileiro de Geografia e Estatística - IBGE, 2007).

Outra representação cartográfica disponível dos solos brasileiros foi compilada pela Embrapa, na escala 1: 5.000.000 em sucessivas edições, sendo a mais recente, no ano de 2011, e pelo nível de detalhe é igualmente adequada para estudos estratégicos. Com abrangência mais restrita, algumas áreas possuem mapeamentos de solos a diferentes níveis, em escalas maiores (MENDONÇA-SANTOS; SANTOS, 2006).

A definição dos solos nos levantamentos oficiais é feita por meio de técnicas

tradicionais de uma classificação formal, onde cada classe identifica solos com características peculiares (OLIVEIRA; CAMARGO, 1992; SANTOS et al., 2013) e que provavelmente têm comportamento semelhante no ambiente.

Arruda et al. (2013) e Ten Caten et al. (2012) declararam que as duas metodologias disponíveis atualmente para mapeamento de solos são complementares e devem ser desenvolvidas simultaneamente: a metodologia tradicional produzindo informações que são utilizadas para treinar modelos e prever classes e propriedades de solos e o MDS, com as técnicas automatizadas, que permite maior rendimento e precisão.

Na metodologia tradicional, a ocorrência das unidades de solo é baseada em fotografia aérea da área, considerando informações do ambiente e a experiência do pedólogo. A partir dessa inferência é elaborado um mapa prévio, no qual são definidos os locais para pesquisa de campo, em perfis que representem os pedons. É adotado um critério para agrupar pedons contíguos, e estabelecer o limite, delineando um polígono no mapa. A premissa básica é de que haja homogeneidade dentro do polígono e heterogeneidade entre polígonos, estabelecendo os limites abruptos entre as unidades (DOBOS; HENGL, 2009; EMBRAPA, 1995; RESENDE et al., 2014).

Hudson (1992) ressaltou a dificuldade de transmissão do conhecimento tácito do pedólogo, habilidade advinda da experiência, que possibilita a identificação dos corpos de solos organizados e mapeáveis aplicando o paradigma definido pela equação de formação do solo (Equação 1) e do modelo solo-paisagem.

O levantamento de solos é baseado em um sistema de classificação. Os sistemas taxonômicos ou naturais de classificação de solos, como a Soil Taxonomy (1999) e a Classificação Brasileira de Solos (SANTOS et al., 2013), têm por objetivo definir as classes de solos, pautando-se pelo reconhecimento de atributos diferenciais que são importantes para a interpretação permanente relativa ao solo, mas o sistema em si é isento de finalidade específica em relação ao solo classificado (LEPSCH, 2011). O produto final do levantamento, nesse caso, permite inferências relativas a uso e manejo do solo, desde que interpretado adequadamente. Minasny e McBratney (2016) ressaltam a dificuldade que pode ocorrer na interpretação do mapa e a legenda de solos resultantes por potenciais usuários.

Quanto ao uso efetivo dos sistemas de classificação de solos, Hartemink (2015)

verificou que mesmo os sistemas como *Soil Taxonomy, World Reference Base* (WRB) e FAO/Unesco são mais utilizados localmente, e mesmo em áreas da ciência do solo ou outras áreas mais fortemente relacionadas a solos, ainda são utilizados termos e denominações relacionados a fatores de formação ou atributos do solo (solos arenosos, solos aluviais, solos florestais, solos vermelhos...), em detrimento dos sistemas de classificação.

Zimback (2007) considerou que o método tradicional de representação cartográfica dos solos não corresponde à ocorrência das variáveis de solo no terreno, que apresenta uma variação espacial contínua, e não mudanças abruptas como nos limites dos polígonos do mapa. Além disso, a variação é grande dentro dos polígonos, em função de limites estabelecidos para as classes dos atributos diagnósticos, como ocorre, por exemplo, com a textura, que pode imprimir comportamento químico, físico e morfológico muito distinto para dois solos que possuam teores de argila de 16% e 34%, próximos aos limites inferior e superior da classe de textura média (a classe de textura média compreendia solos com teores de argila acima de 15% e abaixo de 35%).

Uma dificuldade apontada por Giasson et al. (2011), a respeito dos mapas de solos convencionais, é que tais representações não são suficientes para explicitar o modelo mental do pedólogo da relação solo-paisagem, que fundamenta o trabalho de mapeamento e que dá suporte ao subsequente planejamento de uso da terra.

Os mapas pedológicos oficiais representam as classes de solos para transmitir as informações referentes aos solos (IBGE, 2007; SANTOS et al., 2013), embora a representação dos atributos dos solos tenha sido apontada por pesquisadores como uma forma que permite a interpretação mais precisa e imediata da variável de interesse, o que torna o mapa mais prático e aplicável, especialmente em trabalhos que exigem maiores escalas, como por exemplo, na agricultura de precisão (TEN CATEN, 2011; ZIMBACK, 2001).

Em muitos casos, o objetivo é o mapeamento das propriedades do solo. Segundo Zhu et al. (2013), são necessários mapas de solos funcionais, detalhados e voltados às aplicações, que possam ser utilizados para manejo em locais específicos. Castrignanò et al. (2000) consideraram a caracterização quantitativa e espacial das condições do solo um pré-requisito para práticas agronômicas adequadas, o que proporciona rentabilidade e proteção ambiental em áreas agrícolas.

## 2.5 O mapeamento digital de solos

A pedometria é a área da ciência do solo que trata da aplicação de métodos quantitativos, tanto matemáticos quanto estatísticos, que permitem lidar de uma maneira mais objetiva com a incerteza, nos modelos de previsão de ocorrência de solos (McBRATNEY et al., 2000). Na adoção dos métodos pedométricos busca-se a conciliação dos modelos pedológicos, baseados na relação solo-paisagem e nos processos de formação do solo, com a quantificação da incerteza inerente em processos de modelagem para fenômenos naturais, o que foi proposto por uma abordagem determinístico-estocástica (HENGL, 2009).

Dobos e Hengl (2009) mostraram a evolução dos mapeamentos de solo ao longo do tempo, até os mapas digitais de solos. Os autores explicaram que um aspecto decisivo no surgimento e evolução da pedometria foi o desenvolvimento tecnológico, como por exemplo, os Sistemas de Informações Geográficas – SIG, bancos de dados, sensoriamento remoto, sensores em geral, uso de meio digital e outros. A tarefa de elaboração de mapas passou a utilizar desses recursos, segundo os autores, enfatizando a natureza contínua dos parâmetros dos solos, ao invés de delineamentos das classes em polígonos. Nesse processo de mapeamento de solos a característica marcante é o uso de modelos estatísticos quantitativos, que levam em conta a componente estocástica nos modelos de predição e a possibilidade de analisar a relação entre as variáveis no sistema.

McBratney et al. (2003) consideraram que os estudiosos da ciência do solo buscaram atender à necessidade crescente de informações e dados de solos, para monitoramento e modelagem ambiental, propiciando a elaboração de mapas de solos mais precisos e com a incerteza associada aos modelos. Tal busca se deu pela adoção de modelos preditivos quantitativos, apresentados e discutidos pelos autores.

Embora as bases conceituais e fundamentos dessa abordagem quantitativa não sejam novas (McBRATNEY et al., 2003), uma maior disseminação da pesquisa nessa área pode ser considerada recente, conforme levantamentos de Grunwald (2009) e Ten Caten et al. (2012), utilizando dessas técnicas para gerar o que pode ser considerado mais do que um simples mapa, e conforme Dobos e Hengl (2009), uma representação complexa do meio ambiente. Podem ser constatados resultados

que atendem de forma mais contundente a demanda por mapeamentos técnicos e interpretativos, mais aplicáveis sob o ponto de vista prático (ZHU et al., 2013). MacMillan (2008) atribuiu certa resistência inicial ao uso do MDS aos fatores como: pouca disponibilidade de banco de dados de entrada, preocupação com relação à qualidade dos mapas resultantes e capacidade técnica de pessoal.

McBratney et al. (2000; 2003) apresentaram uma síntese da metodologia geral do mapeamento digital de solos, relacionando solo e outros fatores espacialmente referenciados, com o intuito de aplicar uma função de predição. Foram considerados os mesmos fatores da Equação 1, mas também a correlação espacial das propriedades do solo e a presença de erros autocorrelacionados. O modelo, denominado função de predição espacial de solo com erros espacialmente autocorrelacionados - SCORPAN-SSPFe, foi definido pela Equação 3:

$$Sc \text{ ou } Sa = f(s, c, o, r, p, a, n) \quad (3)$$

Onde  $Sc$  é a classe de solo;  $Sa$  é o atributo de solo;  $s$  refere-se a informação de solo previamente conhecida em um determinado ponto, como por exemplo uma unidade de mapeamento de um mapa pré-existente;  $c$  traduz propriedades climáticas do ambiente no ponto definido;  $o$  representa organismos, vegetação ou ação antrópica;  $r$  é a topografia ou atributos da paisagem;  $p$  é o material de origem, representado pela litologia;  $a$  é o fator tempo e  $n$  a localização ou seja, a posição definida por coordenadas espaciais.

Sob essa perspectiva, segundo McBratney et al. (2000; 2003), são gerados mapas de atributos ou classes de solos, utilizando uma modelagem matemática, considerando os fatores de formação do solo e uma inferência, onde se utilizam informações do solo e ambiente para prever propriedades do solo. Nesse sentido, o mapeamento digital não se contrapõe ou menospreza o mapeamento tradicional de solo, uma vez que são aplicados seus pressupostos básicos e todo o conhecimento inerente, tornando o método mais quantitativo e passível de corroboração e, de acordo com Ten Caten et al. (2012), as abordagens são complementares e devem ser desenvolvidas simultaneamente. Essa adoção de um enfoque mais quantitativo ocorreu também em outras áreas da ciência, como por exemplo, na Geologia, onde Landim (2006) ressaltou que os métodos numéricos

têm o papel de descrever quantitativamente os eventos e não o conjunto de conhecimentos sobre esses eventos, aperfeiçoando a amostragem, análise, comprovação e possibilidade de previsões.

Brevik et al. (2016) entenderam que o modelo scorpan é essencialmente importante em áreas com carência de informações diretas sobre solo, permitindo predições espaciais superiores, pela utilização de covariáveis relacionadas aos fatores tradicionais e pela consideração da autocorrelação.

Na Equação 3, McBratney et al. (2003) realçaram as características de não causalidade, ou seja, classes de solo podem ser preditas com base em propriedades do solo ou o inverso e ainda, da capacidade do modelo de definir vetores probabilidade para a ocorrência de Sc ou Sa. A aplicação da equação estabelece a relação da observação de solo e as variáveis ambientais com significado pedológico, localizadas nas mesmas coordenadas espaciais constituindo os chamados dados de treinamento. Isso representa uma classificação supervisionada, ou um problema de aprendizagem supervisionada. As regras definidas para o conjunto de treinamento são então aplicadas em outros locais, onde somente as variáveis ambientais são conhecidas, predizendo a classe ou os atributos do solo, tornando esse método mais interessante para a representação espacial de atributos (TEN CATEN et al., 2011).

### **2.5.1 O processo e a adoção do MDS**

De acordo com MacMillan (2008), dependendo dos tipos de variáveis de entrada e de variáveis preditas e dos tipos de equações utilizadas, podem ocorrer algumas diferenças nos procedimentos do processo de inferência. Embora não haja protocolos estabelecidos, que atendam amplamente a aplicação do mapeamento digital de solos, é possível estabelecer uma sequência geral do processo, conforme as etapas citadas por Florinsky (2012a):

- a) Conceituação da classe de solo ou propriedade do solo a ser predita;
- b) Identificação e obtenção ou derivação das informações que serão usadas como preditoras (MDTs, dados de sensores remotos, mapas temáticos, dados de levantamentos amostrais em campo);
- c) Desenvolvimento de regras de classificação ou equações de predição usando modelos matemáticos;
- d) Aplicação das regras de classificação ou equações nos dados de entrada;

- e) Avaliação da acurácia ou eficiência de predição em termos de métodos de avaliação adequados;
- f) Produção de mapas finais.

A avaliação do mapa final é feita com base na comparação de valores estimados pelo modelo e valores levantados em campo e, se a acurácia não for considerada satisfatória, podem ser levantados mais dados, os preditores revisados ou as próprias equações revistas, levando sempre em conta os objetivos estabelecidos (MacMILLAN, 2008).

Lima et al. (2013) comentaram que o MDS permite a extrapolação das relações encontradas para um determinado conjunto de covariáveis e  $S_a$  ou  $S_c$  para outras áreas, que possuam condições ambientais semelhantes e para as quais não existam informações de solos. Além disso, novas informações ou alterações ocorridas em termos de conhecimento e tecnologia podem ser rapidamente inseridas na base de dados, permitindo gerar atualizações.

Hengl et al. (2007a) destacaram a importância da conciliação dos conceitos do mapeamento tradicional de solos às técnicas analíticas dos mapeadores digitais de solos e o trabalho conjunto de experientes pedólogos, implicando em metodologias robustas e operacionais. Segundo os autores, essa prática tornaria operacional a utilização de bases de dados existentes para aplicação de técnicas pedométricas, gerando mapas de solos com maior detalhamento, sem necessidade de refazer levantamentos de solos demorados e custosos.

Considerando o uso dessa metodologia no Brasil, Rosa (2012), atribuiu a adoção tardia do MDS pelos pedólogos no Brasil ao atraso na disponibilização e popularização de ferramentas como SIG, sensoriamento remoto e programas, dificuldade de acesso às tecnologias e trabalhos desenvolvidos no exterior nas décadas passadas, carência de pessoal qualificado e um caráter conservador dos pedólogos. O autor constatou ainda um foco maior em mapeamento de classes de solos em detrimento de representações de atributos de solo.

Com relação ao mapeamento de classes de solos, Ten Caten et al. (2012) constataram que o tema MDS se tornou frequente, e a sua divulgação deve ser crescente entre cientistas do solo. Foram discutidos os procedimentos metodológicos adotados nas pesquisas e como expectativas para o futuro, os autores levantaram questões como custos envolvidos nos projetos, padronização na avaliação da qualidade dos mapas gerados e divulgação dos mapas gerados em

formato que possa ser utilizado em um SIG. Ainda segundo os autores, a adoção conjunta de metodologias convencionais e de MDS foi apontada ainda pelos autores como uma solução para viabilizar o mapeamento de classes de solos no Brasil, com melhor aplicação de recursos e geração de mapas pedológicos que atendam a demanda da sociedade. O treinamento de recursos humanos para realização desses mapeamentos também foi considerado prioridade.

### 2.5.2 Técnicas de MDS

São várias as técnicas que podem ser usadas com a finalidade de ajustar as relações quantitativas entre classes ou atributos de solos e as variáveis relativas ao ambiente, ou covariáveis (McBRATNEY et al., 2003). Mendonça-Santos e Santos (2003) sintetizaram em três tipos os modelos adotados em mapeamento digital de solos: os empírico-determinísticos, baseados no tradicional modelo Clorpt (Equação 1), os empírico-estocásticos, baseados em métodos geoestatísticos e o terceiro, apresentado por McBratney et al. (2000), que é considerado o modelo híbrido.

Giasson et al. (2011) consideraram que a associação de técnicas tradicionais de levantamento de solos com as novas tecnologias de mapeamento digital melhorou a elaboração de mapas de solos, exatamente por tornar o processo mais quantitativo.

Na seleção do método, Mendonça-Santos e Santos (2003) ressaltaram a aplicação da técnica de acordo com os objetivos, a resolução espacial desejada e qualidade da informação requerida. Em McBratney et al. (2003) e Florinsky (2012a) são descritos os métodos matemáticos para mapeamento preditivo de propriedades e classes de solo e citados trabalhos com sua aplicação.

Os métodos empírico-determinísticos consistem no uso de modelos lineares, que incluem as regressões e classificações (McBRATNEY et al., 2003). São utilizadas as regressões quando se trata de prever atributos de solos e, as classificações, para estabelecer classes de solos. É aplicado um modelo de regressão linear quando se presume que uma reta descreve a relação entre o atributo de solo e duas ou mais variáveis, geralmente variáveis ambientais, relacionadas aos fatores da Equação 1. Nesse caso, são usadas equações de primeiro grau, mas segundo Milone e Angelini (1995), relações mais complexas podem exigir funções de ordem superior, como por exemplo, quadráticas, ou generalizando, modelos não lineares. Além disso, no caso de previsões de atributos

de solos, geralmente o modelo mais preciso não será uma regressão simples, mas, normalmente são consideradas mais variáveis independentes, implicando em um modelo de regressão linear múltipla, escrita na forma da Equação 4 (McBRATNEY et al., 2003).

$$Sa = Qb + e \quad (4)$$

Onde  $Sa$  = atributo de solo predito, ou o vetor resposta;  $Q$  é a matriz de variáveis preditoras, que explicam parte da variabilidade inerente ao  $Sa$  e podem ser também designadas por covariáveis;  $b$  é o vetor parâmetro da função linear e  $e$  representa os desvios do modelo em relação ao valor observado.

As variáveis preditoras são comumente variáveis contínuas, o ajuste do modelo é feito pelo método dos quadrados mínimos e o grau de associação, expresso pelo coeficiente de determinação ( $R^2$ ), avaliado por um teste estatístico de significância, como o teste F, por exemplo.

Com relação às variáveis preditoras, pode ser utilizada apenas uma, ou possíveis combinações de mais variáveis, sendo avaliada a correlação entre elas, possibilitando o uso, no modelo, daquelas que explicam maior parte da variabilidade da variável predita, usando um método de seleção de variáveis, conforme exposto por Hair et al. (2009).

O método stepwise é um procedimento automático, aplicado para uma avaliação de cada variável individualmente, considerando a sua inclusão e exclusão do modelo, baseado em uma regra de decisão, medido por um teste estatístico, como o teste F (Fisher). Se a variável não contribui para explicar a variável dependente, é considerada como redundante ou não-significativa e descartada (HAIR et al., 2009).

Modelos lineares generalizados e modelos aditivos generalizados são aplicados quando é necessário acomodar uma situação de não linearidade no modelo e, no primeiro caso, é pressuposto que a variável resposta tenha uma distribuição pertencente à família exponencial, segundo McBratney et al. (2003).

Quando ocorrem interações complexas entre atributos, técnicas de inteligência artificial, como árvores de decisão e redes neurais são adotadas (BEHRENS; SCHOLTEN, 2006; SARMENTO et al., 2012; CHAGAS et al., 2011).

Os sistemas de mapeamento fuzzy aplicam escala de transição entre as classes

de solo preditas, separando elementos pertencentes a um determinado conjunto sem estabelecer um limite abrupto entre eles. São definidas regras ou operadores que representam a natureza contínua da variação espacial dos solos em função da variação do ambiente (ZHU et al., 2010) e gerados mapas conciliando o conhecimento qualitativo e quantitativo.

A adoção de métodos estocásticos se refere à aplicação das bases da geoestatística nos estudos de MDS. A elaboração de mapas de atributos de solos para uma extensão territorial resulta da interpolação dos valores conhecidos, em localizações específicas, predizendo os valores para locais de interesse, por estimativa ou simulação. As técnicas adotadas são as várias formas de krigagem, para pixels ou blocos, com estimativa da incerteza na predição, e a co-krigagem, onde variáveis secundárias densamente amostradas dão suporte à estimativa de variáveis amostradas de forma mais esparsa (ODEH et al, 1995). Segundo Omuto e Vargas (2015), os métodos geoestatísticos são amplamente aplicados e com disponibilidade crescente em programas de SIG e de forma especial, a combinação de krigagem com análise de regressão (método híbrido).

Com relação à aplicação das técnicas na modelagem de solos, Malone et al. (2011) atentaram para a imprevisibilidade do solo à simplificação das representações de processos complexos de solos, o que inevitavelmente incorre em erros nas saídas dos modelos. Para os autores é importante determinar qual a extensão da propagação da incerteza no modelo e qual é o efeito nas predições finais de um processo do mundo real. A incerteza final do modelo é a síntese de três principais fontes: a incerteza da estrutura do modelo, incerteza dos parâmetros do modelo e incerteza das entradas do modelo. Essa informação acrescenta credibilidade e segurança no uso do mapa final. Brevik et al. (2016) afirmaram que, além da consistência, o dimensionamento da incerteza é uma característica altamente desejável nos mapas de solos.

Um aspecto a ser ressaltado em relação ao uso de modelos no estabelecimento das relações entre atributos de solo e covariáveis é quanto à complexidade matemática ou estatística, que pode incorrer em distanciamento da interpretação pedológica e ambiental, limitando aplicações práticas (GRUNWALD, 2009).

Minasny e Hartemink (2011) alertaram para a consideração do princípio da parcimônia na geração de modelos, que elege como melhor modelo o que se ajusta ao conjunto de dados com o menor número de parâmetros. Segundo Gray et al.

(2015), técnicas e produtos mais simples servem como uma introdução para produtos mais avançados de MDS, promovendo sua maior aceitação na comunidade de ciência do solo.

Samuel-Rosa et al. (2015) apresentaram três considerações que constituem pontos fracos nas modelagens estatísticas de distribuição de solos: primeiro, a amostragem, que provê os dados pontuais de solos e tem que ser suficiente e com distribuição espacial adequada para a modelagem; segundo, o caráter empírico das relações entre condições do ambiente e propriedades do solo, menos interpretativa das relações solo-paisagem e por último, que as covariáveis utilizadas são somente aproximações das condições ambientais que realmente atuaram na formação dos solos e podem inclusive ser diferentes daquelas. Apesar dessas considerações, os autores afirmaram que o MDS tem permitido, com sucesso, a produção de mapas que levam em conta o padrão de variação espacial dos solos.

De acordo com Ließ (2015), não é possível saber se um mapa, elaborado a partir de um conjunto de amostras coletadas em uma área, representa bem os fatores de formação do solo e sua interação, que são determinantes da variabilidade dos atributos do solo. Segundo os autores, a preocupação com o plano de amostragem adotado é que seja representativo da área, ou seja, se é capaz de capturar a variabilidade dos atributos em estudo.

Ainda com relação à amostragem, Goovaerts (1999) salientou que geralmente são necessários muitos pares de dados para ajuste adequado dos variogramas e avaliação da anisotropia. Segundo Hengl (2009), mapas de qualidade podem ser produzidos a partir de um número considerado reduzido de amostras, se houver forte correlação entre a variável em estudo e as covariáveis preditoras.

Várias questões como adequação de técnicas, escalas de trabalho e resolução, adequação de amostragem, seleção de covariáveis intervenientes no processo e acurácia nos produtos finais são exemplos de questionamentos ainda recorrentes no foco do MDS (DOBOS; HENGL, 2009), bem como a necessidade de automatização e inserção de ferramentas estatísticas contemplando as técnicas de MDS nos SIGs (CASTRIGNANÒ, 2011; HENGL et al., 2007b).

### **2.5.3 A regressão-krigagem**

No âmbito do MDS, o processo de predição espacial leva em consideração a

estimação da variável de interesse em pontos onde essa informação não existe, sendo conhecidas as coordenadas espaciais e um conjunto de variáveis auxiliares (TEN CATEN et al., 2011)

O método híbrido RK é um método de estimação espacial, que possibilita prever atributos do solo associando a modelagem de componentes determinísticos e estocásticos (Equação 2), realizadas separadamente (HENGL et al., 2007b). A componente determinística consiste das tendências de grande-escala e a componente estocástica representa autocorrelação em curta-escala (CASTRIGNANÒ et al., 2000; OMUTO; VARGAS, 2015). As tendências de grande-escala são modeladas por análise de regressão entre a variável de interesse e as covariáveis ambientais e geralmente é múltipla, uma vez que são consideradas mais de uma variável secundária, ou o modelo pode ser não-linear. As tendências de curta-escala são modeladas por análise de krigagem (OMUTO; VARGAS, 2015).

A análise de regressão é utilizada para estabelecer o relacionamento entre duas ou mais variáveis, fornecendo as funções que relacionam as variáveis (MILONE; ANGELINI, 1995). No primeiro caso, ainda segundo esses autores, a relação causal entre a variável dependente e a variável independente é estabelecida por um modelo de regressão simples, do tipo  $Y = f(X)$ , permitindo a definição dos coeficientes “a” e “b” da equação de uma reta, se a relação for linear. No caso de duas ou mais variáveis independentes serem consideradas, são definidos os coeficientes “a” e coeficientes para todos os termos  $X_1, X_2, \dots X_n$ .

Hengl et al. (2003) classificaram as técnicas híbridas krigagem universal, krigagem com deriva externa e regressão-krigagem, aplicadas em casos de não estacionariedade dos dados, enfatizando que há certa confusão entre pesquisadores acerca das diferenças computacionais entre elas.

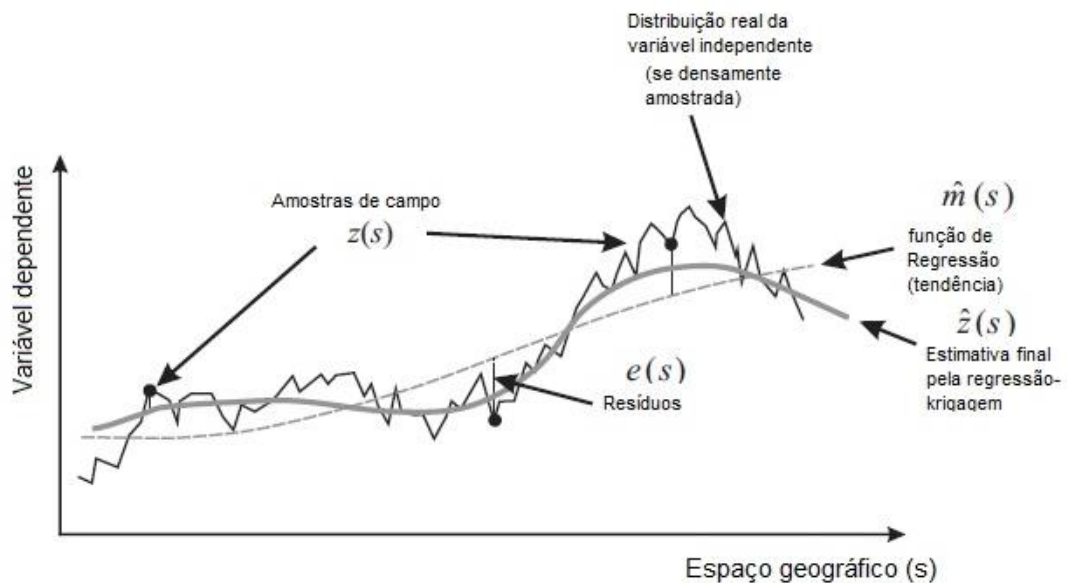
O discernimento é baseado, segundo o autor, nas propriedades dos dados de entrada. No caso da krigagem universal, definida por Matheron, em 1971, a tendência é modelada apenas em função da localização, pautando-se pelos valores das coordenadas ao realizar a regressão. Quando a tendência for definida com base em informações auxiliares, amplamente disponíveis na abrangência do estudo e correlacionadas com a variável de interesse, a técnica é a krigagem com deriva externa ou com tendência externa. Por fim, na regressão-krigagem, a tendência e resíduos são modelados separadamente e somados posteriormente. Conforme Hengl et al. (2004), as técnicas são matematicamente equivalentes, com aplicação

crecente em estudos relacionados a atributos de solos.

Hengl et al. (2007b) concluíram que a regressão-krigagem é um método flexível de mapeamento, que apresenta vantagens conceituais em relação a métodos alternativos, baseado no modelo universal de variação espacial da variável, apresentado na Equação (2). Os autores observaram que, apesar das propriedades atrativas da regressão-krigagem, essa ainda não é uma técnica amplamente usada em geociências, como seria esperado.

Na Figura 1, apresentada por Hengl (2009), são representados os modelos da variável pelas componentes determinística e estocástica, a tendência e os resíduos separadamente, além da estimativa final da variável  $\hat{Z}(s)$ , onde  $s$  representa a posição no espaço, dada por coordenadas conhecidas.

**Figura 1 – Representação esquemática da modelagem espacial de uma variável ambiental por regressão- krigagem**



Fonte: Hengl (2009)

Os dois componentes da predição no modelo da regressão-krigagem representam médias ponderadas. Considerando o modelo geral de regressão linear múltipla (Equação 4) para descrever a tendência dos dados, tem-se, segundo Hengl (2009), na Equação 5:

$$\hat{Z}(x_0) = \sum_{k=0}^p \beta_k \cdot q_k(x_0); \quad q_0(x_0) = 1 \quad (5)$$

Onde  $\hat{Z}(x_0)$  é o valor predito da variável de interesse,  $\beta_k$  os coeficientes da regressão estimados,  $p$  é o número de covariáveis e  $q_k$  os valores das covariáveis no local de interesse.

Na geoestatística, um valor da variável regionalizada pode ser estimado por interpolação baseada na krigagem, dada pela Equação 6 (OLIVER; WEBSTER, 2014).

$$\hat{Z}(x) = \sum_{i=1}^N \lambda_i \cdot Z(x_i) \quad (6)$$

Onde  $\hat{Z}(x)$  é o valor estimado,  $N$  é número de vizinhos medidos,  $\lambda_i$  são os pesos aplicados aos valores de  $Z(x_i)$ . A estimativa é considerada não tendenciosa, com média zero e variância mínima.

Então, considerando o modelo geral da regressão-krigagem, tem-se a igualdade da Equação 7 (HENGL et al., 2007b):

$$\hat{Z}(x_0) = m(x_0) + \varepsilon(x_0) = \sum_{k=0}^p \beta_k \cdot q_k(x_0) + \sum_{i=1}^n \lambda_i \cdot e(x_i) \quad (7)$$

Onde  $m(x_0)$  é a tendência modelada e  $\varepsilon(x_0)$  os resíduos interpolados.  $\beta_k$  são os coeficientes do modelo de tendência,  $\lambda_i$  são os pesos da krigagem determinados pela estrutura de dependência espacial dos resíduos e  $e(x_i)$  é o resíduo no local  $x_i$ .

Inicialmente, os coeficientes de regressão são obtidos a partir dos valores amostrais e informações das covariáveis, por um método de ajuste, como o de quadrados mínimos. A posterior estimativa dos resíduos pode dar-se por método direto, mais utilizado na prática, ou por processo iterativo. Nesse último, depois de obtida a regressão, são definidos os resíduos e realizada a sua modelagem da continuidade espacial. Adotando um método de quadrados mínimos generalizados, novos coeficientes de regressão são ajustados e calculados novos resíduos, podendo ser feito novo ajuste de continuidade espacial, sucedendo novo ciclo de cálculo. No entanto, a primeira iteração pode ser única, podendo ser considerada satisfatória, como apresentado por Hengl et al. (2007b).

Odeh et al. (1995) apresentaram três procedimentos distintos, aos quais

denominaram tipos de regressão-krigagem. No tipo A, aplica-se a regressão e depois é feita a krigagem ordinária (KO) dos valores da regressão; no tipo B, é realizada a krigagem ordinária tanto da regressão quanto dos resíduos e ambos são somados para obtenção da estimativa final e no tipo C, é realizada a krigagem ordinária apenas dos resíduos, somando as duas componentes. Nesse último, as covariáveis ambientais devem ser densamente amostradas e apresentarem correlação com a variável independente.

O modelo de regressão-krigagem foi adotado para mapeamento de variáveis ambientais (CARVALHO JÚNIOR et al., 2014a; GOOVAERTS, 1999; HENGL et al., 2004; 2007a, b), sendo considerado superior a outros métodos (ODEH et al., 1995).

Em função da adoção crescente dessa técnica de MDS, novas propostas de contribuir para aumento da acurácia na sua utilização têm sido apresentadas, como redução do número de pixels vizinhos considerados no processo da interpolação (SUN et al., 2012), inserção de efeitos de agrupamentos na parte determinística da modelagem (OMUTO; VARGAS, 2015).

#### **2.5.4 As covariáveis ambientais**

Aplicando-se as técnicas de MDS e baseadas na equação de Jenny, em 1941, funções matemáticas estabelecem a relação entre classes ou atributos e os fatores de formação do solo, representados por covariáveis ambientais (McBRATNEY et al., 2003). Algumas covariáveis ambientais usadas em MDS são apresentadas na Tabela 1.

Budiman et al. (2003) registraram como preditores dados de análises de solos, descrições de campo e morfologia do solo, variáveis ambientais, espectro eletromagnético do solo e base de dados de solos.

As covariáveis, nos estudos de MDS, são oriundas de material cartográfico, produtos de sensoriamento remoto, de sistemas de posicionamento por satélites, espectralradiometria e outras inovações tecnológicas, selecionadas em função do poder preditor, das inter-relações e efeitos de multicolinearidade entre elas (TEN CATEN, 2011). Du et al. (2015) associaram a seleção das covariáveis ambientais para mapeamento de atributos de solo à natureza da área de estudo e propriedade do solo a ser predita.

**Tabela 1 – Variáveis ambientais intervenientes na pedogênese e na distribuição de solos em região tropical**

Variáveis ambientais	Significância e interpretação	Potencial correlação ambiental
<p><i>Clima</i></p> Temperatura Umidade (precipitação) Radiação solar	Natureza e taxa de intemperismo e atividade biológica	Superfícies climáticas Atributos do terreno
<p><i>Organismos</i></p> Vegetação (índices de sensoriamento remoto) Antropogênese (distúrbios)	Ciclagem de nutrientes e padrões locais de vegetação  Tipo de vegetação e fogo	Sensoriamento remoto multiespectral; índices de vegetação  Sem preditores efetivos
<p><i>Material de origem</i></p> Unidades litológicas	Produtos do intemperismo, mineralogia e tamanho de partículas	Mapa geológico e sensoriamento remoto
<p><i>Relevo</i></p> Elevação  Declividade  Índice topográfico de umidade	Condições climáticas locais, vegetação, energia potencial Taxa de intemperismo, teor de água no solo, erosão/deposição, velocidade escoamento superficial Zonas de concentração de água na paisagem, regime de umidade do solo	Atributos do terreno e sensoriamento remoto geofísico

**Fonte:** Adaptado de McKenzie; Ryan (1999) e Chagas et al (2011)

De acordo com Moore et al. (1993), a superfície pode ser parametrizada por meio de atributos morfológicos como elevação, aspecto, declividade, curvatura planar, perfil de curvatura e fluxo acumulado, obtidos de um modelo digital de elevação – MDE. Hengl e MacMillan (2009), em um trabalho de revisão sobre modelos preditivos de solos, registraram que cerca de 80% das aplicações automatizadas de MDS são baseadas no uso do MDE.

Em escala local ou regional, a distribuição dos solos e vegetação natural pode ser explicada principalmente pelo fator relevo ( $r$ ). Condições topográficas podem alterar regimes climáticos locais e, ainda, definir a intensidade de processos como escoamento superficial, infiltração, erosão e deposição, o que altera várias propriedades do solo (DOBOS; HENGL, 2009).

Para fins de modelagem, o relevo é representado por um modelo numérico, como os modelos digitais de terreno (MDE), representando as cotas ou elevações, que dão origem às variáveis morfométricas. Na geração do MDE podem ser aplicados diferentes métodos e fontes de dados, implicando em incerteza inserida no modelo, o que afetará os índices topográficos derivados (HEBELER; PURVES, 2009; FLORENZANO, 2008).

Chagas et al. (2010) ressaltaram o aumento significativo da demanda por MDE em estudos ambientais no Brasil e a importância da avaliação das limitações em relação a essa representação como fonte de informação espacial. Os autores compararam MDE gerados a partir de curvas de nível de cartas topográficas, que consideraram como fonte mais comum desse tipo de informação e dados oriundos de sensores remotos (*Shuttle Radar Topography Mission* - SRTM e *Advanced Spaceborne Thermal Emission and Reflection Radiometer* - ASTER) e concluíram que o MDE obtido da carta topográfica foi superior aos demais. Um MDE de melhor qualidade foi obtido por Pinheiro et al. (2012) a partir de uma carta topográfica, acrescida de pontos cotados extraídos do modelo de terreno SRTM. A avaliação foi feita por comparação de feições originais e derivadas, utilizando métodos qualitativos.

A resolução do MDE também influencia nos resultados das pesquisas. Na geração da grade de um MDE, Hutchinson (1996) considerou que a escolha da resolução deva ser suficiente para representar o detalhamento dos dados originais. Nessa seleção deve ainda ser levado em conta o objetivo da aplicação, a precisão dos dados e o tempo para sua geração. Cordeiro et al. (2013) ressaltaram que em aplicações que implicam em quantificação de áreas, o MDE de maior resolução pode resultar em maior erro.

Considerando as ferramentas e métodos disponíveis para geração do MDE, Wilson e Gallant (2000) citaram os grids regulares, interpolação em redes de triângulos irregulares (TIN) ou a subdivisão de elementos gerados a partir das curvas de nível e suas ortogonais, esta última mais utilizada em aplicações

hidrológicas. Para Hutchinson (1989), na elaboração de um MDE consistente sob aspecto da hidrologia, gerado a partir de algoritmo baseado em grid regular, são impostas, no cálculo, restrições que asseguram conectividade na rede de drenagem e correto posicionamento de talvegues e divisores de água provenientes de dados de curvas de nível. Em trabalhos de mapeamento de solos, Carvalho Júnior et al. (2011) observou que dentre os parâmetros considerados na modelagem, tal modelo de elevação é muitas vezes considerado a mais importante, possivelmente por descrever com maior fidedignidade os fluxos de água sobre e dentro do solo.

Segundo Valeriano (2008), as variáveis do relevo compreendem atributos primários e secundários do relevo, sendo que os atributos primários são os derivados diretamente (derivada primeira) do MDE, ou seja, declividade, curvatura, curvaturas horizontal e vertical, área de captação ou acumulação de fluxo, etc. Os atributos secundários são derivados de combinações dos atributos primários, representando processos que ocorrem na área de abrangência e interferem no fenômeno estudado. Tais atributos são amplamente utilizados em modelagem de classes e atributos de solos (McBRATNEY et al., 2003), sendo determinantes no comportamento físico-hídrico do solo, de acordo com Florinsky (2012a). Podem ser a única covariável considerada (ODEH et al., 1995) na modelagem.

O relevo está diretamente relacionado à geomorfologia da área. Hengl e MacMillan (2009) indicaram a geomorfologia como um dos principais fatores que controlam processos naturais, como por exemplo, o desenvolvimento e funcionamento dos solos em determinado local pela sua atuação nos gradientes de energia, umidade e nutrientes. Assim, são potentes preditores da distribuição espacial contínua dos atributos de solos.

O fator material de origem define vários atributos do solo e é desejável que seja considerado no MDS. A principal dificuldade para se utilizar a litologia como covariável em um projeto de MDS consiste na disponibilidade dessa informação, uma vez que mapas geológicos são escassos e geralmente elaborados em escalas pequenas, menores que 1:500.000 conforme o Instituto de Pesquisas Tecnológicas – IPT (1981). A inclusão de informação material de origem foi considerada por alguns autores como o fator que seria decisivo para obtenção de melhores resultados (ROSA, 2012; TEN CATEN et al., 2011).

O fator organismos é representado no MDS por covariáveis que exprimem a atuação de fauna e flora e a interferência antrópica na formação do solo. Podem ser

originárias de diversas fontes, incluindo os produtos de sensoriamento remoto, como apresentado por Grunwald (2009). A autora apontou o uso desses produtos como preditores em trabalhos de MDS destacando a capacidade superior dos sensores de prover informação ambiental em abrangência e resolução.

Para estudos relacionados aos inventários e mapeamentos de solos, Jensen (2009) ressaltou a aplicação mais efetiva do sensoriamento remoto quando não há presença de vegetação e o uso de diferentes regiões espectrais. Embora se tenha vantagens no uso dessa tecnologia, ela ainda não é usada como rotina em levantamentos oficiais de solos, constituindo uma área de pesquisa a ser explorada, em todos os seus aspectos (GOMEZ et al., 2012) e considerações a respeito da escala de trabalho e disponibilidade das imagens também devem ser consideradas (MOREIRA, 2007). Segundo Dewitte et al. (2012), imagens de sensoriamento remoto devem ser associadas a dados auxiliares e outras fontes de informação, como observações de campo e análises de laboratório.

Os índices calculados a partir das bandas das imagens exprimem características da vegetação, como o Índice da vegetação pela diferença normalizada (NDVI) (ROUSE et al., 1973) e o Índice de vegetação ajustado ao solo (SAVI) (HUETE, 1988); ou inferências de características e propriedades do solo (BREVIK et al., 2016). Para aplicações em solos, esses preferencialmente devem estar descobertos (MULDER et al., 2011).

Por fim, nos modelos de predição, o fator clima é geralmente representado pelas variáveis: temperatura média anual e precipitação média anual e algumas vezes pela evapotranspiração (McBRATNEY et al., 2003). Embora os processos de formação de solos tenham se dado em época pretérita, conforme salientou Samuel-Rosa et al. (2015), e as condições atuais possivelmente sejam distintas daquelas de quando os solos se formaram, as séries históricas dessas informações (ou dados de sensoriamento remoto) podem melhorar a acurácia dos modelos. Para estudos locais, em função da uniformidade das condições climáticas e generalização dos mapas disponíveis, esse fator tem sido utilizado com menor frequência, ou mesmo desconsiderado nos modelos (TEN CATEN et al., 2012).

## 2.6 Aplicações da regressão-krigagem

Menezes et al. (2016) compararam a aplicação de RK à KO para predição de propriedades físico-hídricas de solos em duas bacias hidrográficas com distintas características solo-paisagem, nas quais diferentes fatores de formação e usos do solo controlam a distribuição das propriedades dos solos. Onde foi observada alta correlação linear entre covariáveis ambientais e atributo do solo, a RK apresentou melhor desempenho, mas em casos de relações mais complexas entre ambiente e atributos de solo, o modelo de regressão linear não se mostrou satisfatório.

A predição de outras variáveis, além das relativas a solos, também foi pesquisada pelo modelo de RK. Mello et al. (2015) avaliaram a aplicação de métodos geoestatísticos na predição do fator erosividade da chuva (R), com base em dados de pluviômetros distribuídos no território brasileiro. Os autores adotaram a altitude como variável secundária e caracterizaram a RK como o método mais indicado para estimação do fator R no Brasil.

Em um estudo com abrangência regional, Carvalho Júnior et al. (2014a) analisaram as relações entre atributos de solo e covariáveis ambientais, em um ambiente de relevo acidentado e clima tropical, no estado do Rio de Janeiro. Os autores avaliaram a predição dos atributos pH, argila, silte, areia e carbono orgânico, a diferentes profundidades, conforme protocolo do projeto *GlobalSoilMap.net*, descrito por Sanchez et al. (2009). As covariáveis utilizadas foram derivadas do MDE e imagens Landsat. Os métodos de MDS adotados foram regressão linear, krigagem ordinária, regressão-krigagem e árvores de regressão. Os autores consideraram que a performance da predição dos atributos de solo foi baixa, e sugeriram que amostragem mais densa e uso de um MDE mais detalhado, proporcionando escala mais detalhada de estudo poderiam melhorar a predição.

Na França, Vaysse e Lagacherie (2014) buscaram atender a demanda no contexto do projeto *GlobalSoilMap.net*, gerando mapas de atributos de solos usando dados legados disponíveis. Os autores aplicaram modelos de mapeamento digital de solos, a partir de mapas legados de solos e a partir de medições realizadas em perfis de solos de levantamentos anteriores, aplicando técnicas diferentes para os dois casos. Modelos pedométricos utilizando regressão e regressão-krigagem foram aplicadas nos dados legados de atributos dos perfis de solos. A krigagem dos

resíduos dos atributos de solos para os perfis permitiu capturar a variabilidade espacial que não havia sido modelada pelo modelo determinístico. As performances das predições apresentaram variação para as propriedades de solos e entre os modelos, com  $R^2$  variando de zero a 0,79. Os autores identificaram as covariáveis predominantes nas previsões das propriedades de solo, como por exemplo, para o pH, o material de origem (mineralogia e índice de textura), clima (temperatura mínima) e relevo (elevação) e para o conteúdo de argila, a geologia (índice de textura). Por fim, a base de dados foi considerada insuficiente para representar as variabilidades de curta-escala, o que representou uma deficiência dos dados para aplicação de MDS.

Avaliando modelos preditivos para elaboração de mapas de propriedades dos solos, Carvalho Júnior et al. (2014b) aplicaram métodos estatísticos e geoestatísticos para estabelecer relações entre as propriedades e os fatores que influenciam sua variabilidade na paisagem. Foram utilizadas covariáveis ambientais contínuas e categóricas na predição. As covariáveis ambientais consistiram de atributos do terreno, relação entre bandas de imagem Landsat 5, e quando o modelo adotado foi a regressão linear múltipla, foram acrescentadas as informações advindas do mapa de uso e cobertura do solo e mapa de solos, verificando se esse acréscimo melhorou a predição. O melhor desempenho foi atribuído ao modelo de predição que explicou 43% da variação da areia na camada de 60-100 cm, usando método de co-krigagem. No geral, foi observado de baixo a médio desempenho, com o coeficiente de determinação em média igual a 0,21. Em função de uniformidade da litologia, as propriedades variaram pouco, indicando que a média poderia representar de forma satisfatória 11 das 30 propriedades estudadas. Os autores julgaram que para regiões montanhosas em clima tropical, é importante avaliar as relações entre propriedades de solos e as covariáveis, uma vez que a densidade amostral, a qualidade das covariáveis e o modelo aplicado são decisivos nos resultados.

Umali et al. (2012) utilizaram métodos regressão, krigagem ordinária e regressão-krigagem, para estabelecer as relações entre propriedades de solos e variáveis ambientais, em um estudo local, de escala grande, realizado dentro de um pomar de maçãs. A área com predomínio de uma única classe de solo permitiu identificar variabilidade nas propriedades em função do manejo adotado para a cultura. A correlação entre parâmetros de terreno e propriedades do solo foi baixa, e a krigagem ordinária apresentou melhores resultados na estimação geoestatística

quando comparada com a regressão-krigagem.

Com o objetivo de elaborar mapas de atributos de solo de acordo com as especificações do *GlobalSoilMap.net*, Ciampalini et al. (2012) utilizaram dados legados de 89 perfis de solos do norte da Tunísia, para testar três funções mais utilizadas em MDS, quais sejam: regressão linear, krigagem ordinária e regressão-krigagem. Foram selecionadas covariáveis ambientais derivadas de MDE e de imagens de satélite. Para 28% das propriedades de solo não foi possível estabelecer correlação com as covariáveis ambientais e não se identificou estrutura espacial do solo. A variabilidade dos solos foi considerada pequena e a variação a curta escala foi atribuída à litologia. Os autores consideraram difícil conciliar dados esparsos com covariáveis ambientais disponíveis em escala global, mas reconheceram o mérito de se utilizar os modelos como indicativo dos problemas que podem ser encontrados na prática do projeto *GlobalSoilMap.net* e justificar maiores investimentos para o mapeamento.

Chagas et al. (2011) avaliaram o MDS, com base em variáveis ambientais com maior detalhamento, no mapeamento de classes de solos na região norte do estado de Minas Gerais. As variáveis do terreno foram derivadas de um MDE com resolução de 5m, mapa de litologia e bandas de uma imagem de alta resolução. O mapa gerado utilizando todas as covariáveis resultou em um desempenho superior, com destaque para a declividade como covariável preditora. Os autores consideraram a metodologia suficiente, mais rápida e mais barata para atender a necessidade de obtenção de mapas de solos, especialmente a escalas maiores que 1:25.000, uma vez que o uso de fotografias aéreas apresenta custo mais elevado.

Hengl et al. (2007a) utilizaram dados de perfis com abrangência nacional para testar a aplicação de métodos de interpolação semiautomáticos, com preditores auxiliares na elaboração de mapas de classes de solos. Os preditores consistiram de parâmetros do relevo, imagens de sensoriamento remoto e mapas de regiões fisiográficas. Foram testadas técnicas de classificação e regressão para geração de mapas preditivos de textura e classes de solos. Os autores consideraram que a melhor predição foi obtida pelo método da regressão-krigagem.

### 3 MATERIAL E MÉTODOS

#### 3.1 Descrição da área de estudo

A área selecionada para o estudo abrange a Fazenda Experimental Edgárdia, pertencente à Faculdade de Ciências Agrônômicas, da Universidade Estadual Paulista, localizada no município de Botucatu, região centro-oeste do estado de São Paulo.

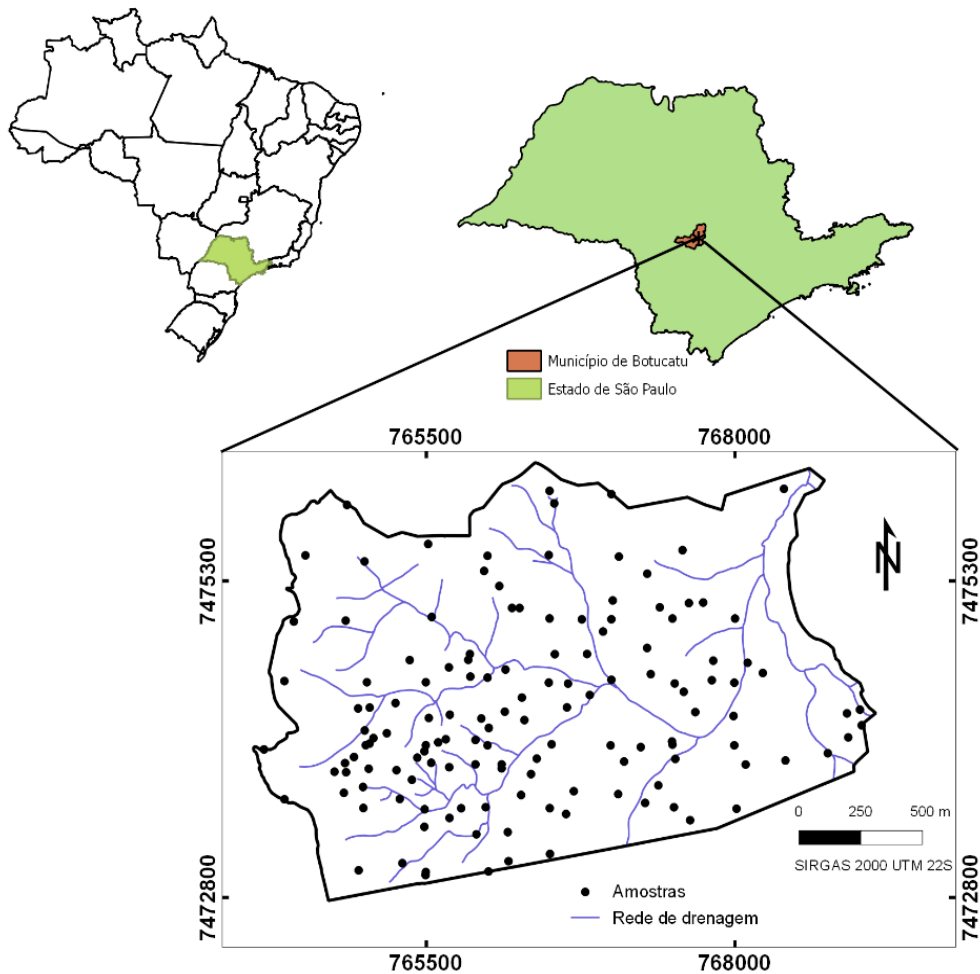
Na Microrregião de Botucatu, o processo de desenvolvimento foi atrelado à exploração agrosilvipastoril, prevalecendo atividade econômica agropecuária (SÃO PAULO, 2010), com importante destaque nas últimas décadas para a intensiva ocupação da terra voltada à produção de matéria prima para biocombustíveis (cultivo da cana) e silvicultura, além de áreas agrícolas e de pastagens.

A Fazenda Experimental Edgárdia é uma unidade de ensino e pesquisa, inserida na Área de Proteção Ambiental (APA) de Botucatu, com remanescentes preservados e em recuperação de vegetação natural, das tipologias Floresta Estadual Semidecidual e cerradão (JORGE; SARTORI, 2002) e áreas com ocupação de agricultura e pastagem.

A fazenda possui uma área de 12,17 km<sup>2</sup>, de acordo com o levantamento para georreferenciamento da propriedade, e sua localização é definida pelas coordenadas geográficas 22°50' a 22°47'30"S e 48°22'30" a 48°26'15"WGr (Figura 2).

A altimetria, segundo Carvalho et al. (1991), varia de mínima em torno de 475m e máxima em torno de 725m.

**Figura 2 – Localização da Fazenda Experimental Edgárdia e distribuição de pontos de amostras de solos.**

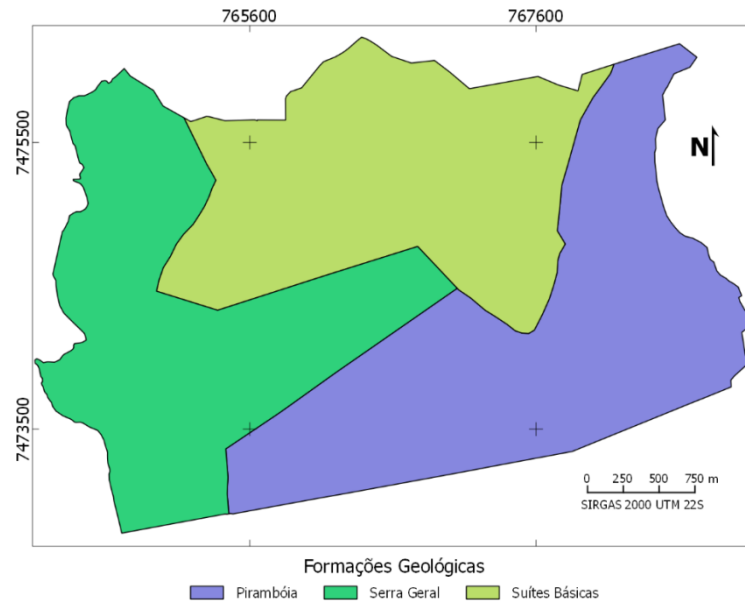


Na região predomina o clima do tipo Cfa, do sistema de classificação climática de Köppen (CUNHA; MARTINS, 2009) com características: temperado quente mesotérmico úmido, com temperatura média anual de 20,3°C, média do mês mais frio de 17,1°C e no mês mais quente acima de 22°C. As precipitações médias nos meses mais chuvoso e mais seco são, de acordo com a normal climatológica, de aproximadamente 200 mm e 35 mm, respectivamente. O índice pluviométrico anual é superior a 1.428,4 mm.

Os domínios geológicos na região remontam ao Mesozóico pertencentes ao Grupo São Bento, correspondendo ao Triássico Inferior, ao Jurássico, incluindo as Formações Pirambóia, Botucatu e Serra Geral e sedimentos aluviais do Holoceno (CARVALHO et al., 1991; IPT, 1981). No local de estudo foram mapeadas as formações Pirambóia, Serra Geral e Suítes Básicas, quando foi realizado o

levantamento geológico no âmbito do convênio DAEE-Unesp, sob a coordenação do prof. Paulo Milton Barbosa Landim e publicado por Landim (1984) (Figura 3).

**Figura 3 – Formações geológicas na área da Fazenda Experimental Edgárdia**



**Fonte:** Adaptado de Landim (1984)

As Formações Pirambóia e Botucatu ocorrem nessa região do estado de São Paulo e são constituídas de arenitos eólico-fluviais e eólicos, respectivamente (IPT, 1981), de coloração avermelhada e granulação fina a média, que ocorrem predominando nos vales, e de maneira generalizada na Depressão Periférica.

O basalto e o diabásio são as principais rochas eruptivas básicas, de coloração escura, textura afanítica e estrutura amigdaloidal presentes na Formação Serra Geral e que podem ser observadas nas partes mais altas da área, como descrito por Carvalho et al. (1991). Esse derramamento basáltico se estabeleceu na forma de lentes intercaladas aos arenitos no front da cuesta da Serra de Botucatu e em função da declividade, resultaram em solos mais rasos.

O limite oeste da área da Fazenda Edgárdia encontra-se no início do reverso da cuesta, abrangendo uma estreita faixa, onde são atingidas as maiores altitudes na área. Nas áreas mais próximas aos rios, principalmente nas margens do Rio Capivara, ocorreram deposições de sedimentos aluviais inconsolidados do Holoceno, e nas bordas da Serra também ocorreram depósitos de materiais sedimentares coluviais recentes provenientes das rochas básicas intemperizadas

(CARVALHO et al., 1991).

Ross e Moroz (1997) delimitaram grandes compartimentos no mapa geomorfológico do Estado de São Paulo e descreveram as unidades morfoestruturais da Bacia Sedimentar do Paraná, que compreende as unidades morfoesculturais Planalto Ocidental Paulista e Depressão Periférica Paulista. Uma das subdivisões menores, com denominação regional, é o Planalto Residual de Botucatu, que faz divisa com a Depressão do Médio Tietê pelo front da cuesta. A Fazenda Edgárdia está localizada nessa área de transição.

O relevo da área de estudo é caracterizado, portanto, por três regiões fisiográficas: o reverso da Cuesta Basáltica com altitudes acima de 700 m, que compreende uma faixa estreita, da parte alta da fazenda; a escarpa arenítico-basáltica, formando o front da cuesta, onde as declividades atingem valores acima de 45% (CARVALHO et al., 1991); e a Depressão Periférica, onde se localiza a parte baixa da área, com altitudes de 400 a 600 m e relevo compreendido de suave a suave-ondulado (CARVALHO et al., 1991; JORGE; MOREIRA, 2000).

As classes de solos presentes na área de estudo são diversificadas, com características intimamente relacionadas ao ambiente, incluindo solos rasos e afloramentos de rochas presentes na Cuesta Basáltica; solos profundos, com e sem horizonte B textural em relevo suave-ondulado e ondulado, na Depressão Periférica; e solos característicos de deposição e gleização na várzea do Rio Capivara (ZIMBACK, 2008; JORGE; SARTORI, 2002). Sob o ponto de vista do uso para agricultura, os solos são geralmente férteis, mas apresentam grande fragilidade quanto à erosão, devido principalmente à declividade.

## **3.2 Material**

### **3.2.1 Seleção da base cartográfica**

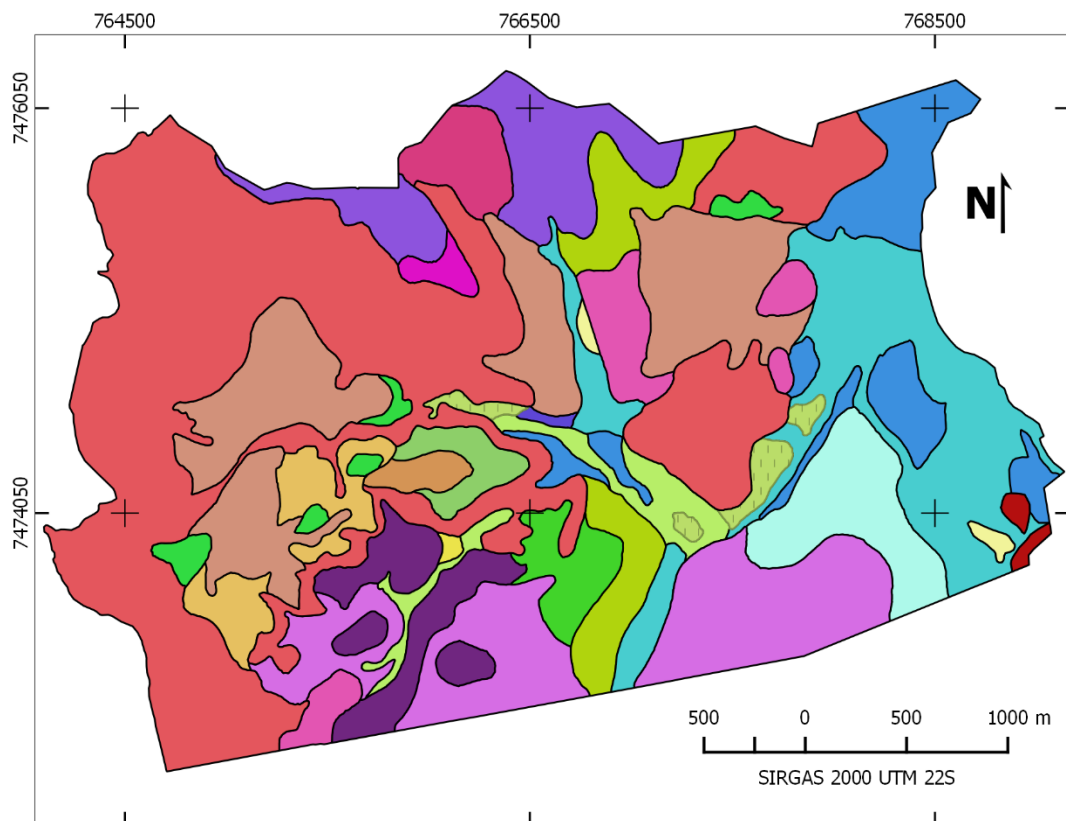
Foi realizado levantamento do material cartográfico que contém a área da Fazenda Edgárdia e, de acordo com o objetivo do estudo, foi adotada como base a carta planialtimétrica do Instituto Geográfico e Cartográfico do Estado de São Paulo (IGC), representada na escala 1:10.000, com equidistância de 5m entre as curvas de nível. A área de estudo está inserida na folha SF-22-Z-B-VI-3-NO-D, denominada Fazenda Lajeado.

### 3.2.2 Mapa de solos

Na escolha da área foi considerada a disponibilidade de dados e de relatório do levantamento semidetalhado de solos, para a aplicação da modelagem digital. As unidades de mapeamento estão representadas na Figura 4 (CARVALHO et al., 1991).

A nomenclatura das classes de solos definidas nesse levantamento foi atualizada, de acordo com a Classificação Brasileira de Solos (SANTOS et al., 2013), sendo adotada uma correspondência entre as classes de solo do sistema atual e o adotado no levantamento, as quais estão descritas na Tabela 2.

**Figura 4 – Mapa de solos da Fazenda Experimental Edgárdia**



**Unidades de mapeamento da Fazenda Experimental Edgárdia**

GXd	LVA	Lvef2	NVe1	PVAe1	PVAe4	RQd2	RYe2	RYe5
GXe1	LVdf	MX1	NVe2	PVAe2	RLe	RQgd	RYe3	
GXe2	Lvef1	MX2	NXe1	PVAe3	RQd1	RYe1	RYe4	

**Fonte:** Adaptado de Carvalho et al. (1991).

**Tabela 2 – Unidades de mapeamento de solos da Fazenda Experimental Edgárdia**

Legenda	Descrição
LVA	LATOSSOLO VERMELHO AMARELO Distrófico A moderado textura média fase floresta tropical subperenifólia relevo suave ondulado
LVdf	Associação de LATOSSOLO VERMELHO Distroférico textura média fase florestatropical subperenifólia/cerrado + LATOSSOLO VERMELHO Eutrófico + NITOSSOLO HÁPLICO Eutrófico Ta abrupto ambos textura arenosa/média fase floresta tropical subperenifólia/cerrado todos A moderado relevo ondulado
LVe1	LATOSSOLO VERMELHO Eutrófico A moderado textura argilosa fase floresta tropical subperenifólia/cerrado relevo suave ondulado
LVe2	LATOSSOLO VERMELHO Eutrófico A moderado textura média fase floresta tropical subperenifólia/cerrado relevo ondulado
NVe1	NITOSSOLO VERMELHO Eutrófico Tb A moderado textura média/argilosa fase floresta tropical subperenifólia relevo suave ondulado
NVe2	NITOSSOLO VERMELHO Eutrófico Ta A moderado textura argilosa fase floresta tropical subperenifólia relevo ondulado
NXe1	NITOSSOLO HÁPLICO Eutrófico Tb A chernozêmico textura média/argilosa fase floresta tropical subperenifólia relevo ondulado
PVAe1	ARGISSOLO VERMELHO-AMARELO Eutrófico Tb abrupto textura arenosa/argilosa fase floresta tropical subperenifólia relevo ondulado
PVAe2	ARGISSOLO VERMELHO-AMARELO Eutrófico abrupto textura arenosa/média fase floresta tropical subperenifólia relevo ondulado
PVAe3	Associação de ARGISSOLO VERMELHO-AMARELO Eutrófico textura arenosa/argilosa + ARGISSOLO VERMELHO-AMARELO Eutrófico textura arenosa/média ambos Tb abrupto fase floresta tropical subperenifólia relevo suave ondulado + LATOSSOLO VERMELHO Distroférico textura média fase floresta tropical subperenifólia/cerrado relevo ondulado + NITOSSOLO HÁPLICO Eutrófico Tb abrupto textura arenosa/média fase floresta tropical subperenifólia relevo suave ondulado todos A moderado
PVAe4	Associação de ARGISSOLO VERMELHO-AMARELO Eutrófico Tb abrupto textura arenosa/argilosa A moderado relevo ondulado + NEOSSOLOS LITÓLICOS Eutróficos Ta a chernozêmico textura média fase relevo forte ondulado substrato basalto e arenito ambos fase floresta tropical subperenifólia + NEOSSOLOS QUARTZARÊNICOS Distróficos A moderado fase floresta tropical subperenifólia relevo suave ondulado + NEOSSOLOS QUARTZARÊNICOS Eutróficos Ta textura arenosa/média ambos A moderado fase campo tropical higrófilo de várzea relevo plano.
MX1	CHERNOSSOLO HÁPLICO textura média/argilosa fase floresta tropical subperenifólia relevo ondulado

(continua...)

---

MX2	Associação de CHERNOSSOLO HÁPLICO textura média/argilosa fase pedregosa relevo ondulado + NEOSSOLO LITÓLICO Eutrófico Ta A chernozêmico textura média relevo forte ondulado substrato basalto e arenito + NITOSSOLO VERMELHO Tb A relevo ondulado fase floresta tropical subperenifólia + NITOSSOLO HÁPLICO Eutrófico Tb A chernozêmico textura média/argilosa fase floresta tropical subperenifólia relevo suave ondulado
GXd	GLEISSOLO HÁPLICO Distrófico Ta textura argilosa/média/arenosa fase campo tropical hidrófilo de várzea relevo plano
GXe1	GLEISSOLO HÁPLICO Ta Eutrófico textura argilosa/média fase campo tropical hidrófilo de várzea relevo plano
GXe2	Associação de GLEISSOLO HÁPLICO Eutrófico + GLEISSOLO Distrófico ambos Ta textura média + GLEISSOLO HÁPLICO Distrófico Tb textura argilosa + GLEISSOLO HÁPLICO Eutrófico Ta textura argilosa todos fase campo tropical hidrófilo de várzea relevo plano
RQd1	NEOSSOLO QUARTZARÊNICO Distrófico A moderado fase floresta tropical subperenifólia relevo suave ondulado
RQd2	Associação de NEOSSOLOS QUARTZARÊNICOS Distróficos A moderado fase floresta tropical subperenifólia relevo ondulado + NEOSSOLOS LITÓLICOS Eutróficos Ta A moderado textura arenosa fase floresta tropical subperenifólia relevo forte ondulado substrato basalto e arenito
RQgd	NEOSSOLO QUARTZARÊNICO HIDROMÓRFICO Distrófico A moderado fase campo tropical hidrófilo de várzea relevo plano
RYe1	NEOSSOLO FLÚVICO Eutrófico Ta A chernozêmico textura média fase campo tropical higrófilo de várzea relevo plano
RYe2	NEOSSOLO FLÚVICO Eutrófico Ta A moderado textura média/arenosa/média fase campo tropical higrófilo de várzea relevo plano
RYe3	NEOSSOLO FLÚVICO Eutrófico Ta típico textura arenosa/média fase campo tropical higrófilo de várzea relevo plano
RYe4	Associação de NEOSSOLO FLÚVICO A chernozêmico textura média + NEOSSOLO FLÚVICO A moderado textura média/arenosa/média ambos fase campo tropical higrófilo de várzea + GLEISSOLO HÁPLICO A moderado textura argilosa/média fase campo tropical hidrófilo de várzea + GLEISSOLO HÁPLICO textura argilosa/ média fase campo tropical hidrófilo de várzea todos Eutróficos Ta relevo plano
RYe5	Associação de NEOSSOLO FLÚVICO textura média + NEOSSOLOS FLÚVICOS textura média/arenosa/média ambos Eutróficos Ta fase campo tropical hidrófilo de várzea + NEOSSOLOS QUARTZARÊNICOS HIDROMÓRFICOS Distrófico fase campo tropical tropical hidrófilo de várzea todos A moderado relevo plano
RLe	NEOSSOLO LITÓLICO Eutrófico Ta chernozêmico textura média fase floresta tropical subperenifólia relevo forte ondulado substrato basalto e arenito

---

**Fonte:** Adaptado de Carvalho et al. (1991)

### 3.2.3 Imagens orbitais

Para gerar os dados auxiliares relativos aos índices de vegetação para a área foram utilizadas bandas da imagem do sistema orbital SPOT sensor HRG2, considerando a compatibilidade da resolução espacial com as demais informações utilizadas. HANSEN et al. (2009) observaram que dos produtos considerados de média resolução espacial, imagens semelhantes foram adequadas para aplicação em âmbito local.

A imagem possui 4 bandas espectrais (Tabela 3) e coordenadas do centro da cena na latitude 23°3'35"S e longitude 48° 17'47" WGr. Data da passagem 27/06/2010 e horário 13:12:26,5.

**Tabela 3 – Características da imagem SPOT sensor HRG2**

Espectro eletromagnético	Pixel (m)	Bandas espectrais (µm)
Pancromática	2,5 ou 5	0,48 - 0,71
B1: verde	10	0,50 - 0,59
B2: vermelho	10	0,61 - 0,68
B3: infravermelho próximo	10	0,78 - 0,89
B4: infravermelho ondas curtas	20	1,58 - 1,75

**Fonte:** <<http://www.geo-airbusds.com/en/194-resolution-and-spectral-bands>>

### 3.2.4 Banco de dados de atributos de solo

Os dados referentes aos atributos físicos e químicos dos solos foram obtidos do relatório do levantamento semidetalhado de solos de Carvalho et al. (1991), aqui denominado "Banco de dados 1" e do banco de dados da bacia hidrográfica do Rio Capivara, que contém a Fazenda Edgárdia, elaborado por Jorge (2009) e Jorge et al. (2014), aqui denominado "Banco de dados 2". As análises físicas e químicas foram realizadas no Laboratório do SNLCS/Embrapa-RJ para as amostras de perfis completos do primeiro levantamento e nos Laboratórios de Solos do Departamento

de Solos e Recursos Ambientais da FCA/Unesp, ambos seguindo metodologia indicada pela Embrapa (1979).

A malha amostral do levantamento semidetalhado de solos (Banco de dados 1) consistiu de 95 pontos, distribuídos irregularmente em campo, de acordo com a representatividade para fins de levantamento, e levando em conta ainda a acessibilidade, uma vez que a área apresenta uma complexidade física, com locais de difícil acesso. Foram descritos e amostrados 25 perfis em trincheiras e cortes de estrada e coleta de amostras efetuada com trado, realizadas em 70 pontos. A localização dos pontos foi elaborada em um mapa de malha quadrática, de 1 cm de lado, baseando-se nas linhas de referência de coordenadas geográficas que cortam a área da fazenda.

A malha amostral do “Banco de dados 2” consistiu de uma grade regular, com 49 pontos, espaçados de 500 m e georreferenciados diretamente no campo, utilizando receptor GPS de navegação.

O estudo foi realizado utilizando dados em duas profundidades, predominando amostras nas profundidades de 0 a 20 cm e 40-60 cm, aqui denominadas camada superficial e subsuperficial, respectivamente.

### **3.2.5. Programas**

ArcGIS (ESRI, 2012): vetorização de dados e geração do MDE.

GRASS GIS v. 6.4 (GRASS DEVELOPMENT TEAM, 2015): processamento de imagens orbitais.

SAGA-GIS v. 2.2.2. (SAGA GIS DEVELOPMENT TEAM, 2014): Geração das covariáveis ambientais relativas ao relevo, derivadas do MDE; processamento dos dados e modelagem da predição.

QGIS v. 2.8.3 (QGIS DEVELOPMENT TEAM, 2015): processamento de dados e finalização de mapas.

## **3.3 Métodos**

### **3.3.1 Pré-processamento de material cartográfico**

A carta base, obtida em formato digital, foi georreferenciada e reprojetada no Sistema Universal Transverso de Mercator, adotando como referência para o trabalho o Datum SIRGAS 2000 e sistema de coordenadas UTM Zona 22S. Essa base digital foi utilizada para vetorização das informações curvas de nível, hidrografia e pontos cotados da área.

O mapa representativo do levantamento semidetalhado de solos da Fazenda Experimental Edgárdia elaborado por Carvalho et al. (1991), na escala 1:10.000, foi digitalizado, georreferenciado e a partir dele foram vetorizadas todas as unidades de mapeamento de solos delimitadas na área. As denominações das classes de solos foram atualizadas de acordo com Santos et al. (2013).

Nos pontos de localização dos perfis e tradagens onde foram coletadas as amostras de solos, registrados no mapa do levantamento, foram resgatados os valores de coordenadas espaciais diretamente no SIG para o banco de dados.

O limite da área da fazenda foi obtido por meio de levantamento de campo, realizado para fins de georreferenciamento do imóvel rural, rastreando sinais do sistema GPS com receptor L1/L2 e pós-processamento dos dados, conferindo precisão centimétrica ao posicionamento.

### **3.3.2 Organização dos dados**

Os dados secundários foram digitados e organizados em planilhas e uniformizados em termos de unidades. Foram selecionados os pontos amostrais que apresentavam resultados para todas as variáveis, sendo eliminados aqueles casos com dados faltantes.

O conjunto total de dados composto por 141 amostras foi separado aleatoriamente em dois subconjuntos, um com 114 amostras, para treinamento e 27 pontos foram reservados para a validação dos modelos. A aleatoriedade na seleção dos pontos foi baseada no uso de ferramenta automática de seleção de pontos no SIG QGIS.

A amostragem de solos foi realizada com o objetivo do levantamento de solos por Carvalho et al., em 1991, e, portanto, não isenta de subjetividade. A densidade amostral variou no espaço geográfico, ressaltando que na área há locais de difícil acesso.

### 3.3.3 Atributos de solo

Os atributos de solo analisados foram selecionados dos bancos de dados, levando-se em conta os critérios de disponibilidade do dado e aplicabilidade da informação para fins de manejo. Foram considerados: Índice de avermelhamento do solo - IAV, pH, capacidade de troca de cátions - CTC (mmolc.dm<sup>-3</sup>), saturação por bases - V (%), teor de areia - Ar (g.kg<sup>-1</sup>) e teor de argila - Arg (g.kg<sup>-1</sup>).

A cor do solo reflete a presença de componentes como óxidos de ferro e matéria orgânica, sendo uma característica usada para diferenciação das classes de solos em sistemas de classificação (SANTOS et al., 2013). Cores avermelhadas, amareladas e brunadas ocorrem pela presença de óxidos e os solos acinzentados se apresentam em ambientes hidromórficos, em situação química de redução, de onde os óxidos de ferro foram transformados e/ou removidos (LEPSCH, 2011).

Na descrição morfométrica do perfil, o atributo cor do solo é determinado utilizando-se a carta de Munsell para solos. A leitura da cor é feita na amostra úmida, identificando os valores de Matiz, Valor e Chroma. A partir desses componentes, registrados no levantamento de solos, foi calculado o índice de avermelhamento - IAV, de acordo com Torrent e Barrón (1993), que quantifica a presença de hematita no solo pela Equação 8:

$$IAV = \frac{(10-M) C}{V} \quad (8)$$

Onde IAV é o Índice de Avermelhamento do solo, M é o índice de acordo com matiz, V= valor e C= chroma. Os valores assumidos para M correspondem a índices correspondentes da tabela de cores de Munsell: 7.5R= 0; 10R=1; 2.5YR=2; 5YR=3; 7.5YR=4; 10YR=5; 25YR=6; 2.5Y=7.

Para os pontos do “Banco de dados 2”, foram assumidos os valores de IAV característicos (modais) da classe de solo envolvente, ou dos pontos mais próximos.

### 3.3.4 Seleção das covariáveis ambientais preditoras

O uso de variáveis auxiliares, ou covariáveis ambientais, no processo de modelagem dos atributos de solo, visa representar o conceito teórico que fundamenta a Equação 1. A seleção das covariáveis preditoras foi baseada nos critérios: facilidade de obtenção, disponibilidade, uso frequente em trabalhos de MDS disponíveis na literatura e, conforme considerado por Vaysse e Lagacherie (2014), a relação lógica entre processo de formação do solo e atributos. Em função das dimensões da área, o fator clima foi considerado uniforme em toda sua extensão. As covariáveis selecionadas foram relacionadas aos fatores: organismos (o), e relevo (r), esse último presumido como o principal fator determinante na definição dos atributos dos solos na área. O fator material de origem (p), embora seja conhecido para a área da Fazenda Edgárdia, baseado no mapa geológico do convênio DAEE-Unesp, coordenado por Landim (1984), não foi considerado na modelagem em função da discrepância da escala de representação (1:250.000), em relação à escala adotada no trabalho e, em um teste preliminar, quando adotado como covariável categórica auxiliar, resultou em única covariável preditora na RLM.

As covariáveis aplicadas no modelo foram de natureza contínua, compreendendo informações do relevo e suas derivadas e índices das imagens de sensoriamento remoto.

#### 3.3.4.1 Obtenção do MDE e variáveis morfométricas

Em função da abrangência local do trabalho, o MDE foi obtido a partir das informações planialtimétricas digitalizadas (curvas de nível, hidrografia e pontos cotados) a partir da carta topográfica, na escala 1:10.000. A representação do MDE foi gerada pela interpolação dos dados, aplicando-se a ferramenta *topo to raster* do programa ArcGIS v. 10, com uso do aplicativo ANUDEM, que provê um MDE consistente do ponto de vista hidrológico (HUTCHINSON, 1993).

Na seleção da resolução espacial a ser adotada, para obtenção das variáveis de relevo a serem utilizadas na modelagem, foram consideradas: a escala da carta base, a compatibilidade com as imagens SPOT, o tamanho da área e a representação adequada da altimetria da área, de acordo com a escala do levantamento pedológico, conforme preconizado por Dobos et al. (2000).

Uma vez que os dados da altimetria foram provenientes de curvas de nível com equidistância de 5m, foi observada a distância entre as curvas nas áreas de maior declividade, para avaliar a adequação da resolução na geração do MDE, buscando maior coerência na operação. Em função de todas as considerações, e buscando avaliar o potencial de um MDE mais detalhado, como considerado por Carvalho Júnior et al. (2014a), foi escolhido o tamanho do pixel das representações *raster* igual a 10 x 10 m.

A qualidade foi avaliada pela sobreposição de curvas de nível obtidas a partir do modelo gerado, em comparação com as digitalizadas da carta e pela avaliação da presença de depressões espúrias (*sinks*), constituídas por pixels em desacordo com o escoamento de fluxo das superfícies, conforme sugerido por Hutchinson e Gallant (2000). Tais depressões resultaram do processo de interpolação e foram corrigidas (DOBOS et al., 2000), por aplicação sucessiva de uma suavização, de modo a contribuir positivamente na qualidade das derivadas de relevo geradas a partir do MDE.

O relevo da Fazenda Edgárdia foi considerado o principal fator dentre os que determinam a variação dos atributos de solos, uma vez que a topografia da área é caracterizada pela presença da cuesta, implicando em um gradiente significativo de altitude e variação geomorfológica. A sobreposição, em ambiente SIG, das camadas de informação MDE e mapa do levantamento semidetalhado de solos possibilitou identificar a relação entre a ocorrência das classes de solos e condições da topografia, conforme o princípio da relação solo-paisagem (RESENDE et al., 2014).

A morfometria da área foi representada pelo próprio MDE, no caso valores de altitudes, representando a superfície do terreno, como destacado por Bock e Köthe (2008) e suas derivadas, variáveis morfométricas primárias e secundárias. Foram adotadas como covariáveis, as variáveis que, de acordo com a literatura, são mais comumente correlacionadas com atributos de solo (CARVALHO JÚNIOR et al., 2014a; MOORE et al., 1993). O conceito e formulação para cálculo dessas variáveis podem ser encontrados em Wilson e Gallant (2000).

Foram obtidas as representações de:

- a) Elevação (Elev): matriz de valores do MDE, com pixels expressos em metros, caracterizando a energia potencial do relevo.
- b) Declividade (Decl): na extração automática é definida para um pixel central e os valores dos vizinhos, calculada pela técnica da média máxima dos

- gradientes de altitude, nas direções X e Y. Tem influência na ocorrência de fluxos subsuperficiais e superficiais e, conseqüentemente, erosão nas encostas.
- c) Curvatura planar ou curvatura horizontal (CPL) e perfil de curvatura ou curvatura vertical (CPE): as curvaturas representam as formas do terreno em tendências convergente, retilínea e divergente no plano horizontal (de projeção) e côncavas, convexas ou planas, no sentido das encostas. As curvaturas estão relacionadas à dispersão ou acúmulo dos fluxos de água de solo, bem como sua velocidade, afetando a formação do solo e a erosão.
- d) Aspecto (Aspc): apresenta uma amplitude de variação possível de 0 a 360°, caracterizando o azimute da vertente. Esta variável tem relação com a insolação na superfície, o que afeta principalmente a abundância e o desenvolvimento de fauna e flora. Para a aplicação do modelo de regressão, os valores angulares de Aspc foram convertidos por meio do cálculo das funções seno e cosseno, sendo ambos avaliados como variáveis auxiliares nas regressões lineares múltiplas (RLM).
- e) Fator LS (FLS): representa uma combinação do comprimento de rampa e declividade, empregado em estudos de erosão dos solos para descrever a influência da topografia nesse processo. Fator usado na Equação Universal de Perda de Solo (EUPS), calculado com base na declividade e área específica de contribuição, pela equação de Moore et al. (1991), adequada a terrenos de topografia complexa (Equação 9).

$$LS = (n + 1) \left( \frac{A_s}{22,13} \right)^n \left( \frac{\sin \beta}{0,0896} \right)^m \quad (9)$$

Onde  $LS$  é o fator LS derivado da teoria da potência do escoamento,  $A_s$  é a área de contribuição específica e  $\beta$  é a declividade de cada pixel.  $n = 0,4$  e  $m = 1,3$  (MOORE et al., 1991).

- f) Índice topográfico de umidade (ITU): índice que reflete a tendência de acúmulo de água no solo e está diretamente relacionado à área de contribuição e inversamente à declividade.

O ITU é uma variável morfométrica secundária, que representa a tendência ao acúmulo de água no terreno, determinado, segundo Beven e Kirkby (1979) para aplicação em modelo hidrológico, pela relação da Equação 10.

$$ITU = \ln\left(\frac{Ac}{\tan\beta}\right) \quad (10)$$

Onde *ITU* é o Índice Topográfico de Umidade, *Ac* é a área de contribuição ao pixel (m<sup>2</sup>) e  $\beta$  é a declividade de cada pixel (radianos).

As derivadas foram geradas de forma automática a partir do MDE corrigido, no programa SAGA GIS versão 2.2.2, que possui vários módulos para estudo de MDE, especialmente parâmetros de terreno na biblioteca de módulos Análise de Terreno (BOCK; KÖTHE, 2008; CONRAD et al., 2015; OLAYA; CONRAD, 2009). Cada derivada do relevo deu origem a uma covariável preditora contínua numérica.

#### 3.3.4.2 Índices de vegetação

Os índices de vegetação foram aplicados como covariáveis nos modelos pedológicos, considerando a relação íntima entre vegetação e solo, que implica em correlação entre índices de vegetação e atributos dos solos, como observado por Zanzarini et al. (2013).

As relações foram obtidas a partir das bandas do vermelho e infravermelho próximo, respectivamente B2 e B3, das imagens multiespectrais SPOT 5, consideradas de média resolução. As imagens foram devidamente georreferenciadas, submetidas à conversão de níveis digitais em valores de reflectância e finalmente, a uma operação de transformação radiométrica, segundo Ponzoni et al. (2012). O processamento das imagens foi realizado no programa GRASS versão 6.4.4.

Foram estabelecidos o índice de vegetação da diferença normalizada (NDVI) e o índice de vegetação ajustado ao solo (SAVI). O NDVI (ROUSE et al., 1973), reflete a quantidade e vigor de biomassa e é amplamente utilizado em estudos de vegetação por sensoriamento remoto, foi calculado pela Equação 11.

$$NDVI = \frac{(NIR-R)}{(NIR+R)} \quad (11)$$

Onde: *NIR* e *R* correspondem aos comprimentos de onda infravermelho próximo (banda 3) e vermelho (banda 2), respectivamente.

O SAVI (HUETE, 1988), reflete a condição de biomassa com um ajuste da influência do solo na resposta espectral (Equação 12):

$$SAVI = \frac{(1+L)(NIR-R)}{L+NIR+R} \quad (12)$$

Onde *NIR* e *R* correspondem aos comprimentos de onda infravermelho próximo (banda 3) e vermelho (banda 2), respectivamente, sendo *L* considerado igual a 0,5.

A resolução espacial das imagens igual a 10 m gerou representações matriciais dos índices de vegetação, com tamanho de pixel compatível às demais informações, de 10 x10 m.

### 3.3.5 Análise exploratória dos dados

O comportamento dos dados utilizados para a modelagem foi inicialmente avaliado e descrito por medidas de posição e dispersão da estatística clássica, como indicado por Landim (2003), além de incluído o parâmetro coeficiente de variação médio (CVm), de acordo com recomendação de Castrignanò (2011), para permitir a comparação entre os atributos, por constituírem diferentes grandezas e unidades.

A distribuição de frequência dos dados foi analisada de forma quantitativa pelos coeficientes de assimetria e curtose e avaliada a ocorrência de dados discrepantes. No entanto, buscou-se manter a integralidade dos bancos de dados em virtude do número restrito de amostras disponíveis.

### 3.3.6 Seleção da predição espacial dos atributos de solo

Os valores georreferenciados dos atributos de solo estudados foram assumidos como realizações pontuais de variáveis aleatórias  $Z(x)$ , como na Equação 2, constituindo dois *shapes* de pontos amostrais georreferenciados, um para treinamento e outro para validação externa.

No ambiente SIG foram inseridos o *shape* de pontos de treinamento e os dados *raster* com a informação das covariáveis ambientais (MDE com as derivadas e os índices de vegetação), com sistema de referência único, propiciando a captura de valores.

Para a aplicação da técnica de modelagem digital dos atributos de solos baseada em variáveis secundárias é presumida correlação entre os atributos e fatores de formação do solo, representados pelas covariáveis ambientais. O teste baseado no coeficiente de correlação momento-produto de Pearson foi aplicado com a finalidade de testar a associação entre os pares de valores amostrais e valores das covariáveis (CARVALHO JÚNIOR et al., 2014a; FLORINSKY, 2012a).

Uma análise prévia de adequação de condições auxiliou na definição do modelo de predição espacial adotado. A seleção do modelo foi baseada em algumas premissas representadas por uma chave (HENGL, 2009), onde basicamente foram avaliados: o conhecimento do modelo físico estudado, a correlação do atributo de solo com as covariáveis ambientais, a autocorrelação das variáveis estudadas e dos resíduos da regressão. A seleção do método preditor para cada atributo de solo, baseada nessas comparações, seguiu uma adaptação da regra simplificada, apresentada por Ciampalini et al. (2012), exposta na Tabela 4.

**Tabela 4 – Regra de decisão para seleção da função de MDS**

		Há correlação com as variáveis ambientais	
		Sim	Não
Há estrutura espacial	Sim	Regressão-krigagem	Krigagem ordinária
	Não	Regressão linear	Média ou IDW

**Fonte:** Adaptado de Ciampalini et al. (2012)

### 3.3.7 Aplicação do modelo híbrido

Nesse modelo, os passos metodológicos compreendem a definição dos coeficientes do modelo determinístico geral do atributo na área e o tratamento geoestatístico dos resíduos desse modelo, com a posterior soma dos dois mapas resultantes, como descrito por Hengl et al. (2007b) e Hengl (2009).

A regressão-krigagem foi aplicada utilizando inicialmente os modelos RLM, para estabelecer a relação entre os atributos de solo, nas duas profundidades e as covariáveis ambientais Elev, Decl, CPL, CPE, Aspc, ITU, FLS, NDVI e SAVI, contemplando a componente determinística da Equação 2. As covariáveis ambientais atuaram como preditores densamente conhecidas na área, combinados linearmente para quantificar a variável resposta, o atributo de solo.

Para a aplicação da krigagem ordinária, foi verificado se os resíduos apresentavam autocorrelação espacial e seguiam a hipótese básica de estacionaridade intrínseca. Essa hipótese expressa que a média e a variância dos resíduos independem da posição absoluta dos dados, sendo dependente unicamente do afastamento entre os pontos ou *lag* (ARMSTRONG, 1998).

O passo inicial foi o ajuste do modelo teórico e definição dos parâmetros do variograma e posterior aplicação desse modelo para predição dos valores em todos os pontos da área. Finalmente, pela soma algébrica dos mapas RLM (modelo determinístico) e KO dos resíduos, foi gerado o mapa final de predição do atributo do solo, conforme Odeh et al. (1995). Do processo de modelagem resultaram um mapa de predição, um de regressão, um dos resíduos e um de qualidade da predição, expresso pelo desvio padrão.

Quando não foi observada correlação do atributo de solo com as covariáveis ambientais, mas houve autocorrelação espacial, a estimação dos valores do atributo ao longo da área foi obtida pela interpolação utilizando a KO dos valores do atributo (MATHERON, 1963; YAMAMOTO; LANDIM, 2013) pela Equação 6.

#### 3.3.7.1 Avaliação das covariáveis e RLM

A seleção de combinações das covariáveis preditoras na aplicação dos modelos foi de acordo com a Regra de Hazor (McBRATNEY et al., 2003), que avaliou suas performances ao gerar os mapas de predições de atributos de solos.

Uma análise prévia relevante para o sucesso de modelagens usando regressão consiste na avaliação de multicolinearidade dos preditores (HENGL, 2009). Foi calculado então o coeficiente de correlação entre as covariáveis e, com objetivo de selecionar as covariáveis a serem consideradas no modelo, foi avaliado o coeficiente de correlação e aplicado o método stepwise (HAIR et al., 2009). Nesse método é avaliada a contribuição individual de cada variável inserida no modelo, para a explicação da variável resposta. Dentre as variáveis consideradas, é realizado um julgamento, para definir se devem ser mantidas todas, ou se alguma deve ser eliminada, de acordo com o valor do Critério de Informação de Akaike (AIC), que considera o número de covariáveis utilizadas e o ajuste do modelo. Se for considerada não significativa ou redundante, a variável é descartada. Todas as variáveis foram consideradas na seleção passo a passo e, o número mínimo de variáveis foi mantido no final do processo, de acordo com o menor valor do AIC. Dessa forma, foi aplicado o princípio da parcimônia nas RLM.

Para o modelo de RLM foi definido, arbitrariamente, conforme mencionado por Moore et al. (1993), um nível de significância, no caso de 5%. A adequação do modelo foi expressa pelo coeficiente de determinação  $R^2$  ajustado da regressão.

Aplicada a seleção e a RLM, o resultado gerou um mapa de predição dos atributos, um mapa de resíduos da RLM e informações sobre o modelo de regressão e coeficiente de determinação. Foi apresentado também a importância dos preditores no modelo, listados de acordo com o percentual de sua participação na explicação da variação do atributo no modelo.

Dos valores estimados pela regressão, em relação aos valores amostrais, foram obtidos os resíduos, os quais contém, como ressaltado por Omuto e Vargas (2015), a autocorrelação de pequena escala e mais os erros de modelagem e os decorrentes da determinação analítica e da amostragem. Os erros de modelagem são devido às deficiências da regressão. Esses resíduos foram submetidos à análise de autocorrelação, pelo cálculo do variograma.

### **3.3.7.2 Krigagem ordinária dos resíduos e mapa de predição**

A realização da inferência de um modelo para descrever o comportamento da variável regionalizada foi baseada em estatística de dois pontos, como descrito em Yamamoto e Landim (2013), avaliando a relação espacial entre pontos amostrais.

De acordo com a técnica selecionada, baseada na Tabela 4, os dados considerados foram os originais do atributo de solo (quando não ocorreu correlação atributo-covariáveis) ou os resíduos da RLM (para aplicação da RK).

A função variância foi adotada para tal fim, buscando-se um ajuste teórico ao modelo experimental. A avaliação da dependência espacial dos atributos de solo foi realizada pelo cálculo e ajuste de variogramas, pressuposta a estacionaridade da hipótese intrínseca (JOURNAL; HUIJBREGTS, 1978), pela Equação 13:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad (13)$$

Onde  $h$  é o vetor distância entre pares de dados,  $N(h)$  é o número de pares e  $Z(x_i)$  e  $Z(x_i+h)$  são os valores medidos da variável aleatória em pontos conhecidos.

A função variograma é definida como o módulo do vetor  $h$  quando são aplicados modelos já conhecidos, como por exemplo, esférico, exponencial, gaussiano, ou mesmo combinações entre eles, ajustando os valores dos parâmetros efeito pepita, patamar e alcance (ARMSTRONG, 1998). No ajuste dos variogramas teóricos aos variogramas experimentais, no intuito de descrever a tendência espacial dos atributos de solo, foi avaliada a adequação dos modelos esférico e exponencial, mais comumente adotados para variáveis de solos, conforme literatura (CARVALHO JÚNIOR et al., 2014b; GOMEZ et al., 2012).

Foi avaliada a consideração de busca global ou local de vizinhos no processo de interpolação, nesse último caso adotando o critério de mínimo de 7 e máximo de 25, de acordo com recomendação de Oliver e Webster (2014). Em função do número e distribuição das amostras, preferencialmente, foi considerada a vizinhança global. O passo do variograma (*lag*) foi inicialmente definido em valor próximo a 10% da metade da distância máxima observada na área.

Os variogramas foram considerados isotrópicos e do melhor ajuste foram definidos: o modelo adotado e os parâmetros patamar, efeito pepita e alcance, ou alcance efetivo (ARMSTRONG, 1998; YAMAMOTO; LANDIM, 2013).

O variograma experimental foi definido estabelecendo-se o valor do *lag* e procedeu-se ao ajuste de um modelo teórico ao variograma, dentre os modelos disponíveis. O método de ajuste adotado foi a sentimento, com julgamento baseado

em análise visual de adequação da curva teórica, alterando os valores dos parâmetros  $C_0$ ,  $C$  e  $A$ , respectivamente, efeito pepita, componente estrutural e alcance. O ajuste foi avaliado pelo coeficiente  $R^2$ , com ênfase na parte inicial do variograma (mais próximo à origem) e avaliando o número de pares de dados para cálculo da variância. O ajuste foi realizado por método iterativo e, com o modelo ajustado, o mesmo foi usado para aplicação da krigagem ordinária dos resíduos.

Da modelagem, foram gerados os mapas finais de predição e de qualidade da predição.

Uma vez que a krigagem é um interpolador exato e não convexo, sua aplicação resultou em valores fora dos limites observados nas amostras e fora da amplitude real dos atributos, produzindo valores incoerentes, para o que foi aplicada uma correção, como preconizado por Hengl et al. (2007b), por meio da ferramenta calculadora *raster* do programa QGIS.

Para o atributo pH foi aplicada a interpolação pelo método do inverso da distância – IDW. Finalmente foi feita a composição dos mapas dos atributos de solo da Fazenda Experimental Edgárdia.

### 3.3.8 Validação

A qualidade do ajuste e, portanto, do modelo escolhido, idealmente deve ser julgada por um método de validação, preferencialmente externa, possibilitando determinar os parâmetros estatísticos do erro médio, erro médio quadrático (OLIVER; WEBSTER, 2014; YAMAMOTO; LANDIM, 2013; CHILÈS; DELPHINER, 2012).

A acurácia dos mapas de atributos de solos resultantes foi avaliada baseada na comparação entre os valores preditos pelo modelo e valores de referência medidos em campo, por meio de validação externa. Os dados de validação foram selecionados de forma aleatória e reservados a partir do banco de dados original, especificamente para essa finalidade, não sendo considerados no estabelecimento dos modelos. Um total de 27 amostras foi destinado à validação.

A avaliação dos modelos se deu por meio do cálculo do Erro Médio da Predição (ME) (Equação 14) e a acurácia da predição, dada pela raiz quadrada do erro quadrático médio (RMSE) (Equação 15). No cálculo desses parâmetros foram

considerados os valores observados nos pontos de validação,  $Z(x_i)$  e os valores estimados por meio do modelo  $\hat{Z}(x_i)$ :

$$ME = \frac{1}{n} \sum_{i=1}^n [\hat{Z}(x_i) - Z(x_i)] \quad (14)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [\hat{Z}(x_i) - Z(x_i)]^2} \quad (15)$$

Onde  $n$  é o número de pontos de validação,  $\hat{Z}(x_i)$  é o valor estimado do atributo e  $Z(x_i)$  o valor obtido por medição.

Um melhor desempenho do modelo resulta em valores de erro médio próximo de zero.

### 3.3.9 Subdivisão da área

Foi adotada a estratificação da área total em duas, mais uniformes em termos altimétricos, para aplicação dos modelos de predição. A altitude de 510 m foi adotada como limite para separação entre área baixa e alta da Fazenda Edgárdia, uma vez que essa curva de nível ocupa uma posição limítrofe entre a área mais íngreme, representada pela Cuesta Basáltica, e a Depressão Periférica, com relevo mais suave e plano.

As RLM e variogramas foram calculados utilizando somente os dados pertencentes a cada subárea, aplicados os modelos de predição e avaliada a diferença da modelagem entre a área total e as duas subáreas, utilizando a estatística descritiva.

## 4 RESULTADOS E DISCUSSÃO

### 4.1 Análise exploratória dos dados

A análise exploratória dos dados de atributos de solo permitiu avaliar o seu comportamento e sua variação. Foram calculadas medidas de posição e dispersão e distribuições de frequências. Os resultados de média, desvio padrão, coeficiente de variação, coeficiente de variação médio, valores máximos e mínimos e coeficientes de forma dos atributos de solo estão apresentados na Tabela 5.

**Tabela 5 – Estatística descritiva dos atributos de solo**

	M	Md	DP	Curtose	As	Min	Max	CV	CVm
IAV1	6,50	7,00	2,56	-0,39	-0,30	0,01	10,67	39,32	3,33
IAV2	9,59	10,50	4,66	-0,75	0,13	0,01	18,00	48,60	4,11
pH1	5,53	5,60	0,74	-0,15	-0,22	3,70	7,30	13,47	1,13
pH2	5,50	5,60	0,83	-0,45	-0,36	3,60	7,20	15,05	1,28
CTC1	165,76	126,00	142,56	1,14	1,37	20,00	602,00	86,00	7,27
CTC2	166,88	121,00	144,63	1,06	1,27	15,00	607,00	86,67	7,32
V1	63,40	66,00	20,68	-0,23	-0,63	11,00	98,00	32,61	2,76
V2	59,64	67,00	26,68	-0,95	-0,59	5,00	95,00	44,74	3,78
Ar1	648,46	703,00	257,43	-0,74	-0,66	10,00	950,00	39,70	3,35
Ar2	591,00	630,00	269,61	-1,31	-0,29	20,00	940,00	45,62	3,86
Arg1	170,50	140,00	132,07	0,87	1,09	10,00	600,00	77,46	6,55
Arg2	235,63	190,00	167,71	-0,74	0,63	10,00	660,00	71,18	6,02

IAV: Índice de avermelhamento; CTC: Capacidade de troca de cátions em mmolc.dm<sup>-3</sup>; V: Saturação por bases em %; Ar: Teor de areia em g.kg<sup>-1</sup>; Arg: Teor de argila em g.kg<sup>-1</sup>; M: média; Md: mediana; DP: Desvio padrão; As: Coeficiente de assimetria; CV: Coeficiente de variação em %; CVm: Coeficiente de variação médio em %; 1 = camada superficial; 2 = camada subsuperficial; obs: número total de amostras = 114.

Embora alguns dados apresentassem valores afastados das médias, implicando em variação considerada alta, apontada pelas medidas de dispersão, evitou-se a eliminação de dados em função de limitado número de pontos amostrais e por considerar real a possibilidade de ocorrência desses valores, relacionada à natureza das formações geológicas locais.

Dentre os atributos, a CTC e teor de Argila foram os de maior variabilidade, sendo considerada alta, na camada superficial e subsuperficial, de acordo com Warrick e Nielsen (1980), em função dos valores do coeficiente de variação acima de 60%. Pelo mesmo critério, todos os demais atributos apresentaram média variabilidade (CV entre 12 e 60%). De modo geral foi observado um incremento nos valores de CV com a profundidade, exceto para a argila. A variabilidade dos atributos estudados pode ser decorrente da heterogeneidade das condições locais, considerando principalmente o relevo e a litologia locais, condicionando os processos de formação dos solos na área.

O atributo pH nas duas camadas foi o que apresentou distribuição mais próxima da normal e valores de coeficiente de assimetria mais próximos de zero (Tabela 5), indicando que os valores médios desses atributos representariam bem o conjunto de dados. O IAV nas duas camadas apresentou distribuição de frequência irregular, mas ainda mantendo simetria, apresentando também baixos valores de coeficiente de assimetria. Os atributos CTC e Argila foram assimétricos à direita, com valores de coeficiente de assimetria considerado alto, acima de 1,0. Areia e V%, principalmente na camada subsuperficial, foram assimétricos à esquerda, com valores de coeficiente de assimetria também mais altos.

Geralmente quando os dados não apresentam distribuição normal, são submetidos a uma transformação, uma vez que ela é presumida em vários métodos estatísticos. Um critério para o julgamento da necessidade de transformação foi apresentado por Webster (2001), segundo o qual um coeficiente de assimetria positiva maior que 0,5 já indica a aplicação da raiz quadrada e acima de 1,0 é indicado a transformação logarítmica. No entanto, após aplicação de tratamento estatístico e estabelecimento dos modelos, é feita transformação inversa, para voltar à escala original dos dados, incorrendo em valores e características estatísticas diversas daquelas do conjunto de valores originais. Tal fato pode justificar o tratamento estatístico dos dados sem realizar a transformação.

Neste estudo foi testada a transformação logarítmica para os atributos que tiveram resultado de coeficiente de simetria acima de 1,0, mas mesmo os dados transformados não apresentaram uma distribuição normal típica.

Winter (2013) comentou que a transformação de dados pode auxiliar na elaboração de modelos em alguns aspectos, mas que não há unanimidade sobre essa prática para aplicação de modelos. Além disso, a aplicação de métodos

geoestatísticos, como a krigagem, estende-se a dados que não seguem essa distribuição (BREGT et al., 1991). Ainda, conforme Silva (2013), parece ser mais coerente preservar a estrutura dos dados em um estudo onde é analisada a sua variabilidade. Diante dessas considerações, optou-se por trabalhar com os dados originais. Distorções foram passíveis de ocorrer na interpretação da incerteza, balizada pelo cálculo do desvio padrão das estimativas, uma vez que essa medida de dispersão é baseada em uma distribuição normal.

De modo geral, a camada subsuperficial apresentou maior variabilidade dos atributos considerados. A rigor, a camada superior dos solos apresenta comumente uma maior variabilidade, mas nas condições do local de estudo, a heterogeneidade do material de origem pode ser a causa dessa ocorrência.

#### **4.2. Covariáveis ambientais**

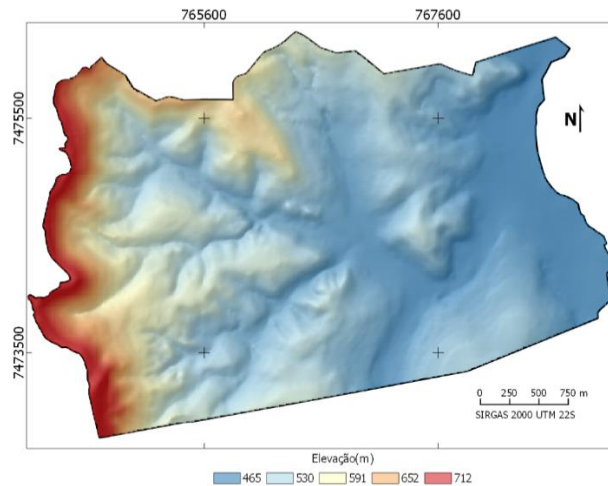
As representações das covariáveis preditoras no formato raster, com resolução de 10 m (pixel de 100 m<sup>2</sup>), estão apresentadas nas Figuras 5 a 9, sendo: a representação da elevação (MDE), seis derivadas do MDE e dois índices derivados das imagens de sensoriamento remoto. Essas variáveis pertencem ao conjunto indicado por Florinsky (2012a) como preditores mais utilizados e foram aplicadas no MDS em diferentes condições pedogenéticas, constando entre as adotadas por Carvalho Junior et al. (2014a, b), Ciampalini et al. (2012), Menezes et al. (2016), Moore et al. (1993) e Pradhan (2010).

Os valores de elevação (Figura 5) variaram de 465 a 712m, com a parte oeste da fazenda apresentando maior altitude, a partir da faixa estreita do planalto ocidental, que constitui o limite da área. Seguindo para leste, o relevo é irregular, a presença do front da cuesta acarreta um decréscimo brusco das altitudes adentrando a Depressão Periférica, com menores altitudes encontradas na várzea do Rio Capivara, no lado leste da fazenda.

O relevo da área é assimétrico, apresentando valores de declividade entre o máximo de 87,95% e o mínimo de 0% (Figura 6). Os valores máximos são predominantemente relacionados ao front da cuesta, mas também são observados valores altos em encostas no interior da área. A predominância do valor de declividade mínima ocorre na várzea do Rio Capivara e ao longo da rede de

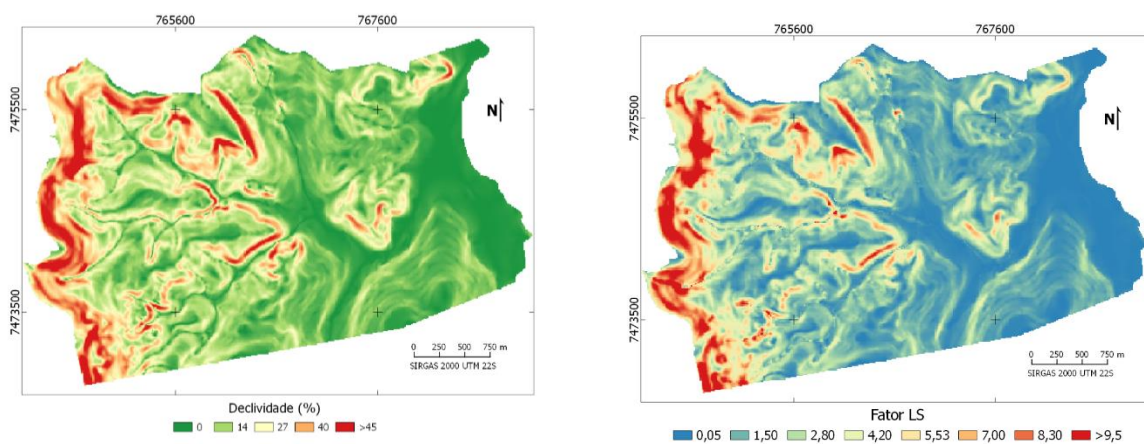
drenagem, com algumas áreas de relevo plano também observadas nos topos de morros.

**Figura 5 – Modelo digital de elevação da Fazenda Experimental Edgárdia**



O FLS (Figura 6) apresentou valores mínimos próximos a zero, na área de baixada, notadamente na várzea do Rio Capivara e valores baixos acompanhando toda a rede de drenagem. Os valores máximos ocorreram ao longo da escarpa da cuesta (da ordem de 18,00) e em algumas vertentes mais íngremes no interior da área. O valor médio calculado resultou em 2,50 e o desvio padrão de 2,48.

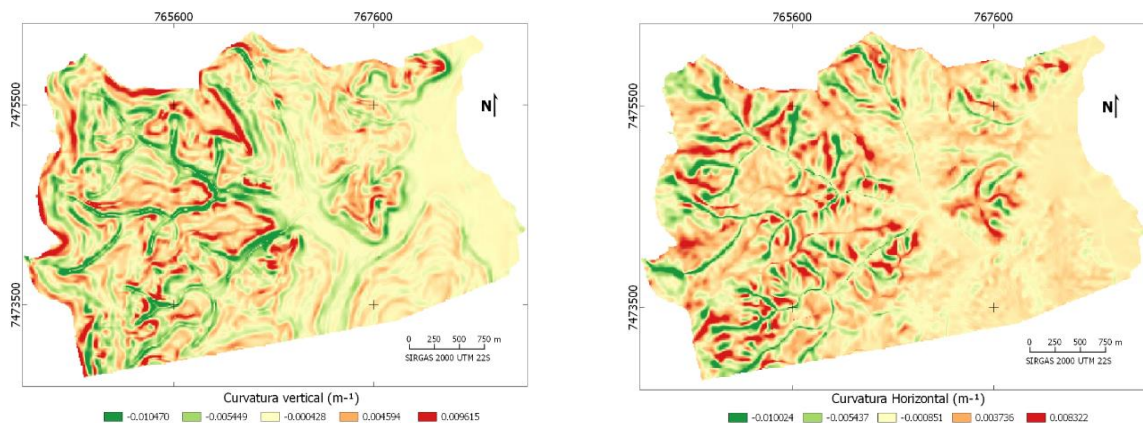
**Figura 6 - Declividade (Decl) e Fator LS (FLS) da Fazenda Experimental Edgárdia**



Apesar de não ocorrerem na natureza superfícies perfeitamente retilíneas, para as curvaturas da área (CPL e CPE), grande parte apresentou valores nulos, caracterizando essa forma de relevo (Figura 7).

Pode ser observada uma combinação de valores positivos e negativos das curvaturas, sobretudo da região central para o oeste da área. Nas áreas de topos de morros as feições apresentaram, principalmente, curvatura vertical convexas e curvatura horizontal divergente, favorecendo a infiltração de água. Nas partes mais baixas das encostas, curvaturas verticais côncavas e horizontais convergentes favorecem a concentração do fluxo de escoamento superficial, propiciando a ocorrência de erosão dos solos. O material erodido das partes altas tende a ser depositado nas áreas baixas, principalmente, próximo à rede de drenagem e nas áreas de várzeas.

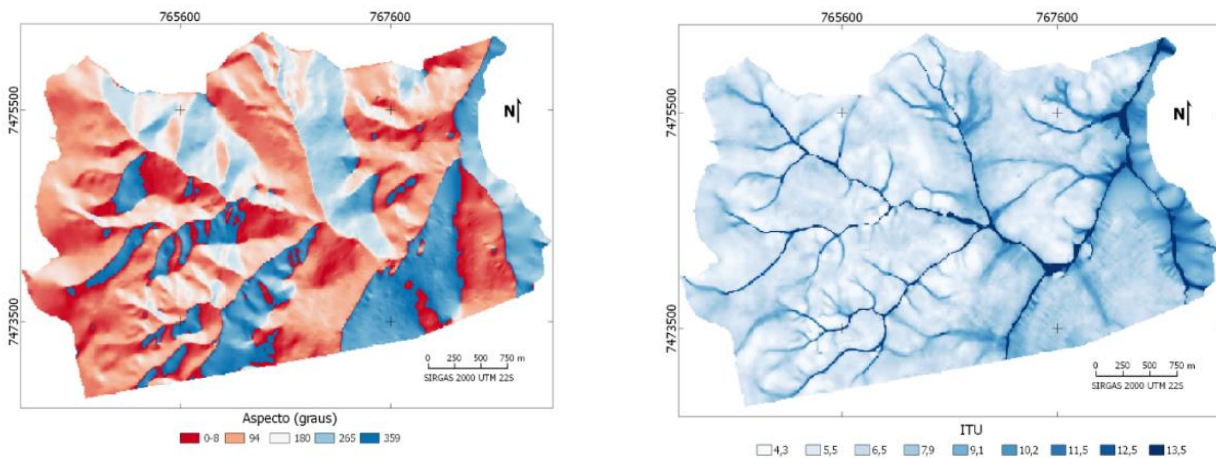
**Figura 7 – Curvatura vertical (CPE) e horizontal (CPL) da Fazenda Experimental Edgárdia**



A caracterização do aspecto (Figura 8) evidenciou predomínio de faces voltadas para o leste, com azimutes variando de 0 a 360°, com um valor médio de 154°. Para avaliação nas regressões lineares múltiplas, foram utilizados os mapas representando as funções  $\sin A$  e  $\cos A$ .

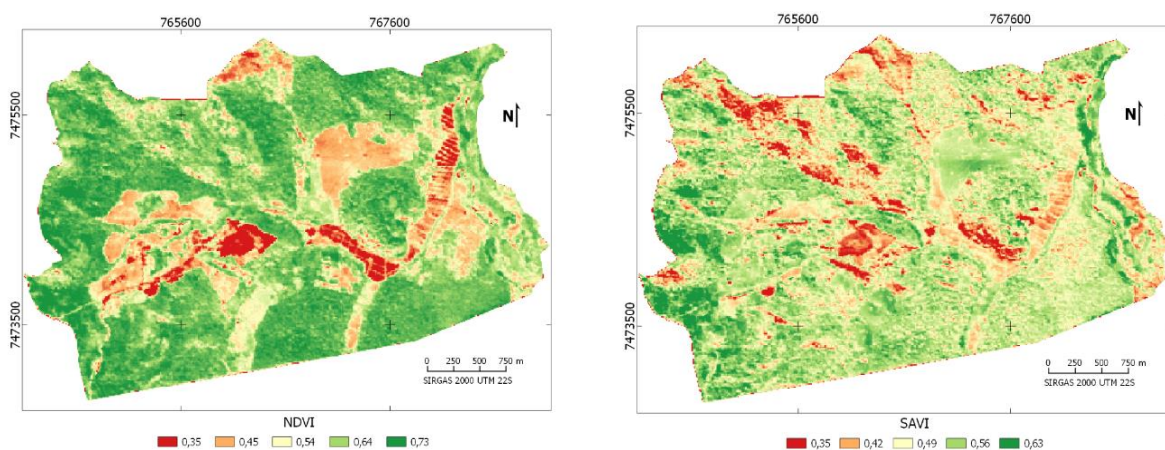
Para o ITU (Figura 8) foi observado um valor médio igual a 6,91 e desvio padrão de 2,15. Os valores mais altos ocorrem em áreas úmidas, nas planícies da rede de drenagem, onde há acúmulo de água no solo em função das condições de relevo. Na várzea do Rio Capivara, os valores foram acima de 10,0 e na região do front da cuesta e nos topos de morros os valores foram menores que 5,0.

**Figura 8 – Aspecto (Aspc) e Índice Topográfico de Umidade (ITU) da Fazenda Experimental Edgárdia**



Os índices NDVI e SAVI foram ambos aplicados como preditores, uma vez que permitiu avaliar também se a influência do solo considerada no SAVI interfere na correlação entre atributos de solo e os índices de vegetação. Tais índices foram considerados importantes preditores em MDS, como observado por Rad et al. (2014). Malone et al. (2009) também consideraram e constataram que o NDVI é um importante preditor, relacionado ao fator de formação organismos, expressando a influência de cobertura e uso da terra. Os valores médios foram de 0,51 e 0,62 com desvios-padrão de 0,06 e 0,09, para SAVI e NDVI, respectivamente (Figura 9).

**Figura 9 – NDVI e SAVI da Fazenda Experimental Edgárdia**



### 4.3 Covariáveis e correlações

Para estabelecer um modelo de regressão de forma adequada, as covariáveis não devem apresentar multicolinearidade e a correlação entre elas deve ser considerada na seleção de variáveis do modelo. Na Tabela 6 é apresentada a relação entre as covariáveis numéricas disponíveis para a modelagem, com os respectivos valores do coeficiente de correlação.

Foi observada correlação alta (acima de 60%) e positiva entre as covariáveis NDVI e SAVI, explicável pela semelhança desses índices, o que indica que, para um determinado modelo, deve ser utilizado apenas um deles. O mesmo ocorreu entre FLS e Decl, com correlação positiva igual a 0,903. De modo geral, os coeficientes de correlação entre as covariáveis numéricas podem ser considerados baixos, apesar de significativos a 5% (Tabela 6).

**Tabela 6 – Matriz de correlação entre as covariáveis ambientais**

	Elev	Decl	CPE	CPL	ITU	NDVI	SAVI	FLS	senA	cosA
Elev	1									
Decl	0,450	1								
CPE	0,284	-0,284	1							
CPL	0,239	-0,064	0,409	1						
ITU	-0,437	-0,457	-0,106	-0,375	1					
NDVI	-0,092	0,285	-0,332	-0,202	-0,123	1				
SAVI	-0,016	0,077	-0,037	-0,080	-0,120	<b>0,653</b>	1			
FLS	0,397	<b>0,903</b>	-0,301	-0,344	-0,215	0,302	0,100	1		
senA	0,121	0,031	0,084	0,207	0,042	-0,085	0,049	0,057	1	
CosA	-0,050	-0,033	0,049	-0,034	-0,152	0,016	0,245	-0,064	-0,123	1

Elev: elevação, Decl: declividade, CPE: curvatura vertical, CPL: curvatura horizontal, senA: senA, cosA: cosA, ITU: índice topográfico de umidade, NDVI: índice de vegetação da diferença normalizada, SAVI: índice de vegetação ajustado ao solo, FLS: Fator LS.

Para avaliar a aplicação da regressão linear múltipla foi calculado inicialmente o coeficiente de correlação de Pearson entre atributos do solo e covariáveis ambientais (Tabela 7). A hipótese nula considerando a não correlação atributo-covariável com valores de p-value <0,005 foi descrita por Carvalho Júnior et al. (2014b). De modo geral, os coeficientes de correlação variaram, de -0,40 a 0,51, abaixo do valor apresentado por Zhu e Lin (2010) como referência para alta correlação, igual em módulo, a 0,6. Foram então avaliadas as correlações até um

nível de significância correspondente ao p-value < 0,05, possibilitando a inclusão de mais preditores nos modelos, como por exemplo, o FLS como variável preditora de IAV e Ar. Apesar dos valores de correlação permanecerem baixos, foram avaliados como significativos pelo teste estatístico (Tabela 7).

**Tabela 7 – Correlações entre atributos do solo e covariáveis**

	IAV		pH		CTC	
	1	2	1	2	1	2
Elev	0,06	0,45**	0,19	0,17	0,04	0,03
Decl	-0,02	0,26*	0,21	0,10	0,06	0,10
CPE	0,18	0,34**	0,02	0,02	-0,09	-0,16
CPL	0,20*	0,25*	0,06	-0,04	0	0
ITU	-0,34**	0,51**	-0,10	-0,04	0,06	0,02
NDVI	-0,04	-0,08	-0,08	-0,22	-0,05	-0,20
SAVI	0,06	0,06	-0,02	-0,13	-0,05	-0,04
FLS	-0,12*	0,15	0,20	0,12	0,10	0,11
senA	-0,15	0,07	0,02	0,10	-0,16	-0,16
cosA	-0,03	-0,07	-0,16	-0,16	-0,07	-0,04

	V%		Ar		Arg	
	1	2	1	2	1	2
Elev	0,18	0,19*	-0,39**	-0,40**	0,26**	0,20*
Decl	0,33**	0,32**	-0,23*	-0,22*	0,08	-0,01
CPE	-0,09	-0,25*	0,03	-0,05	0,09	0,21*
CPL	0,06	0	-0,06	-0,12	0,12	0,20*
ITU	-0,11	-0,05	0,13	0,19*	-0,12	-0,22*
NDVI	-0,04	-0,20*	0,13	0,24*	-0,23*	-0,38**
SAVI	-0,07	-0,20*	0,14	0,19	-0,16	-0,21*
FLS	0,31**	0,13	-0,25*	-0,21*	0,08	-0,05
senA	0,19*	0,33**	-0,12	-0,21*	0,06	0,25*
cosA	-0,16	-0,27**	0,06	0,08	0,06	-0,05

IAV: Índice de avermelhamento; CTC: Capacidade de troca de cátions em mmolc.dm<sup>-3</sup>; V%: Saturação por bases em %; Ar: Teor de areia em g.kg<sup>-1</sup>; Arg: Teor de argila em g.kg<sup>-1</sup>; 1 = camada superficial; 2 = camada subsuperficial. Elev: elevação, Decl: declividade, CPE: curvatura vertical, CPL: curvatura horizontal, Aspc: aspecto, ITU: índice topográfico de umidade, NDVI: índice de vegetação da diferença normalizada, SAVI: índice de vegetação ajustado ao solo, FLS: Fator LS. Destacados: correlações significativas, p-value < 0,005\*\* e p-value < 0,05\*.

Os atributos pH e CTC não apresentaram correlação significativa com nenhuma das covariáveis ambientais. Para esses atributos, a regressão linear múltipla foi desconsiderada, conforme regra prevista na Tabela 4.

Dentre as covariáveis ambientais consideradas, a Elev, Decl, ITU e NDVI foram os melhores preditores pois apresentaram correlação significativa com maior número de atributos do solo. As três primeiras podem indicar a importância preponderante do relevo na formação dos solos, fato apontado como recorrente por Dobos e Hengl (2009) e adotado como pressuposto na hipótese desse estudo. Bohner e Selinge (2006) enfatizaram o resultado satisfatório do uso de variáveis de terreno para inferir atributos de solo principalmente por descrever os processos laterais em topografia acidentada.

A correlação do NDVI com os atributos pode ter favorecido a predição, considerando as características de quantidade de informação e continuidade dos dados de sensoriamento remoto, como foi considerado por Padarian et al. (2014).

Os valores de correlação entre atributos de solo e covariáveis ambientais mesmo sendo significativos são comumente baixos, implicando em valores modestos de  $R^2$  para as RLM estabelecidas, como já constatado em pesquisas (CARVALHO JÚNIOR et al., 2014a; MOORE et al., 1993; SAMUEL-ROSA et al., 2015; UMALI et al., 2012).

Na Tabela 8 está apresentada a seleção do modelo de predição para cada atributo, nas diferentes camadas, conforme as regras previstas na metodologia.

A aplicação de RLM, com seleção das covariáveis pelo método stepwise e ajuste pelos mínimos quadrados ordinário, possibilitou modelar a tendência global da variação dos atributos de solo para a área total da fazenda. Foram estabelecidos os modelos RLM pelas equações, com respectivos coeficientes de determinação e representados espacialmente, para os atributos IAV, V%, Ar e Arg.

A estrutura de dependência espacial dos resíduos gerados na aplicação da RLM pôde ser modelada para a maioria dos atributos, nas duas camadas estudadas, propiciando a obtenção do modelo final por RK, para toda a área.

Para os dados originais da CTC foi definida estrutura de dependência espacial e o modelo aplicado foi a KO.

**Tabela 8 – Resultado da aplicação das regras de seleção do modelo de predição para cada atributo de solo para a área total**

Atributo do solo (Y)	Covariável selecionada na regressão	Estrutura de dependência espacial	Predição
IAV1	ITU, FLS	Sim	RK
IAV2	Elev, CPE, ITU	Sim	RK
V1	Decl	Sim	RK
V2	Decl, senA, cosA, CPE, NDVI	Sim	RK
Areia 1	Elev	Sim	RK
Areia 2	Elev, senA, FLS, NDVI	Sim	RK
Arg 1	Elev, NDVI	Sim	RK
Arg 2	senA, NDVI, ITU	Sim	RK
CTC1	-	Sim	KO
CTC2	-	Sim	KO
pH1	-	Não	IDW
pH2	-	Não	IDW

Elev: elevação, Decl: declividade, CPE: curvatura vertical, CPL: curvatura horizontal, Aspc: aspecto, ITU: índice topográfico de umidade, NDVI: índice de vegetação da diferença normalizada, SAVI: índice de vegetação ajustado ao solo. IAV: índice de avermelhamento do solo, Arg: argila. RK: regressão krigagem; KO: krigagem ordinária; ID: inverso do quadrado da distância; 1 = camada superficial; 2 = camada subsuperficial.

Por sua vez, os dados de pH do solo não foram correlacionados às covariáveis e também não apresentaram estrutura de dependência espacial, sendo a estimativa realizada por um interpolador determinístico, no caso, foi adotado o Método do Inverso da Distância (IDW).

Os mapas de predição dos atributos, resultantes das RLM, de modo geral, expuseram padrões característicos da variação espacial observada nos mapas das covariáveis preditoras, usadas na calibração dos modelos.

#### 4.4 Modelos de predição dos atributos de solo

São apresentados a seguir as equações e mapas de predição, para cada atributo de solo, de acordo com cada modelo aplicado.

#### 4.4.1 Índice de avermelhamento do solo (IAV)

A predição do IAV foi realizada por RK, portanto, pela modelagem das componentes determinística e estocástica, em duas etapas e posterior soma dos dois mapas. As RLM foram estabelecidas, considerando as correlações significativas entre o atributo IAV e os preditores ITU e o FLS, sendo este último de forma inversa, para a camada superficial. Para a camada subsuperficial, o ITU foi mantido com a mesma relação inversa e houve ainda correlação positiva do IAV com a elevação e a curvatura vertical. Os valores de correlação entre IAV e covariáveis predictoras variaram de 0,12 a 0,51 (Tabela 7), sendo verificada maior correlação com o ITU. Os valores desse índice podem ser relacionados à pedogênese por permitirem a identificação de áreas de convergência e fluxos preferenciais, com valores altos nas regiões mais planas, próximas aos cursos d'água. As equações ajustadas, com os respectivos coeficientes e o  $R^2$  ajustado, estão apresentadas na Tabela 9.

Os valores do coeficiente de determinação, iguais a 0,13 para a camada superficial e 0,36 para a camada subsuperficial, foram considerados baixos, principalmente, para a camada superficial. Isso indica que, embora significativo, o modelo RLM não explica amplamente a variabilidade do IAV na área, em função das correlações baixas entre as covariáveis predictoras e o atributo de solo. A variabilidade do IAV deve estar relacionada a covariáveis não consideradas nesse estudo, como por exemplo, a litologia.

A correlação inversa do IAV com o ITU traduz a imposição da condição de umidade determinando de forma decisiva a cor do solo, o que ocorre na área de estudo. Os solos que contém maior quantidade de óxidos de Fe e se encontram em condições de boa drenagem, têm geralmente por característica cores mais vibrantes, em função da predominância de hematita e goethita. Esses óxidos de ferro conferem respectivamente, cores avermelhadas, amareladas ou brunadas ao solo e a coloração mais avermelhada característica da formação de hematita ocorre em condições mais secas que a goethita (CURI; FRANZMEIER, 1984). Em condições de drenagem deficiente, como em áreas de baixadas, o ambiente redutor confere coloração acinzentada ao solo, pela ausência de ferro na forma oxidada, conforme apontado por Lepsch (2011). Na área de estudo tal fato ocorre na proximidade da rede de drenagem, principalmente na várzea do Rio Capivara. Esse aspecto teve melhor representação na camada subsuperficial, em função das outras

covariáveis correlacionadas ao IAV (Elev e CPE), além do ITU, que foi comum para as duas profundidades.

Para a camada subsuperficial, a Elev e a CPE conferiram ao mapa final de predição a identificação de solos com cores mais vivas e avermelhadas nas áreas mais altas e com curvatura vertical convexa. O índice de avermelhamento mais baixo foi atribuído às superfícies mais côncavas e próximas à drenagem natural, onde se localizam os solos de cores mais acinzentadas nas áreas de baixadas. Apesar do IAV, proposto por Torrent e Barón (1993), ser concebido com finalidade de quantificação de hematita e/ou goethita na composição mineralógica do solo (e valores estabelecidos como limites para essa distinção), os valores foram considerados em toda sua extensão.

A correlação com o FLS na camada superficial indicou no mapa de predição a localização de solos com valores baixos de IAV na parte alta da área, próximo ao limite oeste da fazenda, em função dos valores mais elevados dessa covariável preditora, coincidindo com os Neossolos Litólicos. Tal fato não é comprovado em campo, uma vez que nessa região o material de origem é constituído pelo basalto, conferindo ao solo uma cor mais acentuada e avermelhada e não características de cores claras.

**Tabela 9 – Equações da RLM para IAV**

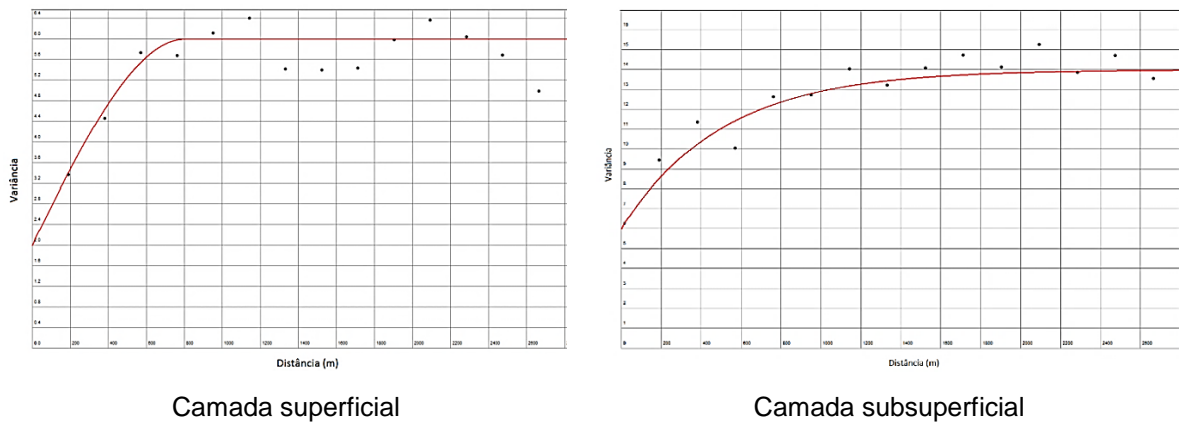
Atributo do solo	Modelo de regressão	R <sup>2</sup> ajustado <sup>1</sup>
IAV1	=9,8151-0,399*ITU-0,2782*FLS	0,13
IAV2	= 4,4930+0,01869*Elev +303,5291*CPE - 0,6958*ITU	0,36

<sup>1</sup> Modelo de regressão significativo a 5%.

IAV: índice de avermelhamento do solo; 1 = camada superficial; 2 = camada subsuperficial; ITU: índice topográfico de umidade, FLS: Fator LS, Elev: elevação, CPE: curvatura vertical.

Os resíduos da RLM foram submetidos à análise da dependência espacial e foi possível ajustar o modelo teórico ao variograma experimental, pelo método visual (a sentimento), observando principalmente o comportamento dos pontos na parte inicial do variograma e o número de pares para a definição do valor calculado da variância. Os modelos ajustados aos variogramas podem ser observados na Figura 10.

**Figura 10 – Variogramas dos resíduos da RLM do atributo IAV**



Para a camada superficial o modelo esférico foi ajustado (Tabela 10), com o efeito pepita igual a 2,0 e patamar igual a 6,0, próximo ao valor da variância dos dados, como é esperado de acordo com a teoria, uma vez que a RLM não contribuiu na explicação da variância, resultando em altos resíduos. Resultado semelhante foi ressaltado por Shi et al. (2011), para atributos químicos de solo.

O modelo exponencial foi ajustado para a camada subsuperficial. O efeito pepita resultou em um valor maior para essa profundidade, igual a 6,0, indicando que deve ocorrer variabilidade não determinada, a uma distância menor que o afastamento entre amostras. Pode ainda, ocorrer erros na determinação do índice, em função do caráter subjetivo na definição dos valores de matiz, valor e chroma, realizadas em campo, ou na atribuição dos valores ao “Banco de dados 2”. O valor do patamar igual a 14,0 indica que, para essa camada, parte da variância foi explicada na determinação da componente determinística, uma vez que a variância das amostras resultou em valor próximo a 22.

Os valores de IDE indicaram uma influência maior da variabilidade a curta escala para a camada subsuperficial, apesar de que para ambas as camadas a estrutura de dependência espacial é considerada moderada, segundo Cambardella et al. (1994). Os valores de IDE exprimem, portanto, uma deficiência da amostragem. Portanto, na camada mais profunda, a variabilidade nos valores do índice é maior, quando comparada à camada superficial.

**Tabela 10 – Modelos e parâmetros dos variogramas isotrópicos ajustados aos resíduos do ajuste de regressão linear para o atributo IAV**

Atributo do solo	Modelo	Efeito Pepita	Patamar	Alcance (m)	IDE
IAV1	Esférico	2,0	6,0	800	0,67
IAV2	Exponencial	6,0	14,0	900	0,57

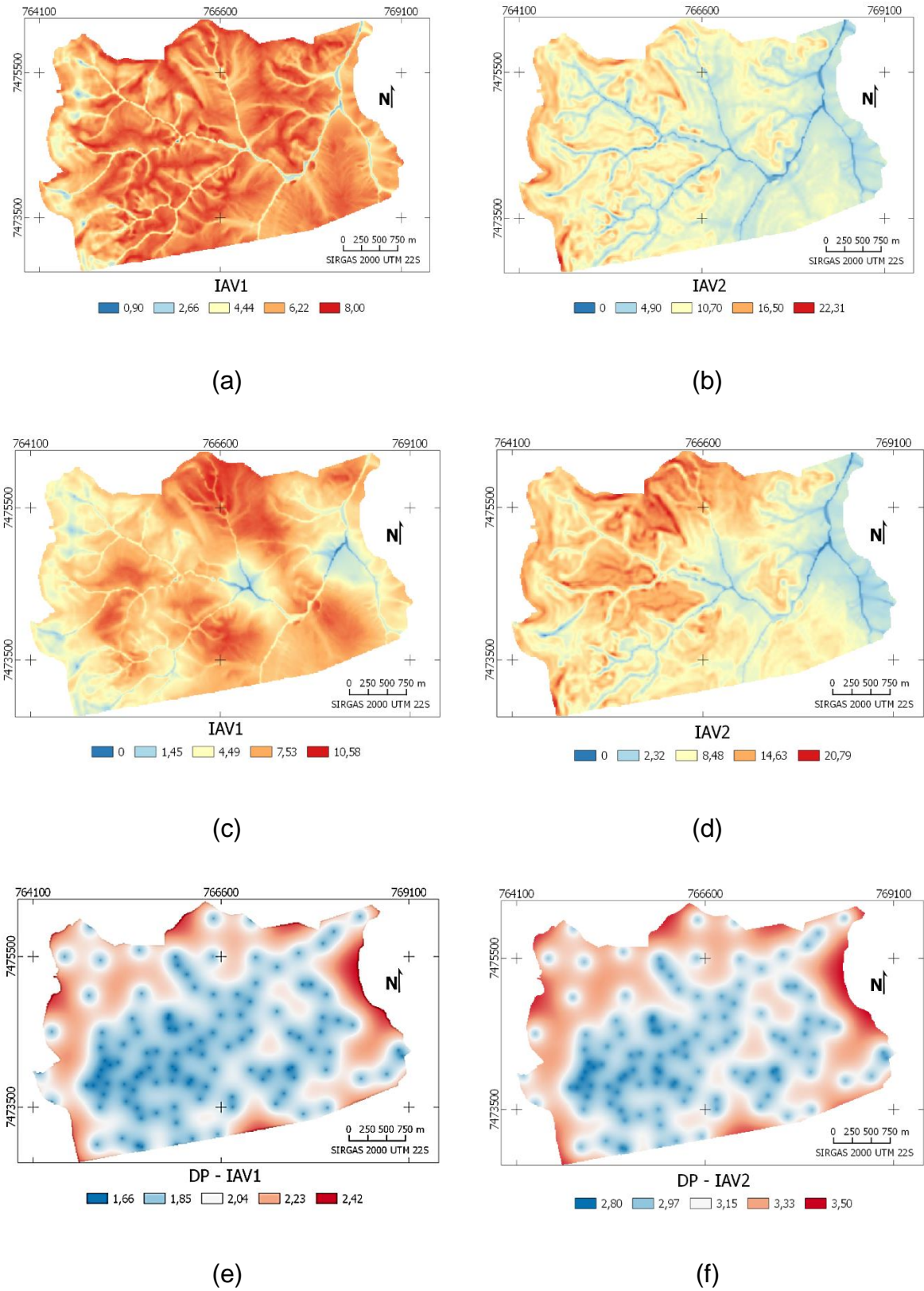
IAV: Índice de avermelhamento; 1: camada superficial; 2: camada subsuperficial.

Os valores de alcance para as duas camadas foram próximos, com a camada subsuperficial apresentando valor ligeiramente maior. Caso sejam realizadas futuras amostragens na área, ou em áreas similares, o espaçamento entre as amostras pode ser definido de acordo com os objetivos e baseado nos alcances encontrados (Webster e Oliver, 2007).

Os mapas resultantes da predição do IAV nas duas profundidades, nas duas etapas da modelagem (RLM e RK) estão apresentados na Figura 11.

A aplicação da técnica da RK foi avaliada como adequada para estimação do IAV na maior parte da área de estudo. Foram constatados valores mais elevados onde o material de origem é o basalto ou suítes básicas e a drenagem favorecida, condição essa preconizada por Curi e Franzmeier (1984), além da presença de vegetação mais densa, como observado por Santana et al. (2013). Observando a localização dos Latossolos no mapa de solos obtido por Carvalho et al. (1991), na Figura 4, verificou-se que a definição da estrutura de dependência espacial e posterior krigagem dos resíduos, para as duas camadas, tornou possível também identificar as áreas onde se localizam essas classes de solo. Essas classes apresentaram maiores valores de IAV, especialmente ao norte da área, onde ocorre o LVe. Essa predição foi mais evidente na camada superficial, para a qual o modelo do variograma apresentou menor valor de efeito pepita. Esse menor valor representa que os dados possuem uma estrutura espacial mais consistente, com menor aleatoriedade favorecendo a elaboração de mapas de krigagem mais confiáveis.

**Figura 11 – Mapas de previsão de IAV por RLM (a) e (b), por RK (c) e (d) e da qualidade da previsão final - DP (e) e (f), da camada superficial (1) e subsuperficial (2) da área total da Fazenda Edgárdia**



A presença de arenitos avermelhados da formação Pirambóia com agentes cimentantes de óxidos de ferro propiciaram a formação de Neossolos Quartzarênicos, de cores avermelhadas e, conseqüentemente, IAV mais elevados (não tanto quanto solos originários de basalto) na parte sul da área de estudo.

As áreas onde se localizam as classes Gleissolos e Neossolos Flúvicos foram relacionadas às regiões de valores mais baixos de IAV no mapa de predição onde ocorre menor presença de óxidos de ferro.

Os mapas representando a incerteza da predição do IAV por RK são apresentados na Figura 11 (e) e (f). Apesar da consideração a respeito da distribuição dos dados, o mapa de DP mostra a influência da distribuição dos pontos amostrais na incerteza da predição, exibindo um padrão semelhante para as duas camadas, mas com maior destaque para a camada subsuperficial. Como é próprio da krigagem, o valor do DP é mínimo nos pontos amostrais e também menor onde há maior concentração de amostras, o que significa que os valores preditos são mais precisos. O DP aumenta nas regiões onde as amostras estão mais dispersas, atingindo os valores máximos próximo ao limite da área. A região próxima às margens do mapa é onde geralmente ocorre maior incerteza de krigagem, mesmo em condições de amostragem regular, conforme Webster e Oliver (2007), uma vez que geralmente nesses locais é retirado número menor de amostras, ou pelo efeito de borda, como alerta Castrignanò (2011).

No processo de krigagem, a amostra com localização em regiões com amostragem deficiente recebe maior peso na interpolação, o que pode resultar em pior qualidade da predição para a área como um todo. Os maiores desvios nas extremidades da área de estudo podem refletir essa situação para os resíduos, como apontado por Webster e Oliver (2007). O mapa de desvio padrão obtido pode ser utilizado para planejamento de amostras adicionais ou ainda para avaliação da necessidade de maiores resoluções ou preditores mais detalhados nas áreas de alto erro de predição, como em Hengl et al. (2007b).

Foram observados valores extrapolando os limites do espaço-característica do atributo, principalmente para valores mínimos de IAV, localizados na extremidade sul e leste da área, na várzea do Rio Capivara. Nesses locais as predições de IAV resultaram em valores mínimos negativos, que não ocorrem na realidade. Os resultados da RK do pH, realizada por Sun et al. (2012), também apresentaram tal extrapolação, prevista e registrada por Hengl et al. (2004), que sugeriu a aplicação

de máscaras para eliminação de valores incoerentes. Tal fato foi relacionado ao efeito de convexidade do modelo de predição espacial por Hengl (2009). Os limites foram corrigidos (Tabela 11), e para que os valores fossem estimados corretamente, necessitaria de mais amostras, inclusive tomadas fora dos limites da área de estudo, o que não ocorreu em função da finalidade do levantamento, relativo à classificação dos solos.

**Tabela 11 – Estatísticas dos dados amostrais e dos modelos ajustados para o atributo IAV para a área total da Fazenda Edgárdia**

Atributo do solo	Máximo	Mínimo	Média	DP	CV (%)
IAV1	10,67	0,01	6,50	2,56	39,38
IAV1_RLM	8,01	0,88	6,37	0,84	13,19
IAV1_RK	10,58	0	6,34	1,76	27,76
IAV2	18,00	0,01	9,59	4,66	48,60
IAV2_RLM	22,31	0,93	9,69	2,73	28,17
IAV2_RK	20,79	0	9,76	3,72	38,11

IAV: Índice de avermelhamento; 1: camada superficial; 2: camada subsuperficial. RLM: modelo por regressão linear múltipla; RK: modelo por regressão krigagem; DP: desvio padrão; CV: coeficiente de variação.

De modo geral, o modelo obtido pela RK apresentou valores de estatísticas de IAV mais próximas das estatísticas dos dados amostrais, quando comparado ao modelo RLM, indicando que o tratamento dos resíduos pela geoestatística e posterior soma ao resultado da RLM proporcionou uma melhora relativa na predição.

A avaliação da predição do IAV por RK foi realizada pelos valores de ME e RMSE iguais a -0,08 e 1,89, na camada superficial e -0,05 e 3,38 na camada subsuperficial. Tais valores indicaram uma melhor representação do atributo na camada superficial.

O IAV não é um atributo comumente considerado em estudos de predição por MDS, carecendo de referências para comparação dos resultados obtidos. No entanto, para inferências em relação ao manejo, a cor do solo pode ser indicativa de características importantes, como por exemplo: comportamento químico em relação a elementos como o fósforo (ALMEIDA et al., 2003; PELUCO et al., 2015), óxidos de

Fe e Al, condição de drenagem e conteúdo de matéria orgânica (SANTANA et al., 2013) e comportamento em relação à erosão (DANTAS et al., 2014).

Brevik et al. (2016), abordando projeções para o futuro da modelagem pedológica, comentaram sobre o desenvolvimento de modelos que descrevam o comportamento do solo e auxiliem na classificação, o que é difícil, uma vez que os sistemas de classificação são baseados em características morfológicas e os modelos atuais não simulam características como, por exemplo, a cor e estrutura do solo. O IAV talvez possa ser um atributo a ser considerado para solos tropicais, especialmente no contexto da cuesta, representando a cor dos solos, principalmente quando baseado em métodos de obtenção de valores mais precisos, como a espectroscopia de reflectância difusa, apontada por Dantas et al. (2014).

#### **4.4.2 Índice de saturação por bases (V%)**

O modelo de predição adotado para a saturação por bases foi a RK, nas duas profundidades. Os coeficientes de correlação entre V% e as covariáveis foram de 0,20 a 0,34 e inicialmente foram obtidas as equações de RLM. Pela seleção stepwise, a covariável Decl foi a única selecionada para o modelo na camada superficial e Decl, Aspc, CPE e NDVI para a camada subsuperficial.

Os modelos RLM restringiram-se a uma parcela limitada da variabilidade do atributo V%, resultando em valores de  $R^2$  ajustado iguais a 0,11 e 0,32 para as camadas em profundidade, como mostrado na Tabela 12. Apesar dos valores baixos, foi estabelecida uma relação positiva entre V% e Decl, que pode ser atribuída à influência do material de origem que fornece cátions ao solo. Uma vez que solos derivados do basalto e diabásio são naturalmente mais ricos em nutrientes e, na Fazenda Edgárdia, estão presentes no domínio da Formação Serra Geral, localizando-se no front da cuesta de Botucatu, que apresenta declive acentuado (Figura 6), implicando em uma faixa de maiores valores de V% nos mapas resultantes da RLM, como mostrado na Figura 13 (a) e (b). Além da influência do material de origem, as áreas mais declivosas favorecem a ocorrência de solos mais rasos em função de menor oportunidade de infiltração de água no solo, proporcionando uma condição climática local que desfavorece o intemperismo e lixiviação, conforme foi apresentado por Dobos e Hengl (2009). Dessa forma, a

região do front da cuesta resultou em valores altos de V% para as duas camadas estudadas.

**Tabela 12 – Equações da RLM para V%**

Atributo do solo	Modelo de regressão	R <sup>2</sup> ajustado <sup>1</sup>
V1	=54,3320+0,65477*Decl	0,11
V2	=84,09740+0,84277*Decl -12,14835*senA - 70,66739*NDVI-8,67234* cosA- 1227,05684*CPE	0,32

<sup>1</sup> Modelo de regressão significativo a 5%.

V: índice de saturação de bases do solo; 1 = camada superficial; 2 = camada subsuperficial; Decl: declividade, A: aspecto, NDVI: índice de vegetação da diferença normalizada, CPE: curvatura vertical.

As demais localizações com valores mais altos de V% no interior da área de estudo, tem por material de origem as Suítes Básicas, mapeadas por Landim (1984), também coincidentes com regiões de maior declividade do terreno e, portanto, pela mesma razão apresentada acima, representam adequadamente os solos eutróficos.

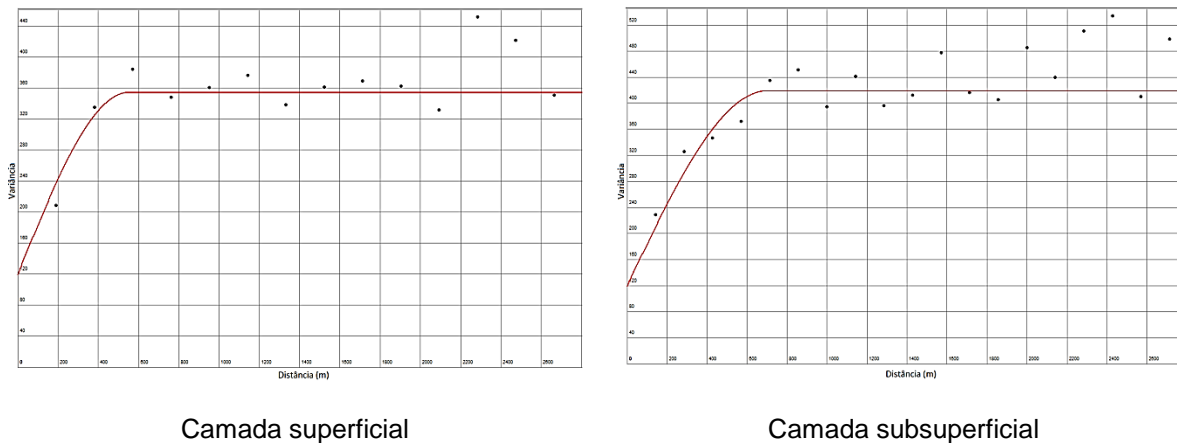
A correlação da covariável Aspc com o atributo V% pode representar a influência da direção da exposição da face do terreno que também atua no sentido de proporcionar um microclima local que afeta a taxa de intemperismo. A consideração do Aspc na predição de atributos de solo, tanto de V% como dos demais atributos para os quais essa covariável foi selecionada, conferiu aos mapas de predição mudanças bruscas, características de onde ocorreu alterações dos valores da covariável.

A inserção de mais variáveis ambientais no modelo RLM, para o V% na camada subsuperficial, promoveu aumento no R<sup>2</sup> ajustado em comparação com a camada superficial, embora para as duas camadas o conjunto de covariáveis selecionadas tenha sido limitado para explicar a variabilidade dos atributos, com valor máximo de R<sup>2</sup> observado de 32%. As covariáveis selecionadas para a equação da RLM na camada subsuperficial referem-se principalmente às derivadas do MDE, que descrevem o fator relevo, mas também o NDVI, que pode ser relacionado ao fator organismos, representado em termos de cobertura e uso da terra, como foi considerado por Malone et al. (2009). Sua inclusão no modelo proporcionou a

predição de valores mais elevados de V% para solos com horizonte B textural (Bt), mas com uso de pastagem (vegetação menos exuberante).

Os resíduos da RLM para V% apresentaram estrutura de dependência espacial nas duas profundidades estudadas. Os modelos teóricos ajustados aos variogramas nas duas camadas foram esféricos e podem ser observados na Figura 12 e seus parâmetros na Tabela 13. Em ambos variogramas, os valores de efeito pepita foram moderados, indicando considerável aleatoriedade nos valores dos resíduos, resultantes do modelo da RLM. Outros erros de amostragem e localização podem ter sido computados nos valores dos resíduos.

**Figura 12 – Variogramas dos resíduos da RLM do atributo V%**



Apesar dos valores de efeito pepita serem iguais e o comportamento em relação à estrutura de dependência espacial também parecida para as duas camadas, um valor de alcance ligeiramente maior foi observado para a camada subsuperficial, indicando que acima de 550 m e 700 m, para a camada superficial e subsuperficial, respectivamente, a variação foi aleatória.

**Tabela 13 – Modelos e parâmetros dos variogramas isotrópicos ajustados aos resíduos do ajuste de regressão linear para a saturação por bases**

Atributo do solo	Modelo	Efeito Pepita	Patamar	Alcance	IDE
V1	Esférico	120	355	550	0,66
V2	Esférico	120	410	700	0,71

V: Saturação por bases em %, 1: camada superficial, 2: camada subsuperficial.

Para as duas camadas estudadas, o modelo final de predição obtido pela RK evidencia as áreas de maior V% em espaços ocupados principalmente por Neossolos Litólicos e alguns solos com horizonte Bt (Figura 13).

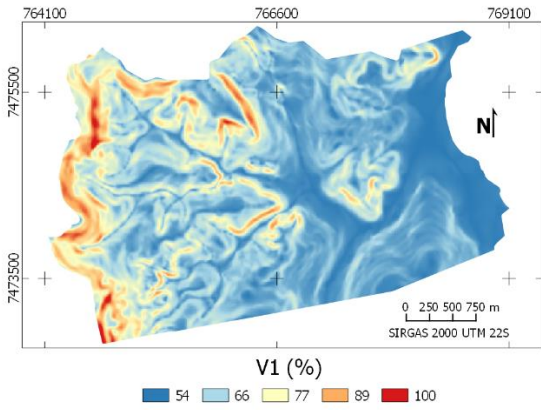
A consideração da covariável Aspc na RLM, associada aos valores de baixa declividade, definiu uma faixa com baixa saturação de bases, coincidente com as classes Neossolo Flúvico Eutrófico e Gleissolo Háptico Eutrófico, no mapa de predição. Tal faixa persistiu no mapa final da RK, apesar de ter sofrido uma suavização com a aplicação da krigagem nos resíduos da RLM. Nessa região, pelo mapa de solos de Carvalho et al. (1991), é registrado caráter eutrófico nos solos ali presentes e, pelo mapa de desvio padrão, essa é a região de maior incerteza da modelagem e, portanto, onde ocorreram maiores erros e ausência de amostras. No entanto, essa covariável foi determinante na localização da classe Neossolos Quartzarênicos, ao sul-sudeste da área.

O modelo possibilitou o mapeamento de áreas com valores mais baixos de V% onde estão localizados os Latossolos, no centro-norte da área, que apesar da litologia nessa região ser constituída principalmente de material básico, a posição no relevo pode ter favorecido o processo de intemperismo, resultando em solos mais pobres.

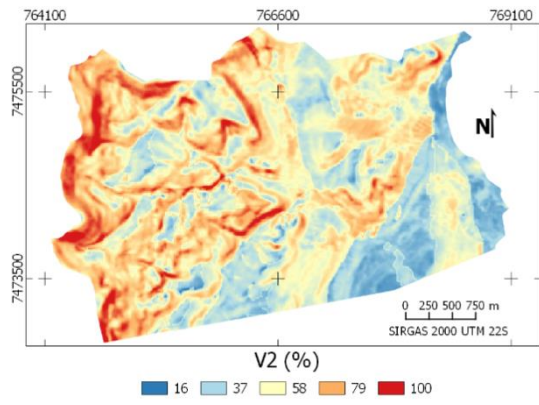
Nas áreas de maior incerteza de predição, a amostragem ficou restrita a poucos pontos, possivelmente por dificuldade de coleta e em função do objetivo, ou seja, o levantamento de solos para a exploração agrícola. De acordo com as classes de solo, a ocorrência de valores mais altos de V% na parte mais plana da fazenda, abrangendo áreas de baixada, principalmente próximo aos cursos d'água, pode ser relacionada à deposição de material advindo da parte alta (colúviação), via processo erosivo, do planalto ocidental ou do front da cuesta, onde o material de origem é mais rico em bases.

As áreas compreendendo as classes de Latossolos distróficos e Neossolos Quartzarênicos foram devidamente representadas no mapa final de predição da RK, com valores estimados de  $V% < 50\%$ .

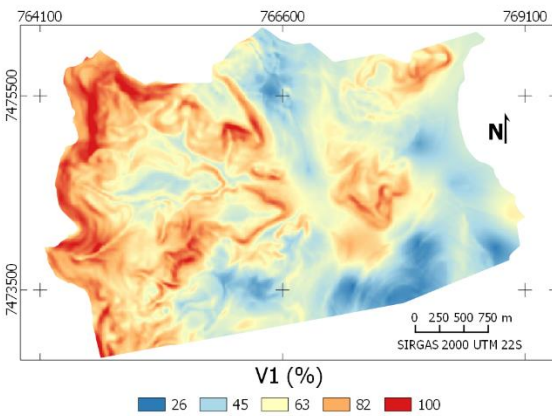
**Figura 13 – Mapas de predição de V% por RLM (a) e (b), por RK (c) e (d), e da qualidade da predição – DP (e) e (f), da camada superficial (1) e subsuperficial (2) da área total da Fazenda Edgárdia**



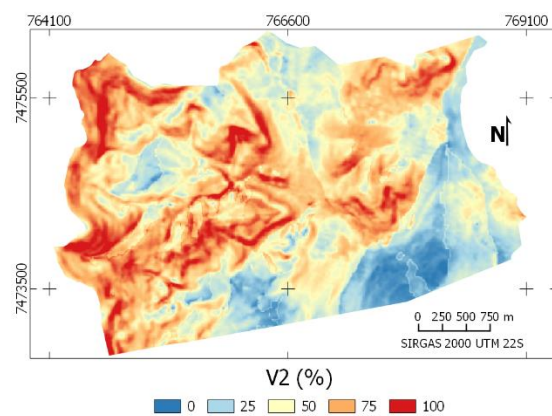
(a)



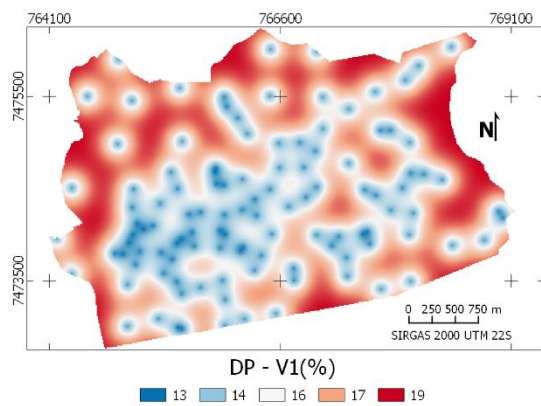
(b)



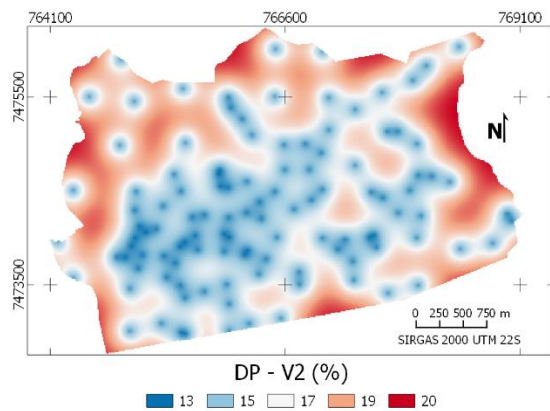
(c)



(d)



(e)



(f)

Em relação à amplitude de variação dos valores preditos de V%, pode ser observada maior variação na camada subsuperficial, principalmente menor valor de V% mínimo. Observando as áreas de ocorrência de valores mínimos de V%, a maior proximidade dos arenitos constituintes da Formação Pirambóia deve influenciar tais valores mais baixos.

Semelhante ao resultado obtido para o IAV, o DP foi mais acentuado próximo aos limites da área, revelando padrão notadamente dependente da ausência de amostragem. Foram observados valores próximos de DP para as duas camadas, ligeiramente mais altos para a camada subsuperficial.

A acurácia de predição do modelo foi limitada, de acordo com os respectivos valores de ME e RMSE, iguais a -1,28% e 16,13% para a camada superficial e -3,50% e 21,50% para a subsuperficial. De modo geral, a maior variabilidade e incerteza da predição em profundidade podem ser associadas à maior proximidade com o material de origem inalterado, uma vez que o material mais superficial é geralmente mais intemperizado e, portanto, mais uniforme. Nos modelos finais de predição, pela RK, os valores mínimos de V% (principalmente nas áreas de Neossolos Quartzarênicos) foram mais baixos na camada subsuperficial. O conteúdo maior de matéria orgânica na superfície, pelo uso da terra com recomposição de mata, ou até mesmo na pastagem, pode ter influenciado esse resultado, proporcionando aumento de bases, uma vez que o material de origem constituído de arenitos tem naturalmente disponibilidade restrita desses elementos.

Os valores médios estimados de V% foram próximos entre si e também próximos aos observados para os dados de treinamento, com a ocorrência de valor mínimo negativo e valor máximo maior que 100, na camada subsuperficial, extrapolando os dados originais, corrigidos para os valores extremos (HENGL et al., 2004). O valor médio de V% para as duas camadas pela RK resultou em torno de 63%, com variação menor que os dados originais (Tabela 14).

Apesar de constituir um dos atributos de menor permanência no solo, dentre os considerados nesse estudo, o que reduz a possibilidade de seu uso em planejamentos a prazos mais longos, o V% tem grande importância no manejo do solo, determinando a necessidade de investimento em insumos.

**Tabela 14 – Estatísticas dos dados amostrais e dos modelos ajustados para o atributo V% para a área total da Fazenda Edgárdia**

Atributo do solo	Máximo	Mínimo	Média	DP	CV (%)
V1	98,00	11,00	63,40	20,68	32,61
V1_RLM	100,00	54,34	77,00	8,69	11,28
V1_RK	100,00	26,20	63,10	14,83	23,45
V2	95,00	5,00	59,64	26,68	44,74
V2_RLM	100,00	16,00	58,00	17,51	30,19
V2_RK	100,00	26,00	63,00	23,13	36,71

V: Saturação por bases em %; 1: camada superficial; 2: camada subsuperficial. RLM: modelo por regressão linear múltipla; RK: modelo por regressão krigagem; DP: desvio padrão; CV: coeficiente de variação.

#### 4.4.3 Areia

Uma vez que as condições locais permitem inferir uma influência significativa da dinâmica de processos erosivos na pedogênese na área da cuesta, a expectativa de que as frações texturais apresentassem correlação com o relevo foi confirmada, apesar desses valores serem considerados baixos.

O teor médio de areia na área da Fazenda Edgárdia pode ser considerado alto, com alta variabilidade ao longo da área, constituindo uma das principais características que refletem a atuação dos fatores de formação na pedogênese local. Apesar de serem aplicadas apenas variáveis relativas ao relevo nesse estudo, a íntima relação do relevo com a litologia e o contexto geomorfológico são determinantes na ocorrência das classes de solo na paisagem e nas variações das frações texturais mapeadas.

A posição no relevo, expressa na RLM pela covariável Elev, associada ao basalto na região do front da cuesta, resultou na presença de material de textura mais fina e, portanto, teores mais baixos ou nulos de areia, com teores crescentes na direção das partes baixas do terreno e nas regiões onde material de origem subjacente é constituído basicamente de arenitos. A possibilidade de ocorrência de arenitos intercalados com as suítes básicas também deve ser considerada, de acordo com a estratigrafia local e descrito em IPT (1981), o que resultou em

Latossolos de textura média, mapeados ao norte da área da fazenda por Carvalho et al. (1991), semelhante aos observados por Arruda et al. (2013) na região de Barra Bonita-SP, com estratigrafia similar.

No entanto, considerando as baixas correlações da areia, pressupõe-se ser difícil aplicar o mapeamento quantitativo baseado em relações solo-paisagem pela técnica de RLM e RK, como foi observado por Carré e Girard (2002).

Na Tabela 15 são apresentadas as equações da RLM para o atributo areia, onde pode-se observar que a covariável Elev foi selecionada para o modelo de predição da areia nas duas camadas, sendo o único preditor da camada superficial. Adotando um nível de significância de 5% para o modelo, o valor do R<sup>2</sup> ajustado para essa equação foi de 0,14, indicando uma limitada capacidade de predição, ou seja, baixo valor da componente determinística nessa camada. Na camada subsuperficial, foram selecionadas 4 covariáveis preditoras, mas ainda assim, o valor do coeficiente de determinação resultou em 0,24, também considerado baixo.

**Tabela 15 – Equações da RLM para Areia**

Atributo do solo	Modelo de regressão	R <sup>2</sup> ajustado <sup>1</sup>
Ar1	=1611,0365-1,783864*Elev	0,14
Ar2	=840,11327-1,07088*Elev+722,34109*NDVI-30,78343*FLS-76,33318*senA	0,24

<sup>1</sup> Modelo de regressão significativo a 5%.

Ar: Teor de areia do solo; 1 = camada superficial; 2 = camada subsuperficial; Elev: elevação, NDVI: índice de vegetação da diferença normalizada, FLS: Fator LS, A: aspecto.

Correlações significativas entre NDVI e frações texturais do solo foram registradas anteriormente, embora seja mais comum sua correlação com a fração argila (OMUTO; VARGAS, 2015). Dobarco et al. (2016) encontraram que o uso da terra foi uma variável importante para modelagem da fração areia em região central da França. Gomez et al. (2012) verificaram que a areia não pôde ser acuradamente predita a partir de dados hiperespectrais.

A correlação do teor de areia com a covariável FLS pode ser atribuída a regiões dissecadas, como é o caso da Fazenda Edgárdia, uma vez que caracteriza os processos de erosão e deposição na formação do solo. Assim também foi considerado no estudo de Moore et al. (1993), que estabeleceu uma analogia entre

FLS e o índice de transporte de sedimentos, sendo o primeiro aplicável de forma bidimensional.

Quanto ao Aspc, não é clara a base física de correlação entre essa covariável e o atributo Ar, no entanto a covariável foi mantida na equação.

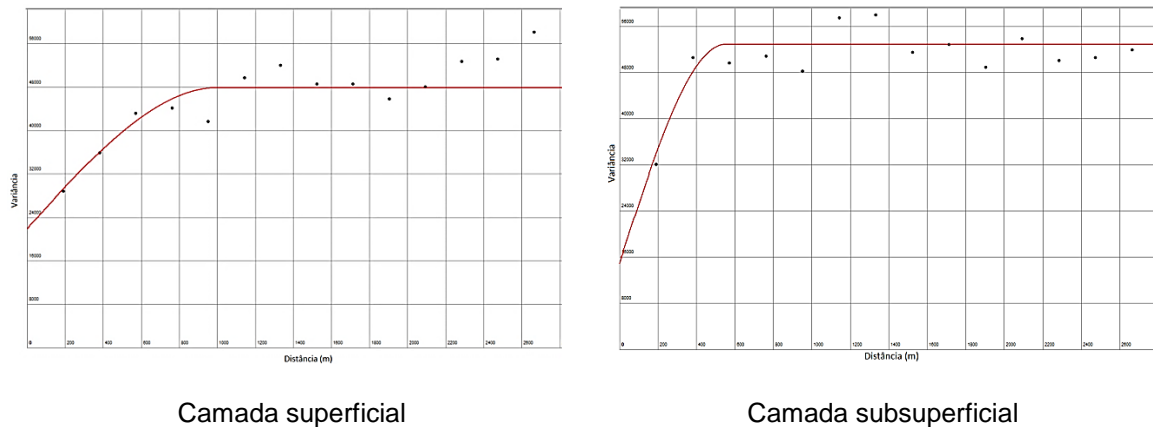
A parte da variabilidade da Ar detectada pela RLM foi discreta, considerando os valores do coeficiente de determinação, embora sejam significativos. Deve-se levar em consideração que na camada superficial, somente a Elev como covariável abordou 14% da variabilidade da Ar, o que pode ser considerado um valor expressivo, em função dos valores baixos de  $R^2$  encontrados de modo geral para as RLM em estudos de solos.

A estimação de frações texturais tem sido realizada por modelos de MDS, uma vez que são atributos de interesse primário em classificação e planejamento de uso e manejo das terras, determinando o comportamento físico, químico e hídrico dos solos. Ciampalini et al. (2012) encontraram correlação significativa do teor de areia apenas com a declividade nas camadas superficiais de solos derivados de arenitos e margas, em uma área de relevo variável na Tunísia. Nas camadas mais profundas, além da covariável elevação, índices morfométricos e a coordenada na direção X também foram correlacionadas.

A análise variográfica dos resíduos de Ar obtidos na RLM, resultou na adequação do modelo esférico para ambas camadas, conforme Figura 14. Os parâmetros dos modelos são apresentados na Tabela 16.

O ajuste do variograma teórico atribuiu valores de patamar inferiores aos valores de variância dos dados, de 66270,20 e 72689,55 respectivamente, para as duas profundidades, indicando que parte da variância foi atribuída à tendência definida pela RLM, como apontado por Hengl et al. (2007b). O IDE, no entanto, também indica que para ambas as camadas, parte da variância é aleatória ou não foi detectada, por estar relacionada à variação de curta escala. Considerando que não ocorreram erros no processo de obtenção dos dados, outras covariáveis relativas aos fatores de formação dos solos, como por exemplo a litologia, podem ser suficientes para melhorar a estimação da fração areia na área da Fazenda Edgárdia.

**Figura 14 – Variogramas dos resíduos da RLM do atributo Areia**



A estrutura de dependência espacial mais forte dos resíduos na camada subsuperficial resultou em menores valores de ME, principalmente no resultado final da RK, sendo os valores para as camadas 1 e 2 iguais a 21,04 g.kg<sup>-1</sup> e -3,34 g.kg<sup>-1</sup> com valores de RMSE iguais a 156 g.kg<sup>-1</sup> e 136 g.kg<sup>-1</sup>, respectivamente. Os valores de ME e RMSE, utilizando dados de validação, são considerados altos, implicando em baixo desempenho dos modelos.

**Tabela 16 – Modelos e parâmetros dos variogramas isotrópicos ajustados aos resíduos do ajuste de regressão linear para o atributo Areia**

Atributo do solo	Modelo	Efeito Pepita	Patamar	Alcance	IDE
Ar1	Esférico	22000	48000	1000	0,54
Ar2	Esférico	15000	53000	550	0,72

Ar: Teor de areia em g.kg<sup>-1</sup>; 1: camada superficial, 2: camada subsuperficial.

Baseado em valores de validação cruzada, Carvalho Júnior et al. (2014a) também consideraram pobre o desempenho obtido pelos modelos RK, KO e RLM para predição de atributos de solo, dentre eles a areia. No entanto, os autores encontraram resultado levemente superior na camada superficial, além de que, os modelos que utilizaram a krigagem foram sempre superiores a RLM, em todos os casos.

Umali et al. (2012) abordaram a fração grosseira do solo e na validação encontraram resultados de erro médio absoluto e RMSE bem próximos para os modelos de regressão e RK e valores ligeiramente inferiores para a KO.

Na Figura 15 está exposta a variação do teor de areia na área da Fazenda Edgárdia, estimado por RLM e RK, nas duas camadas pesquisadas.

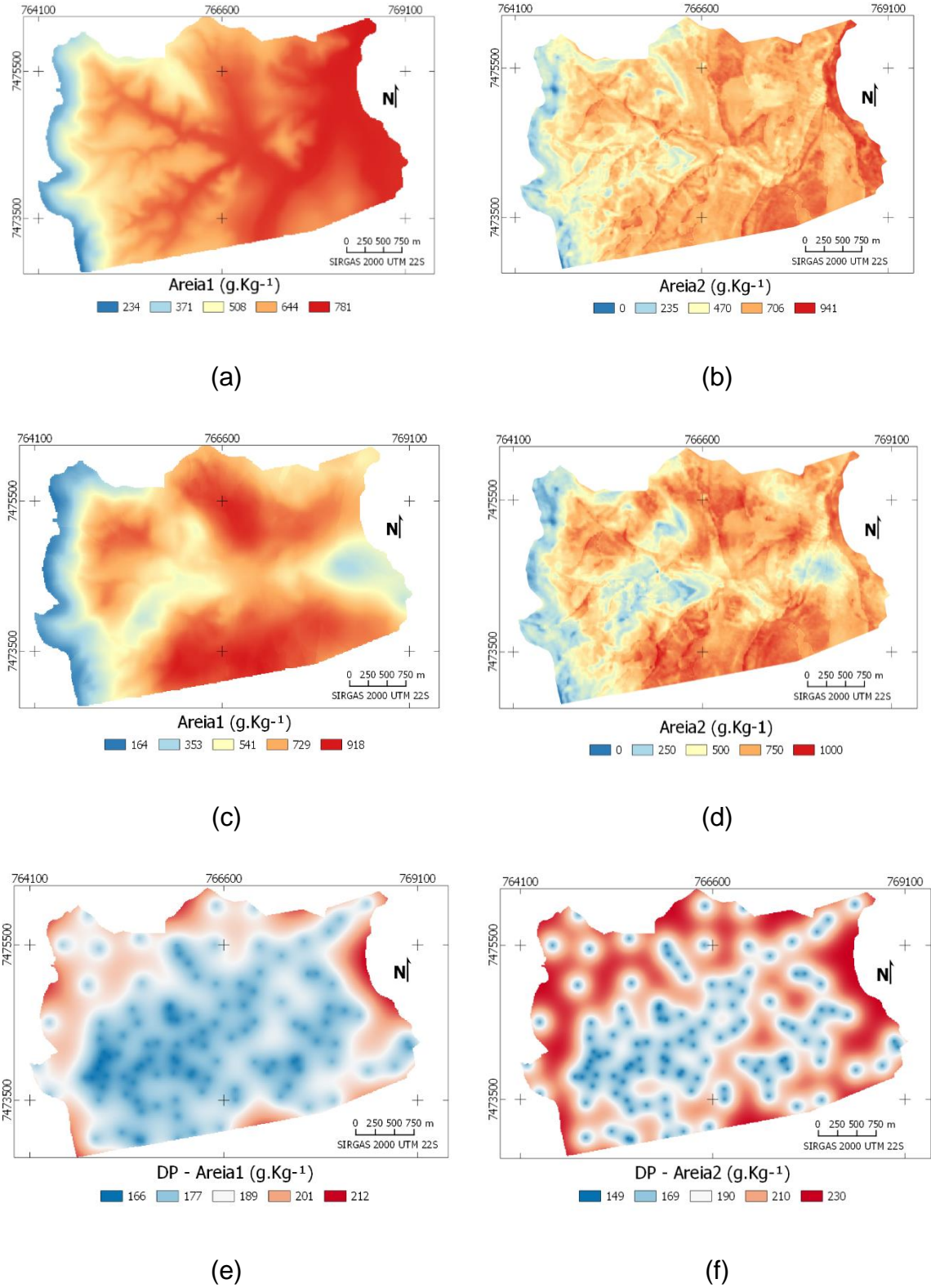
É possível observar o predomínio de áreas com maior teor de areia no leste e sul da área (Figura 15), bem como uma faixa de menor teor de areia no limite oeste da fazenda, região do front da cuesta. Esse padrão é idêntico e perceptível, para os quatro mapas apresentados, mas com maior evidência quando a predição foi realizada pela RLM. A adição da componente estocástica de curta escala ao modelo permitiu a distinção de regiões com menores teores de areia, ao longo da direção Leste-Oeste da área, sendo na porção mais a oeste coincidente com a ocorrência de solos com horizonte Bt.

O padrão geral da distribuição da areia representado no mapa de predição por RK é semelhante ao obtido por Silva et al. (2014), para a mesma área, onde foi utilizada a simulação para representar esse atributo como um dos critérios para uma nova abordagem de classificação de solos. Os mapas das frações texturais areia e argila apresentados no referido estudo indicam as áreas de mais altos e baixos teores concordantes com os obtidos pela modelagem por RK aqui obtidos.

Segundo exposto por Oliver e Webster (2014), o grau de suavização na krigagem depende da relação efeito pepita/patamar, com maiores valores da divisão acarretando maior suavização, subestimando valores mais altos e superestimando valores mais baixos. Essa consideração pode explicar a redução na variação dos valores estimados comparados aos observados, pelos valores de CV na Tabela 17. De qualquer modo, a tendência de ligeiro aumento na variação da Ar em profundidade, obtido nos dados originais, foi mantida nos valores estimados.

Houve extrapolação dos valores em relação aos limites do espaço-característica do atributo, como para outros atributos desse estudo, sendo aplicada correção de acordo com recomendação de Hengl et al. (2004).

**Figura 15 – Mapas de predição da Areia por RLM (a) e (b), por RK (c) e (d), e da qualidade da predição – DP (e) e (f), da camada superficial (1) e subsuperficial (2) da área total da Fazenda Edgárdia**



**Tabela 17 – Estatísticas dos dados amostrais e dos modelos ajustados para o atributo Areia para a área total da Fazenda Edgárdia**

Atributo do solo	Máximo	Mínimo	Média	DP	CV (%)
Ar1	950,00	10,00	648,46	257,43	39,70
Ar1_RLM	780,86	234,72	649,42	110,38	17,01
Ar1_RK	917,73	164,14	657,85	173,00	26,30
Ar2	940,00	20,00	591,00	269,61	45,62
Ar2_RLM	941,00	0	610,70	148,55	24,32
Ar2_RK	1000,00	0	624,88	191,55	30,65

Ar: Teor de areia em g.kg<sup>-1</sup>; 1: camada superficial; 2: camada subsuperficial. RLM: modelo por regressão linear múltipla; RK: modelo por regressão krigagem; DP: desvio padrão; CV: coeficiente de variação.

#### 4.4.4 Argila

Parte da variabilidade do teor de argila na Fazenda Edgárdia foi atribuída às covariáveis Elev e NDVI, na camada superficial e Aspc, ITU e NDVI na subsuperficial, com valores baixos de R<sup>2</sup> ajustado para as RLM (Tabela 18). As correlações com Elev e ITU são representantes do fator relevo, embora nessa região, como foi abordado para o atributo Ar, também foram indiretamente relacionadas ao material de origem. Os Neossolos Litólicos constituem a classe de solos derivados do basalto e demais materiais básicos, com conseqüente textura mais argilosa e, na paisagem, se localizam principalmente no front da cuesta e outras partes mais elevadas e declivosas. Alguns dos solos com Bt, também, localizam-se em partes médias da encosta.

Além disso, os processos erosivos que atuam em função do relevo íngreme no contexto geomorfológico da cuesta, propiciam maior remoção das partes altas e deposição nas baixadas, como foi evidenciado nos mapas de atributo areia. Esse processo pode acarretar acúmulo de argila devido à litologia da área proporcionar a formação de argila em relação às frações de maiores dimensões.

Em predições do atributo argila, utilizando-se covariáveis ambientais, comumente são verificadas correlações com variáveis relacionadas ao relevo e

NDVI. Correlações entre o atributo Arg e os preditores MDE e NDVI foram obtidos também por Omuto e Vargas (2015), além da coordenada X na camada superficial (0-20 cm) de um solo em região do Quênia. No entanto, em comparação ao presente estudo, foram encontrados valores mais elevados de coeficiente de correlação, acima de 0,5.

Na pesquisa de Sun et al. (2012) em um distrito irrigado na Austrália, a Arg também apresentou correlação com a Elev e o NDVI, dentre outras covariáveis, para a camada de 0-10 cm e na camada de 40-50 cm de profundidade, o ITU e o NDVI, associados a outras covariáveis. Carvalho Júnior et al. (2014a) encontraram correlações entre argila e Aspc, NDVI e DEM com valores de 0,13 a 0,28, nas camadas de 30 a 200 cm de profundidade para região montanhosa do estado do Rio de Janeiro.

Em um segmento da paisagem, Umali et al. (2012) encontraram correlações moderadas e fracas entre atributos do solo e parâmetros do terreno, numa amplitude maior (de 0,01 a 0,69) do que a observada nesse estudo. A pesquisa foi de abrangência local, com maior detalhamento e o teor de argila foi relacionado às variáveis elevação, curvaturas, área de contribuição, transporte de sedimentos e índice de umidade, na região entre linhas em um plantio de maçãs.

**Tabela 18 – Equações da RLM para Argila**

Atributo do solo	Modelo de regressão	R <sup>2</sup> ajustado <sup>1</sup>
Arg1	= -41,2446+0,6252*Elev-216,1715*NDVI	0,10
Arg2	= 675,34749+69,984784*senA-15,97858*ITU-603,70901*NDVI	0,26

<sup>1</sup> Modelo de regressão significativo a 5%.

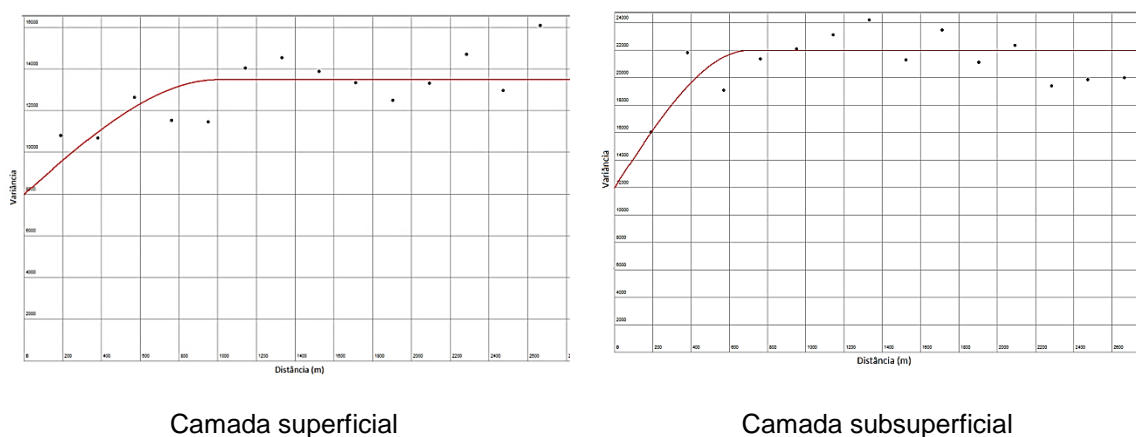
Arg: Teor de argila do solo; 1 = camada superficial; 2 = camada subsuperficial; Elev: elevação, NDVI: índice de vegetação da diferença normalizada, A: aspecto, ITU: índice topográfico de umidade.

Para os variogramas dos resíduos das RLM foram ajustados modelos esféricos nas duas camadas (Figura 16 e Tabela 19), que resultaram em valores de efeito pepita e IDE indicativos de variabilidade não detectada e uma moderada estrutura de dependência espacial (CAMBARDELLA et al., 1994). Essa variação a curta escala, da mesma forma como foi comentado para a Ar, pode ser atribuída às condições do material de origem, naturalmente heterogêneo em função da estratigrafia regional. Valores elevados de C<sub>0</sub> propiciam maiores variâncias na

interpolação e afetam a precisão, como destacado por Castrignanò (2011), resultando em mapas menos suavizados. No entanto, aumentado-se o nível de detalhamento da amostragem, mais informação seria obtida sobre a variabilidade a curta escala, fator esse considerado crítico para estudos dos atributos de solo na Fazenda Edgárdia.

Um menor valor de alcance foi obtido para a camada subsuperficial, comportamento semelhante ao observado para a Ar, definindo menor região de variabilidade estruturada em profundidade, embora tenha que ser considerado, como observado por Oliver e Webster (2014), que o valor desse parâmetro foi obtido por meio subjetivo.

**Figura 16 – Variogramas dos resíduos da RLM do atributo Argila**



Os valores do patamar foram menores que a variância dos dados observados, uma vez que uma componente determinística pôde ser estabelecida em relação a variáveis auxiliares representativas do relevo, apesar dos reduzidos valores de  $R^2$  ajustado.

**Tabela 19 – Modelos e parâmetros dos variogramas isotrópicos ajustados aos resíduos do ajuste de regressão linear para o atributo Argila**

Atributo do solo	Modelo	Efeito Pepita	Patamar	Alcance	IDE
Arg1	Esférico	8000	13500	1000	0,41
Arg2	Esférico	12000	22000	700	0,45

Arg: Teor de argila em  $g.kg^{-1}$ ; 1: camada superficial, 2: camada subsuperficial.

Na Figura 17 é possível observar que os valores de argila de modo geral foram superiores em profundidade, tanto na RLM quanto no modelo final de RK. Em alguns locais, principalmente na porção centro-oeste da área, pode refletir o processo de eluviação da argila, na formação dos solos Bt.

Pelos mapas de predição gerados pode ser inferida a relação da argila com o material de origem. A consideração da Elev na RLM da camada 1 imprimiu um destaque para a faixa de altitudes elevadas no front da cuesta, relacionada à litologia, como comentado anteriormente. Mudanças bruscas perceptíveis nos mapas resultaram da influência da covariável NDVI para a camada superficial e, principalmente, do Aspc na camada subsuperficial.

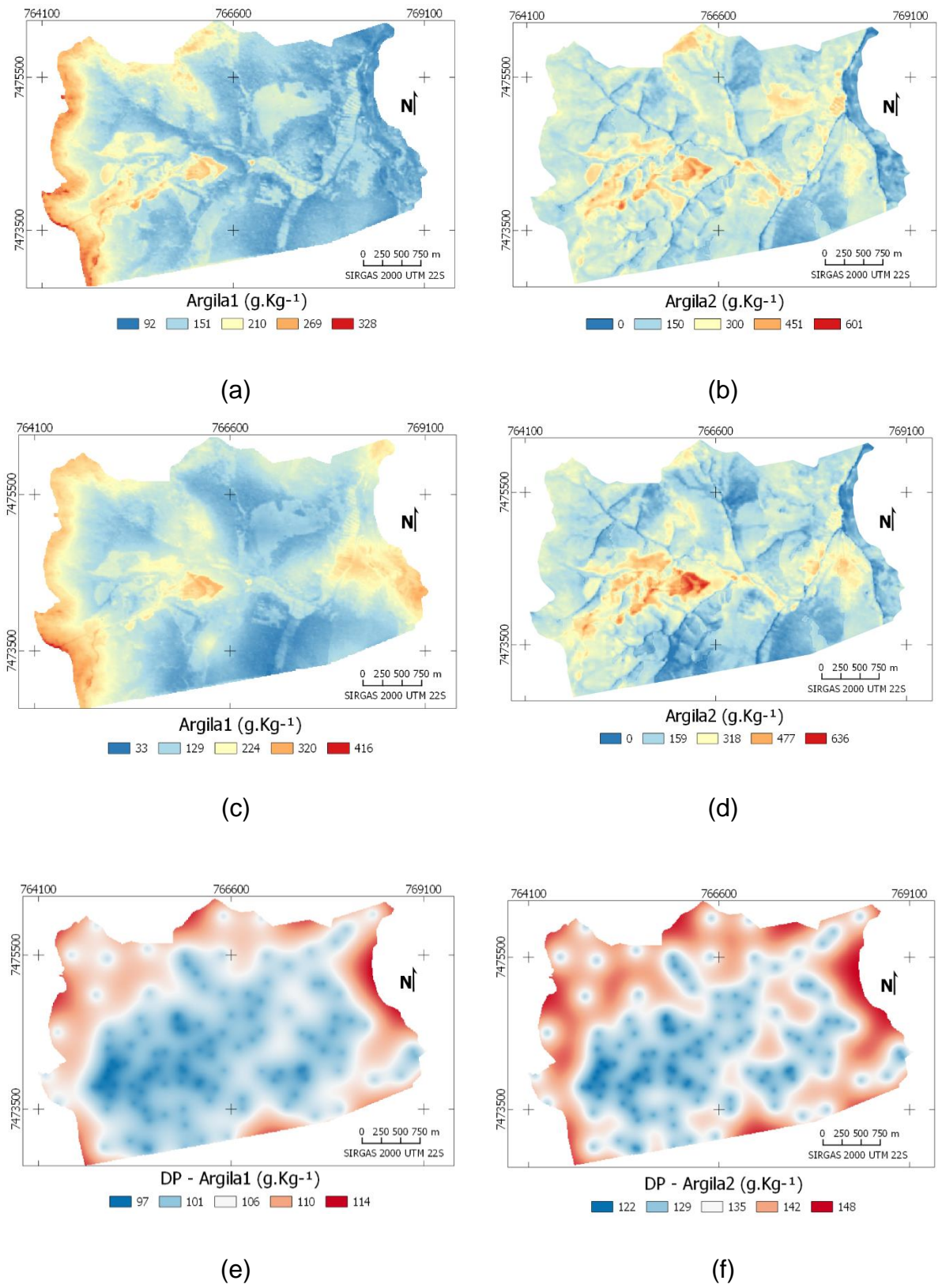
Os valores mais elevados de DP são localizados na região do front da cuesta, em áreas de maior declividade e várzea do Rio Capivara e, como comentado para os demais atributos, podem ser relacionados ao efeito de borda e espaçamento maior das amostras. A incerteza da predição também foi maior na camada subsuperficial.

Para a avaliação da acurácia da predição, com base nos dados de validação, os valores de ME na camada 1 foram aproximadamente de 4 g.kg<sup>-1</sup> para RLM e 2g.kg<sup>-1</sup> para o resultado final da RK e os valores do RMSE foram 140 g.kg<sup>-1</sup> e 93 g.kg<sup>-1</sup>. Para a camada subsuperficial, foram obtidos aproximadamente 30 g.kg<sup>-1</sup> e 38 g.kg<sup>-1</sup> para ME e 133 g.kg<sup>-1</sup> e 113 g.kg<sup>-1</sup> para RMSE, para as etapas de RLM e a soma dos resíduos na RK, respectivamente.

Observou-se, portanto, menores erros para Arg na camada superficial e ligeiro incremento da acurácia entre a RLM e o resultado final da RK, embora os valores desse índice indicaram que ambos os modelos não tiveram desempenho satisfatório.

Vários resultados na literatura indicam dificuldade de acurada predição da fração argila. Os diferentes modelos aplicados por Omuto e Vargas (2015), para elaborar os mapas de estimação da argila, apresentaram desempenho distinto, mas ambos não resultaram acurados pelas estatísticas da validação cruzada.

**Figura 17 – Mapas de predição da Argila por RLM (a) e (b), por RK (c) e (d), e da qualidade da predição – DP (e) e (f), da camada superficial (1) e subsuperficial (2) da área total da Fazenda Edgárdia**



Resultados apontando um pobre desempenho na predição de argila foram também alcançados por Sun et al. (2012), aplicando RK global e local e também co-krigagem adotando covariáveis de relevo e radiométricas e valores de argila obtidos por espectroscopia de reflectância no infravermelho. Os valores de RMSE variaram de 16 a 24 aproximadamente e, como observado para a Fazenda Edgárdia, os maiores valores referiram-se à camada mais profunda, no caso de 40-50 cm, em comparação à camada 0-10 cm. O resultado insatisfatório foi atribuído pelo autor à qualidade dos dados e baixa relação entre o conteúdo de argila e as covariáveis ambientais.

Também Omuto e Vargas (2015) obtiveram baixa acurácia em predição semelhante à deste estudo, quando utilizaram dados de argila de diferentes fontes e covariáveis de chuva, relevo, uso da terra, paisagem, geologia e NDVI para predição de argila na camada de 0-20 cm. Os autores aplicaram o modelo de efeitos mistos, separando grupos de solos, com objetivo de aumentar o desempenho da RK. O resultado desfavorável foi atribuído ao uso de dados secundários, como no presente estudo.

Apesar de obter resultados de baixo desempenho na predição da argila, pela validação cruzada, Samuel-Rosa et al. (2015) observaram que tanto para argila quanto para CTC a etapa da krigagem dos resíduos melhora a acurácia da predição em relação à RLM. Tal fato pôde ser constatado também para a estimativa da argila na Fazenda Edgárdia.

Os valores das estatísticas básicas entre valores observados e estimados estão na Tabela 20. Foram aplicadas correções para eliminar valores mínimos espúrios, que extrapolaram o espaço-característica do atributo. Os modelos promoveram redução da amplitude dos teores de argila em relação aos dados originais, como é próprio da krigagem.

Os altos valores preditos de argila nas áreas de Argissolos, Chernossolos e Nitossolos mapeados por Carvalho et al. (1991), atualmente com uso agrícola e pastagens, e nesse último caso, referente ao centro de pesquisa de bubalinos, revelam a necessidade de manejo adequado para evitar degradação do solo. Nas demais áreas com solos argilosos, as condições de topografia dificultam a utilização da terra com atividades agrícolas. A manutenção da cobertura com mata nativa ou em regeneração auxilia a conservação do solo.

**Tabela 20 – Estatísticas dos dados amostrais e dos modelos ajustados para o atributo Argila para a área total da Fazenda Edgárdia**

Atributo do solo	Máximo	Mínimo	Média	DP	CV (%)
Arg1	600,00	10,00	170,50	132,07	77,46
Arg1_RLM	327,65	91,57	162,37	39,13	24,10
Arg1_RK	415,53	33,03	161,94	60,43	37,32
Arg2	660,00	10,00	235,63	167,71	71,18
Arg2_RLM	600,99	0	213,08	75,55	35,50
Arg2_RK	636,24	0	207,97	92,75	44,60

Arg: Teor de argila em  $g.kg^{-1}$ ; 1: camada superficial; 2: camada subsuperficial. RLM: modelo por regressão linear múltipla; RK: modelo por regressão krigagem; DP: desvio padrão; CV: coeficiente de variação.

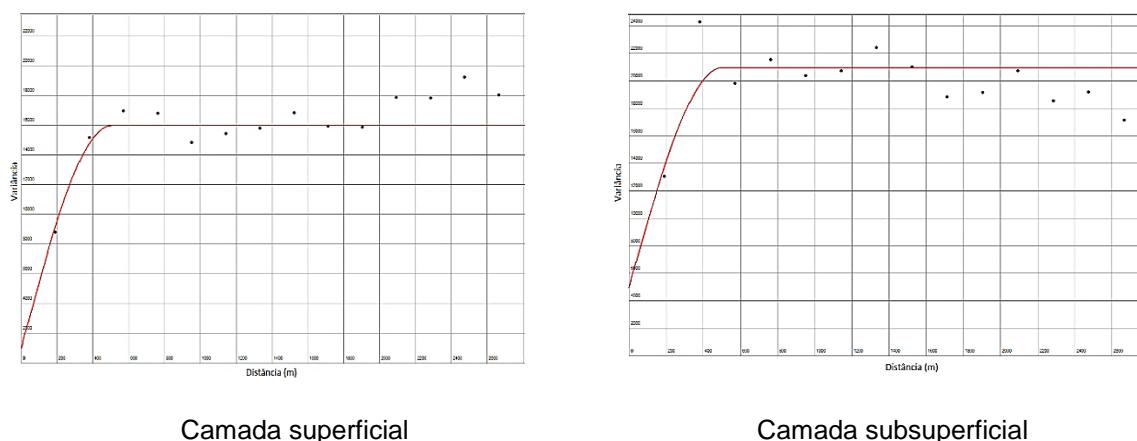
#### 4.4.5 Capacidade de troca de cátions (CTC)

A estimação dos valores da CTC do solo foi realizada pela aplicação da KO, uma vez esse atributo não apresentou correlação com as covariáveis ambientais consideradas. Resultado distinto foi registrado por Samuel-Rosa et al. (2015) para a CTC da camada superficial de solos predominantemente rasos, no sul do Brasil, para os quais houve correlação da CTC com as covariáveis relativas a mapas de solos, geologia, uso da terra, imagem de satélite e DEM. Os modelos pesquisados pelos autores para estimar a CTC do solo foram RLM e modelo misto linear, este último com performance ligeiramente superior.

Ciampalini et al. (2012) verificaram correlações significativas para a CTC e covariáveis ambientais relacionadas à geomorfologia e paisagem, nas profundidades correspondentes às consideradas no presente estudo, além de determinar estrutura espacial da variância desse atributo, o que propiciou adotar a RK como modelo de predição.

Os variogramas experimentais para os dados da CTC do solo exibiram forma similar para as duas camadas, aos quais foram ajustados o mesmo modelo teórico esférico (Tabela 21). Nos variogramas, um possível comportamento distinto parece se impor a longa escala, entre as duas camadas (Figura 18). No entanto, os valores de alcance foram próximos, implicando que amostras afastadas mais de 500 m nas duas camadas são consideradas independentes.

**Figura 18 – Variogramas do atributo CTC**



As razões IDE foram distintas para as camadas superficial e subsuperficial. Na primeira camada essa razão é maior, indicando que são menores as variações devido a fontes de erro aleatórias, relacionadas à amostragem, às determinações analíticas dos atributos e outros erros não detectados. As duas camadas apresentaram valor de patamar inferiores à variância das amostras de treinamento, o que significa que a variância populacional deve ser maior que a variância percebida por meio das amostras e parte da variância não foi captada pelo modelo.

**Tabela 21 – Modelos e parâmetros dos variogramas isotrópicos ajustados aos dados do atributo CTC**

Atributo do solo	Modelo	Efeito Pepita	Patamar	Alcance	IDE
CTC1	Esférico	1000	15900	520	0,94
CTC2	Esférico	5000	21000	500	0,76

CTC: Capacidade de troca de cátions em  $\text{mmolc.dm}^{-3}$ ; 1: camada superficial; 2: camada subsuperficial.

Os mapas de predição para as duas camadas mostram um padrão semelhante de comportamento da CTC para as duas camadas estudadas, com manchas localizadas de valores mais altos próximo ao alinhamento leste-oeste, na parte central da área (Figura 19). Valores baixos de CTC são bem definidos para as áreas onde se localizam os Neossolos Quartzarênicos, ao sul da fazenda, e os Latossolos. Na primeira classe de solo, a baixa superfície específica da fração areia é predominante e, no segundo caso, de solos mais intemperizados, a remoção das

bases são as causas determinantes da baixa ocorrência das bases, em relação aos solos mais argilosos derivados do basalto e dos solos com horizonte Bt, que apresentam maior CTC.

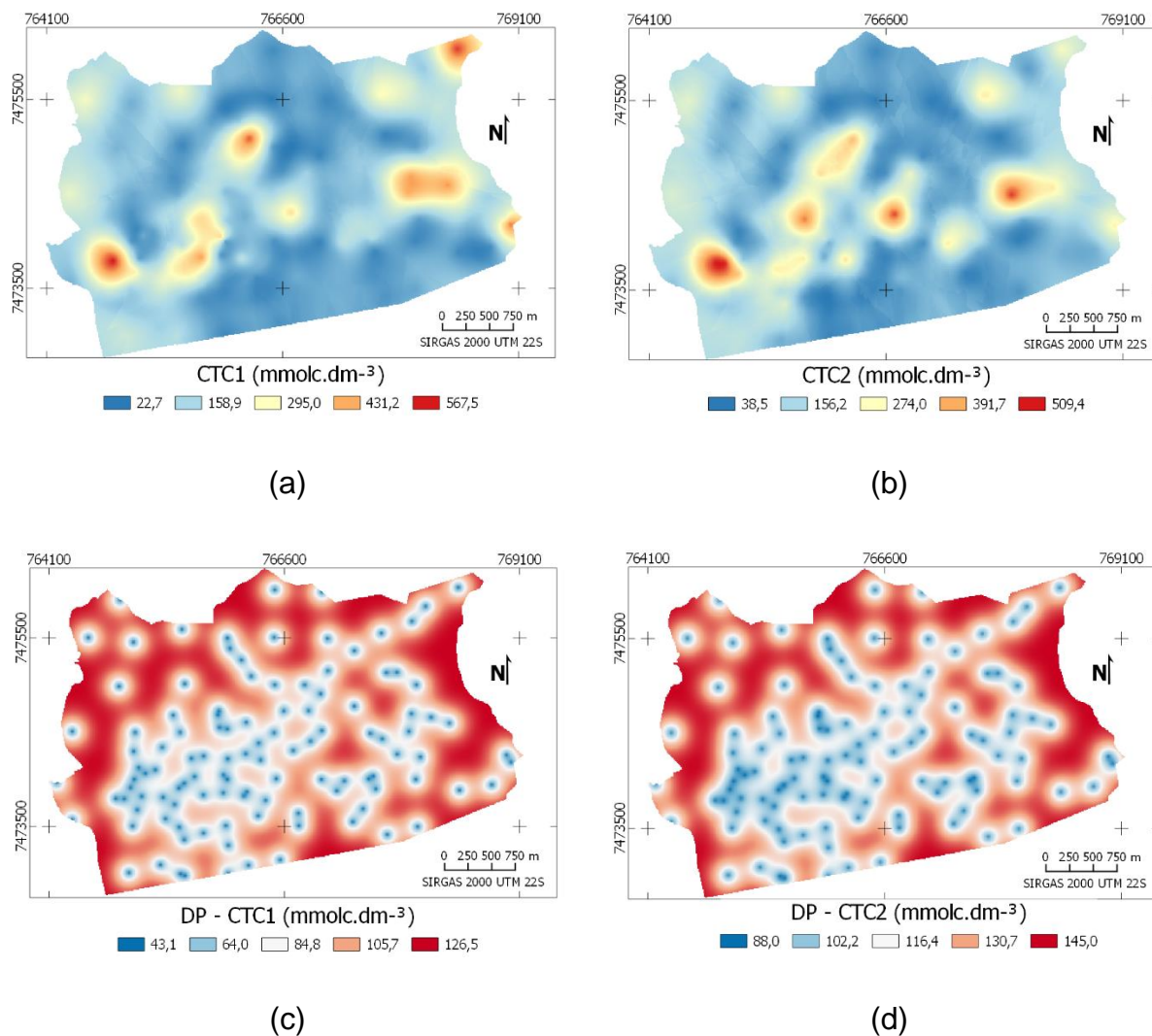
Nos mapas ocorreu alguma formação dos denominados “olhos de boi”, característicos do processo de interpolação, indicativos de locais onde o espaçamento das amostras supera o alcance da estrutura de dependência espacial. De modo geral os mapas exibiram aspecto suavizado.

Os mapas de desvio padrão da predição da CTC do solo apresentaram um padrão único para as duas camadas estudadas, com menores valores nos locais amostrados e na região onde houve maior concentração de amostras. Ao contrário, houve maior discrepância nas partes altas do front da cuesta e nas áreas de várzea. Mesmo no ponto onde foi localizada a amostra de solo a interpolação pela krigagem promoveu uma alteração do valor, não honrando o valor observado e acarretando um erro, que pode ser considerado alto, principalmente na camada subsuperficial.

Os valores observados de ME e RMSE foram iguais a -10,09 e 116,42  $\text{mmolc.dm}^{-3}$  e -9,890 e 114,42  $\text{mmolc.dm}^{-3}$ , para a camada superficial e subsuperficial respectivamente, considerados próximos entre si, indicando desempenho semelhante da modelagem nas duas camadas. No entanto, esses valores são indicadores de baixa capacidade de predição do modelo de estimação e foram similares aos encontrados por Samuel-Rosa et al. (2015) para esse mesmo atributo do solo, na camada superficial.

Os valores de CTC na região dos Latossolos, no extremo norte da área são equiparáveis aos observados por Arruda et al. (2013) no mesmo contexto de material de origem e relevo no município de Barra Bonita-SP.

**Figura 19 – Mapas de predição da CTC por KO (a) e (b) e da qualidade da predição – DP (c) e (d) da camada superficial (1) e subsuperficial (2) da área total da Fazenda Edgárdia.**



A caracterização estatística dos valores estimados pela KO em relação aos valores amostrais está apresentada na Tabela 22. As discrepâncias observadas são similares para as duas camadas estudadas, com estimativas de valores médios inferiores aos observados.

**Tabela 22 – Estatísticas dos dados amostrais e dos modelos ajustados para o atributo CTC para a área total da Fazenda Edgárdia**

Atributo do solo	Máximo	Mínimo	Média	DP	CV (%)
CTC1	602,00	20,0	165,76	132,10	86,00
CTC1_KO	567,50	22,7	153,50	142,56	92,87
CTC2	607,00	15,0	166,88	167,70	86,67
CTC2_KO	509,40	38,5	153,20	144,63	94,40

CTC: Capacidade de troca de cátions em  $\text{mmolc.dm}^{-3}$ ; 1: camada superficial; 2: camada subsuperficial. KO: krigagem ordinária; DP: desvio padrão; CV: coeficiente de variação.

#### 4.4.6 pH do solo

Com o resultado encontrado para o atributo pH, que não apresentou correlações com as covariáveis e também não apresentaram estrutura de dependência espacial, a média seria considerada a melhor representação do atributo ao longo da área, conforme previsto por Ciampalini et al. (2012). No entanto, baseando-se em experiências anteriores no local, sabe-se que os valores de pH variam na extensão da fazenda e, mesmo que a amostragem apresente limitações, a amplitude de variação dos dados amostrais evidencia essa realidade para o atributo em questão. Buscou-se então uma representação mais realista, sendo realizada a interpolação pelo método do inverso da distância – IDW, uma vez que esse modelo mecânico de predição espacial envolve uma ponderação baseada nas distâncias entre as amostras (HENGL, 2009).

A ausência de correlação de pH e indicadores geomorfométricos e de paisagem, além do NDVI, foi verificada por Ciampalini et al. (2012), em estudo regional, utilizando dados legados de solos oriundos de conjuntos esparsos de amostras de levantamentos prévios. A fraca variabilidade do atributo na área de estudo foi presumida como uma das causas da não correlação do pH com a paisagem.

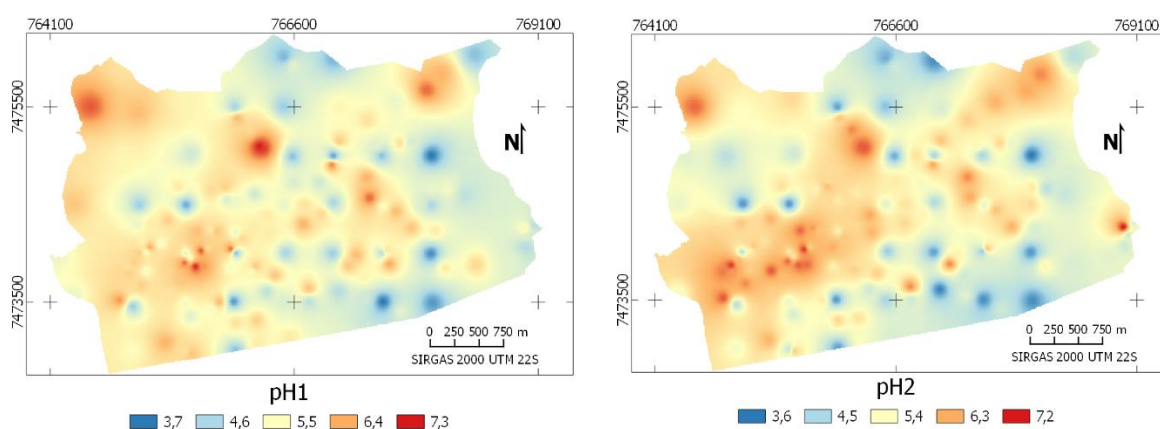
Carvalho Junior et al. (2014) também verificaram não haver correlação do pH com covariáveis do ambiente em profundidades de 30 a 100 cm, e adotaram o valor médio como a melhor representação desse atributo para essas camadas, em ambiente de relevo montanhoso no estado do Rio de Janeiro.

Quando foi observada correlação do pH com os preditores, os valores foram considerados baixos, como constatado por Hengl et al. (2004) ou, ainda, na

pesquisa de Mosleh et al. (2016), com coeficiente de correlação  $<0,10$  entre o pH e o índice de fluxo (*stream power index*) e curvatura horizontal da superfície. Nesse último caso, as condições eram de relevo plano em uma região árida do Irã. Shi et al. (2011) aplicaram RK para dados de pH obtendo boa eficiência de predição, utilizando um modelo linear generalizado (GLM), com base teórica específica para definir a componente determinística e considerando a covariável uso do solo.

Os mapas de predição do pH nas duas camadas são apresentados na Figura 20.

**Figura 20 – Mapas de predição do pH por IDW, da área total da Fazenda Edgárdia na camada superficial (1) e subsuperficial (2)**



É possível observar nos mapas um padrão geral similar da variação do pH nas profundidades estudadas e praticamente a mesma amplitude de valores, com variação nas camadas superficial e subsuperficial de 3,7 a 7,3 e 3,6 a 7,2. O padrão de variação tem relação com os mapas estimados para as frações texturais, com pH mais alto predominando nas áreas com maior teor de argila e valores mais baixos onde ocorrem solos de textura média ou arenosos.

A validação da predição do pH resultou em valores de EM iguais a -0,0006 e -0,0007 e RMSE iguais a 0,003 e 0,002, para as camadas superficial e subsuperficial, respectivamente. Os valores próximos a zero indicam predição mais precisa.

As estatísticas dos dados estimados indicaram valores médios estimados próximos aos observados, sendo o pH médio do perfil igual a 5,4, porém com menor variabilidade para as estimativas, nas duas profundidades analisadas (Tabela 23).

**Tabela 23 – Estatísticas dos dados amostrais e dos modelos ajustados para o atributo pH para a área total da Fazenda Edgárdia**

Atributo do solo	Máximo	Mínimo	Média	DP	CV (%)
pH1	7,30	3,70	5,53	0,74	13,47
pH1_IDW	7,29	3,70	5,49	0,41	7,47
pH2	7,20	3,60	5,50	0,83	15,05
pH2_IDW	7,19	3,60	5,41	0,47	8,69

1: camada superficial; 2: camada subsuperficial. IDW: inverso da distância ponderada; DP: desvio padrão; CV: coeficiente de variação.

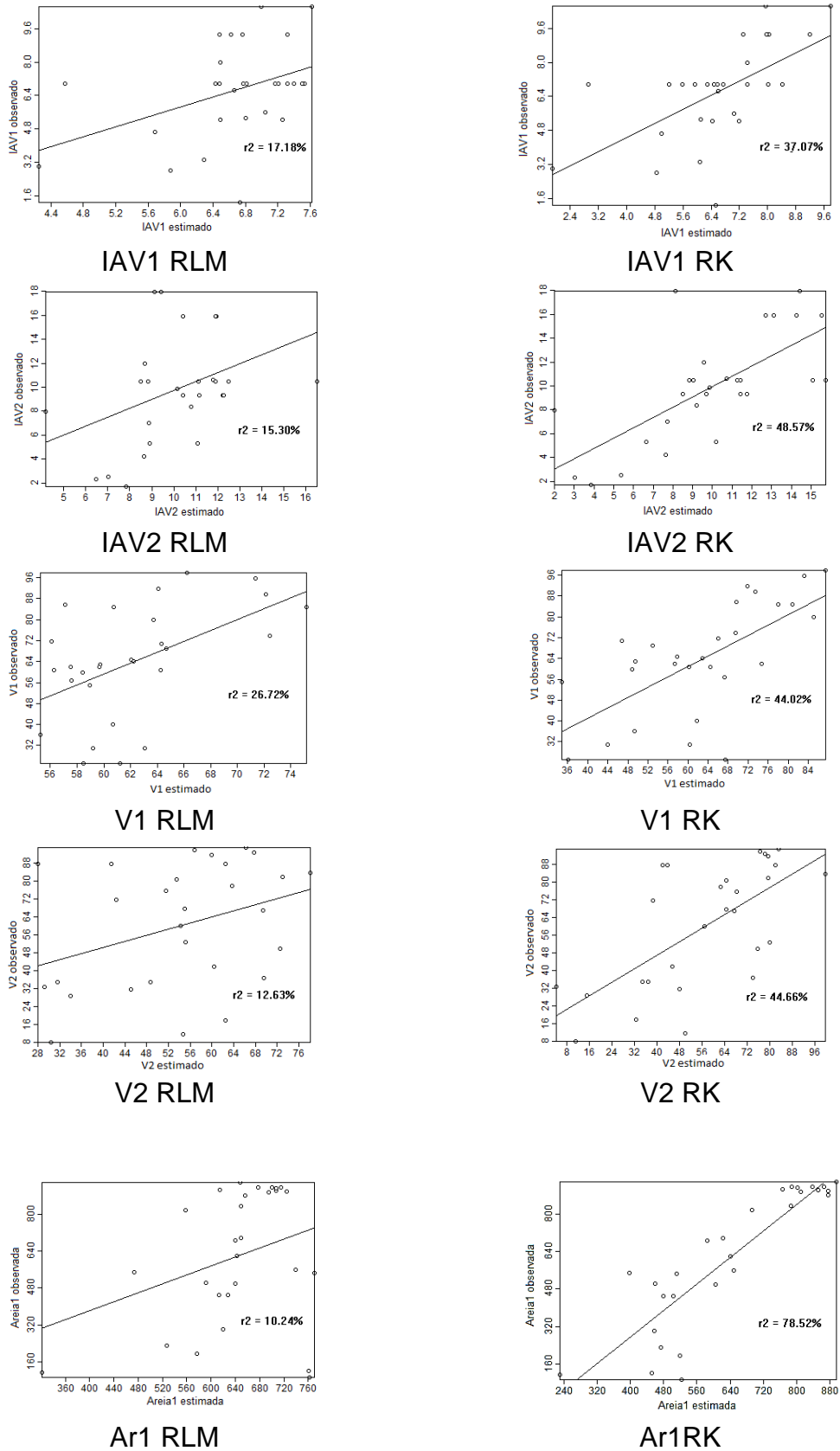
#### 4.5 Representação da validação dos modelos

Segundo Landim (2003), a verificação mais eficiente de um modelo obtido é resultante do confronto de dados estimados e a realidade de campo. Para tal avaliação foram utilizados os dados separados para validação e calculados os valores de ME e RMSE, apresentados na discussão de cada atributo.

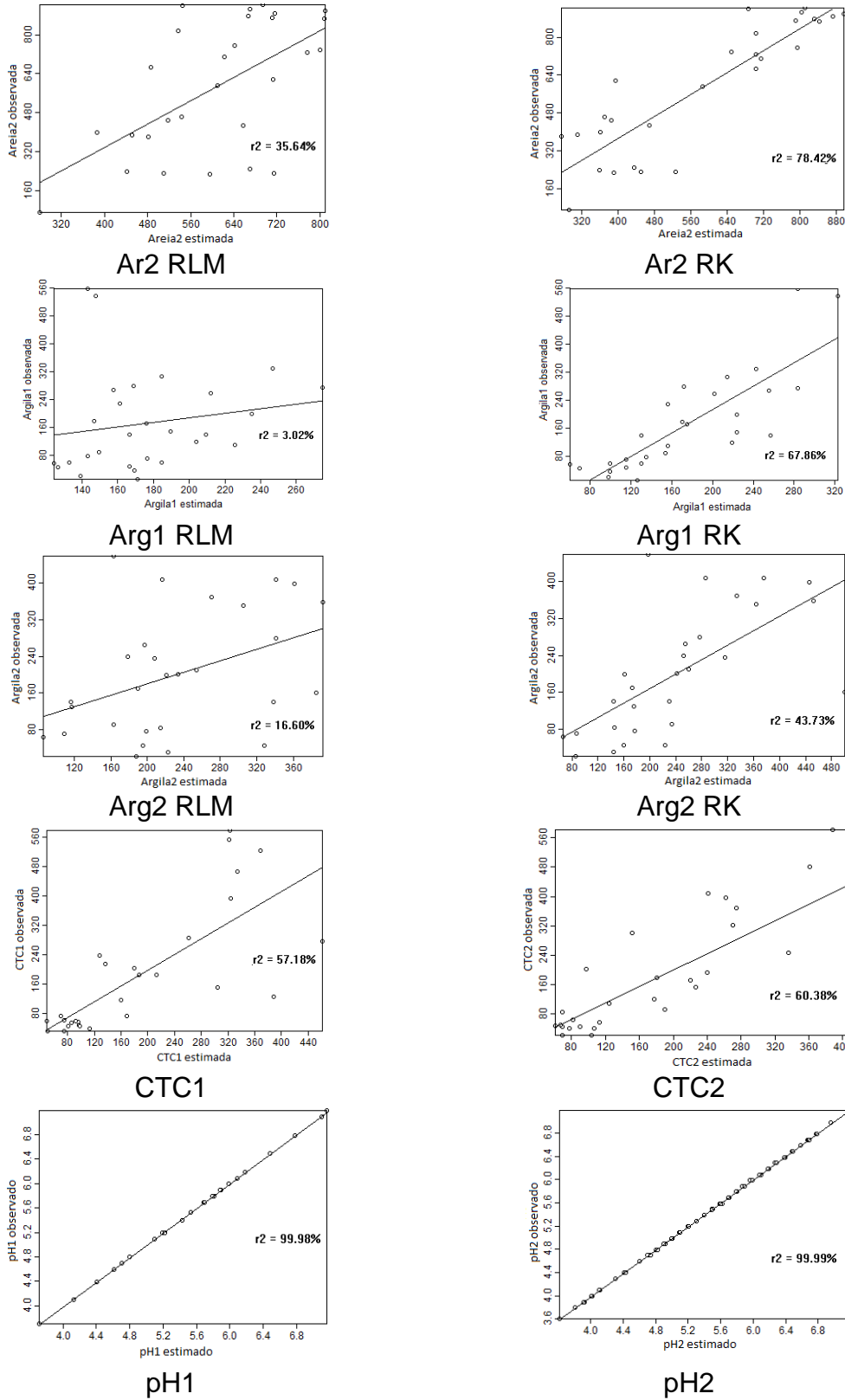
Para visualização dos resultados obtidos, foram elaborados os gráficos das Figuras 21 e 22. Essa apresentação é indicada por Webster e Oliver (2007), como uma forma de avaliar o comportamento da krigagem, ressaltando que os valores deveriam ser os mesmos, mas que esse fato não ocorre na natureza. Carvalho Junior et al. (2014a) adotaram essa comparação para avaliar a predição pelos métodos de RLM, KO e cokrigagem, baseado no valor do  $R^2$  da validação cruzada.

Pelos gráficos, o modelo final de RK apresentou desempenho mais satisfatório para a Arg em condições de superfície, IAV em condições de subsuperfície e V%, Areia, CTC e pH tiveram resultados próximos entre as duas camadas. Para as condições tropicais, Carvalho Junior et al. (2014a) encontraram desempenho semelhante para modelos de estimação em camadas superficial (0-30 cm) e subsuperficial (30-200 cm) para os atributos areia e argila.

**Figura 21 – Representação gráfica da comparação entre valores estimados e observados nos pontos de validação para os atributos IAV, V e Ar do solo da Fazenda Edgárdia. 1: camada superficial; 2: camada subsuperficial.**



**Figura 22 – Representação gráfica da comparação entre valores estimados e observados nos pontos de validação para os atributos Ar, Arg, CTC e pH do solo da Fazenda Edgárdia. 1: camada superficial; 2: camada subsuperficial**



Pode-se observar ainda o incremento no desempenho da estimação quando é aplicada a krigagem nos resíduos da RLM e posterior cômputo do resultado na inferência. A aplicação da RLM teve valores de  $R^2$  sempre baixos, em função das baixas correlações encontradas. A incorporação dos resíduos no processo de predição teve efeito positivo, evidenciando a superioridade da RK em relação à RLM nessas condições.

Observou-se que a alta variabilidade dos dados (e possível ocorrência de *outliers*) nas condições reais, não permitiu um  $R^2$  de 100% entre os valores observados e estimados. A avaliação da similaridade entre os dados observados e estimados pode ser baseada no coeficiente de concordância de Lin, como sugerido por Malone et al. (2009).

#### **4.6 Considerações sobre a modelagem dos atributos de solo para a Fazenda Edgárdia**

A aplicação do modelo híbrido RK em estudos de solos é baseada na distinção e descrição da variabilidade em duas partes: uma parte determinística, relacionada a variáveis ambientais explicativas dos fatores de formação do solo e uma parte aleatória, em curta escala, passível de ser descrita pela estrutura de dependência espacial dos resíduos da tendência (HENGL et al., 2007a, b). Uma vez estabelecidas essas relações, são estimados valores de atributos de solo em pontos onde não foram amostrados, possibilitando a elaboração de um mapa contínuo do atributo em questão.

A hipótese básica desse estudo foi de que o relevo seria o principal fator de formação dos solos na região da cuesta, onde se localiza a Fazenda Edgárdia, uma vez que processos relacionados a deslocamentos de água, partículas sólidas, elementos químicos e sedimentos certamente ocorreram na pedogênese local. No entanto, de modo geral, para a Fazenda Edgárdia, as correlações observadas entre as covariáveis de relevo e os atributos de solo foram baixas, implicando em modelos de equações de regressão linear com limitada capacidade de representar a variabilidade dos atributos, no máximo de 36%. Porém, Carvalho Junior et al (2014) consideraram comum, na literatura, a ocorrência de valores baixos de coeficientes de determinação para esse tipo de modelo estatístico. Os mesmos autores enfatizaram esse fato para regiões declivosas, em ambiente tropical.

A seleção das covariáveis preditoras de forma automática e global pode ter causado a não detecção das correlações atributo-preditores, principalmente pela possibilidade de variações localizadas dessas relações. Como observado por Sun et al. (2012), se ocorrem tais variações locais, RK local deve apresentar melhor resultado, mas seriam necessárias muitas amostras para julgar a estacionaridade da relação entre observações e covariáveis, para decidir entre RK global ou local. De qualquer forma, mais denso conjunto amostral e correlações fortes possibilitariam melhor resultado.

As equações de regressão linear foram aplicadas para tentar representar os processos e fatores de formação dos solos, refletindo, portanto, a complexidade dos processos pedogenéticos na região da Cuesta de Botucatu. Tal complexidade foi atribuída à geologia, relevo variável e demais componentes, que não puderam ser revelados pelas covariáveis selecionadas, ou ainda, como considerado por Sun et al. (2012) e Menezes et al. (2016), pela forma desse relacionamento, que pode ser distinto do linear.

Para a avaliação da variabilidade à curta escala, foram ajustados os modelos teóricos esféricos e exponenciais aos variogramas dos resíduos, que segundo McBratney e Webster (1986), são os mais frequentemente encontrados para os atributos de solos. A estrutura de dependência espacial verificada não foi alta, indicando a deficiência da amostragem utilizada para permitir a representação da variabilidade, seja pela quantidade de amostras ou espaçamento entre elas. Essa aleatoriedade não captada pode ser derivada da complexa estratigrafia local.

A disponibilidade de dados de solos é a componente principal para a obtenção de modelos geoestatísticos robustos. Yigini e Panagos (2014) sugerem que os modelos podem ser melhorados usando mais dados de solo e covariáveis ambientais com alta resolução; uma vez que a resolução adotada no presente trabalho foi de 10 m, considerada adequada para condições de topografia acidentada, como observado por Silva (2013), deduziu-se que seria adequado mais dados de solos. Assim, ficou evidente a necessidade de amostragem mais adequada, considerando a quantidade de amostras e sua distribuição espacial, para permitir aplicação das técnicas geoestatísticas e híbridas. O mapa de erros da predição (mapa de resíduos) pode ser utilizado para localização de amostras adicionais.

Os mapas de desvio padrão de todos os atributos também evidenciaram a necessidade de amostragem adicional para maior acurácia na modelagem, como ressaltaram Hengl et al. (2007b), mormente nas áreas do front da cuesta e de várzea.

Foi observada uma maior variabilidade em profundidade para os atributos e, de modo geral, relação com maior número de covariáveis preditoras em profundidade. Essas observações demonstraram que a proximidade com o material de origem parece estabelecer uma evidência da interação entre a litologia e relevo na formação dos solos.

Portanto, dois aspectos básicos na aplicação de MDS foram considerados limitantes na obtenção de acurácia para o mapeamento dos atributos de solo na área da Fazenda Edgárdia: a amostragem e a heterogeneidade do material de origem, apesar da restrita extensão territorial da fazenda em comparação a aplicações regionais ou globais, comuns na literatura.

A importância cabal desses aspectos foi considerada por Samuel-Rosa et al. (2015) como um ponto fraco nas modelagens estatísticas relativas aos solos. Também Hengl et al. (2007b) ressaltaram que a amostragem e relação física entre variável estimada e preditores são duas das limitações dos modelos híbridos de RK. A quantidade e qualidade das observações de dados reais de solo e das variáveis preditoras foi apontada por McBratney et al. (2003), como fator limitante à aplicação do MDS.

Os mapas resultantes da modelagem, em geral, refletiram o comportamento real dos atributos ao longo da área da Fazenda Experimental Edgárdia, conhecido de pesquisadores e pedólogos com experiência de atuação na área, apesar das baixas correlações dos atributos com as covariáveis nas RLM e imprecisão observada na validação. Portanto, os mapas finais devem ser aprimorados para serem utilizados na aplicação em ações de manejo e tomadas de decisão na área da Fazenda Edgárdia.

#### **4.7 Modelagem considerando duas áreas geomorfologicamente distintas**

Com os resultados obtidos para os atributos modelados pela RK (baixos valores de  $R^2$  nas RLM e altos valores de ME e RMSE nas validações), foi aplicada a modelagem dividindo a área de estudo em duas áreas distintas, mais uniformes em

termos de elevação, no intuito de estratificar os dados e testar a redução da variabilidade para esses atributos. A recomendação de estratificação no processo de modelagem foi comentada por Padarian et al. (2014), em relação a processos complexos de modelagem, no projeto *Globalsoilmap.net*. Na avaliação dos autores, quando os fatores envolvidos na formação dos solos forem heterogêneos, devido às áreas extensas, seria válida a preocupação em relação à modelagem global ou local, seja separando subconjuntos de dados ou incluindo uma covariável extra no modelo, que reflita essa partição.

McBratney et al. (1991) encontraram melhor estimativa em teores de argila quando foi considerada a separação de áreas baseada em uma divisão topográfica natural em New South Wales, Austrália. Os autores consideraram que essa é uma ferramenta útil para aprimorar a representação de atributos de solo mapeados pela geoestatística.

Na separação das áreas da Fazenda Edgárdia, o conjunto total de amostras resultou na separação de 87 amostras para a parte alta e 54 amostras para a parte baixa da fazenda, sendo utilizadas integralmente para a modelagem, pois uma subdivisão desses conjuntos de amostras para treinamento e validação inviabilizaria a aplicação da geoestatística, notadamente na parte baixa da área, com menor número de amostras. Ainda assim, para alguns atributos, não foi possível modelar os variogramas dos resíduos.

Na aplicação do modelo híbrido às duas áreas, foi observado um pequeno acréscimo nos valores de  $R^2$  ajustado para as RLM, de modo geral (Tabela 24).

Foram consideráveis os incrementos nesse coeficiente para as variações da Arg na camada superficial e Ar nas duas camadas, tanto na parte baixa quanto alta, que foram melhor explicadas pelas covariáveis ambientais selecionadas, quando comparadas à área total. O maior aumento foi verificado para a Arg1 que havia resultado em RLM com  $R^2$  ajustado de 0,10, para a área total, considerando as covariáveis Elev e NDVI e apenas para a parte baixa, com a covariável Elev e  $R^2$  ajustado de 0,30.

**Tabela 24 – Modelos de RLM para atributos de solo nas áreas alta e baixa da Fazenda Edgárdia**

Atributo do solo	Modelo	R <sup>2</sup> ajustado <sup>1</sup>
IAV1 baixa	= -36,7266+0,0864*Elev	0,17
IAV1 alta	= 8,076-0,462*Fator LS	0,18
IAV2 baixa	= -57,105+0,127*Elev+0,110*Decl	0,46
IAV2 alta	= 11,2726+375,89*CPE	0,11
V1 baixa	= 50,4395+0,659320*Decl	0,06
V1 alta	= 11,3718+0,525*Decl+0,0848*Elev	0,15
V2baixa	= 138,52049+15,371373*senA- 195,393043*SAVI+0,905843*Decl	0,25
V2 alta	= 91,42288+5,167567*FLS-67,433244*NDVI	0,18
Ar1 baixa	= -3883,78+9,34*Elev	0,26
Ar1 alta	= 1842,9597-2,3102*Elev-8,6100*Decl+392,955*NDVI	0,39
Ar2 baixa	= -2927,061985+7,373603*Elev-80,627477*senA	0,22
Ar2 alta	= 1517,932729-2,213711*Elev+ 89,755152*cosA+714,571407*NDVI-7,786089*Decl	0,37
Arg1 baixa	= 2997,58-5,798*Elev	0,30
Arg1 alta	= -281,173+1,0169*Elev-202,823*NDVI	0,20
Arg2 baixa	= 2884,667950-5,260087*Elev- 14,909538*ITU+46,453073*senA	0,24
Arg2 alta	= 583,832729-605,496893*NDVI+69,905428*senA	0,23

<sup>1</sup> Modelo de regressão significativo a 5%.

IAV: Índice de avermelhamento; V: Saturação por bases em %; Ar: Teor de areia em g.kg-1; Arg: Teor de argila em g.kg-1; 1 = camada superficial; 2 = camada subsuperficial; Elev: elevação, FLS: fator LS, Decl: declividade, CPE: curvatura vertical, A: aspecto, SAVI: índice de vegetação ajustado ao solo, NDVI: índice de vegetação da diferença normalizada, ITU: índice topográfico de umidade.

No tratamento dos resíduos dos atributos, o modelo teórico esférico foi ajustado à maioria dos variogramas, da mesma forma como havia ocorrido para a área toda, sendo observado, no entanto, EPP para alguns atributos, na área baixa da fazenda (Tabela 25).

Na parte baixa, no entanto, foi verificado efeito pepita puro (EPP) para os variogramas do IAV nas duas camadas, além de V1 e Arg2, indicando que, mesmo com menor amplitude de variação da altitude e redução da área, não foi possível explicar a variabilidade, que inclusive apresentou aumento na área baixa (Tabela 26). A redução do número de amostras nos extratos pode ser a principal causa

desse fato. A separação das áreas ocasionou ajustes de variogramas de resíduos com valores de efeito pepita menores e estrutura de dependência espacial mais forte, exceto para V1, V2 e Arg2.

**Tabela 25 – Modelos e parâmetros dos variogramas isotrópicos ajustados aos resíduos da RLM para os atributos de solo nas partes alta e baixa da Fazenda Edgárdia**

Atributo do solo	Modelo	C <sub>0</sub>	C <sub>0</sub> +C	Alcance	IDE
IAV1 alta	esférico	0,30	3,77	500	0,92
IAV1 baixa	EPP	-	-	-	-
IAV2 alta	esférico	3	12,2	580	0,75
IAV2 baixa	EPP	-	-	-	-
V1 alta	esférico	120	300	700	0,60
V1 baixa	EPP	-	-	-	-
V2 alta	esférico	85	465	850	0,64
V2 baixa	esférico	350	550	650	0,36
Ar 1 alta	esférico	10000	40000	700	0,75
Ar 1 baixa	esférico	7000	39000	620	0,82
Ar2 alta	exponencial	15000	50000	400	0,70
Ar2 baixa	esférico	5000	47000	650	0,89
Arg1 alta	esférico	5000	11300	700	0,66
Arg1 baixa	esférico	0	11000	700	0
Arg2 alta	esférico	12000	22000	700	0,45
Arg2 baixa	EPP	-	-	-	-

IAV: Índice de avermelhamento; V: Saturação por bases em %; Ar: Teor de areia em g.kg-1; Arg: Teor de argila em g.kg-1; 1 = camada superficial, 2 = camada subsuperficial; C<sub>0</sub>: Efeito pepita, C<sub>0</sub>+C: patamar; IDE: índice de dependência espacial.

Estabelecendo uma comparação entre as duas áreas, o relevo representou melhor a variabilidade para o IAV e V no caso da camada superficial na parte alta e para a camada subsuperficial na parte baixa da área. Nas partes altas, o fator que deve ter condicionado a variabilidade é o material de origem basalto e não exatamente o relevo. Nas áreas de baixada, como houve deposição de material coluvial na superfície, o relevo refletiu melhor as condições do arenito, material responsável pela formação da maior parte dos solos.

**Tabela 26 – Estatística descritiva das amostras da área total, parte alta e parte baixa da Fazenda Experimental Edgárdia**

Atributo do solo	N	Mín	Máx	1°Q	3°Q	M	Md	DP	CV	CVm	As
IAV1	141	0,01	10,67	4,67	9,33	6,50	7,00	2,56	39	3,33	-0,30
IAV1a	87	3,00	10,67	4,67	9,33	7,04	7,00	2,20	31	3,37	0,22
IAV1b	54	0,01	9,33	3,37	7,80	5,64	5,60	2,87	50	6,99	-0,33
IAV2	141	0,01	18,00	7,00	12,00	9,59	10,50	4,66	48	4,11	0,13
IAV2a	87	2,33	18,00	9,33	16,00	11,66	10,50	4,01	34	3,71	0,15
IAV2b	54	0,01	12,80	2,54	10,21	6,27	5,43	3,61	57	7,91	0,22
pH1	141	3,70	7,30	5,10	6,00	5,53	5,60	0,74	13	1,13	-0,22
pH1a	87	4,20	7,30	5,30	6,05	5,69	5,70	0,68	12	1,29	-0,11
pH1b	54	3,70	6,80	4,73	5,88	5,26	5,25	0,77	15	2,01	-0,15
pH2	141	3,60	7,20	4,90	6,00	5,50	5,60	0,83	15	1,28	-0,36
pH2a	87	3,80	7,00	5,20	6,20	5,67	5,80	0,79	14	1,50	-0,55
pH2b	54	3,60	7,20	4,73	5,78	5,23	5,15	0,82	16	2,15	-0,09
CTC1	141	20,00	602,00	59,00	214,00	165,76	126,00	142,56	86	7,27	1,36
CTC1a	87	20,00	602,00	61,50	215,00	165,29	128,00	137,94	83	9,00	1,43
CTC1b	54	28,00	559,00	48,50	200,75	166,52	125,00	151,02	90	12,46	1,25
CTC2	141	15,00	607,00	51,00	241,00	166,88	121,00	144,63	86	7,32	1,26
CTC2a	87	22,00	607,00	52,50	221,00	164,79	121,00	140,86	85	9,22	1,39
CTC2b	54	15,00	604,00	49,00	284,75	170,24	116,00	151,79	88	12,25	1,07
V1	141	11,00	98,00	54,00	78,00	63,40	66,00	20,68	32	2,76	-0,62
V1a	87	14,00	98,00	57,50	83,50	67,72	69,00	18,78	28	2,99	-0,64
V1b	54	11,00	92,00	39,25	72,75	56,43	61,00	21,84	38	5,32	-0,48
V2	141	5,00	95,00	37,00	81,00	59,64	67,00	26,68	45	3,78	-0,58
V2a	87	7,00	95,00	50,00	84,00	65,11	70,00	23,83	36	3,95	-0,84
V2b	54	5,00	95,00	24,00	78,00	50,81	55,00	28,82	56	7,79	-0,15
Ar1	141	10,00	950,00	450,00	890,00	648,46	703,00	257,43	40	3,36	-0,65
Ar1a	87	75,00	943,00	440,00	870,00	615,89	640,00	256,39	41	4,49	-0,34
Ar1b	54	10,00	950,00	560,00	902,50	700,94	802,00	252,68	36	4,95	-1,22
Ar2	141	20,00	940,00	360,00	862,00	591,00	630,00	269,61	45	3,86	-0,29
Ar2a	87	61,00	937,00	308,50	840,00	538,00	462,00	274,24	51	5,50	0,06
Ar2b	54	20,00	940,00	503,50	873,00	676,39	740,00	240,70	35	4,89	-0,91
Arg1	141	10,00	600,00	60,00	260,00	170,50	140,00	132,07	77	6,55	1,08
Arg1a	87	10,00	600,00	78,50	270,00	181,66	160,00	120,47	66	7,15	0,75
Arg1b	54	30,00	590,00	50,00	197,50	152,54	90,00	148,30	96	13,35	1,50

(continua...)

Arg2	141	10,00	660,00	100,00	370,00	235,63	190,00	167,71	71	6,02	0,63
Arg2a	87	10,00	660,00	110,50	411,50	267,47	223,00	175,65	65	7,08	0,33
Arg2b	54	20,00	580,00	82,25	235,25	184,33	151,00	141,04	76	10,51	1,17

IAV: Índice de avermelhamento; CTC: Capacidade de troca de cátions em mmolc.dm<sup>-3</sup>; V: Saturação por bases em %; Ar: Teor de areia em g.kg<sup>-1</sup>; Arg: Teor de argila em g.kg<sup>-1</sup>; 1: camada superficial, 2: camada subsuperficial; a: parte alta, b: parte baixa; N: nº de observações; Amp: amplitude; M: média; Md: mediana; DP: Desvio padrão; CV: Coeficiente de variação em %; CVm: Coeficiente de variação médio em %; As: Coeficiente de assimetria;

Mosleh et al. (2016) verificaram a dificuldade de aplicar modelos MDS em áreas de baixadas, quando os fatores de formação do solo são mais homogêneos.

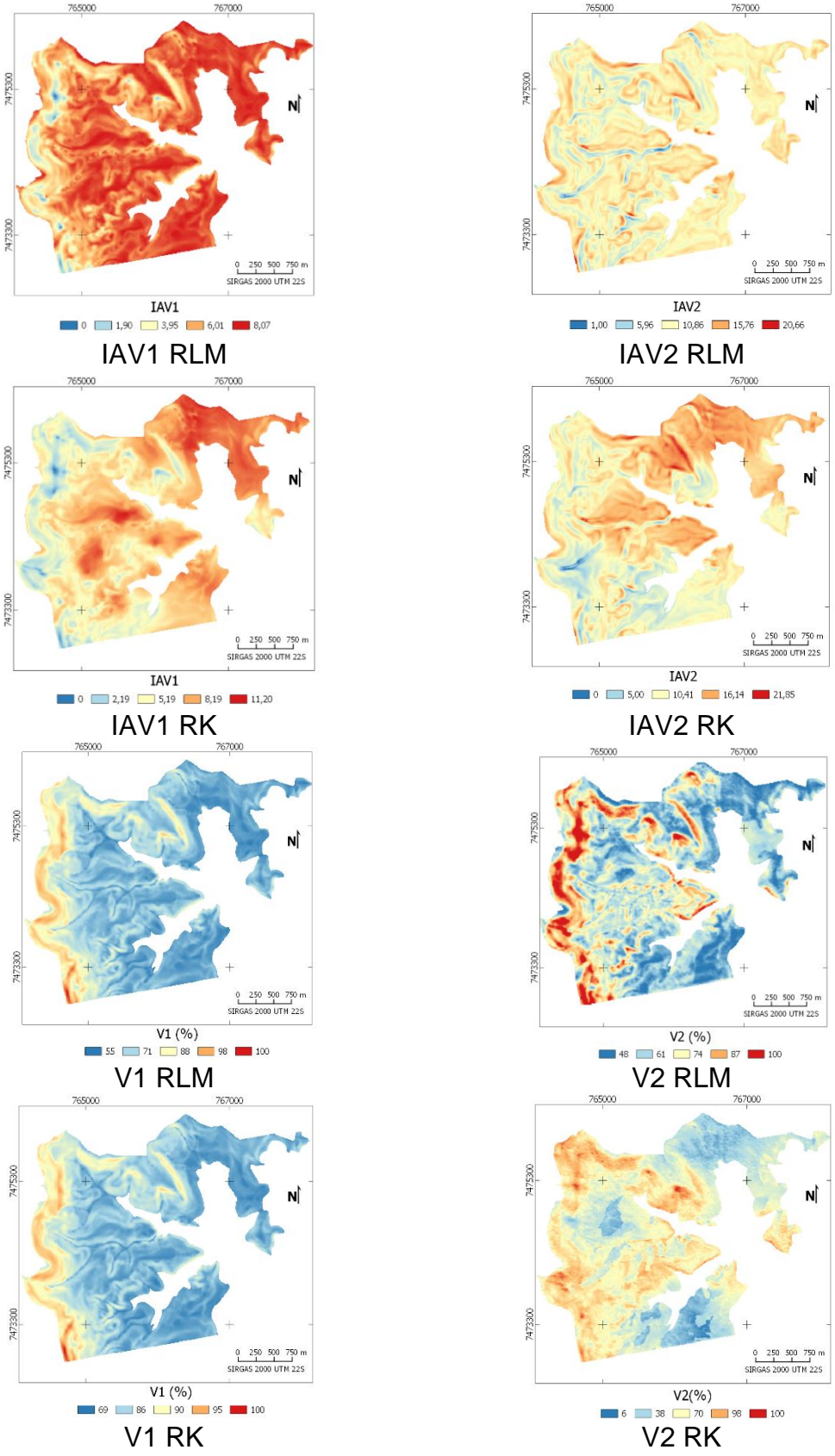
As covariáveis de relevo promoveram maior explicação da variação na parte alta para os teores de Ar, e na parte baixa, para a Arg, ao realizar as RLM. Dessa forma, nos locais de onde foram retirados sedimentos, o relevo teve melhor correlação com a Ar e onde foram depositados os sedimentos, foram verificadas maiores correlações e maiores valores de R<sup>2</sup> para a Arg, ambas nas duas camadas.

Os mapas de predição dos atributos obtidos para as áreas alta e baixa da Fazenda Edgárdia são apresentados nas Figuras 23 a 26.

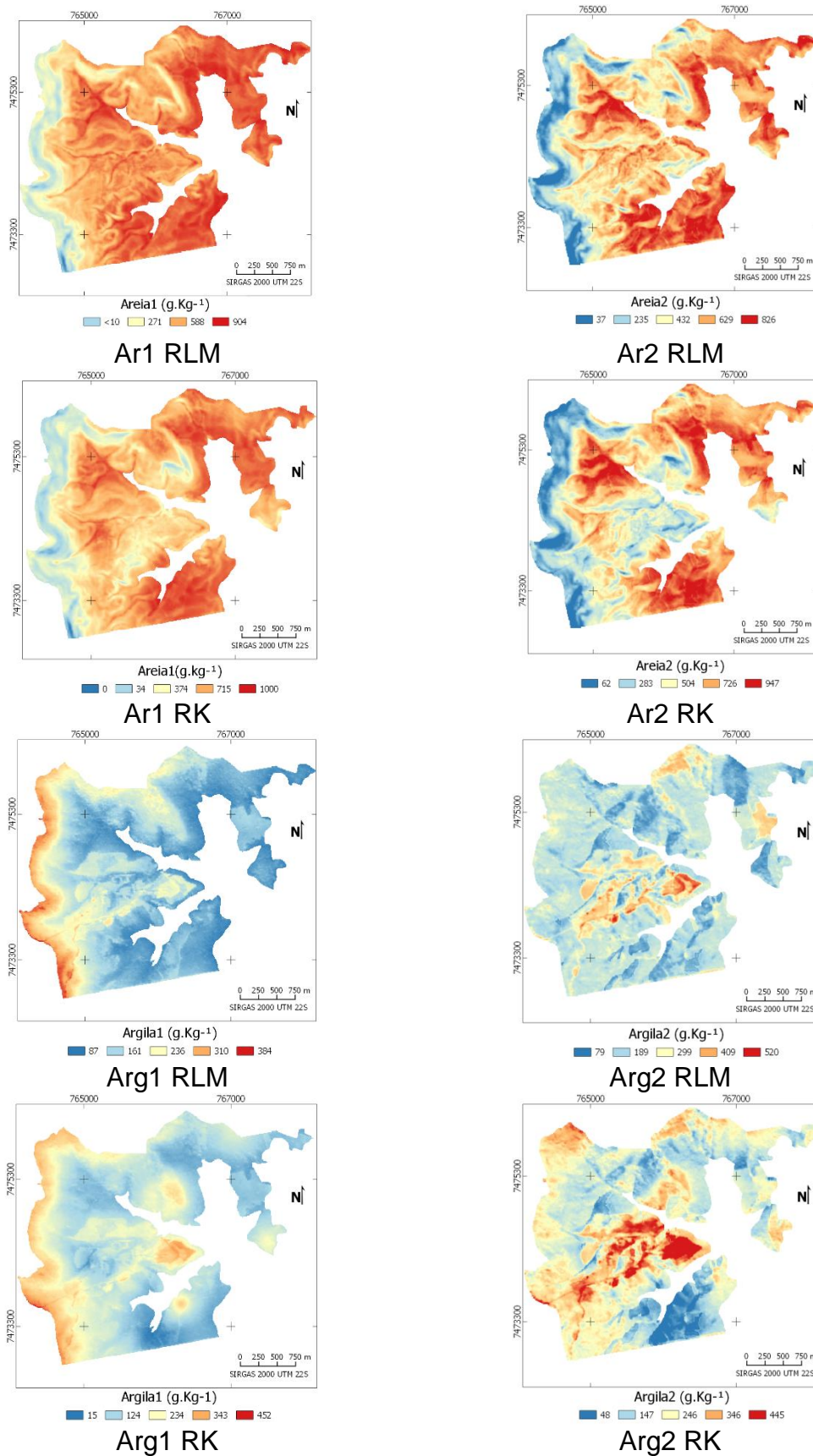
De modo geral, o padrão de variação e os valores estimados nos mapas de atributos não diferiram muito entre a área total e nas áreas divididas. Foram exceção a Ar1 e Arg1, na parte baixa, pois mantiveram a covariável auxiliar Elev, apresentaram valores mais elevados de R<sup>2</sup>, mas os valores dos atributos foram mais distintos. Diante desses resultados, notou-se que a alta variabilidade dos atributos dificultou a estimação, independente da menor amplitude de variação de altitudes, indicando que outras covariáveis, como o material de origem, pode ser o fator preponderante na formação dos solos e variação dos atributos analisados, corroborando a observação de Mosleh et al. (2016).

A variabilidade para a maioria dos atributos na área alta, expressa pelos valores do CVm, foi próxima da observada para a área total (Tabela 26). Para a área baixa, no entanto, foi observado ligeiro aumento dos valores, o que evidencia a característica de maior variabilidade espacial vertical e lateral de terras baixas, como foi observado por Menezes et al. (2016), ao mapear carbono orgânico (CO).

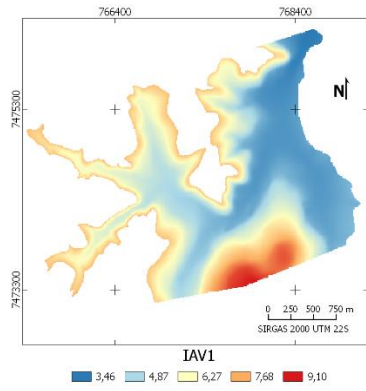
**Figura 23 – Mapas de atributos IAV e V% do solo na parte alta da Fazenda Edgárdia, 1: camada superficial; 2: camada subsuperficial**



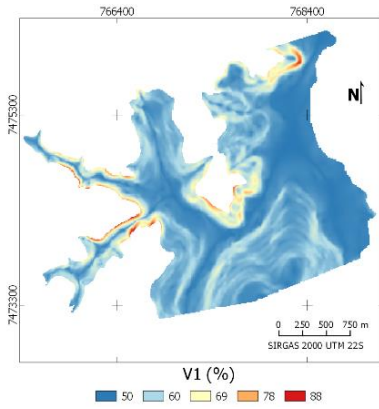
**Figura 24 – Mapas de atributos Areia e Argila do solo na parte alta da Fazenda Edgárdia. 1: camada superficial; 2: camada subsuperficial**



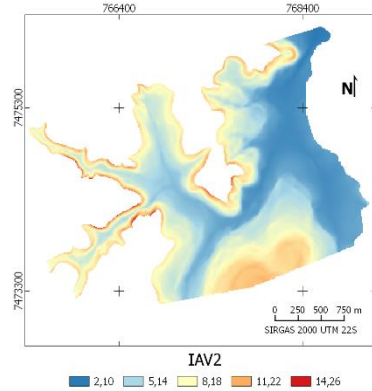
**Figura 25 – Mapas de atributos IAV e V% do solo na parte baixa da Fazenda Edgárdia. 1: camada superficial; 2: camada subsuperficial EPP: Efeito Pepita Puro**



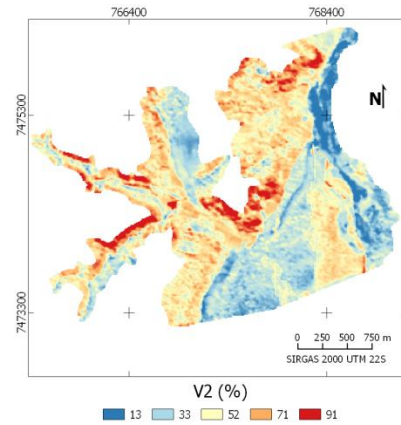
**IAV1 RLM**



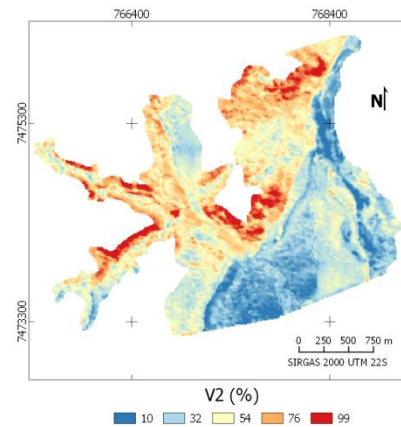
**V1 RLM**



**IAV2 RLM**

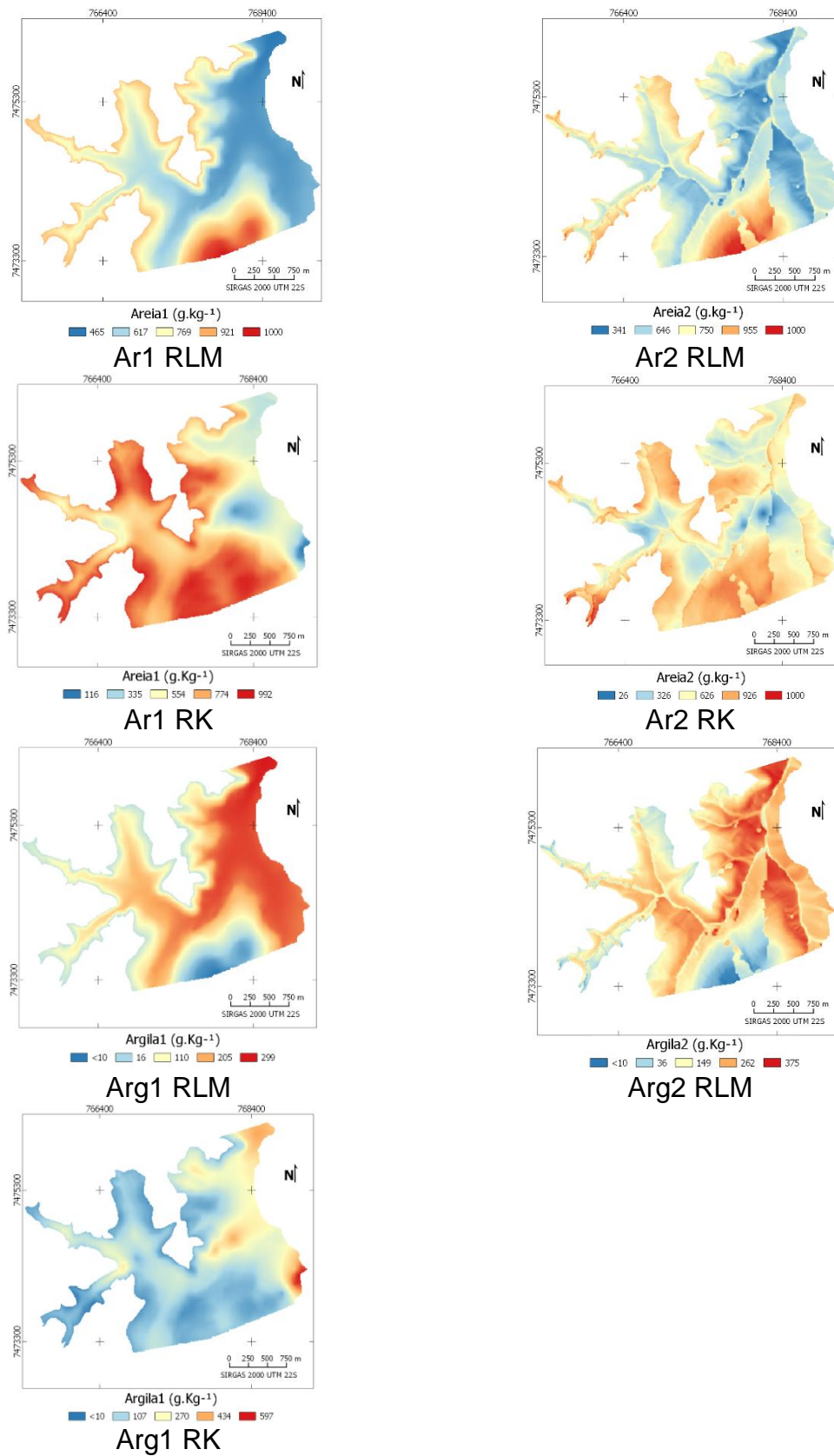


**V2 RLM**



**V2 RK**

**Figura 26 – Mapas de atributos Areia e Argila do solo na parte baixa da Fazenda Edgárdia. 1: camada superficial; 2: camada subsuperficial**



De acordo com Mosleh et al. (2016), o fato das áreas de baixadas serem comumente ocupadas com agricultura extensiva e tratarem-se de ecossistemas frágeis, implica em forte demanda por informações nessas áreas e a aplicação de uma abordagem confiável do MDS, para predição de propriedades de solos, torna-se um desafio. No caso da Fazenda Edgárdia essa consideração é aplicável e a adição de amostras, ou ainda, a adoção de abordagens que considerem o conhecimento e a tradução das interações locais em componentes dos algoritmos dos modelos, podem ser mais adequadas.

Considerando os resultados para as duas áreas em relação à área total, depreende-se a limitação relativa à amostragem, especialmente quando considerados os subconjuntos amostrais, limitação também apontada por McBratney et al. (1991), que vincularam o sucesso da modelagem em subáreas dependente da suficiência da amostragem, uma vez que sofreu uma redução para a modelagem nas novas áreas.

#### **4.8. Considerações finais**

A modelagem numérica na área da Fazenda Experimental Edgárdia teve o importante aspecto de explorar relações atributo de solo-covariáveis, o uso de dados legados, caracterizando a amostragem utilizada, bem como a resolução das informações no estudo. Assim, esse estudo pode ser abalizado como predecessor na obtenção de mapas confiáveis à aplicação em decisões relativas ao uso e manejo da área, ou em áreas nas condições locais de relevo típicas de cuesta.

Com relação ao modelo utilizado, apesar de ser considerado viável para o caso em análise e apontar o incremento de precisão na predição por RK em relação à RLM, ficou evidente o requerimento de altas correlações para sucesso do modelo. Indicadores estatísticos evidenciaram a necessidade de pesquisa de outras covariáveis que atendam esse requisito e resultem em modelos de maior acurácia.

A adoção de funções alternativas, além das relações lineares aqui avaliadas, deve ser fortemente considerada para representar a parte determinística do modelo.

O uso de dados legados, nas condições locais, aponta a questão de ausência de critérios estatísticos para a amostragem tradicional de solo, o que pode resultar em viés nos dados. Uma análise criteriosa dos dados deve revelar sua confiabilidade e, no caso de disponibilidade de recursos para amostragem complementar, a

localização dessas novas amostras deve ser criteriosamente planejada, baseando-se nos resultados obtidos nessa análise prévia.

Outras divisões em subáreas podem ser avaliadas, baseadas por exemplo, na existência de descontinuidades estruturais, como as fraturas presentes na propriedade, passíveis de promover resultados mais satisfatórios dos modelos.

Neste sentido, considera-se que a investigação, para além dos elementos abordados e nas diretrizes apontadas resultará em modelos acurados e aplicáveis de atributos de solos na área da Fazenda Experimental Edgárdia.

## 5. CONCLUSÕES

Pelos resultados obtidos nesse estudo, pode-se concluir que:

- não foram comprovadas as hipóteses de eficiência da aplicação de modelo híbrido multivariado, utilizando dados legados de solos e covariáveis ambientais de relevo e índices de sensoriamento remoto na elaboração de mapas de atributos na área da Fazenda Experimental Edgárdia e ainda, de que o relevo seja o fator preponderante de correlação com os atributos de solo na Cuesta de Botucatu;
- os mapas de predição de atributos gerados pelos modelos foram coerentes com a distribuição real dos atributos de solo na área, o que indica que a técnica é promissora para aplicação no contexto da Cuesta de Botucatu;
- a aplicação do modelo RK foi possível para predição dos atributos Índice de avermelhamento do solo, V%, Areia, Argila, e inviável para CTC e pH do solo;
- a baixa correlação entre os atributos de solo e covariáveis ambientais restringiu a acurácia dos modelos adotados;
- as covariáveis ambientais de relevo Elev, Decl, NDVI e ITU foram mais correlacionadas aos atributos de solo nas condições da modelagem;
- a geologia da área tem indícios de ser um fator preponderante a ser considerado na modelagem de atributos de solos na área;
- o modelo de Regressão-Krigagem apresentou desempenho superior à Regressão Linear Múltipla nas condições do estudo;

- o uso de dados legados de solo foi limitante para modelagem dos atributos de solo, sendo evidente a necessidade de amostragem mais adequada, em quantidade de amostras e distribuição espacial, para permitir aplicação das técnicas geoestatísticas e híbridas;
- o mapa de erros da predição (ou mapa de resíduos) pode ser utilizado para localização de amostras adicionais.

## REFERÊNCIAS

ADHIKARI, K.; HARTEMINK, A. E. Linking soils to ecosystem services – A global review. **Geoderma**, Amsterdam, v. 262, p. 101-111, 2016.

ALMEIDA, J. A.; TORRENT, J.; BARRÓN, V. Cor do solo, formas de fósforo e adsorção de fosfatos em Latossolos desenvolvidos de basalto do extremo sul do Brasil. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 27, n.6, p. 985-1002, 2003.

ARMSTRONG, M. **Basic Linear Geostatistics**. Berlim: Springer, 1998. 153 p.

ARRUDA, G. P.; DEMATTÊ, J. A.; CHAGAS, C. S. Mapeamento digital de solos por redes neurais artificiais com base na relação solo-paisagem. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 37, n. 2, p. 327-338, 2013.

BARTLEY, R. et al. **Land suitability**: technical methods. A technical report to the Australian Government for the Flinders and Gilbert Agricultural Resource Assessment project. Brisbane: CSIRO, 2013.

BEHRENS, T.; SHOLTEN, T. Digital soil mapping in Germany – a review. **Journal of Plant Nutrition Soil Science**, Weinheim, v. 169, p. 434-443, 2006.

BEHRENS, T. et al. Multi-scale digital terrain analysis and feature selection for digital soil mapping. **Geoderma**, Amsterdam, v. 155, n 3-4, p. 175-185, 2010.

BERTOL, I. et al. Propriedades físicas do solo sob preparo convencional e semeadura direta em rotação e sucessão de culturas, comparadas às do campo nativo. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 28, n. 1, p. 155-163, 2004.

BEVEN, K. J.; KIRKBY, M. J. A. Physically based variable contributing area model of basin hydrology. **Hydrological Sciences Bulletin**, Wallingford, v. 24, p. 43-69, 1979.

BOCK, M.; KÖTHE, R. Predicting the depth of hydromorphic soil characteristics influenced by ground water. **Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie**, Hamburg, Heft19, p. 13-22, 2008.

BOCKHEIM, J. G. et al. Soil-forming factors and Soil Taxonomy. **Geoderma**, Amsterdam, v. 226/227, p. 231-237, 2014.

BOHNER, J.; SELIGE, T. Spatial prediction of soil attributes using terrain analysis and climate regionalization. In: BOHNER, J.; McCLOY, K. R.; STROBL, J. SAGA – Analysis and Modelling Applications. **Goettinger Geographische Abhandlungen**, v. 115, p. 13-27, 2006.

BREGT, A.; McBRATNEY, A.; WOPEREIS, M. Construction of isolinear maps of soil attributes with empirical confidence limits. **Soil Science Society of America Journal**, Madison, v. 55, n. 1, p. 14-19, 1991.

BREVIK, E. C. et al., Soil mapping, classification and pedologic modeling: History and future directions. **Geoderma**, Amsterdam, v. 264, p. 256-274, 2016.

BUDIMAN, M. et al. **Revisão sobre funções de pedotransferência (PTFs) e novos métodos de predição de classes e atributos do solo**. Rio de Janeiro: Embrapa Solos, 2003. (Série Documentos, 45).

CAMBARDELLA, C. A. et al. Field-scale variability of soil properties in Central Iowa soils. **Soil Science Society of America Journal**, Madison, v. 58, n.5, p. 1501-1511, 1994.

CAMPOS, M. C. C. et al. Relações solo-superfície geomórfica em uma topossequência várzea-terra firme na região de Humaitá (AM). **Revista Brasileira de Ciência do Solo**, Viçosa, v. 36, n. 2, p. 325-336. 2012.

CARRÉ, F.; GIRARD, M. C. Quantitative mapping of soil types based on regression kriging of taxonomic distances with landform and land cover attributes **Geoderma**, Amsterdam, v. 110 n. 3/4, 241-263, 2002.

CARTA de Brasília. Viçosa: Sociedade Brasileira de Ciência do Solo, 2015. Disponível em: <<http://www.sbcs.org.br/wp-content/uploads/2015/03/carta-Brasilia-1.pdf>>. Acesso em: 03 jun. 2016.

CARVALHO, W. A.; PANOSO, L. A.; MORAES, M. H. **Levantamento semidetalhado dos solos da Fazenda Experimental Edgárdia – Município de Botucatu - SP**. Botucatu: Universidade Estadual Paulista, 1991. 2 v.

CARVALHO JÚNIOR, W. et al. A regional-scale assessment of digital mapping of soil attributes in a tropical hillslope environment. **Geoderma**, Amsterdam, v. 232/234, p. 479-486, 2014a.

CARVALHO JÚNIOR, W. et al. Evaluation of statistical and geostatistical models of digital soil properties mapping in tropical mountain regions. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 38, n. 3, p. 706-717, 2014b.

CARVALHO JÚNIOR, W. et al. Digital soilscape mapping of tropical hillslope areas by neural networks. **Scientia Agricola**, Piracicaba, v. 68, n. 6, p. 691-696, 2011.

CASTRIGNANÒ, A. **Introduction to spatial data processing**. Bari: Aracne, 2011. 108 p.

CASTRIGNANÒ, A. et al. Study of spatial relationships among some soil physico-chemical properties of a field in central Italy using multivariate geostatistics. **Geoderma**, Amsterdam, v. 97, p. 39-60, 2000.

CHAGAS, C. S. et al. Avaliação de modelos digitais de elevação para aplicação em um mapeamento digital de solos. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v. 14, n.2, p. 218–226, 2010.

CHAGAS, C. S.; CARVALHO JÚNIOR, W.; BHERING, S. B. Integração de dados do Quickbird e atributos do terreno no mapeamento digital de solos por redes neurais artificiais. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 35, n. 3, p. 693-704. 2011.

CHIG, L. A.; COUTO, E. G.; AMORIM, R. S. R. Tecnologias para levantamento da variabilidade dos atributos do solo para um programa de agricultura de precisão. **UNIClências**, v. 14, n.2, p. 127-149, 2010.

CHILÈS, J. P.; DELFINER, P. **Geostatistics: modeling spatial uncertainty**. 2nd ed. New York: John Wiley & Sons, 2012. 734 p.

CIAMPALINI, R.; LAGACHERIE, P; HAMROUNI, H. Documenting GlobalSoilMap.net grid cells from legacy measured soil profile and global available covariates in Northern Tunisia. In: MINASNY, B. et al. (Ed.). **Digital soil assessments and beyond**. London: Taylor & Francis Group, 2012. p. 439-444.

COELHO, F. B.; GIASSON, E. Métodos para mapeamento digital de solos com utilização de sistemas de informação geográfica. **Ciência Rural**, Santa Maria, v. 40, n. 10, p. 2099-2106, 2010.

CONRAD, O. et al. System for automated geoscientific analyses – SAGA v. 2.1.4. **Geoscientific Model Development**, Goettingen, v. 8, p. 1991-2007, 2015.

CORDEIRO, L. L. et al. Influência da resolução espacial de Modelos Digitais de Elevação na delimitação automática de Áreas de Preservação Permanente em topos de morros. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 2013, Foz do Iguaçu. **Anais...** São José dos Campos: INPE, 2013. p. 1306-1313.

CUNHA, A. R.; MARTINS, D. Classificação Climática para os municípios de Botucatu e São Manuel – SP. **Revista Irriga**, Botucatu, v. 14, n. 1, p. 1-11. 2009.

CUOMO, S. et al. Physically based modelling of soil erosion induced by rainfall in small mountain basins. **Geomorphology**, Amsterdam, v. 243, p. 106-115, 2015.

CURI, N.; FRANZMEIER, D. P. Toposequence of Oxisols from the Central Plateau of Brazil. **Soil Science Society of America Journal**, Madison, v. 48, n. 2, p. 341-346, 1984.

- DANTAS, J. S. et al. Coeficiente de erodibilidade em sulcos e entressulcos de Argissolos coesos estimado pela cor do solo. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 49, n. 9, p. 700-707, set. 2014.
- DEWITTE, O. et al. Satellite remote sensing for soil mapping in Africa: An overview. **Progress in Physical Geography**, London, v. 36, n. 4, p. 514-538, 2012.
- DOBARCO, M. R. et al. Prediction of soil texture using descriptive statistics and area to point kriging in Region Centre (France). **Geoderma Regional**, Amsterdam, v. 7, n. 3, p. 279-292, 2016.
- DOBOS, E. et al. Use of combined digital elevation model and satellite radiometric data for regional soil mapping. **Geoderma**, Amsterdam, v. 97, p. 367-391, 2000.
- DOBOS, E. et al. **Digital Soil Mapping as a support to production of functional maps**. Luxemburg: Office for Official Publications of the European Communities, 2006. 68 p.
- DOBOS, E.; HENGL, T. Soil Mapping Applications. **Developments of Soil Science**, Madison, v. 33, p. 461-479, 2009.
- DU, F. et al. Soil property variation mapping through data mining of soil category maps. **Hydrological Processes**, Chichester, v. 29, p. 2491-2503, 2015.
- EMBRAPA. Serviço Nacional de Levantamento e Conservação do Solo. **Manual de métodos de análise de solo**. Rio de Janeiro: Embrapa Solos, 1979.
- EMBRAPA. **Procedimentos normativos de levantamentos pedológicos**. Rio de Janeiro: Embrapa, 1995. 101 p.
- ESRI **ArcGIS Desktop**. Software, Versão 10.1 Redlands, CA: Environmental Systems Research Institute, 2012. Disponível em: <http://www.esri.com/software/arcgis/arcgis-for-desktop/free-trial>. Acesso em: 10 jan. 2014.
- FLORENZANO, T. G. **Geomorfologia**: conceitos e tecnologias atuais. São Paulo: Oficina de Textos, 2008.
- FLORINSKY, I. V. **Digital Terrain Analysis in Soil Science and Geology**. Amsterdam: Elsevier, 2012a. 379 p.
- FLORINSKY, I. V. The Dokuchaev hypothesis as a basis for predictive digital soil mapping (on the 125<sup>th</sup> anniversary of its publication). **Eurasian Soil science**, Pochvovedenic, v. 45, n. 4, p. 445-451, 2012b.
- GIASSON, E. et al. Decision trees for digital soil mapping on subtropical basaltic steeplands. **Scientiae Agricola**, Piracicaba, v. 68, n. 2, p.167-174. 2011.

GOMEZ, C.; LAGACHERIE, P.; COULOUMA, G. Regional predictions of eight common soil properties and their spatial structures from hyperspectral Vis-NIR data. **Geoderma**, Amsterdam, v. 189-190, p. 176-185, 2012.

GOOVAERTS, P. Geostatistics in soil science: state-of-the-art and perspectives. **Geoderma**, Amsterdam, v. 89, p. 1-45, 1999.

GRASS DEVELOPMENT TEAM. Geographic Resources Analysis Support System (**GRASS**) Software, Versão 6.4. Open Source Geospatial Foundation. 2015. Disponível em: < <http://grass.osgeo.org> > Acesso em 10 jan. 2015.

GRAY, J. M.; BISHOP, T. F.; YANG, X. Pragmatic models for the prediction and digital mapping of soil properties in eastern Australia. **Soil Research**, Clayton South, v. 53, p. 24-42, 2015.

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

HAIR, J. F. et al. **Análise multivariada de dados**. 6. ed. Porto Alegre: Bookman, 2009. 688 p.

HANSEN, M. K. et al. Inductively mapping expert derived soil-landscape units within dambo wetland catenae using multispectral and topographic data. **Geoderma**, Amsterdam, v. 150, n. 1/2, p. 72-84, 2009.

HARTEMINK, A. E. The use of soil classification in journal papers between 1975 and 2015. **Geoderma Regional**, Maryland Heights, v. 5, p. 127-139, 2015.

HEBELER, F; PURVES, R. S. The influence of elevation uncertainty on derivation of topographic indices. **Geomorphology**, Amsterdam, v. 111, n. 1/2, p. 4-16, 2009.

HENGL, T. et al. A generic framework for spatial prediction of soil variables based on regression-kriging. **Geoderma**, Amsterdam, v. 120, p. 75-93, 2004.

HENGL, T. et al. Methods to interpolate soil categorical variables from profile observations: lessons from Iran. **Geoderma**, Amsterdam, v.140, p. 417-427, 2007a.

HENGL, T.; HEUVELINK, G. B. M.; ROSSITER, D. G. About regression-kriging: from equations to case studies. **Computers and Geosciences**, Amsterdam, v. 33, p. 1301-1315, 2007b.

HENGL, T.; HEUVELINK, G. B. M.; STEIN, A. **Comparison of kriging with external drift and regression-kriging**: Technical note, ITC. Enschede, 2003. Disponível em: < [http://www.itc.nl/library/Papers\\_2003/misca/hengl\\_comparison.pdf](http://www.itc.nl/library/Papers_2003/misca/hengl_comparison.pdf) >. Acesso em 10/01/2017.

HENGL, T. **A practical guide to geostatistical mapping**. Luxembourg: Office for Official Publications of the Europe Communities. 2009. 270 p.

HENGL, T.; MacMILLAN, R. A. Geomorphometry – A key to landscape mapping and modeling. **Developments in Soil Science**. Amsterdam, v. 33, p. 433-460, 2009.

HEUVELINK, G. B. M.; WEBSTER, R. Modelling soil variation: past, present and future. **Geoderma**, Amsterdam, v. 100, n. 3/4, p. 269-301, 2001.

HEUVELINK, G. B. M.; WEBSTER, R. Reply to Comment on “Modelling soil variation: past, present, and future”, by Philippe Baveye. **Geoderma**, Amsterdam, v. 109, n. 3/4, p. 295-297, 2002.

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

HUETE, A. R. A Soil-Adjusted Vegetation Index (SAVI). **Remote sensing of environment**, Amsterdam, v. 25, p. 295-309, 1988.

HUTCHINSON, M. F. A new procedure for gridding elevation and streamline data with automatic removal of spurious pits. **Journal of hydrology**, Amsterdam, v. 106, p. 211-232, 1989.

HUTCHINSON, M. F. A local adaptive approach to the interpolation of digital elevation models. In: THIRD INTERNATIONAL CONFERENCE ON INTEGRATING GIS AND ENVIRONMENT MODELING, 1996, Santa Fé. **Proceedings...** Santa Fé: NCGIA, 1996. p. 390-396.

HUTCHINSON, M. F. Development of a continent-wide DEM with applications to terrain and climate analysis. In: GOODCHILD, M. F. (Ed.). **Environmental Modeling with GIS**. New York: Oxford University Press, 1993. p.392-399.

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

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. **Manual técnico de pedologia**. 2. ed. Rio de Janeiro: IBGE, 2007. 323 p. (Manuais Técnicos em Geociências, 4).

INSTITUTO DE PESQUISAS TECNOLÓGICAS DO ESTADO DE SÃO PAULO S/A – IPT. **Mapa geológico do Estado de São Paulo**. São Paulo: IPT, 1981. Escala 1:500.000.

ISAAKS, E. H.; SRIVASTAVA, R. M. **An introduction to applied geostatistics**. New York: Oxford University Press, 1989. 561p.

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

JENSEN, J. R. **Sensoriamento Remoto do Ambiente: uma perspectiva em recursos terrestres**. 2. ed. São José dos Campos: Parênteses, 2009. 598 p.

JORGE, L. A. B. Soil erosion fragility assessment using an impact model and geographic information system. **Scientia Agricola**, Piracicaba, v. 66, n. 5, p. 658-666, 2009.

JORGE, L. A. B.; MOREIRA, M. P. Padrões da fragmentação do habitat na Cuesta de Botucatu, SP. **Ciência Florestal**, Santa Maria, v. 10, n. 1, p.141-157, 2000.

JORGE, L. A. B.; SARTORI, M. S. Uso do solo e análise temporal da ocorrência de vegetação natural na Fazenda Experimental Edgárdia, em Botucatu-SP. **Revista Árvore**, Viçosa, v. 26, n. 5, p. 585-592, 2002.

JORGE, L. A. B.; MORAES, M. H.; LIMA, S. L. Spatial distribution of runoff depth from a watershed located in a cuesta relief área of São Paulo state, Brazil. **International Journal of Geosciences**, Irvine, v. 5, n. 2, p. 137-145, 2014.

KIM, M. S. et al. Effect of topography and soil parametrization representing soil thicknesses on shallow landslide modelling. **Quaternary International**, Kidlington, v. 384, p. 91-106, 2015. DOI: <http://dx.doi.org/10.1016/j.quaint.2015.03.057>.

KRAVCHENKO, A.; BULLOCK, D. A comparative study of interpolation methods for mapping soil properties. **Agronomy Journal**, Madison, v. 91, n. 3, p. 393-400, 1999.

LANDIM, P. M. B. (Coord.). **Mapa geológico do Estado de São Paulo**. São Paulo: IGCE/Unesp – DAEE, 1984. Escala 1:250.000.

LANDIM, P. M. B. **Análise estatística de dados geológicos**. 2. ed. rev. ampl. São Paulo: Editora Unesp, 2003. 253 p.

LANDIM, P. M. B. Sobre Geoestatística e mapas. **Terrae Didactica**, Campinas, v. 2, n. 1, p. 19-33. 2006. Disponível em: <<http://www.ige.unicamp.br/terraedidactica>>. Acesso em 20 set. 2015.

LEPSCH, I. **19 Lições de pedologia**. São Paulo: Oficina de textos, 2011. 456 p.

LIEB, M. Sampling for regression-based digital soil mapping: closing the gap between statistical desires and operational applicability. **Spatial Statistics**, Amsterdam, v. 13, p. 106-122, 2015.

LIMA, L. A. S. et al. **Mapeamento de solos do tradicional ao digital**. Planaltina: Embrapa Cerrados, 2013. 51 p. (Série Documentos).

MacMILLAN, R. A. Experiences with applied DSM: Protocol, availability, quality and capacity building. In: HARTEMINK, A. E. et al. (Ed.). **Digital soil mapping with limited data**. Wageningen: Springer, 2008. p. 113-135.

MALONE, B. P. et al. Mapping continuous depth functions of soil carbon storage and available water capacity. **Geoderma**, Amsterdam, v. 154, n. 1-2, p. 138-152. 2009.

MALONE, B. P.; McBRATNEY, A. B.; MINASNY, B. Empirical estimates of uncertainty for mapping continuous depth functions of soil attributes. **Geoderma**, Amsterdam, v. 160, p. 614-626. 2011.

MARQUES, S. R. et al. Qualidade física de um Latossolo Vermelho, perdas por erosão e desenvolvimento do milho em dois sistemas de manejo. **Ciência e Agrotecnologia**. Lavras, v. 34, n. 4, p. 967-974, 2010.

MATHERON, G. Principles of Geostatistics. **Economic Geology**, El Paso, v. 58, n. 8, p. 1246-1266, 1963.

MATHERON, G. **The theory of regionalized variables and its applications**. Paris: School of Mines Publication, 1971. 211p.

McBRATNEY, A. B.; WEBSTER, R. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. **Journal of Soil Science**, Chichester, 37, p. 617-639, 1986.

McBRATNEY, A. B.; HART, G. A.; McGARRY, D. The use of partitioning to improve the representation of geostatistically mapped soil attributes. **Journal of Soil Science**, Chichester, v.42, p. 513-532, 1991.

McBRATNEY, A. B.; MENDONÇA SANTOS, M. L.; MINASNY, B. On digital soil mapping. **Geoderma**, Amsterdam, v. 117, n. 1/2, p. 3-52, 2003.

McBRATNEY, A. B. et al. An overview of pedometric techniques for use in soil survey. **Geoderma**, Amsterdam, v. 97, n. 3/4, p. 293-327, 2000.

McKENZIE, N. J.; RYAN, P. J. Spatial prediction of soil properties using environmental correlation. **Geoderma**, Amsterdam, v. 89, n. 1/2, p. 67-94, 1999.

MELLO, C. R. et al. Interpolation methods for improving the RUSLE R-factor mapping in Brazil. **Journal of Soil and Water Conservation**, Ankeny, v. 70, n. 3, p. 182-197. 2015.

MENEZES, M. D. et al. Spatial prediction of soil properties in two contrasting physiographic regions in Brazil. **Scientia Agricola**, Piracicaba, v. 73, n. 3, p. 274-285, May/June 2016.

MENDONÇA-SANTOS, M. L.; SANTOS, H. G. **Mapeamento digital de classes e atributos de solos: métodos, paradigmas e novas técnicas**. Rio de Janeiro: Embrapa Solos, 2003. 19 p.

MENDONÇA-SANTOS, M. L.; SANTOS, H. G. The state of the art of Brazilian soil mapping and prospects for digital soil mapping. In: LAGACHERIE, P.; McBRATNEY, A. B.; VOLTZ, M. (Ed.). **Digital soil mapping: an introductory perspective**. Amsterdam: Elsevier, 2006. p. 39-54.

MILONE, G.; ANGELINI, F. **Estatística aplicada**. São Paulo: Atlas, 1995. 286 p.

MINASNY, B.; HARTEMINK, A. E. Predicting soil properties in the tropics. **Earth Science Reviews**, Amsterdam, v. 106, p. 52-62, 2011.

MINASNY, B.; McBRATNEY, A. B. Digital soil mapping: a brief history and some lessons. **Geoderma**, Amsterdam, v. 264, 301-311. 2016.

MOORE, I. D.; GRAYSON, R. B.; LADSON, A. R. Digital terrain modelling: a review of hydrological, geomorphological and biological applications. **Hydrological Processes**, v. 15, n.1, p. 3-30. 1991.

MOORE, I. D. et al. Soil attribute prediction using terrain analysis. **Soil Science Society of America Journal**. Madison. v. 57, n. 2, p.443-452.1993.

MOREIRA, M. A. **Fundamentos do sensoriamento remoto e metodologias de aplicação**, terceira edição. Viçosa: UFV, 2007. 320 p.

MOSLEH, Z. et al. The effectiveness of digital soil mapping to predict soil properties over low-relief areas. **Environmental Monitoring and Assessment**, Dordrecht, v. 188, n. 3, 195, 2016.

MULDER, V. L. et al. The use of remote sensing in soil and terrain mapping – a review. **Geoderma**, Amsterdam, v. 162, p. 1-19. 2011.

NANDI, A.; SHAKOOR, A. A GIS-based landslide susceptibility evaluation using bivariate and multivariate statistical analyses. **Engineering Geology**, Amsterdam, v. 110, p. 11-20, 2009.

ODEH, I. O. A.; McBRATNEY, A. B.; CHITTLEBOROUGH, D. J. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. **Geoderma**, Amsterdam, v. 67, n. 3/4, p. 215-226, 1995.

OLAYA, V.; CONRAD, O. Geomorphometry in SAGA. In: HENGL, T.; REUTER, H. I. (Ed) Geomorphometry : Concepts, software, applications. **Developments in soil science**, Amsterdam, v. 33, p. 293-308, 2009.

OLIVEIRA, J. B.; CAMARGO, M. N. **Classes gerais de solos do Brasil: guia auxiliar para seu reconhecimento**. Jaboticabal: Funep, 1992. 201 p.

OLIVER, M. A.; WEBSTER, R. A tutorial guide to geostatistics: computing and modelling variograms and kriging. **Catena**, Amsterdam, v. 113, p. 56-69, 2014.

OMUTO, C. T.; VARGAS, R. R. Re-tooling of regression-kriging in R for improved digital mapping of soil properties. **Geosciences Journal**, v. 19, n. 1, p. 157-165, 2015.

PADARIAN, J.; MINASNY, B.; McBRATNEY, A. B. The evolving methodology for global soil mapping. In: ARROUAYS, D. et al. (Ed.). **GlobalSoilMap: basis of the global spatial soil information system**. London: CRC press, 2014. p. 215-220.

PELUCO, R. G. et al. Mapeamento do fósforo adsorvido por meio da cor e da suscetibilidade magnética do solo. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 50, n. 3, p. 259-266, mar. 2015.

PINHEIRO, H. S. K. et al. Modelos de elevação para obtenção de atributos topográficos utilizados em mapeamento digital de solos. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 47, n. 9, p. 1384-1394, 2012.

PONZONI, F. J.; SHIMABUKURO, Y. E.; KUPLICH, T. M. **Sensoriamento remoto da vegetação**. São Paulo: Oficina de textos, 2012. 160 p.

PRADHAN, B. Remote sensing and GIS-based landslide hazard analysis and cross-validation using multivariate logistic regression model on three test areas in Malaysia. **Advances in Space Research**, Kidlington, v. 45, p. 1244-1256, 2010.

QGIS DEVELOPMENT TEAM. **QGIS**. Software, Versão 2.8. 2014. Disponível em: <<http://qgis.org/>> Acesso em 10 ago. 2015.

QUEIROZ NETO, J.P. Geomorfologia e pedologia. **Revista Brasileira de Geomorfologia**, São Paulo, v. 1, n. 1, p. 59-67, 2000.

RAD, M. R. P. et al. Updating soil survey maps using random forest and conditioned Latin hypercube sampling in the loess derived soils of northern Iran. **Geoderma**, Amsterdam, v. 232-234, p. 97-106, nov. 2014.

RESENDE, M. et al. **Pedologia**: base para distinção de ambientes. 6. ed. Lavras: UFLA, 2014. 378 p.

RODOLFO JÚNIOR, F. et al. Relações solo-paisagem em topossequências na Fazenda Água Limpa, Distrito Federal. **Nativa**, Sinop, v. 3, n. 1, p. 27-35, 2015.

ROSS, J. L. S.; MOROZ, I. C. Mapa geomorfológico do Estado de São Paulo. **Revista do Departamento de Geografia**, São Paulo, v. 10, p. 20-32, 1997.

ROUSE, J. W. et al. Monitoring vegetation systems in the Great Plains with ERTS. **Third ERTS Symposium**, NASA SP-351, vol. 1, pp. 309-317, 1973. Disponível em <<http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19740022592.pdf>>. Acesso em 18/11/2015.

ROSA, A. S. **Funções de predição espacial de propriedades do solo**. 2012. 201 f. Dissertação (Mestrado em Ciência do Solo)-Universidade Federal de Santa Maria, Santa Maria, 2012.

SAGA GIS DEVELOPMENT TEAM. System for Automated Geoscientific Analyses (**SAGA GIS**). Software, Versão 2.1. Disponível em <<http://www.sagagis.org/>> Acesso em 12 fev 2015.

SAMUEL-ROSA, A. et al. Do more detailed environmental covariates deliver more accurate soil maps? **Geoderma**, Amsterdam, v. 243-244, p. 214-227, 2015.

SANCHEZ, P. A. et al. Digital soil map of the world. **Science**, Washington, v. 325, p. 680-681, 2009.

SANTANA, O. A. et al. Relação entre o índice de avermelhamento do solo e o estoque de carbono na biomassa aérea da vegetação de cerrado. **Ciência Florestal**, Santa Maria, v. 23, n. 4, p. 783-794, out/dez 2013.

SANTOS, H. G. et al. **Sistema brasileiro de classificação de solos**. 3. ed. rev. ampl. Brasília, DF: Embrapa, 2013. 353 p.

SÃO PAULO (Estado). Secretaria de Economia e Planejamento. **Estudo da morfologia e da hierarquia funcional da rede urbana paulista e regionalização do Estado de São Paulo**: extrato: documento 2: resultados. São Paulo, 2010. 125 p. Disponível em: <[http://deolhonodinheiropublico.jundiai.sp.gov.br/sitetranspv2/biblio.nsf/V03.01/smf\\_c ontas\\_relatorios/\\$file/AUJ.pdf](http://deolhonodinheiropublico.jundiai.sp.gov.br/sitetranspv2/biblio.nsf/V03.01/smf_c ontas_relatorios/$file/AUJ.pdf)>. Acesso em: 02 set 2016.

SARMENTO, E. C. et al. Prediction of soil orders with high spatial resolution: response of different classifiers to sampling density. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 47, n. 9, p. 1395-1403, 2012.

SEIDEL, E. J.; OLIVEIRA, M. S. Novo índice geoestatístico para a mensuração da dependência espacial. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 38, n. 3, p. 699-705, 2014.

SHI, W. et al. Surface modelling of soil properties based on land use information. **Geoderma**, Amsterdam, v. 162, p. 347-357, 2011.

SILVA, A. F. et al. A new approach to soil classification mapping based on the spatial distribution of soil properties. **Geoderma**, Amsterdam, v. 219-220, p. 106-116, 2014.

SILVA, M. A. **Mapeamento digital de atributos do solo e vulnerabilidade ao escoamento superficial, baseado no conhecimento de campo, na sub-bacia das Posses, Extrema, MG**. 2013. 109 f. Tese (Doutorado em Ciência do Solo)-Universidade Federal de Lavras, Lavras, 2013.

SUN, W.; MINASNY, B.; McBRATNEY, A. Analysis and prediction of soil properties using local regression-kriging. **Geoderma**, Amsterdam, v. 171/172, p. 16-23, 2012.

SRIVASTAVA, R. M. Geostatistics: a toolkit for data analysis, spatial prediction and risk management in the coal industry. **International Journal Coal Geology**, Amsterdam, v. 112, p. 2-13. 2013.

TEN CATEN, A. **Mapeamento digital de solos: metodologias para atender a demanda por informação espacial em solos**. 2011. 106 f. Tese (Doutorado em Ciência do Solo)-Universidade Federal de Santa Maria, Santa Maria, 2011.

TEN CATEN, A. et al. Extrapolação das relações solo-paisagem a partir de uma área de referência. **Ciência Rural**, Santa Maria, v. 41, p. 812-816, 2011.

TEN CATEN, A. et al. Mapeamento digital de classes de solos: características da abordagem brasileira. **Ciência Rural**, Santa Maria, v. 42, n. 11, p. 1989- 1997, 2012.

TORRENT, J.; BARRÒN, V. Laboratory measurement of soil color: theory and practice. In: BIGHAM, J. M.; CIOLKOSZ, E. J. (Ed.) **Soil Color**. Madison: SSSA, 1993. p. 21-33. (Special Publication, 31).

UMALI, B. P. et al. The effect of terrain and management on the spatial variability of soil properties in an apple orchard. **Catena**, Amsterdam, v. 93, p. 38-48, 2012.

VALERIANO, M. M. Dados topográficos. In: FLORENZANO, T. G. (Org.). **Geomorfologia: conceitos e aplicações atuais**. São Paulo: Oficina de Textos, 2008. cap. 3, p. 73-104.

VAYSSE, K.; LAGACHERIE, P. Evaluating Digital Soil Mapping approaches for mapping GlobalSoilMap soil properties from legacy data in Languedoc-Roussillon (France). **Geoderma Regional**, Maryland Heights, v. 4, p. 20-30, 2014.

WARRICK, A. W.; NIELSEN, D. R. Spatial variability of soil physical properties in the field. In: HILLEL, D. (Ed.) **Applications of soil physics**. New York: Academic Press, 1980. 385p.

WEBSTER, R. Statistics to support soil research and their presentation. **European Journal Soil Science**, Chichester, v. 52, p.331-340. 2001.

WEBSTER, R.; OLIVER, M. A. **Geostatistics for environmental scientists**. 2nd ed. Chichester: John Willey & Sons, 2007. 315 p.

WHITE, R. E. **Princípios e práticas da ciência do solo: o solo como um recurso natural**. São Paulo: Andrei, 2009. 426 p.

WILSON, J. P.; GALLANT, J. C. Digital Terrain Analysis. In: WILSON, J. P.; GALLANT, J. C. (Ed.) **Terrain analysis: principles and applications**. New York: John Wiley & Sons, 2000. Cap. 1, p. 1-27.

WINTER, B. **Linear models and linear mixed effects models in R with linguistic applications**. Birmingham, 2013. 42 p. Disponível em: <<http://arxiv.org/pdf/1308.5499.pdf>>. Acesso em: 5 nov. 2015.

YAMAMOTO, J. K.; LANDIM, P. M. B. **Geoestatística: conceitos e aplicações**. São Paulo: Oficina de textos, 2013. 215 p.

YIGINI, Y.; PANAGOS, P. Reference area method for mapping soil organic carbon content at regional scale. **Procedia Earth and Planetary Science**, Amsterdam, v. 10, p. 330-338, 2014.

ZANZARINI, F. V. et al. Correlação espacial do índice de vegetação (NDVI) de imagem Landsat/ETM+ com atributos do solo. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v. 17, n. 6, p. 608-614, 2013.

ZEWDU, S.; SURYABHAGAVAN, K. V.; BALAKRISHNAN, M. Geo-spatial approach for soil salinity mapping in Sege Irrigation Farm, South Ethiopia. **Journal of the Saudi Society of Agricultural Sciences**, Riyadh, 2015. In press. DOI: <<http://dx.doi.org/10.1016/j.jssas.2014.12.003>>

ZHU, A. X. et al. Construction of membership functions for predictive soil mapping under fuzzy logic. **Geoderma**, Amsterdam, v. 155, n.3-4, p. 164-174. 2010.

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

ZHU, Q.; LIN, H. S. Comparing ordinary kriging and regression kriging for soil properties in contrasting landscapes, **Pedosphere**, Amsterdam, v. 20, n. 5, p. 594-606, 2010.

ZIMBACK, C. R. L. **Curso de atualização em solos: geoprocessamento, levantamento e mapeamento de solos**. Botucatu: Universidade Estadual Paulista, 2007. 72 p.

ZIMBACK, C. R. L. A Geologia do Município de Botucatu. In: SILVA, R. B.; ORSI, A. C.; CHINELATO, F. C. S. **Lavapés, água e vida: nos caminhos da educação ambiental**. Botucatu: Sabesp, 2008.

ZIMBACK, C. R. L. **Análise espacial de atributos químicos de solos para fins de mapeamento da fertilidade do solo**. 2001. 114 f. Tese (Livre Docência) - Faculdade de Ciências Agronômicas, Universidade Estadual Paulista, Botucatu, 2001.