

**UNIVERSIDADE ESTADUAL PAULISTA – UNESP
CÂMPUS DE JABOTICABAL**

**PLANEJAMENTO AMOSTRAL PARA MAPEAMENTO DE
ARGILA DO SOLO COM COKRIGAGEM E
SUSCETIBILIDADE MAGNÉTICA**

Cristiano de Souza Marchesi

Engenheiro Agrônomo

2018

**UNIVERSIDADE ESTADUAL PAULISTA – UNESP
CÂMPUS DE JABOTICABAL**

**PLANEJAMENTO AMOSTRAL PARA MAPEAMENTO DE
ARGILA DO SOLO COM COKRIGAGEM E
SUSCETIBILIDADE MAGNÉTICA**

Cristiano de Souza Marchesi

Orientador: Prof. Dr. Gener Tadeu Pereira

Coorientador: Prof. Dr. Daniel De Bortoli Teixeira

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Doutor em Agronomia (Ciência do Solo).

2018

M315p Marchesi, Cristiano de Souza
Planejamento amostral para mapeamento de argila do solo com
cokrigagem e suscetibilidade magnética / Cristiano de Souza
Marchesi. -- Jaboticabal, 2018
xvii, 102 p. : il. ; 29 cm

Tese (doutorado) - Universidade Estadual Paulista, Faculdade de
Ciências Agrárias e Veterinárias, 2018

Orientador: Gener Tadeu Pereira

Coorientador: Daniel De Bortoli Teixeira

Banca examinadora: Rafael Montanari, Alan Rodrigo Panosso,
Diogo Mazza Barbieri, Marcílio Vieira Martins Filho

Bibliografia

1. Pedometria. 2. Geoestatística. 3. Proporção amostral. 4.
Validação. 5. Acurácia. 6. Mapa de qualidade. I. Título. II. Jaboticabal-
Faculdade de Ciências Agrárias e Veterinárias.

CDU 631.43

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: PLANEJAMENTO AMOSTRAL PARA MAPEAMENTO DE ARGILA DO SOLO COM COKRIGAGEM E SUSCETIBILIDADE MAGNÉTICA

AUTOR: CRISTIANO DE SOUZA MARCHESI

ORIENTADOR: GENER TADEU PEREIRA

COORIENTADOR: DANIEL DE BORTOLI TEIXEIRA

Aprovado como parte das exigências para obtenção do Título de Doutor em AGRONOMIA (CIÊNCIA DO SOLO), pela Comissão Examinadora:


Prof. Dr. GENER TADEU PEREIRA
Departamento de Ciências Exatas / FCAV / UNESP - Jaboticabal


Prof. Dr. RAFAEL MONTANARI (Videoconferência)
Departamento de Fitossanidade, Engenharia Rural e Solos-FEIS / Ilha Solteira/SP


Prof. Dr. ALAN RODRIGO PANOSSO
Departamento de Ciências Exatas / FCAV-UNESP/Jaboticabal-SP


Dr. DIOGO MAZZA BARBIERI
Athenas Consultoria Agrícola e Laboratório / Jaboticabal/SP


Prof. Dr. MARCÍLIO VIEIRA MARTINS FILHO
Departamento de Solos e Adubos / FCAV / UNESP - Jaboticabal

Jaboticabal, 10 de setembro de 2018

DADOS CURRICULARES DO AUTOR

CRISTIANO DE SOUZA MARCHESI – Filho de Florivaldo Aparecido Marchesi e Sônia Maria de Souza Marchesi, nasceu em Ribeirão Preto – SP no dia 11 de janeiro de 1978. Em 2003 ingressou no Curso de Agronomia pela Universidade do Estado de Mato Grosso – Campus de Cáceres – MT. Foi bolsista de iniciação científica na graduação no período de 2005 a 2006. Foi bolsista monitor da disciplina de Fisiologia Vegetal na graduação no período de 2006 a 2007. Em março de 2008, ingressou no serviço público federal ocupando o cargo de Engenheiro Agrônomo no Instituto Federal de Mato Grosso – Campus Cáceres – MT. Em 2009 obteve o título de Especialista em Fertilidade do Solo e Nutrição de Plantas no Agronegócio pela Universidade Federal de Lavras – Lavras – MG. Em 2013 obteve o título de Mestre em Ciências (Educação Agrícola) pela Universidade Federal Rural do Rio de Janeiro – Campus de Seropédica - RJ. Em 2015 ingressou no Curso de Doutorado em Agronomia (Ciência do Solo) pela Universidade Estadual Paulista – Faculdade de Ciências Agrárias e Veterinárias – Campus de Jaboticabal – SP. Foi bolsista da Fundação de Amparo à Pesquisa do Estado de Mato Grosso a partir de 2016 no doutorado. No dia 10 de setembro de 2018 submeteu-se à banca para a defesa de Tese, sendo aprovado como Doutor em Agronomia (Ciência do Solo).

Epígrafe

“A poesia está guardada nas palavras – é tudo que eu sei.
Meu fado é o de não saber quase tudo.
Sobre o nada eu tenho profundidades.
Não tenho conexões com a realidade.
Poderoso para mim não é aquele que descobre ouro.
Para mim poderoso é aquele que descobre as insignificâncias (do mundo e as nossas).
Por essa pequena sentença me elogiaram de imbecil.
Fiquei emocionado.
Sou fraco para elogios.”

Do sábio poeta pantaneiro: Manoel de Barros (2001)

Dedicatória

A todas as pessoas que passam por esta vida carregando consigo a vontade de fazer o bem não apenas para si (e familiares, amigos ou conhecidos), mas também para as demais pessoas; sem esperar por recompensas.

AGRADECIMENTOS

A Deus por ter me oportunizado fazer parte da existência.

A Jesus Cristo, por ser o caminho, a verdade e a luz.

Ao Espírito Santo, por inspirar e preencher.

Aos meus pais, Florivaldo Aparecido Marchesi (Nino) (*in memoriam*) e Sônia Maria de Souza Marchesi, por me ensinarem como é viver com integridade, retidão, dignidade (trabalho), humildade, amor, esperança e fé. Ou seja, nunca deixaram faltar em momento algum as coisas que são de fato essenciais para um filho.

A meu avô Maurício Marchesi (*in memoriam*), por me tratar como um filho (reforçar o que meus pais fizeram por mim).

A meu irmão Juliano de Souza Marchesi (*in memoriam*), por ter sido uma pessoa permanentemente solidária (não media esforços para ajudar seus familiares, amigos, colegas, conhecidos e mesmo pessoas que nem sequer conhecia).

A meu irmão Reinaldo de Souza Marchesi, por mostrar inúmeras vezes que com o bom uso da inteligência e estudo é possível ir bem mais longe do que a maioria das pessoas acreditam.

A minha esposa Valdira A. dos S. Marchesi, por ser mãe dos meus filhos (Samuel, Alanis e George) e ser uma pessoa extremamente doce, amorosa, companheira e leal. Ou seja, por ser uma esposa que merece ser sempre amada.

Aos meus filhos, por me trazerem muita felicidade (virem ao mundo) em momentos que a tristeza (falecimento do meu pai e do meu irmão) muito me afetou.

A meu sogro (Waldeson), sogra (dona Aparecida) e cunhados (Valéria e Júnior) pelo acolhimento, apoio, torcida e amizade. Em especial a minha sogra, por não medir esforços em ajudar no cuidado dos meus filhos.

A meu tio Sérgio e tia Neide Aparecida de Souza Lehfeld, pelo acolhimento, carinho, torcida, apoio, amizade e generosidade (são pessoas muito boas, me tratam como se fosse um filho).

A meu tio Jader de Souza Júnior e tia Conceição, pela receptividade, carinho, torcida, apoio, amizade e sempre receber bem a minha mãe.

Aos meus primos Lucas Lehfeld, Leandro Lehfeld e Benjamim Monteiro, pela receptividade, atenção (boas conversas), carinho, torcida, apoio e amizade.

Aos professores Gener Tadeu Pereira e Daniel De Bortoli Teixeira, pela orientação, confiança, amizade, incentivo e exemplo de profissionalismo (profundos conhecedores no que fazem e extremamente humildes).

Aos doutores que participaram da Banca de Qualificação, Diego Silva Siqueira e Angélica S. R. de S. Bahia, pelas valiosas contribuições para melhora do presente trabalho; incentivo, amizade e exemplo de profissionalismo (extremamente dedicados, conhecedores e humildes).

Aos doutores que fizeram parte da Comissão Examinadora (Banca de Defesa) Rafael Montanari, Alan Rodrigo Panosso, Diogo Mazza Barbieri, Marcílio Vieira Martins Filho, por se disponibilizarem para a troca de experiências, mostrar seus pontos de vistas sobre a temática (crítico, sugerir, elogiar, etc.) que, por fim, ajudaram muito na melhora do presente trabalho e da minha formação.

Ao coordenador do curso de Pós-Graduação em Agronomia (Ciência do Solo) da UNESP de Jaboticabal, professor José Marques Júnior, e demais profissionais que fizeram parte da equipe de seleção de alunos para ingressarem no Programa no primeiro semestre de 2015, por com atenção me ouvirem, e com total imparcialidade (ninguém me conhecia) me selecionarem para ser aluno do Programa.

A todos os professores da UNESP de Jaboticabal que contribuíram com minha formação (ministraram disciplina ou parte de disciplina que fui matriculado): Célia Regina, Itamar, Marques Júnior, Diego Siqueira, Barbosa, Gener, Daniel Teixeira, Furlani, Teresa Cristina, Renata Costa, Peruzzi, Coutinho, Mara Cristina, Manoel Evaristo, Wanderley, Jairo, Gabriel Maurício.

A todos meus professores da graduação, especialização e mestrado, que contribuíram com minha formação profissional; em especial a Eurípedes Maximiano Arantes, Cassiano Cremon, Dana Kátia Meschede, Reginaldo Antônio Medeiros e Everaldo Zonta, pela confiança, incentivo, amizade e exemplo de profissionalismo (aulas de excelentíssima qualidade).

A todos meus professores do ensino fundamental e médio, por me proporcionarem a base necessária pra minha formação.

As professoras e funcionárias do Centro de Formação Infantil na UNESP de Jaboticabal, pelo acolhimento carinhoso e educacional do meu filho Samuel.

A Richard Webster (pesquisador da Rothamsted, Harpenden, Reino Unido), por gentilmente me fornecer valiosas informações (materiais) e aconselhamentos sobre Geoestatística aplicada à Ciência do Solo.

Aos conhecidos, colegas, amigos e demais pessoas que torcem pelo meu desenvolvimento humano, intelectual e profissional. Em especial aos amigos Flávio de Oliveira, Sinésio, Carlos Alberto, Moacir Horn, Luciano Romano, Ney Ataíde, Ighor Raphael, Minelson, Arildo Alves, Fernando Melo, que ao longo do meu afastamento frequentemente me perguntavam “já concluiu o doutorado? Já está voltando?”. E também a Thiago Águas, Katharine Batista, Hygor Siqueira, Lucas Girio, Antonio Márcio, Leonardo Correia, Thyago Lira, Alysson Jalles, Denise Delarica, Raphael Ferreira, Danilo Olandino, Rodolfo Lizcano, Robson Medeiros, Joel Cabral, Elienay Ferreira, Marcilene Sarah, Dilier Olivera, enfim, colegas de Pós-Graduação que se tornaram amigos.

Ao Instituto Federal de Mato Grosso (IFMT), pela concessão do afastamento para o doutorado para meu aprimoramento profissional e como Servidor Público.

A UNESP de Jaboticabal pelo ensino público de excelentíssima qualidade oferecido.

A Fundação de Amparo à Pesquisa do Estado de Mato Grosso, pela bolsa de doutorado concedida para esta pesquisa (processos nº 357174/2016 referente ao edital nº 009/2016).

Ao grupo de pesquisa Caracterização do Solo Para Fins de Manejo Específico (CSME), pela base material, técnica e humana (contato com profissionais extremamente gabaritados em Pedometria).

A Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo financiamento do projeto de pesquisa (processo nº 2013/17552-6) gerador dos dados.

SUMÁRIO

	Página
RESUMO	xiii
ABSTRACT	xiv
LISTA DE FIGURAS	xv
LISTA DE TABELAS	xvii
1 INTRODUÇÃO	1
2 REVISÃO DE LITERATURA	6
2.1 Variabilidade espacial do solo	6
2.2 Fatores que produzem variabilidade espacial no solo e abordagens	7
2.3 A interpolação para o mapeamento de área-propriedades do solo	13
2.3.1 Geoestatística	13
2.3.1.1 Processos aleatórios	16
2.3.1.2 Estacionariedade	17
2.3.1.3 O variograma	19
2.3.1.3.1 Modelos para o variograma e modelagem	21
2.3.1.3.2 Interpretação variográfica	25
2.3.1.3.3 Amostragem	27
2.3.1.3.3.1 Escala espacial (espaçamento amostral)	27
2.3.1.3.3.2 Configuração amostral	29
2.3.1.3.3.3 “Relaxamento” amostral na cokrigagem	30
2.3.1.3.4 Fatores que afetam a variografia	32
2.3.1.3.4.1 Número de pontos de amostragem e distância de separação	32
2.3.1.3.4.2 Distribuição marginal dos dados (assimetria e outliers)	33
2.3.1.3.4.3 Tendência nos dados	35
2.3.1.3.4.4 Intervalo de lag e abertura angular	36
2.3.1.3.4.5 Anisotropia	37
2.3.1.4 Interpolação geoestatística	38
2.3.1.4.1 Krigagem	38
2.3.1.4.1.1 Krigagem ordinária	39
2.3.1.4.1.2 Cokrigagem ordinária	42

2.3.1.5 Qualidade dos mapas	44
2.3.1.5.1 Validação	46
2.3.1.5.1.1 Validação-cruzada e Validação-externa	48
2.3.2 Argila do solo – uma variável estratégica para o manejo localizado	52
2.3.3 Suscetibilidade magnética e utilização no estudo dos atributos do solo	54
3 MATERIAL E MÉTODOS	60
3.1 Descrição da área de estudo	60
3.2 Amostragem de solo na área de estudo	61
3.3 Análises laboratoriais	62
3.4 Construção dos cenários de amostragem e conjunto para validação	63
3.5 Estatísticas descritivas	65
3.6 Análise geoestatística	66
4 RESULTADOS E DISCUSSÃO	68
4.1 Estatísticas descritivas	68
4.2 Análise variográfica	69
4.3 Interpolação (mapas)	73
4.3.1 Validação dos mapas (acurácia)	78
5 CONCLUSÕES	83
6 REFERÊNCIAS	84
7 CONSIDERAÇÕES FINAIS	101

PLANEJAMENTO AMOSTRAL PARA MAPEAMENTO DE ARGILA DO SOLO COM COKRIGAGEM E SUSCETIBILIDADE MAGNÉTICA

RESUMO – O mapeamento acurado dos atributos do solo apresenta elevado custo, muitas vezes inviabilizador para uso agrícola, principalmente pelo grande número de amostras normalmente exigidas para a coleta e análises laboratoriais. Neste sentido, é necessário o aprofundamento de pesquisas envolvendo metodologias e novos conhecimentos que têm se mostrado eficientes no propósito de geração de mapas de solos com qualidade e possibilidade de redução de custo. Com este propósito, avaliou-se o mapeamento da argila do solo com auxílio da SM no procedimento de interpolação por cokrigagem ordinária (CKO). O objetivo foi usar o método de CKO com o auxílio da SM (medida analítica obtida por um custo relativamente mais baixo do que a da argila) para reduzir a densidade de amostragem da argila e gerar mapas mais acurados. Para tal, usou-se o procedimento de validação com uso de dados independentes (10% dos dados) para checagem dos erros de predição (RMSE) que permitiu avaliar a acurácia dos mapas. A área estudada é um campo agrícola canavieiro que possui 870 ha localizado em Guataparã-SP. Nesta área medidas de SM do solo e do teor de argila do mesmo foram obtidas de 371 amostras coletadas na profundidade de 0-0,25 m de uma malha com pontos separados por distâncias mínimas variando de 145 m a 174 m. Para avaliar a redução da densidade de amostragem da argila, testou-se 14 cenários de diferentes proporções amostrais das duas variáveis. A amostragem da argila correspondeu às densidades de 1 ponto a cada 2,6; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14 e 15 ha, que, em cada uma se incorporou a SM amostrada em 1 ponto a cada 2,6 ha. Os resultados mostraram que esta incorporação proporcionou aumento na acurácia das estimativas em 6 dos 14 cenários estudados. O cenário que apresentou a melhor performance, proporcionou 7,46% de aumento na acurácia com 63% de redução na densidade de amostragem para a argila. Os resultados deste estudo evidenciam o potencial da SM como uma covariável auxiliar junto com o método de CKO no mapeamento da argila com melhor acurácia e considerável economia amostral da mesma.

Palavras-chave: Pedometria, geoestatística, proporção amostral, validação, acurácia, mapa de qualidade.

SAMPLING PLANNING FOR SOIL CLAY MAPPING WITH COKRIGING AND MAGNETIC SUSCEPTIBILITY

ABSTRACT – The accurate mapping of soil attributes presents a high cost, often unfeasible for agricultural use, mainly due to the large number of samples normally required for the collection and laboratory tests. In this sense, it is necessary to deepen research involving methodologies and new knowledge that have proven efficient in the purpose of generating soil maps with quality and possibility of cost reduction. For this purpose, the soil clay mapping was evaluated using magnetic susceptibility (MS) in the ordinary cokriging (OCK) interpolation procedure. The objective was to use the OCK method with the help of MS (analytical measurement obtained at a relatively lower cost than that of clay) to reduce the sampling density of the clay and generate more accurate maps. For this, the validation procedure was used with independent data (10% of the data) to verify the prediction errors (RMSE) that allowed to evaluate the accuracy of the maps. The studied area is a sugar cane field that has 870 ha located in Guatapar-SP. In this area measurements of soil MS and clay content were obtained from 371 samples collected at 0-0.25 m depth of a mesh with points separated by minimum distances varying from 145 m to 174 m. To evaluate the reduction of clay sampling density, 14 scenarios of different sample proportions of the two variables were tested. Sampling of the clay corresponded to densities of 1 point every 2,6; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14 and 15 ha, that in each one was incorporated the sampled MS at 1 point every 2.6 ha. The results showed that this incorporation provided an increase in the accuracy of the estimates in 6 of the 14 scenarios studied. The scenario that presented the best performance, provided a 7.46% increase in accuracy with a 63% reduction in the sampling density for clay. The results of this study evidenced the potential of MS as an auxiliary covariate along with the OCK method in the mapping of clay with better accuracy and considerable sample savings.

Keywords: Pedometry, geostatistics, sample proportion, validation, accuracy, quality map.

LISTA DE FIGURAS

	Página
Figura 1. Diagrama esquemático (em forma de perfis) de possíveis escalas de observação associadas ao domínio (apenas hipotético) de fatores que influenciam a variabilidade espacial de propriedades do solo. Diagrama A: geologia da região; diagrama B: relevo local; diagrama C: elemento da vertente (e nele, microrelevos). Baseado em Burrough (1993)	8
Figura 2. Esquematização visual do comportamento espacial (diagramas A, B e C) típicos ao que os principais modelos de variograma (gaussiano, esférico e exponencial) matematicamente descrevem	23
Figura 3. Comportamentos possíveis do variograma. Gráfico A apresenta patamar; gráfico B não produz patamar; gráfico C apresenta efeito pepita puro. Adaptado de Oliver (2010)	25
Figura 4. Caracterização da área de estudo. Mapas da localização geográfica (país: Brasil; Estado: São Paulo; município: Guatapar) (A); Mapas da rea: (B) geolgico com escala de execuo 1: 500.000 (materiais de origem: SG, DCE e DA), (C) pedolgico com escala de execuo 1: 12.000 (solos dominantes: LVAd, LVd, LVdf, LVef e RQod) e (D) topogrfico (altitude de 519 a 649 m acima do nvel do mar; e os valores no interior do mapa marcam as altitudes intermedirias)	61
Figura 5. Malha de amostragem de solo na rea estudada contendo 371 pontos separados por distncias mnimas variando de 145 m a 174 m, georreferenciados, cobrindo seus 870 ha	62
Figura 6. Ilustrao esquemtica do procedimento de organizao dos dados para execuo do estudo	64

Figura 7. Malha de amostragem de solo na área estudada com a distribuição dos pontos usados na construção dos cenários de amostragem (argila do solo e SM) e conjunto de dados para validação.....	65
Figura 8. Variogramas experimentais estimados da suscetibilidade magnética (SM) e da argila (ARG) do solo e variogramas-cruzados destas para cada cenário amostral estudado	72
Figura 9. Mapas de krigagem da suscetibilidade magnética (SM) ($\times 10^{-8} \text{ m}^3 \text{ kg}^{-1}$) e da argila (ARG) (%) do solo e mapas de cokrigagem desta para cada cenário amostral estudado	77
Figura 10. Mapas de argila do solo (%) da área de estudo: mapa de referência ARG(2,6); mapa com maior ganho relativo de acurácia SM(2,6) \times ARG(7); mapa com perda relativa de acurácia SM(2,6) \times ARG(8) e mapa com maior perda relativa de acurácia SM(2,6) \times ARG(15)	81
Figura 11. Gráficos de quantil-quantil para distribuição de valores observados (verdadeiros) versus distribuição de valores preditos pela cokrigagem da argila (%) para cada cenário amostral estudado	82

LISTA DE TABELAS

	Página
Tabela 1. Estatística descritiva dos dados de suscetibilidade magnética (SM) ($\times 10^{-8}$ m ³ kg ⁻¹) e argila do solo (ARG) (%) para os cenários estudados	68
Tabela 2. Variogramas experimentais estimados da suscetibilidade magnética (SM) e argila (ARG) do solo e variogramas-cruzados destas para cada cenário amostral estudado	71
Tabela 3. Índices de acurácia (RMSE) e de melhoria relativa (MR _{RMSE}) calculados a partir da validação externa para avaliação dos mapas de cokrigagem da argila (%) do solo para cada cenário amostral de estudo.....	73

1 INTRODUÇÃO

Ao longo das últimas décadas no Brasil, a Agricultura de Precisão (ou agricultura de sítio-específico) evoluiu de ser simplesmente uma idéia para uma prática de gestão aceita para a agricultura moderna (Brasil, 2013). O desafio que se coloca diante da comunidade de pesquisa agrícola é identificar como utilizar melhor as tecnologias que ela dispõe em benefício da agricultura e sociedade em geral. Esse desafio inclui não apenas a perspectiva de aumentar a lucratividade agrícola e a produção de alimentos, fibras, etc., mas também aprimorar o uso e conservação dos solos utilizados, proteger melhor os recursos hídricos, melhorar o manejo envolvendo os insumos utilizados (nutrientes, corretivos, água, etc.).

A tecnologia de aplicação à taxa variável (ATV) é um elemento importante da Agricultura de Precisão (Robert, 1993; Pierce e Nowak, 1999; Whelan, 2018). Esta tecnologia permite variar a aplicação de insumos em um campo de modo econômico¹. A fertilização com taxa variável foi projetada para melhorar a rentabilidade da fazenda e a eficiência da gestão de nutrientes em comparação com a aplicação tradicional de fertilizantes à taxa uniforme, direcionando a aplicação de nutrientes para áreas com baixos teores de nutrientes (que podem estar sendo sub-fertilizadas) e reduzindo ou eliminando a aplicação em áreas com concentrações de nutrientes superiores às exigências da cultura e expectativa de produção (Robert, 1993; Pierce e Nowak, 1999; Bongiovanni e Lowenberg-Deboer, 2004). A qualidade da água também pode ter benefícios de um sistema de aplicação com ATV porque o excesso de aplicação de nutrientes em campos podem resultar em perdas de nutrientes para os recursos hídricos incorrendo na diminuição da qualidade da água. As perdas de nutrientes de nossas áreas agrícolas, mesmo que sutis, estão diretamente relacionadas à concentração dos nutrientes, outros elementos químicos presentes, e características dos solos (Bernardi et al., 2015). Assim, reduzir ou eliminar a aplicação de fertilizantes em áreas de campo que já possuem níveis

¹ Econômico num sentido não apenas *stricto* da palavra. Econômico em termos amplo. Ou seja, envolve desde a quantidade do insumo aplicado que deve ser adequada não apenas por motivos de produtividade, mas, também porque muitas vezes é um recurso limitado e não renovável, como por exemplo os fertilizantes minerais a base de fósforo (Cordell e White, 2013). Enfim, há uma relação com as externalidades. A denotação correta para a expressão "...de modo econômico" é que a ação empreendida foi (ou será) racional em seus vários aspectos. Isto normalmente remete, em termos de práticas agrícolas, ao que de fato é sustentável (FAO, 1989; Paterniani, 2001).

suficientes de nutrientes para a produção agrícola, reduzirão a perda potencial de nutrientes no campo e a degradação da qualidade da água, o que é extremamente importante em áreas agrícolas que se situam sobre aquíferos, e mais importante ainda, sobre um grande aquífero, como o aquífero Guaraní, que é considerado um dos mais importantes do mundo².

A presença de variabilidade espacial dos nutrientes e atributos do solo é documentada na literatura e é causada por fatores naturais (revisão sobre este assunto nos itens 2.1 e 2.2) e antrópicos³. Logo, para que os solos permaneçam ou se tornem produtivos nessas áreas, tendo que a atuação desses fatores e as intervenções humanas criam uma distribuição de condições de fertilidades ao longo das paisagens, estas condições necessariamente precisam ser adequadamente analisadas e manejadas (Cambardella e Karlen, 1999).

Embora a aplicação de nutrientes com ATV seja viável⁴ e já esteja sendo usada em diversas fazendas produtivas no Brasil conforme mostra Silva et al. (2011), os produtores rurais e profissionais envolvidos com a produção agrícola, de modo geral, sabem que o sucesso no investimento e uso da ATV depende fundamentalmente de mapas de solo que representem bem a variabilidade espacial dos atributos no campo (Pierce e Nowak, 1999; Oliver, 2010; Grego et al., 2014; Molin et al., 2015). Conscientes disso, a atenção por mapas representativos⁵ da variabilidade espacial de atributos do solo para a utilização com ATV tem crescido bastante nestes últimos anos. Todavia, pesquisas desenvolvidas com o suporte da

² O Aquífero Guaraní é um imenso corpo hídrico subterrâneo que abrange partes dos territórios do Uruguai, Argentina, Paraguai e, principalmente, Brasil, ocupando 1.200.000 km². No Brasil, oito Estados são abrangidos pelo Aquífero Guaraní. Suas reservas abastecem a maior parte das cidades do Oeste paulista. (Wikipedia – Aquífero Guaraní). Sobre esta região do aquífero está uma das áreas agrícolas mais produtivas do país.

³ Exemplo de ações antrópicas que podem produzir variabilidade espacial no solo: conversão de áreas florestadas em áreas de cultivo; aplicação direcional de água como na irrigação por sulco; determinadas operações com máquinas e tratores (enleiramentos, sulcagem larga e profunda, etc.); espécies cultivadas (plantas perenes); sistemas de cultivos (plantio direto, pomares, etc.); aplicações desuniformes (mal operacionalizadas) de calcário, fertilizantes e outros insumos; etc.

⁴ Miller et al. (1999) listam três critérios que ao passo que possam ser satisfeitos justificam o uso de ATV. Estes são: (1) que exista uma variabilidade espacial significativa dentro do campo em fatores que influenciam o rendimento das culturas, (2) que, esta variabilidade possa ser identificada e medida, e (3) que, as informações dessas medições possam ser usadas para modificar as práticas de manejo culturais para aumentar o lucro e/ou diminuir o impacto ambiental.

⁵ Mapa representativo deve ser entendido como sinônimo de mapa com boa qualidade ou alta acurácia, para isto a acurácia precisa ser conhecida (ver item 2.3.1.5). Pouco tem sido mencionado na literatura técnica, por exemplo, que a correspondência entre mapa de recomendação de insumos e mapa de produção (a “resposta”) somente poderá existir havendo representatividade. Ou seja, cabe a pergunta: a acurácia de tais mapas é semelhante?

Geoestatística nas décadas recentes para dar retaguarda a este contexto tecnológico têm mostrado tanto as possibilidades como as dificuldades para elaboração de mapas com boa qualidade (Resende et al., 2010; Teixeira, 2015). Os estudos nesse propósito têm apontado para a necessidade de malhas amostrais relativamente densas, envolvendo a coleta e análises laboratoriais de grande número de amostras, para que assim se consiga captar a estrutura da variação (dependência espacial) dos atributos do solo (Resende et al., 2006; Pereira et al., 2013) o que, conseqüentemente, pode assumir custos acima dos benefícios econômicos proporcionados pela ATV (Resende et al., 2010; Molin et al., 2015). Por esse motivo tem sido comum no âmbito comercial para a Agricultura de Precisão, orientações generalistas para o dimensionamento de malhas de amostragem (Gimenez e Zancanário, 2012; Molin et al., 2015) que não garantem que o quantitativo de dados seja de fato o adequado para elaboração de mapas de solo com boa qualidade.

No entanto, informações auxiliares quantitativas podem ser incorporadas nas predições geoestatísticas de atributos do solo menos densamente amostrados, com a possibilidade de assegurar a representatividade da variabilidade espacial do atributo para o mapa gerado (Webster e Oliver, 2007; Teixeira, 2015; Lark e Minasny, 2018). Para tal, as informações auxiliares necessariamente devem apresentar forte correlação espacial, ser convenientemente de mais fácil obtenção e apresentar menor custo para estimação em relação ao atributo em questão (Mcbratney e Webster, 1983a; Vieira, 2000; Lark e Minasny, 2018). Neste sentido a condutividade elétrica do solo (CE), suscetibilidade magnética (SM), cor, pH, topografia (atributos do relevo) são algumas das propriedades do solo promissoras para este fim porque reúnem as referidas condições. No entanto, é importante destacar que as medidas de campo desses atributos podem sofrer interferências de fatores no solo que devem ser levados em considerações, para que assim se possa definir contexto que se possa utilizá-las. Por exemplo, em alguns trabalhos estão relatados resultados obtidos com a CE que foram considerados insatisfatórios (Tarr et al., 2005; Sudduth et al., 2003). Diante disso é importante se atentar, por exemplo, que a CE apresenta instabilidade em solos cuja mineralogia é rica em óxidos (Wu et al., 2008) o que torna seu emprego em solos tropicais um recurso com pouca

possibilidade de sucesso. Por outro lado, SM, cor, pH, e algumas outras variáveis do solo não apresentam esta instabilidade, o que as tornam mais adequadas para utilização nas condições brasileiras (Siqueira e Marques Júnior, 2014). Como resultado disso, vários trabalhos mostram que a SM vem ganhando cada vez mais espaço nas pesquisas para esta finalidade (Rosa, 2014; Teixeira, 2015; Jordanova, 2017, Teixeira et al., 2017).

Entre os métodos de interpolação que possibilitam a incorporação de informações para auxiliar nas predições espaciais, a cokrigagem ordinária (CKO) é considerado o método mais acurado (Pawlowsky-Glahn e Olea, 2004). A CKO é considerada um avanço em relação aos métodos matemáticos de interpolação simples e também em relação à própria krigagem ordinária (KO, método geoestatístico mais utilizado), exatamente por esta capacidade de aproveitamento extra de informações espaciais correlacionadas (Minasny e McBratney, 2016). A correlação espacial entre os atributos explorada pela CKO, confere a mesma, capacidade preditiva mais acurada mesmo quando a variável a ser mapeada estiver subamostrada (Minasny e McBratney, 2016; Brevik et al., 2017; Lark e Minasny, 2018). As correlações espaciais entre as propriedades do solo são explicadas como resultantes de mesmo fenômeno ou conjunto de processos de formação do solo com atuação regionalizada (McBratney e Webster, 1983a). Assim, a utilização de uma covariável guardando tal característica com a variável a ser mapeada pode permitir diminuição do custo de amostragem com manutenção e/ou mesmo melhoria da acurácia do mapa elaborado (McBratney e Webster, 1983a; Vieira, 2000; Lark e Minasny, 2018). No entanto, a determinação de uma adequada densidade amostral dada pela proporção entre as variáveis é fator fundamental para efetivamente alcançar este objetivo.

Em decorrência da existência de poucas pesquisas realizadas sobre a amostragem para CKO (McBratney, 1992; Kerry et al., 2010) e do potencial de uso da SM como covariável para mapeamento de atributos do solo no Brasil, objetivou-se neste trabalho avaliar cenários de densidade de amostragem para o mapeamento do teor de argila do solo com o auxílio dessa covariável. A hipótese é que o erro de predição (acurácia) obtido pelo procedimento de validação – usando conjunto de

dados independentes – pode indicar a proporção amostral adequada para o mapeamento proposto.

A escolha pela utilização da referida covariável se deu em razão do que foi descrito. Quanto a decisão de se mapear o teor de argila da área utilizada para estudo, a escolha se deu em função de que é um dos atributos do solo considerado dos mais estratégicos para agronomia (Raij et al., 1996; Tomé Júnior, 1997; Lopes e Guilherme, 2000; Resende et al., 2007) fato é que existe orientações técnicas estabelecidas para certos insumos que suas quantidades aplicadas no campo (definição de doses) são ajustadas com base no teor de argila presente no mesmo (Alvarez V et al., 1999; Lopes e Guilherme, 2000; Sousa e Lobato, 2004). Isto coloca a argila do solo e sua possível variabilidade espacial como uma propriedade estratégica que pode ser útil para zonestar o manejo no campo com possibilidade de economia nos custos com obtenção de mapas. Lopes e Guilherme (2000) por exemplo, destacam as seguintes utilizações práticas envolvendo o teor de argila no campo (i) ajustar doses de adubação fosfatada corretiva quando esta é necessária; (ii) ajustar doses de calcário quando há necessidade de correção da acidez e reposição de bases; (iii) ajustar doses de gesso agrícola quando há necessidade de neutralização de alumínio fitotóxico em profundidade; (iv) decidir pelo número de parcelamento das adubações nitrogenadas e potássicas; (v) avaliar probabilidade de problemas com deficiência de micronutrientes para a cultura; (vi) avaliar possibilidade das plantas passarem por estresses hídricos e problemas com drenagem; etc. Neste sentido as informações sobre os teores de argila em campo (que há variabilidade espacial) podem ser utilizadas para operacionalização de práticas localizadas de manejo agrícola com o uso de tecnologia de ATV.

Conforme poderá ser visto, esta Tese está organizada na forma convencional o que inclui a apresentação da revisão de literatura no mesmo corpo antecedendo a parte na qual está relatado a pesquisa realizada. Nesta revisão é apresentada a teoria Geoestatística para modelar e mapear a variabilidade espacial de propriedades do solo dentro do contexto em que foi utilizada neste trabalho, bem como aspectos teóricos e metodológicos para avaliação dos procedimentos e qualidade de mapas. Antes disso, a variabilidade espacial do solo é apresentada em termos conceituais em conjunto com os fatores que teoricamente imprimem

variabilidade espacial no solo e as abordagens gerais para mapeá-la. Como parte final da revisão, são apresentadas mais detalhes sobre a importância das informações sobre a variabilidade espacial da argila e SM do solo para a agricultura moderna.

Embora tudo que está relatado neste trabalho siga o que é de comum para uma Tese (incluindo a obediência as normas), foi feito esforço em termos de redação, na medida do possível, escrever o assunto de forma também acessível a outros públicos que não necessariamente são de pessoas que lidam com a aplicação da Geoestatística. Para a brevidade de cada parte que compõe a Tese, e sem quebrar o fluxo da narrativa, algumas informações – obtidas ao longo das disciplinas cursadas no doutoramento – são fornecidas em forma de nota de rodapé no intuito de enriquecer a temática abordada. Caso o leitor preferir ler apenas os textos, sem a leitura das notas de rodapés, assim poderá fazer sem qualquer prejuízo com perda de sequência ou conectividade do assunto.

2 REVISÃO DE LITERATURA

2.1 Variabilidade espacial do solo

Uma característica intrínseca do solo é a variabilidade espacial de suas propriedades, que é também chamada de heterogeneidade espacial do solo⁶ (Jenny, 1941). Para o contexto científico atual, esta característica do solo apresenta componentes sistemáticos e aleatórios (Wilding e Drees, 1983; Burrough, 1993). O componente sistemático de variação é entendido como a aquelas mudanças espaciais nas propriedades do solo que podem ser explicadas pelos fatores de formação do solo (Jenny, 1941) em uma dada escala de observação; ou seja, originadas de diferenças na topografia, geologia, clima, atividade biológica e tempo relativo. Associado com o componente sistemático de variação espacial, estão diferenças observadas nas propriedades do solo que não permitem relacionar a

⁶ É importante mencionar que existem dois tipos inter-relacionados de heterogeneidade do solo: a heterogeneidade de constituição (as três fases do solo que podem ser constatadas ao se verificar uma porção dele) e a heterogeneidade de distribuição (a variabilidade espacial dos atributos do solo dentro de uma área, paisagem, etc.).

variação à uma causa conhecida. Wilding e Drees (1983) denominaram essa heterogeneidade inexplicável como variação “aleatória” ou “acidental”. Webster e Cuanalo (1975) e Burrough (1993) a denominaram como “ruído”.

É importante ressaltar, com o intuito de evitar equívocos, que apenas as variações que se enquadram como sistemática permite mapeamento. Esta é uma afirmação que é equivocada. De fato as variações “aleatórias” que ocorrem no solo – no sentido que acima foi explicado, ou seja, de estabelecimento de relação de “efeito-causa” – podem não se apresentar evidentemente correlacionadas a fatores causais; porém, isso não significa que tais variações, em termos de distribuição no espaço (estrutura espacial), não estejam correlacionadas e possam ser mapeadas.

2.2 Fatores que produzem variabilidade espacial no solo e abordagens

A variabilidade espacial das propriedades do solo é resultante da atuação dos fatores de formação do solo que operam e interagem entre si no tempo e no espaço (Jenny, 1941). Fatores que espacialmente operam em grandes extensões e por períodos de tempo mais prolongados (por exemplo, o clima, a geologia) tendem a ter seus efeitos modificados por outros fatores que imprimem variabilidade no solo mais localmente (por exemplo, o relevo, os organismos) (Burrough, 1993). Esta natureza aninhada de variabilidade no solo implica que o tipo e as fontes de heterogeneidade expressas pelas características das propriedades do solo tendem a depender, antes de tudo, da escala ou frequência de observação⁷ (Burrough, 1993), como em seguida será explicado e ilustrado na Figura 1. Mudanças na variabilidade espacial do solo com escala crescente depende da propriedade do solo em questão e dos fatores de formação do solo que determinam a mudança espacial (Wilding e Drees, 1983; Burrough, 1993).

Em diversos trabalhos está documentado que algumas propriedades do solo variam muito em curtas distâncias⁸ (Mcintyre, 1967; Protz et al., 1968; Beckett e

⁷ O conceito de escala pode ser confuso, na medida em que possui múltiplos referentes. Escala cartográfica refere-se ao tamanho representado de um recurso em um mapa em relação ao seu tamanho no mundo real. Escala de análise refere-se a frequência (densidade) de observação utilizada no levantamento de algo (solo, geologia, etc.), por exemplo, em um campo agrícola, em uma parte de um município, em uma região de um Estado, etc.

⁸ O teor de fósforo solúvel, o teor de nitrogênio, o conteúdo de matéria orgânica, etc., costumam variar bastante a curtas distâncias.

Webster, 1971; Webster e Butler, 1976; Goderya, 1998), enquanto o oposto é verdadeiro para outras propriedades do solo⁹ (Beckett e Webster, 1971; Webster e Butler, 1976; Goderya, 1998). A mudança pode ser linear ou não linear, onde diferentes processos de formação do solo exercem efeitos dominantes em diferentes escalas espaciais (Webster e Butler, 1976; Nortcliff, 1978; Burrough, 1993). A Figura 1 esquematiza em forma de perfis possíveis escalas de observações associadas ao domínio (apenas hipotético) de dados fatores que influenciam a variabilidade espacial de propriedades do solo, o que permite certas generalizações (tendências) sobre o fator dominante à determinar a estrutura de variação de propriedades do solo.

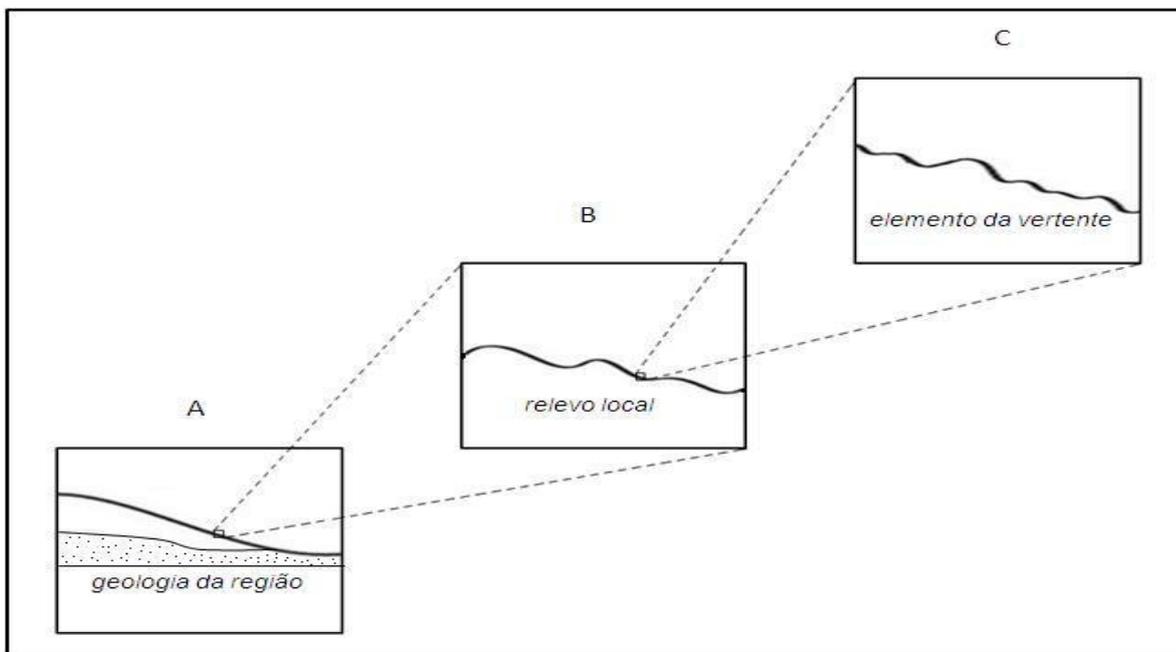


Figura 1. Diagrama esquemático (em forma de perfis) de possíveis escalas de observação associadas ao domínio (apenas hipotético) de fatores que influenciam a variabilidade espacial de propriedades do solo. Diagrama A: geologia da região; diagrama B: relevo local; diagrama C: elemento da vertente (e nele, micro-relevos). Baseado em Burrough (1993).

Conforme pode ser notado na Figura 1 dentro da escala de observação esquematizada no diagrama A, a geologia pode ser considerada o fator dominante à determinar a estrutura de variação das propriedades do solo em termos regional

⁹ O pH, a textura, a estrutura, a suscetibilidade magnética, etc., normalmente variam à média e longas distâncias.

(Ruhe, 1956; Ovalles e Collins, 1986). Já na escala de observação esquematizada no diagrama B (Figura 1), o relevo passa a ser o fator que determinará os padrões de variação das propriedades do solo nos diferentes seguimentos (elementos) da vertente (Brito et al., 2005; Schaetzi e Anderson, 2005). Enquanto que dentro da escala de observação representada no diagrama C (Figura 1) – dentro de um elemento (seguimento) da vertente – o fator que determinará as variações espaciais no solo pode não se apresentar evidentemente explícito e pode até mesmo já não mais dominar (Burrough, 1993). É importante mencionar que os fatores que normalmente operam em extensões equivalentes as unidades de observações A e B (Figura 1), podem definir também variações nas propriedades do solo mesmo a nível local (isto é, em uma parte específica da vertente). Segundo Yaalon (2003) isto pode ocorrer em função de que mudanças abruptas na topografia e na geologia também são possíveis de ocorrerem.

Na realidade as diferenças espaciais encontradas nas propriedades do solo são, em uma primeira análise, devidas à diferenças espaciais que podem estar presentes no próprio material de origem, na topografia (posição, forma do relevo, etc.), no clima local, na ocupação biológica (vegetal, antrópica, etc.) (Burrough, 1993; Jenny, 1994). O que significa que cada fator causal de variação do solo pode não operar independentemente mas sim como uma combinação (interação) com outros fatores de formação do solo (Burrough, 1983a, 1983b, 1993; Jenny, 1994). Embora haja uma tendência a pensar que a mudança gradual nas propriedades do solo pode ser a mais comum em uma escala de campo (dado seguimento da vertente), não se pode assumir esse comportamento como uma regra; pois, isso dependerá do tipo de paisagem e dos processos espaciais locais que estiveram ou estão ali em ação governados por múltiplos fatores (Jenny, 1994; Yaalon, 2003). Além do mais qualquer das propriedades do solo podem exibir variabilidade espacial em curtas distâncias, com mudanças até mesmo abruptas, resultantes por vezes de causas não explícitas como por exemplo dos próprios microrelevos da área (Robertson, 1987). Este autor explica que estas variações na superfície do solo, mesmo que de ocorrência na maioria das vezes de forma muito sutil (não perceptíveis a olhos desarmados), criam condições ambientais (em termos de microclima, orientação da drenagem, intensidade de lixiviação, estado de fertilidade,

atividade biológica, etc.) que durante um longo período de tempo podem promover mudanças em características espaciais de diversas propriedades do solo. Essa condição e outras mais (ver Nota de Rodapé nº 3) revelam a tamanha complexidade que é estudar a variabilidade do solo, seja com o intuito de atribuir possíveis fatores causais (inferências sobre pedogênese), ou mesmo quando o objetivo é usar os possíveis fatores causais (fatores de formação do solo) para auxiliar na predição espacial do solo¹⁰ (Webster, 2008).

Essas informações ora expostas podem até parecer não possuir relevância; no entanto, constituem exatamente o conjunto de informações que permitem um maior senso crítico a cerca das diferentes abordagens para o mapeamento da variabilidade espacial do solo. De modo geral, verifica-se na literatura a existência de dois tipos de abordagens para mapeamento da variabilidade espacial do solo: a abordagem que trata a variabilidade em termos de área-classes de solo, e a que trata a variabilidade em termos de área-propriedades do solo (Burrough, 1998; Heuvelink e Webster, 2001; Rossiter, 2005).

A maioria das informações relativas a variabilidade espacial do solo no território brasileiro foram obtidas pela abordagem área-classes de solo executada por Levantamentos Pedológicos¹¹ (IBGE, 2005; EMBRAPA, 2016; Ker et al., 2017) nos quais à uma unidade de mapeamento – uma área-classe – é atribuída uma classe de solo (ou um domínio de classe de solo ou de classes de solos). Para isso, *grosso modo*, os especialistas em solo utilizam do conhecimento existente de “paisagem de ocorrência”. Ou seja, do arcabouço de conhecimento sobre os tipos possíveis de classes de solo que mais frequentemente ocorrem em dadas partes da paisagem. Este conhecimento vem das relações já estabelecidas entre solo e paisagem com base no conceito de fatores de formação do solo; isto é, a influência

¹⁰ Intuindo enfatizar o quanto é importante o conhecimento sobre os fatores de formação do solo na rotina de execução do mapeamento tradicional de solo Smith (1986) afirmou: “...sabendo sobre os fatores que influenciam a natureza do solo, um pedólogo não tem que perfurar ao acaso e fazer uma grade de suas observações e, em seguida, traçar limites entre os pontos da grade. Isso não só reduz consideravelmente o tempo necessário para o mapeamento, mas aumenta muito a acurácia do mapeamento.” Sobre o mesmo com outras palavras Hudson (1992) afirmou: “Cientistas de solo treinados podem delinear corpos de solo com acurácia na paisagem, examinando diretamente menos de um milésimo do solo abaixo da superfície. Eles podem fazer isso por causa do validado modelo solo-paisagem. Um poderoso paradigma que permite que os cientistas do solo façam previsões muito acuradas...”.

¹¹ Levantamento Pedológico produz mapa de solo e mencionado em inúmeras publicações científicas como mapa convencional de solo, mapa tradicional de solo, mapa clássico de solo, etc. São mapas elaborados incluindo critérios taxonômicos.

conjunta das diferenças espaciais na topografia, geologia, clima, atividade biológica e tempo relativo (tempo presumido da atuação dos fatores anteriores) na formação do solo (Jenny, 1941; Smith, 1986; Hudson, 1992). Com esta interpretação feita auxiliada por certos recursos (por exemplo, uso de estereoscópicos) os polígonos de “solo” no mapeamento pedológico brasileiro puderam ser delineados com base nas variações espaciais percebidas nos fatores de formação (IBGE, 2005) ora explicados. A cada polígono delineado, informações do solo advinda da descrição de um ou mais perfil de solo foram atribuídas. Em conjunto com o referido procedimento de delimitação, checagem das próprias características do solo por intermédio de tradagens foram realizadas (IBGE, 2005) em diversos pontos da área para melhor definir os limites entre polígonos como uma espécie de verificação de similaridade e dissimilaridade das informações e características do solo obtidas das descrições e análises dos perfis mencionados. Consequentemente, à cada polígono do solo foi atribuído um nome que reúne as características extrapoladas das informações advindas da descrição do perfil (IBGE, 2005).

Esta abordagem ora exposta é chamada de abordagem de área-classe porque o nome dado para cada polígono corresponde a “uma” classe de solo supostamente dominante naquela área (polígono), correspondente a um sistema de classificação o qual foi utilizado para enquadrar as informações descritas e analisadas do perfil de solo aberto no interior da área delimitada para o propósito de representá-la (Burrough, 1998; Heuvelink e Webster, 2001; Rossiter, 2005). O nome das área-classes de solo no Brasil na atualidade correspondem ao Sistema Brasileiro de Classificação de Solo (Santos et al., 2006). Todavia, estas explicações feitas nesta revisão sobre a abordagem de mapeamento de área-classes de solo são apenas uma simplificação do procedimento para proporcionar ao leitor uma recapitulação sobre a ideia geral a respeito. Por este motivo é importante mencionar que tal procedimento, na realidade, se desdobra em etapas as quais envolvem parte de campo, parte de laboratório, atividade em escritório, enfim, é um procedimento com diversas facetas que, como temática, não fazem parte do escopo desta Tese abordar. Literaturas como IBGE (2005), Ker et al. (2012) e Curi et al. (2017) permitem ao leitor acesso a informações com um bom nível de detalhes sobre esse assunto.

A outra abordagem para mapeamento do solo é a de área-propriedades do solo (Burrough, 1998; Heuvelink e Webster, 2001; Rossiter, 2005). Esta abordagem trata a variabilidade do solo em termos de distribuição espacial das suas propriedades (atributos ou variáveis) o que significa dizer que aquilo que a abordagem tem por objetivo final é mapear a distribuição espacial de variáveis do solo (Webster, 1985; Rossiter, 2005). Na atualidade existem inúmeros recursos ou metodologias desenvolvidas ou adaptadas para atender tal abordagem. Sua principal diferença em relação a abordagem anterior é que a variabilidade espacial é investigada de modo quantitativo (Webster, 1985; Burrough, 1998; Heuvelink e Webster, 2001; Rossiter, 2005). Logo, características do próprio solo e relações matemáticas e estatísticas entre elas são a base do procedimento (Webster, 1985). Como será visto, esta Tese relata pesquisa desenvolvida usando desta abordagem com a aplicação de metodologias da Geoestatística, por tal, mais detalhes sobre a presente abordagem será tratada em itens seguintes.

Muito embora na atualidade exista sinergias reconhecidas entre a abordagem de mapeamento de área-classes e de área-propriedades do solo, por exemplo, tal que informações de levantamentos de solos existentes (mapeamento Pedológico) possam ser “atualizadas” utilizando métodos Pedométricos (relações matemáticas e outros recursos) (Kempen, 2011, Yang et al., 2011), ainda se encontram em fase embrionária e bastante distante da possibilidade efetiva de atender a agricultura brasileira em relação ao tipo de mapeamento de solo que os avanços tecnológicos passaram a exigir. Referente ao mapeamento de solo com integração das abordagens, assunto que não será tratado nesta Tese, referências como Mendonça-Santos e Manzatto (2007) são bastante esclarecedoras. Uma descrição clara sobre a evolução do conhecimento de solos no Brasil que inclusive contem discussão sobre os resultados práticos dessa integração de abordagens (chamada de Mapeamento Digital de Solo) é fornecida por Ker et al. (2017). Nos itens seguintes as informações apresentadas são sobre a abordagem de mapeamento de área-propriedades do solo utilizando a Geoestatística.

2.3 A interpolação para o mapeamento de área-propriedades do solo

2.3.1 Geoestatística

Conforme foi abordado, o solo cobre parte da Terra mais ou menos continuamente devido a enorme complexidade resultante da interação dos seus fatores de formação (Webster e Oliver, 2007). Conseqüentemente, é possível afirmar que nenhuma descrição de suas propriedades, por criteriosa que seja, medidas diretamente do próprio solo, pode seguir todas as variações ou detalhes (Webster, 2000); mesmo porque as descrições das propriedades do solo em termos de distribuição espacial, normalmente são realizadas por intermédio de observações fragmentárias. Ou seja, estas descrições derivam de observações (análises) provenientes de pequenos volumes de material do solo (amostras com alguns centímetros cúbicos) amplamente separados um dos outros (conjunto de amostras discretas que esparsamente cobrem a área em questão) (Webster e Lark, 2013). Portanto, qualquer representação quantitativa do que espacialmente é o solo, envolve a simplificação (amostragem) e inferências matemáticas e estatísticas (interpolação) de como o solo é entre os pontos de amostragem (Webster e Oliver, 2007). É importante destacar que esse é o procedimento padrão e é realizado dessa forma porque não é prático e nem normalmente viável fazer observações exaustivas do mundo real, principalmente das propriedades do solo, por mais pequena que a área seja, de modo que a amostragem¹² e a interpolação¹³ são consideradas imprescindíveis se o objetivo é conhece-lo espacialmente (mapeá-lo) (Webster e Lark, 2013; Webster e Oliver, 2007). Portanto, para interpolar os valores de variáveis em locais não amostrados, as informações de variabilidade que podem ser capturadas pelo conjunto de pontos amostrados, devem ser usadas (Burgess e Webster, 1980a; Oliver, 1987).

Segundo Burrough e McDonnell (1998) uma questão fundamental na abordagem de mapeamento de área-propriedades do solo é a escolha do método de

¹² Procedimento que se apropria de porção ou porções reduzida(s) de um corpo ou população infinitamente maior, que resume em si mesma determinadas características do todo (Arioli e Andriotti, 2007; Webster e Lark, 2013).

¹³ Procedimento de se “converter” observações pontuais em dados espacialmente contínuos. Essa transição para um campo geográfico requer previsão espacial por meio de cálculos (Li e Heap, 2014).

interpolação. De acordo com esses pesquisadores isso é especialmente crítico quando está em questão os custos e a qualidade dos mapas a serem elaborados. Eles destacaram que, quando os dados são abundantes, a maioria dos métodos de interpolação proporcionam resultados semelhantes. Porém, quando os dados são “escassos”, as suposições subjacentes sobre a variação entre os pontos amostrados podem diferir bastante; com isso a escolha do método de interpolação e as possibilidades proporcionadas com o conjunto de dados (densidade de amostragem) são questões consideradas cruciais (Li e Heap, 2014).

O desenvolvimento recente na teoria estatística permitiu que a dependência espacial das propriedades do solo (estrutura espacial de suas variações) fosse estimada e, diretamente considerada para a interpolação de amostras (Webster e Oliver, 2007). Essa possibilidade agregou ao conjunto de procedimentos de interpolação espacial, e aos procedimentos para amostragem, um desenvolvimento científico considerado revolucionário para o entendimento espacial da variabilidade¹⁴ do solo e de suas propriedades (Webster, 1985; Heuvelink e Webster, 2001). Pode-se considerar que esse desenvolvimento se deu graças à formulação da Teoria das Variáveis Regionalizadas por Matheron¹⁵ em 1963 (Webster, 1985).

Segundo Matheron, uma variável é considerada uma variável regionalizada (VR) se ela varia de um lugar para outro com continuidade (Webster, 1985). É exatamente essa condição que faz com que a VR não possa ser representada adequadamente por funções estatísticas tradicionais (Davis, 1973). A Teoria das VR foi desenvolvida a partir de idéias empíricas de Krige¹⁶ aplicadas na exploração

¹⁴ Segundo alguns estudiosos, por exemplo Virdee e Kottegoda (1984), a inclusão da dependência espacial no procedimento de interpolação não representou avanço em termos de teoria estatística, mas sim um avanço em relação a aplicação da estatística para o contexto em que está sendo utilizado.

¹⁵ Matemático, geólogo, pesquisador e professor francês, Georges François Paul Marie Matheron (Wikipedia – Georges Matheron) mostrou uma capacidade excepcional de formalizar problemas práticos, editar conceitos e métodos viáveis que na atualidade permitem a compreensão quantitativa da distribuição espacial de variáveis como as do solo, do ar, de substratos geológicos, etc. Logo que Matheron iniciou seus trabalhos no Serviço Geológico da França na África, descobriu o trabalho pioneiro da escola sul-africana desenvolvido por Krige nas jazidas de ouro do Witwatersrand, o qual permitiu formalizar os principais conceitos de sua teoria, a Teoria das Variáveis Regionalizadas, que hoje é chamada de Geoestatística (Rivoirard, 2005).

¹⁶ Estatístico, engenheiro de minas, pesquisador e professor sul-africano, Danie Gerhardus Krige (Wikipedia – Danie G. Krige) foi um homem de ideias práticas revolucionárias. Aplicou estatística e matemática na avaliação espacial de recursos minerais de maneira totalmente inovadora (Krige, 1951). Tamanho foi sua importância para o desenvolvimento do conjunto de métodos que compõe a Geoestatística, que a “krigagem” tem tal nome em homenagem a ele. O envolvimento de Krige no ramo da mineração é considerado um “divisor de águas” porque suas ideias deram base para métodos de estimativas que dão maior previsibilidade aos projetos de exploração mineral e que reduzem riscos financeiros no negócio (Braham, 2013).

mineral de ouro na República da África do Sul (Braham, 2013). Ele sugeriu que a estimativa espacial do teor de ouro poderia ser melhorada, levando em consideração o grau de similaridade, ou autocorrelação, entre as amostras. A Teoria das VR leva em consideração as características aleatórias e estruturadas das variáveis espacialmente distribuídas para fornecer ferramentas quantitativas para sua descrição ideal e estimativa não tendenciosa (Journel e Huijbregts, 1978). Esta teoria agora é a base de procedimentos para análise e estimativa de variáveis que o comportamento é espacialmente dependente. Esses procedimentos são conhecidos coletivamente como Geoestatísticas (Journel e Huijbregts, 1978).

A princípio, a Geoestatística, como foi dito, foi desenvolvida para o setor de mineração (Matheron, 1963). Na área de mineração, a Geoestatística tem sido muito útil para engenheiros e geólogos estudarem a distribuição espacial de propriedades importantes, como grau ou espessura ou acumulação de depósitos minerais (Journel e Huijbregts, 1978). No entanto, pelo modo de sua concepção aos propósitos que se visava atender, alguns pedólogos no final da década de 1970 viram na Geoestatística a possibilidade de adaptá-la a abordagem de mapeamento de área-propriedades do solo, ou seja, para a caracterização e mapeamento da variabilidade espacial de atributos do solo (Webster e Oliver, 2007; Lark e Minasny, 2018). Campbell (1978) e Burges e Webster (1980a) têm o mérito desse pioneirismo.

Campbell foi o primeiro a usar o variograma para medir (caracterizar) a variação espacial do solo, e, Burges e Webster foram os primeiros a usarem a krigagem para mapear (estimar) a variação do solo. Um artigo recente de Webster (2015) relata como a geostatística passou a fazer parte da história da Ciência do Solo, com riqueza de detalhes.

A fim de oferecer suporte teórico as metodologias geoestatísticas empregadas no estudo relatado nesta Tese, parte dos itens seguintes foram escritos com esse objetivo em forma de revisão de literatura. Por isso, não apresentam a teoria completa da Geoestatística. Referências para os leitores que pretendem informações com abundância de detalhes sobre a Geoestatística são, por exemplo, Journel e Huijbregts (1978), Cressie (1993), Yamamoto e Landim (2013), e, com aplicação à Ciência do Solo: Goovaerts (1997), Vieira (2000), Webster e Oliver

(2007) os quais são materiais clássicos escritos por cientistas mundialmente renomados.

2.3.1.1 Processos aleatórios

Assim como muitas outras características do ambiente, o solo é o produto da interação de muitos processos físicos, químicos e biológicos, conforme anteriormente foi explicado. O resultado por isso deve ser determinístico. No entanto, tal é a complexidade dos fatores em combinação, a sua variação ao longo do tempo, e a incompletude do nosso conhecimento, que o resultado, o solo, parece-nos como se fosse aleatório (Webster, 2000; Webster e Oliver, 2007).

Esses processos formadores podem até serem determinados fisicamente, mas as suas interações se apresentam tão complexas que as variações imprimidas no solo se apresentam como se fossem aleatórias (Webster, 2000). Isso significa dizer que uma solução determinística ou matemática para quantificar ou registrar as variações que ocorrem no mesmo, está totalmente fora de cogitação no momento (Webster, 2000). A solução lógica para lidar com a “aleatoriedade” na distribuição espacial dos fenômenos naturais, exigiu um salto de imaginação, conforme pode ser visto na Tese seminal de Matheron (Matheron, 1965).

Traduzindo essa ideia de propriedade aleatória em uma matemática, que é chamada de um processo aleatório, a seguinte anotação que se tornou convencional se faz necessária para o entendimento (Oliver e Webster, 2014):

1. O valor de uma propriedade, chamada de z , em qualquer lugar x , equivalente a X_1 , X_2 em duas dimensões, e denotada por $z(x)$ é uma de uma infinidade de valores de uma variável aleatória $Z(x)$ naquele lugar; o qual denominamos uma “realização” do processo.
2. O conjunto de valores aleatórios em todos os lugares, novamente em número infinito, em uma região é um processo aleatório, e, também denotado $Z(x)$.
3. A variável aleatória é espacialmente correlacionada em alguma escala.

Variáveis como, a espessura ou acumulação de depósitos minerais, as alturas dos lençóis freáticos, as concentrações de elementos no solo, temperaturas do ar e precipitação, por exemplo, são consideradas como variáveis espacialmente aleatórias (Oliver e Webster, 2014). Para cada uma, no entanto, tem-se apenas uma única realização; e, conseqüentemente, não se pode calcular estatísticas para a realização ou tirar conclusões a partir dela (Oliver e Webster, 2014). Inferências requer muitas realizações, e assim para superar este impasse, é necessário atender uma outra hipótese, ou seja, que o processo é estacionário.

2.3.1.2 Estacionariedade

Para os geoestatísticos a noção de estacionariedade sustenta a Geoestatística e é com base nela que é possível supor que existe o mesmo grau de variação de lugar para lugar (Oliver e Webster, 2014).

Neste contexto entra a ideia de “previsão baseada em modelo” (Webster e Lark, 2013). Pode-se representar o processo aleatório pelo modelo:

$$Z(\mathbf{x}) = \boldsymbol{\mu} + \varepsilon(\mathbf{x}) \quad (1)$$

em que $\boldsymbol{\mu}$ é a média do processo e $\varepsilon(\mathbf{x})$ é uma quantidade aleatória com média zero e covariância, $C(\mathbf{h})$, dada pela:

$$C(\mathbf{h}) = \mathbf{E}[\varepsilon(\mathbf{x})\varepsilon(\mathbf{x} + \mathbf{h})] \quad (2)$$

o que equivale a:

$$\begin{aligned} C(\mathbf{h}) &= \mathbf{E}[\{Z(\mathbf{x}) - \boldsymbol{\mu}\}\{Z(\mathbf{x} + \mathbf{h}) - \boldsymbol{\mu}\}] \\ &= \mathbf{E}[Z(\mathbf{x})Z(\mathbf{x} + \mathbf{h}) - \boldsymbol{\mu}^2] \end{aligned} \quad (3)$$

em que, \mathbf{h} é a separação entre as amostras em ambas distâncias e direção; $Z(\mathbf{x})$ e $Z(\mathbf{x} + \mathbf{h})$ são os valores de Z em lugares \mathbf{x} e $\mathbf{x} + \mathbf{h}$, e E é o operador da esperança matemática. Porém, se a média não é constante, então a covariância não podem existir, há a necessidade de num pressuposto de estacionaridade considerado menos “exigente” ou mais fraco: a hipótese intrínseca, postulada por Matheron (Oliver e Webster, 2014); em que as diferenças esperadas são iguais a zero, isto é, $E[Z(\mathbf{x}) - Z(\mathbf{x} + \mathbf{h})] = 0$, e a covariância é substituída por metade da variação das diferenças, a semivariância:

$$\begin{aligned}\gamma(\mathbf{h}) &= \frac{1}{2} \text{VAR}[Z(\mathbf{x}) - Z(\mathbf{x} - \mathbf{h})] \\ &= \frac{1}{2} E\{[Z(\mathbf{x}) - Z(\mathbf{x} - \mathbf{h})]^2\}\end{aligned}\quad (4)$$

tal como a covariância, a semivariância depende \mathbf{h} e só em \mathbf{h} , e como uma função de \mathbf{h} é o variograma, $\gamma(\mathbf{h})$. O variograma no geral é mais útil do que a função de covariância por se apoiar num pressuposto mais fraco (hipotese intrínseca) e por isso tornou-se a ferramenta central da Geoestatística (Webster e Oliver, 2007) como mais a frente será abordado. Para processo estacionário de segunda ordem a função de covariância e variograma são equivalentes:

$$\gamma(\mathbf{h}) = C(0) - C(\mathbf{h}) \quad (5)$$

em que, $C(0)$ é a variância (σ^2) de um processo aleatório.

Para fechar este item, é necessário enfatizar que a aleatoriedade e estacionaridade são apenas atributos do modelo de variação assumido; elas não são propriedades, quer do mundo real ou dos dados (Webster, 2000). Isso significa dizer que o solo não é aleatório nem estacionário, tão somente o modelo assumido podem ser um ou outro ou ambos, como bem esclareceu Webster (2000). Todavia, tais suposições são úteis a medida em que ajudam na compreensão da complexidade de como se dá a distribuição espacial de certos fenômenos no mundo real e com isso permitindo prever com mais realismo as condições ou

“continuidades” entre os locais (pontos) esparsamente observados (amostrados) para a variável em questão (Oliver e Webster, 2014).

2.3.1.3 O variograma

A importância do variograma na Geoestatística é de tal ordem que cientistas renomados da Geoestatística se referem a ele como “a pedra ângular da Geoestatística” (Cressie, 1993; Webster e Oliver, 2007). São muitas explicações encontradas na literatura sobre “o que é o variograma?”, mas poucos fornecem uma descrição objetiva em termos práticos sobre o que ele é. Para o estudo do solo pode-se dizer que o variograma é a função mais usada para descrever como uma propriedade do solo varia ao longo da superfície da Terra (Webster e Oliver, 2007). Ele expressa quantitativamente a maneira pela qual a variação de uma propriedade muda à medida que a distância e a direção que separam dois pontos varia.¹⁷

O variograma fornece uma medida de dependência espacial¹⁸, o alcance, que pode ser usado para determinar o espaçamento ideal da amostragem e usado na interpolação geoestatística para estimar o valor de uma propriedade, por exemplo do solo, em locais entre pontos amostrados (Pennock et al., 2007; Webster e Oliver, 2007).

A teoria subjacente à Geoestatística permite utilizar o variograma para estudos intensivos de regiões não extensas, onde a amostragem relativamente densa é viável (Webster, 2008). Parte do pressuposto de que variância depende apenas da separação entre locais e não de suas posições absolutas. Em termos visual, o variograma é um gráfico de valores de semivariância no eixo y em relação a cada uma das distâncias de separação definidas e apresentadas no eixo x do

¹⁷ Além desse significado prático, há diferenças conceituais sobre o variograma (Webster e Oliver, 2007). Há três tipos de variogramas. O primeiro é o variograma do processo aleatório imaginário a partir do qual a realidade é uma realização. O segundo é o variograma regional o qual seria possível computar caso as informações da área estudada fosse completa. Na prática, a informação é fragmentária porque só estão disponíveis observações por amostragem, e é a partir delas que é computado um variograma de amostras, muitas vezes chamado de variograma “experimental”. Para obter uma aproximação ao variograma teórico, é necessário ajustar um modelo de regressão autorizado através dos pontos experimentais obtidos e tratar as flutuações ponto a ponto como efeitos de amostragem.

¹⁸ O termo “dependência espacial” é uma alusão ao comportamento de continuidade espacial que uma série de fenômenos naturais apresenta, como as propriedades do solo, cujos valores em locais que estão próximos no espaço têm maior probabilidade de serem semelhantes do que os mais distantes um do outro, isto é, eles “dependem” uns dos outros em um sentido apenas estatístico (Oliver e Webster, 1990).

gráfico. Essencialmente, os valores de semivariância baseiam-se diretamente em cálculos com dados medidos, enquanto o variograma é derivado dos valores de semivariâncias usando um modelo de regressão ajustado a estes valores de semivariância (Isaaks e Srivastava, 1989; Webster e Oliver, 2007). Apenas um número limitado de modelos de regressão são autorizados para o variograma (modelos gaussiano, esférico, exponencial, e derivados) (Webster e Oliver, 2007), os quais no próximo item são abordados. O variograma é comumente estimado pelo método dos momentos de Matheron (1965) com a solução da equação 6:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})]^2 \quad (6)$$

em que, $\hat{\gamma}(\mathbf{h})$ é a semivariância experimental a uma distância de separação \mathbf{h} ; $z(\mathbf{x}_i)$ é o valor da propriedade no i -ésimo ponto; e $N(\mathbf{h})$ é o número de pares de pontos separados pela distância \mathbf{h} . Em síntese, o variograma descreve a continuidade espacial ou dispersão das variáveis estudadas como uma função da distância entre dois pontos em uma grade.

Uma variante do variograma é o variograma-cruzado (Matheron, 1971; Journel e Huijbregts, 1978; Myers, 1982). Ele é obtido com dados de duas (ou mais) variáveis cujas mesmas devem espacialmente estar correlacionadas. O variograma-cruzado é usado para descrever a dependência espacial conjunta das variáveis correlacionadas (Matheron, 1971; Journel e Huijbregts, 1978; Myers, 1982). A dependência espacial conjunta permite efetuar a interpolação de uma das variáveis por cokrigagem (Matheron, 1971; Journel e Huijbregts, 1978; Myers, 1982) conforme mais a frente será abordado (item 2.3.1.4.1.2). Antes da modelagem do variograma-cruzado são necessários a modelagem individual do variograma de cada uma das variáveis espacialmente correlacionadas. Em termos de equação o variograma-cruzado é apenas um pouco mais complexo em relação a equação do variograma

individual (Matheron, 1965) (equação 6). O variograma-cruzado é estimado pela seguinte equação:

$$\hat{\gamma}_{ij}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{x}_i) - z_i(\mathbf{x}_i + \mathbf{h})] \cdot [z(\mathbf{x}_j) - z_j(\mathbf{x}_j + \mathbf{h})] \quad (7)$$

em que $\hat{\gamma}_{ij}(\mathbf{h})$ é a semivariância cruzada experimental a uma distância de separação \mathbf{h} ; $z(\mathbf{x}_i)$ é o valor da variável primária no i -ésimo ponto; $z(\mathbf{x}_j)$ é o valor da variável secundária no j -ésimo ponto; e $N(\mathbf{h})$ é o número de pares de pontos separados pela distância \mathbf{h} .

Uma condição específica na variografia para a cokrigagem é a modelagem da correionalização (Deutsch e Journel, 1998). Isso envolve o ajuste de modelos tanto para os variogramas como para o variograma-cruzado de modo que esses modelos levem a um sistema de cokrigagem definido positivo. Segundo Rossiter (2007) a maneira mais fácil de garantir isso é ajustar um modelo linear de correionalização. Para isso todos os modelos ajustados nos variogramas e variograma-cruzado devem ter a mesma forma e alcances aproximados. Isso significa que os seus patamares e efeito pepitas podem ser diferentes. Cabe aqui mencionar que esse procedimento será possível, sem forçar a modelagem, se as variáveis em questão (em qualquer número que seja) de fato se apresentarem correlacionadas espacialmente.

2.3.1.3.1 Modelos para o variograma e modelagem

O variograma obtido pelo método dos momentos estima o variograma subjacente, que é uma função contínua, como um conjunto de pontos discretos em intervalos de defasagem específicos (Webster e Oliver, 2007). Essas estimativas são as variâncias existentes entre pontos dentro de cada intervalo, cujas mesmas são

chamadas de valores experimentais (Isaaks e Srivastava, 1989; Webster e Oliver, 2007). Geralmente essas estimativas flutuam de um ponto a outro de intervalo porque estão sujeitas a erros que surgem em grande parte da amostragem. Para descrever a variação espacial que é contínua, é necessário então ajustar um modelo “unindo” os valores experimentais conforme já foi mencionado (nota de rodapé nº 17). Conseqüentemente um modelo de variograma ajustado não poderá ser considerado se não apenas uma representação aproximada da real variação espacial contida no conjunto de dados. Apesar de que isso de fato possa ser visto como sendo uma limitação em termos de representatividade, o modelo corretamente ajustado às semivariâncias expressas no variograma, continua sendo um dos mais poderosos resumos da variabilidade espacial da variável estudada (Webster e Oliver, 2007).

Uma abstração do poder de resumo dos modelos autorizados está esquematizada na Figura 2. No lado esquerdo da figura, estão esquematizados em forma de perfil diagramas dos diferentes comportamentos espaciais de variáveis que apresentam continuidade espacial e teoricamente podem ser modeladas. Do lado direito estão os correspondentes modelos de variograma plausíveis à melhor representar os comportamentos espaciais que foram esquematizados. A informação que se quer passar através desta figura é que para cada comportamento espacial, em termos de variabilidade dentro de uma dada escala, há um modelo que melhor represente este. É possível notar que o diagrama A (Figura 2) esquematiza uma pequena variabilidade espacial para longas distâncias a qual o modelo gaussiano poderá melhor representá-la. Na posição oposta (Figura 2) está o diagrama C que esquematiza dentro de curtas distâncias uma alta variabilidade espacial cuja mesma poderá ser melhor representada pelo modelo exponencial. Já o diagrama B (Figura 2) em uma posição intermediária o mesmo esquematiza moderada variabilidade espacial para curtas distâncias que normalmente é plausível ao modelo esférico melhor representá-la.

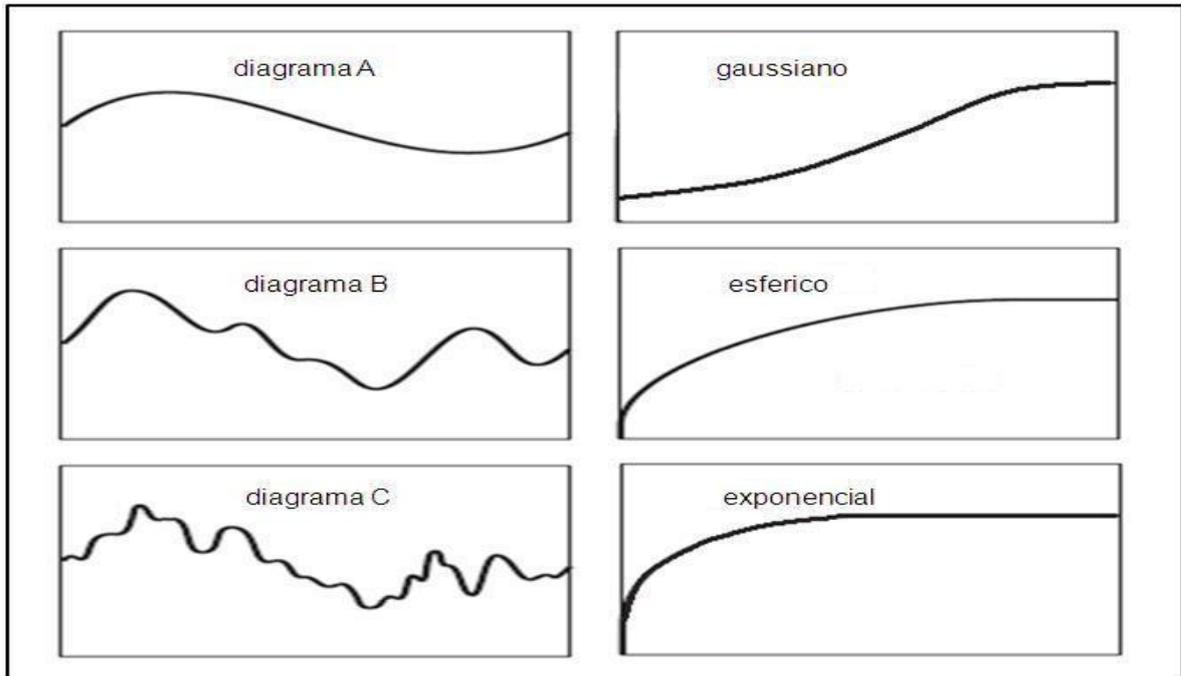


Figura 2. Esquemática visual do comportamento espacial (diagramas A, B e C) típicos ao que os principais modelos de variograma (gaussiano, esférico e exponencial) matematicamente descrevem.¹⁹

Existem algumas funções simples usuais na modelagem variográfica. Elas incluem funções limitadas, que representam processos que são estacionários de segunda ordem, e ilimitados que são intrínsecos apenas (Webster e Oliver, 2007). Webster e Oliver (2007) fornecem explicações de várias funções e há exemplos dos diferentes modelos no livro.

Porém, como mencionado antes, os modelos mais comumente ajustados são o gaussiano, esférico e exponencial; esquematizados na Figura 2, cujas equações respectivamente são 8, 9 e 10:

$$\gamma(h) = C_0 + C_1 \left[1 - \exp\left(-3 \left(\frac{h}{a}\right)^2\right)\right] \quad (8)$$

$$0 < h < d$$

¹⁹ Elaborado com base em explicações sobre a variabilidade espacial do solo em aulas da Pós-Graduação na UNESP Jaboticabal.

em que, C_0 , C_1 e a são os parâmetros de ajuste, h a distância, $\gamma(h)$ o valor da semivariância para o modelo de ajuste (Grego et al., 2014) e d é a máxima distância que o variograma é definido.

$$\gamma(h) = C_0 + C_1 \left[\frac{3}{2} \left(\frac{h}{a} \right) - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right] \quad (9)$$

$$0 < h < a$$

$$\gamma(h) = C_0 + C_1 \quad h > a$$

$$\gamma(h) = C_0 + C_1 \left[1 - \exp \left(-3 \left(\frac{h}{a} \right) \right) \right] \quad (10)$$

$$0 < h < d$$

Olea (2006) e Oliver e Webster (2014) afirmaram que ajustar modelos aos valores experimentais (semivariâncias) do variograma é controverso, e o procedimento mais recomendado é que os profissionais evitem a adaptação visual. Em termos gerais eles explicam que isso ocorre porque a acurácia das semivariâncias varia e o variograma experimental pode flutuar consideravelmente de um ponto para outro. De acordo com estes autores a montagem de um modelo adequado é de fundamental importância na Geoestatística, porque a modelagem afeta todas as análises subsequentes. Uma abordagem de mínimos quadrados ponderados é aconselhável para o ajuste porque leva em conta a acurácia das semivariâncias individuais e a soma de quadrados de resíduos fornece um meio de selecionar a melhor função de ajuste (Webster e Oliver, 2007). No entanto, conforme mais a frente é explanado (itens 2.3.1.5 a 2.3.1.5.1.1) o uso de validação-cruzada é o procedimento que deve ser usado para o veredicto final em relação ao ajuste e tipo de modelo mais adequado a ser escolhido (Davis, 1987; Hengl, 2009; Oliver e Webster, 2014). Isso significa dizer que é necessário ajustar os diferentes modelos plausíveis por mínimos quadrados ponderados, repetir tais ajustes com realização

de alguma possível melhora no encaixe a curta distância, e inspecionar através da validação-cruzada qual de todas essas opções apresenta a melhor performance.

2.3.1.3.2 Interpretação variográfica

Na prática, a interpretação variográfica se dá com a modelagem; ou seja, um modelo ajustado ao variograma experimental e estimativas dos parâmetros do variograma resultante: o efeito pepita, o patamar e o alcance (Oliver e Webster, 1990). A Figura 3 foi esquematizada para uma descrição sucinta sobre estes parâmetros conforme as formas possíveis, as mais comuns, que um modelo ajustado pode assumir em função de como as semivariâncias calculadas podem comportar-se no gráfico. Resumindo, pode-se afirmar que a modelagem é determinada por três etapas: (1) levantamento do comportamento espacial da variável estudada, (2) descrição da distribuição empírica dos dados (melhora na simetria, extração de outliers, etc.) e (3) a própria modelagem em si (raio de busca variográfica, ajuste de modelo, tipo de modelo, anisotropia, etc.).

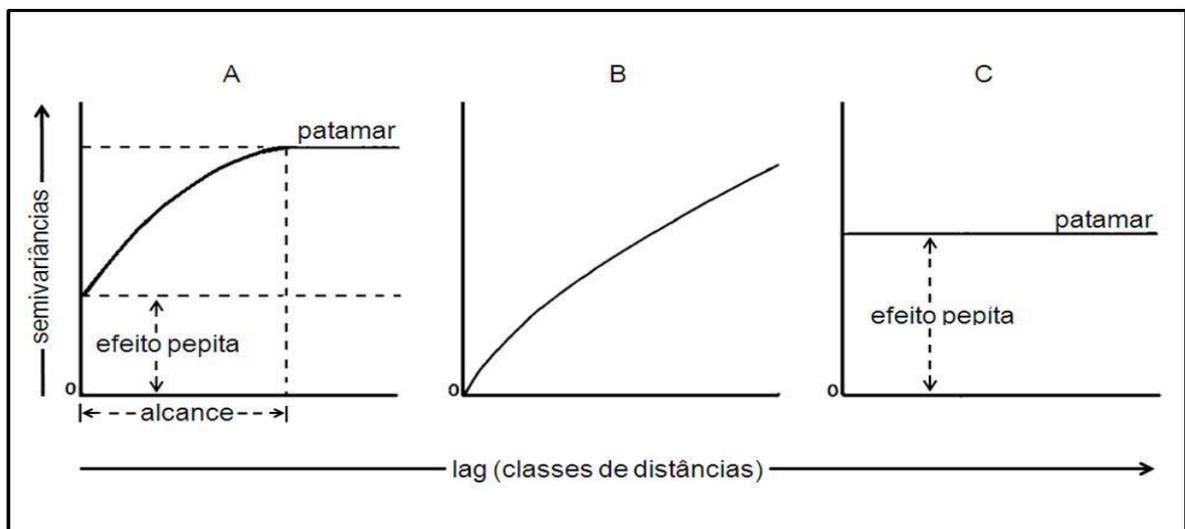


Figura 3. Comportamentos possíveis do variograma. Gráfico A apresenta patamar; gráfico B não produz patamar; gráfico C apresenta efeito pepita puro. Adaptado de Oliver (2010).

A Figura 3, gráficos A e B, mostram o que ocorre na maioria dos casos na variografia; as semivariâncias aumentam com o aumento da distância (Oliver e

Webster, 1990). Isso corresponde a uma correlação mais ou menos forte ou dependência espacial nas distâncias mais curtas, e que enfraquece à medida que a separação aumenta (Oliver e Webster, 1990). Intuitivamente, isso expressa o que se deve esperar: os locais próximos um do outro têm solo semelhante, enquanto os mais afastados tendem a ter um solo que é, em média, mais diferente. Ainda no gráfico A (Figura 3), pode ser visto que as semivariâncias aumentam e depois se aplanam (estabilizam) quando atinge uma variância conhecida como variância do patamar ou simplesmente patamar. Este limite marca a distância em que o patamar é alcançado e, em termos gráficos, eixo x, fornece a medida de alcance de dependência espacial da variável estudada. Alternativamente, o conjunto de semivariâncias podem não formar o patamar e aumentar indefinidamente, como mostrado no gráfico B (Figura 3) – isso significa que a hipótese intrínseca não está sendo satisfeita, ou seja, que as semivariâncias para as distâncias mais longas dentro da área estudada não se estabilizaram acompanhando o comportamento da variância global (Hengl, 2009).

Ainda em relação a Figura 3, normalmente o posicionamento das semivariâncias no gráfico fazem com que o modelo ajustado ao variograma tenha um intercepto positivo na ordenada. Essa faixa positiva na ordenada corresponde à variação interpretada como “aleatória”, e é conhecida como efeito pepita. Graficamente o efeito pepita quando presente marca uma descontinuidade na origem do variograma (Webster e Oliver, 2007). Essa descontinuidade é uma característica muito comum de variogramas. Ela foi observada pela primeira vez em variogramas gerados de dados obtidos para avaliação de depósitos de ouro na África do Sul, onde essa descontinuidade estava associada à presença de pepitas de ouro, e por tal, vindo a se chamar de efeito pepita (Armstrong, 1998). O que se interpretou nesses variogramas, segundo Armstrong (1998), foi que o valor passou abruptamente de zero “fora da pepita” para um valor alto “dentro dela”. Entretanto, a variografia aplicada a avaliação de outros recursos, por exemplo, a variabilidade de propriedades do solo, o efeito pepita reflete outros aspectos; pois, diferentemente de corpos de minério, o solo não contém pepitas dispersas aleatoriamente, e nem de outro modo (Oliver e Webster, 2014). Dessa forma, a Geoestatística aplicada ao contexto da Ciência do Solo, o efeito pepita, quando presente, reflete erros de

medição e também eventuais variações não captadas pelo intervalo de amostragem utilizado ou mesmo “contaminações” (outliers). (Webster e Oliver, 2007). Todavia, é pertinente destacar que o efeito pepita também reflete decisões na modelagem (Oliver e Webster, 2014). E isso pode ser visto quando, por exemplo, em uma dada situação, o ajuste do modelo esférico e do modelo exponencial são igualmente plausíveis: os efeitos pepitas de cada um serão diferentes se ajustados. Dentro desta condição, Oliver e Webster (2014) mostraram que a proporção do efeito pepita do modelo esférico é maior que a do modelo exponencial, e mostraram também implicações decorrentes. Num extremo oposto da possibilidade da ausência de efeito pepita (0% de efeito pepita), está a possibilidade gráfica de se ter feito pepita em proporção total (100% de efeito pepita), conforme ilustrado no gráfico C da Figura 3; ou seja, o variograma se apresentará completamente plano. Isso significa que há ausência total de dependência espacial, ou seja, a dependência espacial, se existir, será manifestada à distância menor do que o menor espaçamento entre amostras do conjunto de dados analisados (Webster e Oliver, 2007). No jargão geoestatístico o variograma apresenta “efeito pepita puro” (Oliver e Webster, 2014). Se nesta circunstância, por exemplo, se tentar krigar, poderá ser verificado que os pesos de krigagem serão todos iguais, e o procedimento de krigagem simplesmente retornará o valor médio dos dados de entorno para qualquer novo ponto estimado. Porém, essa é a predição que se obtém com o uso de um interpolador comum (interpolador que não incorpora a informação de dependência espacial) e, geralmente quando se utiliza a Geoestatística, não é o tipo de previsão que se pretende. A acurácia numa condição como esta tem grande possibilidade de ser baixa (Hengl, 2009). Este é o fato que nos leva à questão da Escala espacial (espaçamento amostral) e da Configuração amostral (disposição amostral).

2.3.1.3.3 Amostragem

2.3.1.3.3.1 Escala espacial (espaçamento amostral)

Um variograma que apresenta efeito pepita puro contraria o nosso senso do que é certo, porque o solo é um *continuum*, exceto nas bordas das falésias, margens

de rios, lagos e afins (Oliver e Webster, 2014). Portanto, existe alguma escala na qual há dependência espacial e, se porventura, o variograma gerado no estudo apresentar efeito pepita puro, ele estará sinalizando que não foi identificada dependência nessa escala e isso significa que se desperdiçou esforços (Webster e Oliver, 2007). Da mesma forma se estará desperdiçando esforços fazendo uma amostragem mais intensiva do que a necessária, em função de não se conhecer a escala aproximada de variação (Webster e Oliver, 2007). Para o pesquisador em determinadas regiões, a distribuição geológica, as características do relevo (elevação), variações obtidas por fotografias aéreas, informações da própria pedologia (ambiente de ocorrência de certas características de solo) – conforme mencionado no item 2.2, fornecem pistas para a escala espacial. Mas, supondo que na região de estudo estes elementos não estão presentes, ou seja, os aspectos da paisagem não oferecem pistas – e é muito comum situação assim – o que se deve fazer para que o plano de amostragem possa capturar a escala espacial para o variograma? Para esta pergunta há mais do que uma resposta.

Informações geradas por pesquisas anteriores podem ser muito úteis para este propósito (Mcbratney e Pringle, 1999; Kerry et al., 2010). Nesse sentido, é conveniente a realização de um levantamento na literatura se há disponível informações sobre a caracterização da variabilidade espacial de propriedades do solo para tal região ou realizada em ambiente com relativa semelhança pedológica. Por exemplo, para o Planalto Ocidental Paulista, existe inúmeros estudos que foram realizados para este propósito, cujo o acesso inclusive é livre por meio da internet (links: <http://www.csme.com.br/> e <https://www.athena.biblioteca.unesp.br/>). Essas informações podem então ser utilizadas para planejar a amostragem do solo na qual o variograma poderá ser estimado com maior possibilidade de precisão dentro do alcance que de fato importa (Kerry et al., 2010). Outra forma para se estabelecer um plano de amostragem ideal para o mapeamento (com a captura da escala espacial de variação) se realiza através de um estudo prévio, procedendo uma análise variográfica baseada em uma amostragem intensiva²⁰ em transeção linear para

²⁰ A amostragem intensiva para este propósito na Agricultura de Precisão, como de bom senso, deve ser de tal modo que o intervalo utilizado não tenha dimensão menor do que a dimensão da área mínima tratável pela tecnologia de aplicação à taxa variável utilizada (Pierce e Nowak, 1999). Ou seja, estimar a variabilidade espacial a distâncias menores do que essa capacidade operativa simplesmente será em vão.

identificar a dependência espacial da variável em questão (McBratney et al., 1981; McBratney e Webster, 1981; McBratney e Webster, 1983). Identificado isso, um intervalo de amostragem econômico poderá ser definido para modelar o variograma com quantidade de dados que realmente importa: sem o desperdiço de esforços com amostragem mais intensiva do que a necessária ou aquém da possibilidade de captar a variabilidade que se pretende mapear (McBratney e Webster, 1983; Kerry e Oliver, 2004). Este esquema de amostragem envolverá a coleta de um número de amostras a distâncias que não ultrapassem o alcance de dependência espacial do variograma previamente estimado para orientação. Todavia, interpolação com chances de predições mais acuradas, requer intervalos mais curtos do que a escala de dependência espacial, normalmente usa-se menos da metade do valor dessa escala (McBratney e Webster, 1986; Kerry e Oliver, 2004). Isso resultará em uma proporção de efeito pepita no variograma relativamente menor, e teoricamente permitirá ao interpolador ponderar melhor os pesos e realizar melhores predições (Cambardella et al., 1994).

2.3.1.3.3.2 Configuração amostral

Conforme mencionado anteriormente, a Geoestatística assume que o valor de uma propriedade de solo, em qualquer local, é função do valor da mesma propriedade em locais próximos (dependência espacial). Por isso é necessário ressaltar que a distância e direção entre locais amostrados são de fundamental importância na estimação da dependência espacial entre os valores de uma propriedade do solo nesses locais. Logo, o uso da Geoestatística exige portanto, cobertura amostral, e, por tal, a configuração ideal de coleta amostral é uma grade (malha) de intervalo regular (Pennock et al., 2007; Webster e Lark, 2013). De acordo com a literatura especializada, uma grade triangular fornecerá as estimativas mais precisas para uma determinada densidade de amostragem, mas, uma grade retângular é mais praticável²¹ e apenas ligeiramente menos eficiente (McBratney et al., 1981; McBratney e Webster, 1983a; Webster e Oliver, 2007; Kerry et al., 2010).

²¹ Praticável deve ser entendido como mais fáceis de ser implementada e também aquilo que é mais usual.

Neste sentido, tudo o que restará levantar para se decidir na amostragem é o intervalo da grade.

2.3.1.3.3.3 “Relaxamento” amostral na cokrigagem

Na prática, se o orçamento estipulado para o propósito de mapeamento tiver valor fixo, isso determinará o número de pontos de amostragem que proporcionalizará a cobertura; e o intervalo de grade se subordinará a isso (Kerry et al., 2010; Webster e Lark, 2013). No entanto, se este não for o caso, e o que se procura é que o mapa retrate bem a variabilidade espacial da variável que se pretende mapear, ou seja, a obtenção de mapas com alta acurácia, então informações de um variograma prévio deverá ser usadas para determinar o intervalo de grade a ser empregado (McBratney et al., 1981; McBratney e Webster, 1981; McBratney e Webster, 1983b) conforme descrito no item 2.3.1.3.3.1.

A partir do exposto fica claro que a informação necessária para a otimização da amostragem na Geoestatística é a da própria dependência espacial da variável investigada. Isso é válido na análise Geoestatística tanto univariada como multivariada (McBratney e Webster, 1983a), ou seja, tanto para krigar como para cokrigar, por exemplo. Porém, no caso da cokrigagem surge a seguinte indagação: como essa afirmação pode estar correta quando levando em consideração que, em última análise, o objetivo prático da cokrigagem é a diminuição amostral da variável a ser mapeada? E que, inclusive, a cokrigagem só proporciona o melhor de si se a densidade amostral da variável que se pretende mapear estiver menos densamente amostrada em relação a covariável auxiliar?

Primeiramente é necessário destacar que para qualquer dos casos, para a otimização amostral, a informação da dependência espacial das variáveis do solo é a base (McBratney et al., 1981). Isso significa que o variograma sempre será ferramenta central (McBratney et al., 1981). Porém, no caso da análise geoestatística para a otimização amostral para a cokrigagem a própria interpolação é parte integrante e uma das ferramentas na análise (Vauclin et al., 1983; Trangmar et al., 1986; Vaughan et al., 1995; Teixeira et al., 2013; Song et al., 2014). Por exemplo, é necessário partir de condições que asseguram variogramas confiáveis

para as variáveis envolvidas. Mas também é necessário considerar que haverá força de uma variável a mais (ou mesmo mais variáveis a mais) na análise de uma variável. Isso permite de certo modo diminuir a exigência numérica de informações espaciais (intensidade amostral) que seria necessária se a análise fosse univariável. Mas, sendo isso possível, como saber o quanto diminuir? McBratney (1992) e Kerry et al. (2010) mencionaram que há falta de informações a este respeito; logo, normalmente ao intuir levantar esta informação, a experimentação com a própria interpolação entra como necessária, como anteriormente destacado (Vauclin et al., 1983; Trangmar et al., 1986; Vaughan et al., 1995; Teixeira et al., 2013; Song et al., 2014).

De acordo com estes estudos a experimentação para otimização amostral para cokrigagem requer cenários de amostragem com densidades suficientes (cobertura ideal) para as variáveis envolvidas: variável alvo e covariável (variável auxiliar). Tendo obtido tais informações (banco de dados), é necessário reconstrução de cenários de amostragem decrescentes apenas da variável a ser mapeada, isto é, com densidades relativamente menores para esta; para que assim se possa realizar verificação da acurácia obtido com os mapas resultantes da interpolação dos diferentes cenários da variável alvo combinados com o cenário da covariável mantido fixo. Qual desses cenários (combinação) permite obter o mapa com maior acurácia? Esta informação é obtida por meio do procedimento de validação dos mapas.

É importante reiterar que o melhor desempenho da cokrigagem requer o “relaxamento” amostral da variável a ser mapeada, isto é, que a densidade de amostragem da variável alvo seja menor do que da amostragem da covariável, como demonstrado em diversos trabalhos (por exemplo, Vauclin et al., 1983; Trangmar et al., 1986; Vaughan et al., 1995; Teixeira et al., 2013; Song et al., 2014). Pois, caso contrário, os resultados obtidos com a cokrigagem serão semelhantes aos que se obtêm com a krigagem ou apenas um pouco melhor. Todavia, é importante observar que estes estudos apontam para a existência de um limite dentro do qual se pode ir. Ou seja, de que há uma relação de densidade de amostragem para cokrigagem (combinação de informações espaciais da variável alvo e de covariável) que permita as melhores predições com a cokrigagem. Isto significa dizer que a incorporação de

informação de uma covariável (ou mais) possui um limite em termos de capacidade de auxiliar na melhora das predições ou qualidade do mapa.

2.3.1.3.4 Fatores que afetam a variografia

A construção do gráfico variograma envolve aspectos que merecem ser abordados. A confiabilidade do mesmo é dependente da modelagem²² e do comportamento dos dados. Uma descrição completa sobre modelagem está além do escopo desta revisão e uma descrição detalhada pode ser encontrada em McBratney e Webster (1986) e Oliver e Webster (2014). Porém, nos subitens seguintes, são revisados cinco fatores que comumente afetam significativamente o variograma e conseqüentemente a análise geoestatística como um todo, por tal, precisam ser muito bem observados:

2.3.1.3.4.1 Número de pontos de amostragem e distância de separação

Muitos pesquisadores registraram em artigos que o variograma experimental computado pelo método dos momentos de Matheron (1962) pode ser significativamente afetado pelo tamanho da amostra e pela distância entre os pontos de amostragem. Recentemente Oliver e Webster (2014) elaboraram um artigo de cunho puramente educacional para mostrar isso. Segundo eles, um variograma confiável requer pelo menos 100 amostras por campo e o intervalo de amostragem deve ter a capacidade de representar a escala da variação. Embora sobre esta informação essa orientação seja a mais difundida, algumas referências utilizam-se da quantidade de pares de pontos para cada estimativa de semivariância como a informação que deve ser observada para obtenção de um variograma confiável (Olea, 2006; Chilès e Delfiner, 2012). Estas referências assumem que são necessários pelo menos 50 pares de pontos para cada estimativa de semivariância.

²² A modelagem adequada não é apenas crucial para se obter uma análise correta da estrutura espacial da variável estudada (variograma confiável) mas também para se produzir economicamente um mapa que de fato seja representativo da variabilidade espacial estudada (interpolação de alta performance) (Webster e Oliver, 2007; Oliver e Webster, 2014).

Todavia, pelo que se verifica na literatura (apenas pelo aspecto visual dos variogramas), muitos estudos parecem desprezar ambas as orientações.

Oliver e Webster (2014) explicam que o intervalo de amostragem diminuindo a medida que o tamanho a densidade de amostragem aumenta, a propagação de valores de $\hat{\gamma}(h)$ se estreita em lags (distância de separação entre pares de pontos) curtos. Ou seja, o variograma apresentará as semivariâncias de forma menos dispersas. Porém, em situação contrária, a dispersão de pontos poderá ser grande e dificultar sobremaneira o ajuste de um modelo que realmente possa representar adequadamente a variabilidade espacial presente na área. Esses autores também explicam que o próprio intervalo de amostragem é importante por outro motivo; que determina a utilidade do variograma experimental: se o intervalo é maior do que o intervalo de correlação do processo ou na realização, conseqüentemente o variograma experimental será plano: “efeito pepita puro” na gíria. E este será inútil para a predição geoestatística. Essa situação apenas sinalizará que toda variação que ocorre, está dentro de uma distância menor do que o intervalo entre pontos de amostragem (Webster e Oliver, 2007).

2.3.1.3.4.2 Distribuição marginal dos dados (assimetria e outliers)

Embora um conjunto de dados simétricos ou próximo disso seja o ideal para as análises geoestatísticas, Cressi (1993) explica que a distribuição normal dos dados não é um pré requisito às análises geoestatísticas. Oliver e Webster (2014) ilustram com dados e resultados de um estudo realizado por Kerry e Oliver (2007) que o comportamento assimétrico dos dados afeta significativamente a confiabilidade dos variogramas. Em razão disso, eles sugerem como rotina na Geoestatística que se verifique a forma da distribuição empírica dos dados (histograma, “box plot”, cálculo de coeficiente de assimetria, etc.) antes de se realizar a análise variográfica. Eles mostram que os dados com comportamento assimétrico podem ser submetidos a alguma transformação (logarítmica, raiz quadrada...) e terem a distribuição aproximadamente normalizada. Kerry e Oliver (2007), Oliver e Webster (2014) e outros pesquisadores consideram plausível a

preocupação com a distribuição dos dados porque como o variograma computado pelo método dos momentos é baseado nas semivariâncias, qualquer desvio da normalidade pode superestimar as semivariâncias e ampliar o limite de confiança que, conseqüentemente, altera a forma do variograma. Portanto, no ponto de vista desses autores se qualquer desvio grande da simetria for observado, os dados podem então ser transformados, de modo que uma aproximação a uma distribuição normal possa ser alcançada e o variograma seja então calculado a partir de dados transformados. Porém, esta é uma preocupação que pode se dizer ser com o “variograma”, e de fato uma preocupação pertinente; pois, o variograma é o “coração” na Geoestatística: é o mensurador da dependência espacial da variável em estudo e é ele o fornecedor dessa informação para o interpolador ser um estimador BLUP²³ (Cressie, 1993; Webster e Oliver, 2007; Oliver e Webster, 2015).

Entretanto, é necessário considerar também a possibilidade de não se obter os melhores resultados com a interpolação usando conjunto de dados o qual se corrigiu a assimetria. Diante disso, surge a pergunta: como resolver este dilema? Para Webster²⁴ não existe uma única resposta simples para esta pergunta. Nadler e Wein (1998) sugerem o mais plausível: usar o erro de predição para decidir o que é conveniente em relação à condução da análise com dados; se com a transformação, ou sem a mesma; ou seja, avaliar se o procedimento de transformação incorre em ganhos ou não em termos de acurácia na interpolação. Para mérito de exemplificação, Goovaerts (2000) observou que a correção da assimetria com a transformação dos dados pode ser problemática, porque, por exemplo, usando-se da exponenciação, pode-se até diminuir a assimetria dos dados, mas, poderá levar os erros de interpolação ficarem exagerados.

Outra situação relacionada com a distribuição dos dados que pode causar graves distorções são os outliers (chamados de valores atípicos). São valores, como destaca Oliver e Webster (2014), que simplesmente não são extremos ou perto de extremos de uma distribuição de frequência, mas são inesperadamente valores grandes ou pequenos em relação ao conjunto (Barnett e Lewis, 1994). Ou seja, são valores que parecem altamente improváveis que pertençam às populações de

²³ O tópico sobre krigagem (item 2.3.1.4.1) contém explicações sobre esta sigla.

²⁴ Webster R (2017) (pesquisador da Rothamsted, Harpenden, Reino Unido). Comunicação Pessoal.

interesse. Portanto, os outliers não são o resultado de valores cujos mesmos façam parte da população que se visa representar.

Oliver e Webster (2014) descrevem que a presença de outliers na análise variográfica pode ser facilmente reconhecida: há um aumento dramático no tamanho do efeito pepita. E alertam que a presença desses valores podem falsificar as descrições dos processos de interesse, e tornando-as completamente equivocadas. Por isso é necessário checagem prévia para constatação, checagem variográfica para visualização, e checagem da acurácia para avaliação se é ou não conveniente modelar e interpolar sem ou com a remoção desses valores. Ou seja, é sempre prudente avaliar qual o procedimento implica em ganhos em termos de acurácia na interpolação.

Enfim, com vista a uma melhor performance geral com a Geoestatística, é como Oliver e Webster (2014) frisaram “devemos tentar todas as alternativas plausíveis para ver qual opção funciona melhor para os dados que estamos trabalhando”.

2.3.1.3.4.3 Tendência nos dados

A presença de tendência são indicadas por semivariâncias que aumentam com a distância de separação entre os pontos de amostragem e não formam um patamar no variograma (Gajem et al., 1981) ou tem um patamar que excede consideravelmente a variância geral dos dados (Bresler et al., 1984). Porém, a quantidade insuficiente de dados também poderá levar a um variograma sem patamar. A inadequada definição do dispositivo de pesquisa para o cálculo do variograma experimental (por exemplo, a faixa de distância na qual se utiliza o variograma) também pode levar a um variograma sem patamar (Yamamoto e Landim, 2013) ou alterar o mesmo. Esses são fatores que eventualmente podem levar erroneamente a um julgamento de presença de tendência. Porém, e se de fato houver presença de tendência? O que é sugerido fazer com os dados quando é necessário interpolar? De acordo com Vieira et al. (2010) uma das maneiras mais práticas de resolver esta questão é subtrair dos dados originais uma função ajustada

por mínimos quadrados, trabalhando (variografia e interpolação) com a função residual resultante. Utilizando desta prática sugerida por Vieira e colaboradores, é necessário converter os valores de predição obtidos com os resíduos para retorno aos valores originais coerente com a unidade de medida da variável em questão.

2.3.1.3.4.4 Intervalo de lag e abertura angular

Como no item anterior foi mencionado, a inadequada definição do dispositivo de pesquisa para o cálculo do variograma experimental (comprimento e ângulo da faixa de distância utilizada pelo variograma, respectivamente h e L) também pode levar a uma análise variográfica não confiável (Yamamoto e Landim, 2013). O dispositivo de pesquisa para o cálculo do variograma experimental definirá a quantidade de pares de pontos a serem trabalhados no cálculo (Yamamoto e Landim, 2013). Segundo estes autores o variograma calculado com maior número de pares é mais suavizado (estável), e isto facilita o trabalho de ajuste do modelo teórico. Eles afirmam também que o variograma experimental representado por um maior número de pares de pontos tem maior significado estatístico. Todavia, segundo Oliver e Webster (2014), a escolha de h e de L requer julgamento. Esses estudiosos explicam que se h for curto e L for estreito, haverá então muitas estimativas de $\gamma(h)$, cada um com base em algumas comparações e sujeitos a grande erro, e o variograma aparecerá “ruidoso”. E, em contrário, se h for grande, e L for grande, poderá haver muito poucas estimativas de $\gamma(h)$ para revelar a forma do variograma. Diante desse dilema, o que se cabe fazer é representar graficamente os valores experimentais para que se possa fazer uma avaliação sensata (Oliver e Webster, 2014). Isso inclui realizar uma verificação global: se cada semivariância experimental está estimada com 50 pares de pontos ou mais; se o variograma experimental apresenta entre 10 a 15 lags; se o variograma se apresenta de forma bem estruturada (suavizado), etc.

2.3.1.3.4.5 Anisotropia

Os dados podem exibir evidências de anisotropia, que devem ser consideradas na modelagem; por poder resultar uma estimativa tendenciosa. Segundo Oliver e Webster (2014) em muitos casos a variação é anisotrópica, e o que frequentemente se percebe, é que os pesquisadores no afã da facilidade insistem em ignorá-la, e, quando não, eles a consideram mas realizam a modelagem de forma inadequada; conseqüentemente seus trabalhos apresentam visões equivocadas das verdadeiras situações. No entanto, em alguns casos, a anisotropia pode ser ignorada para simplificar o ajuste do modelo e para manter alguma consistência entre os variogramas no modelo multivariado (Martinez-Cob, 1996). Esta exceção é possível principalmente quando: (1) a anisotropia não é evidente com a distância de pesquisa especificada; (2) a variável secundária e a variável primária são colocadas, portanto, a influência dos valores circundantes normalmente são pequenos, de modo que a anisotropia acaba fazendo pouca diferença e (3) as direções de variabilidade espacial máxima e mínima para as diferentes variáveis não coincidem. Em seu estudo observando o impacto de variogramas isotrópicos e anisotrópicos no desempenho da interpolação, Haberlandt (2007) verificou que as diferenças não foram significativas no desempenho da predição, embora a anisotropia estivesse claramente presente nos dados. E, paradoxalmente, o autor ressalta que as melhores estimativas foram obtidas usando variogramas isotrópicos. De qualquer forma Haberlandt (2007) deixa uma lição muito válida: usar o erro de predição para decidir o que é conveniente – manter a modelagem considerando a detecção da anisotropia ou modelar sem (como isotrópico)? Ou seja, a decisão que sempre será a mais sensata é optar pelo procedimento que resulte em ganhos em termos de acurácia na interpolação.

2.3.1.4 Interpolação geoestatística

2.3.1.4.1 Krigagem

A krigagem é um termo genérico para uma variedade de métodos geoestatísticos usado para propósitos de interpolação (predições) entre locais (pontos) amostrados (Cressie, 1990). É baseada na suposição de que a propriedade que está sendo interpolada pode ser tratada como uma variável regionalizada (Webster e Oliver, 2007). Uma variável regionalizada é intermediária entre uma variável verdadeiramente aleatória e uma variável completamente determinística, variando de maneira contínua de um local para o outro e, portanto, pontos próximos um do outro possuem certo grau de correlação espacial, mas pontos que são amplamente separados são estatisticamente independentes (Davis, 1986). A krigagem requer um modelo de continuidade espacial ou dependência obtido com a estimativa de variograma para definição de pesos da função krigagem cujos os mesmos têm o papel de minimizar erros das estimativas (Cressie, 1990; Olea, 1999).

É importante ressaltar que existem vários tipos de krigagens na atualidade (Li e Heap, 2014). O que aqui está sendo apresentado sobre a krigagem são descrições gerais (características, origem, panorama, etc.) e neste trabalho apenas os métodos de krigagem ordinária e cokrigagem ordinária são mais a frente revisados com intuito de embasamento teórico dos estudos que nesta Tese é relatado. Para uma abordagem detalhada sobre interpolação geoestatística os mesmos livros recomendados no item 2.3.1 são também indicados .

A origem da krigagem está relacionada aos métodos estatísticos empíricos desenvolvidos por Daniel Krige para prever minério com dados de amostragem em grades, espacialmente correlacionados, nas minas de ouro da África do Sul (Krige, 1951, 1966). Porém, a krigagem só chegou ao meio científico ao fazer parte do corpo principal da Teoria das Variáveis Regionalizadas desenvolvida por Matheron (1965, 1971) e sua equipe (Oliver e Webster, 1990). Foi a partir da década de 1980 que cientistas diversos – pedólogos, hidrólogos, geólogos e cientistas atmosféricos, para citar alguns – começaram a comprovar a aplicabilidade da krigagem em outras

áreas do conhecimento (por exemplo, Burgess e Webster, 1980a; Vauclin et al, 1983; Russo, 1984; Oliver e Webster, 1987).

A krigagem é amplamente aplicada em diversos campos, como meteorologia, hidrologia, engenharia de petróleo, geologia, mineração, poluição ambiental, biologia vegetal e animal, pesca, saúde pública, medicina, economia, Ciência do Solo, Agricultura de Precisão, entre outros. Em síntese, essa popularidade decorre que nenhum outro método de interpolação linear fornece melhores previsões lineares sem viés do que a krigagem, e por isto seu *status* de melhor preditor linear não-viesado (BLUP – sigla em inglês de: Best Linear Unbiased Prediction)²⁵ (Isaaks e Srivastava, 1989; Olea, 1999; Webster; Oliver, 2007; Oliver, 2010). A krigagem é considerada um método de interpolação exato, em um sentido teórico, porque quando a variância estrutural é 100% (o efeito pepita é 0%) a interpolação “honra” os dados. Ou seja, dentro dessa condição a interpolação para locais (pontos) aonde existem medições, a krigagem sempre gerará valores iguais aos valores medidos. Todavia, na verdade, *grosso modo*, o que permite a krigagem ser considerada um interpolador ótimo (BLUP), é sua capacidade de considerar a dependência espacial nos cálculos e minimizar a variação de erro associada à estimativa (Isaaks e Srivastava, 1989; Olea, 1999). Ou seja, para o conjunto de dados vizinhos usados em uma estimativa, o conjunto de pesos calculados com krigagem garante que nenhum outro método gere um erro de estimativa mais baixo.

2.3.1.4.1.1 Krigagem ordinária

Como mencionado, krigagem é um termo genérico para uma variedade de métodos geoestatísticos usados para propósitos de interpolação. A krigagem ordinária (KO) na formulação original de Matheron (1965) é a mais popular em razão de que ela serve bem na maioria das situações, com suas suposições facilmente satisfeitas (Vieira, 2000; Webster e Oliver, 2007). Ela exige apenas o conhecimento

²⁵ É comum na literatura BLUE (Best Linear Unbiased Estimator) ao invés de BLUP (Isaaks e Srivastava, 1989). Decorre que alguns estudiosos consideram que a interpolação produz “estimação” enquanto outros “previsão”. De qualquer modo, estimar ou prever, em um sentido espacial, é apenas uma questão terminológica (Webster e Lark, 2013). Nesta revisão se adotou BLUP porque é o que tem sido mais comum no âmbito da Ciência do Solo; todavia, tanto as palavras “estimação” como “previsão” foram indistintamente utilizadas neste trabalho.

da função variograma e os dados para a sua implementação (Webster e Oliver, 2007). Para esses mesmos autores a KO é também considerada robusta porque normalmente apresenta desempenho razoável mesmo quando a modelagem para o variograma é aquém do ideal. Todavia, é importante destacar que ela só se expressará como BLUP quando a variação espacial é devidamente traduzida em um modelo de variograma correto (Cressie, 1993; Webster e Oliver, 2007) e por isso foi incluído o tópico “Fatores que afetam a variografia” (item 2.3.1.3.4) nesta revisão. A KO resolve um conjunto de equações lineares, conhecidas como sistema de krigagem, o qual faz uso extensivo do variograma (equação 6). O método, KO, em termos de predição se desdobra em krigagem pontual e krigagem em bloco (Oliver e Webster, 2014), cujas equações diferem muito pouca coisa uma em relação a outra. Krigagem em bloco está fora do escopo dessa Tese.

Dessa forma, obtido o variograma da variável e havendo dependência espacial entre as amostras da variável estudada, conforme anteriormente explanado (item 2.3.1.3.2), pode-se interpolar valores em qualquer posição no campo com a krigagem ordinária (KO). A KO é um método de interpolação geoestatístico que consiste em uma média ponderada de amostras vizinhas. Os pesos (λ_i) para cada vizinho são determinados com base no modelo de variograma ajustado, de modo que a variância das estimativas é minimizada levando a um sistema linear de equações, cuja solução é dada pela equação 11:

$$\hat{z}(\mathbf{x}_0) = \sum_{i=1}^N \lambda_i z(\mathbf{x}_i), \quad \text{com a restrição} \quad \sum_{i=1}^N \lambda_i = 1 \quad (11)$$

em que, $\hat{z}(\mathbf{x}_0)$ é o valor predito da propriedade no ponto \mathbf{x}_0 ; N é o número de valores usados para predição; λ_i é a ponderação associada a cada valor; e $z(\mathbf{x}_i)$ é o valor observado no i -ésimo ponto vizinho de $\hat{z}(\mathbf{x}_0)$.

O método de KO tem um diferencial em relação a todos os outros métodos de interpolação comuns. Com a KO é possível estimar a variância do valor interpolado,

$\sigma^2(\mathbf{x}_0)$. Esta estimativa depende de apenas dois fatores. O primeiro é a estrutura espacial da propriedade medida, modelada no variograma, e o segundo, os fatores de ponderação para os vizinhos da localização interpolada. Esses fatores de ponderação dependem, por sua vez, da disposição dos locais amostrados ao redor do local a ser estimado $\hat{z}(\mathbf{x}_0)$. Ela é calculada com a solução da equação 12:

$$\hat{\sigma}^2(\mathbf{x}_0) = \mu + \sum_{i=1}^N \lambda_i \gamma_{i0} \quad (12)$$

em que, $\hat{\sigma}^2$ σ^2 é a variância de estimação, μ é o multiplicador de Lagrange e γ_{i0} é a semivariância na distância de separação entre os locais para os valores vizinhos medidos $z(\mathbf{x}_i)$ a ser interpolados $z(\mathbf{x}_0)$.

É importante algumas considerações sobre a variância de krigagem. Como pode ser vista ao observar a equação que gera a variância de krigagem (equação 12); ela é uma medida que de certa forma é independente dos valores dos dados. Conseqüentemente, ela não reflete a incerteza esperada em um ponto específico na área krigada (Goovaerts, 1997). Ela se caracteriza mais como um índice de classificação da geometria de dados (e densidade) e não é uma medida da disseminação local de erros (Goovaerts, 1997). Alguns autores afirmam que a variância de krigagem é pouco correlacionada com o erro de estimação real. Portanto, enfatizam que a variância da krigagem não pode ser usada sozinha como medida de incerteza local (Goovaerts, 1997). No entanto, a variância de krigagem não pode ser vista como uma informação inútil na Geoestatística. Por exemplo, Voltz e Webster (1990) desenvolveram o índice de validação, MSDR²⁶, o qual é resultante de cálculo que inclui a variância de krigagem e os resíduos da validação-cruzada.

²⁶ O MSDR é a sigla em inglês de “mean squared deviation ratio” cuja tradução livre é “razão de desvio quadrático médio”. MSDR é uma medida que em termos práticos reporta a qualidade do ajuste do modelo no variograma (Oliver e Webster, 2014). O modelo ideal (com o seu ajuste ideal) para o variograma apresentará MSDR igual a 1 (um); por tal, o modelo ajustado a ser escolhido deverá ser aquele que apresentar MSDR o mais próximo possível de 1 (Oliver e Webster, 2014). A equação para obtenção da MSDR pode ser consultada no trabalho desses autores.

Este índice tem se mostrado uma medida das mais seguras para auxiliar a modelagem variográfica. Outra utilidade da variância de krigagem tem sido nos estudos também sobre adequações na densidade de amostragem para atender a krigagem (McBratney et al., 1981; OLIVER, 2010) já que, em última análise, a variância de krigagem para cada ponto estimado é principalmente uma medida da densidade de informação (pontos) em torno deste (Wackernagel, 1998).

2.3.1.4.1.2 Cokrigagem ordinária

A cokrigagem é um dos métodos geoestatísticos de interpolação que utiliza informações sobre os padrões espaciais de duas variáveis diferentes (ou mais), mas espacialmente correlacionadas, para interpolar apenas uma das variáveis (Vauclin et al., 1983). A cokrigagem é uma extensão multivariada do método de krigagem (Pawłowsky-Glahn e Olea, 2004), desenvolvida por Matheron para produzir estimativas com boa acurácia de uma variável que foi menos densamente amostrada com base em sua dependência espacial e na dependência espacial de uma ou mais covariáveis que foram mais intensivamente amostradas (Matheron, 1971; Journel e Huijbregts, 1978; Myers, 1982). Mas, por que isso pode ser importante para a Ciência do Solo ou para a agricultura moderna? Na prática, o custo para o mapeamento de variáveis que são mais difíceis ou mais caras de se realizar medições pode ser substancialmente reduzido se a densidade de amostragem da mesma puder ser “relaxada” em função da possibilidade de suplementação amostral com uma variável que é mais fácil ou menos cara na mensuração (Trangmar et al., 1986; Lark e Minasny, 2018). Além disso, diversos estudos têm mostrado que na maioria dos casos os mapas elaborados com a cokrigagem podem ser mais acurados do que os mapas obtidos por krigagem (Vauclin et al., 1983; Teixeira et al., 2013; Minasny e McBratney, 2016, para citar alguns).

A cokrigagem requer que os modelos de variograma sejam estimados para cada variável medida. Além disso, a cokrigagem requer a estimação do modelo de variograma-cruzado dessas variáveis (equação 7), descrevendo a dependência

especial entre as duas (ou mais) variáveis medidas (Rossiter, 2012). Isto não pode ser feito a menos que medições da variável (ou variáveis) mais intensivamente amostrada esteja disponível em cada um dos locais onde a variável esparsamente amostrada é medida (Rossiter, 2012). Essa é a cokrigagem em sua concepção original (Matheron, 1971), cokrigagem ordinária (CKO), e dentro do escopo desse trabalho. Existe outros tipos de cokrigagem que Wackernagel (2003) apresenta com bastante detalhe.

A interpolação com CKO envolve uma abordagem matemática semelhante à da krigagem, com estimativas de pesos de ponderação para ambas as variáveis medidas. Valores estimados para a variável cokrigada (variável principal, isto é, o atributo esparsamente medido o qual se pretende mapear) são obtidos usando a equação 13:

$$\hat{z}_{CKO}(\mathbf{x}_0) = \sum_{i=1}^{N_1(h)} \lambda_i z_1(\mathbf{x}_i) + \sum_{j=1}^{N_2(h)} \lambda_j z_2(\mathbf{x}_j), \quad \text{sendo } \sum_i \lambda_i = 1 \quad \text{e} \quad \sum_j \lambda_j = 0 \quad (13)$$

em que, $\hat{z}_{CKO}(\mathbf{x}_0)$ é o valor estimado da variável principal por CKO no ponto \mathbf{x}_0 ; $N_1(h)$ é o número valores da variável principal usados para predição; λ_i são os pesos associados à variável principal; $N_2(h)$ é o número valores da variável secundária usados para predição; λ_j são os pesos associados à variável secundária; $z_1(\mathbf{x}_i)$ é o valor da variável principal observada no i -ésimo ponto; e $z_2(\mathbf{x}_j)$ é o valor da variável secundária observada no j -ésimo ponto. Assim, o valor interpolado é simplesmente uma combinação linear dos valores medidos para ambas as variáveis em locais próximos do local sem amostragem. E, como com a krigagem, a predição de CKO é realizada sem viés (tendência) e com variância mínima. Essas três características fazem da krigagem, como também da CKO, um interpolador ótimo (BLUP) (Pawlowsky-Glahn e Olea, 2004; Webster e Oliver, 2007). Isso é garantido na CKO ao exigir que os pesos de ponderação da variável auxiliar somam 0 (zero) e os

pesos de ponderação da variável principal (variável mapeada) somam 1 (um), conforme mostrado na equação anterior (equação 13).

O método de CKO também estima a variância dos valores interpolados (variância de cokrigagem) em todos os locais sem amostragem; ou seja, o método gera medidas de erros associado as suas predições. A variância de cokrigagem (σ_0^2) é calculada com a seguinte equação:

$$\sigma_0^2 = \mu_2 + \sum_{i=1}^{N_1(h)} \lambda_{1i} \gamma_{12,i0} + \sum_{j=1}^{N_2(h)} \lambda_{2j} \gamma_{22,j0} \quad (14)$$

em que, μ_2 é o multiplicador de Lagrange para a propriedade 2, λ_{1i} é o fator de ponderação para a *iésima* localização da variável principal, λ_{2j} é o fator de ponderação para a *jésima* localização da variável secundária, $\gamma_{12,i0}$ é a semivariância cruzada a uma distância de separação entre a *iésima* localização e a localização sem amostragem, e $\gamma_{22,j0}$ é a semivariância para a variável secundária a uma distância de separação entre a *jésima* localização e a localização sem amostragem.

Em termos de importância e aplicabilidade da variância de cokrigagem na geostatística, tudo que foi descrito para variância de krigagem (equação 12) é o mesmo que se pode descrever sobre a variância de cokrigagem (equação 14). Logo, para evitar redundância, o leitor interessado sobre a importância e aplicabilidade da variância de cokrigagem deve retornar ao item 2.3.1.4.1.1 em seu último parágrafo.

2.3.1.5 Qualidade dos mapas

A Geoestatística possui ferramentas analíticas muito poderosas para medir, caracterizar e estimar a distribuição espacial de variáveis sob certas condições

(Webster e Oliver, 2007) como anteriormente foi abordado. A popularidade de alguns dos seus interpoladores (krigagem, cokrigagem, etc.) como “melhor preditor linear não-viesado” ou simplesmente “BLUP”, como já mencionado, não deve ser entendida como garantia de que os mapas gerados com estes²⁷ necessariamente apresentem boa qualidade. Esta ressalva é importante porque, dentre outras coisas, os mapas em geral obtidos com a Geoestatística, seja ele mapa de solo (suas variáveis), ou de qualquer outro recurso ou fenômeno, são apenas representações da realidade, que são baseadas em informações fragmentárias (amostragem esparsas) do mundo real, modelagem e interpolação (Oliver e Webster, 1990; McBratney, 1992; Goovaerts, 1997; Webster, 2000; Heuvelink e Webster, 2001). Isso implica que esta representação (conversão de dados discretos em dados contínuos realizada pelo interpolador) – o fenômeno representado no mapa – sempre se apresentará com algum desvio em relação ao fenômeno observado no mundo real, ou seja, cada mapa conterá erros (Olea, 1999; Hengl e Toomanian, 2006; Brus et al., 2011; Heuvelink, 2018). É exatamente mensurando a magnitude dos erros que se permite determinar a qualidade do mapa²⁸ (Olea, 1999; Kienast-Brown e Libohova, 2017; Brus et al., 2011). Em outras palavras, se na mensuração esta representação não corresponder bem à realidade (o erro for grande), a qualidade ou acurácia do mapa será baixa. Por outro lado, se na mensuração a representação corresponder bem à realidade (o erro for pequeno), a qualidade ou acurácia do mapa será alta.

Há muito tempo os mapas de solo têm sido usados para muitas finalidades (Webster, 1997; Goodchild, 2004; IBGE, 2005; Oliveira, 2008; Hartemink, 2013; Ker et al., 2017; Resende et al., 2017). Porém, na atualidade, com os avanços tecnológicos na agricultura (Agricultura de Precisão), isto é, com a possibilidade de operacionalizar as práticas de manejo de maneira localizada os mapas a serem

²⁷ “O uso da palavra ‘krigagem’ em estatística espacial passou a ser sinônimo de ‘otimamente prever’ ou ‘previsão ideal’ no espaço...” (Cressie, 1990). É importante lembrar que a krigagem é apenas outra forma de interpolação por uma soma ponderada de valores locais em que pretende-se encontrar a melhor combinação de pesos para cada local não amostrado, com base em sua relação espacial com os pontos de amostragem e as relações entre os pontos de amostragem, conforme resumido no variograma (O’Sullivan e Unwin, 2010). Isso significa dizer que mesmo a (co)krigagem sendo o melhor preditor linear não-viesado ela só cumpri esse papel quando a variação espacial é devidamente traduzida em um modelo de variograma correto, daí a importância do que foi discutido nos subitens do item 2.3.1.3.4 desta Tese.

²⁸ Webster R (2018) (pesquisador da Rothamsted, Harpenden, Reino Unido). Comunicação Pessoal.

utilizados²⁹ (que atendem esse tipo de tecnologia) passaram a ter que apresentar qualidade (Robert et al., 1993; Arnold e Wilding, 1991; Webster, 1997; Pierce e Nowak, 1999; Oliver, 2010; Souza e Souza, 2011; Grego et al., 2014; Resende e Coelho, 2014, 2017; Paterson et al., 2018; Whelan, 2018). Porém, não é possível fazer qualquer afirmação sobre ser “este” ou “aquele” mapa o de melhor qualidade, a menos que se verifique ou quantifique a qualidade (Pierce e Nowak, 1999; Olea, 1999; Robinson e Metternicht, 2005; Brus et al., 2011). E nem também apontar ser “esta” ou “aquela” metodologia ou protocolo (de amostragem, ou de interpolação, etc.) que permite os melhores mapas, a não ser que isso tenha sido verificado (Olea, 1999; Robinson e Metternicht, 2005). Para isso, a forma prática que se tem para fazer esta verificação é através do emprego de algum método de validação³⁰ (Pierce e Nowak, 1999; Olea, 1999; Robinson e Metternicht, 2005; Oliver, 2010; Brus et al., 2011; Heuvelink, 2018). Ou seja, com um procedimento que adequadamente possa gerar estatísticas de erros ou de acertos, conforme no item seguinte será abordado. A validação é algo relativamente simples e abrange alguns aspectos que também merecem ser apresentados.

2.3.1.5.1 Validação

Na literatura “validação” possui diversas definições (por exemplo, Oreskes et al., 1994; Addiscott et al., 1995). Existe o entendimento de que etimologicamente o termo é inadequado (Oreskes et al., 1994; Rossiter, 2017; Heuvelink, 2017; Lagacherie, 2017); assim como também existe os que consideram o termo como apropriado (Addiscott et al., 1995; Lark, 2017; Brus, 2017; Minasny, 2017). Porém, não será tentado aqui dar definições das diferentes noções. Apenas optou-se por descrever o que é validação dentro da abrangência que o procedimento foi utilizado para o estudo relatado neste trabalho. Neste sentido, a validação é o procedimento

²⁹ Mesmo o mapeamento de solos norte-americano cujas unidades de mapeamento são Séries de solo – “...considerado como um dos mais eficientes meios de compartimentação do meio físico...” (Lepsch, 2013) – o National Research Council (1997) concluiu que ele não satisfaz os requisitos de dados que atenda a Agricultura de Precisão. Diversos estudiosos documentaram a existência de significativa variabilidade espacial dentro de unidades de mapeamento de solo norte-americano (pureza estatística das unidades é inferior a 50%) (Powell e Springer, 1965; Edmonds e Lentner, 1986; Nordt; Jacob e Wilding, 1991; Young et al., 1997). Além disso, Beckett e Webster (1971) mencionam que a variabilidade difere para cada propriedade do solo.

³⁰ Webster R (2018) (pesquisador da Rothamsted, Harpenden, Reino Unido). Comunicação Pessoal.

no qual as previsões do mapa de solo (os valores preditos) são comparadas com os valores observados (referido também na literatura como valores reais, valores verdadeiros, valores originais, etc.) produzindo, através dessa comparação, medidas de qualidade do mapa (Heuvelink, 2018). Essas medidas podem ser utilizadas para alguns propósitos³¹, mas, em última análise, essas medidas indicam quão acurado o mapa é em média para a área de mapeamento, ou seja, qual é o erro esperado em um local selecionado aleatoriamente na área de mapeamento (Olea, 1999; Brus et al., 2011). Isso significa que as medidas de qualidade do mapa obtidas na validação são medidas globais: cada medida de qualidade dá um valor para todo o mapa (Olea, 1999; Brus et al., 2011). Esta é a abordagem que mais comumente tem sido utilizada porque permite comparação objetiva entre mapas (Keskin e Grunwald, 2018). Porém, como foi mencionado que o procedimento de validação pode atender diferentes propósitos, mais a frente será explanado distinções metodológicas existentes em função dos propósitos.

Embora a validação (acurácia global) permita o que foi descrito, ela não fornece informações sobre como a acurácia varia em toda a superfície do mapa, ou seja, ela não fornece informação da acurácia local (Yang e Hodler, 2000). Yang e Hodler (2000) argumentam que, em muitos casos, medidas de erro ponto a ponto ao longo da superfície podem ser mais úteis do que as medidas de erro global. Principalmente quando na comparação da qualidade entre mapas seus erros globais apresentam-se idênticos. Isso gerará a indagação: qual mapa é o mais confiável? Nesta circunstância (a qual os erros globais são idênticos), aquele mapa que apresentar erros mais uniformemente distribuídos deverá se considerado o mais confiável frente aos que apresentarem os erros de forma concentrada (Yang e Hodler, 2000). De modo geral a seguinte sequência de procedimentos sintetiza o que é encontrado na literatura para avaliação da acurácia local (Yang e Hodler, 2000): o cálculo do residual (erro) para cada um de todos os pontos dos dados; a interpolação desses erros e geração de um mapa por um método exato; e a análise do padrão visual dos erros locais (Webster e Oliver, 2007). Ou, ao invés disso, a

³¹ Como síntese, medidas de qualidade do “mapa” podem ser usadas para, por exemplo, avaliar densidades de amostragem, configurações da amostragem, viabilidade de usar transformações no conjunto de dados, avaliar ajustes de modelos às semivariâncias, avaliar tipos de modelos para o variograma, quantidade de lags no variograma, vizinhança para interpolação, tipos de interpoladores, etc.

acurácia local também pode ser avaliada de outra forma visual. Para isso, é necessário a construção de um gráfico que permita comparar a distribuição dos valores preditos e a distribuição dos valores reais (observados), como por exemplo com um gráfico quantil-quantil. Comparando as distribuições e a inclinação da reta é possível verificação de discrepâncias nas distribuições mencionadas. Em outras palavras, é possível visualizar no gráfico se os erros se apresentam normalmente distribuídos (não enviesados) ou não (grandes desvios localizados) (Isaaks e Srivastava, 1989). Embora este tipo de avaliação em certos contextos seja considerada eficaz, é importante ressaltar que ela se constitui uma avaliação qualitativa e por isso sempre envolverá subjetividade.

Em relação a validação global, os métodos mais comumente utilizados na Geoestatística para este propósito são a validação-cruzada e a validação (Cressie, 1993; Keskin e Grunwald, 2018). Neste trabalho ao invés de “validação” optou-se por chamá-la de “validação-externa”. E, o termo “validação” deve ser entendido nesse trabalho como um termo geral, ou seja, uma menção genérica. Como será visto mais a frente, a validação-cruzada e a validação-externa, são métodos para a validação global cujas informações geradas são usadas para finalidades distintas (específicas).³² Referências como Isaaks e Srivastava (1989), Cressie (1993), Webster e Oliver (2007), Hengl (2009), Brus et al. (2011) contêm mais detalhes sobre o assunto.

2.3.1.5.1.1 Validação-cruzada e Validação-externa

Em relação a validação-cruzada, a idéia básica de como o método realiza o procedimento de validação é muito simples.³³ Primeiro, uma estimativa é obtida pelo interpolador usando todos os dados (conjunto completo de amostras). Em seguida, uma amostra (ou um grupo de amostras) é removida dos dados e uma nova estimativa é obtida usando as amostras restantes. Este processo é repetido até que todos os dados tenham sido removidos. Essa sucessão de novas estimativas permitem realizar a comparação dos valores estimados (preditos) em relação aos

³² Webster R (2018) (pesquisador da Rothamsted, Harpenden, Reino Unido). Comunicação Pessoal.

³³ Webster R (2018) (pesquisador da Rothamsted, Harpenden, Reino Unido). Comunicação Pessoal.

valores verdadeiros (observados), e gerar, no final, índices ou medidas gerais de erros ou acurácia (Oliver e Webster, 2014). A validação-cruzada (VC) requer a solução do seguinte cálculo:

$$VC = Z_{obs} - Z_{est} \quad (15)$$

em que, Z_{obs} é o valor observado, Z_{est} é o valor estimado; VC sendo obtido um a um de valores dos pontos amostrados para modelagem e interpolação (análises geoestatísticas), conforme procedimento anteriormente explicado.

Em relação a validação-externa, a idéia básica é um pouco diferente da validação-cruzada, e pode ser considerada como relativamente mais simples.³⁴ Inclui uma etapa realizada antes do procedimento. Esta etapa consiste em dividir o conjunto de dados medidos originais (conjunto completo de dados) em dois conjuntos de dados: conjunto de dados de treinamento (T) e conjunto de dados de valiação-externa (V). Não existe na literatura uma orientação sobre um índice padrão³⁵ para esta divisão (Keskin e Grunwald, 2018). A proporção de conjunto de dados de “T : V” relatado nos trabalhos publicados variam de “65%:35%” à “90%:10%” (Keskin e Grunwald, 2018; Bernoux et al., 2007) e são resultantes de uma seleção aleatória dos dados (de modo estratificado dentro da área) para a construção do conjunto de V. Na realidade, esta flexibilidade envolvida na definição do tamanho dos conjuntos de dados, não deve desconsiderar a quantidade de amostras mínima requerida para obtenção de um variograma confiável com o conjunto de dados de T. Muito menos desconsiderar a quantidade de amostras (dados de T) que permita cobertura da área em que possa captar a dependência espacial da variável a ser mapeada (Webster, 2008). Dessa forma, com o conjunto maior, ou seja, os dados do conjunto de T, é usado para realizar as predições; e, com o conjunto de dados menor, conjunto V, usa-se seus valores para realizar comparações com os valores preditos.³⁶ Com base nesta confrontação obtem-se os

³⁴ Webster R (2018) (pesquisador da Rothamsted, Harpenden, Reino Unido). Comunicação Pessoal.

³⁵ Um índice que tenha uma fundamentação estatística (McBratney et al., 2003). Ou seja, baseado em estudos que avaliaram qual a divisão mais apropriada (Biswas e Zhang, 2018).

³⁶ Webster R (2018) (pesquisador da Rothamsted, Harpenden, Reino Unido). Comunicação Pessoal.

erros. E, com base nesses erros, gera-se os índices ou medidas gerais de erros ou acurácia. A validação-externa (VE) requer a solução do seguinte cálculo:

$$VE = Z_{obs} - Z_{est} \quad (16)$$

em que, Z_{obs} é o valor observado, Z_{est} é o valor estimado; VE sendo obtido um a um apenas de valores dos pontos amostrados para validação (conjunto de dados de validação; ou seja, dados independentes), conforme procedimento anteriormente explicado.

Apesar da validação ser um assunto presente em muitos trabalhos, e ser o meio pelo qual se avalia as estimativas e previsões estatísticas desde os artigos de Stone (1974) e Geisser (1975), verifica-se que as vantagens e as desvantagens dos métodos de validação utilizados nas pesquisas de mapeamento de solo raramente são apresentadas. Por um lado pragmático, este fato pode ter relação com a facilidade com que se verifica na literatura, um grande número de trabalhos que fazem uso de métodos de validação não adequados para o caso (Goovaerts e Kerry, 2010). Por exemplo, a validação-cruzada não é um método indicado para um diagnóstico final na Geoestatística³⁷ (Davis, 1987). Pelo contrário, a validação-cruzada é um método considerado exploratório³⁸ (Davis, 1987; Hengl, 2009). Ela é indicada para avaliar a viabilidade de certos procedimentos na modelagem (o tratamento de dados atípicos adotado; a transformação dos dados escolhida; a quantidade de lags no variograma assumida; determinado ajuste para o modelo; o modelo de variograma adotado; decisão relativa a anisotropia; etc.) e procedimentos na estimação (o número de vizinhança de procura; o interpolador utilizado; etc.). Por outro lado, a validação-externa é considerado um método complementar na análise geoestatística; porém, de extrema importância (Brus et al., 2011). Este método é indicado para validação dos mapas, ou seja, quantificar a acurácia destes³⁹ (Brus et

³⁷ Webster R (2018) (pesquisador da Rothamsted, Harpenden, Reino Unido). Comunicação Pessoal.

³⁸ Webster R (2018) (pesquisador da Rothamsted, Harpenden, Reino Unido). Comunicação Pessoal.

³⁹ É importante mencionar um procedimento que normalmente é realizado mas que raramente é declarado: a inspeção visual de mapa. É um procedimento que antecede a quantificação da acurácia do mapa (MacEachren, 1995; Yang e Hodler, 2000). Esse procedimento tem por propósito evitar trabalho desnecessário. Por exemplo, se um mapa se apresentar incoerente ou com presença de algum artefato (área de coloração não condizente com a legenda do mapa), não fará nenhum sentido saber sua acurácia. Em certos softwares a inspeção é um procedimento rápido, envolve realizar comparações da distribuição de valores apresentados no mapa com a

al., 2011). Suas informações são consideradas confiáveis para esta finalidade em decorrência de suas características (Brus et al., 2011), cujas mesmas mais a frente serão destacadas. O procedimento da validação-cruzada é considerado vantajoso porque não exige uso de dados adicionais específicos para a validação.⁴⁰ Além do mais, suas informações são consideradas confiáveis para apoio a modelagem⁴¹ (Davis, 1987; Olea, 1999; Oliver e Webster, 2014) conforme também já dito. Esta vantagem se dá em função de que toda a base de dados na qual o modelo foi calibrado é usada no procedimento de validação⁴² (De Paula, 2017). Entre as suas desvantagens está que, usada para a avaliação da qualidade de mapas (diagnóstico final), suas estatísticas podem não se revelarem realistas, ou seja, serem tendenciosas (Heuvelink, 2018). Isso ocorre principalmente porque os dados trabalhados na validação, as “referências”, são os mesmos que alimentaram parte do processo geoestatístico (calibração do modelo) (Heuvelink, 2018). Em contrapartida, já a validação-externa, sua principal vantagem é que suas estimativas têm maior probabilidade de serem imparciais (Heuvelink, 2018; Brus et al., 2011). Isto se dá exatamente porque os dados que servem de base para comparação, os dados de “referência”, são dados independentes (Heuvelink, 2018; Brus et al., 2011). Isto é, são dados que ficam fora (não participam) tanto da variografia como da interpolação. Por isso o nome: “validação externa”. Porém, os custos envolvidos na obtenção de um quantitativo de dados que permita destinar uma parte para finalidade de validação (dados para V) do mapa é considerado sua principal desvantagem.⁴³ Esta desvantagem pode ser melhor notada quando por algum motivo o quantitativo de dados disponíveis para o mapeamento encontra-se no limite mínimo necessário para as análises geoestatísticas. Dentro dessa situação, a decisão em destinar parte dos dados para validação compromete as análises e se

distribuição de valores do mapa de pontos amostrados (conjunto de dados), e, também a verificação da presença ou não de artefato. No entanto, não é suficiente concluir sobre a qualidade de um mapa com apenas inspeção visual, pois, muitos mapas que têm boa aparência visual, podem ter erros muito sérios. A inspeção visual de um mapa pode também ser realizada como um procedimento diferente do apresentado. Ele pode ser realizado para relacionar informações espaciais do mapa elaborado com outras informações espaciais (MacEachren, 1995; Yang e Hodler, 2000). Porém, neste caso, a inspeção visual é feita posterior a verificação da acurácia.

⁴⁰ Webster R (2018) (pesquisador da Rothamsted, Harpenden, Reino Unido). Comunicação Pessoal.

⁴¹ Webster R (2018) (pesquisador da Rothamsted, Harpenden, Reino Unido). Comunicação Pessoal.

⁴² Webster R (2018) (pesquisador da Rothamsted, Harpenden, Reino Unido). Comunicação Pessoal.

⁴³ Webster R (2018) (pesquisador da Rothamsted, Harpenden, Reino Unido). Comunicação Pessoal.

torna sem sentido realizar a própria validação. Nesse contexto, a validação-cruzada, mesmo não sendo o método mais indicado para avaliação da qualidade do mapa, constitui-se como um recurso a ser usado porque não produzirá redução nos dados (Biswas e Zhang, 2018).

Por fim, é importante relatar que existe sugestão diferente da que foi apresentada neste trabalho para obtenção do conjunto de dados para validação. Ao invés da divisão do conjunto de dados totais, em conjuntos de dados de V e T, como foi mencionado, alguns autores sugerem fortemente que se obtenha o conjunto de dados de V por uma amostragem adicional de probabilidade (Brus et al., 2011; Heuvelink, 2018). O argumento que fundamenta essa orientação é de que a amostragem para validação pode ser melhor selecionada por esse tipo de amostragem, de modo que os índices de qualidade do mapa possam ser estimados por inferência baseada em projeto. As estimativas baseadas em projeto são livres de modelo, não são feitas suposições de modelagem e, como consequência, nenhuma objeção pode ser feita contra os índices de qualidade de mapa estimados (Heuvelink, 2018). Todavia, em relação a esta consideração, é necessário lembrar que as estimativas dos índices de qualidade do mapa não são nem mais e nem menos importantes do que as próprias estimativas do mapa (predições espaciais), o que significa dizer que, o critério de amostragem adotado, deve ser de tal modo que, o conjunto de dados totais necessariamente deve possuir dados que pertença a população que se pretende mapear, logo, dividí-la nos conjuntos de dados T e V, ambos pertencerão a mesma população visada, o que torna sem sentido realizar uma segunda amostragem (amostragem adicional) que será usada para checagem desta população.

2.3.2 Argila do solo – uma variável estratégica para o manejo localizado

Nas considerações feitas até agora, foi enfatizada a necessidade de particionar o campo agrícola em seguimentos considerados mais homogêneos com base na caracterização e mapeamento da variabilidade espacial do solo com alta acurácia. Por outro lado, porém, a realidade na agricultura é que normalmente não é

viável realizar o mapeamento de cada uma das variáveis de interesses, como por exemplo, a saturação por bases, o teor de fósforo disponível, o teor de potássio trocável, a concentração de alumínio fitotóxico, etc. Naturalmente, em função disso, surge indagações como: o que mapear então?, ou, qual procedimento adotar para viabilizar um mapeamento que possa representar as variáveis do solo que são de interesses manejar em conjunto, isto é, que potencialmente reúna todas ou as principais informações chaves?

Segundo Whelan (2003), a variabilidade textural do solo é muito importante porque influencia diretamente o potencial de rendimento dos cultivos. Dentre as frações texturais do solo, a argila (partículas com $\varnothing < 0,002$ mm) é considerada como uma das propriedades mais importantes⁴⁴ (Resende et al., 2007) porque a ela está relacionada (i) à capacidade de troca de cátions dos solos (Russell, 1977; White, 2005), (ii) a retenção e disponibilidade de ânions, como o fosfato (Novais e Smyth, 1999), (iii) o controle de propriedades hidráulicas do solo, como armazenamento de água e disponibilidade para plantas cultivadas (Silva et al., 2001; Saxton e Rawls, 2006), (iv) a estabilidade de agregados e qualidade física e condições biológicas do solo (White, 2005; SSSA, 2008), e (v) o estado de fertilidade do solo e, portanto, afeta, pode-se dizer, todo o potencial produtivo do solo em questão (Davey, 1990; Chen et al., 2004). Diante disso, ignorar tais relações com a argila do solo é no mínimo um grande contrasenso.

Como outras variáveis ambientais, o teor de argila do solo muda no espaço. Ou seja, sua quantidade no espaço muda conforme muda a distância. Essa variabilidade espacial da argila no solo pode levar a diferenças espaciais na qualidade do solo (Kettler et al., 2001) conforme se mencionou. Em decorrência disso, a caracterização detalhada e precisa do padrão de variabilidade espacial da argila do solo pode, portanto, ser útil na esquematização de práticas de manejo localizados no campo com maior precisão e redução de custos (Chen et al., 2004). No âmbito da agricultura convencional, informações sobre os teores de argila do solo são consideradas das mais estratégicas há bastante tempo (Raij et al., 1996; Tomé Júnior, 1997; Lopes e Guilherme, 2000; Resende et al., 2007), fato é que

⁴⁴ As propriedades do solo encontram-se de tal forma inter-relacionadas que pode não haver muito sentido em dizer que uma é mais importante do que outra. Até porque, as frações granulométricas de tamanhos maiores que a argila constitui a matéria prima que intemperismo e pedogênese produzem a própria fração argila do solo.

existe orientações técnicas estabelecidas para certos insumos que suas quantidades aplicadas no campo (definição de doses) são ajustadas com base no teor de argila presente no mesmo (Alvarez V et al., 1999; Lopes e Guilherme, 2000; Sousa e Lobato, 2004). Isto coloca a argila do solo e sua variabilidade espacial como uma propriedade estratégica que pode ser usada para zonedar o manejo no campo com economia nos custos com obtenção de mapas. Lopes e Guilherme (2000) por exemplo, pontuam as seguintes utilizações práticas sobre o teor de argila no campo: (i) ajustar doses de adubação fosfatada corretiva quando esta é necessária; (ii) ajustar doses de calcário quando há necessidade de correção da acidez e reposição de bases; (iii) ajustar doses de gesso agrícola quando há necessidade de neutralização de alumínio em profundidade; (iv) decidir pelo número de parcelamento das adubações nitrogenadas e potássicas; (v) avaliar probabilidade de problemas com deficiência de micronutrientes para a cultura; (vi) avaliar possibilidade das plantas passarem por estresses hídricos e problemas com drenagem; etc. Neste sentido as informações espaciais sobre os teores de argila presentes no campo podem ser utilizadas de forma segura para operacionalização de práticas localizadas de manejo agrícola com o uso de tecnologia de aplicação a taxa variável.

2.3.3 Suscetibilidade magnética e utilização no estudo dos atributos do solo

Como anteriormente enfatizado, as propriedades do solo não são estáticas nem homogêneas no espaço e no tempo. Por isso os custos analíticos são frequentemente um fator limitante para qualquer tentativa de abordar a variabilidade do solo, especialmente em aplicações a grandes áreas, envolvendo a elaboração de mapas com alta acurácia, os quais normalmente demandam análises de um grande número de amostras de solos, como ocorre no âmbito da Agricultura de Precisão (Viscarra-Rossel e McBratney, 1998; Plant, 2001). Todavia, na atualidade, os profissionais envolvidos com agricultura têm uma variedade de métodos analíticos à disposição, em comparação com os profissionais de algum tempo atrás que muito poucos recursos podiam contar. Hoje em dia o problema maior tem sido saber o que escolher. Pois, quando os métodos de análises de solo têm o potencial de serem

operacionalizados com economia de tempo, gerando resultados com custos relativamente mais baixos e não serem geradores de resíduos que possam poluir o meio ambiente, normalmente, não produzem resultados analíticos com boa performance em termos de acurácia. À isto, especialistas sobre o assunto têm considerado ser “um passo atrás” na tentativa de promover avanços à Agricultura de Precisão; visto que, muitos profissionais têm feito “vistas grossas” a esta questão no afã equivocado de que as possíveis perdas de acurácia resultantes dos métodos mais baratos escolhidos são normalmente compensadas pelos ganhos que se pode ter ao operacionalizar o manejo agrícola de modo localizado. Este é um assunto aparentemente polêmico, que requer pesquisa específica,⁴⁵ e está fora do escopo dessa revisão.

No entanto, pesquisas públicas realizadas no Brasil têm apontado o uso de métodos para análise da suscetibilidade magnética (SM) do solo como um importante recurso pedométrico para quantificação indireta de propriedades químicas, físicas e mineralógicas do solo, o qual proporciona às análises: economia de tempo, diminuição de custos, isenção de riscos ao meio ambiente e geração de resultados analíticos com boa acurácia (Siqueira et al., 2010; Siqueira e Marques Júnior, 2014).⁴⁶

Diante da afirmação feita, é importante explicar o que é SM. Dearing (1994) afirmou que tudo ao nosso redor é “magnético”. Ele explica que da mesma forma como se pode descrever objetos e materiais pelo seu tamanho, cor ou composição química, também é possível descrevê-los por suas propriedades magnéticas. Isso pode ser uma surpresa para qualquer um que não tenha formação em física e áreas correlatas, porque na vida cotidiana geralmente se explora o magnetismo de forma bastante limitada, em termos apenas, por exemplo, de ímãs e metais, ou quando se utiliza de cartões (cartões como de crédito, cartões de banco, etc.) que as informações neles contidas estão em forma de “trilhas” magnéticas. Dearing (1994) disse algo que de fato normalmente ocorre. Ele disse que no cotidiano as pessoas

⁴⁵ Apesar de ser um assunto verbalmente muito debatido nos eventos (críticas aos métodos e procedimentos pouco acurados), praticamente nada é encontrado (no mesmo “tom”) em formato escrito (artigos científicos, Dissertações, Teses, etc.) sobre este assunto.

⁴⁶ Informações sobre a aplicação da suscetibilidade magnética à Ciência do Solo e agricultura, podem ser consultadas livremente no site: <http://www.csme.com.br/> , incluindo disponibilização de artigos científicos, Dissertações, Teses, etc.

em geral não pensam no comportamento magnético das rochas ou do solo, nem da poeira do ar que elas respiram, muito menos nas propriedades magnéticas da água dos rios ou das folhas das árvores ao seus redores. Porém, mesmo que as pessoas não parem para pensar, a realidade é que toda a matéria é afetada por um campo magnético. O efeito pode ser extremamente fraco ou até negativo, mas existe e pode ser medido facilmente, afirmou Dearing (1994). Diante disso, durante os anos 1970 e 1980, os cientistas perceberam que as propriedades magnéticas eram úteis para descrever e classificar todos os tipos de materiais ambientais. Foi a partir dessa percepção, e de resultados de estudos, que em anos posteriores foram inventados instrumentos para este propósito – considerados relativamente simples – que passaram a ser utilizados em laboratórios, universidades⁴⁷ e outras instituições em muitos locais no mundo (Jordanova, 2017). Com o uso desses aparelhos, passou a ser possível medir a "magnetização" (susceptibilidade magnética) de materiais diversos e estabelecer diversas correlações as quais permitem ser utilizadas para o reconhecimento indireto de certas propriedades inerentes a tais materiais (Dearing, 1994).

De modo sucinto para uma noção a respeito de funcionamento, estes aparelhos produzem um campo magnético e através de um sensor “medem a facilidade” com que o material se magnetiza na presença do referido campo (Dearing, 1994). Essa reação ao campo magnético é possível devido a cinco tipos de comportamentos magnéticos que a nível atômico os materiais podem expressar: diamagnetismo, paramagnetismo, ferromagnetismo, ferrimagnetismo e antiferromagnetismo (Dearing, 1994).

Em relação ao solo, seu comportamento em termos de SM decorre em função de como se comportam os elétrons nas camadas eletrosféricas nos átomos que predominam em cada tipo de mineral que faz parte de sua mineralogia, conforme explicam cientistas do solo (Resende et al., 1988; Siqueira et al., 2010; Siqueira e Marques Júnior, 2014):

⁴⁷ Os trabalhos de Resende et al. (1988), Fontes et al. (2000) e Siqueira et al. (2010) são testemunhos dos esforços científicos no âmbito acadêmico. Mostraram que o instrumento para mensuramento da SM pode ser facilmente construído. Assim como comprovação científica de que o instrumento fornece medidas de SM com alta acurácia (Siqueira et al., 2010).

- a) Nos minerais diamagnéticos (por exemplo, o feldspato, a muscovita, o quartzo), o número de spins eletrônicos alinhados numa direção é igual ao número de spins eletrônicos na direção oposta. O campo magnético externo não exerce influência sobre os spins nesse caso (os campos magnéticos anulam-se), mas inverte o movimento orbital dos elétrons. Eles apresentam respostas ditas fracas (às vezes até negativa) à aplicação do campo magnético.
- b) Nos minerais paramagnéticos (por exemplo, a olivina, piroxênio, anfibólio, biotita), a presença de um campo magnético externo faz com que os spins se alinhem (elétrons giram no mesmo sentido), e mesmo após a retirada do campo magnético, alguns spins permanecem alinhados. Eles não são capazes de preservar a magnetização remanescente. Porém, em meio, muitos silicatos, ao invés de exibirem comportamento paramagnético, exibem comportamento ferromagnético o que lhes confere capacidade de preservar a magnetização remanescente.
- c) Os minerais ferromagnéticos (por exemplo, o ferro e o cobalto) são considerados um caso especial de paramagnetismo. Nesses minerais, após a retirada do campo magnético, os spins permanecem alinhados fazendo com que o mineral possua um grande valor de magnetização remanescente.
- d) Nos minerais ferrimagnéticos (por exemplo, a magnetita e a maghemita), os spins não estão emparelhados, assim prevalece o campo magnético resultante do maior número de spins no mesmo sentido. De acordo com Mullis (1977) os minerais magnéticos com comportamento magnético mais evidente no solo são a magnetita e a maghemita.
- e) Os minerais antiferromagnéticos (por exemplo, a hematita e a goethita) não apresentam propriedades magnéticas. Entretanto, minerais antiferromagnéticos, mesmo que muito fraco, podem se comportar como que minerais ferrimagnéticos (por exemplo, a hematita), isso se dá devido ao antiparalelismo imperfeito no alinhamento de momentos magnéticos opostos.

De qualquer modo, em relação a esta categorização para os minerais do solo, é importante lembrar que os minerais ferri e antiferromagnéticos ocorrem no solo de modo dispersos na forma de partículas muito finas entre os minerais paramagnéticos

e diamagnéticos, de argila e das demais partículas (Ozdemir e Banerjee, 1982 citado por Fontes et al., 2000).

De acordo com Mathé et al. (2006) os minerais magnéticos podem ser considerados como micromarcadores de atributos do solo, e que é possível estabelecer correlações entre a SM e os atributos químicos, físicos e mineralógicos do solo. Dearing et al. (1996) relataram que o uso dessas correlações é ideal para a realização de estudos que requerem a análise de grande número de amostras de solo. Ou seja, ao invés da obtenção de informações analíticas de determinadas propriedades do solo por meio direto com análises convencionais, usa-se da análise de SM do solo para por correlações estimar valores para as propriedades desejadas. A estas correlações conhecidas (faixas de valores de SM associadas a uma dada propriedade do solo) denomina-se “assinatura magnética”.

No Brasil muitos estudos importantes têm sido realizados visando ampliar esse conhecimento. Por exemplo, Fontes et al. (2000) analisaram como diferenças na mineralogia presente em alguns solos são responsáveis por diferenças nas propriedades magnéticas destes. Baseado nas informações que levantaram, esses pesquisadores inferiram que solos com maior fração argila favorecem a formação de minerais de comportamento magnético mais evidente. Isto significa dizer que informações sobre a SM podem ser utilizadas para inferências sobre o teor de argila do solo. Eles também verificaram que a SM variou em função dos tipos de classes de solo que estudaram, assim como dos materiais de origem destes. Eles constataram que solos originados de material basáltico apresentam maior SM.

Peluco et al. (2015) visando meios práticos para realização de mapeamento do fósforo adsorvido no solo, investigaram o uso da SM como medida para quantificação indireta desse íon no solo. Eles concluíram que a SM do solo pode ser usada como medida auxiliar para o mapeamento e na identificação de áreas com diferentes potenciais de adsorção de fósforo.

Em estudo nessa mesma linha de mapeamento de solos Cortez et al. (2011) utilizaram a SM para definição de zonas específicas de manejo em uma área de Latossolo Vermelho distrófico com base em estimativas indiretas de diversos atributos e parâmetros do solo, como por exemplo: estabilidade dos agregados; teor de argila; condutividade hidráulica; macro e microporos; volume total de poros;

densidade do solo; resistência do solo a penetração; teor de água; teor de matéria orgânica; pH; teores de cálcio, magnésio e potássio trocáveis; teor de fósforo disponível.

Siqueira et al. (2010) avaliaram em uma área com domínio de Argissolo Vermelho-Amarelo correlações da SM com o teor de argila, capacidade de troca catiônica, capacidade de adsorção de fósforo e teor de hematita. Em decorrência da alta correlação observada, eles preconizaram que a SM pode ser utilizada com confiança na estimativa indireta das referidas propriedades.

Mais recentemente, em um estudo incluindo investigação de desempenho dos métodos de espectroscopia visível, infravermelho próximo e SM para determinação de atributos do solo, Bahia et al. (2017) observaram entre tais métodos que as previsões por SM apresentaram melhor acurácia para as principais variáveis estudadas: óxido de ferro, argila e carbono total.

No entanto, apesar de toda a perspectiva que é considerada muito positiva em relação à utilização de métodos analíticos considerados rápidos, limpos, acurados e de baixo custo para quantificação indireta de atributos do solo, é importantíssimo levar em consideração suas eventuais limitações em termos de contextos a serem utilizados; como por exemplo, características do solo que são interferentes (ou que às vezes não convenientes), condições ambientais inadequadas, enfim, fatores que podem levar o método a apresentar resultados considerados insatisfatórios (baixa acurácia, instabilidade, etc.). Nesse caso, em relação ao uso da SM do solo – aplicada especialmente na atividade de mapeamento da variabilidade espacial de propriedades do solo – informações contidas em Siqueira e Marques Júnior (2014) e Rosa (2014) podem conduzir os interessados a um maior esclarecimento em relação ao assunto. Além do mais, são materiais preparados com linguagem para fácil compreensão e de disponibilidade livre na Web para qualquer um acessar.

3 MATERIAL E MÉTODOS

3.1 Descrição da área de estudo

Este estudo é uma continuidade de prévios trabalhos (Siqueira, 2013; Teixeira, 2015; Bahia, 2016) desenvolvidos no grupo de pesquisa CSME (Caracterização do Solo para Fins de Manejo Específico) a partir de uma base de dados comum de uma área agrícola localizada no município de Guatapar-SP (Figura 4A). As coordenadas centrais da rea so 21 28' 40" S e 48 01' 38" O. Pela classificao de Thornthwaite (1948) o clima local pode ser definido como B1rB'4a', Tipo Mesotrmico mido, com pequena deficincia hdrica, sendo a evapotranspirao de vero menor que 48% da evapotranspirao anual. A cobertura vegetal original (anteriormente  antropizao) era floresta tropical semidecidual. A rea vem sendo utilizada h mais de 10 anos com cultivo comercial de cana-de-acar (*Saccharum officinarum L.*) com colheita mecanizada.

A rea est inserida na provncia geomorfolgica do Planalto Ocidental Paulista. O material de origem est relacionado  transio Basalto do Grupo So Bento Formao Serra Geral (SG), Depsito Colvio-Eluvionar (DCE) e Depsito Aluvionar (DA) (GEOBANK, 2006) (Figura 4B). Incurses tcnicas para comprovao das informaoes geolgicas foram realizadas na rea. O mapa pedolgico (escala de execuo 1: 12.000) gerado pelo Centro de Tecnologia Canavieira (CTC) documenta os seguintes domnios de solos do Sistema Brasileiro de Classificao de Solos (SiBCS) (Santos et al., 2006) para as unidades de mapeamento: Latossolo Vermelho Amarelo textura mdia (LVAd), Latossolo Vermelho distroftico textura mdia (LVd), Latossolo Vermelho distrfrrico textura argilosa (LVdf), Latossolo Vermelho eutrofrrico textura argilosa (LVef), Neossolo Quartzarnico rtico distrftico textura arenosa (RQod) (Figura 4C). A topografia na rea apresenta um desnvel norte-sul de 130 m (Figura 4D) tendo suas altitudes respectivamente variando de 649 a 519 m acima do nvel do mar.

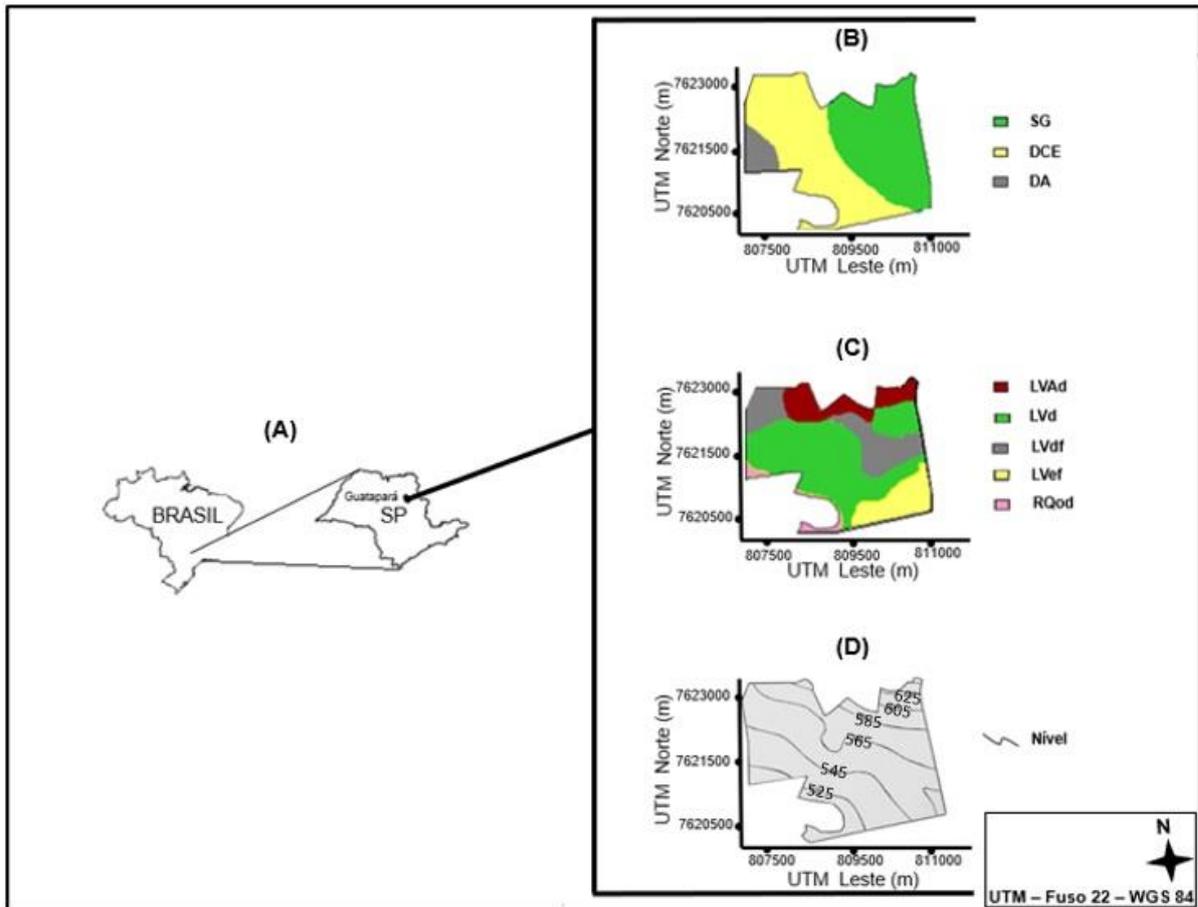


Figura 4. Caracterização da área de estudo. Mapas da localização geográfica (país: Brasil; Estado: São Paulo; município: Guatapará) (A); Mapas da área: (B) geológico com escala de execução 1: 500.000 (materiais de origem: SG, DCE e DA), (C) pedológico com escala de execução 1: 12.000 (solos dominantes: LVAd, LVd, LVdf, LVeF e RQod) e (D) topográfico (altitude de 519 a 649 m acima do nível do mar; e os valores no interior do mapa marcam as altitudes intermediárias).

3.2 Amostragem de solo na área de estudo

A amostragem de solo na área estudada seguiu uma malha regular previamente definida contendo 371 pontos separados por distâncias mínimas variando de 145 m a 174 m, georreferenciados, cobrindo seus 870 ha (Figura 5). A profundidade empregada de coleta nos pontos, 0 - 0,25 m, se deu em função de ser a que tem sido utilizada para a condução do manejo do solo na área em questão. Todas as amostras foram conduzidas para análises em laboratório.

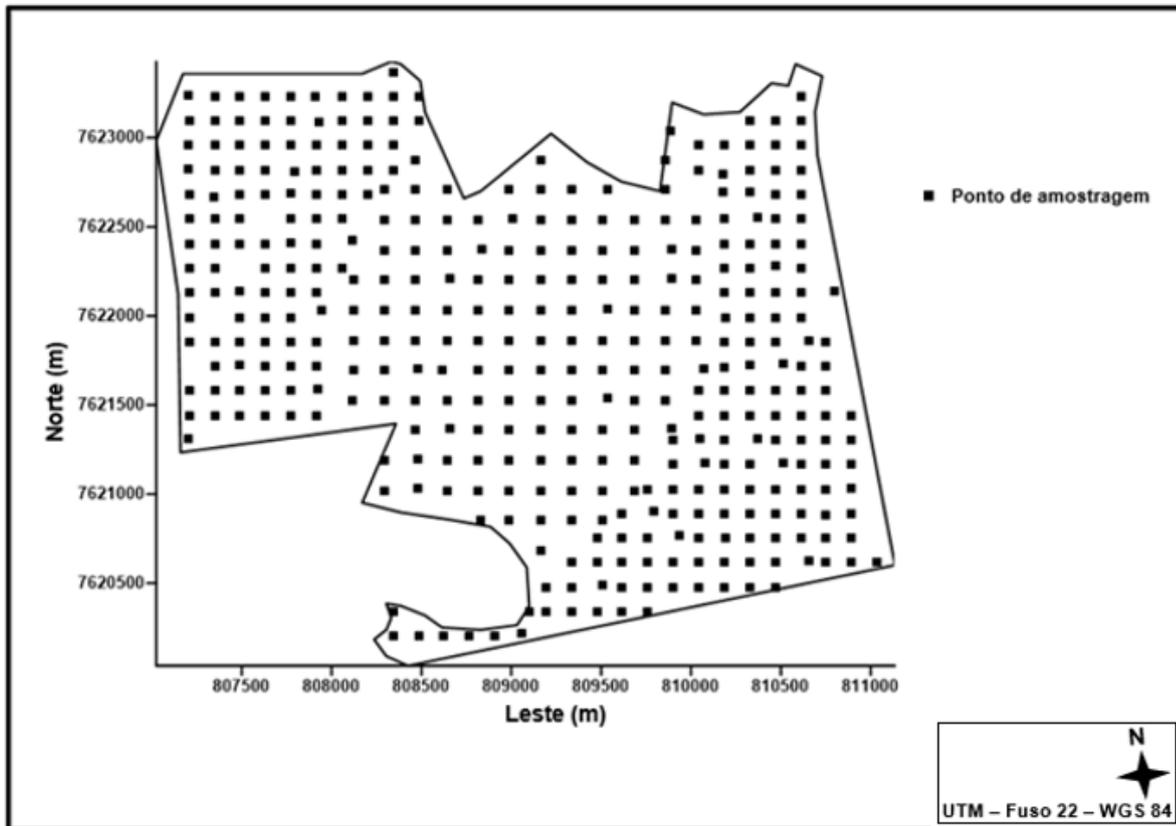


Figura 5. Malha de amostragem de solo na área estudada contendo 371 pontos separados por distâncias mínimas variando de 145 m a 174 m, georreferenciados, cobrindo seus 870 ha.

3.3 Análises laboratoriais

A suscetibilidade magnética (SM) para as 371 amostras de solo foi determinada em 10 g de terra fina seca ao ar utilizando o equipamento Bartington MS2, acoplado ao sensor Bartington MS2B, regulado em baixa frequência (0,47 kHz) conforme proposto por Dearing (1994).

O teor de argila para as 371 amostras de solo foi determinado pelo método da pipeta, utilizando solução de NaOH 0,1 mol L⁻¹ como dispersante químico e agitação mecânica em aparato de baixa rotação pelo tempo preconizado (EMBRAPA, 1997).

3.4 Construção dos cenários de amostragem e conjunto para validação

Para a avaliação da densidade adequada de amostragem para cokrigagem (realização do experimento) foi necessário previamente organizar o banco de dados de ambas as variáveis envolvidas (argila e SM do solo), tanto para a combinação das mesmas (construção dos cenários de amostragem), como para construção do conjunto de dados para finalidade de validação (Figura 6 ilustra esquematicamente o procedimento de organização) para as análises estatísticas e geoestatísticas posteriores. Para essa organização procedeu-se da seguinte forma: partiu-se de cenários de amostragem com densidades teoricamente suficientes para as variáveis envolvidas: argila (variável alvo) e SM (variável auxiliar), ou seja, cada uma ficando com 334 pontos de amostragem (equivale a uma densidade de 1 ponto a cada 2,6 ha, para ambas as variáveis). Este constituiu o primeiro cenário; e utilizou-se para isto 90% dos 371 pontos amostrados (total de dados/pontos que esparsamente cobrem a área). A partir desse primeiro cenário, foi necessária reconstrução de cenários de amostragem decrescentes em relação ao número de pontos apenas da variável alvo, isto é, manteve-se fixa a densidade de amostragem da variável auxiliar, em um 1 ponto para cada 2,6 ha, e reduziu-se a densidade da variável alvo em 290, 218, 174, 145, 124, 109, 97, 87, 79, 73, 67, 62 e 58 pontos, para assim estabelecer cenários que respectivamente apresentam densidades de 1 ponto a cada 3,0; 4,0; 5,0; 6,0; 7,0; 8,0; 9,0; 10,0; 11,0, 12,0, 13,0, 14,0 e 15,0 ha. Com este procedimento organizou-se 14 cenários resultantes das diferentes densidades amostrais da variável alvo e combinação com a densidade amostral da variável auxiliar. Do montante de dados disponíveis, 10% que remaneceram dos dados da variável alvo ficaram reservados constituindo o conjunto de dados para uso no procedimento de validação (37 pontos).

Tanto os pontos que integraram as densidades (diferentes cenários de amostragem organizados) como os que foram utilizados no procedimento de validação (Figura 7) foram aleatoriamente selecionados por procedimento estratificado que assegurou, a cada conjunto, distribuição espacial de pontos com

cobertura equilibrada ao longo da área (item 2.3.1.5.1.1 desta Tese é revisão de literatura sobre esta temática).

Em relação aos diferentes cenários de amostragem, a menor densidade amostral utilizada (1 ponto a cada 15 ha) foi determinada seguindo os princípios da análise geoestatística quanto à necessidade de pelo menos 50 pares de pontos para cada estimativa de semivariância no variograma experimental (Olea, 2006; Chilès e Delfiner, 2012) e, a maior densidade amostral utilizada (1 ponto a cada 2,6 ha) foi determinada com base na cobertura teoricamente ideal (Mcbratney e Webster, 1986; Kerry e Oliver, 2004): amostragem em distâncias inferiores a metade do alcance de dependência espacial para as variáveis envolvidas.

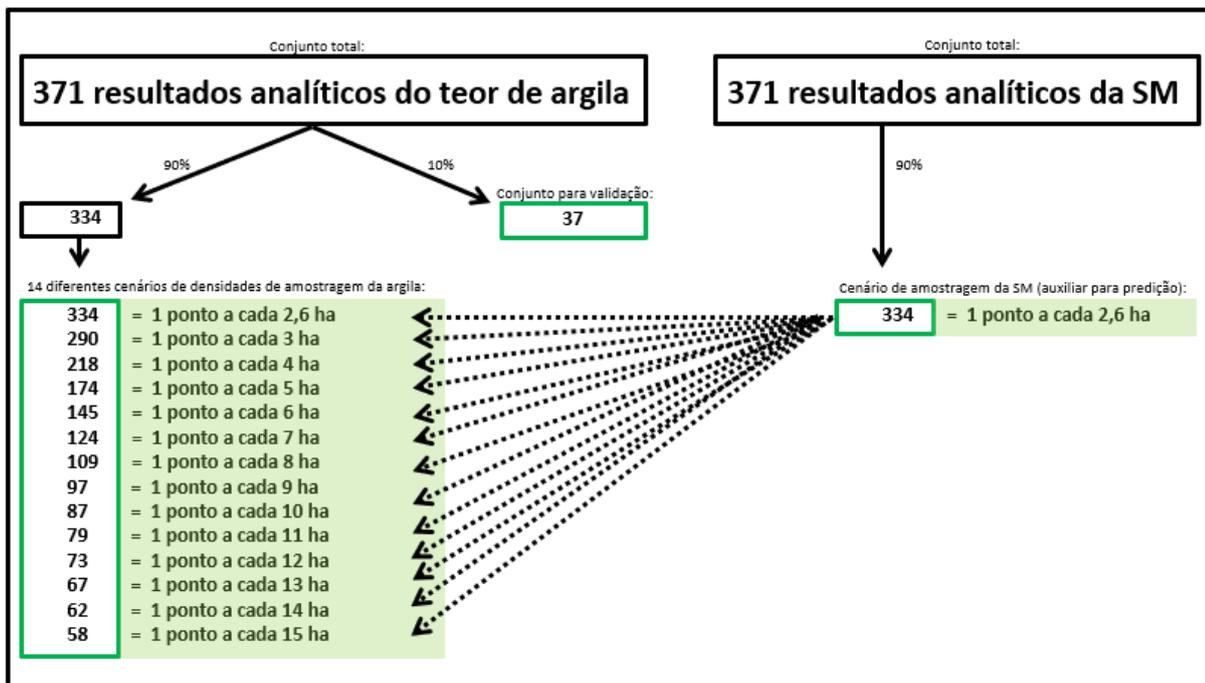


Figura 6. Ilustração esquemática do procedimento de organização dos dados para execução do estudo.

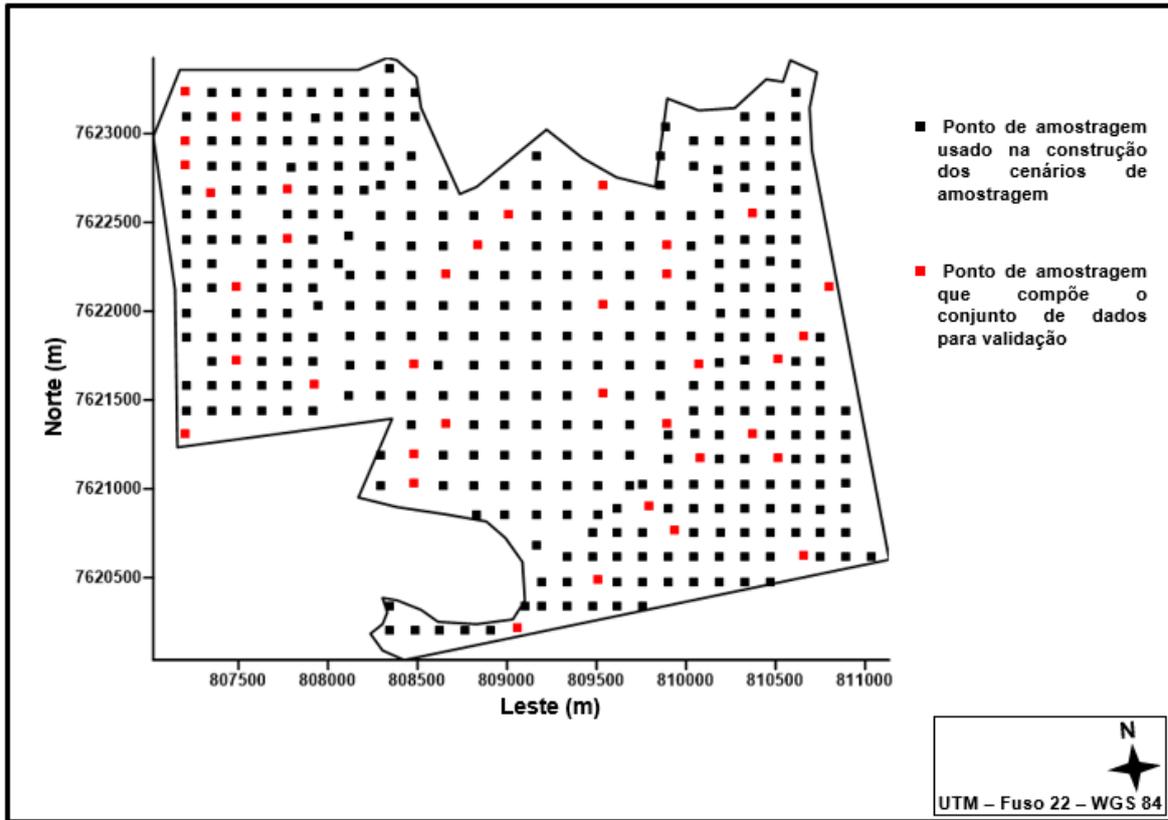


Figura 7. Malha de amostragem de solo na área estudada com a distribuição dos pontos usados na construção dos cenários de amostragem (argila do solo e SM) e conjunto de dados para validação.

3.5 Estatísticas descritivas

As estatísticas descritivas: média, mediana, desvio padrão, coeficiente de variação (CV), valores mínimo e máximo, coeficiente de assimetria e coeficiente de curtose, foram estimadas para avaliar a distribuição de frequência dos atributos antes da análise geoestatística e para verificar a presença de valores atípicos os quais causam as distorções mais graves em Geoestatística (Oliver e Webster, 2014). Estas informações foram estimadas do conjunto de dados da SM (cenário amostral com 1 ponto a cada 2,6 ha) e dos conjuntos de dados da argila (cenários amostrais com 1 ponto a cada 2,6; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14 e 15 ha) para auxiliar nas decisões.

3.6 Análise geoestatística

A dependência espacial foi modelada pelos variogramas e variogramas-cruzados sob a hipótese intrínseca (Oliver e Webster, 2014). Isso foi feito para o conjunto de dados da SM (cenário amostral com 1 ponto a cada 2,6 ha) e para os conjuntos de dados da argila (cenários amostrais com 1 ponto a cada 2,6; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14 e 15 ha). Os variogramas-cruzados foram modelados para cada cenário amostral da argila em combinação com o cenário amostral da suscetibilidade magnética (SM). Os variogramas experimentais (equação 6) foram estimados para descrever a dependência espacial de cada variável dentro de cada cenário amostral. Os variogramas-cruzados (equação 7) foram estimados para descrever a dependência espacial entre as duas variáveis (argila e SM) para cada cenário mencionado. Foi assegurado a positividade da variância das combinações lineares das variáveis conforme preconizado por Rossitter (2007).

Para modelagem de cada variograma foi ajustado diferentes modelos plausíveis (esférico, exponencial e gaussiano) com base na menor soma de quadrado de resíduos (SQR) e no maior coeficiente de determinação (R^2), repetiu-se tais ajustes com realização de possíveis melhoras no encaixe a curta distância, e, posteriormente, dentre estes, foi eleito modelo para cada variograma cujo mesmo apresentou melhor desempenho com base nas estatísticas de validação-cruzada (equação 15), conforme preconizado por Davis (1987), Hengl (2009) e Oliver e Webster (2014). O mesmo se aplicou a cada variograma-cruzado.

Utilizou-se a KO (equação 11) para realizar interpolações com o conjunto de dados da SM (cenário amostral com 1 ponto a cada 2,6 ha) e com o conjunto de dados da argila (cenário amostral com 1 ponto a cada 2,6 ha). Utilizou-se a CKO (equação 13) para realizar interpolações com cada conjunto de dados da argila (cenários amostrais com 1 ponto a cada 2,6; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14 e 15 ha). A cada um dos cenários ora mencionados foi incorporado informações espaciais obtidas do conjunto de dados da SM (cenário amostral com 1 ponto a cada 2,6 ha).

Na avaliação da qualidade das estimativas de CKO em cada cenário de densidade amostral supramencionados, usou-se o procedimento de validação-

externa (equação 16) para obtenção de relatórios de erros de predições. Com base nos relatórios de cada cenário calculou-se o índice de RMSE (raíz do erro quadrático médio) (equação 17) (Hengl, 2009) para cada um desses.

$$RMSE = \left\{ \frac{1}{n} \sum_{i=1}^n [z_{obs}(x_i) - Z_{est}(x_i)]^2 \right\}^{0,5} \quad (17)$$

em que, n é o número de valores utilizados na validação ($n = 37$); $Z_{est}(x_i)$ é o valor estimado da propriedade no i -ésimo ponto; e $z_{obs}(x_i)$ é o valor observado da propriedade no i -ésimo ponto. O índice RMSE fornece uma medida da magnitude dos erros (medidas de acurácia) e, quanto mais próximo de 0 (zero) este índice estiver, mais acurada é a predição. Como o RMSE é sensível a erros tanto sistemáticos quanto aleatórios (Odeh et al., 2007), tem sido uma estatística de resumo muito utilizada para avaliar a acurácia de mapas de propriedade(s) do solo.

Com base nos índices de RMSE de cada cenário, a melhoria relativa (MR_{RMSE}) na acurácia proporcionada pela incorporação de informações secundárias no procedimento de interpolação foi avaliada adotando a seguinte equação:

$$MR_{RMSE} = \frac{RMSE_{KO(referência)} - RMSE_{CKO}}{RMSE_{KO(referência)}} \times 100 \quad (18)$$

em que, $RMSE_{KO(referência)}$ é o valor de RMSE obtido da interpolação com o conjunto de dados da argila cenário amostral com 1 ponto a cada 2,6 ha (mapa de referência); e o $RMSE_{CKO}$ é o valor de RMSE obtido da interpolação por CKO, para os respectivos cenários de amostragem estudados.

As análises variográficas e os cálculos de interpolações foram realizados com o software GS+ versão 9.0 (GAMMA SOFTWARE DESIGN, 2008). Os mapas foram produzidos com o software Surfer versão 9.0 (GOLDEN SOFTWARE, 2009). Os gráficos de quantil-quantil foram gerados com o software R (R DEVELOPMENT CORE TEAM, 2010).

4 RESULTADOS E DISCUSSÃO

4.1 Estatísticas descritivas

A análise estatística das variáveis estudadas (Tabela 1) mostraram que os valores médios (31,90 à 34,80%) e medianos (27,80 à 35,90%) do teor de argila em todos os cenários foram próximos, indicando que as distribuições são aproximadamente simétricas, confirmado pelos respectivos valores de coeficiente de assimetria (-0,08 à 0,27) próximos a 0 para a propriedade argila. Entretanto, a SM, apresentou maior grau de assimetria (0,65). Embora a interpolação geoestatística com dados não gaussianos seja apropriada (Cressie, 1993), os valores preditos a partir deles, podem não ser as melhores estimativas lineares imparciais. Apesar da SM ter sido utilizada como covariável para auxiliar na predição da argila, foi analisado todas as características necessárias (dados com transformação, dados sem transformação, etc.) para seu melhor uso.

Tabela 1. Estatística descritiva dos dados de suscetibilidade magnética (SM) ($\times 10^{-8}$ m³ kg⁻¹) e argila do solo (ARG) (%) para os cenários estudados.

Cenários	Média	Md	DP	CV	Mín.	Máx.	Ass	Curt	r
SM(2,6)	2281,20	1384,00	2035,00	89,36	14,00	7458,00	0,65	-0,84	
ARG(2,6)	33,20	29,70	17,01	53,45	5,50	66,50	0,17	-1,25	0,89
ARG(3)	33,00	29,47	17,12	60,18	5,50	66,50	0,18	-1,24	0,90
ARG(4)	32,80	28,58	17,15	63,74	5,50	66,50	0,21	-1,25	0,89
ARG(5)	31,90	27,80	16,65	62,35	5,50	63,70	0,27	-1,20	0,90
ARG(6)	32,30	27,88	16,94	59,87	5,50	63,70	0,25	-1,23	0,90
ARG(7)	32,20	28,60	16,48	57,10	5,50	63,70	0,23	-1,16	0,89
ARG(8)	32,50	28,28	16,51	54,62	5,70	63,70	0,25	-1,15	0,89
ARG(9)	32,80	29,20	16,65	52,40	5,70	63,70	0,19	-1,19	0,89
ARG(10)	33,10	29,40	16,53	50,07	5,70	63,70	0,15	-1,18	0,89
ARG(11)	32,80	29,60	16,51	48,54	5,70	63,70	0,13	-1,17	0,88
ARG(12)	33,80	31,40	16,90	46,68	5,70	63,70	0,00	-1,22	0,88
ARG(13)	34,30	34,00	16,82	45,10	5,70	63,70	-0,05	-1,19	0,88
ARG(14)	34,80	35,90	16,90	43,53	5,70	63,70	-0,08	-1,19	0,88
ARG (15)	34,70	35,90	16,64	42,17	5,70	63,70	-0,08	-1,18	0,88

Md – Mediana; DP – Desvio Padrão; CV – Coeficiente de Variação; Mín. – Valor mínimo; Máx. – Valor máximo; Ass – Coeficiente de Assimetria; Curt – Coeficiente de Curtose; r – Correlação de Pearson (entre SM e cada conjunto de dados da ARG nas diferentes densidades amostrais).

Observação: O cenário amostral da SM corresponde a densidade de 1 ponto a cada 2,6 ha (valor comunicado entre parênteses); e, os cenários amostrais da ARG correspondem as densidades de 1 ponto à cada 2,6; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14 e 15 ha (valores comunicados entre parênteses) os quais constituem os cenários amostrais de estudo.

Em relação à dispersão dos valores dos dados no entorno de suas médias, os coeficientes de variação (CV) obtidos, de acordo com Warrick e Nielsen (1980), são considerados alto para SM (CV de 89.36%) e médio à alto para argila (CV variando de 42,17 (cenário ARG(15)) à 63,74% (cenário ARG(4))). Outras pesquisas nesta mesma região do Estado também reportam variações com classificação semelhantes para estes atributos (Peluco et al., (2015).

A relação linear (r) entre os dados de SM e os dados de argila nos seus diferentes cenários acusou forte associação positiva entre as duas propriedades, variando de 0,88 à 0,90 e, portanto, estimativas a respeito do teor de argila do solo com o auxílio de informações da SM atende a pressuposição de correlação entre as variáveis para a aplicação da cokrigagem. Embora a correlação estatística não implique correlação espacial, ela sinaliza possibilidade ou não de êxito na interpolação auxiliadas por informações secundárias (Song et al., 2014). Resultados semelhantes aos apresentados neste trabalho foram documentados por outros pesquisadores (Siqueira et al., 2010; Peluco et al., 2013; Marques Júnior et al., 2014; Teixeira, 2015; Siqueira et al., 2016; Teixeira et al., 2017; Bahia et al., 2017). Sobre isso Resende et al., (1988), Fontes et al. (2000) e outros pesquisadores, explicam que por trás da forte correlação entre elevada expressão magnética e maior concentração de argila observada em uma área (assim como o inverso: baixa expressão magnética e menor concentração de argila), estão os tipos de materiais geológicos os quais deram origem a tais solos. Estes pesquisadores também destacam que solos com maior fração argila favorecem a formação de minerais de comportamento magnético mais evidente (maghemita), o que não acontece em solos com baixa concentração de argila.

4.2 Análise variográfica

O modelo teórico que melhor se ajustou aos variogramas experimentais, para ambos os atributos analisados, foi o esférico (Tabela 2), o qual é o mais frequentemente utilizado em Geoestatística (Webster e Oliver, 2007) para estudar o solo (Cambardella et al., 1994). Todos os modelos ajustados exibiram um pequeno efeito pepita indicando que a grade amostral, nos diferentes cenários de densidade

amostral, se manteve apta a capturar a estrutura de dependência espacial de variação da SM e argila do solo.

Os valores de alcance da SM (2080 m) e da argila (2083 à 2395 m) apresentam-se coerentes com valores de alcances levantados em pesquisas de solo para região (Teixeira et al., 2017). Conforme é notado, há uma semelhança expressiva para os alcances das variáveis estudadas, o que torna possível utilizar a SM como variável auxiliar no mapeamento da argila explorando correlações espaciais entre ambas (Hanesch e Scholger, 2005; Siqueira et al., 2010, 2014). Na realidade, em parte, esta semelhança em termos de alcances confirmam a correionalização que é um dos requisitos exigidos no procedimento de interpolação que utiliza de informações secundárias (Rossiter, 2012). De acordo com McBratney e Webster (1983a) a distribuição espacial de qualquer propriedade do solo muitas vezes está estreitamente relacionada com a de outras propriedades afetadas pelo mesmo fenômeno regionalizado (dado processo espacial) e, sob esta situação, tais propriedades são referidas como sendo correionalizadas, o que implica que em termos estatísticos elas guardam dependência espacial aproximadamente semelhantes.

Em relação aos resultados da validação-cruzada apresentados na Tabela 2 os mesmos mostram que no geral, de certo modo, a modelagem dos variogramas e variogramas-cruzados foi bem sucedida, tendo em vista que os coeficientes de regressão entre os valores estimados e os observados são muito próximos de 1 (um) e os coeficientes lineares em sua maioria se afastam muito pouco de 0 (zero); condições estas consideradas como ideais (Vieira et al., 2010) e, que são confirmadas pelos próprios aspectos apresentados pelos variogramas e variogramas-cruzados (Figura 8). Conforme pode ser notado nestes, no geral, os modelos mostram-se bem ajustados ao conjunto de semivariâncias (nuvem de pontos de valores experimentais) que compõe cada variograma e variograma-cruzado, o que indica qualitativamente o quão bem os modelos correspondem aos dados estudados (Oliver e Webster, 2014).

Tabela 2. Variogramas experimentais estimados da suscetibilidade magnética (SM) e argila (ARG) do solo e variogramas-cruzados destas para cada cenário amostral estudado*.

Variograma	Cenário amostral	Mod	C ₀	C ₀ +C ₁	A (m)	SQR	R ²	Parâmetro da reta da V.C.	
								a	b
Experimental	SM(2,6)	Esf.	0,0001	0,1	2080	4,6E-	0,9	-0,95	1,00
	ARG(2,6)	Esf.	12,0	315,7	2085	8924,0	0,9	-0,15	1,00
Cruzado	SM(2,6)xARG(2,6)	Esf.	0	4,7	2085	2,45	0,9	-0,16	1,03
	SM(2,6)xARG(3)	Esf.	0,01	4,7	2083	3,00	0,9	-0,04	1,03
	SM(2,6)xARG(4)	Esf.	0	4,5	2085	2,58	0,9	-0,21	1,03
	SM(2,6)xARG(5)	Esf.	0,0	4,3	2087	2,11	0,9	-0,95	1,02
	SM(2,6)xARG(6)	Esf.	0	4,6	2098	3,25	0,9	-0,43	1,04
	SM(2,6)xARG(7)	Esf.	0	4,5	2102	2,00	0,9	-0,10	1,03
	SM(2,6)xARG(8)	Esf.	0	4,5	2109	2,02	0,9	-0,02	1,03
	SM(2,6)xARG(9)	Esf.	0,01	4,7	2114	3,02	0,9	-0,63	1,01
	SM(2,6)xARG(10)	Esf.	0	4,5	2118	3,69	0,9	-0,39	1,01
	SM(2,6)xARG(11)	Esf.	0	4,5	2118	3,95	0,9	-0,37	1,01
	SM(2,6)xARG(12)	Esf.	0	4,9	2206	4,49	0,9	-0,37	1,01
	SM(2,6)xARG(13)	Esf.	0	4,4	2139	4,13	0,9	-0,38	1,01
	SM(2,6)xARG(14)	Esf.	0	4,9	2385	3,61	0,9	-0,47	1,01
	SM(2,6)xARG(15)	Esf.	0	4,7	2395	3,05	0,9	-0,37	1,01

Mod = Modelo; Esf. = Esférico; C₀ = Efeito Pepita; C₀+C₁ = Patamar; A (m) = Alcance; SQR = Soma de Quadrado de Resíduos; R² = Coeficiente de Determinação; V.C. = Validação Cruzada; a = Coeficiente de Linear; b = Coeficiente Regressão; * O cenário amostral da SM corresponde a densidade de 1 ponto a cada 2,6 ha (valor comunicado entre parênteses); e, os cenários amostrais da ARG correspondem as densidades de 1 ponto à cada 2,6; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14 e 15 ha (valores comunicados entre parênteses) os quais constituem os cenários amostrais de estudo que, em cada um deles, incorporou-se informação da SM amostrada em 1 ponto a cada 2,6 ha.

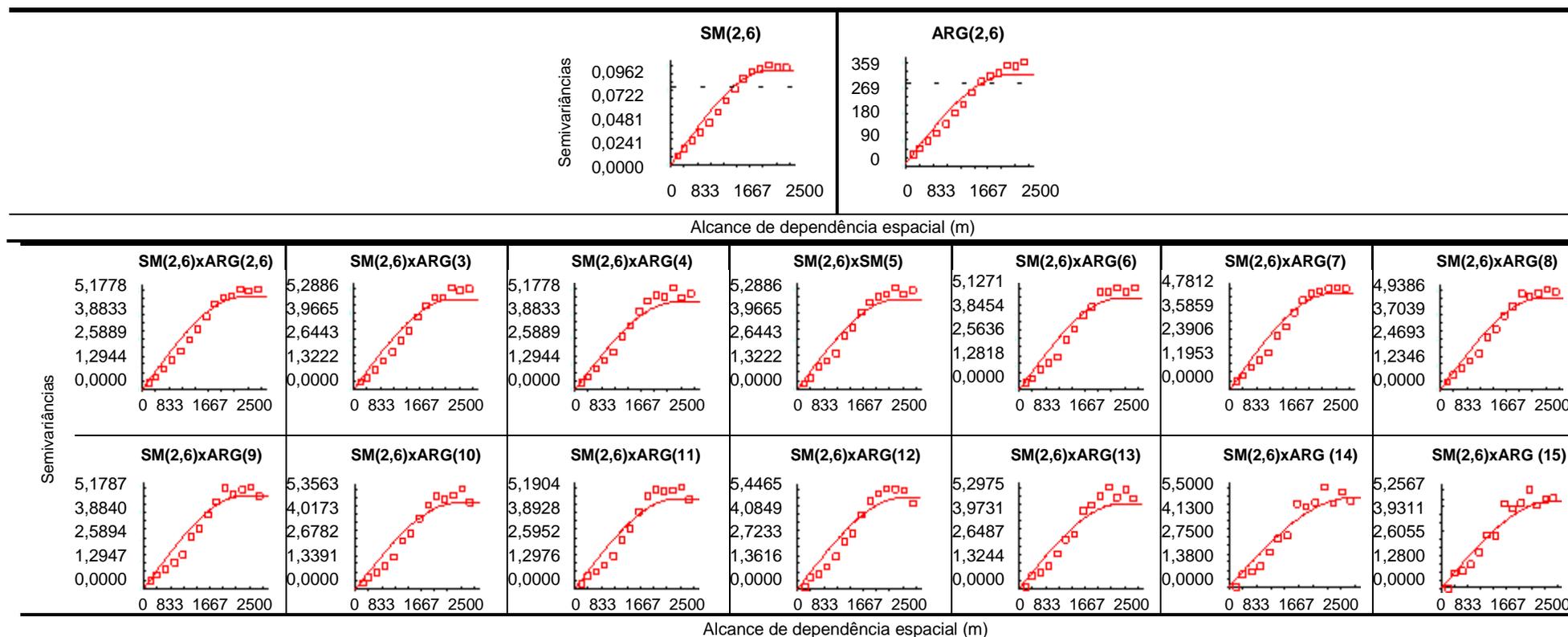


Figura 8. Variogramas experimentais estimados da suscetibilidade magnética (SM) e da argila (ARG) do solo e variogramas-cruzados destas para cada cenário amostral estudado.

Observação: O cenário amostral da SM corresponde a densidade de 1 ponto a cada 2,6 ha (valor comunicado entre parênteses); e, os cenários amostrais da ARG correspondem as densidades de 1 ponto à cada 2,6; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14 e 15 ha (valores comunicados entre parênteses) os quais constituem os cenários amostrais de estudo que, em cada um deles, incorporou-se informação da SM amostrada em 1 ponto a cada 2,6 ha.

4.3 Interpolação (mapas)

Na Figura 9 são apresentados todos os mapas de KO, da SM e da argila do solo (ambas amostradas em 1 ponto a cada 2,6 ha), e os mapas de CKO da argila amostrada nas densidades estudadas, ou seja, amostrada em 1 ponto à cada 2,6; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14 e 15 ha os quais constituem os cenários amostrais de estudo que, em cada um deles, incorporou-se informação da SM amostrada em 1 ponto a cada 2,6 ha. O que pode ser notado de uma verificação conjunta desses mapas contidos na Figura 9, tanto nos de KO da SM e da argila como nos de CKO de argila amostrada nas diferentes densidades estudadas, é que em termos de características gerais os mesmos apresentam padrões de variabilidade espacial muito semelhante. Esta constatação traz à tona a informação de que a distribuição espacial da argila do solo está estreitamente relacionada com a distribuição espacial da SM na área estudada, confirmando as análises de correlação e de variografia e a hipótese de que tais propriedades são originadas ou fortemente influenciadas por fatores de formação do solo que operam e interage dentro de escala geográfica semelhante o que coloca as referidas propriedades do solo como variáveis corregionalizadas (McBratney e Webster; 1983a).

Normalmente a forte correlação espacial entre elevada expressão magnética e maior concentração de teor de argila observada em uma área (assim como o inverso: baixa expressão magnética e menor concentração de argila) tem sido apontada como uma resultante das características mineralógicas do material geológico (presença de magnetita) as quais conferem características magnéticas aos solos. Porém, sobre isso, outras condições subjacentes podem estar relacionadas a tal correlação. Segundo Fontes et al. (2000) solos que cuja composição granulométrica predominam a fração argila, favorecem a formação de minerais de comportamento magnético expressivo (maghemita) podendo conferir aos mesmos maior SM. Ainda em relação a possibilidade conferida por solos de composição mais argilosa, está a maior capacidade destes de acumular matéria orgânica. De acordo com Hanesch et al. (2006) a presença de matéria orgânica no solo também favorece o aumento da SM do mesmo, uma vez que proporciona

condições favoráveis para redução do ferro, o que dificulta a transformação de magnetita em hematita, tendo esta menor expressão magnética nos solos (Santana et al., 2001). Além disso Hanesch et al. (2006) afirmaram que a matéria orgânica favorece a transformação de Fe pedogenético e proporciona ambiente para a neoformação de minerais com característica magnética (maghemita).

Fazendo um resumo do exposto das diferentes literaturas, pode-se considerar que a SM do solo é pedologicamente covariativa do material de origem, da textura e da matéria orgânica do mesmo. Todavia, é importante destacar que áreas que se encontram situadas em zonas que constituem transição entre materiais de origem, como a de estudo deste trabalho, as inferências acima apresentadas podem ser difíceis de serem confirmadas, pois, nesse caso existem grandes chances de em certas partes da área possuir solos (ou parte destes, por exemplo, camadas superiores) aloctônes (que não são originados(as) da própria rocha-matriz, mas sim do transporte de sedimentos de outras rochas) ou mistura de atributos de solos próximos, porém provenientes de rochas diferentes. Por este motivo é necessário se ter cautela com as inferências sobre fatores causais, mesmo baseadas em correlações, para que as mesmas não se constituam meras especulações.

Com atenção ao que foi mencionado, baseando-se em confrontação das informações espaciais contidas nos mapas da Figura 9 com as documentadas no mapa B da Figura 4, é possível visualizar que as regiões que apresentaram valores mais altos de SM (acima de $5000 \times 10^{-8} \text{ m}^3 \text{ kg}^{-1}$) (parte superior direita dos mapas da Figura 9) estão localizadas na região de material de origem derivado de rocha basáltica mostrada no mapa B da Figura 4 (parte superior direita do mapa). No entanto, nota-se que em algumas partes da área que estão localizadas na região que se localiza rochas sedimentares, conforme mostrado no mapa B da Figura 4, também apresentaram valores de SM do solo (lado esquerdo no centro em qualquer dos mapas da Figura 9) que são considerados altos (acima de $3000 \times 10^{-8} \text{ m}^3 \text{ kg}^{-1}$) em relação aos reportados por outros estudiosos (Preetz et al., 2008; Camargo et al., 2014). Além da possibilidade dos solos localizados em áreas situadas em zonas que constituem transição entre materiais de origem estarem sujeitos a se desenvolverem de material de origem que são, na realidade, misturas de materiais de origens distintas, há a possibilidade de certas características destes solos serem

resultantes de ações antropogênicas, como por exemplo, de sucessivas queimas da cobertura vegetal. Por exemplo, em relação aos altos valores que se mencionou para SM, estes podem estar relacionados aos ciclos de safras que anteriormente ocorriam na área nos quais os procedimentos de colheitas se davam após o submetimento da cultura (palhada e folhas da cana de açúcar) à queima porque as colheitas eram realizadas de forma manual. Segundo Schwertmann (1985), a presença do fogo pode promover a transformação de óxidos de ferro pedogenéticos e a subsequente neoformação da maghemita, que possui alto potencial magnético. Logo, tal explicação pode ser plausível porque segundo estimativas da temperatura da camada superficial do solo durante o processo de queima da cana esta pode variar entre 160 a 200 °C (Ripoli e Ripoli, 2004) e a formação de maghemita pode ocorrer em temperaturas abaixo de 250 °C (Liu et al., 2010).

Voltando a atenção ao que é proposto neste estudo em relação à avaliação de cenários de amostragem para o mapeamento da argila do solo por CKO com a incorporação de informações espaciais da SM do solo, outras verificações nos mapas apresentados na Figura 9 são importantes. Para isto é necessário uma apreciação visual de cada um dos mapas gerados por CKO e uma confrontação destes com o mapa de argila (mapa ARG(2,6)) produzido com o conjunto de dados em número de pontos de amostragem (densidade de amostragem) considerado teoricamente suficiente (argila amostrada em 1 ponto a cada 2,6 ha). Esta verificação não visa aquilatar sobre qualidade dos mapas mesmo porque se sabe que a capacidade humana, mesmo para uma simples comparação visual de mapas, é muito limitada (Berry, 2002). Entretanto, MacEachren (1995) e Yang e Hodler (2000) ressaltam que este tipo de verificação ao ser feita com bastante atenção, permite encontrar eventuais distorções, detalhes acrescentados ou faltantes nos mapas, etc. Ou seja, estes aspectos fornecem uma ideia sobre coerências e incoerências contidas nos mapas, o que Yang e Hodler (2000) chamam de fidelidade visual (o quão próxima ou semelhante em termos visual é a imagem de uma superfície interpolada em relação a superfície real ou de referência).

Considerando o exposto, observa-se que as informações espaciais contidas nos mapas de CKO da argila, provenientes dos cenários que a mesma foi amostrada em 1 ponto a cada 2,6; 3; 4; 5; 6 e 7 ha, são muito parecidas com as informações

espaciais apresentadas no mapa de referência (mapa ARG(2,6)). Entretanto, o mesmo não se pode dizer ao observar os mapas de CKO da argila cujos mesmos a mesma foi amostrada em 1 ponto a cada 8; 9; 10; 11; 12; 13; 14 e 15 ha. Nota-se das informações espaciais contidas nestes mapas (gerados com menores densidades de amostragens), certa discrepância visual em relação aos detalhes espaciais contidos no mapa de referência. Dentre outros aspectos visuais, observa-se nesses mapas que as isolinhas apresentam-se com angulosidades. É possível observar ainda que entre estes (mapas produzidos de cenários menos densamente amostrados) que o que possui a maior cobertura de pontos (cenário SM(2,6)×ARG(8)) é o que as angulosidades apresentadas nas isolinhas são as mais suaves (quase não perceptíveis), enquanto o de menor cobertura amostral (cenário SM(2,6)×ARG(15)), é o que as angulosidades nas isolinhas apresentam-se mais acentuadas (são claramente evidentes). Sobre esta questão é necessário destacar que em decorrência de certas condições subjacentes aos procedimentos geoestatísticos, a interpolação pode gerar imagens enganosas da variabilidade espacial (Oliver e Webster, 2014). De acordo com alguns estudiosos a alta densidade de informação da variável secundária em relação a variável alvo pode resultar em instabilidade no sistema de equações lineares da cokrigagem ordinária (Isaaks e Srivastava, 1989; Goovaerts, 1997; Wackernagel, 2003).

Estas constatações relatadas são provas de que o procedimento de validação se faz necessário na Geoestatística quando se visa uma efetiva avaliação da qualidade dos mapas (Olea, 1999; Oliver, 2010; Brus et al., 2011; Heuvelink, 2018), mesmo quando utilizando método de interpolação que permite incorporação de informações correlacionadas para suplementação amostral (Rossiter, 2012), como se utilizou neste trabalho.

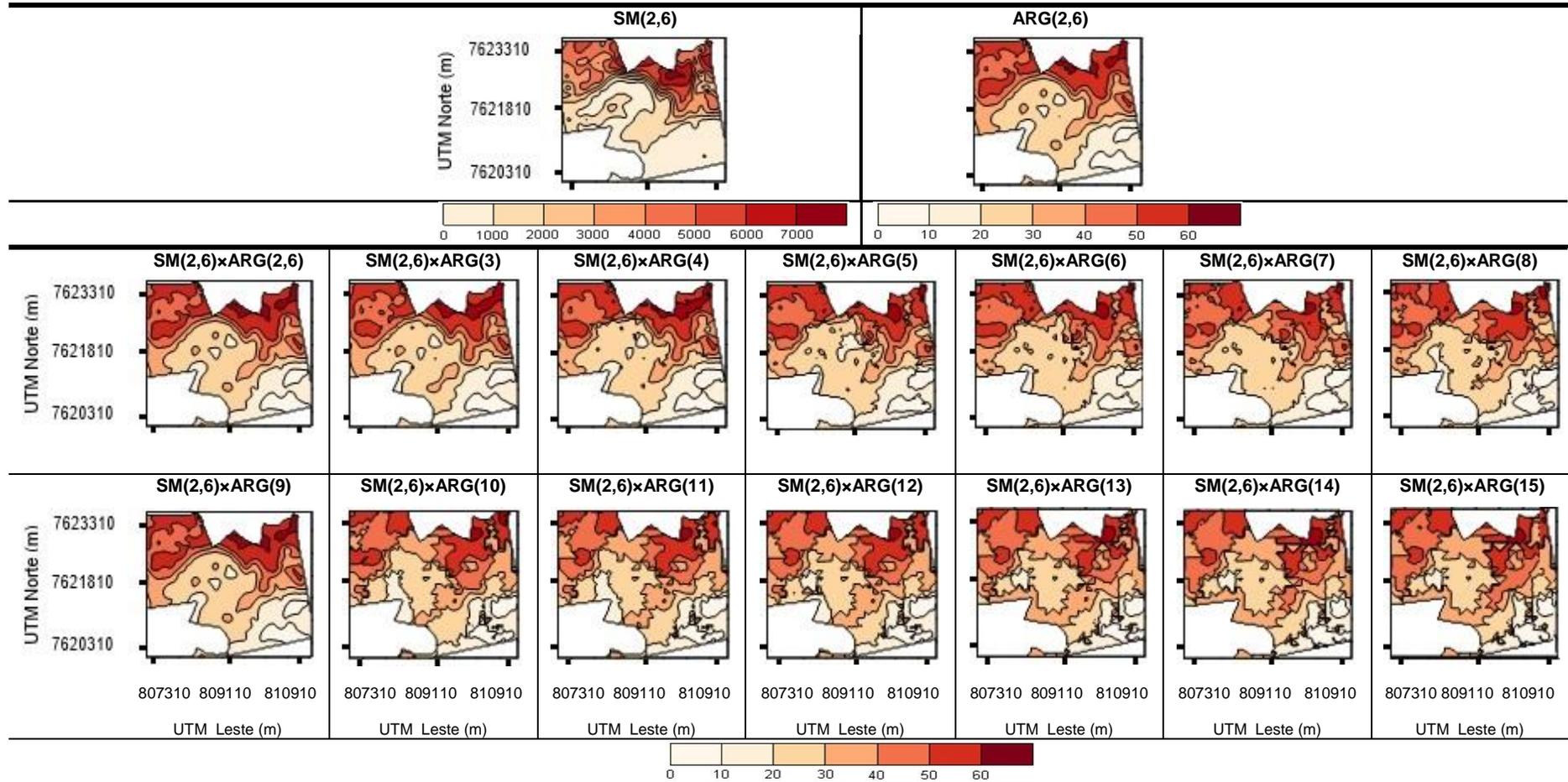


Figura 9. Mapas de krigagem da suscetibilidade magnética (SM) ($\times 10^{-8} \text{ m}^3 \text{ kg}^{-1}$) e da argila (ARG) (%) do solo e mapas de cokrigagem desta para cada cenário amostral estudado.

Observação: O cenário amostral da SM corresponde a densidade de 1 ponto a cada 2,6 ha (valor comunicado entre parênteses); e, os cenários amostrais da ARG correspondem as densidades de 1 ponto à cada 2,6; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14 e 15 ha (valores comunicados entre parênteses) os quais constituem os cenários amostrais de estudo que, em cada um deles, incorporou-se informação da SM amostrada em 1 ponto a cada 2,6 ha.

4.3.1 Validação dos mapas (acurácia)

Quando não se tem dados amostrais suficientes de uma variável que se necessita krigar (mapear) é muito comum na literatura a ideia de que a cokrigagem se constitui como solução quando se pode contar com dados (informações) de uma variável espacialmente correlacionada que foi densamente amostrada. De fato a cokrigagem pode ser a solução, mas, a falha na literatura é não mencionar que o bom desempenho da cokrigagem depende de uma adequada combinação amostral para ambas as variáveis. No presente trabalho as informações apresentadas na Tabela 3 mostram exatamente isto para o mapeamento da argila do solo usando a suplementação amostral da SM. Nesta tabela o índice de acurácia (RMSE) obtido com a solução da equação 17 é apresentado para o mapa de argila gerado a partir de uma amostragem com densidade teoricamente ideal para argila (mapa ARG(2,6)), isto é, sem a necessidade de suplementação amostral (incorporar informação espacial) de uma outra variável. Nesta tabela também é apresentado os índices de acurácia (RMSE) obtidos com a solução da equação 17 para cada um dos mapas de argila gerados com densidades amostrais relativamente menores para a argila mas que receberam suplementação amostral da SM com cobertura amostral suficiente (mapas SM(2,6)×ARG(3) à SM(2,6)×ARG(15)). Por fim, na Tabela 3 está contido os índices de melhoria relativa (MR_{RMSE}) calculados a partir da solução da equação 18 para cada mapa de cokrigagem, os quais expressam percentualmente o quanto cada mapa apresenta de ganho ou de perda de acurácia em relação (comparativamente) ao mapa de argila gerado a partir da amostragem com densidade teoricamente ideal para argila.

Observando o que proporcionou a CKO pela Tabela 3, nota-se que o cenário com o maior número de dados, SM(2,6)×ARG(2,6), não foi o que apresentou melhores resultados de acurácia. Ainda assim houve certo ganho em acurácia (0,91%) em relação a ARG(2,6) (mapa de referência). Isto já era esperado porque a densidade de amostragem da argila utilizada na elaboração deste mapa, corresponde a um cenário de amostragem teoricamente suficiente para KO.

Tabela 3. Índices de acurácia (RMSE) e de melhoria relativa (MR_{RMSE}) calculados a partir da validação externa* para avaliação dos mapas de cokrigagem da argila (%) do solo para cada cenário amostral de estudo.

Mapas		RMSE	MR_{RMSE}
Referência	ARG(2,6)	6.57	
Cenários amostrais de estudo**	SM(2,6)×ARG(2,6)	6.51	0,91
	SM(2,6)×ARG(3)	6.29	4,26
	SM(2,6)×ARG(4)	6.16	6,24
	SM(2,6)×ARG(5)	6.14	6,54
	SM(2,6)×ARG(6)	6.09	7,30
	SM(2,6)×ARG(7)	6.08	7,46
	SM(2,6)×ARG(8)	7.88	-19,94
	SM(2,6)×ARG(9)	8.85	-34,70
	SM(2,6)×ARG(10)	9.70	-47,64
	SM(2,6)×ARG(11)	9.89	-50,53
	SM(2,6)×ARG(12)	9.94	-51,29
	SM(2,6)×ARG(13)	10.83	-64,84
	SM(2,6)×ARG(14)	10.94	-66,51
	SM(2,6)×ARG(15)	10.97	-67,00

RMSE = Raiz quadrada do quadrado médio do erro; MR_{RMSE} = índice de Melhoria Relativa ou Melhora Relativa (%); * Conjunto de dados com 37 pontos; ** Estes correspondem as densidades de amostragem da argila em 1 ponto à cada 2,6; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14 e 15 ha (valores comunicados entre parênteses) que, em cada um deles, incorporou-se informação da SM amostrada em 1 ponto a cada 2,6 ha.

Na realidade, como já foi mencionado, a CKO é uma adaptação do método de KO que Matheron desenvolveu para produzir estimativas (mapa) com boa acurácia de uma variável que foi menos densamente amostrada com base em sua dependência espacial e na dependência espacial de uma ou mais covariável que foram mais intensivamente amostradas (Matheron, 1971; Journel e Huijbregts, 1978; Myers, 1982). Vieira (2000) afirma que em um cenário de amostragem já suficiente é supérfluo ou desnecessário utilizar a CKO, porque haverá o que se pode chamar de “redundância” nas observações. E, ressalta que, a vantagem do uso da CKO (ganhos significantes em acurácia) só existirá apenas quando a variável a ser estimada apresentar uma menor densidade de amostragem do que a variável associada (Vieira, 2000). Todavia, como poderá ser visto, isto só ocorre dentro de certo limite.

Os cenários SM(2,6)×ARG(3), SM(2,6)×ARG(4), SM(2,6)×ARG(5), SM(2,6)×ARG(6) e SM(2,6)×ARG(7) confirmam exatamente isso. Pois, estes foram os que permitiram ganhos com o método de CKO. Estes cenários de amostragem

permitiram que a CKO expressasse seus dois principais benefícios, reduzir a densidade amostral da variável alvo em 13, 35, 48, 57 e 63% respectivamente; e, aumentar a acurácia de predição em respectivamente 4,26; 6,24; 6,54; 7,30 e 7,46% (Tabela 3).

Além da escassez de trabalhos sobre planejamento amostral para CKO (Kerry et al., 2010) nota-se que a literatura técnica é falha ao não informar que a cokrigagem necessita de uma combinação amostral adequada para que o método apresente um bom desempenho. Entre os estudos desenvolvidos para o aperfeiçoamento da amostragem para CKO, nota-se na literatura que os mesmos trabalharam com um número relativamente limitado de cenários de densidades de amostragem, normalmente não passaram de cinco, que, talvez seja insuficiente para levantar esta informação. No presente estudo, foi levantado esta informação: constatou-se que do mapa $SM(2,6) \times ARG(7)$ para $SM(2,6) \times ARG(8)$ passou-se de ganhos em acurácia para perdas em relação a acurácia, na ordem de 27,40% ao comparar os dois cenários. Isto significa que a incorporação de informação auxiliar apresenta uma combinação de densidades de amostragem que possui um limiar entre bom desempenho e ruim. De acordo com Isaaks e Srivastava (1989) e Goovaerts (1997) os valores desconhecidos dos dados primários, nos locais a serem estimados, são preditos como uma combinação linear ponderada de valores primários próximos de cada ponto estimado juntamente com a combinação linear ponderada de valores secundários próximos. Portanto, o valor médio dos dados secundários e dos dados primários deve ser igual na vizinhança de procura da interpolação (Isaaks e Srivastava, 1989; Goovaerts, 1997; Wackernagel, 2003) o que pode não ocorrer quando a cobertura de dados das variáveis são numericamente contrastantes. Isto impõe, para uma aplicação técnica da CKO, a necessidade de informações prévias para o ajuste da densidade de amostragem para que a CKO garanta benefícios, e não o contrário. A Figura 10 reúne os principais mapas apresentados na Figura 9, porém com maior tamanho, o que permite uma melhor visão do efeito da interpolação ora discutido e anteriormente (item 4.3).

Em relação aos resultados de validação registrados neste trabalho, é importante mencionar que os índices RMSE das predições de argila por CKO

relatados, em média, são consistentes com valores de RMSE documentados por outros pesquisadores (Liao et al., 2013).

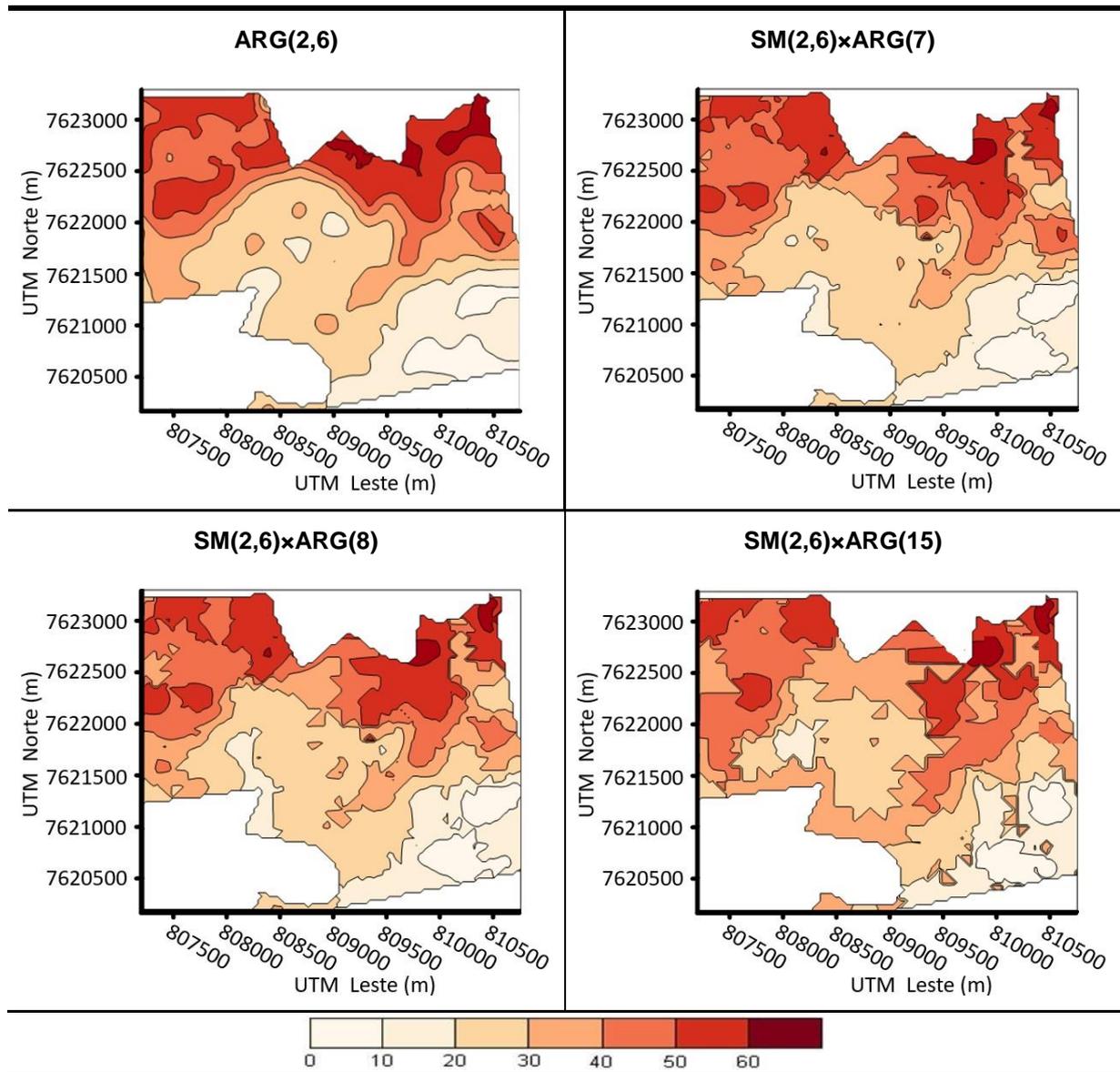


Figura 10. Mapas de argila do solo (%) da área de estudo: mapa de referência ARG(2,6); mapa com maior ganho relativo de acurácia SM(2,6)×ARG(7); mapa com perda relativa de acurácia SM(2,6)×ARG(8) e mapa com maior perda relativa de acurácia SM(2,6)×ARG(15).

Observação: Os números comunicados entre parênteses correspondem as densidades de amostragem (quantidade de hectares por ponto de amostragem) da variável para o mapeamento.

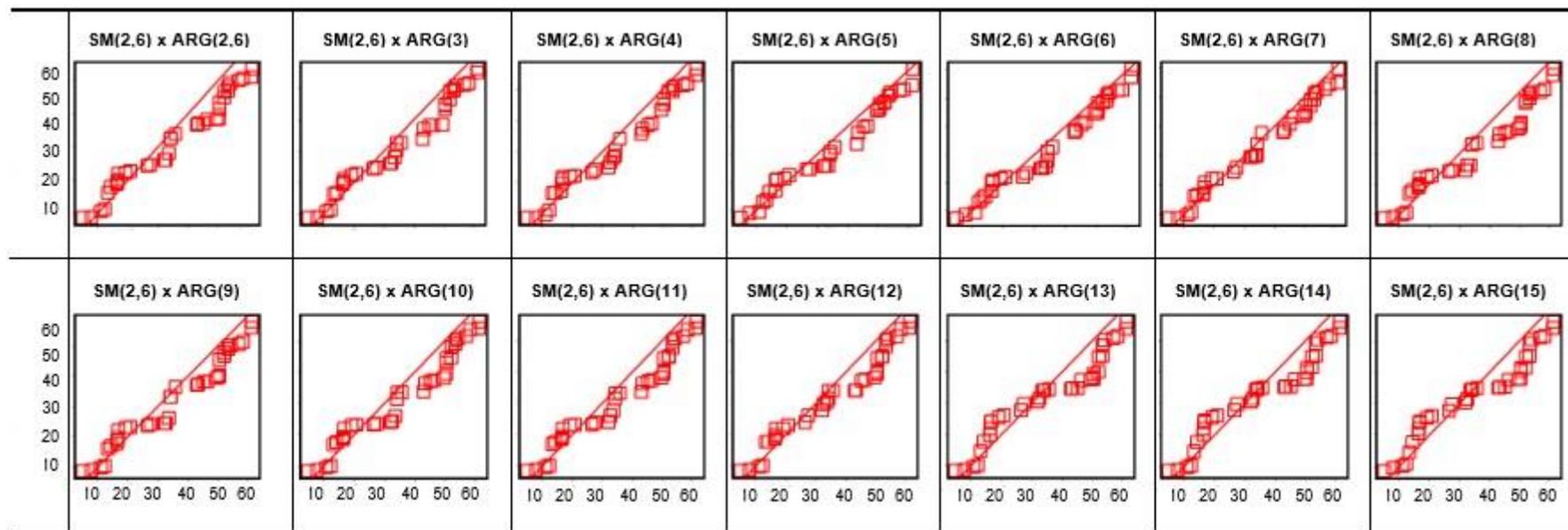


Figura 11. Gráficos de quantil-quantil para distribuição de valores observados (verdadeiros) versus distribuição de valores preditos pela cokrigagem da argila (%) para cada cenário amostral estudado.

Observação: O cenário amostral da SM corresponde a densidade de 1 ponto a cada 2,6 ha (valor comunicado entre parênteses); e, os cenários amostrais da ARG correspondem as densidades de 1 ponto à cada 2,6; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12; 13; 14 e 15 ha (valores comunicados entre parênteses) os quais constituem os cenários amostrais de estudo que, em cada um deles, incorporou-se informação da SM amostrada em 1 ponto a cada 2,6 ha. Nas abscissas os valores observados, nas ordenadas os preditos.

Os gráficos de quantil-quantil apresentados na Figura 11 ilustram o que foi afirmado nos parágrafos anteriores, ou seja, a partir da densidade de 8 amostras a cada hectare os desvios dos quantis das distribuições dos dados (observados) versus estimados (preditos) pela CKO se afastam mais da reta em relação aos demais cenários de estudo. Em suma, os desvios dos valores preditos em relação aos valores observados (nos 37 locais de validação), verificados nos gráficos, apenas emolduram qualitativamente as considerações que ora foram traçadas para os desempenhos relativos a acurácia para os cenários de amostragem estudados. Logo, a importância de se usar o gráfico de quantil-quantil em um estudo como este – que usou medidas globais (índices de RMSE) para comparações – é mostrar que os erros ponto a ponto nos 37 locais de validação não apresentam-se enviesados (grandes desvios localizados) (Isaaks e Srivastava, 1989), como pode ser notado ao se observar cada gráfico. Uma outra informação que com estes gráficos é possível verificar, é que o conjunto de dados usados (37 pontos) para validação dos mapas contempla o espectro de variação do atributo mapeado dentro da área, isto é, desde teor de argila próximo do valor mínimo (5,50%), a teor de argila no entorno do valor máximo (66,50%) (ver Tabela 1); o que demonstra que o conjunto de dados usado na avaliação da qualidade dos mapas de argila (quantificar a acurácia) possui uma boa distribuição espacial dentro da referida área (ver Figura 7).

5 CONCLUSÕES

A incorporação de informações espaciais da SM no processo de mapeamento do teor de argila do solo proporcionou aumento na acurácia das estimativas de 6 dos 14 cenários estudados: SM(2,6)×ARG(2,6), SM(2,6)×ARG(3), SM(2,6)×ARG(4), SM(2,6)×ARG(5), SM(2,6)×ARG(6) e SM(2,6)×ARG(7).

O cenário SM(2,6)×ARG(7) foi o que apresentou a maior melhoria relativa (MR_{RMSE}) das estimativas (7,46%) e propiciou considerável redução de densidade de amostragem para a argila (63%).

Os desempenhos apresentados pelos cenários SM(2,6)×ARG(7) e SM(2,6)×ARG(8) evidenciam a existência de um limite de densidade de amostragem da variável alvo que permite afirmar ser o adequado para o mapeamento com o uso

do método de CKO, pois, a CKO no cenário SM(2,6)×ARG(7) permitiu ganhos relativos de acurácia (7,46%) e redução amostral (63%) na obtenção do mapa, enquanto no cenário SM(2,6)×ARG(8), com a redução amostral em 67%, levou o mapa produzido apresentar diminuição de acurácia (19,94% de perda relativa).

6 REFERÊNCIAS

Addiscott T, Smith J, Bradbury N (1995) Critical evaluation of models and their parameters. **Journal of environmental quality** 24:803-807

Alvarez V VH, Novais RF, Barros NF, Cantarutti RB, Lopes AS (1999) Interpretação dos resultados das análises de solos. In: Ribeiro AC, Guimarães PTG, Alvarez V VH, (Eds.) **Recomendações para uso de corretivos e fertilizantes em Minas Gerais - 5º Aproximação**. Viçosa: Comissão de Fertilidade do Solo do Estado de Minas Gerais, p.25-32.

Arioli EE, Andriotti JLS (2007) Representatividade da amostragem na prospecção geoquímica. In: Licht OAB, Mello CSB, Da Silva CR. (Eds.) **Prospecção Geoquímica - depósitos minerais metálicos, não metálicos, óleo e gás**. Rio de Janeiro: SBGq, p. 594-596.

Armstrong M (1998) **Basic Linear Geostatistics**. Berlin: Springer-Verlag, 165 p.

Arnold RW, Wilding LP (1991) The need to quantify spatial variability. In.: Mausbach MJ, Wilding LP (eds.) **Spatial variabilities of soils and landforms**. Madison: SSSA, p.1-8.

Bahia ASRS (2016) **Estimação de atributos do solo por espectroscopia de reflectância difusa e suscetibilidade magnética no contexto da paisagem**. 133 f. Tese (Doutorado em Agronomia) – Unesp, Jaboticabal.

Bahia ASRS, Marques Júnior J, La Scala N, Cerri CEP, Camargo LA (2017) Prediction and mapping of soil attributes using diffuse reflectance spectroscopy and magnetic susceptibility. **Soil Sci. Society of America Journal** 81:1450-1462.

Beckett PHT, Webster R (1971) Soil variability: A review. Harpenden: **Soil Fertilizer** 34:1-15.

Bernardi ACC, Bettiol GM, Grego CR, Andrade RG, Rabello LM, Inamasu RY (2015) Ferramentas de agricultura de precisão como auxílio ao manejo da fertilidade do solo. Brasília: **Cadernos de Ciência & Tecnologia** 32:211-227.

Bernoux M, Arrouays D, Cerri CEP, Cerri CC (2007) Regional organic carbon storage maps of the western Brazilian Amazon based on prior soil maps and geostatistical interpolation. In: Lagacherie P, McBratney AB, Voltz M (Eds.) **Developments in Soil Science**. Amsterdam: Elsevier, p.497-507.

Berry JK (2002) Quantitative Methods for Analyzing Map Similarity and Zoning. In: Proceedings of GeoTech Conference on Geographic Information Systems. **Web page...** Ontario: GeoTech, 2002. Disponível em: http://www.innovativegis.com/basis/present/GIS02_similarity/GIS02_similarity.htm. Acesso em: 20 jul. 2016.

Biswas A, Zhang Y (2018) Sampling Designs for Validating Digital Soil Maps: A Review. **Pedosphere** 28:1-15.

Bongiovanni R, Lowenberg-Deboer J (2004) Precision agriculture and sustainability **Precision Agriculture** 5:359–387.

Braham S, Krige D (2013) South Africa's Giant of Geostatistics. **J. S. Afr. Inst. Min. Metall.** 113:01-03.

BRASIL (2013) **Ministério da Agricultura, Pecuária e Abastecimento. Agricultura de Precisão**. Brasília: MAPA/ACS. 36 p.

Brus DJ (2017) A valid map is a map that is made by a valid model. **Pedometron** 41:26-28.

Brus DJ, Kempen B, Heuvelink GBM (2011) Sampling for validation of digital soil maps. **European Journal of Soil Science** 62:394–407.

Burgess TM, Webster R (1980a) Optimal interpolation and isarithmic mapping of soil properties. I. The semi-variogram and punctual kriging. **J. Soil Sci.** 31:315-331.

Burrough PA (1993) Soil variability: a late 20th century view. **Soils and Fertilizers** 56:529–562.

Burrough PA (1983b) Multiscale sources of spatial variation in soil. II. A non-Brownian fractal model and its application in soil survey. **J. Soil Sci.** 34:599-620.

Burrough PA (1983a) Multiscale sources of spatial variation in soils. I. The application of fractal concepts to nested levels of soil variation. **J. Soil Sci.** 34:577-597.

Burrough PA, McDonnell RA (1998) **Principles of Geographical Information Systems**. New York: Oxford University Press.

Camargo LA, Marques Júnior J, Pereira GT, Bahia ASRS (2014) Clay mineralogy and magnetic susceptibility of oxisols in geomorphic surfaces. **Sci. Agric.** 71:244–256.

Cambardella CA, Karlen DL (1999) Spatial analysis of soil fertility parameters. **Precision Agriculture** 1:1-14.

Cambardella CA, Moorman TB, Novak JM, Parkin TB, Karlen DL, Turco RF, Konopka AE (1994) Field-scale variability of soil properties in central Iowa soils. **Soil Sci. Society of America Journal** 58:1501-1511.

Campbell JB (1978) Spatial variation of sand content and pH within single contiguous delineation of two soil mapping units. **Soil Sci. Society of America Journal** 42:460-464.

Chen F, Kissel DE, West LT, Adkins W, Clark R, Rickman D, Luvall JC (2004) Field scale mapping of surface soil clay concentration. **Precision Agriculture** 5:7-26.

Chilès J.-P, Delfiner P (2012) **Geostatistics: Modeling Spatial Uncertainty** New York: Wiley, 699p.

Cordell D, White S (2013) Sustainable Phosphorus Measures: Strategies and Technologies for Achieving Phosphorus Security. **Agronomy** 3:86-116.

Cortez LA, Marques Júnior J, Peluco RG, Teixeira DDB, Siqueira DS (2011) Suscetibilidade magnética para identificação de áreas de manejo específico em citricultura. In: SGeA - II Simpósio de Geoestatística Aplicada em Ciências Agrárias. **Anais...** Botucatu: Unesp. Disponível em: <http://www.fca.unesp.br/Home/Instituicao/Departamentos/CienciadoSolo/gepag/> Acesso em: 15 nov. 2016.

Cressie N (1990) The origins of kriging. **Mathematical Geology** 22:239–252.

Cressie N (1993) **Statistics for spatial data**. New York: Wiley, 923 p.

Curi N, Ker JC, Novais RF (Eds.) (2017) **Pedologia: Solos dos Biomas Brasileiros**. Viçosa: SBCS, 597p.

Davey BG (1990) The chemical properties of soils. In.: Campbell KO, Bowyer JW (Eds.) **The scientific basis of modern agriculture**. Sydney: Sydney University Press, p. 54-78.

Davis BM (1987) Uses and Abuses of Cross Validation in Geostatistics. **Mathematical Geology** 19:241-248.

Davis JC (1986) **Statistics and Data Analysis in Geology**. Toronto: John Wiley & Sons, 1986. 645 p.

Davis JC (1973) **Statistics and Data Analysis in Geology**. New York: John Wiley and Sons, Inc. 1973. 550 p.

De Paula R (2017) **Análise comparativa de modelos de estatística multivariada aplicados à previsão de níveis de poluentes atmosféricos**. 93 f. Dissertação (Mestrado em Engenharia Química) – Escola Politécnica da USP. São Paulo.

Dearing JA (1994) **Environmental magnetic susceptibility. Using the Bartington MS2 System**. England: British Library. 104 p.

Dearing JA, Hay KL, Baban SMK, Huddleston AS, Wellington EMH, Loveland PJ (1996) Magnetic susceptibility of soil: an evaluation of conflicting theories using a national data set. **Geophysical Journal International** 127:728–734.

Edmonds WJ, Lentner M (1986) Statistical evaluation of the taxonomic composition of three soil maps in Virginia. **Soil Sci. Society of America Journal** 50:997–1001.

EMBRAPA (1997) Empresa Brasileira de Pesquisa Agropecuária. Centro Nacional de Pesquisa de Solos. Manual de métodos de análises de solo. Rio de Janeiro: **EMBRAPA Solos**, 212 p.

EMBRAPA (2016) Programa Nacional de Solos do Brasil (PronaSolos). Rio de Janeiro: **EMBRAPA Solos**, 53 p.

FAO (1989) Sustainable agricultural production: implications for international agricultural research. Rome: **FAO** 131 p.

Fontes MPF, Oliveira TS, Costa LM, Campos AAG (2000) Magnetic separation and evaluation of magnetization of Brazilian soils from different parent materials. **Geoderma** 96:81-99.

GAMMA DESIGN SOFTWARE (2008) **GS+ Geostatistics for the Environmental Sciences. 9.0 version**. Michigan, 176 p. (User's Guide).

Geisser S (1975) The predictive sample reuse method with applications. **Journal of the American Statistical Association** 70:320-328.

GEOBANK (2006) Serviço Geológico do Brasil – CPRM. **Web Page...** Disponível em: <http://geobank.sa.cprm.gov.br/>; Acesso em: 03 mar. 2017.

Jimenez L, Zancanaro L (2012) Monitoramento da fertilidade de solo com a técnica da amostragem em grade. **IPNI - Informações Agronômicas** 138:19-25.

Goderya FS (1998) Field scale variations in soil properties for spatially variable control: a review. **J. Soil Cont.** 7:243–264.

Goodchild MF (2004) GIScience, geography, form, and process. **Annals of the Association of American Geographers** 94:709–14.

Goovaerts P (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. **Journal of Hydrology** 228:113–129.

Goovaerts P (1997) **Geostatistics for natural resources evaluation**. Oxford: Oxford University Press. 483 p.

Goovaerts P, Kerry R (2010) Using ancillary information data to improve precision of soil and crop attributes in precision agriculture. In.: Oliver MA (Ed.) **Geostatistical applications for precision agriculture**. Nova York: Springer, p.167-194.

Grego CR, Oliveira RP, Vieira SR (2014) Geoestatística aplicada a agricultura de precisão. In.: Bernardi ACC, Naime JM, Resende AV, Bassoi LH, Inamasu RY (EdS.) **Agricultura de precisão: resultados de um novo olhar**. Brasília: Embrapa, p.74-83.

Haberlandt U (2007) Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. **Journal of Hydrology** 332:144-157.

Hanesch M, Scholger R (2005) The influence of soil type on the magnetic susceptibility measured throughout soil profiles. **Geophysical Journal International** 161:50–56.

Hanesch M, Stanjek H, Petersen N (2006) Thermomagnetic measurements of soil iron minerals: The role of organic carbon. **Geophysical Journal International** 165:53–61.

Hartemink AE, Krasilnikov P, Bockheim JG (2013) Soil maps of the world. **Geoderma** 207:256–267.

Hengl T (2009) **A practical guide to geostatistical mapping**. Amsterdam: University of Amsterdam, 290 p.

Hengl T, Toomanian N (2006) Maps are not what they seem: representing uncertainty in soil-property maps. In.: Caetano M, Painho M (Eds.) **International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences**. Lisboa: Instituto Geográfico Português, p. 805-813.

Heuvelink G (2017) We can avoid confusion and misunderstanding if we make the change. **Pedometron** 41:24-24.

Heuvelink GBM (2018) Uncertainty and Uncertainty Propagation in Soil Mapping and Modelling. In.: McBratney AB, Minasny B, Stockmann U (Eds.) **Pedometrics, Progress in Soil Science**. Gewerbestrasse: Springer Nature, p. 439-461.

Heuvelink GBM, Webster R (2001) Modelling soil variation: past, present, and future. **Geoderma** 100:269-301.

Hudson BD (1992) The soil survey as paradigm-based science. **Soil Sci. Society of America Journal** 56:836–841

IBGE (2005) **Instituto Brasileiro de Geografia e Estatística - Coordenação de Recursos Naturais e Estudos Ambientais. Manual técnico de pedologia**. Rio de Janeiro: IBGE. 300 p.

Isaaks EH, Srivastava RM (1989) **Applied Geostatistics**. New York: Oxford University Press. 575 p.

Jenny H (1941) **Factors of soil formation: a system of quantitative pedology**. Toronto: Dover Publications. 281 p.

Jenny H (1994) **Factors of soil formation: a system of quantitative pedology**. New York: McGraw-Hill. 281 p.

Jordanova N (2017) **Soil Magnetism. Applications in Pedology, Environmental Science and Agriculture**. Amsterdam/Boston: Elsevier/Academic Press. 475 p.

Journel AG, Huijbregts CJ (1978) **Mining geostatistics**. London: Academic. 610 p.

Kempen B (2011) **Updating soil information with digital soil mapping**. 218 f. Thesis (Doutorado) Wageningen University, Wageningen.

Ker JC, Curi N, Schaefer CEGR, Vidal-Torrado P (Eds.) (2012) **Pedologia: Fundamentos**. Viçosa: SBCS. 343 p.

Ker JC, Motta PEF, Oliveira VA (2017) Levantamentos pedológicos e a evolução dos conhecimentos dos solos no Brasil. In.: Curi N, Ker JC, Novais RF, Vidal-Torrado P,

Schaefer CEGR (Eds.) **Pedologia - Solos dos Biomas Brasileiros**. Viçosa: SBCS. 597 p.

Kerry R, Oliver MA (2007b) Determining the effect of asymmetric data on the variogram. II. Outliers. **Comput. Geosci.** 33:1233–1260.

Kerry R, Oliver MA (2004) Average variograms to guide soil sampling. **Intern. Journal of Applied Earth Observ. Geoinformation** 5:307-325.

Kerry R, Oliver MA, Frogbrook ZL (2010) Sampling in Precision Agriculture. In.: Oliver MA (Ed.) **Geostatistical Applications for Precision Agriculture**. Heidelberg: Springer-Verlag. p. 35-63.

Keskin H, Grunwald S (2018) Regression kriging as a workhorse in the digital soil mapper's toolbox. **Geoderma** 326:22–41.

Kettler TA, Doran JW, Gilbert TL (2001) Simplified method for soil particle-size determination to accompany soil-quality analyses. **Soil Sci. Society of America Journal** 65:849–852.

Kienast-Brown S, Libohova Z, Boettinger J (2017) Digital Soil Mapping. **Web page...** Washington: USDA. Disponível em: https://www.nrcs.usda.gov/wps/PA_NRCSCconsumption/download?cid=nrcseprd1333022&ext=pdf. Acesso em 12 jan. 2018.

Krige DG (1951) A statistical approach to some basic mine valuation problems on the Witwatersrand. **Journal of the South African Institute of Mining and Metallurgy** 52:119–139.

Krige DG (1966) Two-dimensional weighted moving average trend surfaces for ore-evaluation. **Journal of the South African Institute of Mining and Metallurgy** 66:13-38.

Lagacherie P (2017) A problem well-stated is a problem half solved. **Pedometron** 41:25-25.

Lark M (2017) Careful writing should make that clear, adoption of a half-baked linguistic rule will not. **Pedometron** 41:22-23.

Lark RM, Minasny B (2018) Classical Soil Geostatistics. In.: McBratney AB, Minasny B, Stockmann U (Eds.) **Pedometrics, Progress in Soil Science**. Gewerbestrasse: Springer Nature, p. 291-340.

Lepsch IF (2013) As necessidades de efetuarmos levantamentos pedológicos detalhados no Brasil e de estabelecermos as séries de solos. **Revista Tamoios** 9:03-15.

Li J, Heap AD (2014) Spatial interpolation methods applied in the environmental sciences: A review. **Environmental Modelling & Software** 53:173–189.

Liu XM, Shaw J, Jiang JZ, Bloemendal J, Hesse P, Rolph T, Mao XG (2010) Analysis on variety and characteristics of maghemite. **Sci. China Earth Sci.** 53:1153–1162.

Lopes AS, Guilherme LRS (2000) **Uso de fertilizantes e corretivos agrícolas: aspectos agronômicos**. São Paulo: Anda. 69 p.

MacEachren AM (1995) **How maps work: Representation, visualization, and design**. New York: The Guilford Press. 526 p.

Marques Júnior J, Siqueira DS, Camargo LA, Teixeira DDB, Barron V, Torrent J (2014) Magnetic susceptibility and diffuse reflectance spectroscopy to characterize the spatial variability of soil properties in a Brazilian Haplustalf. **Geoderma** 219:63-71.

Martinez-Cob A (1996) Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. **Journal of Hydrology** 174:19-35.

Mathé V, Lévêque F, Mathé PE, Chevallier C, Pons Y (2006) Soil anomaly mapping using a caesium magnetometer: Limits in the low magnetic amplitude case. **Journal of Applied Geophysics** 58:202–217.

Matheron G (1962) **Traité de géostatistique appliqué**. Paris: Editions Technip. 171 p.

Matheron G (1963) Principles of geostatistics. **Economic Geology** 58:1246–1266.

Matheron G (1965) **Les variables régionalisées et leur estimation**. Paris: École Nationale Supérieure des Mines. 306 p.

Matheron G (1971) **The theory of regionalized variables and its applications. Les Cahiers du Centre de Morphologie; Mathématique de Fontainebleau**, Paris: Ecole des Mines de Paris. 211 p.

McBratney AB (1992) On variation, uncertainty and informatics in environmental soil management. **Australian Journal of Soil Research** 30:913–935.

McBratney AB, Pringle MJ (1999) Estimating average and proportional variograms of soil properties and their potential use in precision agriculture. **Precision Agriculture** 1:125–152.

McBratney AB, Webster R (1981) The design of optimal sampling schemes for local estimation and mapping of regionalized variables. II. Program and examples. **Computers and Geosciences** 7:335–365.

McBratney AB, Webster R. (1983a) Optimal interpolation and isarithmic mapping of soil properties. V. Co-regionalization and multiple sampling strategy. **Journal of Soil Sci.** 34:137–162.

McBratney AB, Webster R (1983b) How many observations are needed for regional estimation of soil properties. **Journal of Soil Sci.** 135:177-183.

Mcbratney AB, Webster R (1986) Choosing functions for semivariograms of soil properties and fitting them to sampling estimates. **Journal of Soil Sci.** 37:617–639.

McBratney AB, Webster R, Burgess TM (1981) The design of optimal sampling schemes for local estimation and mapping of regionalized variables. I. Theory and method. **Computers and Geosciences** 7:331–334.

Mcbratney A, Mendonça-Santos MDL, Minasny B (2003) On digital soil mapping. **Geoderma** 117:3-52.

Mcintyre GA (1967) Soil sampling for soil testing. **Journal Aust. Inst. Agric. Sci.** 33:318-320.

Mendonça-Santos ML, Manzatto CV (2007) Pedologia ou Pedometria? Modelo discreto ou contínuo? Mapeamento tradicional ou digital? **Boletim Informativo da SBCS** 32:21-26.

Miller RO, Pettygrove S, Denison RF, Jackson L, Cahn M, Plant R, Kearny T (1999) Site-specific relationships among flag leaf nitrogen, SPAD meter values and grain protein in irrigated wheat. In.: Robert PC, Rust RH, Larson WE (Eds.) **Proceedings of the Fourth International Conference on Precision Agriculture**. Minnesota: American Society of Agronomy. p. 113-122.

Minasny B (2017) What's in a name?. **Pedometron** 41:28-28.

Minasny B, McBratney AB (2016) Digital soil mapping: A brief history and some lessons. **Geoderma** 264:301–311.

Molin JP, Amaral LR, Colaço AF (2015) **Agricultura de Precisão**. São Paulo: Oficina de Textos. 238 p.

Myers DE (1982) Matrix formulation of co-kriging. **Journal Int. Assoc. Math. Geol.** 14:249–257.

Nadler I, Wein R (1998) Spatial interpolation of climatic Normals: test of a new method in the Canadian boreal forest. *Agricultural and Forest Meteorology* 92:211-225.

NATIONAL RESEARCH COUNCIL (1997) **Precision agriculture in the 21 st Century**. Washington: National Academy Press. 149 p.

Nordt LC, Jacob JS, Wilding LP (1991) Quantifying map unit composition for quality control in soil survey. In.: Mausbach MJ, Wilding LP (Eds.) **Spatial Variability of Soils and Landforms**. Soil Science Society of America 28:183–197.

Novais RF, Smyth TJ (1999) **Fósforo em solo e planta em condições tropicais**. Viçosa: UFV. 399 p.

Odeh IOA, Crawford M, Mcbratney AB (2007) Digital mapping of soil attributes for regional and catchment modelling, using ancillary covariates, statistical and geostatistical techniques. In: Lagacherie P, McBratney AB, Voltz M (Eds.) **Digital soil mapping: An introductory perspective**. Amsterdam: Elsevier p. 437-453.

O'Sullivan D, Unwin D (2010) **Geographic Information Analysis**. Hoboken: John Wiley & Sons, Inc. 404 p.

Olea RA (1999) **Geostatistics for engineers and earth scientists**. Kluwer: Springer-Science+Business Media. 303 p.

Olea RA (2006) A six-step practical approach to semivariogram modeling. **Stoch Environ Res Risk Assess** 20:307-318.

Oliveira JB (2008) **Pedologia Aplicada**. Piracicaba: Fealq. 592 p.

Oliver MA (1987) Geostatistics and its application to soil science. **Soil Use and Management** 3:8-20.

Oliver MA (2010) An overview of geostatistics and precision agriculture. In.: Oliver MA (Ed.) **Geostatistical Applications for Precision Agriculture**. New York: Springer. p.1-34.

Oliver MA (2010) **Geostatistical Applications for Precision Agriculture**. New York: Springer. 345 p.

Oliver MA, Webster R (2014) A tutorial guide to geostatistics: computing and modelling variograms and kriging. **Catena** 113:56–69.

Oliver MA, Webster R (2015) **Basic steps in geostatistics: The variogram and kriging**. New York: Springer. 100 p.

Oliver MA, Webster R (1990) Kriging: A Method of Interpolation for Geographical Information Systems. **Intern. Journal of Geographic Inform. Systems** 4:313–332.

Oreskes N, Shrader-Frechette K, Belitz K (1994) Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. **Science** 263:641-646.

Ozdemir O, Banerjee SK (1982) A preliminary magnetic study of soil samples from west-central Minnesota. **Earth Planet. Sci. Lett.** 59:393–403.

Paterniani E (2001) Agricultura sustentável nos trópicos. **Estudos Avançados** 15:303-326.

Paterson S, McBratney AB, Minasny B, Pringle MJ (2018) Variograms of Soil Properties for Agricultural and Environmental Applications. In.: McBratney AB, Minasny B, Stockmann U (Eds.) **Pedometrics, Progress in Soil Science**. Gewerbestrasse: Springer Nature, p.439-461.

Pawlowsky-Glahn V, Olea RA (2004) **Geostatistical analysis of compositional data**. New York: Oxford University Press. 304 p.

Peluco RG, Marques Júnior J, Siqueira DS, Cortez LA, Pereira GT (2013) Magnetic susceptibility in the prediction of soil attributes in two sugarcane harvesting management systems. **Engenharia Agrícola** 33:1134-1143.

Peluco RG, Marques Júnior J, Siqueira DS, Pereira GT, Barbosa RS, Teixeira DDB (2015) Mapeamento do fósforo adsorvido por meio da cor e da suscetibilidade magnética do solo. **Pesquisa Agropecuária Brasileira** 50:259-266.

Pennock D, Yates T, Braidek J (2007) Soil sampling designs. In.: Carter MR, Gregorich EG (Eds.) **Soil sampling and methods of analysis**. Boca Raton: CRC Press. p. 1-14.

Pereira GT, Souza ZM, Teixeira DDB, Montanari R, Marques Júnior J. Optimization of the sampling scheme for maps of physical and chemical properties estimated by kriging. **R. Bras Ci Solo**. 2013; 37:1128–1135.

Pierce FJ, Nowak P (1999) Aspects of precision agriculture. **Advances in Agronomy** 67:1-85.

Plant RE (2001) Site-specific management The application of information technology to crop production. **Comput. Electron. Agric.** 30:9–29.

Powell JC, Springer ME (1965) Composition and precision of classification of several mapping units of the Appling, Ceciel, and Lloyd series in Walton county, Georgia. **Soil Sci. Society of America Journal** 29:454– 458.

Preetz H, Altfelder S, Igel J (2008) Tropical soils and landmine detection – an approach for a classification system. **Soil Sci. Society of America Journal** 72:151–159.

Protz R, Presant EW, Arnold RW (1968) Establishment of the modal profile and measurement of variability within a soil landform unit. **Canadian Journal of Soil Science** 48:7-19.

R DEVELOPMENT CORE TEAM (2010) R: A language and environment for statistical computing. **Web page...** Vienna: R Foundation for Statistical Computing., 2010. Disponível em: <http://www.R-project.org> Acesso em: 25 nov. 2016

Raij BV, Cantarella H, Quaggio JA, Furlani AMC (eds.) (1996) **Recomendações de adubação e calagem para o Estado de São Paulo (Boletim Técnico, 100)**. Campinas: Instituto Agrônômico & Fundação IAC. 285 p.

Resende AV, Coelho AM (2017) Amostragem para mapeamento e manejo da fertilidade do solo na abordagem de agricultura de precisão. Piracicaba: **IPNI - Informações Agrônômicas**. 159:1-8.

Resende AV, Coelho AM (2014) Muestreo para mapeo y manejo de la fertilidad del suelo In.: Mantovani EC, Magdalena C (Eds.) **Manual de Agricultura de Precisión**. Montevideo: IICA. 176 p.

Resende AV, Shiratsuchi LS, Coelho AM, Corazza EJ, Vilela MF, Inamasu R, Bernardi ACC, Bassoi LH, Naime JM (2010) Agricultura de precisão no Brasil: avanços, dificuldades e impactos no manejo e conservação do solo, segurança alimentar e sustentabilidade.. In: XVIII REUNIÃO BRASILEIRA DE MANEJO E CONSERVAÇÃO DO SOLO E DA ÁGUA - Novos Caminhos para Agricultura Conservacionista no Brasil, **Anais...** Teresina: Embrapa Meio Norte/SBCS, p. 1-23.

Resende M, Curi N, Rezende SB (2017) Uso das informações pedológicas agrícola e não agrícola. In.: Curi N, Ker JC, Novais RF, Vidal-Torrado P, Schaefer CEGR (Eds.) **Pedologia - Solos dos Biomas Brasileiros**. Viçosa: SBCS, p. 47-110.

Resende M, Curi N, Rezende SB, Correa GF (2007) **Pedologia: bases para distinção de ambientes**. Lavras: UFLA, 2007. 322 p.

Resende M, Santana DP, Rezende SB (1988) Suscetibilidade magnética em Latossolos do sudeste e Sul do Brasil. In: Anais da III Reunião de Classificação, Correlação de Solos e Interpretações de Aptidão Agrícola, **Anais...** Rio de Janeiro: EMBRAPA-SNLCS/SECS, p. 233–258.

Resende AV, Shiratsuchi LS, Sena MC, Krahl LL, Oliveira JVF, Corrêa RF, Oro T. Grades amostrais para fins de mapeamento da fertilidade do solo em área de cerrado. **Anais...** São Pedro: Congresso Brasileiro de Agricultura de Precisão, 2006. 1 CD-ROM.

Ripoli TCC, Ripoli MLC (2004) **Biomassa de cana-de-açúcar: Colheita, energia e ambiente**. Piracicaba: Barros & Marques Eletrônica, 302p.

Rivoirard J (2005) Concepts and methods of geostatistics. In: Bilodeau M, Meyer F, Schmitt M (Eds.) **Space, Structure and Randomness**. New York: Springer-Verlag, p. 17–37.

Robert P (1993) Characterisation of soil conditions at the field level for soil specific management. **Geoderma** 60:57–72.

Robert PC, Rust RH, Larson WE (Eds.) (1993) **Proceedings of Site Specific Crop Management: A Workshop on Research and Development Issues**. Minneapolis: Soil Sci. Society of America Journal.

Robertson GP (1987) Geostatistics in ecology: interpolating with known variance. **Ecology** 68:744–748.

Robinson TP (2005) Metternicht, G; Comparing the Performance of Techniques to Improve the Quality of Yield Maps. **Agricultural Systems** 85:19-41.

Rosa G (2014) O cio da terra. In: Marques Neto JC, Julião A, Rosa G (Eds.) Terra à vista – Brasil ainda conhece pouco o potencial produtivo dos seus solos. Novos métodos de análise e mapeamento do terreno já permitem uma agricultura de alta precisão e sustentável. **Report...** São Paulo: Editora Unesp. p. 18-25. Disponível em: <https://issuu.com/unesp-ciencia/docs/uc52> Acesso em: 10 mar. 2015.

Rossiter DG (2017) Maps and models are never valid, but they can be evaluated. **Pedometron** 41:19-21.

Rossiter DG (2012) Technical Note: **Co-kriging with the gstat package of the R environment for statistical computing Version 2.3**. Enschede: University of Twente. 81 p.

Rossiter DG (2005) Digital soil mapping: towards a multiple-use soil information system. **Anal Geogr.** 32:7-15.

Russell EJ (1977) **Soil conditions and plant growth**. London: Longmans Group Ltd. 850 p.

Santana GP, Fabris JD, Goulart AT, Santana DP (2001) Magnetite and its transformation to hematite in a soil derived from steatite. **RBSC** 25:33-42.

Santos HG, Jacomine PKT, Anjos LHC, Oliveira VA, Oliveira JB, Coelho MR, Lumberras JF, Cunha TJF (2006) **Sistema Brasileiro de Classificação de Solos**. Rio de Janeiro: EMBRAPA Solos. 306 p.

Saxton K, Rawls WJ (2006) Soil water characteristic estimates by texture and organic matter for hydrologic solutions. **Soil Sci. Society American Journal** 70:1569-1578.

Schaetzl R, Anderson S (2005) **Soils: Genesis and geomorphology**. Cambridge: Cambridge University Press 817 p.

Schwertmann U (1985) The effect of environments on iron oxide minerals. **Adv. Soil Sci.** 1:172–200.

Silva AP, Nadler A, Kay B (2001) Factors contributing to temporal stability in spatial patterns of water content in the tillage zone. **Soil and Tillage Research** 58:207-218, 2001.

Silva CB, Moraes MAFD, Molin JP (2011) Adoption and use of precision agriculture technologies in the sugarcane industry of São Paulo state, Brazil. **Precision Agriculture** 12:67-81.

Siqueira DS (2010) **Suscetibilidade magnética para a estimativa de atributos do solo e mapeamento de áreas sob cultivo de cana-de-açúcar**. 75 f. Dissertação (Mestrado em Agronomia) – Unesp, Jaboticabal.

Siqueira DS (2013) **Mapeamento de atributos e planejamento amostral para latossolos utilizando suscetibilidade magnética, cor e relação solo-paisagem**. 109 f. Tese (Doutorado em Agronomia) – Unesp, Jaboticabal.

Siqueira DS, Marques Júnior J (2014) Suscetibilidade Magnética nos Solos Tropicais: uma Alternativa Promissora. **Newsletter da Comissão de Pedometria da SBCS** 3:2-5.

Siqueira DS, Marques Júnior J, Matias SSR, Barrón V, Torrent J, Baffa O, Oliveira LC (2010) Correlation of properties of Brazilian Haplustalfs with magnetic susceptibility measurements. **Soil Use and Management** 26:425-431.

Siqueira DS, Marques Júnior J, Pereira GT, Barbosa RS, Teixeira DDB, Peluco RG (2014) Sampling density and proportion for the characterization of the variability of Oxisol attributes on different materials. **Geoderma** 232:172-182.

Siqueira DS, Marques Júnior J, Teixeira DDB, Matias SSR, Camargo LA, Pereira GT (2016) Magnetic susceptibility for characterizing areas with different potentials for sugarcane production. **Pesquisa Agropecuária Brasileira** 51:1349-1358.

Smith G (1986) **The Guy Smith interviews: Rationale for concepts in soil taxonomy**. Washington: SMSS. 259 p.

SSSA (2008) **SOIL SCIENCE SOCIETY OF AMERICA – Glossary of soil science terms**. Madison: SSSA. 88 p.

Song G, Zhang J, Wang K (2014) Selection of Optimal Auxiliary Soil Nutrient Variables for Cokriging Interpolation. **Plos One** 9:1-9.

Sousa DMG, Lobato E (2004) **Cerrado: correção do solo e adubação**. Brasília: Embrapa Informação Tecnológica. 416 p.

Souza ZM, Souza GS (2011) Número de amostras e seus efeitos na análise geoestatística e krigagem de atributos do solo. In: SGeA - II Simpósio de Geoestatística Aplicada em Ciências Agrárias. **Anais...** Botucatu: Unesp. Disponível em: <http://www.fca.unesp.br/Home/Instituicao/Departamentos/CienciadoSolo/gepag/> Acesso em: 05 fev. 2017.

Stone M (1974) Cross-validatory choice and assessment of statistical predictions. **Journal of the Royal Statistical Society** 36:111-133

Teixeira DDB (2015) **Suscetibilidade magnética e análise de incertezas em mapeamento de solos**. 88 f. Tese (Doutorado em Agronomia) – Unesp, Jaboticabal.

Teixeira DDB, Bicalho ES, Panosso AR, Cerri CEP, Pereira GT, La Scala N (2013) Spatial variability of soil CO₂ emission in a sugarcane area characterized by secondary information. **Scientia Agricola** 70:195-203.

Teixeira DDB, Marques Júnior J, Siqueira DS, Vasconcelos V, Carvalho Júnior AO, Martins ES, Pereira GT (2017) Sample planning for quantifying and mapping magnetic susceptibility, clay content, and base saturation using auxiliary information. **Geoderma** 305:208-218.

Thornthwaite CW (1948) An approach towards a rational classification of climate. **Geographical Review** 38:55-94.

Tomé Júnior JB (1997) **Manual para interpretação de análise de solo**. Guaíba: Agropecuária. 247 p.

Trangmar BB, Yost RS, Uehara G (1986) Spatial dependence and interpolation of soil properties in West Sumatra, Indonesia: II. Co-regionalization and co-kriging. **Soil Sci. Society of America Journal** 50:1396–1400.

Vauclin M, Vieira SR, Vauchaud G, Nielsen DR (1983) The use of cokriging with limited field observation. **Soil Sci. Society of America Journal** 47:175-184.

Vaughan PJ, Lesch SM, Corwin DL, Cone DG (1995) Water content effect on soil salinity prediction: a geostatistical study using cokriging. **Soil Sci. Society of America Journal** 59:1146–1156.

Vieira SR (2000) Geoestatística em estudos de variabilidade espacial do solo. In: Novais RF, Alvarez V VH, Schaefer CEGR (Eds.) **Tópicos em ciência do solo. vol.1**. Viçosa: SBCS. p.1-54.

Vieira SR, Carvalho JRP, Ceddia MB, Gonzalez AP (2010) Detrending non stationary data for geostatistical applications. **Bragantia** 69:01-08.

Vieira SR, Carvalho JRP, Paz-González A (2010) Jack knifing for semivariogram validation. **Bragantia** 69:97-105.

Virdee TS, Kottegoda NT (1984) A brief review of kriging and its application to optimal interpolation and observation well selection. **Hydrol. Sci. J.** 29:367–387.

Viscarra-Rossel RA, McBratney AB (1998) Soil chemical analytical accuracy and costs Implications from precision agriculture. **Aust. J. Exp. Agric.** 7:765–775.

Voltz M, Webster R (1990) A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. **Journal of Soil Science** 41:473–490.

Wackernagel H (1998) **Multivariate Geostatistics**. Berlin: Springer-Verlag. 291 p.

Wackernagel H (2003) **Multivariate Geostatistics**. Berlin: Springer-Verlag. 387 p.

Warrick AW, Nielsen DR (1980) Spatial variability of soil physical properties in the field. In: Hillel D (Ed.) **Applications of soil physics**. New York: Academic. p. 319-344.

Webster R (2008) Geostatistics. In: McKenzie NJ, Grundy MJ, Webster R, Ringrose-Voase AJ (Eds.) **Guidelines for Surveying Soil and Land Resources**. Melbourne: CSIRO Publishing. p. 369–382.

Webster R (1985) **Quantitative spatial analysis of soil in the field**. New York: Springer-Verlag. 70 p.

Webster R (1997) Soil resources and their assessment. **Phil. Trans. R. Soc. Lond.** 352:963–973.

Webster R (2000) Is soil variation random? **Geoderma** 97:149–163.

Webster R (2015) Technological developments for spatial prediction of soil properties, and Danie Krige's influence on it. **The J. S. African Instit Min. and Metall.** 115:165-172.

Webster R, Butler BE (1976) Soil classification and survey studies at Ginninderra. **Soil Research** 14:1-24.

Webster R, Cuanalo HE (1975) Soil transect correlograms of North Oxfordshire and their interpretation. **Journal of Soil Sci.** 26:176-194.

Webster R, Lark RM (2013) **Field sampling for environmental science and management**. London: Routledge. 200 p.

Webster R, Oliver MA (2007) **Geostatistics for environmental scientists**. Chichester: John Wiley & Sons Ltd, 328 p.

Whelan BM (2018) Site-Specific Crop Management. In.: McBratney AB, Minasny B, Stockmann U (Eds.) **Pedometrics, Progress in Soil Science**. Gewerbestrasse: Springer Nature, p. 597-622.

Whelan BM (2003) **Precision Agriculture: An Introduction to Concepts, Analysis and Interpretation — A Training Course for Graduate and Industrial Professional**. Sydney: University of Sydney. p. 11–153.

White RE (2005) **Principles and practice of soil science: the soil as a natural resource**. Oxford: Blackwell. 384 p.

WIKIPEDIA – Aquífero Guarani. **Web page...** Disponível em: https://pt.wikipedia.org/wiki/Aqu%C3%ADfero_Guarani Acesso em: 15/11/2017.

WIKIPEDIA – Danie G. Krige. **Web page...** Disponível em: https://en.wikipedia.org/wiki/Danie_G._Krige Acesso em: 12/12/2017.

WIKIPEDIA – Georges Matheron. **Web page...** Disponível em: https://pt.wikipedia.org/wiki/Georges_Matheron. Acesso em: 12/12/2017.

Wilding LP, Drees LR (1983) Spatial variability and pedology. In: Wilding LP, Smeck NE, Hall GF (Eds.) **Pedogenesis and soil taxonomy. I. Concepts and interactions**. Amsterdam: Elsevier Science. p. 83-116.

Yaalon DH (2003) Are soils spatially a continuum? **Pedometron** 14:3-4.

Yamamoto JK, Landim PMB (2013) **Geoestatística: conceitos e aplicações**. São Paulo: Oficina de textos. 215 p.

YANG L, JIAO Y, FAHMY S, ZHU A-X, HANN S, BURT JE, QI F (2011) Updating conventional soil maps through digital soil mapping. **Soil Sci. Society of America Journal** 75:1044–1053.

Yang X, Hodler T (2000) Visual and statistical comparisons of surface modeling techniques for point-based environmental data. **Cartography and Geographic Inform. Sci.** 17:165-175.

Young FJ, Hammer RD, Williams F (1997) Estimation of map unit composition from transect data. **Soil Sci. Society of America Journal** 61:845– 861.

7 CONSIDERAÇÕES FINAIS

Por uma perspectiva prática (possibilidade de extrapolação das informações deste trabalho), as informações levantadas neste estudo se constituem importantes para o uso no mapeamento de argila do solo incorporando informações espaciais da SM no procedimento de CKO. Pelas informações levantadas têm-se como adequado para CKO da argila com auxílio da SM, respectivamente, a proporção amostral das variáveis em 1 : 2,69 (isto é, para cada 1 ponto amostral da argila usa-se 2,69 pontos amostrais da SM), com redução amostral da argila em 63% em relação a cobertura amostral ideal indicada por variograma prévio (análise variográfica baseada em uma amostragem intensiva em transeção linear para identificação da dependência espacial).

Com base no que foi levantado na elaboração deste trabalho (incluindo a parte de revisão de literatura) pode-se afirmar que as correlações existentes entre

argila do solo e fatores relacionados à produção vegetal (por exemplo, retenção e disponibilidade de água, nutrientes, outras substâncias) podem ser melhor exploradas em termos espaciais para operacionalização de práticas de manejo localizadas no campo (usar tecnologia de aplicação à taxa variável), tendo em vista ser um atributo que pode ser mapeado de forma mais barata (pela redução amostral) e com melhor acurácia (pela CKO com auxílio da SM). Assim como podem ser também melhor explorada as correlações existentes entre SM do solo e outros atributos do solo (como a constatada para o teor de argila do solo), já que a SM é uma medida do solo que pode ser obtida (determinada) por um procedimento analítico considerado limpo, simples, barato e acurado. Já a CKO dispensa considerações sobre seus potenciais benefícios, exceto em relação a omissão presente nos principais materiais didáticos (livros textos) que descrevem a respeito do método, mas normalmente não informam que: o bom desempenho da cokrigagem depende de uma adequada combinação (proporção) amostral para as variáveis envolvidas, como se mostrou neste trabalho.