UFRRJ

INSTITUTO DE AGRONOMIA

PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA - CIÊNCIA DO SOLO

DISSERTAÇÃO

Mapas Rápidos e Precisos de CEa do Solo para Pesquisa com EM38-MK2: Otimização do Espaçamento entre Linhas, Densidade Amostral e Resolução Espacial de Saída

Hugo Machado Rodrigues

2020

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

MAPAS RÁPIDOS E PRECISOS DE CEA DO SOLO PARA PESQUISA COM EM38-MK2: OTIMIZAÇÃO DO ESPAÇAMENTO ENTRE LINHAS, DENSIDADE AMOSTRAL E RESOLUÇÃO ESPACIAL DE SAÍDA

HUGO MACHADO RODRIGUES

Sob a Orientação do Professor
Marcos Bacis Ceddia

e co-orientação do Pesquisador
Gustavo de Mattos Vasques

Dissertação submetida como requisito parcial para obtenção do grau de **Mestre**, no Programa de Pós-Graduação em Agronomia – Ciência do Solo, Área de Concentração em Manejo do Solo e Qualidade Ambiental.

Seropédica, RJ
Fevereiro de 2020

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

Ficha catalográfica elaborada
Com os dados fornecidos pelo(a) Autor(a)

R865m	Rodrigues, Hugo Machado, 1992- Mapas Rápidos e Precisos de CEa do Solo para Pesquisa com EM38-MK2: Otimização do Espaçamento entre Linhas, Densidade Amostral e Rsolução Espacial de Saída. – Seropédica, 2020. 50 f. : il. Orientador: Marcos Bacis Ceddia. Dissertação (Mestrado). -- Universidade Federal Rural do Rio de Janeiro, Pós-Graduação em Agronomia – Ciência do Solo, 2020. 1. Senoriamento Proximal do Solo. 2. Geoestatística. 3. Krigagem. I. Ceddia, Marcos Bacis, 1968-, orient. II. Universidade Federal Rural do Rio de Janeiro. III. Título.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

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

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

HUGO MACHADO RODRIGUES

Dissertação submetida como requisito parcial para obtenção do grau de **Mestre** no Programa de Pós-Graduação em Agronomia – Ciência do Solo, na área de Concentração em Manejo do Solo e Qualidade Ambiental.

DISSERTAÇÃO APROVADA EM 28/02/2020.

Marcos Bacis Ceddia. Dr. UFRRJ
(Orientador)

Helena Saraiva Koenow Pinheiro. Dra. UFRRJ

Ronaldo Pereira de Oliveira. Dr. Embrapa Solos

AGRADECIMENTOS

À minha companheira Carolina Daltoé da Cunha

Aos meus pais e irmão, que me ensinaram de muitas maneiras.
Aos meus sogros Ary e Alizete.
Ao meu amigo e orientador Marcos Bacis Ceddia.
Ao meu amigo e co-orientador Gustavo de Mattos Vasques.
Ao meu amigo e conselheiro Marcelo Wermelinger Aguiar Lemes.
Aos meus amigos e conselheiros Sílvio Roberto de Lucena Tavares, Ronaldo Pereira de Oliveira, Luís Carlos Hernani e Ricardo de Oliveira Dart.
Ao meu amigo Juacy Campbell de Oliveira Correa.
À Embrapa Solos e seus profissionais.
À UFRRJ – PPGA-CS e seus profissionais.
À CAPES.

RESUMO

RODRIGUES, H. M. **Mapas rápidos e precisos de CEa do solo para pesquisa com EM38-MK2: Otimização do espaçamento entre linhas, densidade amostral e resolução espacial de saída**. 2020. 50f. Dissertação (Mestrado em Agronomia-Ciência do Solo). Instituto de Agronomia, Universidade Federal Rural do Rio de Janeiro, Seropédica, 2020.

Estabelecer um plano de delineamento amostral é um dos mais importantes estágios para o reconhecimento detalhado do solo e é um desafio para que interpoladores geoestatísticos, como krigagem, sejam utilizados em sistemas informatizados de suporte e apoio à decisão quando o tempo de processamento dos dados é alto. Este estudo teve o objetivo de avaliar quatro níveis de espaçamentos (40, 80, 150 e 300 m) entre linhas de coleta de dados de condutividade elétrica aparente (CEa) com o sensor Geonics EM38-MK2, em uma área de 72 ha sob pivô de irrigação, localizada no município de Itaí, São Paulo. Para avaliar o efeito do distanciamento entre as linhas nas incertezas dos mapas de CEa produzidos por krigagem ordinária, utilizou-se o índice de raiz do erro quadrado médio (REQM) e associaram-se os níveis de espaçamentos a sete resoluções espaciais para mapeamento (100, 80, 50, 30, 20, 10 e 5 m), a fim de identificar o menor tempo de processamento para produção de cada mapa via krigagem. Todas as análises foram realizadas no software R. Os mapas que utilizaram as resoluções 100, 80 e 50 m para os quatro níveis de espaçamento não demonstraram compatibilidade com a extensão da área de estudo, visto que inviabilizaram a interpretação da variação espacial da CEa. O mapa com resolução de 5 m com dados espaçados de 40 m demandou 25 minutos para execução e apresentou o menor REQM entre todos os mapas avaliados, 0,67 mS/m. O alto tempo de processamento associado ao excessivo detalhamento para operações em agricultura de precisão tornam a resolução de 5 m uma escolha inadequada. As resoluções restantes à escolha foram 30, 20 e 10 m. Observando-se que os valores de REQM aumentam conforme espaçam-se as linhas de dados e à medida que a resolução diminui, o REQM de 1,09 mS/m foi o menor valor encontrado nessas resoluções, a partir do espaçamento entre linhas de 80 m e resolução de 10 m. Para reduzir o tempo de 77 s necessários para a confecção do mapa, utilizando-se a combinação do conjunto de dados com 80 m de espaçamento e a resolução de 10 m, removeram-se, aleatoriamente, 25, 50, 75 e 95% dos pontos. O valor de REQM para o mapa com remoção de 25% foi 1,09 mS/m, e o tempo para a produção do mapa reduziu-se para 42 s. Os mapas com os dados removidos em 50 e 75% apresentaram REQM 1,21 mS/m, e tempos de processamento 20 e 4,9 s, respectivamente. Ao produzir o mapa com os dados removidos em 95% dos pontos, o tempo não ultrapassou 1 s, contudo, o REQM de 1,54 mS/m indica incertezas superiores ao se utilizar espaçamento entre linhas de 150 m. Portanto, o delineamento amostral contendo 512 pontos, referente ao conjunto com 75% de dados removidos a partir do espaçamento de 80 m e a resolução espacial de 10 m permitiram a caraterização detalhada da variação espacial da CEa do solo na área de estudo, com REQM 1,21 mS/m e tempo de processamento inferior a 10s.

Palavras-chave: Sensoriamento proximal do solo. Geoestatística. Krigagem. Incerteza.

ABSTRACT

RODRIGUES, H. M. **Fast and accurate soil ECa maps for EM38-MK2 survey: Optimizing transect spacing, sample density and output spatial resolution**. 2020. 50p. Dissertation (Master in Agronomy Soil Science). Instituto de Agronomia, Universidade Federal Rural do Rio de Janeiro, Seropédica, 2020.

Choosing a sampling design is one of the most important stages for detailed soil recognition and is a challenge for geostatistical interpolators, such as kriging, to be used in automated decision support systems when the data processing time is large. This study aimed to evaluate four levels of spacing (40, 80, 150 and 300 m) between data transects of apparent electrical conductivity (CEa) measured by the Geonics EM38-MK2 sensor, in an area of 72 ha under irrigation pivot, located in the municipality of Itaí, São Paulo, Brazil. To assess the effect of the distance between the transects in the uncertainty of the CEa maps produced by ordinary kriging, the root mean square error index (RMSE) was used, and the spacing levels were associated with seven spatial resolutions for the output map (100 , 80, 50, 30, 20, 10 and 5 m), in order to identify the shortest processing time for producing each map by kriging. The maps at resolutions 100, 80 and 50 m for the four spacing levels did not demonstrate compatibility with the extension of the study area since they made the interpretation of the spatial variation of CEa unfeasible. The 5 m resolution map made from data with 40 m spacing took 25 minutes to execute and had the lowest RMSE among all the evaluated maps, namely 0.67 mS/m. The large processing time associated with excessively detailed maps for precision agriculture operations makes the 5 m resolution a suboptimal choice. Thus, the remaining resolutions to choose from were 30, 20 and 10 m. Observing that the RMSE values increase as the transect spacing increases and resolutions degrade, the RMSE reached 1.09 mS/m as the lowest value among these resolution options, from data taken at 80 m spacing using 10 m resolution. To reduce the time of 77 s required for making the map, using the combination of the data set with 80 m spacing and resolution of 10 m, 25, 50, 75 and 95% of the points were randomly removed from this set. The RMSE value for the map with 25% removal was 1.09 mS/m, and the time for producing the map was reduced to 42 s. The maps with 50 and 75% of the data removed presented RMSE of 1.21 mS/m, and the processing times were 20 and 4.9 seconds, respectively. When producing the map with 95% of the data removed, the time did not exceed 1 s, however, the RMSE of 1.54 mS/m indicates greater uncertainty when using spacing between transects of 150 m. Therefore, the sample design containing 512 points, referring to the set with 75% of the data removed from the 80 m spacing data, and a 10 m spatial resolution, allowed assessing in detail the spatial variation of soil CEa in the study area, with a RMSE of 1.21 mS/m and processing time lower than 10 s.

Keywords: Proximal soil sensing. Geostatistics. Kriging. Uncertainty.

LISTA DE ABREVIAÇÕES

CEa	Condutividade elétrica aparente
CElab	Condutividade elétrica medida em laboratório
GCM	Medidores da condutividade do solo
GPS	Sistema de posicionamento global
IEM	Indução eletromagnética
SMa	Suscetibilidade magnética aparente
SPS	Sensores proximais, ou próximos, do solo
REQM	Índice da raiz do erro quadrado médio
PI	Profundidade de interesse

LISTA DE FIGURAS

Figura 1. Parâmetros de ajuste do semivariograma, onde: γ(h) é a semivariância; h é a distância entre os pontos amostrais; Range é o alcance; Sill é o patamar; nugget é o efeito pepita; e Fitted Semivariogram Model é o semivariograma ajustado (KWON et al., 2019). 12

Figura 2. Localização da área de estudo no município de Itaí, SP. 13

Figura 3. EM38-MK2 acoplado à caminhonete. Fonte: Hugo Machado Rodrigues. 14

Figura 4. Delineamento amostral para medição da CEa pelo SPS EM38-MK2. Referência espacial: UTM 22S, WGS 84. 15

Figura 5. Esquema ilustrativo das etapas do trabalho: Etapa 1: Seleção da melhor combinação de espaçamento entre linhas de caminhamento com EM38-MK2 e resolução espacial de saída do mapa de CEa; e Etapa 2: Seleção do melhor raleamento de pontos a partir da melhor combinação selecionada na Etapa 1. 17

Figura 6. Histogramas da condutividade elétrica aparente do solo (CEa, em mS/m) partir dos diferentes níveis de espaçamento: a) Distribuição de frequência sem transformação; b) Distribuição de frequência após a transformação log na base 10. 22

Figura 7. Semivariogramas observados (pontos) e ajustados (linhas), pelo modelo esférico, da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de espaçamento. 24

Figura 8. Mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m), resolução de 5 m (A) e 10 m (B). 26

Figura 9. Mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m), resolução de 20 m (A) e 30 m (B). 27

Figura 10. Mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m), resolução de 50 m (A) e 80 m (B). 28

Figura 11. Mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m), resolução de 100 m. 29

Figura 12. Desvio padrão da krigagem dos mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m) – Resolução de 5 m (A) e 10 m (B). 30

Figura 13. Desvio padrão da krigagem dos mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m) – Resolução de 20 m (A) e 30 m (B). 30

Figura 14. Desvio padrão da krigagem dos mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m) – Resolução de 50 m (A), 80 m (B) e 100 m (C). 31

Figura 15. Gráfico do índice da raiz do erro quadrado médio (REQM) e tempo de processamento para os dados de condutividade elétrica aparente do solo (CEa, em mS/m) a partir dos diferentes níveis de espaçamentos agrupados por resolução espacial e nível de espaçamento. 33

Figura 16. Diferenças entre os valores dos mapas dos dados de condutividade elétrica aparente do solo (CEa, em mS/m) – Resolução de 5 m (A) e 10 m (B). Da esquerda para direita: Nível 1 menos Nível 2; Nível 1 menos Nível 3; Nível 1 menos Nível 4. 34

Figura 17. Diferenças entre os valores dos mapas dos dados de condutividade elétrica aparente do solo (CEa, em mS/m) – Resolução de 20 m (A) e 30 m (B). Da esquerda para direita: Nível 1 menos Nível 2; Nível 1 menos Nível 3; Nível 1 menos Nível 4. 35

Figura 18. Diferenças entre os valores dos mapas dos dados de condutividade elétrica aparente do solo (CEa, em mS/m) – Resolução de 50 m. Da esquerda para direita: Nível 1 menos Nível 2; Nível 1 menos Nível 3; Nível 1 menos Nível 4. 35

Figura 19. Diferenças entre os valores dos mapas dos dados de condutividade elétrica aparente do solo (CEa, em mS/m) – Resolução de 80 m (A) e 100 m (B). Da esquerda para direita: Nível 1 menos Nível 2; Nível 1 menos Nível 3; Nível 1 menos Nível 4. 36

Figura 20. Histogramas da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de raleamento: a) Distribuição de frequência ; b) Distribuição de frequência após a transformação log na base 10..... 38

Figura 21. Semivariogramas observados (pontos) e ajustados (linhas), pelo modelo esférico, da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de raleamento. 39

Figura 22. Mapas da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos quatro níveis de raleamento utilizando resolução de 10 m..... 40

Figura 23. Desvio padrão da krigagem da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de raleamento utilizando resolução de 10 m. 41

Figura 24. Gráfico da REQM e tempo de processamento da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de raleamento utilizando resolução de 10 m. 42

LISTA DE TABELAS

Tabela 1. Níveis de amostragem da condutividade elétrica aparente do solo (CEa, em mS/m) partir dos diferentes níveis de espaçamento. 16

Tabela 2. Estatística descritiva da condutividade elétrica aparente do solo (CEa, em mS/m) a partir dos diferentes níveis de espaçamento. 22

Tabela 3. Parâmetros de ajuste dos semivariogramas, pelo modelo esférico, da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de espaçamento. 24

Tabela 4. Índice de erro de validação externa (raiz do erro quadrado médio; REQM) e tempo médio de processamento dos mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m) a partir dos diferentes níveis de espaçamento. 32

Tabela 5. Estimativa de tempo para coleta de dados de condutividade elétrica aparente do solo (CEa, em mS/m) a partir dos diferentes níveis de espaçamentos. 36

Tabela 6. Estatística descritiva da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de raleamento. 38

Tabela 7. Parâmetros de ajuste dos semivariogramas dos dados de condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de raleamento utilizando modelo esférico. 39

Tabela 8. Índices de erro (raiz do erro quadrado médio; REQM) da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de raleamento utilizando resolução de 10 m. 41

SUMÁRIO

1. INTRODUÇÃO 1
2. REVISÃO DE LITERATURA 4
2.1. Sensores Proximais do Solo 4
2.2. Condutividade Elétrica de Sensor Proximal e de Laboratório: Relações e Interações Ambientais 5
2.3. Delineamento Amostral 8
2.4. Variabilidade Espacial do Solo e Geostatística 9
3. MATERIAL E MÉTODOS 12
3.1. Área de Estudo 12
3.2. Sensor Proximal de Solo 13
3.3. Densidade Amostral e Resolução Espacial 15
3.4. Divisão do Processo Metodológico em Etapas 16
3.5. Etapa 1 - Avaliação do Espaçamento entre Linhas e Resolução Espacial 18
3.6. Etapa 2 - Avaliação do Adensamento Amostral 19
3.7. Estimativa de Tempo e Distância de Percurso por Nível de Espaçamento 20
4. RESULTADOS E DISCUSSÃO 21
4.1. Etapa 1 - Avaliação do Espaçamento entre Linhas e Resolução Espacial 21
4.1.1. Estatística descritiva 21
4.1.2. Semivariogramas 22
4.1.3. Krigagem ordinária 24
4.1.4. Tempo de processamento e incerteza dos mapas a partir das combinações entre níveis de espaçamento entre linhas e resolução espacial 31
4.1.5. Diferenças espaciais entre valores estimados de CEa do solo 33
4.1.6. Estimativa de tempo de percurso por nível de espaçamento 36
4.1.7. Critérios para decisão de nível de espaçamento e resolução espacial para avanço à etapa 2 37
4.2. Etapa 2 - Avaliação do Adensamento Amostral 37
4.2.1. Estatística descritiva 37
4.2.2. Semivariogramas 38
4.2.3. Krigagem ordinária 40
4.2.4. Seleção do nível de espaçamento, resolução espacial e nível de raleamento para produção do mapa final de CEa 42
5. CONCLUSÕES 43
6. CONSIDERAÇÕES FINAIS 44
7. REFERÊNCIAS BIBLIOGRÁFICAS 45

1. INTRODUÇÃO

A caracterização das propriedades do solo em alta resolução espacial é cada vez mais exigida para dar suporte ao gerenciamento sustentável da terra, carecendo, para isso, de técnicas para mensuração, caracterização e mapeamento das propriedades do solo de maneira rápida e economicamente viável. Mapas de atributos do solo têm sido usados cada vez mais em sistemas de suporte à decisão para direcionar o uso e manejo do solo na propriedade agrícola. No entanto, os métodos convencionais de obtenção de dados de atributos do solo baseados em análises laboratoriais são laboriosos, dispendiosos, e, por isso, muitas vezes limitam o número de amostras em função do custo da coleta e das análises laboratoriais das amostras, o que limita as opções de mapeamento desses atributos e a qualidade dos mapas gerados. Isso posto, é necessário implementar ou testar novos e mais eficientes métodos de obtenção e análise de dados de atributos do solo.

A obtenção de dados de solo por sensoriamento remoto é uma abordagem promissora que recebeu considerável atenção na presente década última década (MULDER et al., 2011). Apesar de ser uma abordagem sem contato e não requerer acesso ao campo, muitas vezes o sensoriamento remoto não permite o reconhecimento detalhado dos solos em alta resolução espacial e temporal, devido ao alto custo das imagens, sendo inadequado para o gerenciamento em escalas de fazenda ou talhão (>1:5000) (MCBRATNEY; MENDONÇA SANTOS; MINASNY, 2003), além da interferência de nuvens, sombreamento, cobertura vegetal e resíduos de culturas.

Para suprir a escala espacial e temporal requerida para o nível de fazenda ou talhão (>1:5000), têm-se utilizado sensores proximais ou próximos do solo (SPS), que são equipamentos portáteis que podem ser transportadas por veículos comuns, embarcados em tratores ou carregados manualmente. Comparados aos métodos convencionais de coleta e análise em laboratório, os SPS permitem a obtenção de um número elevado de amostras em campo em pouco tempo, com menor custo por amostra, e podem em alguns casos dispensar análises em laboratório, tornando o processo de mapeamento mais eficiente como um todo. A análise de solo baseada em SPS oferece diversas vantagens em relação aos métodos convencionais de coleta e análise em laboratório, como menor custo por amostra, maior eficiência para obtenção destas, geração de resultados em menor tempo e possibilidade de obtenção de um conjunto consideravelmente maior de medições em campo em comparação à amostragem de solo convencional.

Para a obtenção de dados de atributos de solo, os SPS usam diferentes princípios geofísicos de funcionamento (USDA, 2017; VISCARRA ROSSEL et al., 2011). Entre tais princípios, citam-se os sensores que atuam por indução eletromagnética (IEM), incluindo medidores de condutividade elétrica aparente (CEa) e suscetibilidade magnética aparente (SMa), por exemplo o SPS EM38-MK2 (Geonics, Mississauga, Ontário, Canadá), utilizado na presente pesquisa, além dos espectrômetros, que incluem a reflectância, no visível e infravermelho, de raios gama (γ) ou de fluorescência de raio X portáteis (ADAMCHUK et al., 2018). Esses sensores trabalham em conjunto com sistemas de posicionamento global (GPS) e, quando implementados em sistemas de informação geográfica, potencializam a interpretação desses dados e a sua conversão em informações úteis – mapas de atributos do solo – para caracterizar o padrão espacial das propriedades do solo em áreas agrícolas (SUDDUTH; DRUMMOND; KITCHEN, 2001), auxiliando a tomada de decisão pelo produtor.

Os dados de CEa têm sido utilizados para identificar a variabilidade das concentrações de água (KACHANOSKI; GREGORICH; VAN WESENBEECK, 1988; ISLAM et al., 2011)

e argila no solo (FULTON et al., 2011; PIIKKI; SÖDERSTRÖM; STENBERG, 2013; RODRIGUES; BRAMLEY; GOBBETT, 2015; SUDDUTH et al., 2005), além de servirem ao monitoramento do processo de salinização em áreas agrícolas (BENNETT; GEORGE; RYDER, 1995; GUO et al., 2015; HUANG et al., 2015; LI et al., 2013).

Os dados de CEa obtidos por SPS precisam ser processados para a geração de mapas, sendo comum o uso de interpoladores espaciais, como o interpolador pelo inverso da distância (ALMASI; JALALIAN; TOOMANIAN, 2014; NOURI et al., 2018; PIIKKI; SÖDERSTRÖM; STENBERG, 2013; RAWLINS et al., 2009), e de métodos geoestatísticos, sendo o mais comum a krigagem (COCKX; VAN MEIRVENNE; DE VOS, 2007; GHOLIZADEH et al., 2011; HUANG; PROCHAZKA; TRIANTAFILIS, 2016; JAHKNWA et al., 2014; KESKIN et al., 2018; VIEIRA; MILLETE; TOPP, 2002; YAO et al., 2013).

Contudo, o número de leituras (amostras) obtidas pelo SPS EM38-MK2 é normalmente grande, dado que esse sensor faz uma leitura a cada segundo, o que pode limitar o uso da krigagem em sistemas informatizados de suporte e apoio à decisão, devido ao alto tempo de processamento; limitando, também, o seu uso em aplicativos e sistemas *on-the-go*, por exemplo. Como o número de amostras controla o tempo de processamento na krigagem, é importante identificar o delineamento amostral e o número ideal de amostras que permitam gerar mapa de CEa com boa acurácia e qualidade visual, e com tempo de processamento razoavelmente curto que permita a adoção da krigagem em sistemas informatizados de suporte à decisão, aplicativos móveis com visualização *on-the-fly* de mapas e sistemas *on-the-go* de aplicação de adubos, corretivos e defensivos agrícolas.

Outro fator que influencia tanto a qualidade do mapa gerado quanto o tempo de processamento é a resolução espacial de saída do mapa. Ela influencia o aspecto visual e o seu uso, seja para a caracterização e interpretação dos padrões espaciais da CEa, seja para suporte para a tomada de decisão quanto ao uso e manejo do solo. No caso da resolução espacial, quanto maior o tamanho do *pixel* do mapa de saída, menor o tempo de processamento e menos refinado é o aspecto visual do mapa. Portanto, é importante identificar o maior tamanho de *pixel* que preserve qualidade visual e acurácia do mapa aceitáveis para o interesse da pesquisa, cabendo posteriormente avaliar se o tempo de processamento correspondente é fator limitante.

O trabalho norteia-se a partir da seguinte questão científica: Quais são as melhores combinações de delineamento amostral e resolução espacial do mapa de CEa quando se deseja conjugar a maior qualidade do mapa com o menor tempo de processamento para a geração desse mapa? Para isso, a dissertação teve por objetivo geral identificar as melhores combinações de delineamento amostral e resolução espacial de saída para mapear a CEa do solo por krigagem ordinária, a partir de dados obtidos pelo SPS Geonics EM38-MK2, em uma área de 72 ha sob irrigação por pivô central no município de Itaí, São Paulo.

Os objetivos específicos foram:

a. Mapear a CEa do solo usando krigagem ordinária a partir de leituras do SPS EM38-MK2 obtidas em caminhamento livre, simulando ziguezague, com diferentes números de linhas de caminhamento (transetos) e distâncias entre essas linhas, sendo elas 40, 80, 150 e 300 m;

b. Usando a abordagem acima, mapear a CEa do solo em diferentes resoluções espaciais (tamanho de *pixel*) de saída, sendo: 100, 80, 50, 30, 20, 10 e 5 m;

c. Identificar as melhores combinações de delineamento amostral e resolução espacial de saída do mapa, considerando o tempo de processamento de máquina, a acurácia das predições de CEa e o aspecto visual dos padrões espaciais obtidos em relação à variação espacial do solo na área de estudo;

d. Reduzir o tempo de processamento para menos de 10 s, por meio da remoção aleatória (raleamento) de pontos amostrais da melhor combinação escolhida acima, mantendo a acurácia e qualidade visual do mapa de CEa.

2. REVISÃO DE LITERATURA

2.1. Sensores Proximais do Solo

De maneira similar aos sensores embarcados nas diversas plataformas, como satélites ou aeronaves, os SPS apresentam variedades quanto aos princípios de funcionamento envolvidos para coleta de dados. Entre os comumente utilizados na agricultura, podem ser citados os que atuam por indução eletromagnética (IEM), como os condutivímetros e susceptibilímetros, que atuam baseados no comportamento eletromagnético dos elementos constituintes do solo (CORWIN; LESCH, 2005; FULTON et al., 2011; SUDDUTH et al., 2010; VISCARRA ROSSEL; BOUMA, 2016). Como a IEM não requer contato direto com o solo, a coleta de dados é relativamente fácil, rápida e menos custosa. Isso permite um número maior de medições e uma cobertura mais abrangente de lugares, do que seria possível realizar utilizando-se métodos tradicionais de coleta e análise de solo.

O método de IEM utiliza medidores da condutividade do solo (GCM, do inglês *Ground Conductivity Meters*). Esses medidores consistem em uma bobina transmissora e outra receptora (ou múltiplas bobinas), espaçadas no equipamento em distâncias conhecidas. Os GCMs produzem correntes elétricas alternadas no solo que iniciam nas bobinas e geram um campo eletromagnético primário variável no tempo. Este campo primário induz as correntes de Foucault (ou correntes parasitas) a fluírem através do solo e, assim, gerarem um campo eletromagnético secundário. A amplitude e a fase dos campos eletromagnéticos primários e secundários são medidas pela(s) bobina(s) receptora(s) (USDA, 2017). Em condições conhecidas como "operação em baixos números de indução" (MCNEILL, 1980), o campo secundário é proporcional à corrente de terra e é usado para calcular a condutividade elétrica "aparente" ou "em massa" (CEa) do solo, que é comumente expressa em unidades de milisiemens por metro (mS/m). As vantagens da IEM incluem a ação não invasiva ao solo, a velocidade rápida de operação e o registro contínuo de dados georreferenciados. A grande quantidade desses dados pode ser coletada de forma rápida e economicamente viável e permite uma caracterização mais completa da variabilidade das propriedades do solo, inclusive, a curtas distâncias (DE BENEDETTO et al., 2013; RODRIGUES; BRAMLEY; GOBBETT, 2015).

A eficácia do método de IEM depende do grau em que as diferenças nos valores de CEa correspondem às diferenças nas propriedades físico-químicas do solo estudado. (DOOLITTLE; BREVIK, 2014; MA et al., 2011). Vale lembrar que, por definição, a CEa é uma medida da média ponderada da condutividade elétrica em profundidade para uma coluna de materiais argilosos a uma profundidade específica (GREENHOUSE; SLAINE, 1983). Portanto, as variações na CEa são produzidas majoritariamente por mudanças na condutividade elétrica dos materiais argilosos. A condutividade elétrica dos solos é afetada principalmente pelo tipo e concentração de íons em solução, pela quantidade e tipo de argila na matriz do solo e pelo teor, temperatura e fase da água do solo (MCNEILL, 1980).

Atualmente, o mapeamento da CEa é reconhecido como um dos métodos mais promissores na agricultura para medir a variabilidade espacial das propriedades do solo nas escalas de campo e paisagem (ADAMCHUK et al., 2018; ALLRED et al., 2016). A profundidade de investigação para medições de CEa feitas com GCM é geralmente tomada como a profundidade de 70% da resposta cumulativa. A profundidade de investigação (PI) é dependente da condutividade elétrica do solo, da frequência eletromagnética que o SPS possui, da orientação e do espaçamento entre as bobinas dos GCMs (USDA, 2017). Tal profundidade, para os GCMs comumente usados em investigações do solo, pode variar de cerca de 30 a 300 cm. Também existem GCMs comercialmente disponíveis que alcançam PIs de 300 a 600 cm.

Os SPSs estão relacionados com os teores dos minerais presentes no solo, como as concentrações de minerais de argilas, a distribuição dos poros e a densidade aparente, e com a concentração de umidade (AKRAMKHANOV et al., 2008; HEIL; SCHMIDHALTER, 2018; LANDRUM et al., 2015). Os SPSs têm sido apontados como ideais para estudos em nível da escala de fazenda ou talhão (>1:5000) (HARTEMINK; MINASNY, 2014; MULDER et al., 2011; VISCARRA ROSSEL et al., 2011).

O SPS EM38-MK2 (Geonics Ltd., Mississauga, Ontário, Canadá) fornece um conjunto de dados densos e pode ser usado para atingir o objetivo de recobrir áreas com maior densidade amostral, sendo o sensor por IEM mais utilizado na agricultura (DOOLITTLE; BREVIK, 2014; SUDDUTH et al., 2013). Por sua vez, a CEa do solo, medida pelo EM38-MK2, exibe padrões associados à variabilidade espacial de outras propriedades do solo e permite, indiretamente, o reconhecimento de atributos que, por demandarem coleta e análise em laboratório, são referidos como de difícil mensuração (VISCARRA ROSSEL; BOUMA, 2016).

2.2. Condutividade Elétrica de Sensor Proximal e de Laboratório: Relações e Interações Ambientais

O princípio para a determinação da quantidade de sais presentes no solo inicia-se na possibilidade de mensurar os sais solúveis nos solos. Essa mensuração, como pode ser melhor descrita em Teixeira et al (2017) consiste da medição dos cátions e ânions no extrato aquoso após a preparação da amostra seguindo os procedimentos descritos pelos mesmos autores supracitados, permitindo a obtenção do produto necessário para a determinação da condutividade em laboratório, que é o extrato da pasta saturada. A partir de então, torna-se possível mensurar a condutividade elétrica no extrato da pasta saturada, e essa é normalemente referida na bibliografia como condutividade elétrica de laboratório.

Deve-se entender que o método utilizado para mensuração da condutividade elétrica em laboratório (CElab) demanda, primeiramente, amostragem de solo e em seguida medição laboratorial da salinidade da amostra, usando o método da pasta saturada. A técnica laboratorial demanda cuidadosa saturação da amostra por água destilada e repouso da amostra por determinado tempo (geralmente uma noite) para que os sais se dissolvam. A amostra então é submetida à sucção para extrair alguma quantidade de extrato, cuja condutividade elétrica (CE) é medida por um medidor de condutividade de laboratório convencional, como um eletrodo. Ainda assim, é possível mensurar a CElab para profundidades específicas do solo, uma vez que as análises são performadas em extratos de pastas saturadas de amostras de solos e essas podem ser coletadas em diferentes profundidades.

Outra medida de salinidade comumente usada é a condutividade elétrica do extrato de água do solo 1:2 (CE 1:2). Neste procedimento, que é mais rápido que a medição de CE para o extrato da pasta saturada descrito no parágrafo acima, é considerada a medida de salinidade por unidade de volume de solo, sendo que uma parte do solo seco é misturada com duas partes em peso de água destilada. A mistura é então agitada durante 30 minutos, e por fim mensura-se a CE da solução com um medidor de laboratório convencional. A proporção de CE da pasta saturada para para CE 1:2, que de certa forma é dependente da estrutura do solo, é de aproximadamente dois (McNeil, 1980).

Infelizmente, esses dois procedimentos são demorados (cerca de 30 min para CE 1:2) e caros de executar. Ainda deve-se considerar questões acerca da representatividade dos locais amostrados e se esses são realmente representativos de para a área investigada.

Em contraposição às análises de CElab, a mensuraçao da CE por sensores proximais apresenta pontos positivos frente à CElab uma vez que o sensor possibilita a coleta de maior quantidade de pontos. Além de dispensar a aquisição de amostras para extração da pasta

saturada. De modo que o fluxo de corrente é totalmente eletrolítico, ocorrendo através da água do solo, e uma vez que os solos normais contêm uma concentração insignificante de minerais eletricamente condutores, a CE do solo mensurada *in situ* é, portanto, uma função da CE da água do solo, que é diretamente proporcional à salinidade do solo (McNeil, 1980). No entanto, a condutividade *in situ* também é uma função do grau de saturação do solo (que controla a conectividade dos caminhos que a corrente elétrica fará), a porosidade do solo (que controla efetivamente a tortuosidade dos caminhos pelos quais a corrente flui), a temperatura do solo (por exemplo: a temperatura do eletrólito, cuja condutividade elétrica aumenta em cerca de 2% por grau *celsius*) e, finalmente, o conteúdo argiloso do solo (desde argilas com alta capacidade de troca iônica, como a montmorilonita, podem fornecer íons adicionais para melhorar a condutividade elétrica) (McNeil, 1980).

Ainda que o princípio de mensuração elétrica seja o parecido, utilizando um campo elétrico primário e após a interação da eletricidade com os sais solúveis no solo (ou na pasta saturada) o campo elétrico secundário seja recebido pela bobina receptora acoplada no condutivímetro do sensor proximal (ou pelo condutivímetro de bancada no caso da CElab), a abordagem de mensuração é diferente. Por exemplo, se estivesse sendo medida a CE da água destilada o valor de resposta dessa condutidade elétrica seria zero (ou muito próximo desse valor) para o método do condutivímetro de bancada, pois esse método exige um recipiente para conter a água e nesse a interação com o eletrodo ocorre somente entre eletrodo e água. Contudo, se fosse realizada a mensuração da CE desse mesmo copo utilizando um sensor proximal, outros materiais teriam sua CE medida, como os materiais que compõem a mesa onde o copo está apoiado e o próprio copo onde está inserida a água.

Ressalta-se que o estudo com os condutivímetros acoplados à sensores proximais apresentam potencialidades que superam o método de bancada pela potencialidade dos sensores coletarem um número de pontos em campo em órdens de grandeza superiores aos de bancada.

Inicialmente, o EM38-MK2 foi desenvolvido para monitorar a salinidade na zona das raízes (RHOADES; CORWIN, 1981) e, ao longo do tempo, muitas equações de ajustes de predição de condutividade elétrica medida em laboratório (CElab) via dados de EM38-MK2 foram propostas (LESCH et al., 1992; McKENZIE; CHOMISTEK; CLARK, 1989; SLAVICH, 1990; WOLLENHAUPT et al., 1986).

Além dos autores brevemente citados acima, podemos citar o trabalho de Rhoades e Corwin (1981) que buscaram correlações entre a salinidade medida por CElab e CEa medida pelo EM38. Os autores utilizaram uma série de parcelas de 4,3 por 4,3 m e essas foram cuidadosamente ajustadas à diferentes valores de salinização por diferentes intensidades de lixiviação utilizando água com diferentes concentrações de sal. A lixiviação foi controlada para produzir um perfil de salinidade que não variou com a profundidade. A medição da CEa foi realizada com vários espaçamentos para o arranjo de *Wenner* (método de indução de sondas eletromagnéticas) e no modo dipolo horizontal para o EM38. Os dados mostram que as duas técnicas (comparando dois tipos de sensores proximais) produzem resultados praticamente idênticos quando a condutividade elétrica do solo é uniforme com a profundidade e estabelecem que os medidores de condutividade do solo medem com precisão a condutividade do solo, apresentando r^2 de 0,96 entre os dados de CElab e CEa para ambos sensores.

Outros autores que encontraram correlações elevadas entre CE 1:2 e CEa foram Slavich e Read (1983). Eles, acidentalmente, buscando correlações entre dados de produtividade de cereais e CEa medida por EM38, encontraram valores de r^2 0,88 entre CE 1:2 e CEa medida por sensor proximal. Ainda, Cameron et al (1981), estudando 16 ha de uma área da província de Saskatchewan no Canadá, encontraram correlações de r^2 0,93 entre CEa e CElab.

Van der Lelij (1983) buscaram correlações entre CElab e CEa medida por EM38 em uma seção transversal irrigada de 5000 metros na planície do rio *Murrumbidgee*, Austrália. Esse autor realizou mensurações de CEa por EM38 a cada 20 m, e identificou o efeito da microtopografia na variação dos dados de CEa, que respondeu por algumas, mas não à todas variações de CEa (as regiões mais elevadas geralmente estão associadas à menores valores de CE).

Interpretando os dados de CEa para casos gerais, a salinidade e a condutividade do solo variam com a profundidade e permitem inferir que é possível compreender a variação da salinidade a partir da relação entre CElab e CEa de modo que a interação entre essas varie ligeiramente. Portanto, se o objetivo da pesquisa for obter valores de CElab (ou CE 1:2), o procedimento sugerido por McNeil (1980) é o de realizar a pesquisa com o sensor EM38 de condutividade elétrica do solo e, em seguida, coletar amostras de solo nos locais que a pesquisa se mostrar mais significativos, com o interesse de corroborar os dados previamente coletados por sensor proximal. Por exemplo, se o CEa variar ligeiramente em toda a área da pesquisa, serão coletadas amostras em locais que estejam associados à toda a faixa de variação. Se, por outro lado, a maior parte da área exibir baixos valores de condutividade elétrica para a maior parte da área, mas em alguns locais pontuais exibirem valores elevados para a variável de CEa, recomenda-se que a maior das amostras destinadas à CElab deva ser coletada nesses locais.

Apesar da relação entre a CEa medida pelo EM38-MK2 e a CElab ser amplamente estudada, o sensor EM38-MK2 também tem sido testado em outras frentes de pesquisa em Ciência do Solo, como, por exemplo, para a determinação da umidade do solo (HUANG et al., 2014; KACHANOSKI; GREGORICH; VAN WESENBEECK, 1988; SHEETS; HENDRICKX, 1995), determinação da capacidade de troca catiônica (RODRIGUES; BRAMLEY; GOBBETT, 2015; TRIANTAFILIS; SANTOS, 2010), estimativa e representatividade tridimensional da profundidade de drenagem (HUANG et al., 2017), determinação da textura do solo (COCKX; VAN MEIRVENNE; DE VOS, 2007; HEIL; SCHMIDHALTER, 2017; MACHADO et al., 2006) e para a distinção de tipos de solo (BECEGATO; FERREIRA, 2005; SÖDERSTRÖM et al., 2013; SUDDUTH et al., 2005).

Islam et al. (2011) utilizaram o SPS Geonics EM38-MK2 acoplado a um veículo denominado *FloSSy* (Floating Soil Sensing System), adaptado para dirigir sobre água. O estudo teve como objetivo geral identificar regiões de maior ocorrência de drenagem em uma área de plantação de arroz inundada de 1,4 ha no campo experimental da Universidade de Bangladesh, em Mymensingh. Os solos da área foram caracterizados como *floodplain soils* desenvolvidos em depósitos aluviais do rio Brahmaputra. Após a coleta dos dados de EM38-MK2, utilizaram-se *fuzzy* k-médias para agrupar duas classes de drenagem. Encontraram-se significativas diferenças entre as classes de CEa agrupadas, de modo que baixos valores de CEa representavam altas taxas de infiltração e vice-versa. Segundo o autor, essa maior infiltração estaria associada à diferença de texturas do solo, especificamente, na concentração de areia que favorecia a maior drenagem.

Anderson-Cook et al. (2002), utilizando um SPS Geonics EM38-MK2 associado a dados de produtividade de cevada, milho, trigo e soja coletados em uma área experimental de 24 ha, sob rotação de cultura nos planaltos da Virgínia – EUA, identificaram diferenças entre quatro ordens de solos. Para isso, utilizaram classificadores estatísticos não paramétricos chamados de árvores de classificação binária. As leituras de CEa e os dados de produtividade foram positivamente correlacionados às ordens de solos. Nas situações em que os valores de CEa e produtividade eram reduzidos, ocorreram solos *Bojac (Hapludults)*; ao passo que, em elevados valores de CEa e produtividade, ocorreram solos *Wickham (Ultic Hapludalfs).* As árvores de classificação utilizando os dados de CEa obtiveram 85% de acurácia para distinguir as duas ordens de solos citadas.

Embora seja perceptível a publicação de pesquisas utilizando SPS (especificamente o EM38-MK2) para o mapeamento indireto de teores de água, argila e outros atributos do solo, por exemplo, a capacidade de troca catiônica (RODRIGUES; BRAMLEY; GOBBETT, 2015) e os teores de carbono orgânico (GOMEZ; VISCARRA ROSSEL; MCBRATNEY, 2008), ainda é incipiente o estudo acerca do número ideal de pontos coletados por SPS e do adequado delineamento amostral para fins de mapeamento.

2.3. Delineamento Amostral

O delineamento amostral desempenha um papel crucial no fornecimento de informações confiáveis para a confecção de mapas de atributos do solo e no aumento da eficiência para interpolação e análise espacial. O delineamento amostral, além de ser uma etapa prévia à ida ao campo para coleta das amostras, considera o orçamento financeiro disponível e a variabilidade espacial da área de estudo. Desse modo, é um procedimento de seleção para identificação de um conjunto de locais para amostragem que recubra, de forma abrangente, o espaço geográfico da área de estudo. Caracteriza-se, ainda, como otimização para o equilíbrio das distribuições das localizações dos pontos para amostragem, considerando informações previamente disponíveis, como, por exemplo, antigos dados coletados que utilizaram densidade inferiores ao que se almeja, informações quanto a classificações geológicas e geomorfológicas (caso estejam em escala adequada), informações representativas do relevo, como modelos digitais de elevação ou terreno, e outras informações que possam ser consideradas para otimizar a distribuição dos pontos para coleta de amostras.

Atualmente, no campo de estudo em Ciência do Solo, especificamente na produção de mapas de atributos do solo obtidas por SPS, não existe uma padronização quanto à densidade de pontos coletados, bem como dos delineamentos para caminhamento com o SPS pela área de estudo, embora exista avaliação quanto às incertezas e acurácias dos mapas produzidos a partir de diferentes delineamentos amostrais, conforme apresentado no livro de De Gruijter et al. (2006). Nessa obra, são descritas diferentes estratégias para amostragens no espaço. Alguns exemplos citados aqui são as amostragens agrupadas, amostragens sistemáticas e transecto em linha para amostragem aleatória. Além do delineamento no espaço, encontram-se também delineamentos recomendados, se o interesse da coleta de dados for ao longo do tempo, e podem-se citar as estimativas de médias temporais e tendências a partir de diferentes delineamentos. Ao combinar espaço e tempo, são citados alguns exemplos de delineamentos como quantidade global espaço-temporal, quantidade global atual, médias espaciais e tendências temporais.

A partir do livro de De Gruijter et al. (2006), Brus (2019) realizou testes quanto à acurácia avaliando diferentes propostas de delineamento amostrais, de quatro áreas de estudos diferentes: 1. *Cotton Research Field, Khorezm province,* Uzbequistão; 2. *Hunter Valley, New South Wales*, Austrália; 3. *Xuancheng, Anhui province*, China; 4. *The woredas Alefa, Chilga and Dembia*, Etiópia. Na pesquisa 1, foram avaliados os dados de EM38-MK2; na pesquisa 2, foram avaliados os dados de covariáveis ambientais: elevação, curvatura, aspecto, índice topográfico composto e índice de diferença de vegetação normalizado; nas pesquisas 3 e 4, avaliaram-se os dados de carbono orgânico do solo coletados no horizonte A. Os testes realizados por Brus (2019) são acompanhados por um roteiro para que o leitor possa realizar os mesmos testes, de modo a avaliar qual seria o delineamento adequado para a área de estudo do interessado. Isso demonstra a pertinência do presente estudo para recomendação de melhor delineamento amostral para o mapeamento de propriedades do solo.

Hengl (2006) e Samet et al. (2017) complementaram o conceito de "área mínima mapeável", de modo que essa abordagem inclua as avaliações de delineamentos amostrais, pois, ao combinar o delineamento para reduzir incertezas e otimizar recursos com a resolução

espacial adequada para cada extensão de área de estudo, pode-se encontrar, para cada área de estudo em específico, uma combinação adequada entre delineamento e resolução espacial.

Outros autores, como De Carvalho Júnior et al. (2014) e Ramirez-Lopez et al. (2014), utilizaram o Hipercubo Latino Condicionado (MINASNY; MCBRATNEY, 2006) para otimizar o delineamento amostral, uma vez que esse método considera covariáveis ambientais e/ou outras informações espaciais para que a distribuição das amostras contemple, o máximo possível, a heterogeneidade das áreas de interesse. Por exemplo, é possível considerar a possibilidade de acesso às áreas onde ocorrerão as coletas de amostras ao se utilizar mapas de estradas ou ao se limitar que os pontos sejam distribuídos em áreas remotas, como cumes rochosos, utilizando-se *rasters* de geologia e geomorfologia, ou mesmo um modelo digital de elevação.

Biswas e Zhang (2018) conduziram um extenso trabalho de revisão para identificar trabalhos científicos no repositório *Web of Science* (Thomson Reuters), buscando delineamento amostral e mapeamento digital de solos como termos indexadores. Dos 146 trabalhos retornados pelo sistema, selecionaram 95 pela maior relevância. Trinta e um trabalhos apresentavam estratégias de desenhos amostrais, utilizando dados legados de atributos do solo, para que esses fossem incorporados ao banco de dados e, portanto, permitissem que outras regiões na área de estudo fossem sugeridas para amostragem. Mesmo com os critérios de seleção, encontraram 14 trabalhos que não discorreram sobre o delineamento amostral utilizado e oito que descreveram de maneira ambígua. Entre os 42 trabalhos restantes, encontraram diferentes estratégias para delineamento e amostragem: amostragens aleatórias; utilizando malhas regulares; por meio de transecto; aninhando pontos para representar melhor alguma região onde ocorre maior variância de uma propriedade; amostragens que considerassem distribuição regular espacialmente; amostragens aleatórias estratificadas; amostragem utilizando hipercubo latino e suas diversas formas; e amostragem por *fuzzy* k-médias.

Conjuntamente aos esforços para desenvolvimento e avaliação de desenhos amostrais para coleta de amostras em campo de maneira otimizada, iniciativas para implementação dessas estratégias de amostragens espaciais em programas como R, por exemplo (R CORE TEAM, 2019), demonstram que o estado da arte de delineamento amostral evolui em conjunto com as técnicas computacionais. Os delineamentos são implementados em funções e inseridos em ferramentas carregadas no *software* R, no formato de pacotes, que trazem consigo diversas funcionalidades agrupadas.

Alguns pacotes, envolvendo delineamento amostral e otimização, podem ser citados, como o *clhs* (ROUDIER, 2011), que permite implementar a amostragem por hipercubo latino e carregar inúmeras covariáveis ambientais para compor o delineamento amostral; o *spsann* (SAMUEL-ROSA, 2019), que permite implementar amostragens otimizadas, utilizando-se o conceito de *Simulated annealing* (tradução livre: recozimento simulado); e o *spcosa* (WALVOORT; BRUS; DE GRUIJTER, 2010) que, além de permitir a criação de malhas regulares de modo muito simples, gera delineamentos amostrais, considerando as localizações de dados legados para compor um delineamento amostral otimizado que reduza custos ao evitar amostras de áreas para as quais já existem dados.

2.4. Variabilidade Espacial do Solo e Geostatística

A variabilidade espacial dos atributos do solo é estudada desde o início do século XX (SMITH, 1909; MONTGOMERY, 1913). Utilizando estatística convencional, Smith (1909) buscou eliminar o efeito do tipo de solo no rendimento da variedade de trigo. Para avaliar a influência dos tipos de solo sobre a disponibilidade de nitrogênio para o rendimento de cultura de trigo, ele utilizou 224 parcelas experimentais em diferentes tipos de solo. Robinson e

Lloyd (1915) e Pendleton (1919) focaram em verificar os erros de amostragem para avaliar a diferença nos solos de uma mesma classe taxonômica. Waynick (1918) e Waynick e Sharp (1919) estudaram a variabilidade do nitrogênio e do carbono no solo, utilizando um grande número de amostras e variando os tipos de delineamentos amostrais. Todos os exemplos de pesquisas supracitadas utilizaram estatística clássica como método de análise, embora tenham começado a perceber que seria necessário considerar a variabilidade espacial de atributos do solo, uma vez que influenciavam diretamente no rendimento dos cultivares (VIEIRA; MILLETE; TOPP, 2002).

Foi com Harris (1920), com esforços para analisar uma grande quantidade de dados, que surgiram equações matemáticas, que se assemelham às utilizadas para análise de variância em blocos. Mesmo com a percepção da necessidade de levar em consideração a localização das amostras coletadas, com a utilização de repetições e casualização, foi crescente a aceitação de funções de distribuição que considerassem amostragens ao acaso e, consequentemente, a desconsideração da influência das coordenadas das amostras para interpretação estatística do tema estudado. Isso foi impulsionado por procedimentos estatísticos, como a distribuição normal de frequência, para assumir independência entre as amostras e assim permitir o uso da média e do desvio-padrão como métodos e conceitos que representassem bem o fenômeno estudado (FISHER, 1956).

Segundo Yamamoto e Landim (2013), embora tenha sido muito utilizada, a distribuição de frequência não permitia afirmar a independência entre as amostras. E foi nesse momento que se começou a considerar importante a distância e, consequentemente, a posição espacial entre as amostras, pois, como o cálculo da frequência de distribuição desconsidera a sua posição, começava a não fazer sentido, já que as propriedades medidas em diferentes posições apresentavam autocorrelação espacial. Esse desafio foi posto em evidência em 1951, na África do Sul, quando o engenheiro de minas Daniel G. Krige, trabalhando com dados de concentração de ouro (KRIGE, 1951), concluiu que não conseguia encontrar sentido na variância, caso não levasse em consideração a distância entre as amostras, ou seja, ele se deparou com a necessidade de considerar a dependência espacial para compreender a distribuição das suas amostras de interesse, dando início ao desenvolvimento de uma estatística espacial, a Geoestatística (BREVIK et al., 2016; MILLER; SCHAETZL, 2014, 2016; ROSSITER, 2018; MCBRATNEY et al., 2000).

Matheron (1963) inspirado nos trabalhos de De Wijs (DE WIJS, 1951, 1953), na Holanda, e no de Daniel Krige (KRIGE, 1951), apresentou nos anos 1960 importantes trabalhos que viriam a formalizar a Teoria das Variáves Regionalizadas, base da Geoestatística (MATHERON, 1963). De acordo com Trangmar et al. (1985) , uma variável aleatória é uma medida que varia de acordo com alguma lei de distribuição de probabilidade, sendo caracterizada pelos momentos da distribuição, tais como valor esperado (média) e variância. Quando a variável aleatória assume diferentes valores em função da localização em que é amostrada no campo, ela se torna uma variável regionalizada (WEBSTER; OLIVER, 2007).

Segundo Matheron (1971 *apud* YAMAMOTO; LANDIM, 2013), uma variável regionalizada é uma função $f(x)$ de um ponto x, mas também é uma função irregular na qual se tem dois aspectos contraditórios ou complementares: um aspecto aleatório, cuja irregularidade não permite prever as variações de um ponto a outro, e um aspecto estruturado, que reflete as características do fenômeno regionalizado.

Matheron desenvolveu um método de interpolação geoestatística e o nomeou krigagem, em homenagem a Daniel Krige (MATHERON, 1965). Esse método utiliza a dependência espacial entre as amostras, identificada no semivariograma, para que assim seja possível estimar valores em regiões da área de estudo onde não houve coleta de dados, sem que haja tendência e com variância mínima. Essas duas características fazem da krigagem um

interpolador ótimo (BURGESS; WEBSTER, 1980). A predição obtida por krigagem consiste em uma média ponderada dos vizinhos que estão dentro da distância de dependência ou autocorrelação espacial, identificada pelo alcance no semivariograma.

A krigagem parte do princípio de que pontos próximos no espaço tendem a ter valores mais parecidos do que pontos mais afastados, que é o mesmo princípio da Primeira Lei da Geografia de Tobler (TOBLER, 1970), que diz que unidades de análise mais próximas entre si são mais parecidas do que unidades mais afastadas. Dessa forma, a krigagem utiliza funções matemáticas para atribuir pesos maiores para as posições mais próximas aos pontos a serem preditos e pesos menores para as posições mais distantes, gerando predições a partir da média ponderada dos valores da variável nessas posições, ou seja, via combinação linear dos valores nas amostras vizinhas ao local de predição, em que os pesos são calculados a partir do semivariograma ajustado, conforme equação abaixo apresentada (Equação 1).

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(n)} [z(u_\alpha + h) - z(u_\alpha)]^2 \qquad \text{Equação 1}$$

Onde: $\gamma(h)$ é a semivariância entre os N pares de pontos separados a uma distância h; h é a distância euclidiana entre as observações; N é o número de pares de pontos separados por uma distância h; e $z(u_\alpha)$ e $z(u_\alpha + h)$ são os valores da variável observados em posições separadas por uma distância h, onde u_α representa o vetor de coordenada espacial. Adaptado de (KWON et al., 2019).

Para obter esses pesos, é necessário caracterizar a estrutura de autocorrelação – ou dependência espacial – da variável, por meio do semivariograma (ou variograma), que é o gráfico que representa o comportamento da semivariância da variável em função da distância entre as observações coletadas em campo (Figura 1). Ou seja, a análise da semivariância em função da distância permite modelar a dependência espacial da variável e, posteriormente, gerar mapas por meio da interpolação dos dados observados via krigagem.

Os parâmetros do semivariograma caracterizam a estrutura de dependência espacial da variável e atribuem pesos às observações para permitirem a interpolação dos dados e posterior mapeamento. Entre eles, o alcance (*range*) do semivariograma (Figura 1) representa a distância máxima dentro da qual as observações estão dependentes espacialmente entre si, ou autocorrelacionadas (JOURNEL; HUIJBREGTS, 1978). O valor da semivariância na distância máxima de autocorrelação espacial (ou seja, no alcance) é o denominado patamar (*sill*), que se aproxima da variância dos dados. O efeito pepita (*nugget effect*) é outro parâmetro do semivariograma. Indica a variância não explicada em distâncias menores do que a menor distância entre observações no campo, englobando a variação espacial não amostrada e os erros analíticos. Esses parâmetros podem ser interpretados para a tomada de decisão em relação à amostragem de dados no espaço e para a interpretação de padrões de distribuição espacial e das interações entre as variáveis regionalizadas.

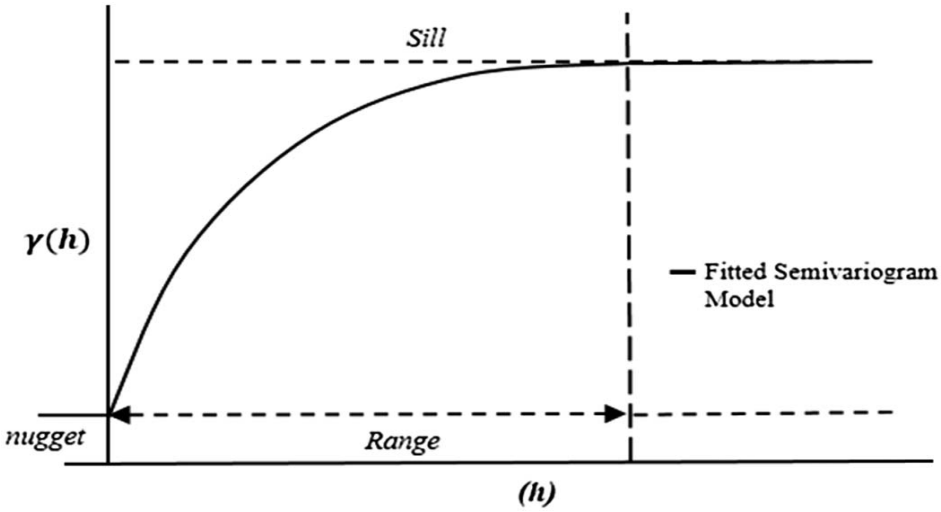

Figura 1. Parâmetros de ajuste do semivariograma, onde: γ(h) é a semivariância; h é a distância entre os pontos amostrais; Range é o alcance; Sill é o patamar; nugget é o efeito pepita; e Fitted Semivariogram Model é o semivariograma ajustado (KWON et al., 2019).

3. MATERIAL E MÉTODOS

3.1. Área de Estudo

O estudo foi realizado em uma área de produção irrigada, equipada com implemento do tipo pivô central de irrigação, possuindo uma área de 72 ha com coordenadas centrais 23,5854° de latitude Sul e 48,9395° de longitude Oeste, com elevação de aproximadamente 685 m de altitude, localizado no município de Itaí, estado de São Paulo, Brasil (Figura 2). A fazenda produz feijão, soja, trigo e aveia, em rotação de culturas, e adota sistema de plantio direto. A área de estudo está localizada em unidade de mapeamento do tipo associação de LATOSSOLO VERMELHO Distrófico típico com horizonte A moderado, textura média, álico + LATOSSOLO VERMELHO Distrófico típico com A moderado, textura argilosa, álico, ambos fase relevo ondulado e suave ondulado (ROSSI, 2017). A média pluviométrica anual para a região de Itaí – SP é de 119 mm, e as temperaturas médias anuais variam de 16 a 26 °C (RAMOS; SANTOS; FORTES, 2009).

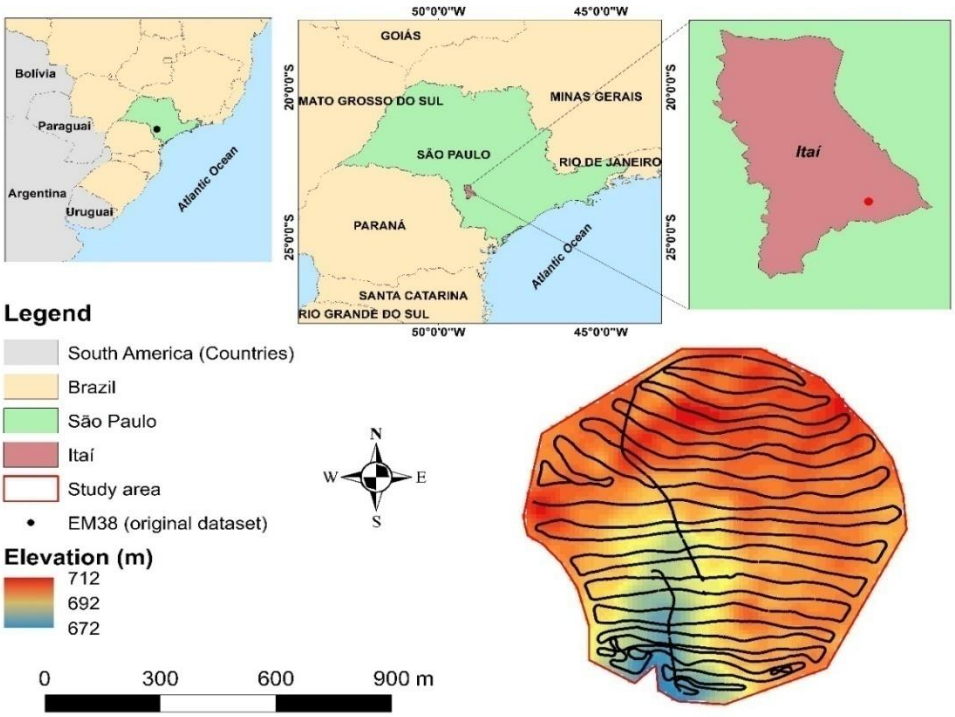

Figura 2. Localização da área de estudo no município de Itaí, SP.

3.2. Sensor Proximal de Solo

Utilizou-se o SPS EM38-MK2 (Geonics Limited, Mississauga, Canadá). Esse SPS permite realizar leituras de CEa e SMa, ambas propriedades com caráter aparente (característica representada pela inclusão do sufixo "a" às siglas), pois os valores dessas propriedades representam uma medida realizada em um volume aproximado de solo. O SPS EM38-MK2 é equipado com três bobinas fixas que permitem a medição da CEa e SMa em dois intervalos de profundidade simultaneamente. Quando se utiliza a distância entre bobinas de 0,5 m, as leituras das propriedades eletromagnéticas refletem a profundidade do solo de aproximadamente 0,75 cm, enquanto o espaçamento de 1 m entre bobinas proporciona valores referentes à profundidade de 1,5 m (BECEGATO; FERREIRA, 2005; MULLINS, 1977). Além da possibilidade de coleta de dados usando as duas bobinas, é possível operar o EM38-MK2 em duas orientações, vertical e horizontal.

Para a coleta dos dados de SPS EM38-MK2 na área de estudo, o equipamento foi colocado em uma caixa de madeira construída sem a presença de peças metálicas, como pregos ou parafusos. As junções para conformação da caixa foram feitas utilizando cola. A caixa foi presa sobre uma esteira de borracha de alta resistência, de aproximadamente 1 cm de espessura, e esse conjunto foi preso a cabos de nylon. Os cabos foram conectados à traseira de uma caminhonete, distando 3 m da mesma a fim de evitar interferência pela carroceria metálica (Figura 3).

Figura 3. EM38-MK2 acoplado à caminhonete. Fonte: Hugo Machado Rodrigues.

O EM38-MK2 foi colocado na caixa de madeira, e suas laterais foram escoradas com espuma (Figura 3), de modo que o sensor não se movimentasse para os lados, a fim de realizar as leituras utilizando a posição vertical. Uma vez que o sensor é arrastado pela área de estudo para obtenção dos dados, o modo como ele realiza as medições de CEa e SMa é por meio da configuração de um tempo constante para cada leitura. Sendo assim, configurou-se o EM38-MK2 para realizar leituras a cada 1 s e o espaçamento entre as bobinas escolhido foi o de 1 m de distância. Para aquisição do posicionamento dos pontos de CEa e SMa coletados por EM38, utilizou-se um GPS pareado ao computador de mão por meio de Bluetooth©.

Após o EM38-MK2 configurado e acoplado à caminhonete, percorreu-se a área de estudo com uma velocidade aproximada de 14 km/h e, com a finalidade de recobrir toda a área, executou-se um delineamento amostral a partir de um caminhamento livre, que buscou simular um ziguezague. A distância média entre cada linha de caminhamento foi aproximadamente 40 m (Figura 4). À época da coleta dos dados, a área encontrava-se em uma janela de plantio e o solo estava coberto por palhada de trigo. Ao todo, percorreu-se a área por aproximadamente 1h30min a partir das 13h e coletaram-se 5788 pontos em um total de 26 linhas de caminhamento com o SPS EM38-MK2.

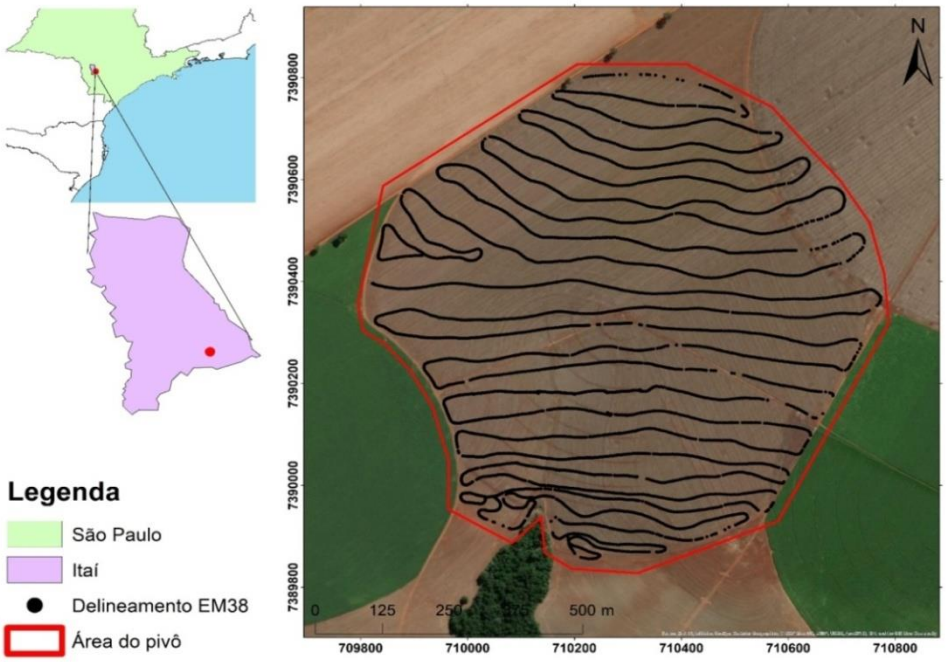

Figura 4. Delineamento amostral para medição da CEa pelo SPS EM38-MK2. Referência espacial: UTM 22S, WGS 84.

3.3. Densidade Amostral e Resolução Espacial

Os dados coletados de CEa utilizando o SPS EM38-MK2 totalizaram 5788 pontos. Para a simulação de transetos em diferentes níveis de espaçamento, eliminaram-se algumas linhas não paralelas, bem como alguns pontos próximos ao pivô de irrigação, cujos valores de CEa eram exageradamente altos, visto que o pivô de irrigação é de ferro, portanto altamente condutivo.

Do total de pontos restantes após a eliminação das linhas não regulares, tomou-se como o novo total de pontos coletados pelo SPS EM38-MK2 o conjunto de 5196 pontos. Desse total, separaram-se 525 pontos, utilizando-se seleção aleatória, para composição de um grupo de dados destinados à validação externa dos mapas produzidos.

Para avaliar a influência de diferentes delineamentos amostrais na qualidade dos mapas e tempo de processamento, estipularam-se quatro níveis de espaçamentos entre linhas de coleta de EM38-MK2 a partir do grupo de dados restantes, que seriam os dados de ajuste (treinamento) dos semivariograma, ou seja, os 5196 pontos originalmente coletados, menos os 525 pontos destinados à validação externa. O Nível 1 é o grupo de dados mais detalhado, possui 26 linhas de caminhamento, distando entre si aproximadamente 40 m e totaliza 4671 pontos (5196 – 525 = 4671 pontos) (Tabela 1; Figura 5). O Nível 2 foi obtido mantendo-se a primeira linha do Nível 1 e eliminando-se a próxima e assim sucessivamente. Ele possui 13 linhas de caminhamento, distando entre si aproximadamente 80 m e totaliza 2047 pontos. O Nível 3 foi obtido da mesma forma anterior, a partir do Nível 2, e possui 7 linhas de caminhamento com aproximadamente 150 m entre si e 1040 pontos. O Nível 4, obtido a partir do Nível 3, possui 4 linhas, distando entre si aproximadamente 300 m e totalizando 548 pontos.

Tabela 1. Níveis de amostragem da condutividade elétrica aparente do solo (CEa, em mS/m) partir dos diferentes níveis de espaçamento.

Nível de amostragem	Número de linhas de caminhamento	Espaçamento aproximado entre linhas de caminhamento (m)	Número de observações
1	26	40	4671
2	13	80	2047
3	7	150	1040
4	4	300	548

Para testar a influência da resolução espacial na qualidade dos mapas e tempo de processamento, além desses quatro níveis de amostragem, estipularam-se sete tamanhos de *pixel* (resolução espacial) do mapa de saída: 100, 80, 50, 30, 20, 10 e 5 m. A partir dos quatro níveis de amostragem e sete resoluções espaciais, respectivamente, confeccionaram-se 28 mapas (quatro níveis de amostragem x sete resoluções = 28 combinações).

3.4. Divisão do Processo Metodológico em Etapas

Para avaliar a influência de duas abordagens de delineamentos amostrais e o impacto de cada uma no produto em formato de mapa, decidiu-se separar a metodologia utilizada para cada delineamento em duas etapas e essas serão mais bem descritas nos tópicos a seguir. Nessa seção, são resumidas as etapas metodológicas do trabalho, sendo que a figura 5 mostra esquema da metodologia utilizada em cada etapa.

A Etapa 1 consiste na avaliação da influência de quatro diferentes conjuntos de dados de CEa com espaçamentos variados. Esses conjuntos foram definidos como *níveis de espaçamento* e a partir do conjunto de dados originalmente coletado, foram removidas linhas alternadamente, de modo que o nível seguinte apresentasse aproximadamente a metade de linhas do conjunto anterior. Esses Níveis de espaçamento foram então combinados à sete resoluções espaciais para confecção de 28 mapas utilizando interpolação por krigagem ordinária que será mais bem descrita nas seções seguintes.

Todos esses mapas foram avaliados quanto às suas incertezas por meio de validação externa dos mapas a partir de amostras de CEa previamente reservadas para isso e, também, pelo tempo necessário para o processamento de cada mapa, considerando os diferentes conjuntos de níveis de espaçamento utilizados como entradas de dados (quanto menor o espaçamento maior a quantidade de dados de CEa e vice-versa). Dentre os 28 mapas, a combinação entre nível de espaçamento e resolução espacial que apresentou menor tempo de processamento e menor índice de erro foi selecionada e com objetivo de reduzir ainda mais o tempo necessário para o seu processamento, o que foi feito na Etapa 2.

A Etapa 2 consistiu em utilizar a combinação entre nível de espaçamento e resolução espacial que exibiu menor incerteza quanto à validação externa do mapa e o menor tempo necessário para o processamento na Etapa 1. O conjunto de dados selecionado foi reduzido a partir da remoção aleatória de pontos, gerando quatro novos conjuntos de dados de CEa a partir de quatro níveis de raleamento (remoção) de pontos. Sendo assim, a Etapa 2 avaliou a influência de quatro *níveis de raleamento* de pontos e, de maneira similar à Etapa 1, avaliou-se a incerteza de quatro mapas por meio de validação externa com o conjunto de pontos de CEa previamente reservados e o tempo de processamento para geração de cada mapa. Ressalta-se que a Etapa 2 avaliou somente quatro mapas, pois nessa etapa não houve variação de resolução espacial do mapa de saída, uma vez que essa foi definida na Etapa 1.

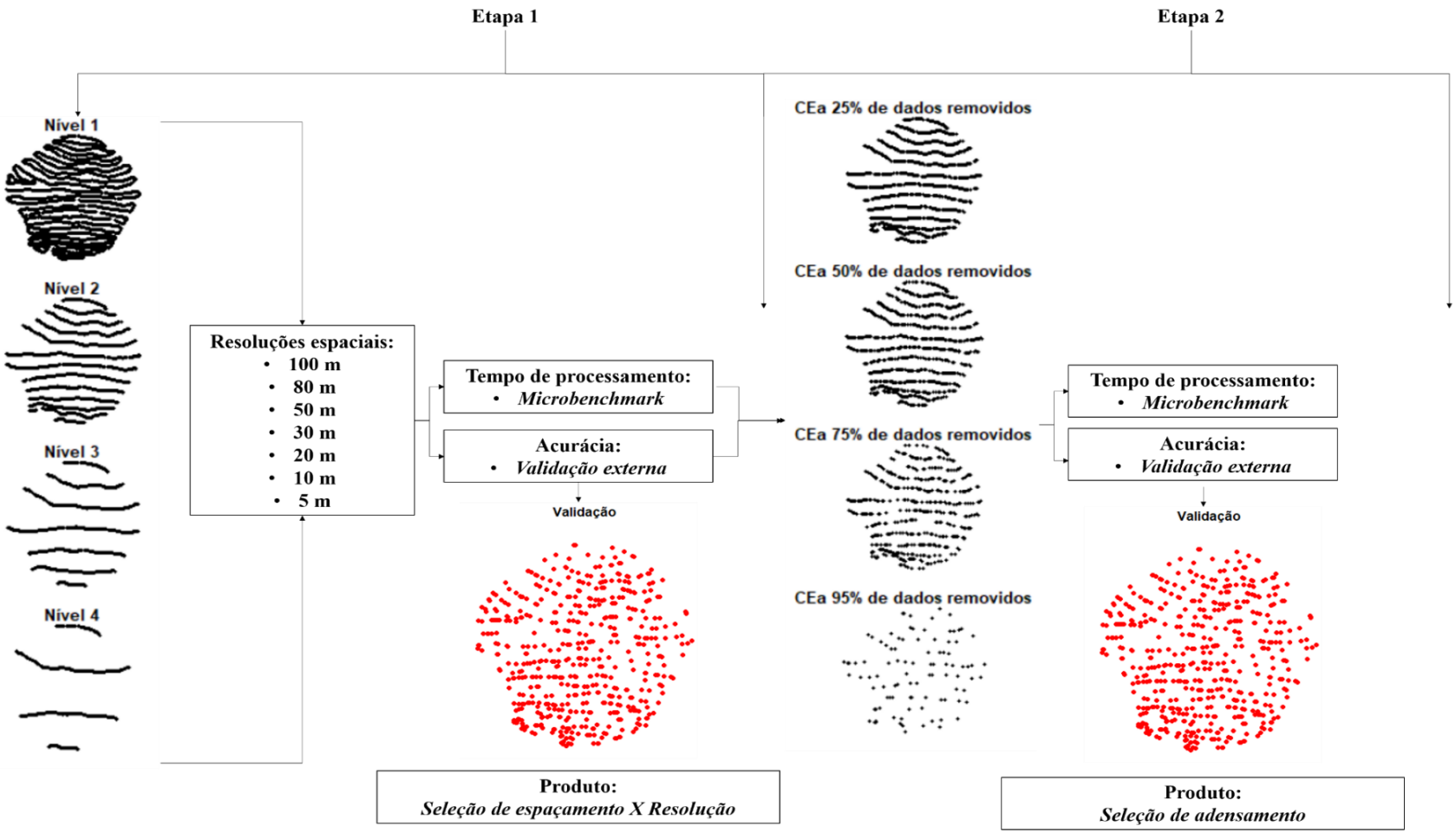

Figura 5. Esquema ilustrativo das etapas do trabalho: Etapa 1: Seleção da melhor combinação de espaçamento entre linhas de caminhamento com EM38-MK2 e resolução espacial de saída do mapa de CEa; e Etapa 2: Seleção do melhor raleamento de pontos a partir da melhor combinação selecionada na Etapa 1.

3.5. Etapa 1 - Avaliação do Espaçamento entre Linhas e Resolução Espacial

Foi realizada análise estatística descritiva dos dados de CEa, utilizando-se o *software* R (R CORE TEAM, 2019). Para realizar o mapeamento da CEa medida pelo EM38 foi utilizado o método de krigagem ordinária (KO). Esse método de interpolação é o mais amplamente utilizado é permite estimar o valor de um ponto em uma região onde esse não foi medido por meio do ajuste do variograma, utilizando os dados vizinhos a esse (CEDDIA et al 2015, *apud* WACKERNAGEL, 2003). Para avaliar os diferentes desenhos amostrais que serão propostos para a Etapa 1 (e em seguida para Etapa 2), a variável primária Z_i (representada pela CEa), mensurada a cada um segundo em campo para cada localização u será utilizada para estimar o valor de CEa em localizações não visitadas ($Z^*(u)$). A estacionaridade dos valores da média será assumida apenas para dentro da vizinhança W(u), centralizada na localização que está sendo estimada u, como descrito por Ceddia et al (2015). Para o método de interpolação por KO o valor da média é considerado constante, embora esse seja desconhecido, por exemplo, pode ser descrito como m(u') = valor constante, mas desconhecido, $\forall u'$ pertencente ao W(u). Portanto, a KO é um estimador escrito como uma combinação linear de n(u) para os dados Z(u_α) com uma simples condição não enviesada (equação 2):

$$Z^*(u) = \sum_{a=1}^{n(u)} \lambda_a\,(u)[Z_1(u)] \;\text{ com }\; \sum_{a=1}^{n(u)} \lambda_a = 1 \qquad \text{Equação 2}$$

Ainda sobre a equação 1, o valor da média desconhecida para os locais m(u) o é filtrada por um estimador linear e forçando os pesos atribuídos à krigagem somarem o valor 1 (CEDDIA et al, 2015).

Também se avaliaram os dados de CEa quanto aos seus padrões de distribuição para avaliar se estes atendiam aos pressupostos de normalidade, para que os dados pudessem ser submetidos à interpolação por meio da krigagem ordinária. Todos os dados separados em subgrupos para as Etapas 1 e 2 foram transformados para seu valor em logaritmo na base 10 ($\log_{10}$).

Cada nível de amostragem teve um semivariograma ajustado pelo modelo esférico (Equação 3) e, em seguida, interpolaram-se esses dados utilizando-se krigagem ordinária (Equação 4) (WEBSTER; OLIVER, 2007) a partir da combinação entre nível de espaçamento (quatro níveis) e resolução espacial (sete resoluções) (Figura 5, Etapa 1).

$$Y(h) = \begin{Bmatrix} c\left\{\frac{3h}{2a} - \frac{1}{2}(h/a)^3\right\} \\ c \end{Bmatrix}, \begin{matrix} h \leqslant a \\ h \geq a \end{matrix} \qquad \text{Equação 3}$$

Onde: c_0 é o efeito pepita, c é a variância do patamar (sill) e a é o alcance (range). h é o vetor de separação de pares de pontos no qual $h = x_i - x_j$ e onde $C(x_i, x_j) = E\left[\{Z(x_i - \mu)\}\{Z(x_j) - \mu\}\right]$ que é uma constante para qualquer h. Essa constância da média, variância e covariância dependem apenas da separação e não da posição absoluta.

$$Z_{KO}^{\circ}(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) \qquad \text{Equação 4}$$

Onde: $Z_{KO}^{\circ}(x_0)$ é o valor da variável aleatória a ser estimado por krigagem ordinária (KO), λ_i são os pesos ótimos calculados sob duas condições de restrição: a) o estimador não é enviesado; e b) a variância da estimativa seja mínima, e $Z(x_i)$ são os valores da variável aleatória nos n pontos amostrais.

Para avaliação quanto à acurácia de cada um dos 28 mapas de CEa produzidos, calculou-se o índice da raiz do erro quadrado médio (REQM, Equação 5) de cada mapa, utilizando-se os 525 pontos de CEa previamente separados aleatoriamente para validação externa. Os semivariogramas foram ajustados manualmente, e a krigagem ordinária realizada usando a função krigeTg do pacote *gstat* (PEBESMA, 2004), do software R (R CORE TEAM, 2019), pois essa função permite trabalhar com dados transformados em $\log_{10}$ e realiza no próprio algoritmo a conversão dos valores em $\log_{10}$ para os valores na unidade de medida original.

$$REQM = \sqrt{\sum_{i=1}^{N} \frac{(y-\hat{y})^2}{N}} \qquad \text{Equação 5}$$

Onde: N é o número de observações, y é o valor observado e $\hat{y}$ é o valor predito

Além da acurácia estimada pela REQM, calculada a partir dos mapas reconvertidos para a unidade original, avaliou-se o tempo de processamento de máquina necessário para produção de cada um dos mapas de CEa via krigagem ordinária, a partir de cada combinação entre nível de amostragem e resolução espacial de saída do mapa. A avaliação do tempo médio de processamento foi baseada em 10 repetições do processo de krigagem ordinária a partir de cada combinação, utilizando-se a função *microbenchmark* do pacote *microbenchmark* (MERSMANN, 2019), no software R (R CORE TEAM, 2019), de maneira que essas repetições permitissem que o tempo estimado para o processamento de cada mapa fosse reflexo da média de 10 repetições para a confecção de cada mapa. O tempo foi medido somente na aplicação da krigagem ordinária propriamente dita, usando a função krigeTg.

Para verificar as consequências do aumento do espaçamento entre linhas e da diminuição da resolução espacial, respectivamente, na qualidade visual e incerteza do mapa, tomou-se novamente o mapa que utilizou 26 linhas para sua produção, como fator de comparação e subtraiu-se do mesmo cada um dos mapas produzidos a partir de 13, 7 e 4 linhas (Níveis 2, 3 e 4), de modo que as diferenças entre os valores de CEa estimados nos diferentes níveis de espaçamento e resoluções fossem espacialmente exibidas.

3.6. Etapa 2 - Avaliação do Adensamento Amostral

A partir da escolha da melhor combinação entre o nível de espaçamento entre linhas de caminhamento e a resolução espacial, utilizando-se os critérios a qualidade visual do mapa gerado, o tempo de processamento e a acurácia do mapa medida pela REQM, foi possível avançar para a etapa de avaliação quanto ao raleamento aleatório de pontos (Figura 5, Etapa 2).

A segunda etapa avaliou diferentes níveis de raleamento (ou adensamento) de pontos amostrais, simulando densidades amostrais menores a partir da exclusão aleatória de pontos do conjunto amostral selecionado na melhor combinação de nível de amostragem e resolução espacial na etapa anterior. Aqui o objetivo foi obter mapas de CEa com qualidade visual e acurácia similares àquele obtido a partir da melhor combinação na etapa anterior, diminuindo o tempo de processamento via redução do tamanho amostral para menos de 10 s.

Assim, uma vez definida a melhor combinação de nível de espaçamento entre linhas de caminhamento e resolução espacial na primeira etapa, o conjunto de dados presente nessa combinação foi raleado pela remoção aleatória de pontos, obtendo-se quatro níveis de raleamento: Nível 1 com remoção aleatória de 25% dos pontos; Nível 2 com remoção aleatória de 50% dos pontos; Nível 3 com remoção aleatória de 75% dos pontos; e Nível 4

com remoção aleatória de 95% dos pontos, restando apenas 5% dos pontos originais para ajuste do variograma, interpolação e geração do mapa de CEa.

Cada nível definido de raleamento foi submetido à estatística descritiva para os dados de CEa, utilizando-se o *software* R (R CORE TEAM, 2019). Também se avaliaram os dados de CEa definidos nos quatro níveis de raleamento quanto aos seus padrões de distribuição, para avaliar se eles atendiam aos pressupostos de normalidade para que os dados pudessem ser submetidos à interpolação por meio da krigagem ordinária.

Para o procedimento de interpolação, primeiro, ajustaram-se semivariogramas para cada um dos quatro níveis de raleamento, e depois geraram-se os mapas na mesma resolução espacial selecionada na melhor combinação da etapa anterior por krigagem ordinária, usando-se o pacote *gstat* (PEBESMA, 2004) do software R (R CORE TEAM, 2019).

Uma vez que nessa etapa estipulou-se uma resolução espacial única para a confecção dos mapas, avaliaram-se ao total quatro mapas: quanto à sua acurácia, utilizando-se o índice de REQM a partir dos 525 pontos previamente separados para validação externa; quanto à sua qualidade visual para representar os padrões espaciais da CEa do solo na área de estudo e quanto ao tempo de processamento de máquina, necessário para krigagem e produção dos quatro mapas de CEa, calculado em 10 repetições pela função *microbenchmark*, como na etapa anterior.

Na Figura 5 é apresentado um fluxograma para ilustrar a primeira etapa, em que se selecionou a melhor combinação de espaçamento de linhas de caminhamento e resolução espacial de saída do mapa de CEa e, também, para ilustrar a segunda etapa, em que se selecionou o melhor nível de raleamento para produção do mapa final de CEa, a partir da combinação escolhida na etapa anterior.

3.7. Estimativa de Tempo e Distância de Percurso por Nível de Espaçamento

Uma vez que as variações de conjuntos de dados presentes na Etapa 1 são simulações de caminhamento provenientes do conjunto de dados originalmente coletado em campo, sendo ele representado pelos dados de CEa presentes no Nível 1 somados aos 525 pontos reservados para validação dos mapas, não foi possível medir o tempo necessário para coleta dos dados de CEa referentes aos níveis de espaçamento subsequentes, sendo eles os Níveis 2, 3 e 4.

Para estimar o tempo de caminhamento necessário para cada nível de espaçamento, primeiro foi necessário identificar a extensão total caminhada por cada nível subsequente ao Nível 1. Para isso, utilizou-se a função *LineLength* do pacote *sp* presente no software R. Com o interesse de estimar o tempo utilizado para percorrer a área em cada nível de espaçamento, primeiro estimou-se a velocidade média da caminhonete dividindo-se a distância total percorrida nas 26 linhas do Nível 1 pelo tempo de percurso, considerando como tempo de percurso 5196 s, já que o sensor realizou 4671 leituras no Nível 1 somados aos 525 pontos removidos para validação com a configuração de 1 leitura/s. Tendo em vista que o sensor foi calibrado para realizar a mensuração de cada ponto de CEa baseando-se em tempo, e esse tendo sido configurado para um segundo, utilizou-se a equação 5 para estimativa da velocidade.

$$Velocidade = \left(\frac{\Delta_s}{\Delta_t}\right) \qquad \text{Equação 5}$$

Onde: Δ_s é o espaço percorrido e Δ_t é o tempo de percurso.

4. RESULTADOS E DISCUSSÃO

4.1. Etapa 1 - Avaliação do Espaçamento entre Linhas e Resolução Espacial

4.1.1. Estatística descritiva

As estatísticas descritivas da CEa do solo para cada nível de espaçamento, bem como para os dados de CEa destinados à validação externa dos mapas, são apresentadas na Tabela 2. Os valores mínimos não variaram entre os diferentes níveis de espaçamentos (3,28 mS/m), com exceção do Nível 1, que apresentou valor de 2,62 mS/m. O valor máximo de CEa para os Níveis 1, 2 e 3 foi 31,48 mS/m, enquanto para o Nível 4 foi 26,25 mS/m. Devido aos subsequentes níveis de espaçamento serem derivados do Nível 1, os valores de mínimo e máximo se assemelharam em alguns níveis.

Todos os níveis de espaçamento apresentaram valores de média e mediana próximos a 10 mS/m, indicando que, embora os espaçamentos aumentem as distâncias entre linhas conforme os Níveis avançam do 1 ao 4, os dados ainda possuem similaridades quanto a seus valores médios. O valor de desvio padrão para o Nível 1 foi 3,85 mS/m. Os valores de desvio padrão dos Níveis 2, 3 e 4 diminuíram devido à redução da quantidade de pontos em comparação ao Nível 1 (Tabela 2).

As assimetrias foram moderadas e podem ser classificadas como *assimétricas positivas* para todos os níveis de espaçamento, indicando que a maioria dos valores são menores ou se concentram à esquerda da média aritmética e que a "cauda" do histograma é mais longa à direita, como pode ser visto na Figura 6a.

Os coeficientes de curtose para todos os níveis, inclusive para os dados de validação, apresentaram valores maiores que 0,3, indicando que o achatamento do padrão de distribuição dos dados pode ser classificado como *leptocúrtico*, enquanto valores de curtose próximos a 0, indicariam uma distribuição normal dos dados.

Assim, como os valores de curtose e assimetrias sugerem que a distribuição dos dados para todos os níveis não seja normal – sendo esse um preceito para o prosseguimento para as análises geoestatísticas, como ajustes de semivariogramas e posterior krigagem – optou-se por transformar os dados para que os padrões de distribuição se normalizassem. Para isso, todos os dados foram transformados para logaritmo na base 10.

Após a transformação dos conjuntos de dados de CEa para os quatro níveis de espaçamento entre linhas, o aspecto do histograma (Figura 6b) apresentou melhoria quanto à distribuição dos dados e aproximou essa distribuição a uma normalidade. Dessa forma, prosseguiu-se para os ajustes de semivariogramas e interpolação por krigagem, utilizando-se os dados de CEa para todos os níveis transformados para logaritmo na base 10.

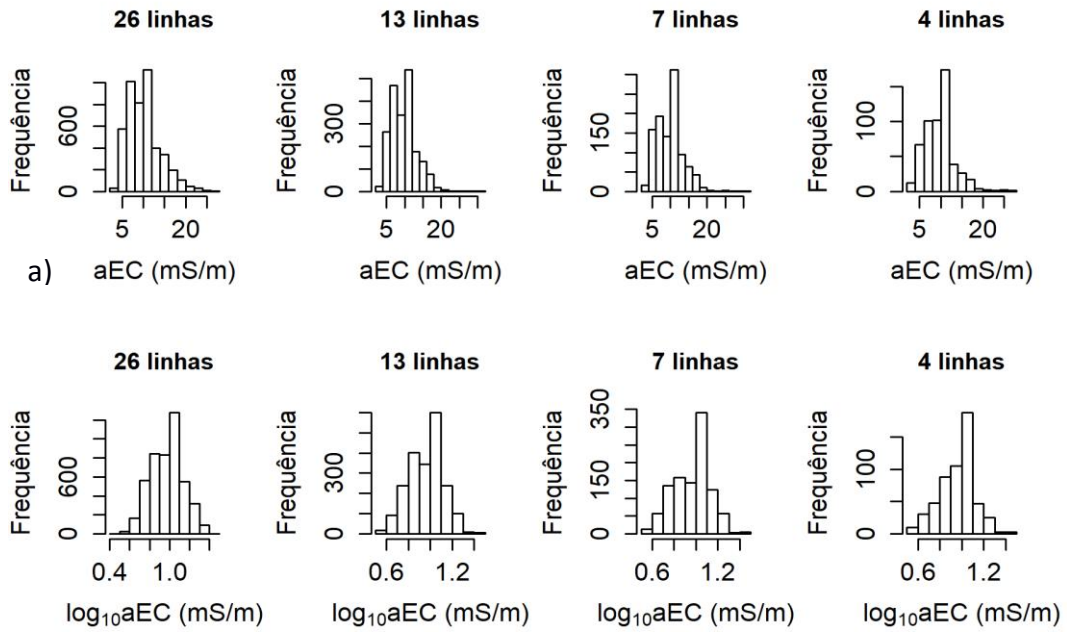

Figura 6. Histogramas da condutividade elétrica aparente do solo (CEa, em mS/m) partir dos diferentes níveis de espaçamento: a) Distribuição de frequência sem transformação; b) Distribuição de frequência após a transformação log na base 10.

Tabela 2. Estatística descritiva da condutividade elétrica aparente do solo (CEa, em mS/m) a partir dos diferentes níveis de espaçamento.

Estatística	CEa (mS/m)				Validação
	Treinamento				
	Nível 1	Nível 2	Nível 3	Nível 4	
Observações	4671	2047	1040	548	525
Mínimo	2,62	3,28	3,28	3,28	3,44
Máximo	31,48	31,48	31,48	26,25	24,18
Média	10,10	9,73	9,84	9,75	10,13
Mediana	9,81	9,61	10,12	9,90	9,96
Desvio Padrão	3,85	3,48	3,64	3,43	3,76
Assimetria	0,89	0,82	0,84	0,90	0,84
Curtose	0,77	1,43	2,04	2,27	0,70

4.1.2. Semivariogramas

Após os dados atenderem à distribuição normal requerida para a krigagem, ajustaram-se manualmente os semivariogramas para cada nível de espaçamento, e os parâmetros de ajustes são apresentados na Tabela 3. Os semivariogramas ajustados pelo modelo esférico são exibidos na Figura 7. O efeito pepita para o Nível 1, considerando-o como o "melhor" espaçamento, já que possui mais linhas e menor distância entre elas, apresentou menor valor (Tabela 3). O efeito pepita representa a parte da semivariância que se traduz como erro ou acaso, ou seja, é a parte da semivariância não estrutural e que não é explicada pelo modelo ajustado ao semivariograma experimental. Conforme aumenta-se esse efeito, diretamente se aumenta a aleatoriedade. Quanto maior for o valor do efeito pepita, maior será a variância nos mapas krigados. Ao espaçar as linhas e avançar para análise dos Níveis 2, 3 e 4, o valor dos

efeitos pepita aumentam, indicando que a redução do número de pontos e o distanciamento entre linhas contendo os dados de CEa aumentam a incerteza dos mapas.

O alcance diminui progressivamente, conforme reduzem-se os números de pontos e aumentam-se os espaçamentos entre as linhas do Nível 1 ao 4. A partir dos conjunto de dados com menor espaçamento entre linhas (Nível 1) o valor de alcance foi 533 m, enquanto o conjunto com maior espaçamento entre linhas (Nível 4) apresentou valor de alcance de 475 m. É possível observar na tabela 3 que os valores de alcance são reduzidos ligeiramente a medida que reduzem-se o número de linhas por conjunto de dados de CEa (representado pelos níveis de espaçamento). Dessa forma, pode ser identificada a influência da quantidade de linhas e, conforme o distanciamento é crescente entre essas, reduz-se o reconhecimento detalhado da CEa do solo da região.

Lopes e Montenegro (2019) trabalharam com EM38 nas posições vertical e horizontal e, utilizando as duas distâncias entre bobinas, buscaram correlações entre medidas de CEa obtidas pelo sensor e os valores de umidade e condutividade hidráulica do solo. A área de estudo dos autores possui 2,8 ha com três usos diferentes de cobertura do solo em Parnamirim, região semiárida de Pernambuco, Brasil. Os autores trabalharam com um grid amostral de 20 x 10 m, totalizando 80 pontos amostrais para coleta de amostras para definição da umidade e condutividade hidráulica. Um grid secundário de 10 x 10 m contendo 152 pontos foi utilizado para mensuração da CEa pelo EM38. As leituras de CEa foram realizadas em diferentes alturas de posicionamento do mesmo sobre o solo (0, 0,3, 0,6, 1,2 e 1,5 m). Para cada altura de leitura de CEa ajustaram modelos de regressões para as medidas de umidade e condutividade hidráulica. Os valores de alcance para os dados de CEa medida por EM38 foram 410, 377, e 359 m, respectivamente, para CEa medida nas alturas de 0-0,3; 0,3-0,6; e 0,6-0,9 m do solo. Ainda que as distâncias entre os pontos de CEa para Lopes e Montenegro tenha sido 10 m (medidas realizadas em grid 10 x 10 m), o fenômeno da condutividade elétrica apresentou valores similares para alcance quando comparados ao presente estudo (Tabela 3).

Akramkhanov, Brus e Walvoort (2014), trabalhando com quatro repetições de levantamento de dados de CEa obtidos por EM38 (anos de 2008, 2009, 2010 e 2011), em uma área de 80 ha na estação de pesquisa de algodão na província de *Khorem*, Uzbequistão, encontraram correlações log-log significativas (r^2 = 0,44) entre as medidas de sensor e avaliações sobre salinidade do solo realizadas por CE de laboratório obtidas em 142 localizações. Os valores de resíduos dos modelos de predição apresentaram valores de alcance de aproximadamente 120 m.

Para dimensionar o impacto do efeito pepita na representação final da semivariância, utiliza-se a razão efeito pepita / patamar, apresentada na Tabela 3. Para melhor compreensão e discussão, multiplicaram-se os valores dessa razão por 100 para que seja representada em porcentagem. A razão Efeito pepita/Patamar indica que o efeito pepita para o Nível 1 contribui com 3% na semivariância; enquanto, progressivamente crescente, no Nível 4 o efeito pepita representa 7% da semivariância.

Lopes e Montenegro (2019) encontraram menores valores Efeito pepita/Patamar, sendo 2,91, 1,41 e 1,17 %, respectivamente para as alturas de CEa do solo 0-0,3; 0,3-0,6; e 0,6-0,9 m, quando comparados ao presente estudo (Tabela 3).

Tabela 3. Parâmetros de ajuste dos semivariogramas, pelo modelo esférico, da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de espaçamento.

Nível de espaçamento	Efeito pepita	Patamar	Efeito pepita/Patamar (%)	Alcance (m)
1 – 26 linhas	9,0E-04	3,0E-02	3	533
2 – 13 linhas	1,4E-03	2,7E-02	5	530
3 – 7 linhas	1,6E-03	2,8E-02	5	489
4 – 4 linhas	1,6E-03	2,2E-02	7	475

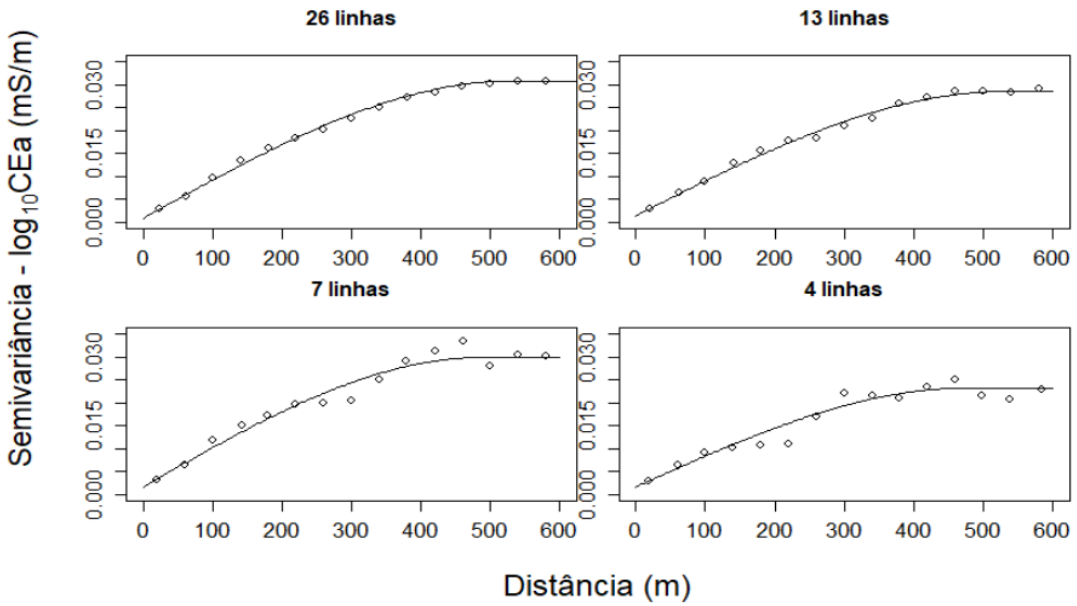

Figura 7. Semivariogramas observados (pontos) e ajustados (linhas), pelo modelo esférico, da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de espaçamento.

4.1.3. Krigagem ordinária

Os mapas interpolados utilizando krigagem ordinária a partir dos diferentes níveis de espaçamento e resoluções espaciais são apresentados nas Figuras 8 a 11. O mapa produzido a partir do Nível 1 (26 linhas) e com resolução de saída de 5 m (Figura 8A) será considerado como base para comparação das combinações entre diferentes níveis de espaçamento e resolução espacial. Percebe-se que somente a partir da resolução espacial de 30 m (Figura 9B) torna-se possível a interpretação e reconhecimento da variação espacial da CEa na área de estudo. Sendo assim, as resoluções mais "grosseiras" que 30 m (Figuras 10 e 11) foram interpretadas como não adequadas para produção de mapas de CEa, ainda mais quando se considera o tamanho da área de estudo de 72 ha.

Utilizando o Nível 1 e a resolução de 5 m (Figura 8A), é possível notar que os maiores valores de CEa encontram-se na parte sudoeste do mapa. Nessa região de maior concentração dos valores de CEa, o terreno forma um anfiteatro onde existe um canal de drenagem efêmero, logo, mesmo que a coleta dos dados de EM38-MK2 tenha sido realizada em um período seco (Julho), sem a presença do canal e não havendo irrigação ativa (pois era uma janela de plantio), esses maiores valores devem-se à maior concentração de umidade e,

possivelmente, argila, como é relatado nos trabalhos de McNeill (1980); Sudduth et al. (2005); Sudduth, Drummond e Kitchen (2001).

Por outro lado, os menores valores de CEa encontram-se na parte norte da área de estudo, sendo essa uma região com menor umidade do solo, já que drena para a parte mais baixa do terreno localizada na região sudoeste do mapa. Pode-se comparar esses resultados aos encontrados por Islam et al. (2011) e Triantafilis e Lesch (2005), em que as menores concentrações de argila e umidade estavam associadas a menores valores de CEa e vice-versa.

Avaliando a distribuição espacial da CEa do mapa produzido a partir do Nível 2 (13 linhas) com resolução espacial de 5 m (Figura 8A), percebe-se que na parte norte e sudeste do mapa houve ligeiras suavizações quanto aos padrões de CEa apresentados devido à redução do número de pontos e aumento do espaçamento entre linhas. Comparando-se esse mapa com o produzido a partir do Nível 1, com mesma resolução de 5 m, percebe-se pouca diferença quanto aos padrões de CEa majoritariamente exibidos para toda a área do pivô, o que é confirmado pelos mapas de diferenças, apresentados e discutidos posteriormente (Figuras 16 a 19). O mapa produzido a partir do Nível 3 combinado à resolução de 5 m apresentou variabilidade espacial da CEa similar ao mapa mais detalhado (Nível 1) utilizando a mesma resolução (5 m). No entanto, regiões com menores valores de CEa presentes no norte e no sudeste do mapa mais detalhado (Nível 1) produzido na resolução mais refinada (5 m – Figura 8A) foram divididas e suavizadas, respectivamente, norte e sudeste, quando observado o mapa de CEa que utilizou o Nível 3 com mesma resolução. O mapa utilizando os dados de CEa do Nível 3 e resolução 5 m (Figura 8A) ainda exibiram, nitidamente, o canal de drenagem efêmero localizado na parte sudoeste do mapa, em que se localizam os maiores valores de CEa.

O mapa de CEa produzido a partir do Nível 4 (4 linhas) e resolução de 5 m (Figura 8A) reconheceu a variabilidade da CEa na região, identificando a localização dos maiores e menores valores de CEa. Contudo, houve intensa generalização dessas regiões de variabilidade, somente sendo possível reconhecer tendências de crescimento e decrescimento dos valores de CEa. O canal de drenagem foi reconhecido pelo conjunto de dados de CEa do Nível 4, contudo, em virtude da maior distância entre as linhas de caminhamento, houve uma menor capacidade de caracterizar a região com o canal e com maiores valores de CEa (sudoeste) mostrando o canal efêmero com comprimento encurtado quando comparado aos Níveis 1, 2 e 3.

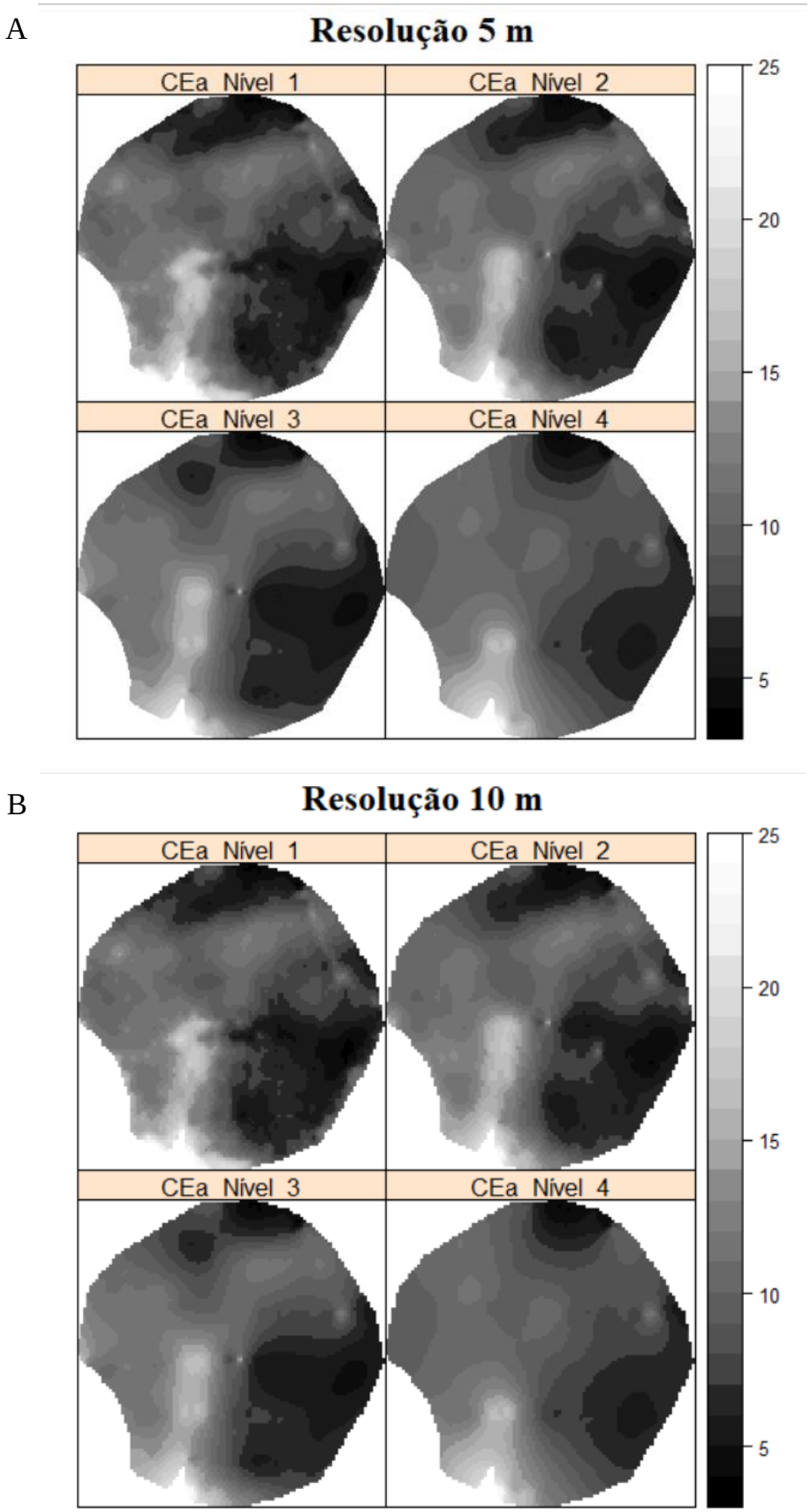

Figura 8. Mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m), resolução de 5 m (A) e 10 m (B).

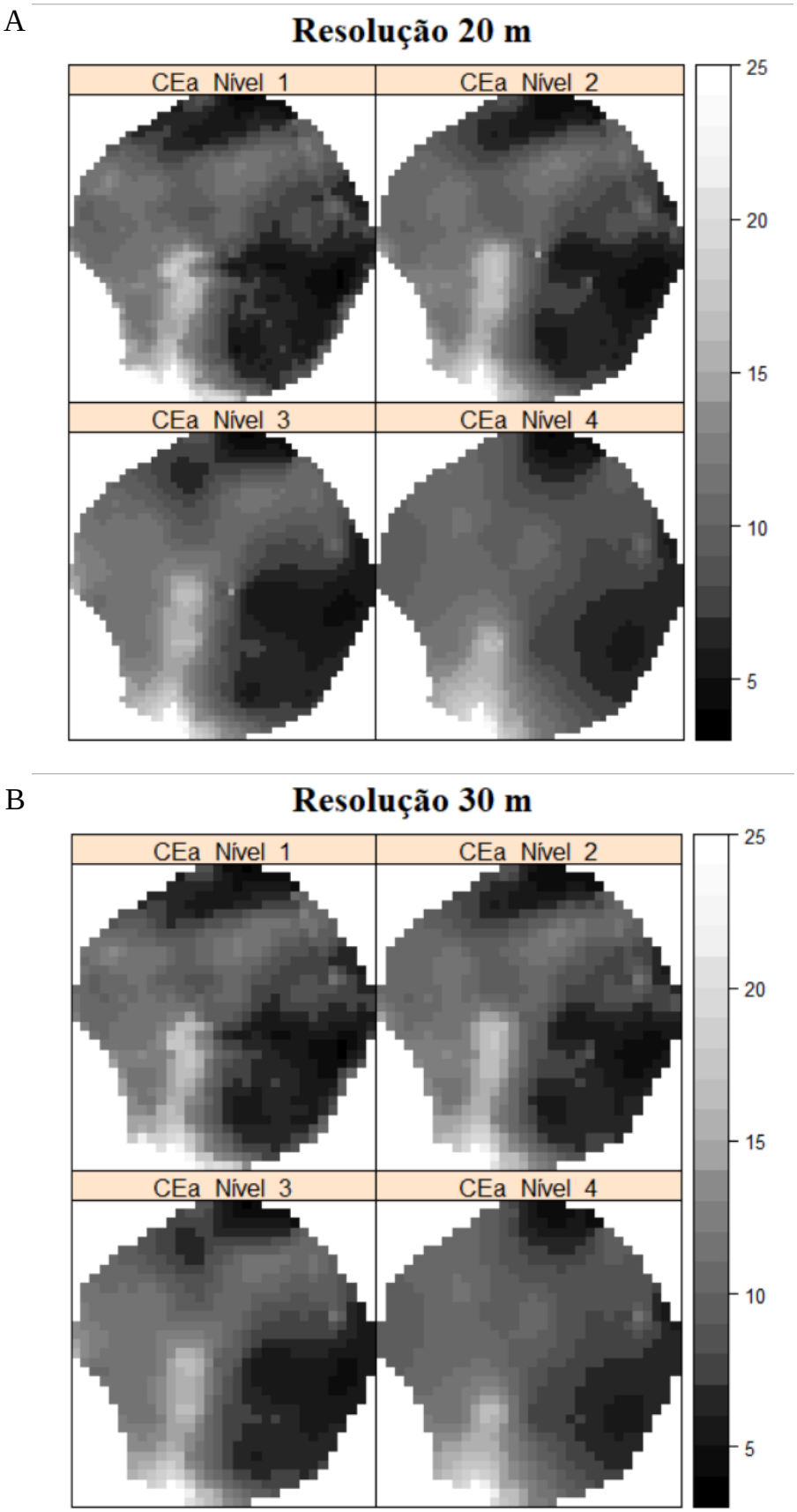

Figura 9. Mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m), resolução de 20 m (A) e 30 m (B).

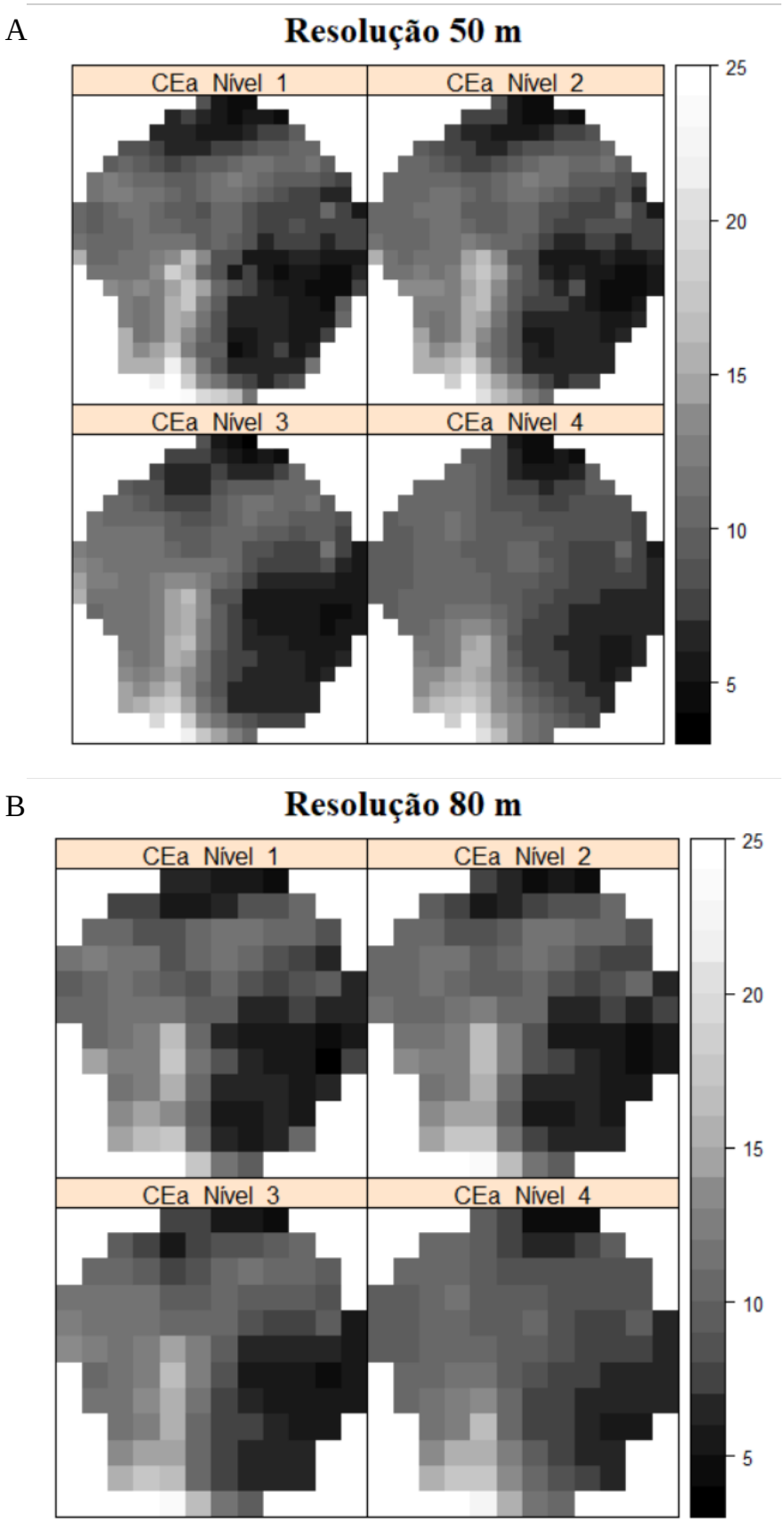

Figura 10. Mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m), resolução de 50 m (A) e 80 m (B).

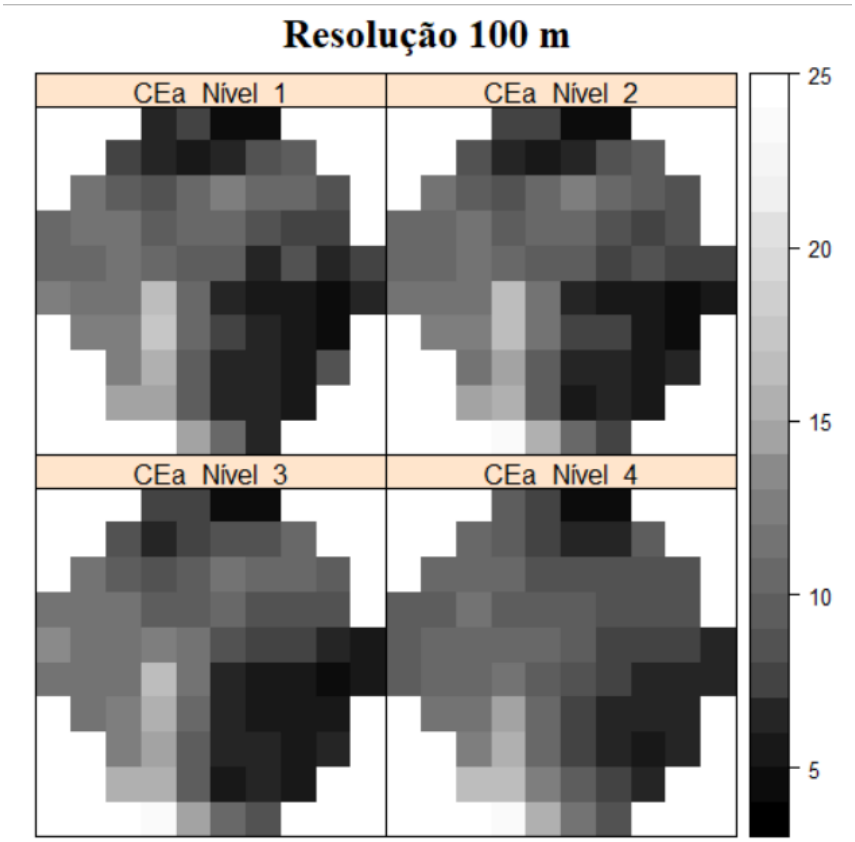

Figura 11. Mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m), resolução de 100 m.

De modo a complementar a análise visual apresentada nas Figuras 8 a 11, calculou-se o desvio padrão da krigagem de cada um dos 28 mapas produzidos (quatro níveis x sete resoluções). Os Níveis 1 e 2 apresentaram desvios padrões similares entre si e menores quando comparados aos desvios padrões exibidos para os mapas produzidos pelos Níveis 3 e 4, possivelmente em detrimento do maior distanciamento entra linhas de caminhamento conforme avança-se do Nível 1 ao Nível 4.

A partir da resolução espacial de 50 m (Figura 10A), é perceptível o decréscimo dos valores dos desvios padrões para os *pixels* próximos às coordenadas dos pontos de coleta de CEa para cada nível. É evidente que a maior quantidade de linhas produz menores desvios padrões, e, ao comparar a distribuição espacial dos desvios padrões (Figuras 12 a 14) com os valores apresentados por conjuntos de dados na Tabela 2 (Nível 1: 3,85 (mS/m); Nível 2: 3,48 (mS/m); Nível 3: 3,64 (mS/m); e Nível 4: 3,43 (mS/m)), percebe-se que, ainda que os valores de desvio padrão para o conjunto de dados sejam próximos, quando observados em formato espacializado, observa-se que há significativa variação, e essa é crescente a medida que ocorre o decréscimo na quantidade de linhas de caminhamento por conjunto de dados de CEa.

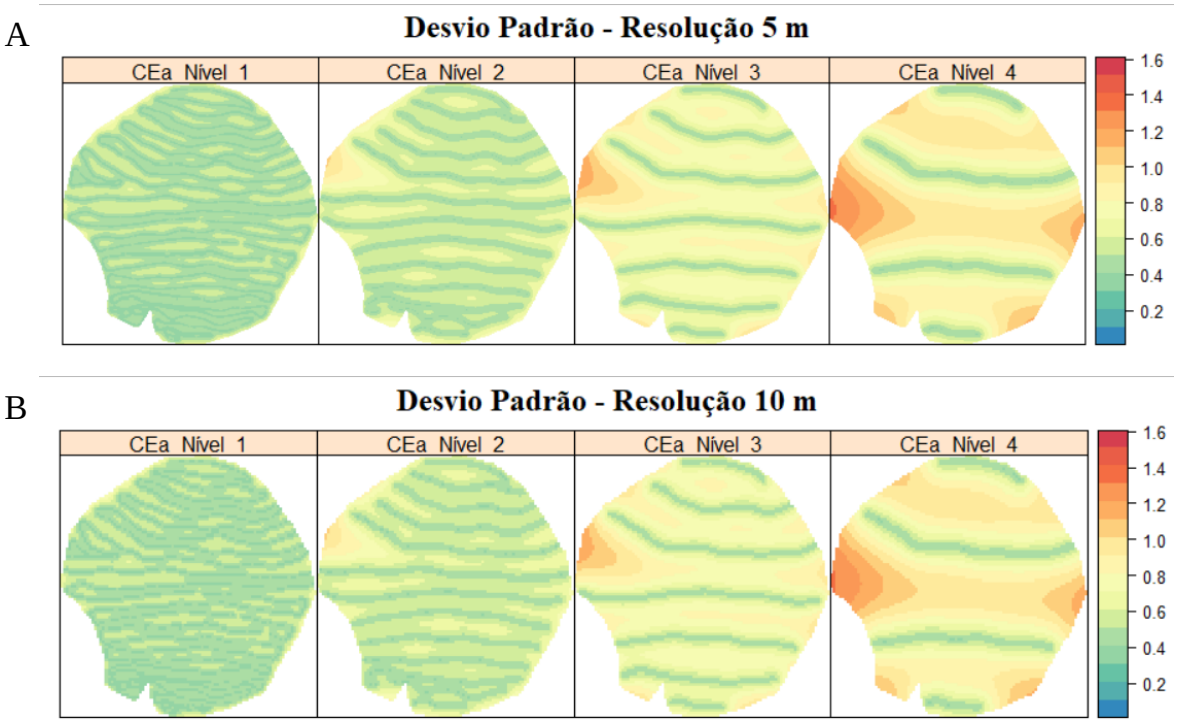

Figura 12. Desvio padrão da krigagem dos mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m) – Resolução de 5 m (A) e 10 m (B).

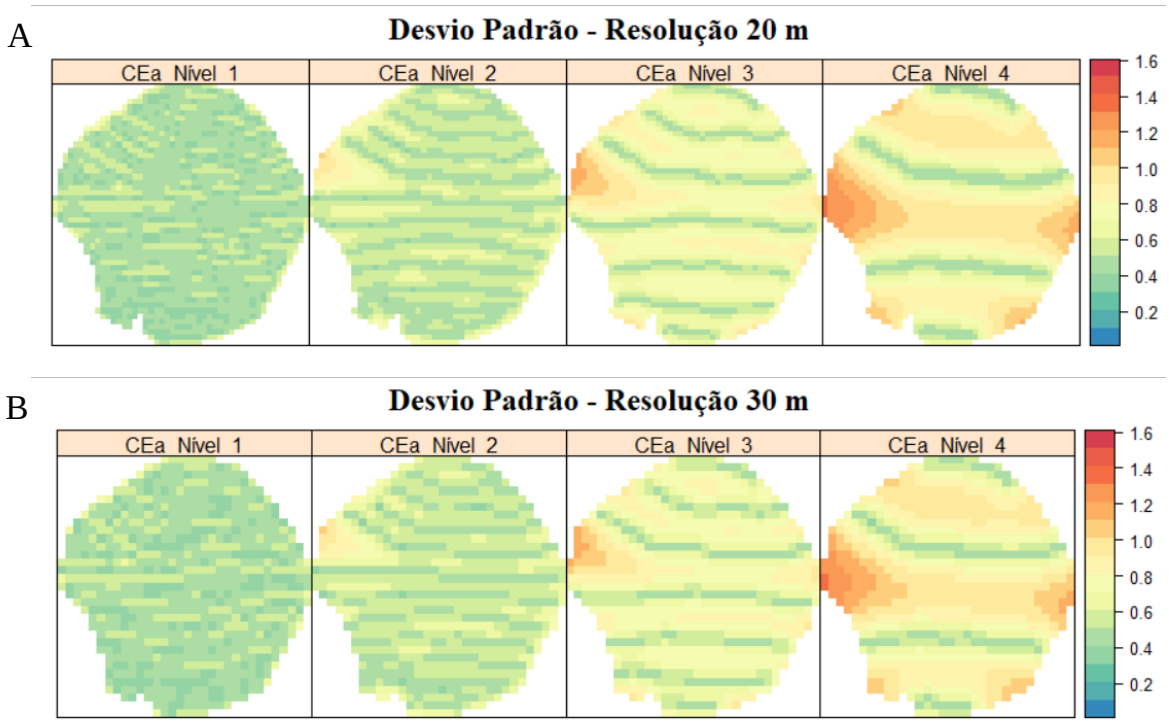

Figura 13. Desvio padrão da krigagem dos mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m) – Resolução de 20 m (A) e 30 m (B).

Figura 14. Desvio padrão da krigagem dos mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m) – Resolução de 50 m (A), 80 m (B) e 100 m (C).

4.1.4. Tempo de processamento e incerteza dos mapas a partir das combinações entre níveis de espaçamento entre linhas e resolução espacial

Os tempos de processamento para cada mapa são exibidos na Tabela 4 e Figura 15. Para a confecção do "melhor" mapa (Nível 1 e resolução de 5 m) o tempo necessário foi de 1512 s, cerca de 25 minutos. Esse tempo de processamento é muito alto, o que inviabiliza a adoção dessa combinação de espaçamento entre linhas e resolução espacial em sistemas automatizados de apoio à decisão e sistemas *on-the-go* conectados a implementos agrícolas. Portanto, uma combinação alternativa que atinja tempo de processamento menor do que 10 s, conforme estipulado por esse trabalho, seria o ideal. Esse tempo de processamento, aliado à boa qualidade visual e acurácia (REQM) do mapa, só foi atingido na Etapa 2, após remoção aleatória (raleamento) de parte dos dados, conforme apresentado posteriormente.

Embora o tempo necessário para a confecção dos mapas de CEa, utilizando todos os níveis de espaçamento entre linhas combinados às resoluções 100, 80, e 50 m, seja conveniente, uma vez que demandaram, no máximo, 70 s para a confecção do mapa de CEa, essas resoluções, mesmo no melhor nível de espaçamento (Nível 1), não permitem perceber

claramente os padrões espaciais da CEa e tornam esses mapas inviáveis para interpretação e utilização, ainda mais quando se considera o tamanho reduzido da área, 72 ha, que demanda mapas de maior detalhe para tomada de decisão.

Tabela 4. Índice de erro de validação externa (raiz do erro quadrado médio; REQM) e tempo médio de processamento dos mapas interpolados da condutividade elétrica aparente do solo (CEa, em mS/m) a partir dos diferentes níveis de espaçamento.

Resolução espacial	Nível de espaçamento entre linhas	REQM (mS/m)	Tempo médio de processamento (s)
5 m	1 – 40 m	0,67	1512,4
	2 – 80 m	1,10	293,0
	3 – 150 m	1,29	83,6
	4 – 300 m	1,93	18,2
10 m	1 – 40 m	0,69	436,2
	2 – 80 m	1,09	77,9
	3 – 150 m	1,29	20,6
	4 – 300 m	1,93	4,7
20 m	1 – 40 m	0,81	184,8
	2 – 80 m	1,15	22,2
	3 – 150 m	1,32	5,5
	4 – 300 m	1,95	1,7
30 m	1 – 40 m	0,96	104,0
	2 – 80 m	1,16	13,1
	3 – 150 m	1,31	2,9
	4 – 300 m	1,93	0,7
50 m	1 – 40 m	1,28	77,4
	2 – 80 m	1,29	8,3
	3 – 150 m	1,38	1,8
	4 – 300 m	1,93	0,4
80 m	1 – 40 m	1,72	73,5
	2 – 80 m	1,67	6,6
	3 – 150 m	1,80	1,1
	4 – 300 m	2,17	0,3
100 m	1 – 40 m	1,87	69,0
	2 – 80 m	1,73	6,2
	3 – 150 m	1,79	1,0
	4 – 300 m	2,17	0,2

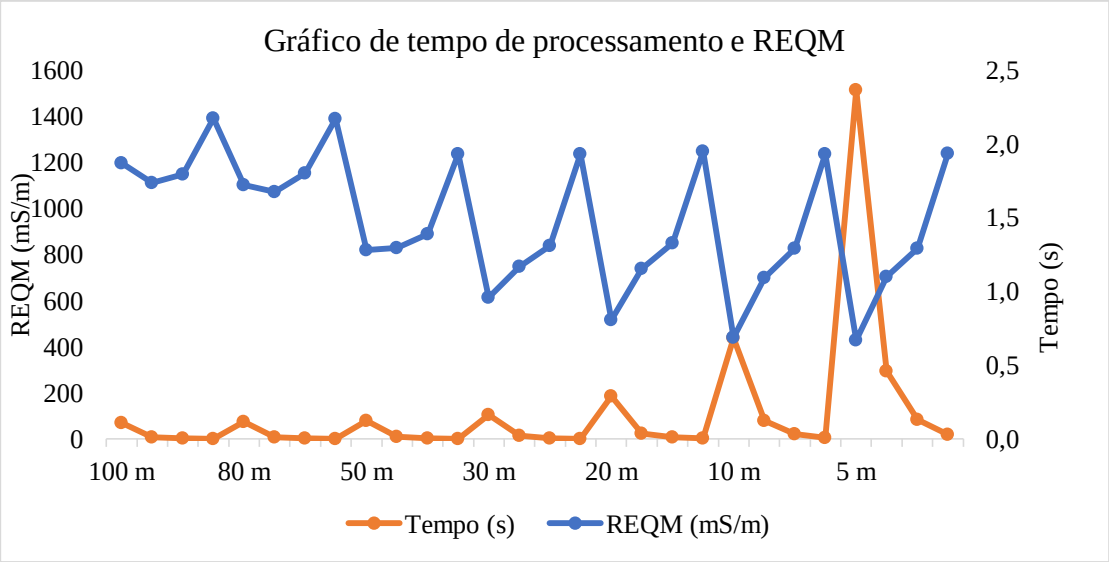

Figura 15. Gráfico do índice da raiz do erro quadrado médio (REQM) e tempo de processamento para os dados de condutividade elétrica aparente do solo (CEa, em mS/m) a partir dos diferentes níveis de espaçamentos agrupados por resolução espacial e nível de espaçamento.

Por outro lado, embora a resolução 5 m seja refinada e adequada para interpretação visual dos mapas, para atividades agrícolas essa pode ser incompatível com as diversas atividades aplicadas em agricultura de precisão. Por exemplo, mesmo que exista na fazenda um aplicador de insumo a taxas variadas, um mapa com 5 m de resolução não permitiria aplicação a taxas variadas a cada 5 m caso os implementos não possuam tal precisão (5 m), devido, por exemplo, a: erros no posicionamento do GPS; larguras de implementos e eixos de manobras de tratores que podem ultrapassar os 5 m; e atrasos no tempo de resposta do implemento agrícola.

O valor "ótimo" de REQM de validação utilizado para comparações entre os mapas de CEa produzidos a partir de diferentes níveis e resoluções é 0,67 mS/m, encontrado no mapa de CEa construído a partir do Nível 1 e resolução espacial de 5 m (Tabela 4 e Figura 15).

Como se desconsideram as resoluções espaciais de 100, 80 e 50 m, elas não serão extensamente discutidas e podem ser conferidas na Tabela 4. É esperado que os menores valores de REQM encontrados para os mapas de CEa para as resoluções de 30, 20 e 10 m são os que utilizaram o Nível 1 (Tabela 4).

As resoluções de 30, 20 e 10 m combinadas aos Níveis 1 ao 4 apresentaram valores de REQM progressivamente crescentes (ideal é que esses valores sejam reduzidos e os "melhores" são aqueles próximos ao valor 0) demonstrando a influência negativa da redução da quantidade de linhas em relação aos valores de incerteza dos mapas de CEa.

4.1.5. Diferenças espaciais entre valores estimados de CEa do solo

Para avaliar as diferenças entre os valores de CEa estimados pela krigagem, subtraíram-se os valores de CEa apresentados nos mapas produzidos pelos Níveis 2, 3 e 4 do mapa de CEa construído a partir das 26 linhas (Nível 1) na mesma resolução espacial, respectivamente.

As resoluções de 100, 80 e 50 m apresentam pouca diferença visual entre os valores de CEa mapeados, sendo esses muito próximos entre si. Ainda, quando subtraídos os mapas produzidos utilizando poucas linhas de caminhamento dos mapas utilizando muitas (por

exemplo: Nível 1 com 26 linhas – Nível 4 com 4 linhas), não foi possível visualizar o impacto (ou influência) do decréscimo do número de linhas de CEa na confecção dos mapas, possivelmente devido ao tamanho da área de estudo (72 ha) não ser compatível com as resoluções 100, 80 e 50 m para apresentar um mapa sem um aspecto *"pixelado*".

Nos mapas de diferenças confeccionados nas resoluções de 30 a 5 m, percebe-se nas bordas a ocorrência das maiores diferenças entre os valores, utilizando-se conjuntos com diferentes níveis de espaçamento. Provavelmente, isso decorre devido à presença de menos amostras vizinhas para interpolação nas bordas, podendo ser inclusive um efeito de extrapolação das estimativas que, nas Figuras 16 a 19, apresenta-se pelas cores vermelhas (diferenças positivas) e azuis (diferenças negativas).

Ao subtrair-se o mapa confeccionado a partir do Nível 4 (4 linhas) do "melhor" mapa (Nível 1 – 26 linhas), ocorrem maiores diferenças entre os valores estimados de CEa. O mapa com maior diferença entre valores de CEa estimada é o confeccionado a partir do Nível 4 (4 linhas) para as resoluções de 30 a 5 m. Podem-se notar as intensas diferenças positivas (coloridas em vermelho – Figuras 16 a 19) e negativas (coloridas em azul – Figura 16 a 19) em maior ocorrência e intensidade, quando comparadas aos mapas de diferenças entre o "melhor" mapa e os dos Níveis 2 e 3, respectivamente.

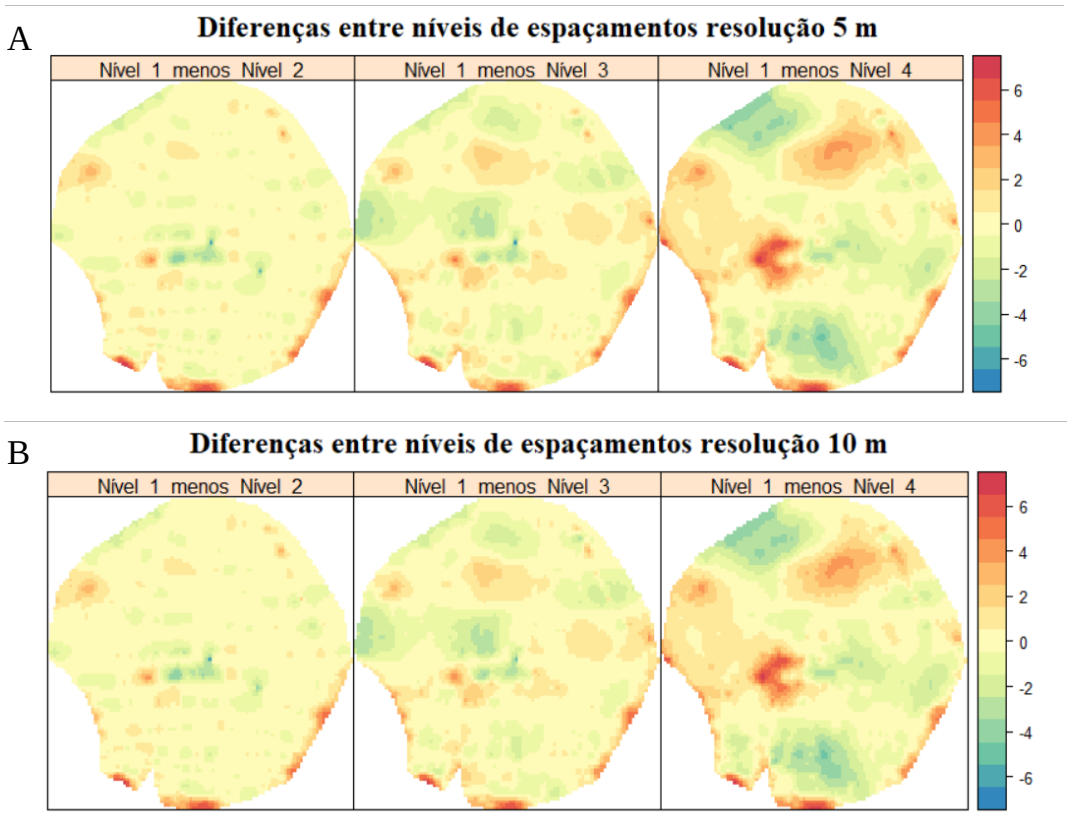

Figura 16. Diferenças entre os valores dos mapas dos dados de condutividade elétrica aparente do solo (CEa, em mS/m) – Resolução de 5 m (A) e 10 m (B). Da esquerda para direita: Nível 1 menos Nível 2; Nível 1 menos Nível 3; Nível 1 menos Nível 4.

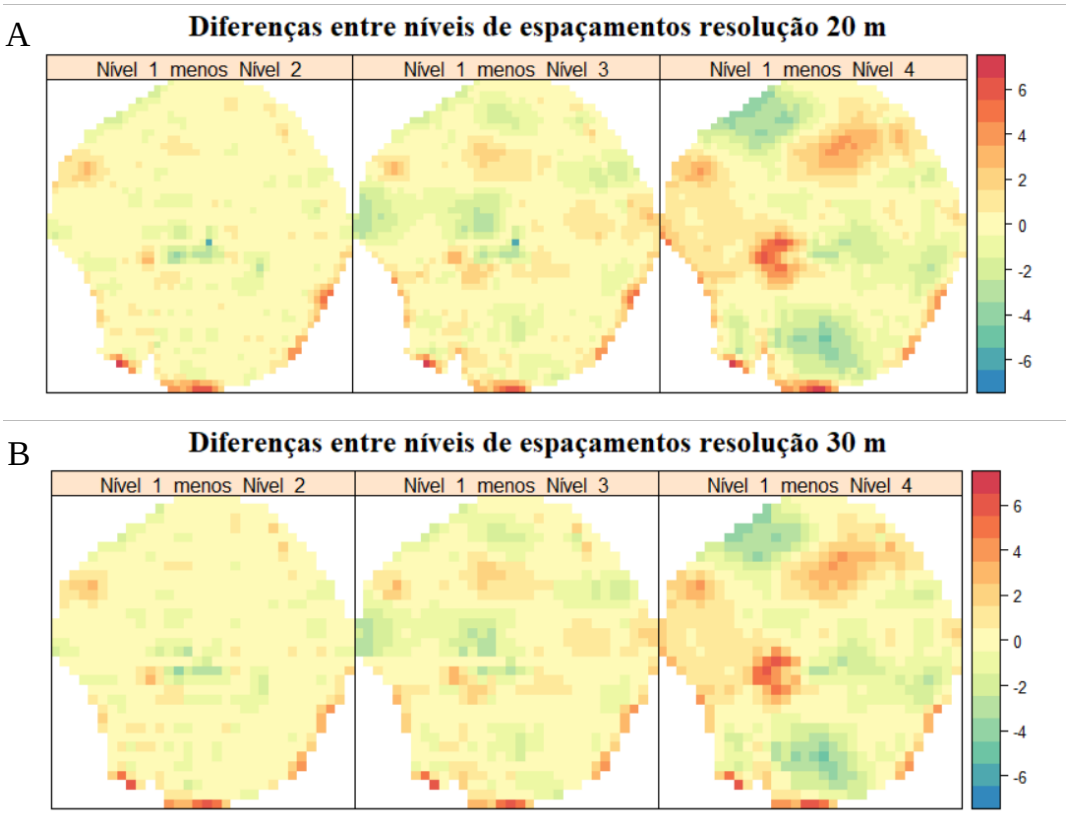

Figura 17. Diferenças entre os valores dos mapas dos dados de condutividade elétrica aparente do solo (CEa, em mS/m) – Resolução de 20 m (A) e 30 m (B). Da esquerda para direita: Nível 1 menos Nível 2; Nível 1 menos Nível 3; Nível 1 menos Nível 4.

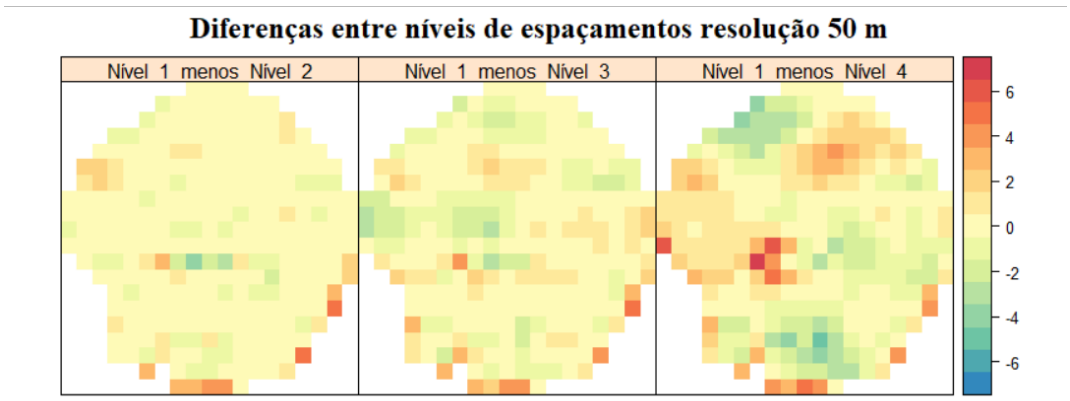

Figura 18. Diferenças entre os valores dos mapas dos dados de condutividade elétrica aparente do solo (CEa, em mS/m) – Resolução de 50 m. Da esquerda para direita: Nível 1 menos Nível 2; Nível 1 menos Nível 3; Nível 1 menos Nível 4.

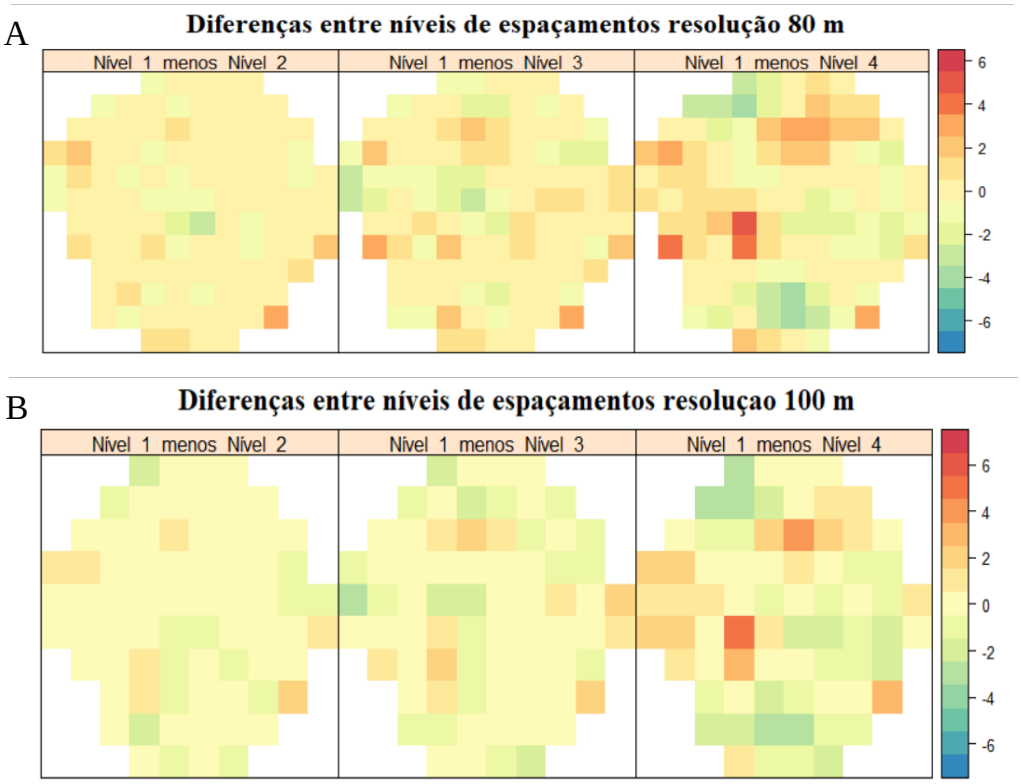

Figura 19. Diferenças entre os valores dos mapas dos dados de condutividade elétrica aparente do solo (CEa, em mS/m) – Resolução de 80 m (A) e 100 m (B). Da esquerda para direita: Nível 1 menos Nível 2; Nível 1 menos Nível 3; Nível 1 menos Nível 4.

4.1.6. Estimativa de tempo de percurso por nível de espaçamento

A distância percorrida calculada foi 19.833 m. Sendo assim, a velocidade média de percurso estimada foi: 19.833 m / 5196 s = 3,8 m/s (13,7 km/h). Note que essa velocidade média desconsidera as manobras realizadas entre as linhas de caminhamento, sendo, portanto, o tempo gasto efetivamente para realizar as leituras de CEa.

Em seguida, para a estimativa do tempo de percurso necessário para realizar as leituras em cada nível de espaçamento, dividiu-se a distância total percorrida em cada nível pela velocidade estimada de 3,8 m/s, sendo os tempos de percurso estimados mostrados, em minutos, na Tabela 5. O tempo de percurso variou de cerca de 19 minutos no Nível 4 até 87 minutos no Nível 1.

Tabela 5. Estimativa de tempo para coleta de dados de condutividade elétrica aparente do solo (CEa, em mS/m) a partir dos diferentes níveis de espaçamentos.

Nível de espaçamento	Distância percorrida (m)	Tempo de percurso (minutos)
1 – 26 linhas	19833	87
2 – 13 linhas	16378	71
3 – 7 linhas	8407	37
4 – 4 linhas	4248	19

4.1.7. Critérios para decisão de nível de espaçamento e resolução espacial para avanço à etapa 2

- O aspecto visual dos mapas de CEa produzidos a partir das resoluções de 100, 80 e 50 m demonstraram que essas são inadequadas para o estudo, bem como para continuarem à Etapa 2 da pesquisa;
- O Nível 4 de espaçamento suaviza em excesso os padrões de variação da CEa e não delimita corretamente a área mais baixa do terreno onde ocorre um canal de drenagem efêmero, sendo, portanto, inadequado;
- Por outro lado, a resolução de 5 m demonstra-se inadequada por ser excessivamente "*refinada*" às atividades práticas para as quais o presente estudo está direcionando esforços, como atividades envolvendo tratores, GPS e técnicas de Agricultura de Precisão;
- Para os valores de tempo de processamento para interpolação dos dados e confecção do mapa final, os Níveis 2 e 3, associados às resoluções de 30, 20 e 10 m demonstraram-se combinações adequadas; enquanto o processamento a partir do Nível 1 utilizando as resoluções 30, 20 e 10 m demonstrou-se excessivamente demorado, sendo, portanto, desconsiderado para a próxima etapa;
- Os valores dos desvios padrões da krigagem permitiram reconhecer que somente o Nível 2 possui uma distância entre linhas de caminhamento que mantêm o desvio padrão relativamente baixo, quando comparados aos mapas de desvio padrão que utilizam o subsequente e mais espaçado Nível 3;
- A qualidade visual e a acurácia dos mapas produzidos a partir do Nível 2 são superiores quando utilizada a resolução de 10 m àquelas dos mapas nas resoluções de 20 e 30 m.

4.2. Etapa 2 - Avaliação do Adensamento Amostral

4.2.1. Estatística descritiva

Uma vez estipulada a combinação entre o Nível 2, que contém 13 linhas espaçadas a 80 m, com a resolução espacial de 10 m para a confecção dos mapas, removeram-se aleatoriamente 25, 50, 75 e 95% dos dados do Nível 2. A partir de agora, os Níveis não serão mais referenciados às variações nos espaçamentos entre linhas, mas sim às quantidades de pontos removidos, sendo, portanto, considerados *níveis de raleamento*. O Nível 1, com 25% dos pontos removidos, ficou com 1536 pontos; o Nível 2, com 50% dos pontos removidos, ficou com 1024 pontos; o Nível 3, com 75% dos pontos removidos, ficou com 512 pontos; e o Nível 4, com 95% dos pontos removidos, ficou com 103 pontos.

As estatísticas descritivas dos novos conjuntos de dados são apresentadas na Tabela 6. Os valores mínimos e máximos para os Níveis 1, 2 e 3 de raleamento foram 3,4 e 28,2 mS/m, respectivamente; enquanto para o Nível 4 o mínimo foi 4,18 e o máximo 19,5 mS/m. O valor do desvio padrão para o grupo de dados de CEa, contemplando-se 100% dos pontos, foi 3,48 mS/m (Tabela 6); enquanto para todos os níveis de raleamento subsequentes, os valores mantiveram-se próximos (Tabela 6).

De maneira similar às análises estatísticas realizadas na Etapa 1, realizou-se transformação para logaritmo na base 10 após avaliação quanto à distribuição de frequência dos dados de CEa (Figura 20).

Tabela 6. Estatística descritiva da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de raleamento.

Estatística	CEa (mS/m)				
	Nível 1 25% removidos	**Nível 2 50% removidos**	**Nível 3 75% removidos**	**Nível 4 95% removidos**	**Nível 0 (Sem ralear)**
Observações	1536	1024	512	103	2047
Mínimo	3,40	3,40	3,40	4,18	3,28
Máximo	28,24	28,24	28,24	19,5	31,48
Média	9,77	9,82	9,74	9,94	9,73
Mediana	9,64	9,84	9,77	10,00	9,61
Desvio Padrão	3,50	3,42	3,46	3,37	3,48
Assimetria	0,76	0,61	0,81	0,37	0,82
Curtose	1,10	0,58	1,51	-0,51	1,43

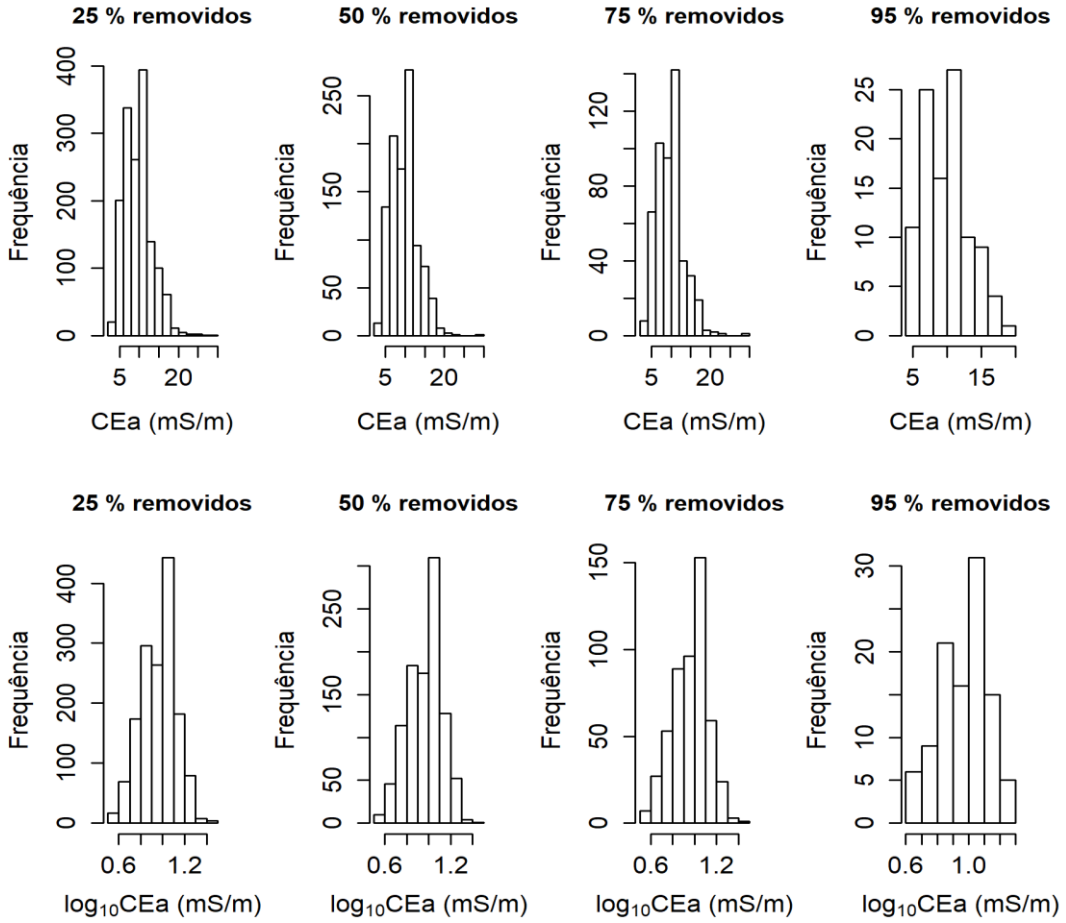

Figura 20. Histogramas da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de raleamento: a) Distribuição de frequência ; b) Distribuição de frequência após a transformação log na base 10.

4.2.2. Semivariogramas

Embora o Nível 1 de raleamento possua o maior adensamento e número de pontos, ele não foi o que apresentou o menor efeito pepita (Tabela 7; Figura 21). O Nível 4 teve efeito pepita 0 e os outros níveis efeito pepita muito próximo de 0, sendo a diferença entre eles

visível somente ao se comparar as razões efeito pepita/patamar (Tabela 7). Ainda que baixos, o valor dos efeitos pepita aumentaram do Nível 1 até o Nível 3.

A razão Efeito pepita/Patamar indica que o efeito pepita para o Nível 1 representa 4% da semivariância, enquanto essa razão é progressivamente crescente até o Nível 3, chegando a 9%. No Nível 4, por ter efeito pepita 0, a razão foi 0 também, ou seja, 100% da semivariância é devido à dependência espacial estruturada.

Assim como o discutido na seção 3.1.2 e apresentado na Tabela 3, o alcance dos semivariogramas para os dados de CEa para os diferentes níveis de raleamento (Tabela 7) apresentaram valores próximos à 530 m, sendo esse o valor apresentado pelo Nível 0 de raleamento (Nível 2 de espaçamento da Etapa 1 referente ao conjunto de dados com 13 linhas). Dessa maneira, observa-se que embora reduzam-se as quantidades de pontos à medida que se avança do Nível 1 ao 4 de raleamento, os valores dos alcances mantêm-se próximos ao Nível 0 de raleamento.

Tabela 7. Parâmetros de ajuste dos semivariogramas dos dados de condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de raleamento utilizando modelo esférico.

Nível de raleamento	Efeito pepita	Patamar	Efeito pepita/Patamar (%)	Alcance (m)
1 – 25% removidos	1,2E-03	2,9E-02	4	527
2 – 50% removidos	1,4E-03	2,8E-02	5	512
3 – 75% removidos	2,4E-03	2,7E-02	9	476
4 – 95% removidos	0,0E+00	2,7E-02	0	550

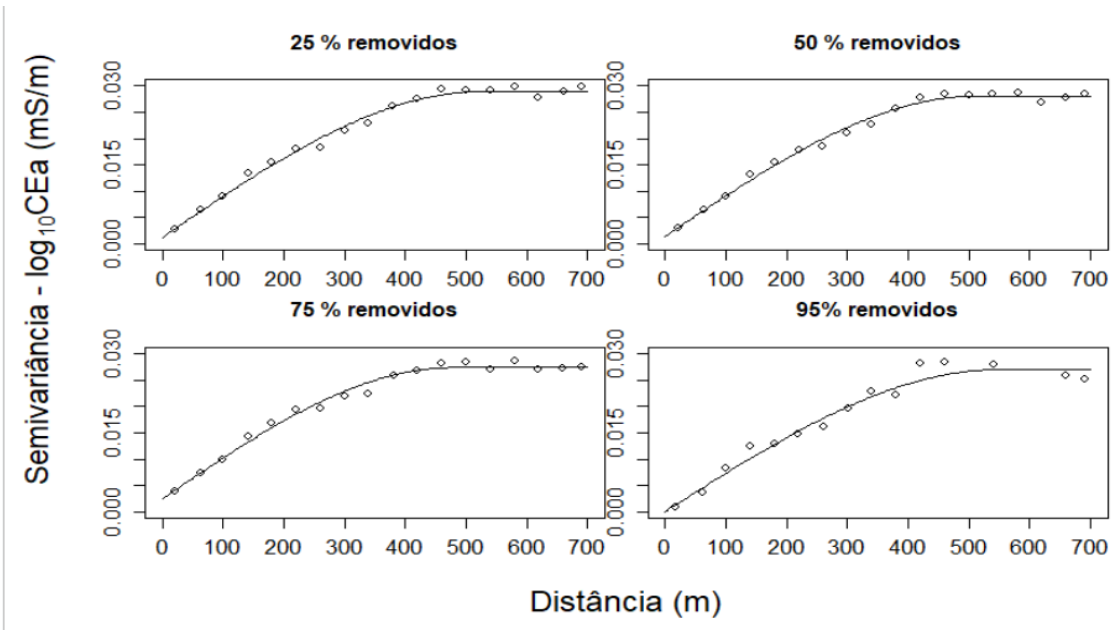

Figura 21. Semivariogramas observados (pontos) e ajustados (linhas), pelo modelo esférico, da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de raleamento.

4.2.3. Krigagem ordinária

Os mapas interpolados que utilizam a krigagem ordinária a partir dos diferentes níveis de raleamento e com resolução espacial de 10 m são apresentados na Figura 22. A partir da avaliação visual dos mapas produzidos pelos quatro níveis de raleamento de pontos, percebe-se que todos representaram as principais variações de CEa na área de estudo.

Em todos os mapas da Figura 22 foi possível identificar a região do canal de drenagem efêmero na parte sudoeste da área de estudo, onde existem os maiores valores de CEa. Em contrapartida, as áreas com menores concentrações de CEa, localizadas ao norte e à sudeste na área de estudo, mantiveram-se claramente delimitadas, salvo pelo Nível 4 que suavizou e generalizou ligeiramente os valores de CEa na região sudeste do mapa.

Ao avaliar os desvios padrões da krigagem na Figura 23, é possível perceber que o Nível 2 consegue manter valores de desvios padrões próximos ao valor do conjunto de dados com 100% das observações (Figura 9; Nível 2 de espaçamento na resolução espacial de 10 m), embora desvios maiores ocorram na maior parte da área de estudo no Nível 4 de raleamento, exceto ao redor dos 103 pontos amostrais.

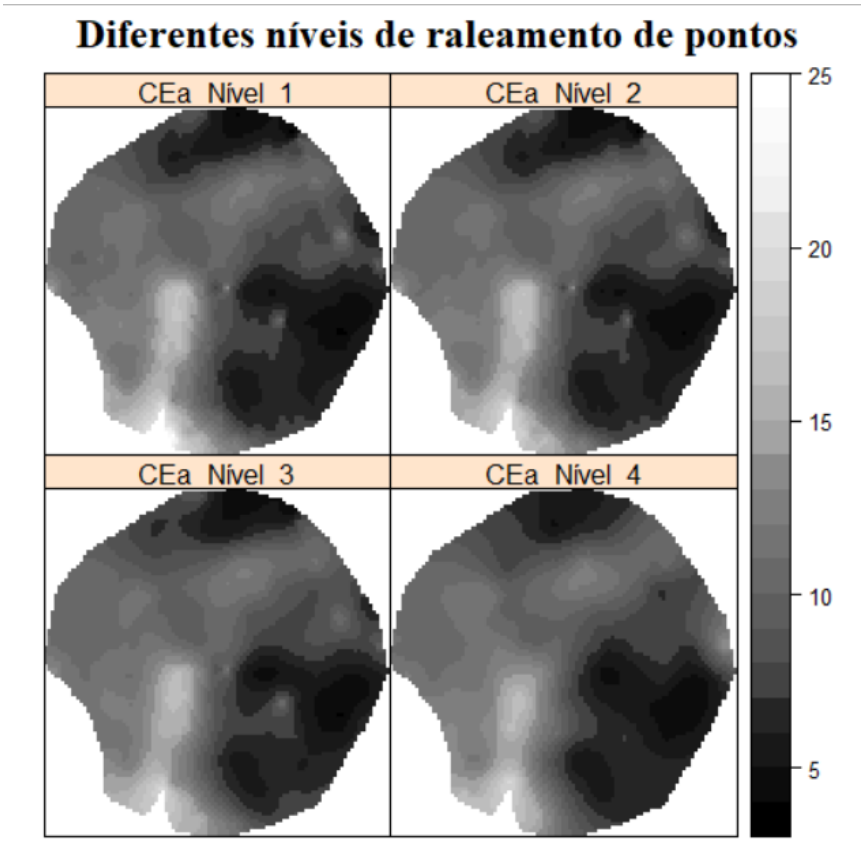

Figura 22. Mapas da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos quatro níveis de raleamento utilizando resolução de 10 m.

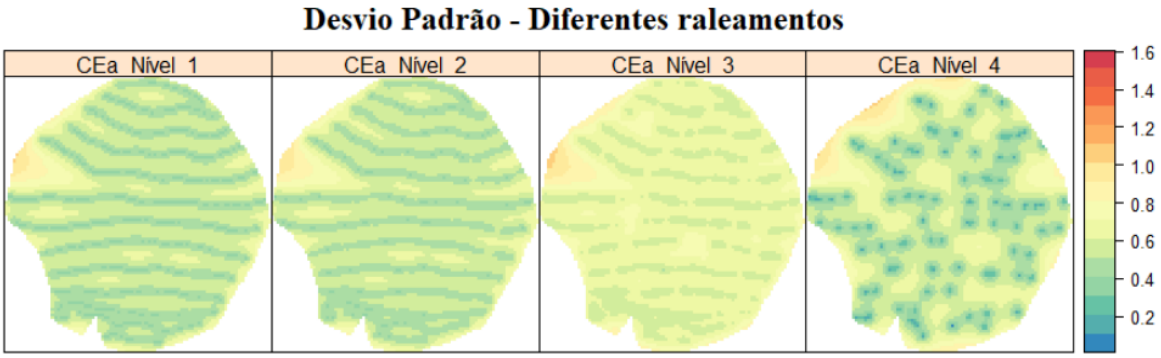

Figura 23. Desvio padrão da krigagem da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de raleamento utilizando resolução de 10 m.

O menor valor de REQM de validação externa (525 pontos) foi obtido para o Nível 1, com 25% os pontos removidos, se assemelhando ao valor de REQM obtido quando utilizando 100% dos pontos (Tabela 4; Nível 2 de espaçamento na resolução espacial de 10 m). O REQM obtido nos Níveis 2 e 3 de raleamento, com respectivamente 50 e 75% dos pontos removidos, foi de 1,21 mS/m, enquanto para o Nível 4, com remoção de 95% dos pontos, a REQM foi 1,54 mS/m.

O tempo de processamento necessário para a confecção dos mapas diminuiu conforme o número de amostras reduziu-se, embora não tenha ultrapassado 45 s para a confecção do mapa com 1536 pontos (Nível 1) (Tabela 8). O Nível 3, com somente 512 pontos, precisou de 4,9 s para produzir um mapa com REQM similar ao mapa produzido por 1024 pontos (Nível 2), que levou 20 s para ser produzido. Na Figura 24, é possível observar o decréscimo do tempo de processamento à medida que se reduz o número de pontos, enquanto os índices de erro REQM são inversamente proporcionais ao tempo, já que aumentam à medida que os dados são raleados.

Tabela 8. Índices de erro (raiz do erro quadrado médio; REQM) da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de raleamento utilizando resolução de 10 m.

Nível de raleamento	REQM	Tempo médio de processamento (s)
1 - 25% removidos	1,09	44,2
2 - 50% removidos	1,21	20,3
3 - 75% removidos	1,21	4,9
4 - 95% removidos	1,54	0,4

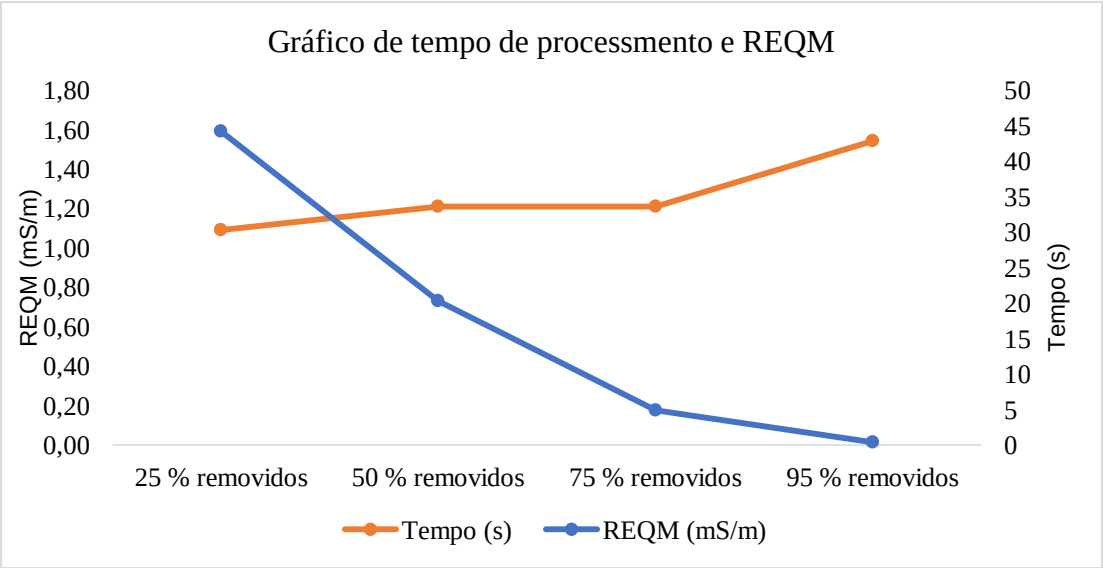

Figura 24. Gráfico da REQM e tempo de processamento da condutividade elétrica aparente do solo (CEa, em mS/m) no formato log10 a partir dos diferentes níveis de raleamento utilizando resolução de 10 m.

4.2.4. Seleção do nível de espaçamento, resolução espacial e nível de raleamento para produção do mapa final de CEa

As opções escolhidas para o mapeamento da CEa do solo na área de estudo, após a comparação de diferentes combinações de níveis de espaçamento e resolução espacial (Etapa 1), e de diferentes níveis de raleamento de pontos (Etapa 2), considerando a meta de reduzir o tempo de processamento para menos de 10 s, foram as seguintes:

- Para nível de espaçamento, elegeu-se o Nível 2 de espaçamento entre linhas de caminhamento (13 linhas, 2258 pontos), com aproximadamente 80 m de distância entre linhas;
- A resolução espacial de saída do mapa selecionada foi 10 m;
- Para o raleamento dos pontos e diminuição do tamanho amostral, com o interesse de reduzir o tempo de processamento para menos de 10 s, elegeu-se o Nível 3 de raleamento (remoção de 75% dos pontos) a partir do Nível 2 de espaçamento, resultando em 512 pontos;
- O valor de 4,9 s foi considerado um tempo adequado para o processamento do mapa de CEa dentre as combinações de níveis de raleamento e resolução de 10 m testadas, estando essa abaixo do limite de 10 s anteriormente estabelecido como alvo.

5. CONCLUSÕES

Após a análise dos dados desse estudo, as seguintes conclusões podem ser apresentadas:

a) Os levantamentos de dados de CEa em campo usando o sensor Geonics EM38-MK2 não devem ultrapassar 80 m de distância entre cada linha coletada, quando em caminhamento livre, simulando um *zigue-zague*;

b) A variação das resoluções espaciais dos mapas de saída demonstrou que para uma região com 72 ha é possível utilizar resoluções com no máximo 30 m sem que ocorra grande perda de compreensão visual;

c) A resolução de 5 m demonstrou-se excessiva para o objetivo de mapeamento de dados de CEa; além dessa ter demandado maior tempo para processamento dos mapas, a acurácia não foi significativamente maior do que as obtidas com resoluções mais grosseiras, como 10 e 20 m;

d) Para redução do tempo de processamento, mantendo-se as incertezas dos mapas, desde que se percorra toda a extensão da área com o sensor, foi possível remover em até 75 % dos dados aleatoriamente sem que houvesse diminuição da acurácia e do aspecto visual dos mapas.

6. CONSIDERAÇÕES FINAIS

A redução do tempo de processamento dos mapas permite que o pesquisador possa avançar sobre outros problemas envolvendo mapeamento. Ainda, as combinações de espaçamento entre linhas de amostragem, raleamento aleatório de pontos amostrais e resolução espacial do mapa de saída identificadas nesse trabalho norteiam futuros trabalhos e protocolos para coleta e processamento de dados de CEa para fornecer informação rápida e de qualidade para embasar decisões em escritório e no campo, por exemplo em operações on-the-go de tratores.

7. REFERÊNCIAS BIBLIOGRÁFICAS

ADAMCHUK, V.; JI, W.; VISCARRA ROSSEL, R.; GEBBERS, R.; TREMBLAY, N. Proximal Soil and Plant Sensing. In: SHANNON, D. K.; CLAY, D. E.; KITCHEN, N. R. (Eds.). **Precision Agriculture Basics**. Hoboken: John Wiley & Sons, 2018. p. 119-140.

AKRAMKHANOV, A.; BRUS, D. J.; WALVOORT, D. J. J. Geostatistical monitoring of soil salinity in Uzbekistan by repeated EMI surveys. **Geoderma**, v. 21, p. 600-607, 2014.

AKRAMKHANOV, A.; SOMMER, R.; MARTIUS, C.; HENDRICKX, J. M. H.; VLEK, P. L. G. Comparison and sensitivity of measurement techniques for spatial distribution of soil salinity. **Irrigation & Drainage Systems Engineering**, v. 22, p. 115-126, 2008.

ALLRED, B. J.; ADAMCHUK, V. I.; ROSSEL, VISCARRA, R. A.; DOOLITTLE, J.; FREELAND, R. S.; GROTE, K. R.; CORWIN, D. L. Geophysical Methods. In: LAL, R. (Eds.). **Encyclopedia of Soil Science**. 3. ed. Boca Raton: Taylor & Francis, 2016.

ALMASI, A.; JALALIAN, A.; TOOMANIAN, N. Using OK and IDW methods for prediction the spatial variability of A horizon depth and OM in soils of Shahrekord, Iran. **Journal of Environment and Earth Science**, v. 4, p. 17-28, 2014.

ANDERSON-COOK, C. M.; ALLEY, M. M.; ROYGARD, J. K. F. F.; KHOSLA, R.; NOBLE, R. B.; DOOLITTLE, J. A. Differentiating soil types using electromagnetic conductivity and crop yield maps. **Soil Science Society of America Journal**, v. 66, p. 1562-1570, 2002.

BECEGATO, V. A.; FERREIRA, F. J. F. Gamaespectrometria, resistividade elétrica e susceptibilidade magnética de solos agrícolas no noroeste do estado do Paraná. **Revista Brasileira de Geofísica**, v. 23, p. 371-405, 2005.

BENNETT, D. L.; GEORGE, R.; RYDER, A. **Soil Salinity Assessment Using the EM38: Field Operating Instructions and Data Interpretation.** Report 4/95. Perth: Department of Agriculture and Food, Western Australia, 1995.

BISWAS, A.; ZHANG, Y. Sampling designs for validating digital soil maps: A review. **Pedosphere**, v. 28, p. 1-15, 2018.

BREVIK, E. C.; CALZOLARI, C.; MILLER, B. A.; PEREIRA, P.; KABALA, C.; BAUMGARTEN, A.; JORDÁN, A. Soil mapping, classification, and pedologic modeling: History and future directions. **Geoderma**, v. 264, p. 256-274, 2016.

BRUS, D. J. Sampling for digital soil mapping: A tutorial supported by R scripts. **Geoderma**, v. 338, p. 464-480, 2018.

BURGESS, T. M.; WEBSTER, R. Optimal interpolation and isarithmic mapping of soil properties. **Journal of Soil Science**, v. 31, p. 333–341, 1980.

CAMERON, D. R.; READ, D. W. L.; JONG, E. DE; OOSTERVELD, M. Mapping Salinity Using Resistivity and Electromagnetic Inductive Techniques. **Canadian Journal of Soil Science**, v. 61, n. 1, p. 67-78, 1981.

CEDDIA, M. B.; VILLELA, A. L. O.; PINHEIRO, É. F. M.; WENDROTH, O. Spatial variability of soil carbon stock in the Urucu river basin, Central Amazon-Brazil. **Science of the Total Environment**, v. 526, p. 58-69, 2015.

COCKX, L.; VAN MEIRVENNE, M.; DE VOS, B. Using the EM38DD soil sensor to delineate clay lenses in a sandy forest soil. **Soil Science Society of America Journal**, v. 71, p. 1314-1322, 2007.

CORWIN, D. L.; LESCH, S. M. Apparent soil electrical conductivity measurements in agriculture. **Computers and Electronics in Agriculture**, v. 46, p. 11-43, 2005.

DE BENEDETTO, D.; CASTRIGNANÒ, A.; DIACONO, M.; RINALDI, M.; RUGGIERI, S.; TAMBORRINO, R. Field partition by proximal and remote sensing data fusion. **Biosystems Engineering**, v. 114, p. 372-383, 2013.

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

DE GRUIJTER, J. J.; BRUS, D. J.; BIERKENS, M. F. P.; KNOTTERS, M. **Sampling for Natural Resource Monitoring**. Berlim: Springer, 2006.

DE WIJS, H. J. Statistics of ore distribution. Part I: Frequency distribution of assay values. **Netherlands Journal of Geosciences**, v. 13, p. 365-375, 1951.

DE WIJS, H. J. Statistics of ore distribution. Part II: Theory of binomial distribution applied to sampling and engineering problems. **Netherlands Journal of Geosciences**, v. 15, p. 12-24, 1953.

DOOLITTLE, J. A.; BREVIK, E. C. The use of electromagnetic induction techniques in soils studies. **Geoderma**, v. 223-225, p. 33-45, 2014.

FISHER, R. A. **Statistical Methods and Statistical Iinference**. Edimburgo: Oliver and Boyd, 1956.

FULTON, A.; SCHWANKL, L.; LYNN, K.; LAMPINEN, B.; EDSTROM, J.; PRICHARD, T. Using EM and VERIS technology to assess land suitability for orchard and vineyard development. **Irrigation Science**, v. 29, p. 497-512, 2011.

GHOLIZADEH, A.; AMIN, M. S. M.; ANUAR, A. R.; AIMRUN, W. Apparent electrical conductivity in correspondence to soil chemical properties and plant nutrients in soil. **Communications in Soil Science and Plant Analysis**, v. 42, p. 1447-1461, 2011.

GOMEZ, C.; VISCARRA ROSSEL, R. A.; MCBRATNEY, A. B. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. **Geoderma**, v. 146, p. 403-411, 2008.

GREENHOUSE, J. P.; SLAINE, D. D. The use of reconnaissance electromagnetic methods to map contaminant migration. **Groundwater Monitoring & Remediation**, v. 3, p. 47-59, 1983.

GUO, Y.; HUANG, J.; SHI, Z.; LI, H. Mapping spatial variability of soil salinity in a coastal paddy field based on electromagnetic sensors. **PLoS ONE**, v. 10, p. 1-12, 2015.

HARRIS J. A. Practical universality of field heterogeneity as a factor influencing plot yields. **Journal of Agricultural Research**, v. 19, p. 279-314, 1920.

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

HEIL, K.; SCHMIDHALTER, U. The application of EM38: Determination of soil parameters, selection of soil sampling points and use in agriculture and archaeology. Sensors, v. 17, p. 2540, 2017.

HENGL, T. Finding the right pixel size. **Computers & Geosciences**, v. 32, p. 1283–1298, 2006.

HUANG, J.; LARK, R. M.; ROBINSON, D. A.; LEBRON, I.; KEITH, A. M.; RAWLINS, B.; TYE, A.; KURAS, O.; RAINES, M.; TRIANTAFILIS, J. Scope to predict soil properties at within-field scale from small samples using proximally sensed γ-ray spectrometer and EM induction data. **Geoderma**, v. 232-34, p. 69-80, 2014.

HUANG, J.; TAGHIZADEH-MEHRJARDI, R.; MINASNY, B.; TRIANTAFILIS, J. Modeling soil salinity along a hillslope in Iran by inversion of EM38 data. **Soil Science Society of America Journal**, v. 79, p. 1142-1153, 2015.

HUANG, J.; MCBRATNEY, A. B.; MINASNY, B.; TRIANTAFILIS, J. 3D soil water nowcasting using electromagnetic conductivity imaging and the ensemble Kalman filter. **Journal of Hydrology**, v. 549, p. 62-78, 2017.

HUANG, J.; PROCHAZKA, M. J.; TRIANTAFILIS, J. Irrigation salinity hazard assessment and risk mapping in the lower Macintyre Valley, Australia. **Science of the Total Environment**, v. 551-552, p. 460-473, 2016.

ISLAM, M. M.; MEERSCHMAN, E.; SAEY, T.; DE SMEDT, P.; VAN DE VIJVER, E.; VAN MEIRVENNE, M. Comparing apparent electrical conductivity measurements on a paddy field under flooded and drained conditions. **Precision Agriculture**, v. 13, p. 384-392, 2012.

ISLAM, M. M.; SAEY, T.; MEERSCHMAN, E.; DE SMEDT, P.; MEEUWS, F.; VAN DE VIJVER, E.; VAN MEIRVENNE, M. Delineating water management zones in a paddy rice field using a floating soil sensing system. **Agricultural Water Management**, v. 102, p. 8-12, 2011.

JAHKNWA, C. J.; RAY, H. H.; ZEMBA, A. A.; ADEBAYO, A. A.; WUYEP, S. Z. Spatial heterogeneity of salinity parameters in vertisols of Kerau, Guyuk area of Adamawa state,

Nigeria. **International Research Journal of Agriculture and Soil Science**, v. 4, p. 5-12, 2014.

JOURNEL, A.; HUIJBREGTS, C. J. Mining Geostatistics. Londres: Academic Press, 1978.
KACHANOSKI, R. G.; GREGORICH, E. G.; VAN WESENBEECK, I. J. Estimating spatial variations of soil water content using noncontacting electromagnetic inductive methods. **Canadian Journal of Soil Science**, v. 68, p. 715-722, 1988.

KESKIN, H.; GRUNWALD, S.; HARRIS, W. G. Digital mapping of soil carbon fractions with machine learning. **Geoderma**, v. 339, p. 40-58, 2019.

KRIGE, D. G. A statistical approach to some basic mine valuation problems on the Witwatersrand. **Journal of the Chemical Metallurgical & Mining Society of South Africa**, v. 52, p. 119-139, 1951.

KWON, T. J.; MURESAN, M.; FU, L.; USMAN, T. Development of zonal-specific semivariograms for a strategic RWIS network optimization: Case study. **Journal of Infrastructure Systems**, v. 25, p. 1-9, 2019.

LANDRUM, C.; CASTRIGNANÒ, A.; MUELLER, T.; ZOURARAKIS, D.; ZHU, J.; DE BENEDETTO, D. An approach for delineating homogeneous within-field zones using proximal sensing and multivariate geostatistics. **Agricultural Water Management**, v. 147, p. 144-153, 2015.

LESCH, S. M.; RHOADES, J. D.; LUND, L. J.; CORWIN, D. L. Mapping soil salinity using calibrated electromagnetic measurements. **Soil Science Society of America Journal**, v. 56, p. 540–548, 1992.

LI, H. Y.; SHI, Z.; WEBSTER, R.; TRIANTAFILIS, J. Mapping the three-dimensional variation of soil salinity in a rice-paddy soil. **Geoderma**, v. 195-196, p. 31-41, 2013.

LOPES, I; MONTENEGRO, A. A. A. Spatialization of electrical conductivity and physical hydraulic parameters of soils under different uses in an alluvial valley. **Revista Caatinga**, v. 32, p. 222-233, 2019.

MA, R.; MCBRATNEY, A.; WHELAN, B.; MINASNY, B.; SHORT, M. Comparing temperature correction models for soil electrical conductivity measurement. **Precision Agriculture**, v. 12, p. 55-66, 2011.

MACHADO, P. L. O. A.; BERNARDI, A. C. C.; VALENCIA, L. I. O.; MOLIN, J. P.; GIMENEZ, L. M.; SILVA, C. A.; ANDRADE, A. G.; MADARI, B. E.; MEIRELLES, M. S. P. Mapeamento da condutividade elétrica e relação com a argila de Latossolo sob plantio direto. **Pesquisa Agropecuária Brasileira**, v. 41, p. 1023-1031, 2006.

MATHERON, G. **Les Variables Régionalisées et Leur Estimation**. Une Application de la Théorie des Fonctions Aléatoires aux Siences de la Nature. Paris: Masson et Cie, 1965.

MATHERON, G. Principles of geostatistics. **Economic Geology**, v. 58, p. 1246-1266, 1963.

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

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

MCKENZIE, R. C.; CHOMISTEK, W.; CLARK, N. F. Conversion of electromagnetic inductance readings to saturated past extract values in soils for different temperature, texture, and moisture conditions. **Canadian Journal of Soil Science**, v. 69, p. 25-32, 1989.

MCNEILL, J. D. **Electromagnetic Terrain Conductivity Measurement at Low Induction Numbers**. Ontario: Geonics Limited, 1980.

MERSMANN, O. microbenchmark: Accurate Timing Functions. **R Package Version** 1.4-7. 2019.

MILLER, B. A.; SCHAETZL, R. J. The historical role of base maps in soil geography. **Geoderma**, v. 230-231, p. 329-339, 2014.

MINASNY, B.; MCBRATNEY, A. B. A conditioned Latin hypercube method for sampling in the presence of ancillary information. **Computers & Geosciences**, v. 32, p. 1378-1388, 2006.

MONTGOMERY, E. G. **Experiments in Wheat Breeding**: Experimental Error in the Nursery and Variation in Nitrogen and Yield. Washington: United States Department of Agriculture, 1913.

MULDER, V. L.; DE BRUIN, S.; SCHAEPMAN, M. E.; MAYR, T. R. The use of remote sensing in soil and terrain mapping - A review. **Geoderma**, v. 162, p. 1-19, 2011.

MULLINS, C. E. Magnetic susceptibility of the soil and its significance in soil science – A review. **Journal of Soil Science**, v. 28, p. 223-246, 1977.

MYERS, D. B.; KITCHEN, N. R.; SUDDUTH, K. A.; GRUNWALD, S.; MILES, R. J.; SADLER, E. J.; UDAWATTA, R. P. Combining Proximal and Penetrating Soil Electrical Conductivity Sensors for High-resolution Digital Soil Mapping. In: VISCARRA ROSSEL, R. A.; MCBRATNEY, A. B.; MINASNY B. (Eds.). **Proximal Soil Sensing**. Dordrecht: Springer, 2010. p. 233-243.

NOURI, H.; BORUJENI, S. C.; ALAGHMAND, S.; ANDERSON, S. J.; SUTTON, P. C.; PARVAZIAN, S.; BEECHAM, S. Soil salinity mapping of urban greenery using remote sensing and proximal sensing techniques; The case of Veale Gardens within the Adelaide Parklands. **Sustainability**, v. 10, p. 1-14, 2018.

PEBESMA, E. J. Multivariable geostatistics in S: the gstat package. **Computers & Geosciences**, v. 30, p. 683-691, 2004.

PENDLETON, R. L. **Are soils mapped under a given type name by the bureau of soils method closely similar to one another?** Berkeley: University of California Press, 1919.

PIIKKI, K.; SÖDERSTRÖM, M.; STENBERG, B. Sensor data fusion for topsoil clay mapping. **Geoderma**, v. 199, p. 106-116, 2013.

R CORE TEAM. R: A Language and Environment for Statistical Computing. Viena: R **Foundation for Statistical Computing**, 2019.

RAMIREZ-LOPEZ, L.; SCHMIDT, K.; BEHRENS, T.; VAN WESEMAEL, B.; DEMATTÊ, J. A. M.; SCHOLTEN, T. Sampling optimal calibration sets in soil infrared spectroscopy. **Geoderma**, v. 226-227, p. 140-150, 2014.

RAMOS, A. M.; SANTOS, L. A. R.; FORTES, L. T. G. (Eds.). **Normais Climatológicas do Brasil 1961-1990.** Versão Revista e Ampliada. Brasília: Instituto Nacional de Meteorologia, 2009.

RAWLINS, B. G.; MARCHANT, B. P.; SMYTH, D.; SCHEIB, C.; LARK, R. M.; JORDAN, C. Airborne radiometric survey data and a DTM as covariates for regional scale mapping of soil organic carbon across Northern Ireland. **European Journal of Soil Science**, v. 60, p. 44-54, 2009.

RHOADES, J. D.; CORWIN, D. L. Determining soil electrical conductivity-depth relations using an inductive electromagnetic soil conductivity meter. **Soil Science Society of America Journal**, v. 45, p. 255-260, 1981.

ROBINSON, G. W.; LLOYD, W. E. On the probable error of sampling in soil surveys. **The Journal of Agricultural Science**, v. 7, p. 144-153, 1915.

RODRIGUES, F. A.; BRAMLEY, R. G. V.; GOBBETT, D. L. Proximal soil sensing for Precision Agriculture: Simultaneous use of electromagnetic induction and gamma radiometrics in contrasting soils. **Geoderma**, v. 243-244, p. 183-195, 2015.

ROSSI, M. **Mapa Pedológico do Estado de São Paulo:** Revisado e Ampliado. 1. ed. São Paulo: Instituto Florestal, 2017.

ROSSITER, D. G. Past, present & future of information technology in pedometrics. **Geoderma**, v. 324, p. 131-137, 2018.

ROUDIER, P. clhs: a R Package for Conditioned Latin Hypercube Sampling. **Software R Package**. 2011.

SAMET, R.; ÇELIK, E.; TURAL, S.; ŞENGÖNÜL, E.; ÖZKAN, M.; DAMCI, E. Using interpolation techniques to determine the optimal profile interval in ground-penetrating radar applications. **Journal of Applied Geophysics**, v. 140, p. 154-167, 2017.

SAMUEL-ROSA, A. spsann: Optimization of Sample Configurations Using Spatial Simulated Annealing. **Software R Package Version 2.2.0.** 2019.

SHEETS, K. R.; HENDRICKX, J. M. H. Noninvasive soil water content measurement using electromagnetic induction. **Water Resources Research**, v. 31, p. 2401-2409, 1995.

SLAVICH, P. G. Determining ECa-depth profiles from electromagnetic induction measurements. **Australian Journal of Soil Research**, v. 28, p. 443-452, 1990.
SLAVICH, P. G.; READ B. J. Assessment of electromagnetic induction measurements using an inductive electromagnetic soil conductivity meter. **Soil Science Society of America Journal**. v. 45, p. 255-260, 1983.

SMITH, L. H. Plot arrangement for variety experiments with corn. **Agronomy Journal**, v. 1, p. 84-89, 1909.

SÖDERSTRÖM, M.; ERIKSSON, J.; ISENDAHL, C.; ARAÚJO, S. R.; REBELLATO, L.; PAHL SCHAAN, D.; STENBORG, P. Using proximal soil sensors and fuzzy classification for mapping Amazonian Dark Earths. **Agricultural and Food Science**, v. 22, p. 380-389, 2013.

SUDDUTH, K. A.; KITCHEN, N. R.; WIEBOLD, W. J.; BATCHELOR, W. D.; BOLLERO, G. A.; BULLOCK, D. G.; CLAY, D. E.; PALM, H. L.; PIERCE, F. J.; SCHULER, R. T.; THELEN, K. D. Relating apparent electrical conductivity to soil properties across the north-central USA. **Computers and Electronics in Agriculture**, v. 46, p. 263-283, 2005.

SUDDUTH, K. A.; DRUMMOND, S. T.; KITCHEN, N. R. Accuracy issues in electromagnetic induction sensing of soil electrical conductivity for precision agriculture. **Computers and Electronics in Agriculture**, v. 31, p. 239-264, 2001.

SUDDUTH, K. A; KITCHEN, N. R.; MYERS, D. B.; DRUMMOND, S. T. Mapping depth to argillic soil horizons using apparent electrical conductivity. **Journal of Environmental and Engineering Geophysics**, v. 15, p. 135-146, 2010.

SUDDUTH, K. A.; MYERS, D. B.; KITCHEN, N. R.; DRUMMOND, S. T. Modeling soil electrical conductivity-depth relationships with data from proximal and penetrating ECa sensors. **Geoderma**, v. 199, p. 12-21, 2013.

TEIXEIRA, P. C.; DONAGEMMA, G. K.; FONTANA, A.; TEIXEIRA, W. G. **Manual de Métodos de Análise de Solo**. 3. ed. rev ed. Brasília, DF: Embrapa, 2017.

TOBLER, W. R. A Computer movie simulating urban growth in the Detroit region. **Economic Geography**, v. 46, p. 234-240, 1970.

TRANGMAR, B. B.; YOST, R. S.; UEHARA, G. Application of geostatistics to spatial studies of soil properties. **Advances in Agronomy**, v. 36, p. 45-94, 1985.

TRIANTAFILIS, J.; LESCH, S. M. Mapping clay content variation using electromagnetic induction techniques. **Computers and Electronics in Agriculture**, v. 46, p. 203237, 2005.

TRIANTAFILIS, J.; SANTOS, F. A. M. Resolving the spatial distribution of the true electrical conductivity with depth using EM38 and EM31 signal data and a laterally constrained inversion model. **Australian Journal of Soil Research**, v. 48, p. 434-446, 2010.

USDA (UNITED STATES DEPARTMENT OF AGRICULTURE). **Soil Survey Manual**. Washington, D.C.: United States Department of Agriculture, 2017.

VAN DER LELIJ, A., **Use of an electromagnetic induction instrument (type EM38) for mapping of soil salinity**, Internal Report Research Branch, Water Resources Commission, NSW, Australia, 1983.

VIEIRA, S. R.; MILLETE, J.; TOPP, G. Handbook for Geostatistical Analysis of Variability in Soil and Climate Data. In: ALVAREZ, V. H.; SCHAEFER, C. R.; BARROS, N. F.; MELLO, J. W. V.; COSTA, L. M. (Eds.). **Tópicos em Ciência do Solo**. Volume 2. Viçosa: Sociedade Brasileira de Ciência do Solo, 2002. p. 1-45.

VISCARRA ROSSEL, R. A.; ADAMCHUK, V. I.; SUDDUTH, K. A.; MCKENZIE, N. J.; LOBSEY, C. Proximal Soil Sensing: An Effective Approach for Soil Measurements in Space and Time. In: SPARKS, D. (Ed.). **Advances in Agronomy**. Volume 113. Amsterdã: Elsevier, 2011. p. 243-291.

VISCARRA ROSSEL, R. A.; BOUMA, J. Soil sensing: A new paradigm for agriculture. **Agricultural Systems**, v. 148, p. 71-74, 2016.

WACKERNAGEL, H. **Multivariate Geoestatistics:** An Introduction with applications. 3. ed. Verlag Berlin Heidelberg: Springer, 2003.

WALVOORT, D. J. J.; BRUS, D. J.; DE GRUIJTER, J. J. An R package for spatial coverage sampling and random sampling from compact geographical strata by k-means. **Computers & Geosciences**, v. 36, p. 1261-1267, 2010.

WAYNICK, D. D. Variability in soils and its significance to past and future soil investigations: I. A statistical study of nitrification in soils. **University of California Publications in Agricultural Science**, v. 3, p. 243-270, 1918.

WAYNICK, D. D.; SHARP, L. T. Variability in soils and its significance to past and future soil investigations: II. Variability in nitrogen and carbon in field soils and their relation to the accuracy of field trials. **University of California Publications in Agricultural Science**, v. 4, p. 121-139, 1919.

WEBSTER, R.; OLIVER, M. A. **Geostatistics for Environmental Scientists**. 2. ed. Chichester: John Wiley & Sons, 2007.

WOLLENHAUPT, N. C.; RICHARDSON, J. L.; FOSS, J. E.; DOLL, E. C. A rapid method for estimating weighted soil salinity from apparent soil electrical conductivity measured with an aboveground electromagnetic induction meter. **Canadian Journal of Soil Science**, v. 66, p. 315-321, 1986.

YAMAMOTO, J. K.; LANDIM, P. M. B. **Geoestatística:** Conceitos e Aplicações. São Carlos: Oficina de Textos, 2013.

YAO, X.; FU, B.; LÜ, Y.; SUN, F.; WANG, S.; LIU, M. Comparison of four spatial interpolation methods for estimating soil moisture in a complex terrain catchment. **PLoS ONE**, v. 8, p. 1-13, 2013.